Oberwolfach Reports

OWR Vol.1, Iss.1
First quarter of 2004

Contents:

Mathematisches Forschungsinstitut Oberwolfach

Report No. 1/2004

Combinatorics

Organised by
László Lovász (Redmond)
Hans Jürgen Prömel (Berlin)

January 4th – January 10th, 2004

Introduction by the Organisers

The conference was organized by László Lovász (Redmond) and Hans Jürgen Prömel (Berlin). The programme consisted of 15 lectures, supplemented by 21 shorter contributions, and covered many areas in Combinatorics such as partition theory, discrete geometry, homomorphisms and lattices, extremal combinatorics, graph theory, random structures, and additive number theory. The aim of the workshop was to emphasize the underlying methods that are common to many of these combinatorial branches and that act as both driving forces and organizing principles of the field. The diversity of the topics and participants stimulated a lot of fruitful discussion between the different branches and gave rise to new collaborations, in particular for the younger generation of researchers.

In total, 51 scientists participated in this meeting; almost 40 came from countries other than Germany. The organizers and participants thank the *Mathematisches Forschungsinstitut Oberwolfach* for providing an inspiring setting for this conference. In the following we include the abstracts in alphabetical order.

Workshop on Combinatorics

Table of Contents

Abstracts

CutNorm, Grothendieck's Inequality, and Approximation Algorithms for Dense Graphs

Noga Alon
(joint work with Assaf Naor)

The *cut-norm* $||A||_C$ of a real matrix $A = (a_{ij})_{i\in R, j\in S}$ with a set of rows indexed by R and a set of columns indexed by S is the maximum, over all $I \subset R,\ J \subset S$, of the quantity $|\sum_{i\in I, j\in J} a_{ij}|$. This concept plays a major role in the work of Frieze and Kannan on efficient approximation algorithms for dense graph and matrix problems, [3] (see also [1] and its references). Although the techniques in [3] enable the authors to approximate efficiently the cut-norm of an n by m matrix with entries in $[-1,1]$ up to an **additive** error of ϵnm, there is no known polynomial algorithm that approximates the cut-norm of a general real matrix up to a constant multiplicative factor.

Let CUT NORM denote the computational problem of computing the cut-norm of a given real matrix. Here we first observe that the CUT NORM problem is MAX SNP hard, and then provide an efficient approximation algorithm for the problem. This algorithm finds, for a given matrix $A = (a_{ij})_{i\in R, j\in S}$, two subsets $I \subset R$ and $J \subset S$, such that $|\sum_{i\in I, j\in J} a_{ij}| \geq \rho ||A||_C$, where $\rho > 0$ is an absolute constant. We first describe a deterministic algorithm that supplies a rather poor value of ρ, and then describe a randomized algorithm that provides a solution of expected value greater than 0.56 times the optimum.

The algorithm combines semidefinite programming with a novel rounding technique based on (the proofs of) Grothendieck's Inequality. This inequality, first proved in [6], is a fundamental tool in Functional Analysis, and has several interesting applications in this area. We will actually use the matrix version of Grothendieck's inequality, formulated in [10]. In order to apply semidefinite programming for studying the cut-norm of an n by m matrix $A = (a_{ij})$, it is convenient to first study another norm,

$$||A||_{\infty\mapsto 1} = \max \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}x_i y_j,$$

where the maximum is taken over all $x_i, y_j \in \{-1, 1\}$.

It is not difficult to show, that for every matrix A,

$$4||A||_C \geq ||A||_{\infty\mapsto 1} \geq ||A||_C,$$

and hence a constant approximation of any of these norms provides a constant approximation of the other.

The value of $||A||_{\infty \mapsto 1}$ is given by the following quadratic integer program

$$\text{Maximize} \quad \sum_{ij} a_{ij} x_i y_j \tag{1}$$

$$\text{subject to} \quad x_i, y_j \in \{-1, 1\} \quad \text{for all} \quad i, j.$$

The obvious semidefinite relaxation of this program is

$$\text{Maximize} \quad \sum_{ij} a_{ij} u_i \cdot v_j \tag{2}$$

$$\text{subject to} \quad ||u_i|| = ||v_j|| = 1,$$

where here $u_i \cdot v_j$ denotes the inner product of u_i and v_j, which are now vectors of (Euclidean) norm 1 that lie in an arbitrary Hilbert space. Clearly we may assume, without loss of generality, that they lie in an $n+m$-dimensional space.

This semidefinite program can be solved, using well known techniques (see [5]) within an additive error of ϵ, in polynomial time (in the length of the input and in the logarithm of $1/\epsilon$.) The main problem is the task of rounding this solution into an integral one. A first possible attempt is to imitate the technique of Goemans and Williamson in [7], that is, given a solution u_i, v_j to the above program, pick a random vector z and define $x_i = \text{sign}(u_i \cdot z)$ and $y_j = \text{sign}(v_j \cdot z)$. It is easy to check that the expected value of $x_i y_j$ satisfies $E(x_i y_j) = \frac{2}{\pi} \arcsin(u_i \cdot v_j)$, and as $\arcsin(t)$ and t differ only in constant factors for all $-1 \leq t \leq 1$, one could hope that this will provide an integral solution whose value is at least some absolute constant fraction of the value of the optimal solution. This reasoning is, unfortunately, incorrect, as some of the entries a_{ij} may be positive and some may be negative, (in fact, the problem is interesting only if this is the case, since otherwise either $x_i = y_j = 1$ or $x_i = -y_j = 1$ for all i, j supplies the required maximum). Therefore, even if each single term $a_{ij} u_i \cdot v_j$ is approximated well by its integral rounding $a_{ij} x_i y_j$, there is no reason to expect the sum to be well-approximated, due to cancellations. We thus have to compare the value of the rounded solution to that of the semidefinite program on a global basis. Nesterov [11] obtained a result of this form for the problem of approximating the maximum value of a quadratic form $\sum_{ij} b_{ij} x_i x_j$, where $x_i \in \{-1, 1\}$, but only for the special case in which the matrix $B = (b_{ij})$ is positive semidefinite. While his estimate is global, his rounding is the same simple rounding technique of [7] described above. As explained before, some new ideas are required in our case in order to get any nontrivial result.

Luckily, there is a well known inequality of Grothendieck, which asserts that the value of the semidefinite program (2) and that of the integer program (1) can differ only by a constant factor. The precise value of this constant, called

Grothendieck's constant and denoted by K_G, is not known, but it is known that its value is at most $\frac{\pi}{2\ln(1+\sqrt{2})} = 1.782...$ ([8]) and at least $\frac{\pi}{2} = 1.570...$ ([6]). Stated in other words, the integrability gap of the problem is at most K_G. (Krivine mentions in [8] that he can improve the lower bound, but such an improvement has never been published).

It follows that the value of the semidefinite program (2) provides an approximation of $||A||_{\infty\mapsto 1}$ up to a constant factor. This, however, still does not tell us how to round the solution of the semidefinite program into an integral one with a comparable value. Indeed, this task requires more work, and is carried out in the full paper a preliminary version of which will appear in the proceedings of STOC 2004.

We describe three rounding techniques. The first one is a deterministic procedure, which combines Grothendieck's Inequality with some facts about four-wise independent random variables, in a manner that resembles the technique used in [2] to approximate the second frequency moment of a stream of data under severe space constraints. The second rounding method is based on Rietz' proof of Grothendieck's Inequality [12]. This proof supplies a better approximation guarantee for the special case of positive semidefinite matrices A, where the integrality gap can be shown to be precisely $\pi/2$, and implies that Nesterov's analysis for the problem he considers in [11] is tight.

The third technique, which supplies the best approximation guarantee, is based on Krivine's proof of Grothendieck's Inequality. Here we use the vectors u_i, v_j which form a solution of the semidefinite program (2) to construct some other unit vectors u'_i, v'_j, which are first shown to exist in an infinite dimensional Hilbert space, and are then found, using another instance of semidefinite programming, in an $n+m$-dimensional space. These vectors can now be rounded to $\{-1, 1\}$ in order to provide an integral solution for the original problem (1) in a rather simple way. We note that there are several known techniques for modifying the solution of a semidefinite program before rounding it, see [13], [9], [4]. Here, however, the modification seems more substantial.

We believe that our techniques will have further applications, as they provide a method for handling problems in which there is a possible cancellation between positive and negative terms. It seems that there are additional interesting problems of this type. Moreover, unlike the semidefinite based approximation algorithms for MAX CUT, MAX 2SAT and related problems, suggested in the seminal paper of [7] and further developed in many subsequent papers, the problem considered here has no known constant approximation algorithm, and the semidefinite programming and its rounding appear to be essential in order to obtain any constant approximation guarantee, and not only in order to improve the constants ensured by appropriate combinatorial algorithms.

References

[1] N. Alon, W. F. de la Vega, R. Kannan and M. Karpinski, Random sampling and approximation of MAX-CSP problems, Proc. of the 34^{th} ACM STOC, ACM Press (2002), 232–239.

[2] N. Alon, Y. Matias and M. Szegedy, The space complexity of approximating the frequency moments, Proc. of the 28^{th} ACM STOC, ACM Press (1996), 20-29. Also; J. Comp. Sys. Sci. 58 (1999), 137-147.

[3] A. M. Frieze and R. Kannan, Quick Approximation to matrices and applications, Combinatorica **19** (2) (1999), 175–200.

[4] U. Feige and M. Langberg, The RPR^2 Rounding Technique for Semidefinite Programs, Proceedings of the 28th International Colloquium on Automata, Languages and Programming, Crete, Greece, 2001, 213–224.

[5] M. Grötschel, L. Lovász and A. Schrijver, The ellipsoid method and its consequences in combinatorial optimization, Combinatorica 1 (1981), 169–197.

[6] A. Grothendieck, Résumé de la théorie métrique des produits tensoriels topologiques, Bol. Soc. Mat. Sao Paulo 8 (1953), 1–79.

[7] M. X. Goemans and D. P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, J. ACM 42 (1995), 1115–1145.

[8] J. L. Krivine, Sur la constante de Grothendieck, C. R. Acad. Sci. Paris Ser. A-B 284 (1977), 445–446.

[9] M. Lewin, D. Livnat and U. Zwick, Improved rounding techniques for the MAX 2-SAT and MAX DI-CUT problems, Proceedings of the 9th Conference on Integer Programming and Combinatorial Optimization, Cambridge, Massachusetts, 2002, 67–82.

[10] J. Lindenstrauss and A. Pełczyński, Absolutely summing operators in L_p-spaces and their applications, Studia Math. 29 (1968), 275-326.

[11] Y. E. Nesterov, Semidefinite relaxation and nonconvex quadratic optimization, Optimization Methods and Software 9 (1998), 141–160.

[12] R. E. Rietz, A proof of the Grothendieck inequality, Israel J. Math. 19 (1974), 271–276.

[13] U. Zwick, Outward rotations: a tool for rounding solutions of semidefinite programming relaxations, with applications to MAX CUT and other problems, Proceedings of the 31th Annual ACM Symposium on Theory of Computing (STOC), Atlanta, Georgia, 1999, 679–687.

Blockers, Ideals and some Turán-type Questions

Anders Björner

(joint work with Axel Hultman, Irena Peeva and Jessica Sidman [1, 2])

The point of departure are the theorems of Li & Li and Kleitman & Lovász (from 1981) describing generators for certain ideals, see [3]. The immediate combinatorial interest of these theorems is that they in a useful way describe ideals with the property that (upper) bounded independence number and (lower) bounded chromatic number of a given graph are equivalent to membership of the corresponding graph polynomial in these ideals. But the theorems are also interesting from a ring-theoretic point of view, since they suggest a combinatorial procedure for constructing generators for vanishing ideals of subspace arrangements.

The work presented was:

(1) The *blocker construction* $A \mapsto A^*$ for antichains in finite posets, generalizing the well-known concept in Boolean lattices (set clutters). Particularly how to compute blockers for symmetric antichains in the partition lattice Π_n. This procedure involves both the refinement order and the dominance order on the set of all number partitions of n.

(2) The construction of the *blocker ideal* $B_{\mathcal{A},\mathcal{H}}$ for a subspace arrangement $\mathcal{A}$ embedded in a hyperplane arrangement $\mathcal{H}$. This ideal is contained in the vanishing ideal $I_{\mathcal{A}}$ for the union of the subspaces in $\mathcal{A}$, and

$$B_{\mathcal{A},\mathcal{H}} = I_{\mathcal{A}} \quad \Rightarrow \quad \mathcal{A}^{**} = \mathcal{A},$$

where $\mathcal{A}^*$ denotes the blocker of $\mathcal{A}$ w.r.t. the intersection lattice of $\mathcal{H}$.

(3) The fact that $B_{\mathcal{A},\mathcal{H}} = I_{\mathcal{A}}$ implies that a minimal blocking set for $\mathcal{A}$ has size equal to the minimal size of a flat in the blocker $\mathcal{A}^*$. Some extremal results (e.g. Turán's theorem) can be deduced this way.

References

[1] A. Björner and A. Hultman, *A note on blockers in posets*, preprint, 2004. (http://arXiv.org/abs/math.CO/0403094)

[2] A. Björner, I. Peeva and J. Sidman, *Subspace arrangements defined by products of linear forms*, preprint, 2003. (http://arXiv.org/abs/math.CO/0401373)

[3] László Lovász, *Stable sets and polynomials*, Discrete Math. **124** (1994), 137–153.

Random Geometric Graphs

Béla Bollobás
(joint work with Paul Balister, Amites Sarkar and Mark Walters)

Random geometric graphs were introduced by Gilbert [6] in 1961, and in the past forty years many variants of them have been studied in great detail (see Meester and Roy [7], Penrose [8]). The aim of the talk is to present a number of recent results obtained jointly with Paul Balister, Amites Sarkar and Mark Walters on a variety of geometric random graphs.

Gilbert's *disc model* G_r is defined as follows. Place points $\{x_i\}$ in $\mathbb{R}^2$ according to a Poisson process with intensity 1 and let G_r be the random graph with vertex set $\{x_i\}$ and edges x_ix_j whenever $|x_i - x_j| \le r$. Equivalently, let D_r be the disc of radius r with centre the origin, and join each x_i to every x_j in the disc $x_i + D_r$ of radius r centred at x_i. There is a critical area a_c such that if $|D_r| = \pi r^2 < a_c$ then a.s. G_r has no infinite component (G_r does not percolate), while if $|D_r| > a_c$ then G_r percolates a.s. The proven bounds on a_c are still rather weak, with almost a factor 5 between the upper and lower bounds. In the talk we present the result due to Balister, Bollobás and Walters [4] that $4.508 < a_c < 4.515$ *with probability* 99.99%. (The probability is due to the uncertainty of numerically evaluating a large integral.) For the critical area s_c of a square rather than a disc, defined analogously, the corresponding bounds are $4.392 < s_c < 4.398$.

Problems concerning *ad hoc* networks of radio transceivers inspire the following considerable extension of the disc model. Place points $\{x_i\}$ in $\mathbb{R}^d$ according to a Poisson process with intensity 1. Then, independently for each x_i, choose a bounded region A_{x_i} from some fixed distribution and let $\mathcal{G}$ be the random directed graph with vertex set $\{x_i\}$ and edges $\vec{x_ix_j}$ whenever $x_j \in x_i + A_{x_i}$. The main result of Balister, Bollobás and Walters [3] states that for any $\eta > 0$, if the regions $x_i + A_{x_i}$ do not overlap too much (i.e., satisfy a somewhat technical precise condition), then $\mathcal{G}$ has an infinite directed path provided the expectation of the area $|A_{x_i}|$ of the domain A_{x_i} is at least $1 + \eta$. (It is trivial that the area has to be at least 1.) One example where these conditions hold, and we obtain percolation, is in dimension d

with A_{x_i} a ball of volume $1+\eta$, where η tends to zero as d tends to infinity. Another example is in two dimensions, where the A_{x_i} are randomly oriented sectors of a disk of angle $2\pi\varepsilon$ and area $1+\eta$. In this case we can let η tend to zero as ε tends to zero. Yet another special case of this theorem is the result proved independently in [2] and by Franceschetti et al [5] that, given $\eta > 0$, if $\varepsilon > 0$ is small enough, in $\mathbb{R}^2$ we may take each A_{x_i} to be a 'thin' annulus $A = \{x \in \mathbb{R}^2 : \ r(1-\varepsilon) \leq |x| \leq r\}$ of area $1+\eta$.

In the talk we shall examine some *finite* geometric random graphs as well. Let $\mathcal{P}$ be a Poisson process of intensity one in a square S_n of area n. We construct a random geometric graph $G_{n,k}$ by joining each point of $\mathcal{P}$ to its k nearest neighbors. Recently, Xue and Kumar [9] proved that if $k = 0.074 \log n$ then the probability that $G_{n,k}$ is connected tends to zero as $n \to \infty$, while if $k = 5.1774 \log n$ then the probability that $G_{n,k}$ is connected tends to one as $n \to \infty$. They conjectured that the threshold for connectivity is $k = \log n$. Recently, Balister, Bollobás, Sarkar and Walters [1] have improved these lower and upper bounds to $k = 0.3043 \log n$ and $k = 0.5139 \log n$, respectively, disproving this conjecture, and have proved reasonably good bounds for some generalizations of this problem.

References

[1] P. Balister, B. Bollobás, A. Sarkar and M. Walters, Connectivity of random geometric graphs, to appear

[2] P. Balister, B. Bollobás and M. Walters, Continuum percolation with steps in an annulus, *Advances of Applied Probability,* to appear

[3] P. Balister, B. Bollobás and M. Walters, Random transceiver networks, to appear

[4] P. Balister, B. Bollobás and M. Walters, Continuum percolation in the square and the disc, to appear

[5] M. Franceschetti, L. Booth, M. Cook, R. Meester and J. Bruck, *IEEE Trans. on Information Theory,* to appear

[6] E.N. Gilbert, Random plane networks, *Journal of the Society for Industrial Applied Mathematics* **9** (1961), 533-543.

[7] R. Meester and R. Roy, *Continuum Percolation,* Cambridge University Press, 1996, x+238 pp.

[8] M.D. Penrose, *Random Geometric Graphs,* Oxford University Press, 2003, xiii+330 pp.

[9] F. Xue and P.R. Kumar, The number of neighbors needed for connectivity of wireless networks, *Wireless Networks*, to appear

The Number of Linear Extensions of the Boolean Lattice

Graham Brightwell
(joint work with Prasad Tetali [1])

Let $L(P)$ denote the number of linear extensions of a poset P. A natural problem is to estimate $L(P)$ when P is the Boolean lattice Q^t, consisting of the subsets of $\{1, 2, \ldots, t\}$, ordered by inclusion. This problem was apparently first posed by Richard Stanley, although it has also been raised by several others independently.

A trivial lower bound on $L(Q^t)$ is $\prod_{j=0}^{t} \binom{t}{j}!$, and a simple upper bound is $\binom{t}{\lfloor t/2 \rfloor}^{2^t}$; these bounds can be written as

$$\log \binom{t}{\lfloor t/2 \rfloor} - \frac{3}{2} \log e + o(1) \leq \frac{\log(L(Q^t))}{2^t} \leq \log \binom{t}{\lfloor t/2 \rfloor}.$$

(All logarithms are base 2.)

The only previous improvement on these trivial bounds was made by Sha and Kleitman [4], who improved the upper bound to

$$L(Q^t) \leq \prod_{j=0}^{t} \binom{t}{j}^{\binom{t}{j}} \leq \prod_{j=0}^{t} \binom{t}{j}! \exp(2^t),$$

yielding

$$\frac{\log(L(Q^t))}{2^t} \leq \log \binom{t}{\lfloor t/2 \rfloor} - \frac{1}{2} \log e + o(1).$$

In fact, the Sha-Kleitman bound can be generalised to any ranked poset satisfying the LYM condition (see [1]).

We prove the following result, which shows that (as was generally expected) the trivial lower bound gives the correct constant term in the asymptotic expansion:

$$\frac{\log(L(Q^t))}{2^t} = \log \binom{t}{\lfloor t/2 \rfloor} - \frac{3}{2} \log e + O\left(\frac{\ln t}{t}\right).$$

Our proof is based on what seems to be emerging as an “entropy method” developed by Jeff Kahn [2], and used by him [3] to give a short and natural proof of the Kleitman-Markowsky bound for Dedekind's problem concerning the number of antichains in the Boolean lattice.

In the case where the poset P is bipartite, a small adaptation of Kahn's proof from [2] yields an extremal result. For P a bipartite poset on n elements, with two ranks A and B, such that every element of A is below exactly u elements of B, and every element of B is above exactly d elements of A, we have

$$L(P) \leq n! \binom{d+u}{u}^{-n/(d+u)}.$$

This result is best possible for n a multiple of $d+u$.

References

[1] G.R. Brightwell and P. Tetali, The number of linear extensions of the Boolean lattice, to appear in *Order*.

[2] J. Kahn, An entropy approach to the hard-core model on bipartite graphs, *Combin. Prob. Comp.* **10** (2001), 219–237.

[3] J. Kahn, Entropy, independent sets and antichains: a new approach to Dedekind's problem, *Proc. Amer. Math. Soc.* **130** (2002), 371–378.

[4] J. Sha and D. J. Kleitman, The number of linear extensions of subset ordering, *Discrete Math.*, **63** (1987), 271–279.

Coloring Claw-free Graphs

Maria Chudnovsky

(joint work with Paul Seymour)

A graph is called *claw-free* if it has no induced subgraph isomorphic to $K_{1,3}$. Line graphs are a well-known class of claw-free graphs, but there are others, such as *circular interval graphs* and subgraphs of the *Schläfli graph* (a circular interval graph is obtained from a collection of circular intervals and points on a circle by making two points adjacent if they belong to the same interval). Recently we were able to prove that all claw-free graphs in which every vertex is in a stable set of size three, can be built from the classes mentioned above, together with some others, by combining them in prescribed ways (this work is described in another paper in this issue).

Claw-free graphs being a generalization of line graphs, it is natural to ask what properties of line graphs can be extended to all claw-free graphs. Vizing's theorem [1] gives a bound on the chromatic number, χ, of a line graph, in terms of the size of a maximum clique, ω, namely $\chi \leq \omega + 1$. Is there a similar bound for all claw-free graphs? Does there exist a function f such that if G is a claw-free graph then $\chi(G) \leq f(\omega(G))$? It is easy to see that such f exists, in fact $\chi(G) \leq \omega(G)^2$ (the neighborhood of a vertex in a clique of size ω is the union of at most ω cliques).

One might hope to get closer to Vizing's bound, asking whether f is a linear function. Unfortunately the answer to this question is negative. If G is the complement of a triangle free graph, then $\chi(G) \geq \frac{|V(G)|}{2}$, and yet $\omega(G)$ may be of order $\sqrt{(|V(G)|)}$. However, if we insist that G contains a stable set of size three, and is connected (to prevent taking disjoin union with large complement triangle-free graphs), then a much stronger result is true. We prove:

Theorem 1 *Let G be a connected, claw-free graph and assume that G contains a stable set of size three. Then $\chi(G) \leq 2\omega(G)$.*

This bound is best possible. The proof of 1 uses the structure theorem mentioned above: first we verify the result for the basic classes of claw-free graphs, and then prove that it is preserved under the operations. This proves the theorem for those claw-free graphs that satisfy the hypotheses of the structure theorem, namely claw-free graphs where every vertex is in a stable set of size three. But it turns out that having proved the result for the part of the graph where every vertex is in a stable set of size three, one can always figure out the "important" information about vertices not in stable sets of size three, and finish the proof.

There is a slightly worse, but still linear bound on χ in terms of ω, that has a short proof, without using the structure theorem, and we include it here.

Theorem 2 *Let G be a connected, claw-free graph and assume that G contains a stable set of size three. Then $\chi(G) \leq 4\omega(G)$.*

In fact, we prove the following stronger statement that clearly implies 2. This was conjectured by N. Linial during the Oberwolfach meeting.

Theorem 3 *Let G be a connected, claw-free graph and assume that G contains a stable set of size three. Then every vertex of G has degree at most $4\omega(G)$.*

Proof. We use induction on $|V(G)|$. Let v be a vertex of maximum degree in G and let N be the set of neighbors of v. Since G is claw-free and contains a stable set of size three, $V(G) \neq N \cup \{v\}$ and there exists a vertex $u \in V(G) \setminus (N \cup \{v\})$ such that the graph $G \setminus u$ is connected. We may assume $G \setminus u$ does not contain a stable set of size three, for otherwise the result follows inductively. Let A be the set of neighbors of u in G and B the set of non-neighbors. Since G contains a stable set of size three, and $G \setminus u$ does not, it follows that there exist two non-adjacent vertices b_1, b_2 in B. Since G is connected, A is non-empty. For $i = 1, 2$ let N_{b_i} be the set of neighbors of b_i in A. Since every vertex in $N_{b_1} \cap N_{b_2}$ would be the center of a claw in G, $N_{b_1} \cap N_{b_2} = \emptyset$. Since $G \setminus u$ contains no stable set of size three, $A \setminus N_{b_i}$ is a clique for $i = 1, 2$, and A is the union of two cliques. Also since $G \setminus u$ contains no stable set of size three, $N_{b_1} \cup N_{b_2} = A$. So for every pair of non-adjacent vertices in B, the sets of their neighbors in A partition A. It follows that $G|B$ does not contain the complement of an odd cycle, and so $G|B$ is the complement of a bipartite graph, in particular B is the union of two cliques. But now G is the union of four cliques, so $\omega \geq \frac{|V(G)|}{4}$, and the theorem holds. This proves 3. ∎

References

[1] D.B. West, Introduction to Graph Theory, (Prentice Hall, 2001),

The Homology of a Locally Finite Graph with Ends

Reinhard Diestel

When one studies the homology aspects of an infinite graph – in graph-theoretic language, the properties of its cycle space – one can observe a curious phenomenon: while all the basic properties of the cycle space of a finite graph remain true (and trivial) also for infinite graphs, few of the less trivial theorems carry over.

Surprisingly, the situation can be remedied simultaneously for all those theorems that fail in the infinite case by using a different homology for locally finite graphs: not the simplicial homology of the graph itself, but a variant of the singular homology of its Freudenthal compactification.

Our approach permits the extension to locally finite infinite graphs of the following finite theorems, whose infinite analogues all fail with the usual simplicial homology:

- Tutte's theorem that the peripheral (ie., non-separating and induced) cycles of a 3-connected graph generate its cycle space;
- Whitney's theorem that a graph has a combinatorial dual if and only if it is planar;

- Euler's theorem that a connected graph admits an Euler tour iff its edge set lies in its cycle space (the infinite analogue of an Euler tour being a closed topological curve in the compacification that traverses every edge exactly once);
- Gallai's theorem that the vertex set of a graph can be partitioned into two sets each inducing an element of its cycle space;
- MacLane's theorem that a graph is planar iff its cycle space has a set of generators such that every edge lies in at most two of these;
- Tutte's theorem that a 3-connected graph is planar iff every edge lies on at most two peripheral cycles;
- the Tutte - Nash-Williams *tree-packing* theorem that a graph has k edge-disjoint spanning trees iff every vertex partition, into ℓ sets say, is crossed by at least $k(\ell - 1)$ edges;
- the *4-colour-theorem* (expressed dually in terms of 4-flows) that the edge set of a planar bridgeless graph is a union of two elements of its cycle space (ie., has a 4-flow).

Furthermore, the following easy facts about the cycle space of a finite graph extend to non-trivial theorems about locally finite graphs with this new cycle space:

- Every element of the cycle space is an edge-disjoint union (not just a sum) of cycles.
- A non-empty set of edges lies in the cycle space iff it meets every finite cut in an even number of edges, and it lies in the cocycle space (ie., is a cut) iff it meets every finite element of the cycle space in an even number of edges.
- The fundamental cycles of any spanning tree generate the cycle space (the generalization is based on *topological spanning trees*, path-connected subspaces containing all the vertices and ends but no continuous 1–1 image of S^1; note that these 'trees' need not induce connected subgraphs, as their path-connectedness can result from topological paths including ends).
- A set of edges lies in the cycle space iff in the subgraph it induces all vertex degrees are even.

The generalization of the last statement involves the definition of 'degrees' also for ends. An end has *degree* k if there are k but not $k + 1$ edge-disjoint infinite paths converging to it. If there is no such k, it has *infinite degree*. Infinite end degrees are also classified into 'odd' and 'even' in a more complicated way, which however is essential for the generalization of the above statement.

The new notion of end degrees motivated by these results seems to open up new possibilities for an 'extremal' branch of infinite graph theory. For example, is there a function $f\colon \mathbb{N} \to \mathbb{N}$ such that every locally finite graph whose vertices and ends all have degree at least $f(k)$ contains a k-connected subgraph? (Note that since infinite trees can have large minimum degree, vertex degrees alone do not force any dense substructures.)

Another natural area of application lies in Hamiltonicity problems. Define a *Hamilton circle* in a graph G as a homeomorphic image of S^1 in its Freudenthal compactification that contains all its vertices. Does every 4-connected planar locally finite graph have a Hamilton circle (extending Tutte's theorem)? Does the square of every 2-connected locally finite graph have a Hamilton circle (extending Fleischner's theorem)?

See [6] for an introductory overview of these results and numerous further problems.

References

[1] H. Bruhn, The cycle space of a 3-connected locally finite graph is generated by its finite and infinite peripheral circuits, *J. Combin. Theory B* (to appear).

[2] H. Bruhn and R. Diestel, Duality in infinite graphs, in preparation.

[3] H. Bruhn, R. Diestel and M. Stein, Cycle-cocycle partitions and faithful cycle covers for locally finite graphs, preprint 2003.

[4] H. Bruhn and M. Stein, MacLane's planarity criterion for locally finite graphs, preprint 2003.

[5] H. Bruhn and M. Stein, Characterizing the cycle space of a locally finite graph by vertex and end degrees, in preparation.

[6] R. Diestel, The cycle space of an infinite graph, *Comb. Probab. Computing* (to appear).
`http://www.math.uni-hamburg.de/home/diestel/papers/CyclesExpository.pdf`

[7] R. Diestel and D. Kühn, On infinite cycles I, *Combinatorica* (to appear).

[8] R. Diestel and D. Kühn, On infinite cycles II, *Combinatorica*, (to appear).

[9] R. Diestel and D. Kühn, Topological paths, cycles and spanning trees in infinite graphs, *Europ. J. Combinatorics* (to appear).

All the above papers are available as preprints at

`http://www.math.uni-hamburg.de/math/research/preprints/hbm.html`

Graph Products, Fourier Analysis and Spectral Techniques

Ehud Friedgut
(joint work with Noga Alon, Irit Dinur and Benny Sudakov)

We consider powers of regular graphs defined by the weak graph product and give a characterization of maximum-size independent sets for a wide family of base graphs which includes, among others, complete graphs, line graphs of regular graphs which contain a perfect matching and Kneser graphs. In many cases this also characterizes the optimal colorings of these products.

We show that the independent sets induced by the base graph are the only maximum-size independent sets. Furthermore we give a qualitative stability statement: any independent set of size close to the maximum is close to some independent set of maximum size.

Our approach is based on Fourier analysis on Abelian groups and on Spectral Techniques. To this end we develop some basic lemmas regarding the Fourier transform of functions on $\{0 \dots r-1\}^n$, generalizing some useful results from the $\{0,1\}^n$ case.

Consider the following combinatorial problem:

Assume that at a given road junction there are n three-position switches that control the red-yellow-green position of the traffic light. You are told that whenever you change the position of *all* the switches then the color of the light changes. Prove that in fact the light is controlled by only one of the switches.

The above problem is a special case of the problem we wish to tackle in this paper, characterizing the optimal colorings and maximal independent sets of products of regular graphs. The configuration space of the switches described above can be modeled by the n-fold product of K_3. Let us begin by defining the weak graph product of two graphs.

The *weak product* of G and H, denoted by $G \times H$ is defined as follows: the vertex set of $G \times H$ is the Cartesian product of the vertex sets of G and H. Two vertices (g_1, h_1) and (g_2, h_2) are adjacent in $G \times H$ if g_1g_2 is an edge of G and h_1h_2 is an edge of H. The "times" symbol, $\times$, is supposed to be reminiscent of the weak product of two edges: $| \times - = \times$. In this paper "graph product" will always mean the weak product.

In the first part of the paper we consider the interesting special case of the product of complete graphs on $r > 2$ vertices,

$$G = K_r^n = \times_{j=1}^n K_r.$$

We then discuss a more general setting, considering other r-regular graphs as well.

When $G = K_r^n$, we identify the vertices of G in the obvious way with the elements of $\mathbb{Z}_r^n$. Recalling the definition of the product, two vertices are adjacent in G iff the corresponding vectors differ in every coordinate. Clearly one can color G with r colors by choosing a coordinate i and coloring every vertex according to its ith coordinate. The following theorem asserts that if $r > 2$ then these are the only r-colorings. Here, and in what follows, we denote by $|H|$ the number of vertices of a graph H.

Theorem 1 *Let $G = K_r^n$, and assume $r \geq 3$. Let I be an independent set with $|I| = |G|/r$. Then there exists a coordinate $i \in \{1 \dots n\}$ and $k \in \{0 \dots r-1\}$ such that*

$$I = \{v : v_i = k\}.$$

Consequently, the only colorings of G by r colors are those induced by colorings of one of the factors K_r.

Greenwell and Lovász [2] proved the above theorem (and actually, a somewhat stronger statement) more than a quarter of a century ago. The novelty in this paper is the proof we supply that uses Fourier analysis on the group $\mathbb{Z}_r^n$. Our approach also allows us to deduce a stability version of the above theorem:

Theorem 2 *For every $r \geq 3$ there exists a constant $M = M(r)$ such that for any $\epsilon > 0$ the following is true. Let $G = K_r^n$. Let J be an independent set such that $\frac{|J|}{|G|} = \frac{1}{r} - \epsilon$. Then there exists an independent set I with $\frac{|I|}{|G|} = \frac{1}{r}$ such that $\frac{|J \triangle I|}{|G|} < M\epsilon$.*

Here "$\triangle$" denotes the symmetric difference. What the above theorem tells us is (in conjunction with Theorem 1) that any independent set that is *close* to being of maximum-size is close to being determined by one coordinate. We do not know of any purely combinatorial proof of this result.

The results in both theorems above can be extended to other base graphs. Let $\alpha(G)$ denote the maximum possible size of an independent set in a graph G. The following observation determines $\alpha(H^n)$ for any vertex transitive base graph H, in terms of $\alpha(H)$ and $|H|$.

Proposition 3 *For any vertex transitive graph H and for any integer $n \geq 1$, if $G = H^n$ then*

$$\frac{\alpha(G)}{|G|} = \frac{\alpha(H)}{|H|}.$$

After the simple proof of this proposition (some special cases of which are proved in [1]), we will provide some examples showing that the above equality does not necessarily hold without the transitivity assumption.

The relevance of graph eigenvalues to independent sets in graphs is well known and can be traced back to the old result that the independence number of any regular graph H on r vertices in which the eigenvalues of the adjacency matrix are $\mu_1 \geq \mu_2 \cdots \geq \mu_r$, is at most $-r\mu_r/(\mu_1 - \mu_r)$. A proof of this fact, as well as of the related results on the connection between the Shannon capacity of a graph and its eigenvalues, can be found in [3]. This bound is tight for many graphs H including, for example, complete graphs and the Petersen graph. It turns out that the results in Theorem 1 and in Theorem 2 can be extended to any connected non-bipartite regular base graph H for which the above bound is tight.

Theorem 4 *Let H be a connected d-regular graph on r vertices and let $d = \mu_1 \geq \mu_2 \geq \cdots \geq \mu_r$ be its eigenvalues. If*

$$\frac{\alpha(H)}{r} = \frac{-\mu_r}{d - \mu_r} \tag{1}$$

then for every integer $n \geq 1$,

$$\frac{\alpha(H^n)}{r^n} = \frac{-\mu_r}{d - \mu_r}.$$

Moreover, if H is also non-bipartite, and if I is an independent set of size $\frac{-\mu_r}{d-\mu_r} r^n$ in $G = H^n$, then there exists a coordinate $i \in \{1, 2, \ldots, n\}$ and a maximum independent set J in H, such that

$$I = \{v \in V(H)^n : \ v_i \in J\}.$$

Remark: Note that for any H and n, $\chi(H^n) = \chi(H)$. If H satisfies the conditions of the last Theorem and if, in addition, $\chi(H) = \frac{r}{\alpha(H)}$ then every optimal coloring of H^n is induced by a coloring of one of the multiplicands, since it is a partition of H^n into maximum-size independent sets. Such a partition can only be consistent if each color class is induced by the same coordinate. The assumption $\chi(H) = \frac{r}{\alpha(H)}$ holds for many of the interesting classes of graphs to which Theorem 4 applies.

Theorem 5 *Let H be a d-regular, connected, non-bipartite graph on r vertices, let $d = \mu_1 \geq \mu_2 \geq \cdots \geq \mu_r$ be its eigenvalues and suppose its independence number satisfies (1). Then, there exists a constant $M = M(H)$ such that for any $\epsilon > 0$ the following holds. Let $G = H^n$ and let I be an independent set such that $\frac{|I|}{|G|} = \frac{\alpha(H)}{|H|} - \epsilon$. Then there exists an independent set I' with $\frac{|I'|}{|G|} = \frac{\alpha(H)}{|H|}$ such that $\frac{|I' \triangle I|}{|G|} < M\epsilon$.*

References

[1] J. I. Brown, R. J. Nowakowski and D. Rall, The ultimate categorical independence ratio of a graph, SIAM J. Discrete Math. 9 (1996), 290-300.

[2] D. Greenwell, L. Lovász, Applications of Product Colorings, Acta Math. Acad. Sci. Hungar. 25 (3-4) (1974) pp. 335-340

[3] L. Lovász, *On the Shannon capacity of a graph*, IEEE Transactions on Information Theory IT-25, (1979), 1-7.

Triple Systems Not Containing a Fano Configuration and other Turán-type Problems

Zoltán Füredi

Given a 3-uniform hypergraph $\mathcal{F}$, let $\mathrm{ex}_3(n, \mathcal{F})$ denote the maximum possible size of a 3-uniform hypergraph of order n that does not contain any subhypergraph isomorphic to $\mathcal{F}$. The **Fano configuration** $\mathbb{F}$ (or Fano plane, or finite projective plane of order 2, or Steiner triple system, $STS(7)$, or blockdesign $S_2(7,3,2)$) is a hypergraph on 7 elements, say $\{x_1, x_2, x_3, a, b, c, d\}$, with 7 edges $\{x_1, x_2, x_3\}$, $\{x_1, a, b\}$, $\{x_1, c, d\}$, $\{x_2, a, c\}$, $\{x_2, b, d\}$, $\{x_3, a, d\}$, $\{x_3, b, c\}$. D. de Caen and Z. Füredi [2] proved a conjecture of Vera T. Sós [11] that

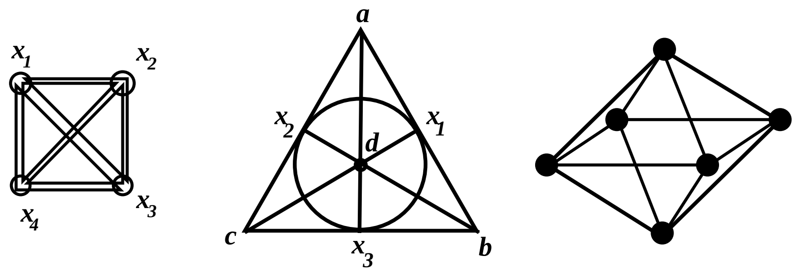
Theorem 1

$$\mathrm{ex}_3(n, \mathbb{F}) = \frac{3}{4}\binom{n}{3} + O(n^2).$$

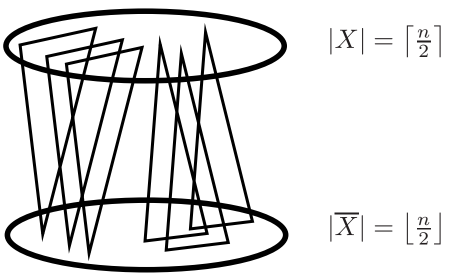
The **tetrahedron**, $K_4^{(3)}$, i.e., a complete 3-uniform hypergraph on four vertices, has four triples $\{x_1, x_2, x_3\}$, $\{x_1, x_2, x_4\}$, $\{x_1, x_3, x_4\}$, $\{x_2, x_3, x_4\}$. An averaging argument shows [7] that the ratio $\mathrm{ex}_3(n, \mathcal{F})/\binom{n}{3}$ is a non-increasing sequence. Therefore

$$\pi(\mathcal{F}) := \lim_{n\to\infty} \mathrm{ex}_3(n, \mathcal{F})/\tbinom{n}{3}$$

exists. The determination of $\pi(K_4^{(3)})$ is one of the oldest problems of this field, due to Turán [12], who published a conjecture in 1961 that this limit value is 5/9, and Erdős [4] offered \$1000 for a proof. The best upper bound, $.5935\ldots$, is due to Fan Chung and Linyuan Lu [3]. The limit $\pi(\mathcal{H})$ is known only for very few cases when it is non-zero.



The Complete 4-graph, the Fano hypergraph, and the Octahedron



T. Sós also conjectured that the following hypergraph, $\mathcal{H}^n$, gives the exact value of $\mathrm{ex}_3(n, \mathbb{F})$. Let $\mathcal{H}(X, \overline{X})$ be the hypergraph obtained by taking the union of two disjoint sets X and $\overline{X}$ as the set of vertices and define the edge set as the set of all triples meeting both X and $\overline{X}$. For $\mathcal{H}^n$ we take $|X| = \lceil n/2 \rceil$ and $|\overline{X}| = \lfloor n/2 \rfloor$, (i.e., they have nearly equal sizes). Then

$$e(\mathcal{H}^n) = \binom{n}{3} - \binom{\lfloor n/2 \rfloor}{3} - \binom{\lceil n/2 \rceil}{3}.$$

The chromatic number of a hypergraph $\mathcal{H}$ is the minimum p such that its vertex set can be decomposed into p parts with no edge contained entirely in a single part. It is well known and easy to check that the Fano plane is not two-colorable, its chromatic number is 3. Therefore $\mathbb{F} \not\subseteq \mathcal{H}(X, \overline{X})$. Thus $\mathcal{H}^n$ supplies the lower bound for $\mathrm{ex}_3(n, \mathbb{F})$ in Theorem 1, implying that $\pi(\mathbb{F}) \geq \frac{3}{4}$.

Theorem 2 (Füredi and Simonovits [6]) *There exist a $\gamma_2 > 0$ and an n_2 such that the following holds. If $\mathcal{H}$ is a triple system on $n > n_2$ vertices not containing the Fano configuration $\mathbb{F}$ and*

$$\deg(x) > \left(\frac{3}{4} - \gamma_2\right)\binom{n}{2}$$

holds for every $x \in V(\mathcal{H})$, then $\mathcal{H}$ is bipartite, $\mathcal{H} \subseteq \mathcal{H}(X, \overline{X})$ for some $X \subseteq V(\mathcal{H})$.

This result is a distant relative of the following classical theorem of Andrásfai, Erdős and T. Sós [1]. Let G be a triangle-free graph on n vertices with minimum degree $\delta(G)$. If $\delta(G) > \frac{2}{5}n$, then G is bipartite. The blow up of a five-cycle C_5 shows that this bound is the best possible.

Using the method of [2] Mubayi and Rödl [9] determined the limit π for a few more 3-uniform hypergraphs, for all of them $\pi = 3/4$. It is very likely that the extremal hypergraphs are 2-colorable in those cases, too.

Turán [12] also conjectured that the 2-colorable triple system $\mathcal{H}^n$ is the largest $K_5^{(3)}$-free hypergraph. Sidorenko [10] disproved this conjecture, in this sharp form, for odd values $n \geq 9$. But it is still conjectured that it is true for all even values and it seems that $\pi(K_5^{(3)}) = 3/4$ holds as well. However this question seems to be extremely difficult.

De Caen and Füredi [2] applied some multigraph extremal results of Füredi and Kündgen [5]. To prove Theorem 2 we use colored multigraph extremal results.

A corollary of Theorem 2, namely that $\mathcal{H}(X, \overline{X})$ is extremal, was proved independently and in a fairly similar way by Keevash and Sudakov [8]. Our Theorem 2 is stronger.

References

[1] Andrásfai, B., Erdős, P. and T. Sós, V. (1974) On the connection between chromatic number, maximal clique and minimal degree of a graph. *Discrete Math.* **8** 205-218.

[2] de Caen, D. and Füredi, Z. (2000) The maximum size of 3-uniform hypergraphs not containing a Fano plane. *J. Combin. Theory Ser. B* **78** 274–276.

[3] Chung, F. R. K. and Lu, Linyuan (1999) An upper bound for the Turán number $t_3(n, 4)$. *J. Combin. Theory Ser. A* **87** 381–389.

[4] Erdős, P. (1981) On the combinatorial problems which I would most like to see solved. *Combinatorica* **1** 25–42.

[5] Füredi Z. and Kündgen, A. (2002) Turán problems for integer-weighted graphs. *J. Graph Theory* **40** 195–225.

[6] Füredi Z. and Simonovits, M. (2004) Triple systems not containing a Fano configuration. *Combin. Prob. and Comput.* to appear.

[7] Katona, G., Nemetz, T. and Simonovits, M. (1964) On a problem of Turán in the theory of graphs. *Mat. Lapok* **15** 228–238.

[8] Keevash, P. and Sudakov, B. The exact Turán number of the Fano plane. To appear in *Combinatorica*.

[9] Mubayi, D. and Rödl, V. (2002) On the Turán number of triple systems. *J. Combin. Th. Ser. A* **100** 136–152.

[10] Sidorenko, A. (1995) What we know and what we do not know about Turán numbers. *Graphs and Combin.* **11** 179–199.

[11] T. Sós, V. (1976) Some remarks on the connection of graph theory, finite geometry and block designs. *in:* Teorie Combinatorie, Tomo II., Accad. Naz. Linzei, Roma, pp. 223–233.

[12] Turán, P. (1961) Research problems. *MTA Mat. Kutató Int. Közl.* **6** 417-423.

Entropy and Graph Homomorphisms

David Galvin
(joint work with Prasad Tetali [3])

Let G be an n-regular, N-vertex bipartite graph on vertex set $V(G)$, and let H be a fixed graph on vertex set $V(H)$ (perhaps with loops). Set

$$Hom(G,H) = \{f : V(G) \to V(H) \;:\; u \sim v \Rightarrow f(u) \sim f(v)\}.$$

That is, $Hom(G,H)$ is the set of graph homomorphisms from G to H.

When $H = H_{ind}$ consists of one looped and one unlooped vertex connected by an edge, an element of $Hom(G, H_{ind})$ can be thought of as a specification of an independent set (a set of vertices spanning no edges) in G. Our point of departure is the following result of Kahn [4], bounding the size of $\mathcal{I}(G)$, the set of independent sets of G.

Theorem 1 *For any n-regular, N-vertex bipartite graph G,*

$$|\mathcal{I}(G)| \leq (2^{n+1} - 1)^{N/2n}.$$

Note that $|Hom(K_{n,n}, H_{ind})| = 2^{n+1} - 1$ (where $K_{n,n}$ is the complete bipartite graph with n vertices on each side), so we may paraphrase Theorem 1 by saying that $|Hom(G, H_{ind})|$ is maximum when G is a disjoint union of $K_{n,n}$'s. Our main result is a generalization of this statement (and our proof is a generalization of Kahn's).

Proposition 2 *For any n-regular, N-vertex bipartite G, and any H,*

$$|Hom(G,H)| \leq |Hom(K_{n,n}, H)|^{N/2n}.$$

We also consider a weighted version of Proposition 2. Following [1], we put a measure on $Hom(G, H)$ as follows. To each $i \in V(H)$ assign a positive "activity" λ_i, and write Λ for the set of activities. Give each $f \in Hom(G, H)$ weight $w^{\Lambda}(f) = \prod_{v \in V(G)} \lambda_{f(v)}$. The constant that turns this assignment of weights on $Hom(G, H)$ into a probability distribution is

$$Z^{\Lambda}(G, H) = \sum_{f \in Hom(G,H)} w^{\Lambda}(f).$$

When all activities are 1, we have $Z^{\Lambda}(G, H) = |Hom(G, H)|$, and so the following is a generalization of Proposition 2.

Proposition 3 *For any n-regular, N-vertex bipartite G, any H, and any system Λ of positive activities on $V(H)$,*

$$Z^{\Lambda}(G, H) \leq \left(Z^{\Lambda}(K_{n,n}, H)\right)^{N/2n}.$$

We may put this result in the framework of a well-known mathematical model of physical systems with "hard constraints" (see [1]). We think of the vertices of G as particles and the edges as bonds between pairs of particles, and we think of the vertices of H as possible "spins" that particles may take. Pairs of bonded vertices of G may have spins i and j only when i and j are adjacent in H. Thus the legal spin configurations on the vertices of G are precisely the homomorphisms from G to H. We think of the activities on the vertices of H as a measure of the likelihood of seeing the different spins; the probability of a particular spin configuration is proportional to the product over the vertices of G of the activities of the spins. Proposition 3 concerns the "partition function" of this model — the normalizing constant that turns the above-described system of weights on the set of legal configurations into a probability measure.

Our proofs are based on entropy considerations, and in particular on a lemma of Shearer (see [2, p. 33]) bounding the entropy of a random vector.

References

[1] G. Brightwell and P. Winkler, Graph homomorphisms and phase transitions, *J. Combin. Theory Ser. B* **77** (1999), 221–262.

[2] F.R.K. Chung, P. Frankl, R. Graham and J.B. Shearer, Some intersection theorems for ordered sets and graphs, *J. Combin. Theory Ser. A.* **48** (1986), 23–37.

[3] D. Galvin and P. Tetali, On weighted graph homomorphisms, to appear in AMS volume on DIMACS/DIMATIA workshop *Graphs, Morphisms and Statistical Physics*, March 2001.

[4] J. Kahn, An entropy approach to the hard-core model on bipartite graphs, *Combin. Prob. Comp.* **10** (2001), 219–237.

Random Planar Graphs

Stefanie Gerke
(joint work with Colin McDiarmid [3])

Given $0 < p < 1$ and a positive integer n, let $G_{n,p}$ denote the random graph with nodes $v_1, \ldots, v_n$ in which the $\binom{n}{2}$ possible edges appear independently with probability p. We denote by $R_{n,p}$ the random graph $G_{n,p}$ conditioned on it being planar. (We may think of repeatedly sampling a graph $G_{n,p}$ until we find one that is planar.) Also, let us denote $R_{n,\frac{1}{2}}$ by R_n. Thus R_n is uniformly distributed over all labelled planar graphs on n nodes.

Rather little is known about random planar graphs, even about the number of edges in such graphs, which is our focus here. Let us denote the number of edges in a (simple) graph G by $m(G)$. Thus we are interested in the random variable $m(R_n)$ and more generally in $m(R_{n,p})$. Of course $m(G) \leq 3n - 6$ for any planar graph G on n nodes. The expected value $\mathbf{E}[m(R_n)]$ is at least $(3n-6)/2$ – see [2]. It is shown in [1] that $m(R_n) \leq 2.54n$ asymptotically almost surely (aas), that is with probability tending to 1 as $n \to \infty$. This result slightly improves the upper bound of 2.56 in [6]. We will show here in particular that $m(R_n) \geq \frac{13}{7}n + o(n)$ aas, thereby improving on the result from [2] mentioned above.

We now introduce two functions $f(\alpha)$ and $g(p)$ which are needed to state our two main results – see also Figure 1.

Given $1 < \alpha \leq 3$, let $k = k(\alpha) = \lfloor \frac{2\alpha}{\alpha-1} \rfloor$, and let

$$f(\alpha) = \frac{1}{4}\left(k^2 + k + 6 - (k^2 - 3k + 6)\alpha\right).$$

It is not hard to verify that $f(\alpha)$ is continuous and decreasing on $1 < \alpha \leq 3$, and satisfies $f(\alpha) \to \infty$ as $\alpha \to 1$ and $f(3) = 0$, see also the end of Section 4. (The function f is also piecewise-linear and convex.) For $0 < p < 1$ we may define $g(p)$ to be the unique value $\rho \in (1,3)$ such that $f(\rho)/\rho = (1-p)/p$. The function g is continuous and increasing on $0 < p < 1$, and satisfies $g(p) \to 1$ as $p \to 0$, $g(\frac{1}{2}) = \frac{13}{7}$ and $g(p) \to 3$ as $p \to 1$. We are now able to state our theorem concerning the number of edges of random planar graphs.

Theorem 1 *Let $0 < p < 1$. Then as $n \to \infty$,*

$$\mathbf{E}[m(R_{n,p})] \geq g(p)n + o(n);$$

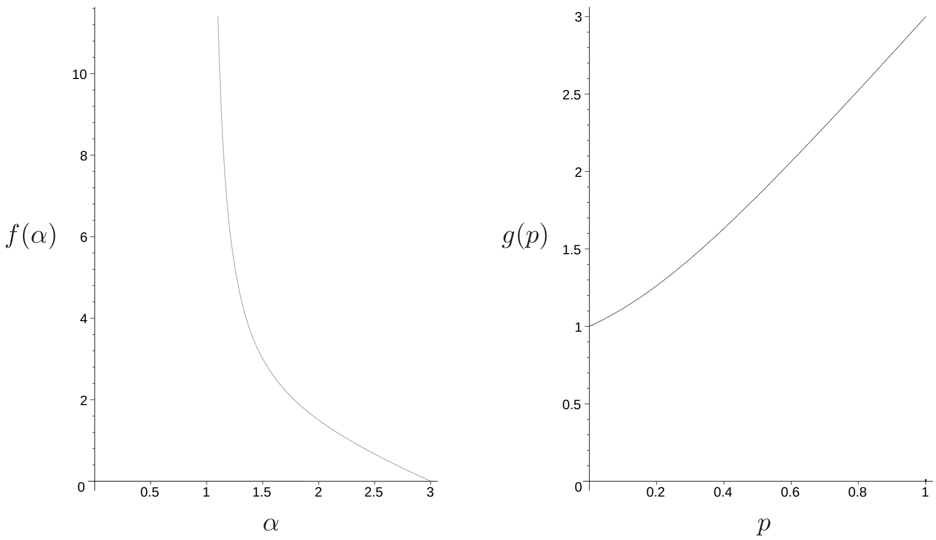


Figure 1: The functions f and g

and indeed for any $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\mathbf{Pr}(m(R_{n,p}) < (g(p) - \varepsilon)n) = o(e^{-\delta n}).$$

In particular, since $g(\frac{1}{2}) = \frac{13}{7}$, this theorem shows that the expected number of edges in a planar graph sampled uniformly at random from all labelled planar graphs on n nodes is at least about $\frac{13}{7}n$.

To prove this result we will consider the number of edges that can be added to a planar graph of n nodes and m edges while keeping the graph planar. Given a planar graph G, we call a non-edge f *addable* in G if the graph $G + f$ obtained by adding f as an edge is still planar; and we let $\text{add}(G)$ denote the set of addable non-edges of G. Let $\mathcal{P}(n)$ denote the set of all (simple) planar graphs with n nodes $v_1, \ldots, v_n$; let $\mathcal{P}(n, m)$ denote the set of all graphs $G \in \mathcal{P}(n)$ with m edges; and let $\text{add}(n, m)$ denote the minimum value of $|\text{add}(G)|$ over all graphs $G \in \mathcal{P}(n, m)$. Observe that by Kuratowski's theorem, if $m \leq 7$ then $\text{add}(n, m) = \binom{n}{2} - m$, and if $n \geq 6$ and $m \geq 8$ then $\text{add}(n, m) < \binom{n}{2} - m$. Also, $\text{add}(n, m) > 0$ if $m < 3n - 6$ and $\text{add}(n, 3n - 6) = 0$.

Theorem 2 *Let $1 < \alpha \leq 3$, and suppose that $m = m(n) = \alpha n + O(1)$ as $n \to \infty$. Then add$(n, m) = f(\alpha)n + O(1)$.*

It was shown in [5] that a.a.s. the random planar graph contains any fixed connected planar graph. If one chooses a graph uniformly at random from $\mathcal{P}(n, m)$ with $m = \lfloor qn \rfloor$ $1 < q < 3$ then the same statements holds:

Theorem 3 (G., McDiarmid, Steger, Weißl [4]) *Let $1 < q < 3$. Then a.a.s. the random planar graph on n nodes and $\lfloor qn \rfloor$ edges contains any fixed connected planar graph.*

References

[1] N. Bonichon, C. Gavoille and Nicolas Hanusse, *An information-theoretic upper bound on planar graphs using triangulation,* in Proc. of the 20th Annual Symp. of Theoret. Comp. Science (2003) 499-510.

[2] A. Denise, M. Vasconcellos and D. Welsh, *The random planar graph,* Congr. Numer., **113** (1996) 61–79.

[3] S. Gerke and C. McDiarmid, On the number of edges in random planar graphs, Combin. Probab. Comput., to appear.

[4] S. Gerke, C. McDiarmid, A. Steger and A. Weissl, Random planar graphs with a fixed number of edges, manuscript.

[5] C. McDiarmid, A. Steger and D. Welsh, *Random planar graphs*, manuscript.

[6] D. Osthus, H.J. Prömel and A. Taraz, *On random planar graphs, the number of planar graphs and their triangulations,* J. Combin. Theory Ser. B, **88**(1) (2003) 119–134.

Low-dimensional Faces of Random 0/1-Polytopes

Volker Kaibel

Investigations of special classes of 0/1-polytopes (convex hulls of subsets of $\{0,1\}^d$) have not only lead to beautiful insights into combinatorial (optimization) problems during the last decades, but also powerful algorithms have emerged from them. Consequently, there has been some desire to learn more about the geometrical and combinatorial structure of 0/1-polytopes in general. Here, the study of random 0/1-polytopes has turned out to be particularly fruitful,

A quite fascinating result in this direction has been obtained by Dyer, Füredi, and McDiarmid in 1992, who proved in [2] that the expected volume $\mathbb{E}[\text{Vol } P]$ of a d-dimensional random 0/1-polytope P with n vertices has a threshold at $2^{(1-(\log e)/2)d}$ (i.e., for each $\varepsilon > 0$, $\text{Vol } P = \text{o}(1)$ if $n \leq 2^{(1-(\log e)/2-\varepsilon)d}$ and $\text{Vol } P =$

$1 - \mathrm{o}(1)$ if $n \geq 2^{(1-(\log e)/2+\varepsilon)d}$). Building on the methods developed in Dyer, Füredi, and McDiarmid' work, Bárány and Pór proved in 2000 that a random 0/1-polytope (within a certain range of vertex numbers) has a super-exponential (in the dimension) number of facets [1].

While Bárány and Pór's work sheds some light on the highest dimensional faces of 0/1-polytopes, in my recent work (partly together with Anja Remshagen) I have investigated the lowest dimensional faces of random 0/1-polytopes. In [4] we proved that the expected graph density of a d-dimensional random 0/1-polytope P with n vertices has a threshold at $2^{(1/2)d}$. In [3] this result has been extended to the density of arbitrary (fixed) dimensional faces in the following way.

Denote by $\nu_r(P)$ the quotient of the number of faces of P with exactly r vertices and $\binom{n}{r}$ (the *r-density* of P). In [3], for each $r \geq 3$, we establish the existence of a sharp threshold for the r-density and determine the values of the threshold numbers τ_r such that, for all $\varepsilon > 0$,

$$\mathbb{E}\left[\nu_r(P)\right] \; = \; \begin{cases} 1 - \mathrm{o}(1) & \text{if } n \leq 2^{(\tau_r-\varepsilon)d} \text{ for all } d \\ \mathrm{o}(1) & \text{if } n \geq 2^{(\tau_r+\varepsilon)d} \text{ for all } d \end{cases}$$

holds for the expected value of $\nu_r(P)$.

In particular, these results indicate that the high densities often encountered in polyhedral combinatorics (e.g., the cut-polytope of the complete graph has both 2- and 3-density equal to one) are due to the geometry of 0/1-polytopes rather than to the special combinatorics of the underlying problems.

The threshold values τ_r (for $r \geq 3$) nicely extend the results for $r = 2$, while the proof becomes more involved and needs a heavier machinery (the one developed in the above mentioned paper by Dyer, Füredi, and McDiarmid). As a pay-back, however, it reveals several interesting insights into the geometry of (random) 0/1-polytopes.

References

[1] Imre Bárány and Attila Pór, *On* 0-1 *polytopes with many facets*, Adv. Math. **161** (2001), no. 2, 209–228.

[2] Martin E. Dyer, Zoltan Füredi, and Colin McDiarmid, *Volumes spanned by random points in the hypercube*, Random Structures Algorithms **3** (1992), no. 1, 91–106.

[3] Volker Kaibel, *Low-dimensional faces of random 0/1-polytopes*, Technical Report, TU Berlin, arXiv: math.CO/0311393 (extended abstract, 10 pages +3 pages appendix).

[4] Volker Kaibel and Anja Remshagen, *On the graph-density of random 0/1-polytopes*, Approximation, Randomization, and Combinatorial Optimization (Proc. RANDOM 2003) (S. Arora, K. Jansen, J.D.P. Rolim, and A. Sahai, eds.), Lecture Notes in Computer Science, vol. 2764, Springer, 2003, pp. 318–328.

Excluded Subposets in the Boolean Lattice

Gyula O.H. Katona

Introduction. Let $[n] = \{1, 2, \ldots, n\}$ be a finite set, families $\mathcal{F}, \mathcal{G}$, etc. of its subsets will be investigated. If $\mathcal{F}$ is a family let f_i denote the number of its i-element members. Let P be a poset. The goal of the present investigations is to determine the maximum size of a family $\mathcal{F}$ (in $[n]$) which does not contain P as a (non-necessarily induced) subposet. This maximum is denoted by $\mathrm{La}(n, P)$.

The easiest example is the case when P consist of two comparable elements (subsets of [1]). Then we are actually looking for the largest family without inclusion. The well-known Sperner theorem ([6]) gives the answer, the maximum is $\binom{n}{\lfloor \frac{n}{2} \rfloor}$.

The following sharpening, the so called YBLM inequality ([8], [1], [4], [5]) is also important.

Theorem 1 *If $\mathcal{F}$ is a family of subsets of $[n]$ without inclusion then*

$$\sum_{i=0}^{n} \frac{f_i}{\binom{n}{i}} \leq 1$$

holds

We say that the distinct sets $A, B_1 \ldots, B_r$ form an *r-fork* if they satisfy $A \subset B_1, \ldots, B_r$.

The first result in this direction of generalizing the Sperner theorem was the following one ([3]).

Theorem 2 *Suppose that $\mathcal{F}$ contains no 2-fork. Then*

$$\binom{n}{\lfloor \frac{n}{2} \rfloor} \left(1 + \frac{1}{n} + o\left(\frac{1}{n}\right)\right) \leq |\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor} \left(1 + \frac{2}{n} + o\left(\frac{1}{n}\right)\right)$$

holds.

The case of $r+1$-forks was considered in [7].

Theorem 3 *Suppose that $\mathcal{F}$ contains no $r+1$-fork. Then*

$$\binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1+\frac{r}{n}+o\left(\frac{1}{n}\right)\right) \leq |\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1+2\frac{r^2}{n}+o\left(\frac{1}{n}\right)\right)$$

holds.

Let us remark that the lower estimates in the previous results and in the new results of the next section are all based on a code construction of [2] and its generalizations.

New results. The second term of the upper estimate is too weak in Theorem 1.3. We were recently able to improve this result.

Theorem 4 (A. de Bonis, G.O.H. Katona) *Suppose that $\mathcal{F}$ contains no $r+1$-fork. Then*

$$\binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1+\frac{r}{n}+o\left(\frac{1}{n}\right)\right) \leq |\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1+\frac{2r}{n}+o\left(\frac{1}{n}\right)\right).$$

This is best possible in the sense that the coding problem what is used in the construction contains an undecided multiplicative factor 2.

The proof of the upper bound in the above theorem is based on the following YBLM-type inequality.

Theorem 5 (A. de Bonis, G.O.H. Katona) *Suppose that $\mathcal{F}$ contains no $r+1$-fork $(0<r)$ and all members $F \in \mathcal{F}$ satisfy $|F| \leq m$. Then*

$$\sum_{i=0}^{n} \frac{f_i}{\binom{n}{i}} \leq 1+\frac{r}{n-m+1}.$$

Let us now try to maximize the size of a family $\mathcal{F}$ containing no $r+s+1$ distinct members satisfying $A_1,\dots,A_s \subset B_1,\dots,B_{r+1}$. Let $P_{r+1,s}$ denote the poset with two levels, s element on the lower, $r+1$ elements on the upper level, every lower one is in relation with every upper one. It is easy to see that our condition can be formulated in the way that we are looking for the maximum number of the elements in the Boolean lattice of subsets of $[n]$ (defined by inclusion) without containing $P_{r+1,s}$ as a subposet.

Theorem 6 (A. de Bonis, G.O.H. Katona) *Suppose that $2 \leq s, 2 \leq r$ and $s \leq r+1$ hold. Then*

$$\begin{aligned}\binom{n}{\lfloor \frac{n}{2} \rfloor}\left(2+\frac{r}{n}+o\left(\frac{1}{n}\right)\right) &\leq \mathrm{La}(n, P_{r+1,s}) \\ &\leq \binom{n}{\lfloor \frac{n}{2} \rfloor}\left(2+2\frac{r+s-2}{n}+o\left(\frac{1}{n}\right)\right).\end{aligned}$$

Surprisingly, we have an exact result in the case $s = 2, r = 1$.

Theorem 7 (A. de Bonis, G.O.H. Katona, K. Swanepoel) *If* $5 \leq n$ *and* $\mathcal{F}$ *contains no four distinct members* A_1, A_2, B_1, B_2 *such that* $A_i \subset B_j, i, j = 1, 2$ *then the maximum of* $|\mathcal{F}|$ *is the sum of the two largest binomial coefficients of order* n.

References

[1] B. Bollobás, On generalized graphs, *Acta Math. Acad. Sci. Hungar* **16**(1965), 447–452.

[2] R.L. Graham and H.J.A. Sloane, Lower bounds for constant weight] codes, *IEEE IT* **26** 37-43.

[3] G.O.H. Katona and T. Tarján, Extremal problems with excluded subgraphs in the n-cube, *Lecture Notes in Math.* **1018**, 84-93.

[4] Lubell, A short proof of Sperner's lemma, *J. Combin. Theory* **1**(1966), 299.

[5] L.D. Meshalkin, A generalization of Sperner's theorem on the number of a finite set (in Russian), *Teor. Verojatnost. i Primen.* **8**(1963), 219–220.

[6] E. Sperner, Ein Satz über Untermegen einer endlichen Menge, *Math. Z.* **27**(1928) 544–548.

[7] Hai Tran Thanh, An extremal problem with excluded subposets in the Boolean lattice, *Order* **15**(1998) 51-57.

[8] K. Yamamoto, Logarithmic order of free distributive lattices, *J. Math. Soc. Japan* **6**(1954) 347–357.

Local Chromatic Number and Sperner Capacity

János Körner

(joint work with Concetta Pilotto and Gábor Simonyi)

Colouring the vertices of a graph so that no adjacent vertices receive identical colours gives rise to many interesting problems and invariants. The best known among all these invariants is the chromatic number, the minimum number of colours needed for such proper colourings. the following interesting variant was introduced by Erdős, Füredi, Hajnal, Komjáth, Rödl, and Seress [5] (cf. also [7]).

Definition 1 ([5]) *The local chromatic number $\psi(G)$ of a graph G is the maximum number of different colours appearing in the closed neighbourhood of any vertex, minimized over all proper colourings of G. Formally,*

$$\psi(G) := \min_{c:\ V(G)\to N} \max_{v\in V(G)} |\{c(u) : u \in \Gamma_G(v)\}|,$$

where N is the set of natural numbers, $\Gamma_G(v)$, the closed neighborhood of the vertex $v \in V(G)$, is the set of those vertices of G that are either adjacent or equal to v and $c : V(G) \to N$ runs over those functions that are proper colourings of G.

It was proved in [5] that there exist graphs with $\psi(G) = 3$ and $\chi(G)$ arbitrarily large.

Throughout our paper [12] the present extended abstract is referring to, we are interested in chromatic invariants as upper bounds for the Shannon capacity of undirected graphs and its natural generalization Sperner capacity for directed graphs. We treat Shannon capacity in terms that are complementary to Shannon's own, (cf. [15], [14] and [9], [11]). In this language Shannon capacity describes the asymptotic growth of the clique number in the co-normal powers of a graph. Shannon proved (although in different terms) that the Shannon capacity $c(G)$ of a graph is upper bounded by its fractional chromatic number.

We show that $\psi(G)$ is bounded from below by the fractional chromatic number of G. This proves, among other things, that $\psi(G)$ is always an upper bound for the Shannon capacity $c(G)$ of G, but it is not a very useful one since it is always weaker than the fractional chromatic number itself. Thus the situation is rather different in the case of directed graphs. We introduce an analog of the local chromatic number for directed graphs and show that it is always an upper bound for the Sperner capacity of the digraph at hand. To illustrate the usefulness of this bound we apply it to show, for example, that an oriented odd cycle with at least two vertices with outdegree and indegree 1 always has its Sperner capacity equal to that of the single-edge graph K_2. We introduce fractional versions that further strengthen our bounds.

The definition of the directed version of $\psi(G)$ is straightforward.

Definition 2 *The local chromatic number $\psi_d(G)$ of a digraph G is the maximum number of different colours appearing in the closed out-neighbourhood of any vertex, minimized over all proper colourings of G. Formally,*

$$\psi_d(G) := \min_{c:\ V(G)\to N} \max_{v\in V(G)} |\{c(w) : w \in \Gamma_G^+(v)\}|$$

where N is the set of natural numbers, $\Gamma_G^+(v)$, the closed out-neighbourhood of the vertex $v \in V(G)$, is the set of those vertices $w \in V(G)$ that are either equal to v or else are endpoints of directed edges $(v, w) \in E(G)$, originated in v, and $c : V(G) \to N$ runs over those functions that are proper colourings of G.

Our main goal is to prove that $\psi_d(G)$ is an upper bound for the Sperner capacity of digraph G.

Definition 3 *For directed graphs $G = (V, E)$ and $H = (W, L)$, the co-normal (or disjunctive or OR) product $G \cdot H$ is defined to be the following directed graph:*

$$V(G \cdot H) = V \times W$$

and

$$E(G \cdot H) = \{((v, w), (v', w')) : (v, v') \in E \quad \text{or} \quad (w, w') \in L\}.$$

The n^{th} co-normal (or disjunctive or OR) power G^n of digraph G is defined as the n-fold co-normal product of G with itself, i. e., the vertex set of G^n is $V^n = \{\mathbf{x} = (x_1 \dots x_n) : x_i \in V\}$, while its edge set is defined as

$$E(G^n) = \{(\mathbf{x}, \mathbf{y}) : \exists i \ (x_i, y_i) \in E(G)\}.$$

(A pair (a, b) always means an oriented edge in this paper as opposed to undirected edges denoted by $\{a, b\}$.)

Definition 4 ([9])
A subgraph of a digraph is called a symmetric clique if its edge set contains all ordered pairs of vertices belonging to the subgraph and we denote the size (number of vertices) of the largest symmetric clique by $\omega_s(G)$. The (non-logarithmic) Sperner capacity of a digraph G is defined as

$$\sigma(G) = \sup_n \sqrt[n]{\omega_s(G^n)}.$$

It is obvious that Sperner capacity is a generalization of Shannon capacity. It is a true generalization in the sense that there exist digraphs the Sperner capacity of which is different from the Shannon capacity ($c(G)$ value) of its underlying undirected graph. Denoting by G both an arbitrary digraph and its underlying undirected graph, it follows from the definitions that $\sigma(G) \leq c(G)$ always holds. The smallest example with strict inequality in the previous relation is a cyclically oriented triangle, cf. [4], [3].

Shannon capacity is is difficult to determine, and it is unknown for many relatively small and simple graphs, for example, for all odd cycles of length at least 7. This shows that Sperner capacity cannot be easy to determine either. There is an interesting and important connection between Sperner capacity and extremal set theory, introduced in [13] and fully explored in [10]. Several problems of this flavour are also discussed in [11].

Alon [1] proved that for any digraph G

$$\sigma(G) \leq \min\{\Delta_+(G), \Delta_-(G)\} + 1$$

where $\Delta_+(G)$ is the maximum out-degree of the graph G and similarly $\Delta_-(G)$ is the maximum in-degree. The proof relies on a linear algebraic method similar to the one already used in [3] for a special case (cf. also [6] for a strengthening and cf. [2] for a general setup for this method in case of undirected graphs). We also use this method for proving the following stronger result.

Theorem 5

$$\sigma(G) \leq \psi_d(G).$$

We call an oriented cycle alternating if it has at most one vertex of outdegree 1. (In stating the following results we follow the convention that an oriented graph is a graph without oppositely directed edges between the same two points, while a general directed graph may contain such pairs of edges.) Clearly, in any oriented cycle the number of vertices of outdegree 2 equals the number of vertices of outdegree 0. Thus, in particular, a $2k+1$ length oriented odd cycle is alternating if it has k points of outdegree zero, k points of outdegree 2 and only 1 point of outdegree 1. It takes an easy checking that up to isomorphism there is only one orientation of C_{2k+1} which is alternating.

Theorem 6 *Let G be an oriented odd cycle that is not alternating. Then*

$$\sigma(G) = 2.$$

The Sperner capacity of an alternating odd cycle can indeed be larger than 2. This is obvious for C_3, where the alternating orientation produces a transitive clique of size 3. A construction proving that the Sperner capacity of the alternating C_5 is at least $\sqrt{5}$ is given in [8]. The construction is clearly best possible by the celebrated result of Lovász [14] showing $c(C_5) = \sqrt{5}$.

References

[1] N. Alon, On the capacity of digraphs, *European J. Combin.*, **19** (1998), 1–5.

[2] N. Alon, The Shannon capacity of a union, *Combinatorica*, **18** (1998), 301–310.

[3] A. Blokhuis, On the Sperner capacity of the cyclic triangle, *J. Algebraic Combin.*, **2** (1993), 123–124.

[4] R. Calderbank, P. Frankl, R. L. Graham, W. Li, L. Shepp, The Sperner capacity of the cyclic triangle for linear and non-linear codes, *J. Algebraic Combin.*, **2** (1993), 31–48.

[5] P. Erdős, Z. Füredi, A. Hajnal, P. Komjáth, V. Rödl, Á. Seress, Coloring graphs with locally few colors, *Discrete Math.*, **59** (1986), 21–34.

[6] E. Fachini and J. Körner, Colour number, capacity and perfectness of directed graphs, *Graphs Combin.*, **16** (2000), 389–398.

[7] Z. Füredi, Local colourings of graphs (Gráfok lokális színezései), manuscript in Hungarian, September 2002.

[8] A. Galluccio, L. Gargano, J. Körner, G. Simonyi, Different capacities of a digraph, *Graphs Combin.*, **10** (1994), 105–121.

[9] L. Gargano, J. Körner, U. Vaccaro, Sperner theorems on directed graphs and qualitative independence, *J. Combin. Theory Ser. A*, **61** (1992), 173–192.

[10] L. Gargano, J. Körner, U. Vaccaro, Capacities: from information theory to extremal set theory, *J. Combin. Theory Ser. A*, **68** (1994), 296–316.

[11] J. Körner, A. Orlitsky, Zero-error information theory, *IEEE Trans. Inform. Theory*, **44**, No. 6 (October 1998, commemorative issue), 2207–2229.

[12] J. Körner, C. Pilotto, G. Simonyi, Local chromatic number and Sperner capacity, *J. Combin. Theory, Ser. B*

[13] J. Körner, G. Simonyi, A Sperner–type theorem and qualitative independence, *J. Combin. Theory Ser. A*, **59** (1992), pp. 90–103,

[14] L. Lovász, On the Shannon capacity of a graph, *IEEE Trans. Inform. Theory*, **25** (1979), 1–7.

[15] C. E. Shannon, The zero-capacity of a noisy channel, *IRE Trans. Inform. Theory*, **2** (1956), 8–19.

On H-linked Graphs

Alexandr Kostochka
(joint work with Gexin Yu)

In this talk, we introduce the notion of H-linked graphs and find suffiicient minimum degree conditions for a graph to be H-linked. This improves known conditions for a graph to be k-ordered.

Let H be a graph. An *H-subdivision* in a graph G is a pair of mappings $f : V(H) \to V(G)$ and g: $E(H)$ into the set of paths in G such that:

(a) $f(u) \neq f(v)$ for all distinct $u, v \in V(H)$;

(b) for every $uv \in E(H)$, $g(uv)$ is an $f(u)f(v)$-path in G, and distinct edges map into internally disjoint paths in G.

Say that a graph G is *H-linked* if every injective mapping $f : V(H) \to V(G)$ can be extended to an H-subdivision in G. This is a natural generalization of k-linkage.

Recall that a graph is *k-linked* if for every list of $2k$ vertices

$$\{s_1, \dots, s_k, t_1, \dots, t_k\} ,$$

there exist internally disjoint paths $P_1, \dots, P_k$ such that each P_i is an s_i, t_i-path. From the definitions of k-linked and H-linked graphs, we immediately see that *a graph G is k-linked if and only if G is H-linked for every graph H with $|E(H)| = k$.*

It is known that to check that a graph on at least $2k$ vertices is k-linked it is enough to check only the lists $\{s_1, \dots, s_k, t_1, \dots, t_k\}$, where all s_i and t_i are distinct. Thus, *a graph G on at least $2k$ vertices is k-linked if and only if G is M_k-linked, where M_k is the matching with k edges.*

Let B_k denote the (multi)graph with 2 vertices and k parallel edges. By Menger's Theorem, *a graph G on at least $k + 1$ vertices is k-connected if and only if G is B_k-linked.*

A graph is *k-ordered*, if for every ordered sequence of k vertices, there is a cycle that encounters the vertices of the sequence in the given order. Let C_k denote the cycle of length k. Clearly, *a graph G is k-ordered if and only if G is C_k-linked.*

After Chartrand introduced the notion of k-ordered graphs, several authors (see, e.g., [4, 8, 7, 5]) studied sufficient degree conditions for a graph to be k-ordered. Recall that Dirac [2] found sufficient conditions for a simple graph G to be Hamiltonian in terms of the minimum degree, $\delta(G)$, and Ore [9] found similar conditions in terms of $\sigma_2(G)$, the minimum value of the sum $\deg(u) + \deg(v)$ over all pairs $\{u, v\}$ of non-adjacent vertices in G. Let $D_0(n, k)$ denote the minimum positive integer d such that every n-vertex simple graph with minimum degree at least d is k-ordered. Similarly, let $R_0(n, k)$ denote the minimum positive integer r such that every n-vertex simple graph G with $\sigma_2(G) \geq r$ is k-ordered. Improving on results in [4, 8], it was shown in [5] that $R_0(n, k) = n + \lceil (3k - 9)/2 \rceil$ for every $3 \leq k \leq n/2$. This implies that $D_0(n, k) \leq \lceil (2n + 3k - 9)/4 \rceil$ for every $3 \leq k \leq n/2$. Moreover, Kierstead et al. [7] showed that $D_0(n, k) = \lceil \frac{n}{4} \rceil + \lfloor \frac{k}{2} \rfloor - 1$ for $3 \leq k \leq \frac{n+3}{11}$. Observe that these bounds demonstrate the interesting phenomenon: $R_0(n, k) > 2D_0(n, k)$ for k small with respect to n. It is also known that $D_0(n, k) > \lceil \frac{n}{4} \rceil + \lfloor \frac{k}{2} \rfloor - 1$ for $k > n/3$, but the value of $D_0(n, k)$ was not known for $\frac{n+3}{11} < k < \frac{2n}{5}$.

The main result of the talk gives the minimum degree conditions for a graph to be H-linked if $\delta(H) \geq 2$. This results extends the result of Kierstead et al. [7] in two directions: for a larger scope of k and for much more general H.

Theorem 1 *Let H be a simple graph with k edges and $\delta(H) \geq 2$. Every graph G of order $n \geq 5k$ with $\delta(G) \geq \lceil (n+k)/2 \rceil - 1$ is H-linked. If H is the cycle C_k with k edges, then every graph G of order $n \geq 5k$ with $\delta(G) \geq \lceil n/2 \rceil + \lfloor k/2 \rfloor - 1$ is H-linked. The minimum degree conditions are sharp.*

In particular, Theorem 1 yields $D_0(n,k) = \lceil \frac{n}{4} \rceil + \lfloor \frac{k}{2} \rfloor - 1$ for $k \leq n/5$.

Note that $\delta(G) \geq \lceil (n+k)/2 \rceil - 1$ is exactly the minimum degree condition that provides the k-connectivity of G. Thus, an evident degree condition for a graph to be k-connected, provides that a graph is H-linked for many H. If one drops the condition $\delta(H) \geq 2$, then this degree restriction is not sufficient in general. In a joint work with Kawarabayashi [6], we considered similar problem for k-linked graphs. Let $D(n,k)$ be the minimum positive integer d such that every n-vertex graph with minimum degree at least d is k-linked. Also, let $R(n,k)$ denote the minimum positive integer r such that every n-vertex graph G with $\sigma_2(G) \geq r$ is k-linked.

Theorem 2 [6] *If $k \geq 2$, then*

$$R(n,k) = \begin{cases} 2n-3, & n \leq 3k-1; \\ \lfloor \frac{2(n+5k)}{3} \rfloor - 3 & 3k \leq n \leq 4k-2; \\ n+2k-3, & n \geq 4k-1, \end{cases} \tag{1}$$

and

$$D(n,k) = \left\lceil \frac{R(n,k)}{2} \right\rceil = \begin{cases} n-1, & n \leq 3k-1; \\ \lfloor \frac{n+5k}{3} \rfloor - 1 & 3k \leq n \leq 4k-2; \\ \lceil \frac{n-3}{2} \rceil + k, & n \geq 4k-1. \end{cases} \tag{2}$$

Egawa et al. [3] considered a closely related problem, but the answers differ, especially for $\sigma_2(G)$. The bounds of Theorem 2 and of Egawa et al. [3] are helpful in estimating $f(k)$ — the minimum positive integer f such that every f-connected graph is k-linked. After a series of papers by Jung, Larman and Mani, Mader, and Robertson and Seymour, the first linear upper bound for f, namely, $f(k) \leq 22k$ was proved by Bollobás and Thomason [1]. Very recently, Thomas and Wollan [10] improved this bound to $f(k) \leq 16k$. In [6] we show how to apply Theorem 2 in the Thomas-Wollan proof to improve their bound to $f(k) \leq 12k$. Thomas and Wollan informed us that elaborating our idea they are able to improve the bound even further: to $f(k) \leq 10k$.

References

[1] B. Bollobás and A. Thomason, Highly linked graphs, *Combinatorica*, **16** (1996), 313–320.

[2] G. Dirac, Some theorems on abstract graphs, *Proc. London Math. Soc.*, **2** (1952), 69–81.

[3] Y. Egawa, R. J. Faudree, E. Györi, Y. Ishigami, R. H. Schelp, and H. Wang, Vertex-disjoint cycles containing specified edges, *Graphs and Combinatorics*, **16** (2000), 81–92.

[4] J. Faudree, R. Faudree, R. Gould, M. Jacobson, L. Lesniak, On k-ordered graphs, *J. Graph Theory*, **35** (2000), 69–82.

[5] R. J. Gould, A. V. Kostochka, L. Lesniak, I. Schiermeyer, and A. Saito, Degree conditions for k-ordered hamiltonian graphs, *J. Graph Theory*, **42** (2003), 199–210.

[6] K. Kawarabayashi, A. Kostochka, and G. Yu, On sufficient degree conditions for a graph to be k-linked, submitted.

[7] H. A. Kierstead, G. Sárközy, S. Selkow, On k-ordered hamiltonian graphs, *J. Graph Theory*, **32** (1999), 17–25.

[8] L. Ng and M. Schultz, k-Ordered hamiltonian graphs, *J. Graph Theory*, **2** (1997), 45–57.

[9] O. Ore, Note on Hamilton circuits, *Amer. Math. Monthly*, **67** (1960), 55.

[10] R. Thomas, P. Wollan, An Improved Linear Edge Bound for Graph Linkage, 2003, submitted.

Spanning Triangulations in Graphs with Large Minimum Degree

Daniela Kühn

(joint work with Deryk Osthus)

In [6] and [5] we investigated the following extremal problem: given a function $m = m(n)$, how large does the minimum degree of a graph G of order n have to be in order to guarantee a planar subgraph with at least $m(n)$ edges? The main result of [5] determines the minimum degree which is necessary to force a planar subgraph with the maximum possible number of edges, i.e. a planar triangulation.

Theorem 1 *There exists an integer n_0 such that every graph G of order $n \geq n_0$ and minimum degree at least $2n/3$ contains a triangulation as a spanning subgraph.*

Our proof of Theorem 1 can easily be extended to obtain a spanning triangulation of an arbitrary surface. Theorem 1 improves a result from [6] where the minimum degree was required to be at least $2n/3 + \gamma n$ (here $\gamma > 0$ can be chosen to be arbitrary small and $n_0 = n_0(\gamma)$).

The following example shows that Theorem 1 is best possible for all integers n which are divisible by 3. Consider the graph G^* which is obtained from two disjoint cliques A and B of order $n/3$ by adding an independent set C of $n/3$ new vertices and joining each of them to all the vertices in the two cliques. So G^* has minimum degree $2n/3 - 1$. Observe that any spanning triangulation in G^* would have two facial triangles T_1 and T_2 which share an edge and are such that T_1 contains a vertex of A and T_2 contains a vertex of B. But this is impossible since every triangle of G^* containing a vertex of A (respectively B) can have at most one vertex outside A (respectively B), namely in C. One can extend this example slightly to show that for all n a minimum degree of $\lceil 2n/3 \rceil - 1$ does not ensure a spanning triangulation (see [5]).

The spanning triangulation guaranteed by Theorem 1 can be found in polynomial time. In other words, the maximum planar subgraph problem (which in a given graph G asks for a planar subgraph with the maximum number of edges) can be solved in polynomial time for graphs G of minimum degree at least $2n/3$. In general this problem was shown to be Max SNP-hard by Călinescu et al. [1], i.e. there exists a positive ε for which there cannot be a polynomial time approximation algorithm with approximation ratio better than $1 - \varepsilon$, unless $P = NP$. The best known approximation algorithm has an approximation ratio of 4/9 [1].

Our proof of Theorem 1 relies on Szemerédi's Regularity lemma, the Blow-up lemma of Komlós, Sárközy and Szemerédi [4] and several ideas which were introduced in [3] by the same authors. (In [3] they proved the related result that every graph of sufficiently large order n and minimum degree at least $2n/3$ contains the square of a Hamilton cycle.)

In the remainder we discuss how Theorem 1 might perhaps be strengthened. Obviously a minimum degree of $2n/3$ will not force every given triangulation P of order n as a subgraph. For example, G might be 3-partite, which implies that we can only hope for triangulations P with chromatic number 3. Of course, we cannot guarantee all of these either, as there are triangulations whose chromatic number is 3 and whose maximum degree is $n - 2$. However, in view of our proof of Theorem 1, it might be helpful to restrict one's attention to triangulations P of bounded band-width, as this imposes a linear structure on P. (The *band-width* of a graph H is the smallest integer k for which there exists an enumeration $v_1, \dots, v_{|H|}$ of the vertices of H such that every edge $v_i v_j \in H$ satisfies $|i - j| \leq k$.) Bollobás and Komlós [2] conjectured that for every $\gamma > 0$ and all $r, \Delta \in \mathbb{N}$ there are $\alpha > 0$

and $n_0 \in \mathbb{N}$ such that every graph G of order $n \geq n_0$ and minimum degree at least $(1 - \frac{1}{r} + \gamma)n$ contains a copy of every graph H of order n whose chromatic number is at most r, whose maximum degree is at most Δ and whose band-width is at most αn.

This would imply that every sufficiently large graph of minimum degree at least $(2/3 + \gamma)n$ contains every 3-chromatic triangulation of bounded band-width. Even in this special case the error term γn cannot be omitted completely: there are 3-chromatic triangulations whose colour classes have different sizes. These obviously do not embed into the complete 3-partite graph whose vertex classes have equal size. However, it might be true that for all integers b there exists a constant $C = C(b)$ such that every graph of order n and minimum degree at least $2n/3 + C$ contains every 3-chromatic triangulation of order n and band-width at most b as a subgraph.

Also, we do not know whether one can strengthen Theorem 1 in the following way. Given n, is there a triangulation P_n of order n which is contained in *every* graph G of order n and minimum degree at least $2n/3$? When n is divisible by 3, then the preceding arguments show that P_n would have to be 3-chromatic with equal size colour classes. Moreover, P_n would have to contain induced cycles of many different lengths. To see the latter, consider a graph G which is similar to the graph G^* defined earlier. This time the cliques have order $n/3 - 1$, the independent set C has order $n/3 + 2$ and we insert a 2-factor into C. One can show that every spanning triangulation of G must contain one of the cycles in $G[C]$.

References

[1] G. Călinescu, C.G. Fernandes, U. Finkler and H. Karloff, A better approximation algorithm for finding planar subgraphs, *J. Algorithms* **28** (1998), 105-124.

[2] J. Komlós, The Blow-up Lemma, *Comb. Probab. Comput.* **8** (1999), 161–176.

[3] J. Komlós, G. N. Sárközy and E. Szemerédi, On the square of a Hamilton cycle in dense graphs, *Random Struct. Alg.* **9** (1996), 193–211.

[4] J. Komlós, G. N. Sárközy and E. Szemerédi, Blow-up Lemma, *Combinatorica* **17** (1997), 109–123.

[5] D. Kühn and D. Osthus, Spanning triangulations in graphs, submitted.

[6] D. Kühn, D. Osthus and A. Taraz, Large planar subgraphs in dense graphs, submitted.

Revisiting Two Theorems of Curto and Fialkow on Moment Matrices
Monique Laurent

The moment problem

Given a probability measure μ on $\mathbb{R}^n$, the quantity $y_\alpha := \int x^\alpha \mu(dx)$ is called its *moment of order* α. The moment problem concerns the characterization of the sequences $y = (y_\alpha)_{\alpha \in \mathbb{Z}_+^n}$ that are the sequences of moments of some nonnegative measure μ; in that case one says that μ is a *representing measure* for y and μ is a probability measure if $y_0 = 1$. The results of Curto and Fialkow that we consider here deal with moment sequences of finite atomic measures, i.e., measures of the form $\mu = \sum_{i=1}^r \lambda_i \delta_{x_i}$ with $\lambda_1, \ldots, \lambda_r \neq 0$ and $x_1, \ldots, x_r \in \mathbb{R}^n$. Here, δ_x is the *Dirac measure* at $x \in \mathbb{R}^n$ (with mass 1 at x and 0 elsewhere), whose moment sequence is the *zeta vector* $\zeta_x := (x^\alpha)_{\alpha \in \mathbb{Z}^n} \in \mathbb{R}^{\mathbb{Z}_+^n}$.

Given $y \in \mathbb{R}^{\mathbb{Z}_+^n}$, its *moment matrix* is the symmetric matrix $M(y)$ indexed by $\mathbb{Z}_+^n$ whose (α, β)th entry is equal to $y_{\alpha+\beta}$, for $\alpha, \beta \in \mathbb{Z}_+^n$. A well known necessary condition for y to have a representing measure μ is the positive semidefiniteness of its moment matrix. Moreover, the support of μ is contained in the set of common zeros of the polynomials belonging to the kernel of $M(y)$ and the rank of $M(y)$ is at most the number of atoms in the support of μ.

The cone $\mathcal{M}$ consisting of the sequences y having a representing measure, and the cone $\mathcal{P}$ consisting of the polynomials nonnegative on $\mathbb{R}^n$, are dual of each other (Haviland [5]). Moreover, the cone $\mathcal{M}_+$ consisting of the sequences y whose moment matrix $M(y)$ is positive semidefinite, and the cone Σ^2 consisting of all sums of squares of polynomials, are dual of each other (Berg et al. [1]). Thus the moment problem can be cast - via duality - as the problem of characterizing nonnegative polynomials. The inclusion: $\Sigma^2 \subseteq \mathcal{P}$ is an equality for $n = 1$ and it is strict for $n \geq 2$, as already noticed by Hilbert in the 1890s. Equivalently, the inclusion: $\mathcal{M} \subseteq \mathcal{M}_+$ is an equality for $n = 1$ (this is Hamburger's theorem) and it is strict for $n \geq 2$.

There are, however, some cases when the implication: $y \in \mathcal{M}_+ \Longrightarrow y \in \mathcal{M}$ holds. Berg, Christensen and Ressel [1] show that this is true when y is bounded. Curto and Fialkow [2] show that this is true when $M(y)$ has finite rank.

Theorem 1 *[2] If $M(y) \succeq 0$ and $M(y)$ has finite rank r, then y has a unique representing measure, which is r-atomic.*

As a direct application of Theorem 1, the reverse implication also holds: If y has a r-atomic representing measure, then $M(y) \succeq 0$ and rank $M(y) = r$.

Curto and Fialkow's proof for Theorem 1 is along the following lines. (See chapter 4 in [2].) Assume $M(y) \succeq 0$ and rank $M(y) = r$. Then, the kernel $I := \{p \in \mathbb{R}[x_1, \ldots, x_n] \mid M(y)p = 0\}$ of $M(y)$ is an ideal in $\mathbb{R}[x_1, \ldots, x_n]$ and the quotient vector space $A := \mathbb{R}[x_1, \ldots, x_n]/I$ has finite dimension r. Define an inner product on A by setting $\langle p, q\rangle := p^T M(y) q$. In this way, A is a Hilbert space of dimension r. For $q \in \mathbb{R}[x_1, \ldots, x_n]$, consider the multiplication operator $\varphi_q : A \to A$ defined by $\varphi_q(p) = pq$. Obviously, the operators $\varphi_{x_1}, \ldots, \varphi_{x_n}$ pairwise commute. Curto and Fialkow use then the spectral theorem and the Riesz representation theorem for proving the existence of a representation measure for y. This type of proof based on functional analytic tools is often used for proving results about the moment problem. See, e.g., Fuglede [4], Schmüdgen [8].

The first main contribution of this paper is an alternative more elementary proof for Theorem 1. Our proof uses Hilbert's Nullstellensatz and, beside this algebraic result, it uses only elementary linear algebra. Our starting point is to observe that the kernel I of $M(y)$ is a *radical* ideal. Hence, the variety $V(I)$ (consisting of the common complex roots of all polynomials in I) has cardinality r. Say, $V(I) = \{v_1, \ldots, v_r\}$. Note that a complex point v belongs to $V(I)$ if and only if its conjugate $\overline{v}$ belongs to $V(I)$. Thus, one can write: $V(I) = S \cup T \cup \overline{T}$, where $S := V(I) \cap \mathbb{R}^n$ and $\overline{T} := \{\overline{v} \mid v \in T\}$.

Let $p_{v_1}, \ldots, p_{v_r} \in \mathbb{C}[x_1, \ldots, x_n]$ be interpolation polynomials at the points of $V(I)$; that is, $p_{v_i}(v_j) = 1$ if $i = j$ and $p_{v_i}(v_j) = 0$ if $i \neq j$, for $i, j = 1, \ldots, r$. One can assume that p_v is real valued for $v \in S$ and that $p_{\overline{v}} = \overline{p_v}$ for $v \in T$.

Let Z be the matrix whose columns are the zeta vectors $\zeta_{v_1}, \ldots, \zeta_{v_r}$, and let $\tilde{Z}$ be the matrix whose rows contain the coefficient vectors of the interpolation polynomials $p_{v_1}, \ldots, p_{v_r}$. Thus, $\tilde{Z}Z = I_r$. Theorem 1 now follows from the next three lemmas.

Lemma 2 $M(y) = Z\text{diag}(\tilde{Z}y)Z^T$.

Lemma 3 $V(I) \subseteq \mathbb{R}^n$.

Lemma 4 $M(y) = \sum_{i=1}^r p_{v_i}^T M(y) p_{v_i} \zeta_{v_i} \zeta_{v_i}^T$ *and* $\mu := \sum_{i=1}^r p_{v_i}^T M(y) p_{v_i} \delta_{v_i}$ *is the unique measure representing* y*.*

The F-moment problem

Curto and Fialkow [3] study the F-moment problem for truncated sequences. That is, given a sequence $y \in \mathbb{R}^{S_{2t}}$, decide whether y has a representing measure supported by a given set $F \subseteq \mathbb{R}^n$. Here, for an integer $t \geq 1$, S_t denotes the set of $\alpha \in \mathbb{Z}_+^n$ with $\sum_i \alpha_i \leq t$. Consider the case when F is a basic closed semialgebraic set, of the form

$$F := \{x \in \mathbb{R}^n \mid h_1(x) \geq 0, \ldots, h_m(x) \geq 0\}, \tag{1}$$

where $h_1, \ldots, h_m \in \mathbb{R}[x_1, \ldots, x_n]$; set

$$d_j = \lceil \deg(h_j)/2 \rceil, \ d := \max_{j=1}^m d_j. \tag{2}$$

Necessary conditions can be formulated in terms of positive semidefiniteness of the localizing matrices of y. Given $h \in \mathbb{R}[x_1, \ldots, x_n]$, $h * y$ denotes the vector whose αth entry is $(h * y)_\alpha := \sum_\beta h_\beta y_{\alpha+\beta}$; its moment matrix is a *localizing matrix* of y. Moreover, $M_t(y)$ is the matrix indexed by S_t whose (α, β)th entry is $y_{\alpha+\beta}$. One can easily verify that, if $y \in \mathbb{R}^{S_{2t}}$ has a representing measure supported by the set F, then $M_t(y) \succeq 0$ and $M_{t-d_j}(h_j * y) \succeq 0$ for all $j = 1, \ldots, m$. Curto and Fialkow [3] show that, under certain rank assumptions, these necessary conditions are also sufficient for the existence of a representing measure supported by F. The following is the main result of [3] (Theorem 1.6 there).

Theorem 5 *[3] Let F be the set from (1) and let $d_1, \ldots, d_m, d$ be as in (2). Let $y \in \mathbb{R}^{S_{2t}}$ and $r := \mathrm{rank}\ M_t(y)$. The following assertions are equivalent.*

(i) y has a r-atomic representing measure whose support is contained in F.

(ii) $M_t(y) \succeq 0$ and y can be extended to a vector $y \in \mathbb{R}^{S_{2(t+d)}}$ in such a way that $M_{t+d}(y)$ is a flat extension of $M_t(y)$ and $M_t(h_j * y) \succeq 0$ for $j = 1, \ldots, m$.

The second main contribution of our paper is a very short proof of this result. Assume (ii) holds. Then, by Theorem 6 below, y has a representing measure $\mu = \sum_{i=1}^r \lambda_i \delta_{v_i}$, where $r = \mathrm{rank}\ M_t(y)$. Hence, it suffices to show that all v_i's belong to the set F. This follows from the assumption that $M_t(h_j * y) \succeq 0$, after observing that, as $r = \mathrm{rank}\ M_t(y)$, one can find interpolation polynomials at $v_1, \ldots, v_r$ having degree at most t.

Theorem 6 *[2] Given $y \in \mathbb{R}^{S_{2t}}$, assume that $M_t(y) \succeq 0$ and that $\mathrm{rank}\ M_t(y) = \mathrm{rank}\ M_{t-1}(y)$. Then one can extend y to a vector in $\mathbb{R}^{\mathbb{Z}_+^n}$ having a representing measure which is $(\mathrm{rank}\ M_t(y))$-atomic.*

Our study of the moment problem is partly motivated by its application to optimization. Indeed, Lasserre [7] shows how to construct asymptotic converging sequences of semidefinite relaxations using moment matrices, for the problem of minimizing a polynomial over a basic closed semi-algebraic set. Curto and Fialkow's results are used for proving, in some cases, the finite convergence. See also Henrion and Lasserre [6].

References

[1] C. Berg, J.P.R. Christensen, and C.U. Jensen. A remark on the multidimensional moment problem. *Math. Ann.* **243** (1979), 163–169.

[2] R.E. Curto and L.A. Fialkow. Solution of the truncated complex moment problem for flat data. *Mem. Amer. Math. Soc.* vol. 568, Amer. Math. Soc., Providence, RI, 1996.

[3] R.E. Curto and L.A. Fialkow. The truncated complex K-moment problem. *Trans. Amer. Math. Soc.* **352** (2000), 2825–2855.

[4] B. Fuglede. The multidimensional moment problem. *Expo. Math.* **1** (1983), 47–65.

[5] E.K. Haviland. On the momentum problem for distributions in more than one dimension. *Amer. J. Math.* **57** (1935), 562–568.

[6] D. Henrion and J.-B. Lasserre. Detecting global optimality and extracting solutions in GloptiPoly. Preprint, 2003.

[7] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* **11** (2001), 796–817.

[8] K. Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Math. Ann.* **289** (1991), 203–206.

Partition Regular Equations

Imre Leader

(joint work with N. Hindman, P.A. Russell and D. Strauss)

An $n \times m$ matrix A, with rational entries, is called *partition regular* if whenever the natural numbers are finitely coloured there is a monochromatic vector x (meaning that all entries of x have the same colour) such that $Ax = 0$. We may also speak of the 'system of equations' $Ax = 0$ being partition regular.

The aim of this talk is to review some previous knowledge about the important notion of 'consistency', to be defined below, and then to go on to some more recent work. This recent work is joint with Hindman and Strauss [3],[4] and joint with Russell [5].

Many of the classical theorems of Ramsey Theory, such as Schur's theorem and van der Waerden's theorem, may naturally be interpreted at statements that certain matrices are partition regular. The partition regular matrices were characterised by Rado [7] in the 1930s. His characterisation had the following important consequence: if A and B are partition regular then so is their diagonal sum. In other words, if we can always solve $Ax = 0$ in one colour class, and we can always solve $By = 0$ in one colour class, then in fact we can solve $Ax = 0$ and $By = 0$ in the *same* colour class. We say that the matrices A and B are *consistent*.

This is important because it can be used to prove some 'universal' results. For example, whenever the natural numbers are finitely coloured, some class must contain solutions to *all* partition regular systems.

Let us now pass to the infinite case, where, in contrast to Rado's theorem, the whole picture is very much not yet understood. Which infinite systems of equations are partition regular? One very simple example, coming straight from Ramsey's theorem, is as follows: whenever the natural numbers are finitely coloured there exist $x_1, x_2, \ldots$ such that the set $\{x_i + x_j : i \neq j\}$ is monochromatic. (This is not quite given in the form of a solution to $Ax = 0$ for some suitable A, but it can easily be converted into that form if desired.) More generally, Ramsey's theorem implies that, for any fixed $a_1, \ldots, a_m$ positive integers, whenever the natural numbers are finitely coloured there exist $x_1, x_2, \ldots$ such that the set $\{a_1 x_{i_1} + \ldots + a_m x_{i_m} : i_1 < \ldots < i_m\}$ is monochromatic. We call this simple system a 'Ramsey' system. It is worth pointing out that one cannot relax the condition on $i_1, \ldots, i_m$ to the condition that they are merely distinct: for this system there are bad colourings.

The first non-trivial example of an infinite partition regular system was given by Hindman [2], who showed that whenever the natural numbers are finitely coloured there exist $x_1, x_2, \ldots$ such that the set

$$FS(x_1, x_2, \ldots) \;=\; \{\sum_{i \in I} x_i : \; 0 < |I| < \infty\}$$

is monochromatic. This was extended by Milliken [6]and Taylor [8], who showed that, for any fixed $a_1, \ldots, a_m$ positive integers, whenever the natural numbers are finitely coloured there exist $x_1, x_2, \ldots$ such that the set

$$FS_{a_1,\ldots,a_m}(x_1, x_2, \ldots) \;=\; \{a_1 \sum_{i \in I_1} x_i \; + \ldots + a_m \sum_{i \in I_m} x_i\}$$

is monochromatic, where we allow all finite nonempty $I_1, \ldots, I_m$ such that $\max I_r < \min I_{r+1}$ for all r. However, it is important to point out that not too many other examples of infinite partition regular systems are known.

It was proved by Deuber, Hindman, Leader and Lefmann [1] that unfortunately, in the infinite case, consistency does not always hold. Indeed, two different Milliken-Taylor systems are, except in trivial cases, always inconsistent. This left as a vexing open problem the question of whether or not the simple Ramsey systems were consistent. This was open for some time, being eventually solved by Hindman, Leader and Strauss [3]. The proof uses a large amount of machinery from the Stone-Cech compactification of the natural numbers (the space of ultrafilters), together with a new notion related to this space called 'central partition regularity'.

Recently, however, Leader and Russell [5] have found a very short proof of the consistency of Ramsey systems.

One other interesting development has concerned Ramsey systems with negative entries. Here we allow some of the a_i to be negative (although of course the final coefficient a_m must be positive, to have any hope of finding solutions in the natural numbers). One might imagine that this is just generalisation for its own sake, but curiously enough when one allows negative entries one suddenly obtains some much simpler proofs of inconsistency than were needed in [1]. This work is presented in [4].

References

[1] W. Deuber, N. Hindman, I. Leader and H. Lefmann, Infinite partition regular matrices, *Combinatorica* **15** (1995), 333-355.

[2] N. Hindman, Finite sums from sequences within cells of a partition of $\mathbb{N}$, *J. Combinatorial Theory (A)* **17** (1974), 1–11.

[3] N. Hindman, I. Leader and D. Strauss, Infinite partition regular matrices – solutions in central sets, *Trans. Amer. Math. Soc.* **355** (2003), 1213-1235.

[4] N. Hindman, I. Leader and D. Strauss, Separating Milliken-Taylor systems with negative entries, *Proc. Edinburgh Math. Soc.* **46** (2003), 45-61.

[5] I. Leader and P.A. Russell, Consistency for partition regular equations, submitted.

[6] K.R. Milliken, Ramsey's theorem with sums or unions, *J. Combinatorial Theory (A)* **18** (1975), 276-290.

[7] R. Rado, Studien zur Kombinatorik, *Math. Zeit.* **36** (1933), 242-280.

[8] A. Taylor, A canonical partition relation for finite subsets of ω, *J. Combinatorial Theory (A)* **21** (1976), 137-146.

Lifts, Discrepancy and Nearly Optimal Spectral Gaps

Nati Linial

(joint work with Yonatan Bilu)

Let G be a graph on n vertices. A 2-lift of G is a graph H on $2n$ vertices, with a covering map $\pi : H \to G$. It is not hard to see that all eigenvalues of G are also eigenvalues of H. In addition, H has n "new" eigenvalues. We conjecture that

every d-regular graph has a 2-lift such that all new eigenvalues are in the range $[-2\sqrt{d-1}, 2\sqrt{d-1}]$ (If true, this is tight , e.g. by the Alon-Boppana bound). Here we show that every graph of maximal degree d has a 2-lift such that all "new" eigenvalues are in the range $[-c\sqrt{d\log^3 d}, c\sqrt{d\log^3 d}]$ for some constant c. This leads to a polynomial time algorithm for constructing arbitrarily large d-regular graphs, with second eigenvalue $O(\sqrt{d\log^3 d})$.
The proof uses the following lemma (Lemma 5): Let A be a real symmetric matrix with zeros on the diagonal. Let d be such that the l_1 norm of each row in A is at most d. Let α be such that for every $x, y \in \{0,1\}^n$ with $< x, y >= 0$ it holds that $\frac{|xAy|}{||x||||y||} \leq \alpha$. Then the spectral radius of A is $O(\alpha(\log(d/\alpha) + 1))$. An interesting consequence of this lemma is a converse to the Expander Mixing Lemma.

Definitions

Let $G = (V, E)$ be a graph on n vertices, and let A be its adjacency matrix. Let $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n$ be the eigenvalues of A. We say that G is an $(n, d, \mu)-$ *expander* if G is d-regular, and $\max_{i=2,\ldots,n} |\mu_i| \leq \mu$. When $\mu = 2\sqrt{d-1}$ we say that such a graph is *Ramanujan*. When $\mu = \tilde{O}(\sqrt{d})$ we say that such a graph is *Quasi-Ramanujan*.

A *signing* of the edges of G is a function $s : E(G) \to \{-1, 1\}$. The *signed adjacency matrix* of a graph G with a signing s has rows and columns indexed by the vertices of G. The (x, y) entry is $s(x, y)$ if $(x, y) \in E$ and 0 otherwise.
A 2-*lift* of G, associated with a signing s, is a graph $\hat{G}$ defined as follows. Associated with every vertex $x \in V$ are two vertices, x_0 and x_1, called the *fiber* of x. If $(x, y) \in E$, and $s(x, y) = 1$ then the corresponding edges in $\hat{G}$ are (x_0, y_0) and (x_1, y_1). If $s(x, y) = -1$, then the corresponding edges in $\hat{G}$ are (x_0, y_1) and (x_1, y_0). The graph G is called the *base graph*, and $\hat{G}$ a 2-lift of G. By the *spectral radius of a signing* we refer to the spectral radius of the corresponding signed adjacency matrix. When the spectral radius of a signing of a d-regular graph is $\tilde{O}(\sqrt{d})$ we say that the signing (or the lift) is *Quasi-Ramanujan*.

Quasi-Ramanujan 2-Lifts and Quasi-Ramanujan Graphs

Preliminaries. The eigenvalues of a 2-lift of G can be easily characterized in terms of the adjacency matrix and the signed adjacency matrix:

Lemma 1 *Let A be the adjacency matrix of a graph G, and A_s the signed adjacency matrix associated with a 2-lift $\hat{G}$. Then every eigenvalue of A and every eigenvalue of A_s are eigenvalues of $\hat{G}$. Furthermore, the multiplicity of each eigenvalue of $\hat{G}$ is the sum of its multiplicities in A and A_s.*

Consider the following scheme for constructing (n, d, λ)-expanders. Start with $G_0 = k_{d+1}$, the complete graph on $d+1$ vertices * . Its eigenvalues are d, with multiplicity 1, and -1, with multiplicity d. We want to define G_i as a 2-lift of G_{i-1}, such that all new eigenvalues are in the range $[-\lambda, \lambda]$. Assuming such a 2-lifts always exist, the G_i constitute an infinite family of (n, d, λ)-expanders.
It is therefore natural to look for the smallest $\lambda = \lambda(d)$ such that every graph of degree at most d has a 2-lift, with new eigenvalues in the range $[-\lambda, \lambda]$. In other words, a signing with spectral radius $\leq \lambda$. We note that $\lambda(d) \geq 2\sqrt{d-1}$ follows from the Alon-Boppana bound.
Quasi-Ramanujan 2-lifts for every graph. Based on extensive computer simulations we conjecture that every graph has a signing with small spectral radius:

Conjecture 2 *Every d-regular graph has a signing with spectral radius at most $2\sqrt{d-1}$.*

In this section we show a close upper bound:

Theorem 3 *Every graph of maximal degree d has a signing with spectral radius $O(\sqrt{d \cdot \log^3 d})$.*

The theorem follows immediately from the following two lemmata (along with Lemma 1). The first shows that with positive probability the Rayliegh quotient is small for vectors in $v, u \in \{-1, 0, 1\}^n$. The second shows that this is essentially a sufficient condition for all eigenvalues being small.

Lemma 4 *For every graph of maximal degree d, there exists a signing s such that for all $v, u \in \{-1, 0, 1\}^n$ the following holds:*

$$\frac{|v^t A_s u|}{||v||||u||} \leq 10\sqrt{d \log d}, \tag{1}$$

where A_s is the signed adjacency matrix.

Lemma 5 *Let A be an $n \times n$ real symmetric matrix such that the l_1 norm of each row in A is at most d. Assume that for any two vectors, $u, v \in \{0, 1\}^n$, with $supp(u) \cap supp(v) = \emptyset$:*

$$\frac{|uAv|}{||u||||v||} \leq \alpha,$$

and that all diagonal entries of A are, in absolute value, $O(\alpha(log(d/\alpha)+1))$. Then the spectral radius of A is $O(\alpha(log(d/\alpha) + 1))$.

---

*We could start with any small d-regular graph with a large spectral gap. Such graphs are easy to find.

Note 6 *Lemma 5 is tight up to constant factors. To see this, consider the n-dimensional vector x whose i'th entry is $1/\sqrt{i}$. Let A be the outer product of x with itself, that is, the matrix whose (i,j)'th entry is $1/\sqrt{i \cdot j}$. Clearly x is an eigenvector of A corresponding to the eigenvalue $||x||^2 = \Theta(\log(n))$. Also, the sum of each row in A is $O(\sqrt{n})$. To prove that the lemma is essentially tight, we need to show that $max_{u,v\in\{0,1\}^n} \frac{uAv}{||u||||v||}$ is constant. Indeed, fix $k,l \in [n]$. Let $u,v \in \{0,1\}^n$ be such that $||u|| = k$ and $||v|| = l$. As the entries of A are decreasing along the rows and the columns, uAv is maximized for such vectors when their support is the first k and l coordinates. For these optimal vectors, $uAv = \Theta(\sqrt{k \cdot l})$. Thus,*

$$max_{u,v\in\{0,1\}^n} \frac{uAv}{||u||||v||} = \Theta(1).$$

An explicit construction of quasi-Ramanujan graphs. For the purpose of building expanders, it is enough to prove a weaker version of Theorem 3. Roughly, that every *expander* graph has a 2-lift with small spectral radius. In this subsection we show that when the base graph is a good expander (in the sense of the definition below), then w.h.p. a random 2-lift has a small spectral radius. We then derandomize the construction to get a deterministic polynomial time algorithm for constructing arbitrarily large expander graphs.

Definition 7 *We say that a graph G on n vertices is (β,t)-sparse if for every $u,v \in \{0,1\}^n$, with $|S(u,v)| \leq t$,*

$$uAv \leq \beta||u||||v||.$$

Lemma 8 *Let A be the adjacency matrix of a d-regular $(\gamma(d), \log n)$-sparse G graph on n vertices, where $\gamma(d) = 10\sqrt{d \log d}$. Then for a random signing of G (where the sign of each edge is chosen uniformly at random) the following hold w.h.p.:*

1. $\forall u,v \in \{-1,0,1\}^n : |uA_s v| \leq \gamma(d)||u||||v||$.
2. *$\hat{G}$ is $(\gamma(d), 1+\log n)$-sparse*

where A_s is the random signed adjacency matrix, and $\hat{G}$ is the corresponding 2-lift.

Corollary 9 *Let A be the adjacency matrix of a d-regular $(\gamma(d), \log n)$-sparse G graph on n vertices, where $\gamma(d) = 10\sqrt{d \log d}$. Then there is a deterministic polynomial time algorithm for finding a signing s of G such that the following hold:*

1. *The spectral radius of A_s is $O(\sqrt{d \log^3 d})$.*
2. *$\hat{G}$ is $(\gamma(d), 1+\log n)$-sparse,*

where A_s is the signed adjacency matrix, and $\hat{G}$ is the corresponding 2-lift.

A converse to the Expander Mixing Lemma

There are several approaches to expansion in graphs. A combinatorial definition says that a d-regular graph on n vertices is an (n, d, c)-vertex expander if every set of vertices, W, of size at most $n/2$, has at least $c|W|$ neighbors outside itself. An algebraic definition says that such a graph is an (n, d, λ)-expander if all eigenvalues but the largest are, in absolute value, at most λ.
The two notions are closely related. For example, it is known (cf. [2]) that an (n, d, λ)-expander is also an $(n, d, \frac{d-\lambda}{2d})$-vertex expander. Conversely, Alon shows in [1] that an (n, d, c)-vertex expander is also an $(n, d, d - \frac{c^2}{4+2c^2})$-expander. Roughly, these results show that one type of expansion implies the other. However, in all such results one implication (from combinatorial to algebraic expansion) is much weaker than the other.
For two subsets of vertices, S and T, let $e(S, T)$ denote the number of edges between them. A very useful property of (n, d, λ)-expanders is known as the Expander Mixing Lemma (cf. [2]): For every two subsets of vertices, A and B, of an (n, d, λ)-expander:

$$|e(A, B) - d|A||B|/n| \leq \lambda\sqrt{|A||B|}.$$

Lemma 5 also implies a converse to this well known fact:

Corollary 10 *Let G be a d-regular graph on n vertices. Suppose that for any $S, T \subset V(G)$, with $S \cap T = \emptyset$*

$$|e(S, T) - \frac{|S||T|d}{n}| \leq \alpha\sqrt{|S||T|}$$

Then all but the largest eigenvalue of G are bounded, in absolute value, by $O(\alpha(1 + \log(d/\alpha)))$.

It is known that for a random d-regular graph, w.h.p., the condition in Corollary 10 holds with $\alpha = O(\sqrt{d})$ (cf. [3]). Hence, it follows from the corollary that w.h.p., such a graph is an $(n, d, O(\sqrt{d}\log d))$-expander. While this result is weaker than previous ones ([6, 5, 4]), the proof here is somewhat shorter and simpler.
Acknowledgments. We thank László Lovász for insightful discussions, and Efrat Daom for help with computer simulations. We thank Eran Ofek for suggesting that Corollary 10 might be used to bound the second eigenvalue of random d-regular graphs.
We appreciate the helpful comments we got from Alex Samorodnitsky, Eyal Rozenman and Shlomo Hoory.

References

[1] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986. Theory of computing (Singer Island, Fla., 1984).

[2] N. Alon and J. H. Spencer. *The probabilistic method.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience [John Wiley & Sons], New York, second edition, 2000. With an appendix on the life and work of Paul Erdős.

[3] Y. Bilu and S. Hoory. On codes and hypergraphs. To appear in European Journal of Combinatorics.

[4] J. Friedman. A proof of alon's second eigenvalue conjecture. manuscript.

[5] J. Friedman. On the second eigenvalue and random walks in random d-regular graphs. *Combinatorica*, 11(4):331–362, 1991.

[6] J. Friedman, J. Kahn, and E. Szemeredi. On the second eigenvalue of random regular graphs. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 587–598. ACM Press, 1989.

Expected Length of the Longest Common Subsequence for Large Alphabets

Jiří Matoušek

(joint work with Marcos Kiwi and Martin Loebl)

We investigate the distribution of the length L of the longest common subsequence of two randomly uniformly and independently chosen n character words $u = u_1 u_2 \dots u_n$ and $v = v_1 v_2 \dots v_n$ over a k-ary alphabet. That is, L is the maximum integer such that there exist indices $i_1 < i_2 < \cdots < i_L$ and $j_1 < j_2 < \cdots < j_L$ with $u_{i_q} = v_{j_q}$, $q = 1, 2, \dots, L$. This problem has emerged more or less independently in several remarkably disparate areas, including the comparison of versions of computer programs, cryptographic snooping, and molecular biology. An extended abstract of this work appears in *Proc. 6th Latin American Theoretical Informatics Symposium* (LATIN 2004), LNCS series, Springer, Berlin. A full version is available at the web page of the author.

By a well-known subadditivity argument, $\mathbf{E}[L]/n$ converges to a constant γ_k. The value of γ_k is not known for any particular value of k, although much effort has been spent in finding good upper an lower bounds for it (see, for example, [2] and references therein).

We analyze the behavior of γ_k for $k \to \infty$, and more generally, we consider the expectation of L when k is an (arbitrarily slowly growing) function of n and

$n \to \infty$. In particular, we prove a conjecture of Sankoff and Mainville from the early 80's [7] stating that

$$\lim_{k\to\infty} \gamma_k \sqrt{k} = 2. \tag{1}$$

(See [6, § 6.8] for a discussion of work on lower and upper bounds on γ_k as well as a stronger version, due to Arratia and Steele, of the above stated conjecture.)

The constant 2 in (1) arises from a connection with another celebrated problem, the distribution of LIS_N, the length of the longest increasing subsequence in a (uniform) random permutation of $\{1, 2, .., N\}$. Hammersley [4] proved the existence of $\beta = \lim_{N\to\infty}(\mathbf{E}[\mathrm{LIS}_N]/\sqrt{N})$ and conjectured that $\beta = 2$. Later, Logan and Shepp [5], based on a result by Schensted [8], proved $\beta \le 2$, and finally, Vershik and Kerov [10] showed $\beta = 2$. In a major recent breakthrough Baik, Deift, Johansson [3] described explicitly the asymptotic distribution of LIS_N (for $N \to \infty$). For a detailed account of these results, history, and related work we can recommend the surveys of Aldous and Diaconis [1] and Stanley [9]; the methods used in attacking this problem are of remarkable beauty and diversity.

Our main result about the longest common subsequence can be stated as follows.

Theorem 1 *For every $\varepsilon > 0$ there exist k_0 and C such that for all $k > k_0$ and all n with $n/\sqrt{k} > C$ we have*

$$(1-\varepsilon)\cdot\frac{2n}{\sqrt{k}} \;\le\; \mathbf{E}[L] \;\le\; (1+\varepsilon)\cdot\frac{2n}{\sqrt{k}}$$

where, as above, L is the length of the longest common subsequence of two random words of length n over an alphabet of size k. Moreover, there is an exponentially small tail bound; namely, for every $\varepsilon > 0$ there exists $c > 0$ such that for k and n as above,

$$\mathbf{P}\left[\left|L - \frac{2n}{\sqrt{k}}\right| \ge \varepsilon\frac{2n}{\sqrt{k}}\right] \le e^{-cn/\sqrt{k}}.$$

In the rest of this extended abstract, we outline the main tools and ideas of the proof, referring to the full version for precise formulations and further details.

First we reformulate the problem a little. Given the two random words $u = u_1u_2\dots u_n$ and $v = v_1v_2\dots v_n$, let us draw two horizontal lines in the plane and place n points $a_1, a_2, \dots, a_n$ in this order on the top line and n points $b_1, b_2, \dots, b_n$ in this order on the bottom line. Then we connect a_i to b_j by an edge (straight segment) iff $u_i = v_j$, obtaining a drawing of a bipartite graph G (which is a disjoint union of complete bipartite graphs). A common subsequence of the words u and v corresponds to a *planar matching* in G (a matching in which no two edges cross).

Although we want to deal mainly with the case of n arbitrarily large compared to k, which is the situation in the Sankoff–Mainville conjecture, we first consider a seemingly opposite setting: when k is large and n is also large but considerably smaller than n. For definiteness, we set $n = k^{0.7}$. Then we expect G to have about $n^2/k = k^{0.4}$ edges, and most of these edges connect vertices of degree 1. If we let G' be the subgraph of G obtained by deleting all edges incident to vertices of degree greater than 1, then G' is a matching (plus some isolated vertices). The number N of edges of G' is typically very close to $k^{0.4}$. The matching determines a permutation of $\{1, 2, \ldots, N\}$, and by a symmetry argument, it can be seen that, for a given N, all permutations of $\{1, 2, \ldots, N\}$ have the same probability of being obtained in this way. Moreover, the longest increasing subsequence of the permutation corresponds exactly to the largest planar matching in G'. Therefore, up to a small error, the longest common subsequence of u and v is distributed as LIS_N. Then one can derive from the known results about LIS_N that $\mathbf{E}[L] = (2 + o(1))n/\sqrt{k}$ holds in this situation. For the rest of the proof, we also need tail estimates for large deviations of L, and these are conveniently obtained from Talagrand's inequality applied to L (we cannot directly use known tail estimates for LIS_N, for example because of the vertices of degree larger than 1 in G).

Now we consider n very large compared to k (and k still large). A lower bound for $\mathbf{E}[L]$ is straightforward: We partition the words u and v into segments of length $k^{0.7}$ each, and we use the previously derived result separately for each block (the ith block consists of the ith segment of the word u plus the ith segment of the word v). Thus, the lower bound is provided by a common subsequence, or planar matching in the graph language, that never crosses a block boundary.

An upper bound for $\mathbf{E}[L]$ is more demanding, since the largest planar matching need not respect any partition into blocks fixed in advance; there could be "very skew" edges. Our strategy is to simultaneously consider many different partitions into blocks. The blocks have upper and lower segments of size about $k^{0.7}$, but they can be very skew; the segment of u starting at a position i can form a block with a segment of v starting at position j, with i and j differing by a large amount. Supposing that there is a planar matching with at least $m = (1+\varepsilon)2n/\sqrt{k}$ edges, it "fits" at least one of the block partitions, meaning that it respects its block boundaries. For each fixed block partition and each fixed distribution of the numbers of edges of the planar matching among the blocks, we bound above the probability that there is a planar matching with m edges that fits that block partition; this relies on independence among the blocks. Then we sum up over all possible block partitions and show that with high probability, there is no planar matching with m edges at all.

References

[1] D. Aldous and P. Diaconis. Longest increasing subsequences: From patience sorting to the Baik–Deift–Johansson theorem. *Bulletin of the AMS*, 36(4):413–432, 1999.

[2] R. Baeza-Yates, G. Navarro, R. Gavaldá, and R. Scheing. Bounding the expected length of the longest common subsequences and forests. *Theory of Computing Systems*, 32(4):435–452, 1999.

[3] J. Baik, P. A. Deift, and K. Johansson. On the distribution of the length of the longest increasing subsequence of random permutaions. *J. Amer. Math. Soc.*, 12:1119–1178, 1999.

[4] J. M. Hammersley. A few seedlings of research. In *Proc. Sixth Berkeley Sympos. Math. Stat. Prob.*, pages 345–394, Berkeley, Calif., 1972. Univ. of California Press.

[5] B. F. Logan and L. A. Shepp. A variational problem or random Young tableaux. *Adv. in Math.*, 26:206–222, 1977.

[6] P.A. Pevzner. *Computational Molecular Biology: An Algorithmic Approach*. MIT Press, 2000.

[7] D. Sankoff and J. Kruskal, editors. *Common subsequences and monotone subsequences*, chapter 17, pages 363–365. Addison–Wesley, Reading, Mass., 1983.

[8] C. Schensted. Longest increasing and decreasing subsequences. *Canad. J. Math.*, 13:179–191, 1961.

[9] R. P. Stanley. Recent progress in algebraic combinatorics. *Bulletin of the AMS*, 40(1):55–68, 2002.

[10] A. M. Vershik and S. V. Kerov. Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux. *Dokl. Akad. Nauk SSSR*, 233:1024–1028, 1977.

On the Power of Two Choices in Continuous Time

Colin McDiarmid
(joint work with Malwina Luczak)

Balls-and-bins processes have been useful for analysing a wide range of problems, in discrete mathematics and computer science, and in particular for problems which involve load sharing and resource balancing, see [8]. Here is one central result, from Azar, Broder, Karlin and Upfal (1994 [1],1999 [2]), concerning the 'power of two choices'. Let d be a fixed positive integer. Suppose that there are n bins, and n balls arrive one after another: each ball picks d bins uniformly at random and is placed in a least loaded of these bins. Then with probability tending to 1 as $n \to \infty$ (*aas*), the maximum load of a bin is $\ln n / \ln \ln n + O(1)$ if $d = 1$, and is $\ln \ln n / \ln d + O(1)$ if $d \geq 2$. Thus there is a dramatic drop when we move from 1 to 2 choices.

In some recent work, balls have been allowed to 'die' – see [2, 3, 9] – which is of course desirable when modelling telephone calls. For example, suppose that we start with n balls in n bins: at each time step, one ball is deleted uniformly at random, and one new ball appears and is placed in one of d bins as before. It is shown in [2] that as $n \to \infty$, at any given time $t \geq cn^2 \ln \ln n$, aas the maximum load of a bin is at most $\ln \ln n / \ln d + O(1)$.

Let us consider here a simple and natural 'immigration-death' balls-and-bins model in continuous time. Indeed let us consider two such models, one involving bins and one involving queues, first the bins.

Let d be a fixed positive integer. Suppose that there are n bins. Balls arrive in a Poisson process at rate λn, where $\lambda > 0$ is a constant. Upon arrival each ball chooses d random bins (with replacement), and is placed into a least-loaded bin among those chosen. (If there is more than one chosen bin with least load, the ball is placed in the first such bin chosen.) Balls have independent exponential lifetimes with unit mean.

Probabilists have proved various detailed weak-convergence results for such models, see for example [4, 9, 10], but these results seem not to say anything about quantities like the equilibrium maximum load. Using mainly combinatorial methods, we can establish concentration results, which apply to the fraction of bins with load at least k at time t; these concentration results may then be used to analyse a balance equation involving these quantities. We are thus able to handle random variables like the maximum load, over long periods of time. The system mixes rapidly, so let us focus on the stationary behaviour. (In fact, it is because the system mixes rapidly that we are able to prove our concentration results.)

Theorem 1 ([5]) *Let d be a fixed positive integer, and suppose that the n-bin* *system is in the stationary distribution. Then there is an integer-valued function* $m(n)$ *such that aas the maximum load is* $m(n)$ *or* $m(n)-1$*: if* $d=1$ *then* $m(n) = (1+o(1))\ln n/\ln\ln n$*, and if* $d \geq 2$ *then* $m(n) = \ln\ln n/\ln d + O(1)$*.*

Now consider a second continuous-time model, the *supermarket model.* This is as before except that now bins are replaced by queues, each with a single unit-rate server, and $\lambda < 1$. There are similar results for this model.

Theorem 2 ([6]) *Let* d *be a fixed positive integer, and suppose that the* n*-queue system is in the stationary distribution. If* $d=1$*, then aas the maximum queue length is within* $\omega(n)$ *of* $\ln n/\ln(1/\lambda)$*, where* $\omega(n)$ *is any function tending to* ∞*), and it is not concentrated on a bounded interval. If* $d \geq 2$ *then there is an integer valued function* $m(n) = \ln\ln n/\ln d + O(1)$ *such that aas the maximum queue length is* $m(n)$ *or* $m(n)-1$*.*

This is all joint work with Malwina Luczak. It arose from our endeavour to establish rigorous continuous-time results for routing in networks analogous to the discrete-time results in [7]. The 'bins' part of this work has recently been written up, the queues part nearly so: results on routing will follow later.

References

[1] Azar, Y., Broder, A., Karlin, A., Upfal, E. Balanced Allocations, *Proc. 26th ACM Symp. Theory Comp.* (1994) 593 – 602.

[2] Azar, Y., Broder, A., Karlin, A., Upfal, E. Balanced Allocations, *SIAM J. Comput.* **29** (1999) 180 – 200.

[3] Czumaj, A. Recovery time of dynamic allocation processes, *Proc. 10th Annual ACM Symp. on Parallel Algorithms and Architectures* (1998) 202 – 211.

[4] Graham, C. Kinetic Limits for Large Communication Networks, In *Modelling in Applied Sciences: A Kinetic Theory Approach*, N. Bellomo and M. Pulvirenti, eds., Birkhauser, 1999.

[5] Luczak, M.J. and McDiarmid, C. On the power of two choices: balls and bins in continuous time, preprint, 2003.

[6] Luczak, M.J. and McDiarmid, C. On the maximum queue length in the supermarket model, in preparation, 2004.

[7] Luczak, M.J., McDiarmid, C. and Upfal, E. On-line routing of random calls in networks, *Probab. Theory Relat. Fields.* **125** (2003) 457 – 482.

[8] Mitzenmacher, M., Richa, A.W. and Sitaraman, R. The power of two random choices: a survey of techniques and results, in preparation, 2002.

[9] Turner, S.R.E. The effect of increasing routeing choice on resource pooling, *Probability in Engineering and Informational Sciences* **12** (1998) 109 – 124.

[10] Vvedenskaya, N.D., Dobrushin, R.L., and Karpelevich, F.I. Queueing system with selection of the shortest of two queues: An asymptotic approach, *Problems of Information Transmission* **32** (1) (1996) 15–27.

Homomorphism Duality: On Short Answers to Exponentially Long Questions

Jaroslav Nešetřil

(joint work with Claude Tardif)

We give a new and more efficient construction of duals for general finite relational structures of a given type. We complement this by proving the superpolynomial lower bound for the size of the dual core. This bound is achieved even for duals of paths (i.e. for the type (2). This solves the main problem of [9].

Coloring problems belong to some of the central problems of combinatorics. Perhaps being encouraged by applications (such as channel assignement problems or Constraint Satisfaction type problems (CSP)) the recent revival of interest led to the investigation of many variants and far reaching generalizations, see e.g. [5, 3, 13]. The following problem captures both the difficulty and generality of some of this development:

H-coloring problem

Instance: A graph G;

Question: Does there exists a homomorphism $G \longrightarrow H$.

Recall, that a *homomorphism* $G \longrightarrow H$ is any mapping $f : V(G) \longrightarrow V(H)$ satisfying $f(x)f(y) \in E(H)$ whenever $xy \in E(G)$.

Thus for any complete graph $H = K_k$ the H-coloring problem reduces to the question whether the chromatic number $\chi(G)$ of graph G is $\leq k$. All CSP-problems may be expressed in a similar way as H-coloring problems for relational structures:

Let $\Delta = (\delta_i; i \in I)$ be a sequence of positive integers. A *relational structure of type* Δ (shortly Δ-structure) is a pair $(X, (R_i; i \in I))$ where X is a finite set and R_i is a δ_i-nary relation on X (i.e. we have $R_i \subset X^{\delta_i}$). Given a type Δ and

δ-systems $A = (X, (R_i; i \in I))$ and $A' = (X', (R'_i; i \in I))$ a *homomorphism* is a mapping $f : X \longrightarrow X'$ satisfying for every $i \in I$

$$(f(x_1), f(x_2), \ldots, f(x_{\delta_i})) \in R'_i \text{ whenever } (x_1, x_2, \ldots, x_{\delta_i}) \in R_i.$$

Given a structure H of type Δ we define the H-coloring problem in the complete analogy to graphs (yes, despite using for Δ-systems symbols A, B and the like, we still want to reserve H for the template of the coloring problem).

Viewing all this one expects that H-coloring problems are difficult to handle and that such problems tend to be computationally hard. This is indeed the case for undirected graphs. But for other types, and already for type (2) corresponding to the directed graphs, the situation is very difficult and there are many polynomial instances and the whole problem seems to be presently very difficult: there are many polynomial instances and even more hard cases, see e.g.[3, 2, 1].

This paper is devoted to the study of polynomial instances of H-coloring problems. Among those perhaps the simplest are those coloring problems which can be characterized by a simple obstruction set, by forbidden structures of a single type. This is expressed by the notion of the *(singleton) homomorphism duality*:

We say that a pair (F, H) of Δ-structures is a *dual pair* if the following equivalence holds for every Δ-structure A:

$$F \not\longrightarrow A \text{ iff } A \longrightarrow H.$$

The Δ-structure H is also called the *dual of* F and it is denoted by D_F. Note that up to homomorphism equivalence the dual D_F is uniquelly determined. One also sees easily that the only dual pair for undirected graphs (up to the homomorphism equivalence) is the pair (K_2, K_1), see [8] where this notion was first isolated. However one should not be discouraged by this as the richness of dualities lies in relational structures. Already for directed graphs (i.e. the type (2)) the duality pairs include pairs (P_k, T_k) where P_k is the monotone path of length k (i.e. with $k + 1$ vertices) and T_k is the transitive tournament with k vertices. One can see easily that these duality pairs correspond to the Hasse-Galai-Roy theorem: an undirected graph G has chromatic number $> k$ if and only if every orientation of G contains a monotone path of length k. Dualities represent a suprisingly rich scheme and many more dualities (and thus polynomial instances of coloring problems) were found [6, 7, 11]. Finally [9] characterize all homomorphism dualities (recall that a *core* of Δ structure A is the minimal structure which is homomorphism equivalent to A):

Theorem 1 *For every type Δ and for every Δ-tree T there exists a dual Δ-structure D_T. There are no other dual pairs.*

Viewing the difficulty of the classification of polynomial instances of H-coloring (already) for directed graphs it is perhaps surprising that one can achieve the full characterization of homomorphism dualities for general Δ-systems. The abundance of polynomial instances leads to the question about the nature of dual graphs. The proof given in [9] rests on some algebraic construction (such as the graph exponentiation) and on the reformulation of dualities in terms of homomorphism partial order $\mathcal{C}$ ("gaps" in $\mathcal{C}$). Thus dual structures D_T are complicated (and constructed indirectly) and their properties are non-trivial (and sometimes surprising, [10]). Thus it is desirable to have simpler explicite construction. Such a construction was provided in [11] for the case of directed graphs. This has been recently used in [12] to prove that the construction of the dual D_T is connected appart from isolated vertices.

In this paper we give a new construction of the dual for a general type Δ. This new construction is also more efficient: for a Δ-tree T with n vertices it produces the dual D_T of size $2^{n\log(n)}$ (as opposed to the double exponential bound which follows from [9]).

We complement this by providing examples which yield superpolynomial lower bound for cores of D_T. This improves the result of [9] and solves the main open problem left there. Perhaps surprisingly, in order to prove this lower bound we use relational structures (for large Δ).

The super polynomial lower bound for the size of core duals can be interpreted in the positive terms:

Corollary 2 *There are directed core graphs H such that $|V(H)| \geq 2^n$, and for every directed graph G, G is H-colorable if and only if every subgraph of G with at most $n\log(n)$ vertices is H-colorable.*

This an introduction to a paper by the same authors and the same title which is being submitted. It is available electronically at ITI Series and KAM-DIMATIA Series.

References

[1] A. Bulatov, A. Krokhin, P. Jeavons, Constraint Satisfaction Problems and Finite Algebras, Lecture Notes in Computer Science, vol. 1853, Springer Verlag 2000,p. 272–282.

[2] T. Feder, M. Vardi: The computational structure of monotone monadic SNP and constraint satisfaction: A study through datalog and group theory. SIAM J. of Computing 28 (1998), 57–104.

[3] P. Hell, J. Nešetřil, Graphs and Homomorphisms, Oxford University Press, 2003 (to appear).

[4] P. Hell, J. Nešetřil, X. Zhu, Duality and polynomial testing of tree homomorphisms. *Trans. Amer. Math. Soc.* **348** (1996), 1281–1297.

[5] T. Jensen, B. Toft: Graph Coloring Problems, Wiley-Interscience, New York , 1995.

[6] P. Komárek, Good characterisations in the class of oriented graphs (in czech), Ph. D. Thesis, Charles University, Prague, 1987.

[7] P. Komárek, Some new good characterizations for directed graphs. *Časopis Pěst. Mat.* **109** (1984), 348–354.

[8] J. Nešetřil, A. Pultr: On Classes of Relations and Graphs Determined by Subobjects and Factorobjects, Discrete Math. 22 (1978), 287–300.

[9] J. Nešetřil, C. Tardif, Duality Theorems for Finite Structures (Characterizing Gaps and Good Characterizations), *J. Combin. Theory Ser B* **80** (2000), 80–97.

[10] J. Nešetřil, C. Tardif, A dualistic approach to bounding the chromatic number of a graph, ITI Series 2001-036 (to appear).

[11] J. Nešetřil, C. Tardif, Density via Fuality, Theoretical Comp. Sci. 287, 2 (2002), 585–591.

[12] I. Švejdarová, Colouring of graphs and dual objects (in Czech), thesis, Charles University 2003.

[13] X. Zhu: Circular chromatic number: a survey, Discrete Math. 229 (2001), 371–410.

Extremal Connectivity for Topological Cliques

Deryk Osthus
(joint work with Daniela Kühn)

Given a natural number s, let $d(s)$ be the smallest number such that every graph of average degree $> d(s)$ contains a subdivision of the complete graph K_s of order s. The existence of $d(s)$ was proved by Mader [6]. As first observed by Jung [3], the complete bipartite graph $K_{t,t}$ with $t := \lfloor s^2/8 \rfloor$ shows that $d(s) \geq$

$\lfloor s^2/8 \rfloor$. Bollobás and Thomason [2] as well as Komlós and Szemerédi [4] showed that s^2 is the correct order of magnitude for $d(s)$. More precisely, it is known that

$$(1+o(1))\frac{9s^2}{64} \leq d(s) \leq (1+o(1))\frac{s^2}{2}. \tag{1}$$

The upper bound is due to Komlós and Szemerédi [4]. As observed by Łuczak, the lower bound is obtained by considering a random subgraph of a complete bipartite graph with edge probability 3/4. It is widely believed that the lower bound gives the correct constant, i.e. that random graphs provide the extremal graphs. If true, this would mean that the situation is similar as for ordinary minors. Indeed, Thomason [8] was recently able to prove that random graphs are extremal for minors and Myers [7] showed that all extremal graphs are essentially disjoint unions of pseudo-random graphs.

In [5] we showed that the lower bound in (1) is correct if we restrict our attention to bipartite graphs whose connectivity is close to their average degree:

Theorem 1 *Given $s \in \mathbb{N}$, let $c_{bip}(s)$ denote the smallest number such that every $c_{bip}(s)$-connected bipartite graph contains a subdivision of K_s. Then*

$$c_{bip}(s) = (1+o(1))\frac{9s^2}{64}.$$

In Theorem 1 the condition of being bipartite can be weakened to being H-free for some arbitrary but fixed 3-chromatic graph H. The proof of Theorem 1 builds on results and methods of Komlós and Szemerédi [4]. For arbitrary graphs, the best current upper bound on the extremal connectivity is the same as in (1). The proof of Theorem 1 yields the following improvement [5].

Theorem 2 *Given $s \in \mathbb{N}$, let $c(s)$ denote the smallest number such that every $c(s)$-connected graph contains a subdivision of K_s. Then*

$$(1+o(1))\frac{9s^2}{64} \leq c(s) \leq (1+o(1))\frac{s^2}{4}.$$

The lower bounds in Theorems 1 and 2 are provided by the random bipartite graphs mentioned above (since their connectivity is close to their average degree). Thus at least in the case of highly connected bipartite graphs they are indeed extremal.

By using methods as in the proof of Theorem 1, in [5] we also obtain a small improvement for the constant in the upper bound in (1).

Theorem 3 *Given $s \in \mathbb{N}$, let $d(s)$ denote the smallest number such that every graph of average degree $> d(s)$ contains a subdivision of K_s. Then*

$$(1+o(1))\frac{9s^2}{64} \leq d(s) \leq (1+o(1))\frac{10s^2}{23}.$$

The example of Łuczak mentioned above only gives us extremal graphs for Theorem 1 whose connectivity is about $3n/8$, i.e. whose connectivity is rather large compared to the order n of the graph. However, in [5] we showed that there are also extremal graphs whose order is arbitrarily large compared to their connectivity. In contrast to this, the situation for ordinary minors is quite different. In general a connectivity of order $s\sqrt{\log s}$ is needed to force a K_s minor, but the connectivity of the known extremal graphs is linear in their order. In fact, confirming a conjecture of Thomason [9], Böhme, Kawarabayashi and Mohar [1] proved that for all integers s there is an integer $n_0 = n_0(s)$ such that every graph of order at least n_0 and connectivity at least $45s$ contains the complete graph K_s as minor. Thus a linear connectivity suffices to force a K_s minor if we only consider sufficiently large graphs.

References

[1] T. Böhme, K. Kawarabayashi, B. Mohar, Linear connectivity forces dense minors, preprint 2003.

[2] B. Bollobás and A. Thomason, Proof of a conjecture of Mader, Erdős and Hajnal on topological complete subgraphs, *Eur. J. Comb.* **19** (1998), 883–887.

[3] H.A. Jung, Eine Verallgemeinerung des n-fachen Zusammenhangs für Graphen, *Math. Ann.* **187** (1970), 95–103.

[4] J. Komlós and E. Szemerédi, Topological cliques in graphs II, *Combin. Probab. Comput.* **5** (1996), 79–90.

[5] D. Kühn and D. Osthus, Extremal connectivity for topological cliques in bipartite graphs, submitted.

[6] W. Mader, Homomorphieeigenschaften und mittlere Kantendichte von Graphen, *Math. Annalen* **174** (1967), 265–268.

[7] J.S. Myers, Graphs without large complete minors are quasi-random, *Combin. Probab. Comput.*, to appear.

[8] A. Thomason, The extremal function for complete minors, *J. Combin. Theory B* **81** (2001), 318–338.

[9] A. Thomason, Extremal function for graph minors, preprint.

Constructions of Non-Principal Families in Extremal Hypergraph Theory

Oleg Pikhurko
(joint work with Dhruv Mubayi)

Here, we prove the non-principality phenomenon for the classical extremal problems for k-uniform hypergraphs. The main motivation is to study the qualitative difference between the cases $k = 2$, and $k \geq 3$, and our results for the Turán problem exhibit this difference.

Given a a family $\mathcal{F}$ of k-graphs, let $\mathrm{ex}(n, \mathcal{F})$ be the maximum size of an $\mathcal{F}$-free k-graph G on n vertices. Let $\pi(\mathcal{F})$ be the limit of $\mathrm{ex}(n, \mathcal{F})/\binom{n}{k}$ as $n \to \infty$. We call $\pi(\mathcal{F})$ the *Turán density* of $\mathcal{F}$.

Mubayi and Rödl [11] conjectured that there is a family $\mathcal{F}$ of 3-graphs such that

$$\pi(\mathcal{F}) < \min\{\pi(F) \mid F \in \mathcal{F}\}, \tag{1}$$

and commented that the result should even hold for a family $\mathcal{F}$ of size two. Balogh [1] proved the conjecture, calling this phenomenon the *non-principality* of the Turán function. This is in sharp contrast with the case of graphs ($k = 2$) where the Erdős-Stone-Simonovits Theorem [4, 2] applies.

However, Balogh's family has many graphs. Here we show how the so-called stability results lead to families $\mathcal{F}$ satisfying (1) and consisting of two k-graphs only. This approach succeeds for all even $k \geq 4$ and for $k = 3$, since it depends on stability results which are known only in these cases.

Non-Principal Families of Size 2

To obtain the *cone* $\mathrm{cn}(F)$ of a k-graph F, enlarge each edge of F by a new common vertex x:

$$\mathrm{cn}(F) := \{\{x\} \cup D \mid D \in F\}.$$

We call two order-n k-graphs F and G *ε-close* if we can make F isomorphic to G by adding and removing at most $\varepsilon\binom{n}{k}$ edges. A k-graph G is *F-extremal* if it is a maximum F-free k-graph of order $v(G)$. Let us call a k-graph F *stable* if any F-free k-graph G of order n with at least $(\pi(F) - o(1))\binom{n}{k}$ edges is $o(1)$-close to an F-extremal k-graph.

Lemma 1 *Let F be a stable k-graph. Suppose that we can find a k-graph H of order h such that $\pi(H) \geq \pi(F)$ and any F-extremal k-graph of order n contains*

$\Omega(n^h)$ copies of H. Then

$$\pi(\{F,H\}) < \min(\pi(F),\pi(H)). \tag{2}$$

Proof. Suppose on the contrary that $\pi(\{F,H\}) \ge \pi(F)$. Then there is an $\{F,H\}$-free k-graph G of order n and size $(\pi(F)-o(1))\binom{n}{k}$. Since F is stable, G is $o(1)$-close to an F-extremal k-graph G'. By hypothesis, G' contains $\Omega(n^h)$ copies of H. But each edge belongs to $O(n^{h-k})$ H-subgraphs, so we cannot destroy all of them by removing $o(n^k)$ edges. This is a contradiction to $G \not\subset H$. ∎

Theorem 2 *For even $k \ge 4$ and for $k=3$ there are k-graphs F and H satisfying (2).*

Proof. Let $k=2l$ be even. Let $F=\{A\cup B, A\cup C, B\cup C\}$, where A,B,C are disjoint l-sets. Frankl [5] showed that $\pi(F)=\frac{1}{2}$. Keevash and Sudakov [9, Theorem 3.4] showed that F is stable. Every extremal k-graph G' for F on $n \ge n_0$ vertices has vertex partition $X \cup Y$, $|X| \approx |Y| \approx \frac{n}{2}$, and consists of all edges intersecting X (and also Y) in an odd number of vertices.

Let us take $H=\mathrm{cn}(K_m^{k-1})$ where $m=m(k)$ is a sufficiently large integer to satisfy $\frac{k!}{m^k}\binom{m}{k} > \frac{1}{2}$. The latter implies that $\pi(H) > \frac{1}{2}$, because the blown-up K_m^k does not contain H. As G' contains $(2+o(1))\,\frac{n}{2}\binom{n/2}{m}$ copies of H, Lemma 1 implies that the family $\{F,H\}$ has the required properties.

For $k=3$ we can use the stability result either for the Fano plane, (established independently by Füredi and Simonovits [7] and by Keevash and Sudakov [8]), or for $F_{3,2}$, established by Füredi, Pikhurko, and Simonovits [6]. In both cases we can take $H=\mathrm{cn}(K_m^2)$ for some sufficiently large m. ∎

Concluding Remarks

For the case of odd $k \ge 5$, we can build upon the ideas in [1] and construct a non-principal k-graph family $\mathcal{F}$ for every $k \ge 3$, see [10]. The obtained family consists of finitely many k-graphs; however, this approach does not seem to give $|\mathcal{F}|=2$.

One can also consider the *Ramsey-Turán* density $\rho(\mathcal{F})$ where in addition to being $\mathcal{F}$-free we require that the maximum size of an independent set of G is $o(n)$. (This problem was introduced by Erdős and Sós [3].) One can show that for $k \ge 3$ if $\mathcal{F}$ is a non-principal k-graph family with respect to the Turán density then $\mathcal{F}(2)$ is non-principal with respect to the Ramsey-Turán density, see [10]. Here $\mathcal{F}(2)$ is obtained by blowing-up each member of $\mathcal{F}$ by factor of 2.

Curiously, the situation with graphs remains open.

Problem 3 *Do there exist* 2*-graphs* G_1, G_2 *for which*

$$\rho(\{G_1, G_2\}) < \min\{\rho(G_1), \rho(G_2)\}?$$

What about if we require $\rho(\{G_1, G_2\}) > 0$ *as well?*

References

[1] J. Balogh. The Turán density of triple systems is not principal. *J. Combin. Theory* (A), 100:176–180, 2002.

[2] P. Erdős and M. Simonovits. A limit theorem in graph theory. *Stud. Sci. Math. Hungar.*, pages 51–57, 1966.

[3] P. Erdős and V. T. Sós. On Ramsey-Turán type theorems for hypergraphs. *Combinatorica*, 2:289–295, 1982.

[4] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[5] P. Frankl. Asymptotic solution of a Turán-type problem. *Graphs Combin.*, 6:223–227, 1990.

[6] Z. Füredi, O. Pikhurko, and M. Simonovits. Triple systems with empty neighborhoods. Submitted, 2004.

[7] Z. Füredi and M. Simonovits. Triple systems not containing a Fano configuration. Submitted, 2003.

[8] P. Keevash and B. Sudakov. The exact Turán number of the Fano plane. Submitted, 2003.

[9] P. Keevash and B. Sudakov. On a hypergraph Turán problem of Frankl. Submitted, 2003.

[10] D. Mubayi and O. Pikhurko. Constructions of non-principal families in extremal hypergraph theory. Submitted, 2003.

[11] D. Mubayi and V. Rödl. On the Turán number of triple systems. *J. Combin. Theory* (A), 100:135–152, 2002.

The Phase Transition in the Uniformly Grown Random Graph has Infinite Order

Oliver Riordan
(joint work with Béla Bollobás and Svante Janson)

The emergence of a giant component is one of the most frequently studied phenomena in the theory of random graphs. Much of the interest is due to the fact that a giant component in a finite graph corresponds to an infinite component, or 'infinite cluster', in percolation on an infinite graph. In fact, it can be argued that it is more important and more difficult to study detailed properties of the emergence of the giant component than to study the corresponding infinite percolation near the critical probability.

The quintessential example of the emergence of a giant component is in the classical random graph model $G_{n,p}$, the graph with vertex set $\{1, 2, \ldots, n\}$ in which each pair of vertices is joined with probability p, independently of all other pairs. Let us say that an event holds *with high probability* (**whp**), if it holds with probability tending to 1 as $n \to \infty$. In 1960, Erdős and Rényi [7, 8] showed that the critical probability for $G_{n,p}$ is $1/n$: if $c < 1$ is a constant then **whp** the largest component of $G_{n,c/n}$ has $O(\log n)$ vertices, while there is a function $\theta(c) > 0$ such that for constant $c > 1$, **whp** $G_{n,c/n}$ has a component of order $(\theta(c) + o(1))n$, and no other component of order larger than $O(\log n)$. The proper 'window' of the phase transition was found much later by Bollobás [1] and Łuczak [10]. In $G_{n,c/n}$ the giant component emerges rather rapidly: the right-derivative of $\theta(c)$ at $c = 1$ is 2; this makes the study of the phenomenon manageable.

Our task here is considerably harder, since in the model we shall study the giant component emerges much more slowly. Our model, $G_n(c)$, is the finite version of a model first proposed by Dubins in 1984 (see [9, 11]): it is parametrized by n, the number of vertices, and a constant $c > 0$ to which edge probabilities are proportional, just as for $G_{n,c/n}$. It can be read out of results of Kalikow and Weiss [9] and Shepp [11] that there is a critical value $c = 1/4$ above which a giant component is present. In $G_n(c)$, the transition from having no giant component to having a giant component is rather tantalizing, since it is very slow indeed. It turns out that for any c less than $1/4$, **whp** the largest component of $G_n(c)$ already contains $n^{\Theta(1)}$ vertices, which is *much* larger than the $O(\log n)$ we have in $G_{n,a/n}$, $a < 1$. For $c > 1/4$, **whp** there is a giant component of order proportional to n, and the other components are small. In fact, there is a function $\phi(c)$, equal to 0 for $c \leq 1/4$ but positive for $c > 1/4$, such that **whp** the largest component of $G_n(c)$ has order $(\phi(c) + o(1))n$. However, rather than having positive right-derivative at the critical point, in this case (if the derivatives exist) every derivative of $\phi(c)$ at $c = 1/4$ is zero. This phenomenon is often called a phase transition of *infinite* *order.* Somewhat surprisingly, in spite of this extremely gentle growth of the giant component, we can give good bounds on $\phi(c)$ from above and below, showing, in particular, that $\phi(1/4+\epsilon) = o(\epsilon^k)$ for every k.

A somewhat similar, although less surprising, phenomenon was studied in [2], where for a different model it was shown that for *every* positive value of the appropriate parameter c there is a giant component, but its normalized size has all derivatives zero at $c = 0$. Nevertheless, a gentle increase at the very beginning is considerably less suprising than a 'sudden' gentle increase in a function which is zero up to some positive value.

Turning to the model, in [3], Callaway, Hopcroft, Kleinberg, Newman and Strogatz introduced a simple new model (which we shall call the CHKNS model) for random graphs growing in time. They gave heuristic arguments to find the critical point for the percolation phase transition in this graph, and numerical results (from integrating an equation, rather than just simulating the graph) to suggest that this transition has infinite order. Heuristic arguments for an infinite order phase transition in this and other models have been given by Dorogovtsev, Mendes and Samukhin [4].

Here we consider an even simpler and more natural model, the *uniformly grown random graph*, or '*$1/j$-graph*'. This is the finite version of a model proposed by Dubins in 1984. We define the $1/j$-graph $G_n^{1/j}$ as the random graph on $\{1, 2, \ldots, n\}$ in which each pair $i < j$ of vertices is joined independently with probability $1/j$. We may think of $G_n = G_n^{1/j}$ as a graph growing in time, where each vertex joins to a set of earlier vertices chosen uniformly at random, the set itself having a random size, which is essentially Poisson with mean 1. We study the random subgraph $G_n(c)$ of G_n obtained by selecting edges independently with probability $c < 1$. Of course, $G_n(c)$ can be defined directly by specifying that each pair $i < j$ is joined independently with probability c/j. With this definition, values of c greater than one make sense, provided we replace c/j by $\max\{c/j, 1\}$.

Kalikow and Weiss [9] showed that for $c < 1/4$ the infinite version $G_\infty(c)$ of $G_n(c)$ is disconnected with probability one. It is implicit in their work that **whp** the largest component in the finite graph $G_n(c)$, $c < 1/4$, has order $o(n)$. In the other direction, Shepp [11] showed that for $c > 1/4$, $G_\infty(c)$ is connected with probability 1; his proof involved showing that $G_n(c)$ has a component of order $\Theta(n)$ with probability bounded away from zero. Hence, the threshold for the emergence of a giant component in $G_n(c)$ is at $c = 1/4$. A similar result for a considerably more general model was proved by Durrett and Kesten [6].

Here we study the size of the giant component above the threshold, showing that the giant component emerges very slowly.

Theorem 1 *There is a function $\phi(c)$ such that as $n \to \infty$ with $c \geq 0$ fixed,* **whp** *the largest component of $G_n^{1/j}(c)$ contains $(\phi(c) + o(1))n$ vertices.*

Furthermore, $\phi(c) = 0$ *for* $c \leq 1/4$*, and*

$$\phi(c) = \exp\left(-\frac{\pi + o(1)}{2\sqrt{c - 1/4}}\right)$$

as c *tends to* $1/4$ *from above.*

In particular, $\phi(1/4 + \epsilon) = o(\epsilon^k)$ for any k, and the phase transition is of 'infinite order'.

Although we work with the $1/j$-graph, as it has a simpler and more natural static description, all our results carry over to the CHKNS model. As pointed out independently by Durrett [5], this is also true of the earlier threshold results, which predate the CHKNS model by 10 years!

References

[1] B. Bollobás, The evolution of random graphs, *Trans. Amer. Math. Soc.* **286**, 257–274 (1984).

[2] B. Bollobás and O. Riordan, Robustness and vulnerability of scale-free random graphs, *Internet Mathematics* **1** (2003), 1–35.

[3] D.S. Callaway, J.E. Hopcroft, J.M. Kleinberg, M.E.J. Newman and S.H. Strogatz, Are randomly grown graphs really random?, *Phys. Rev. E* **64** (2001), 041902.

[4] S.N. Dorogovtsev, J.F.F. Mendes and A.N. Samukhin, Anomalous percolation properties of growing networks, *Phys. Rev. E* **64** (2001), 066110.

[5] R. Durrett, Rigorous result for the CHKNS random graph model, Proceedings, Discrete Random Walks 2003, Cyril Banderier and Christian Krattenthaler, Eds. *Discrete Mathematics and Theoretical Computer Science* **AC** (2003), 95–104. `http://dmtcs.loria.fr/proceedings/`

[6] R. Durrett and H. Kesten, The critical parameter for connectedness of some random graphs, in *A tribute to Paul Erdős*, Cambridge Univ. Press, Cambridge, 1990, pp 161–176,

[7] P. Erdős and A. Rényi, On the evolution of random graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5**, 17–61 (1960).

[8] P. Erdős and A. Rényi, On the evolution of random graphs, *Bull. Inst. Internat. Statist.* **38**, 343–347, (1961).

[9] S. Kalikow and B. Weiss, When are random graphs connected?, *Israel J. Math.* **62** (1988), 257–268.

[10] T. Łuczak, Component behavior near the critical point of the random graph process, *Random Structures and Algorithms* **1**, 287–310 (1990).

[11] L.A. Shepp, Connectedness of certain random graphs. *Israel J. Math.* **67** (1989), 23–33.

The Regularity Method for k-uniform Hypergraphs

Vojtěch Rödl
(joint work with Brendan Nagle, Mathias Schacht and Jozef Skokan)

The Regularity Lemma of Szemerédi [20], proved to be a powerful tool in Combinatorics. This lemma states that all sufficiently large graphs can be approximated, in some sense, by random graphs. Since "random-like" graphs are often easier to handle than arbitrary graphs, the Regularity Lemma is especially useful in situations when the problem in question is easier to prove for random graphs.

Let $G = (V, E)$ be a graph and A, $B \subseteq V$ be a pair of disjoint sets of vertices of G. Denote by $e(A, B)$ the number of edges of G between A and B. The *density* of the pair (A, B) is defined by $d(A, B) = e(A, B)/(|A||B|)$. The pair is called *$\varepsilon$-regular* if for any $A' \subseteq A$, $B' \subseteq B$ with $|A'| \geq \varepsilon|A|$, $|B'| \geq \varepsilon|B|$, we have $|d(A, B) - d(A', B')| < \varepsilon$.

Theorem 1 (Szemerédi's Regularity Lemma) *For every $\varepsilon > 0$ there exist a T_0 such that the vertex set $V(G)$ of any graph G can be partitioned into $t \leq T_0$ classes $V(G) = V_1 \cup \cdots \cup V_t$, so that all but εt^2 pairs (V_i, V_j) are ε-regular.*

Many applications of the Regularity Lemma are based on its accompanying Counting Lemma (see, e.g., [9, 10] for a survey).

Theorem 2 (Counting Lemma) *If G is an ℓ-partite graph with $V(G) = V_1 \cup \cdots \cup V_\ell$ and $|V_i| = n$ for all $i \in [\ell]$, and all pairs (V_i, V_j) are ε-regular of density d for $1 \leq i < j \leq \ell$, then G contains $(1 \pm f_\ell(\varepsilon))d^{\binom{\ell}{2}} \times n^\ell$ cliques K_ℓ of order ℓ, where $f_\ell(\varepsilon) \to 0$ as $\varepsilon \to 0$.*

We discuss a generalization of Szemerédi's Regularity Lemma from graphs to k-uniform hypergraphs, which allows us to prove an accompanying Counting Lemma.

Unlike for graphs, there are several "natural ways" to define "regularity" for k-uniform hypergraphs. Consequently, various forms of a Regularity Lemma for hypergraphs have been already considered in [1, 2, 4, 6, 13]. None of these Regularity Lemmas seemed to admit a companion counting result (i.e., a corresponding generalization of Theorem 2). The first attempt of developing a Hypergraph Regularity Lemma with a corresponding Counting Lemma was undertaken by Frankl and the speaker in [5] for 3-uniform hypergraphs. Recently, the speaker in collaboration with Skokan [17] established a generalization of this Regularity Lemma to k-uniform hypergraphs for any $k \geq 3$.

Analogously to the feature that Szemerédi's Regularity Lemma decomposes a given graph into an ε-regular partition, this Hypergraph Regularity Lemma decomposes the edge set of a given k-uniform hypergraph into constantly many "blocks", almost all of which are, in a specific sense, "quasi-random". The concept of hypergraph regularity which plays the analogous role of the ε-regular pair is, unfortunately, considerably more technical than its graph counterpart, and we cannot give the precise definitions here.

Just as Theorem 2, the Counting Lemma, is an important companion statement to Szemerédi's Regularity Lemma, most applications of the Hypergraph Regularity Lemma from [17] require a similar companion lemma - the "general Counting Lemma". Analogously to Theorem 2, the general Counting Lemma estimates the number of copies of the clique $K_\ell^{(k)}$ (i.e., the complete k-uniform hypergraph on ℓ vertices) contained in an appropriate collection of "dense and regular blocks" within a regular partition provided by the Hypergraph Regularity Lemma. Such a Counting Lemma was established for special cases ($k = 3$, $\ell > 3$ and $k = 4$, $\ell = 5$) in [5, 11, 16]. Recently, in [12] Nagle, Schacht and the speaker, succeeded to prove the general Counting Lemma for any $\ell > k \geq 2$, reducing it to an earlier result from [8]. This Counting Lemma together with the Hyergraph Regularity Lemma of [17] can be viewed as a generalization of the Regularity Method from graphs to uniform hypergraphs. A similar extension was independently obtained by Gowers [7].

These generalizations can be applied to several extremal hypergraph problems. In particular, answering a question of Erdős, Frankl, and speaker [3], we proved the following theorem in [15]

Theorem 3 *Suppose an n-vertex k-uniform hypergraph $\mathcal{H}$ contains only $o(n^\ell)$ copies of $K_\ell^{(k)}$. Then one can delete $o(n^k)$ edges of $\mathcal{H}$ to make it $K_\ell^{(k)}$-free.*

It is known that this theorem can be used to give an alternative proof the well-known Density Theorem of Szemerédi regarding the upper density of sets containing no arithmetic progression of fixed length (see [5, 15]). Moreover, it can also be used to derive combinatorial proofs to some of the density theorems of Furstenberg and Katznelson (see [7, 14, 18]).

References

[1] F. R. K. Chung, *Regularity lemmas for hypergraphs and quasi-randomness*, Random Structures Algorithms **2** (1991), no. 2, 241–252.

[2] A. Czygrinow and V. Rödl, *An algorithmic regularity lemma for hypergraphs*, SIAM J. Comput. **30** (2000), no. 4, 1041–1066 (electronic).

[3] P. Erdős, P. Frankl, and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin. **2** (1986), no. 2, 113–121.

[4] P. Frankl and V. Rödl, *The uniformity lemma for hypergraphs*, Graphs Combin. **8** (1992), no. 4, 309–312.

[5] ———, *Extremal problems on set systems*, Random Structures Algorithms **20** (2002), no. 2, 131–164.

[6] A. Frieze and R. Kannan, *Quick approximation to matrices and applications*, Combinatorica **19** (1999), no. 2, 175–220.

[7] W. T. Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, manuscript.

[8] Y. Kohayakawa, V. Rödl, and J. Skokan, *Hypergraphs, quasi-randomness, and conditions for regularity*, J. Combin. Theory Ser. A **97** (2002), no. 2, 307–352.

[9] J. Komlós, A. Shokoufandeh, M. Simonovits, and E. Szemerédi, *The regularity lemma and its applications in graph theory*, Theoretical aspects of computer science (Tehran, 2000), Lecture Notes in Comput. Sci., vol. 2292, Springer, Berlin, 2002, pp. 84–112.

[10] J. Komlós and M. Simonovits, *Szemerédi's regularity lemma and its applications in graph theory*, Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993), Bolyai Soc. Math. Stud., vol. 2, János Bolyai Math. Soc., Budapest, 1996, pp. 295–352.

[11] B. Nagle and V. Rödl, *Regularity properties for triple systems*, Random Structures Algorithms **23** (2003), no. 3, 264–332.

[12] B. Nagel, V. Rödl, and M. Schacht, *The counting lemma for regular k-uniform hypergraphs*, manuscript.

[13] H. J. Prömel and A. Steger, *Excluding induced subgraphs. III. A general asymptotic*, Random Structures Algorithms **3** (1992), no. 1, 19–31.

[14] V. Rödl, M. Schacht, E. Tengan, and N. Tokushige, *Density theorems and extremal hypergraph problems*, manuscript.

[15] V. Rödl and J. Skokan, *Applications of the regularity lemma for uniform hypergraphs*, manuscript.

[16] ———, *Counting subgraphs in quasi-random 4-uniform hypergraphs*, submitted.

[17] ———, *Regularity lemma for k-uniform hypergraphs*, to appear in Random Structures Algorithms.

[18] J. Solymosi, *Note on a question of Erdős and Graham*, to appear in Combin. Probab. Comput.

[19] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245, Collection of articles in memory of Juriĭ Vladimirovič Linnik.

[20] ———, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), CNRS, Paris, 1978, pp. 399–401.

Graph Parameters and Reflection Positivity

Alexander Schrijver

(joint work with Michael H. Freedman and László Lovász [1])

We characterize which real-valued (undirected) graph parameters are of the following type, where H is a graph and $\alpha : VH \to \mathbb{R}_+$ and $\beta : EH \to \mathbb{R}$:

$$f_{H,\alpha,\beta}(G) := \sum_{\substack{\phi: VG \to VH \\ \phi \text{ homomorphism}}} \Big(\prod_{v \in VG} \alpha_{\phi(v)}\Big)\Big(\prod_{uv \in EG} \beta_{\phi(u)\phi(v)}\Big). \tag{1}$$

Here $\phi : VG \to VH$ is a *homomorphism* if $\phi(u)\phi(v) \in EH$ for all $uv \in EG$. (So if $\phi(u) = \phi(v)$, then H has a loop at $\phi(u)$.) To reduce technicalities, it has turned out to be convenient to assume that G has no loops but may have multiple edges, while H has no multiple edges but may have loops.

Several graph parameters are indeed of this type. A first example of such a parameter is $f(G) :=$ the number of k-vertex-colourings of G (for some fixed k). Then we can take $H = K_k$ (the complete loopless graph on k vertices), and α and β the all-one functions on VG and EG respectively. More generally, by taking any graph H and $\alpha \equiv 1$ and $\beta \equiv 1$, $f(G)$ counts the number of homomorphism of G into H. By taking H to be a two-vertex graph with one edge connecting the two vertices and a loop at one of the two vertices, $f(G)$ then counts the number of stable sets of G.

Other examples are given by the partition functions of several models in statistical mechanics. Then H can be taken to be a complete graph with all loops attached, and VH is interpreted as the set of states certain elements of a system G can adopt. The function $\beta : EH \to \mathbb{R}$ describes the energy of the interaction of two neighbouring states, while $\alpha : VH \to \mathbb{R}_+$ can be the external energy of the different states, or, alternatively, if $\sum_{v\in VH} \alpha_v = 1$, α_v may be the probability that an element is in state v. Then any function $\phi : VG \to VH$ is a *configuration* of system G, and $f_{H,\alpha,\beta}(G)$ is the total or average energy of the system. (A different interpretation of this model is in economics, where β gives the profit or cost of certain interactions, and $f_{H,\alpha,\beta}$ gives the expected profit or cost.)

It will follow from our theorem (but also a direct construction based on characters can be made) that also the following graph parameters are of the type above. Let Γ be a finite abelian group and let S be a subset of Γ with $-S = S$ (i.e., $-s \in S$ if $s \in S$). For any graph G, fix an arbitrary orientation. Call a function $x : EG \to \Gamma$ an *S-flow* if all values of x are in S and x satisfies the flow conservation law at each vertex v of G: the inflow is equal to the outflow. Let $f(G)$ be the number of S-flows. (Since $-S = S$, this number is independent of the orientation chosen.) A well-known example is when Γ is the cyclic group with k elements and $S = \Gamma \setminus \{0\}$. Then an S-flow corresponds to a nowhere-zero k-flow, and Tutte's nowhere-zero 5-flow conjecture says that $f(G) > 0$ if $k = 5$ and G has no bridges. (It can be shown that for the case of nowhere-zero k-flows, we can take for H the complete graph on k vertices with all loops attached, and set $\alpha(v) = 1/k$ for each $v \in VH$, $\beta(e) = k - 1$ for each nonloop edge e of H, and $\beta(e) = -1$ for each loop e of H.)

The question of characterizing the graph parameters of form (1) is motivated, among others, by the question of the physical realizability of certain graph parameters. It turns out that two conditions on certain matrices derived from the graph parameter are necessary and sufficient: restricted (namely exponential) growth of the ranks and positive semidefiniteness — a condition that corresponds to the well-known *reflection positivity* in statistical mechanics.

These matrices are described as follows. For any integer $k \geq 0$, let $\mathcal{G}_k$ be the set of graphs in which k of the vertices are labeled $1, \dots, k$, while the remaining vertices are unlabeled. For $G, G' \in \mathcal{G}_k$, let GG' denote the graph obtained by first taking the disjoint sum of G and G', and next identifying equally labeled vertices. (So GG' has $|VG| + |VG'| - k$ vertices.) For any graph parameter f, let $M_{f,k}$ be the (infinite) $\mathcal{G}_k \times \mathcal{G}_k$ matrix whose entry in position G, G' is equal to $f(GG')$.

Then for any graph parameter f (where K_0 is the graph with no vertices and edges):

Theorem 1 *There exist H, $\alpha : VH \to \mathbb{R}_+$ and $\beta : EH \to \mathbb{R}$ such that $f = f_{H,\alpha,\beta}$ if and only if $f(K_0) = 1$ and there exists a c such that each $M_{f,k}$ is positive semidefinite and has degree at most c^k.*

Necessity can be shown rather straightforwardly. The method for proving sufficiency is based on considering each $\mathcal{G}_k$ as a semigroup (taking GG' above as multiplication), making the semigroup algebra over $\mathcal{G}_k$, and taking the quotient algebra over the null-space of $M_{f,k}$, thus obtaining a finite-dimensional Banach algebra, which has a basis of idempotents. The interaction of the idempotents between these algebras for different values of k gives us the combinatorics to find H and the functions α and β.

Extension of this method gives similar results for directed graph and hypergraph parameters, and more generally for any parameter for systems that have a certain semigroup structure.

References

[1] M.H. Freedman, L. Lovász, A. Schrijver, Reflection positivity, rank connectivity, and homomorphisms of graphs, preprint, 2003.

Claw-free Graphs

Paul Seymour

(joint work with Maria Chudnovsky)

A graph is *claw-free* if no induced subgraph is isomorphic to the complete bipartite graph $K_{1,3}$. We give a structural description of all claw-free graphs with the additional property that every vertex is in a 3-vertex stable set.

One way to formulate our result is that, for every claw-free graph G, either G belongs to one of (about ten) well-understood basic classes of graphs, or G admits one of (about five) types of decomposition, or some vertex is not in a stable set of size 3. Having proved that, we can stand back and ask, what does this tell us about the "global structure" of G? And there is indeed a "structure theorem", but we are still working on its precise formulation, and for this abstract we confine ourselves to the decomposition theorem.

First, here are a few kinds of claw-free graphs.

- **Line graphs.** If H is a graph, its *line graph* $L(H)$ is the graph with vertex set $E(H)$, in which distinct $e, f \in E(H)$ are adjacent if and only if they have a common end in H.
- **The icosahedron.** This is the unique planar graph with twelve vertices all of degree five.
- **The Schläfli graph.** Let G be the graph with 27 vertices $a_{i,j,k}$ ($1 \leq i,j,k \leq 3$), and with adjacency as follows. $a_{i,j,k}$ is adjacent to $a_{i',j',k'}$ if and only if either

  - $k' = k$ and either $i' = i$ or $j' = j$, or
  - $k' = k + 1 \pmod 3$ and $j' \neq i$, or
  - $k' = k + 2 \pmod 3$ and $i' \neq j$.

- **Circular interval graphs.** Let Σ be a circle and let $F_1, \dots, F_k$ be subsets of Σ, each homeomorphic to the closed interval $[0, 1]$, and no three with union Σ. Let V be a finite subset of Σ, and let G be the graph with vertex set V in which $v_1, v_2 \in V$ are adjacent if and only $v_1, v_2 \in F_i$ for some i.

- **XX-configurations.** Let G be the graph with vertex set $\{v_1, \dots, v_{13}\}$, with adjacency as follows. v_1-$\cdots$-v_6 is a hole in G of length 6. Next, v_7 is adjacent to v_1, v_2; v_8 is adjacent to v_4, v_5, and possibly to v_7; v_9 is adjacent to v_6, v_1, v_2, v_3; v_{10} is adjacent to v_3, v_4, v_5, v_6, v_9; v_{11} is adjacent to $v_3, v_4, v_6, v_1, v_9, v_{10}$; v_{12} is adjacent to $v_2, v_3, v_5, v_6, v_9, v_{10}$; and v_{13} is adjacent to $v_1, v_2, v_4, v_5, v_7, v_8$.

- **An extension of $L(K_6)$.** Let H be the graph with seven vertices $h_0, \dots, h_6$, in which $h_1, \dots, h_6$ are pairwise adjacent and h_0 is adjacent to h_1. Let G be the graph obtained from the line graph $L(H)$ of H by adding one new vertex, adjacent precisely to the members of $V(L(H)) = E(H)$ that are not incident with h_1 in H.

- **The graph of crosses.** Let $k \geq 1$. Let G have vertex set the union of nine disjoint sets $A_{i,j}$ $(1 \leq i, j \leq 3)$, where $A_{2,1}$,$A_{2,3}$, $A_{1,2}$,$A_{3,2}$ all have cardinality k, and the other five have cardinality 1. Let every vertex of $A_{i,j}$ be adjacent to every vertex of $A_{i',j'}$ if either $i = i'$ or $j = j'$, and otherwise let there be no edges between $A_{i,j}$ and $A_{i',j'}$. Now we need to change the adjacency between the four sets $A_{2,1}, A_{2,3}, A_{1,2}, A_{3,2}$. Order each of these four sets. If u is the ith vertex of one of these four sets, say $A_{a,b}$, and v is the jth vertex of another of these sets, say $A_{c,d}$, let u, v be adjacent if either
  - $i = j$ and $a \neq c$ and $b \neq d$, or
  - $i \neq j$ and either $a = c$ or $b = d$.

- **The path of triangles.** Let G have vertices $v_1, \dots, v_n$ with n odd, in which for $i < j$, v_i is adjacent to v_j if either $j - i = 1$, or $j - i = 2$ and i is odd, or $j - i \geq 3$ and $j - i = 2 \bmod 3$.

For each of these types of graph, we regard the graphs of that type and all their induced subgraphs as forming one of our basic classes. These are the nicest of our classes; there are a few others, quite similar, that we omit. (We shall not attempt a precise statement of the theorem here.)

Next, decompositions. Two subsets X, Y of $V(G)$ with $X \cap Y = \emptyset$ are *complete* to each other if every vertex of X is adjacent to every vertex of Y, and *anticomplete* if no vertex in X is adjacent to a member of Y.

Distinct vertices u, v of G are *twins* (in G) if they are adjacent and have exactly the same neighbours in $V(G) \setminus \{u, v\}$. Admitting twins is the first of our decompositions.

Now let A, B be disjoint subsets of $V(G)$. The pair (A, B) is called a *homogeneous pair of cliques* if

- A, B are both cliques
- every vertex $v \in V(G) \setminus (A \cup B)$ is either A-complete or A-anticomplete and either B-complete or B-anticomplete, and
- A is neither complete nor anticomplete to B.

The third kind of decomposition is a *1-join*. Suppose that V_1, V_2 partition $V(G)$, and for $i = 1, 2$ there is a subset $A_i \subseteq V_i$ such that:

- for $i = 1, 2$, A_i is a clique, and $A_i, V_i \setminus A_i$ are both nonempty
- A_1 is complete to A_2
- every edge between V_1 and V_2 is between A_1 and A_2.

In these circumstances, we say that (V_1, V_2) is a *1-join*.

Next, suppose that V_0, V_1, V_2 are disjoint subsets with union $V(G)$, and for $i = 1, 2$ there are subsets A_i, B_i of V_i satisfying the following:

- for $i = 1, 2$, A_i, B_i are cliques, $A_i \cap B_i = \emptyset$ and A_i, B_i and $V_i \setminus (A_i \cup B_i)$ are all nonempty
- A_1 is complete to A_2, and B_1 is complete to B_2, and there are no other edges between V_1 and V_2, and
- V_0 is a clique; and for $i = 1, 2$, V_0 is complete to $A_i \cup B_i$ and anticomplete to $V_i \setminus (A_i \cup B_i)$.

We call the triple (V_1, V_0, V_2) a *2-join*. (This is closely related to, but not quite the same as, what has been called a 2-join in other papers.)

The fifth and last decomposition is the following. Let (V_1, V_2) be a partition of $V(G)$, such that for $i = 1, 2$ there are cliques $A_i, B_i, C_i \subseteq V_i$ with the following properties:

- For $i = 1, 2$ the sets A_i, B_i, C_i are pairwise disjoint and have union V_i

- V_1 is complete to V_2 except that there are no edges between A_1 and A_2, between B_1 and B_2, and between C_1 and C_2.

- V_1, V_2 are both nonempty.

In these circumstances we say that G is a *hex-join* of $G|V_1$ and $G|V_2$. Note that if G is expressible as a hex-join as above, then the sets $A_1 \cup B_2, B_1 \cup C_2$ and $C_1 \cup A_2$ are three cliques with union $V(G)$, and consequently no graph G with $\alpha(G) > 3$ admits a hex-join. ($\alpha(G)$ denotes the size of the largest stable set in G.)

Let us say a *triad* in G is a stable set of vertices with cardinality 3. Our main theorem, then, says:

Theorem 1 *For every connected claw-free graph in which every vertex belongs to a triad, either G belongs to one of the basic classes, or G admits either twins, a homogeneous pair of cliques, a 1-join, a 2-join or a hex-join.*

It is convenient to break the proof (and indeed, the full statement of the theorem) into four cases:

- $\alpha(G) \geq 4$

- $\alpha(G) \leq 3$, but there are four vertices so that only one pair of them is adjacent

- for every triad, every vertex not in X has exactly two neighbours in X, and every vertex is in a triad

- for every triad, every vertex not in X has exactly two neighbours in X, and some vertex is not in any triad.

In each case (except the fourth, where we have nothing to say), we have a result that "either G belongs to a basic class or G admits a decomposition", but the basic classes and decompositions are different for different types. We omit further details here. Some of these results are written in [1, 2].

References

[1] Maria Chudnovsky and Paul Seymour, "Claw-free graphs. II. Circular interval graphs", manuscript, October 2003.

[2] Maria Chudnovsky and Paul Seymour, "Claw-free graphs. III. Sparse decomposition", manuscript, October 2003.

Paradoxical Decompositions and Growth Properties

Vera T. Sós

The theory of paradoxical decompositions arose in connection with the existence of non-Lebesgue measurable sets.

The non-existence of isometry-invariant finitely additive measure in $\mathbb{R}^3$ was proved by Banach and Tarski (1924) [1] by means of paradoxical decomposition. They proved that it is possible to partition the unit ball in $\mathbb{R}^3$ into finitely many pieces and to rearrange them by rigid motions (using isometric transformations) to form two unit balls. This "duplication", this "paradoxical decomposition" of the ball at first seems to be impossible.

The analysis of this surprising phenomenon led to the concept of amenable groups introduced and studied first by von Neumann (1929) [10]. Since that time the subject developed into a field which has importance beside analysis, group theory and geometry in discrete mathematics and computer science (e.g., in the theory of random walks, percolation, expanders).

The Hausdorff-Banach-Tarski paradoxical decompositions of the ball (or of the sphere) in $\mathbb{R}^d$ exist for $d = 3$ (and also for $d > 3$), but do not exist for $d = 1$ and $d = 2$.

Von Neumann discovered that these different phenomena are due to the difference between the isometry groups of $\mathbb{R}^1, \mathbb{R}^2$ and $\mathbb{R}^3$, the latter one is more *"rich"*. He considered a general setting where the basic notions are the finitely additive group invariant measure (or invariant mean) and the paradoxical groups (or amenable groups=non-paradoxical groups).

The objective of the present talk is to give some illustrations and indications of the wide range of topics which developed from the subject mentioned above, providing some motivation of the particular problem considered in the paper of Deuber, Simonovits and Sós [3] and some of its aftermath.

In the paper [3] - - for an arbitrary metric space the concept of *wobbling transformations* (called more recently also *bounded perturbation of the identity*) is introduced.

Definition. Let (X, d) be a metric space, $A, B \subseteq X$. A bijection $f : A \to B$ is called a *wobbling bijection* if

$$\sup_{x \in A} d(x, f(x)) < \infty .$$

$A, B \subseteq X$ are called *wobbling equivalent* if there is a wobbling bijection $f : A \to B$.

Definition. The set $A \subseteq X$ is called *wobbling paradoxical* if there is a decomposition

$$A = A_1 \cup A_2, \qquad A_1 \cap A_2 = \emptyset$$

such that A, A_1, A_2 are pairwise wobbling equivalent.

In [3] wobbling paradoxicity is characterized by the following growth condition: For $A \subset X$, $k > 0$ let $N_k(A)$ denote the k-neighbourhood of A:

$$N_k(A) = \{x \in X : d(x, A) \leq k\} \ .$$

Definition. The metric space (X, d) is *doubling*, if there is a $k > 0$ such that

$$|N_k(A)| \geq 2|A| \text{for every finite} A \subset X.$$

Theorem 1 *Let (X, d) be a discrete and countable metric space. (X, d) is wobbling paradoxical if and only if it is doubling.*

In the lecture we surveyed the connection of wobbling paradoxicity to the amenability of groups, to theory of random walks on graphs and groups and some recent applications of the doubling property and wobbling paradoxicity.

A survey paper written jointly with Gábor Elek will appear in a Volume dedicated to the memory of Walter Deuber.

For detailed information and references about the extremely wide area the reader is referred to the excellent books like of Gromov [5], de la Harpe [6], Lubotzky [9], Paterson[11], Wagon [13], Woess [14], and the many survey papers on these subjects, e.g., by Ceccherini-Silberstein, Grigorchuk and de la Harpe [2], Laczkovich [7], [8], Thomassen and Woess [12].

References

[1] Banach, S. and Tarski, A. [1924] Sur la decomposition des ensembles de point en parties respectivement congruents, *Fund. Math.* **6**, 244-277.

[2] Ceccherini-Silberstein, T., Grigorchuk, R. I. and de la Harpe, P. [1999] Amenability and paradoxical decompositions for pseudogroups and discrete metric spaces (Russian), *Tr. Mat. Inst. Steklova* **224**, 68-111.

[3] Deuber, W. A., Simonovits, M. and Sós, V. T. [1995] A note on paradoxical metric spaces, *Studia Sci. Math. Hungar.* **30**, 17-23.

[4] Elek, G. and T. Sós, V., Paradoxical Decompositions and Growth Properties, (2004) submitted

[5] Gromov, M. [1993] Asymptotic invariants of infinite groups, *Volume2 of Geometric Group Theory* Ed. Niblo, G.A., and Roller, M.A. *London Math. Soc. Lecture Notes Series* **182**

[6] de la Harpe, P. [2000] Topics in geometric group theory, *Th University of Chicago Press*

[7] Laczkovich, M., [1994] Paradoxical decompositions: a survey of recent results, in: *First European Congress of Mathematics (Paris, July 6-10, 1992).* Progress in Mathematics, No. 120, Birkhäuser, 1994, Volume II, 159-184.

[8] Laczkovich, Miklós, Paradoxes in measure theory. Handbook of measure theory, Vol. I, II, 83–123, North-Holland, Amsterdam, 2002.

[9] Lubotzky, A. [1994] *Discrete Groups, Expanding Graphs and Invariant Measureas.* Progress in Mathematics, No. 125, Birkhäuser, 1994.

[10] von Neumann, J. [1929] Zur allgemeinen Theorie des Masses, *Fund. Math.* **13**, 73-116.

[11] Paterson, A.L.T. [1988] *Amenability.* Mathematical Surveys and Monographs No. 29, American Mathematical Society, Providence, Rhode Island, 1988.

[12] Thomassen, C. and Woess, W. [1994] Vertex transitive graphs and accessibility, *J. Comb.Theory B* **26**, 1-60.

[13] Wagon, S. [1986] *The Banach-Tarski paradox.* Second edition, Cambridge Univ. Press, 1986. First paperback edition, 1993.

[14] Woess, W., [2001] Random walks on infinite graphs and groups, *Cambridge Tracts in Mathematics*, **138**, Cambridge University Press.

On the Sparse Regularity Lemma

Angelika Steger
(joint work with S. Gerke, Y. Kohayakawa, V. Rödl)

Over the last decades Szemerédi's regularity lemma [17] has proven to be a very powerful tool in modern graph theory. Roughly speaking, the regularity lemma asserts that one can partition a graph G into a constant number of equal-size parts in such a way that most parts are pairwise ε-regular; see [1, 2, 14] for the precise statement of Szemerédi's regularity lemma and some applications. Unfortunately, in its original setting it only gives nontrivial results for dense graphs, that is graphs with $\Theta(n^2)$ edges. In 1996 Kohayakawa [10] and independently Rödl introduced a variant which holds for sparse graphs, provided they satisfy some additional structural conditions (which essentially mean that the graph does not contain too dense spots). However, using this sparse regularity lemma to prove extremal and Ramsey type results similar to the known results in the dense case, requires an additional key step, as Łuczak showed that one cannot directly generalise the methods used for dense graphs to the sparse case, see [12]. The missing step has been formulated as a conjecture by Kohayakawa, Łuczak and Rödl [11], see also [12]. Over the last few years this conjecture has already attracted considerable attention; see [9] and the references therein. One reason for the popularity of the conjecture is its connection with Turán-type problems in random graphs: if the KŁR conjecture is true for a graph H, then asymptotically almost surely (a.a.s.) the number of edges in any H-free subgraph of a binomial random graph $G_{n,p}$ is at most $((1-1/(\chi(H)-1)+\varepsilon)\binom{n}{2}p$ for any $\varepsilon > 0$ as long as $p > C(\varepsilon, H)n^{-1/d_2(H)}$. Here $\chi(H)$ denotes the chromatic number of H, and $d_2(H)$ denotes the 2-density of H. Observe that the bound on the number of edges in an H-free subgraph is essentially best possible since every graph G contains a $(\chi(H)-1)$-partite subgraph with $(1-1/(\chi(H)-1))|E(G)|$ edges. Also the result is not true for much smaller p since then a.a.s. the number of copies of H in $G_{n.p}$ is much smaller than the number of edges; see [9].

This Turán-type result has been established in a series of papers for various special cases, each requiring its own a tailor-made proof. It is now known when $H = K_3$ is a triangle [3], H is a cycle of arbitrary length [4, 7, 8], and when $H = K_4$ is the complete graph on four vertices [11]. If one only considers denser random graphs, where p is about the square root of the conjectured value, then the result is also known to be true for all complete graphs [13, 16].

In fact in their paper [11] Kohayakawa, Łuczak and Rödl not only proved the Turán problem for $H = K_4$, but also outlined a proof strategy based on the sparse regularity lemma which would prove the Turán result for general graphs H, if one could prove an equivalent of the well known embedding lemma for dense graphs in the sparse context as well. They formulated this requirement as a conjecture — the above mentioned KŁR-conjecture.

In the remainder of this abstract we first state the KŁR-conjecture precisely and then report on recent achievements.

Definition 1 *A bipartite graph $B = (U \dot\cup W, E)$ is called (ε, p)-regular if for all $U' \subseteq U$ and $W' \subseteq W$ with $|U'| \geq \varepsilon n$ and $|W'| \geq \varepsilon n$,*

$$\left| \frac{|E(U', W')|}{p \cdot |U'| \cdot |W'|} - \frac{|E(U, W)|}{p \cdot |U| \cdot |W|} \right| \leq \varepsilon.$$

If instead all such $U' \subseteq U$ $W' \subseteq W$ just satisfy

$$|E(U', W')| \ \geq \ p \cdot \lambda \cdot |U'| \cdot |W'|$$

for some constant $\lambda > 0$ the graph $B = (U \dot\cup W, E)$ is called $(\varepsilon, p, \lambda)$-lower regular.

Definition 2 *Let H be a graph on l vertices. An l-partite graph $G = (V_1 \cup \ldots \cup V_l, E)$ on l pairwise disjoint vertex sets V_i of size n each is called (H, n, m, ε)-regular if the graph induced by V_i, V_j is $(\varepsilon, m/n^2)$-regular whenever $\{i, j\} \in E(H)$ and there are no edges between V_i and V_j otherwise. The set of all (H, n, m, ε)-regular graphs is denoted by $\mathcal{S}(H; n, m, \varepsilon)$, and $\mathcal{F}(H; n, m, \varepsilon)$ is the set of all graphs in $\mathcal{S}(H; n, m, \varepsilon)$ not containing H as a subgraph.*

The KŁR-conjecture can now be formulated as follows.

Conjecture 3 *Let H be an arbitrary graph and $\beta > 0$, then there exist positive constants ε_0, C, n_0 such that*

$$|\mathcal{F}(H, n, m, \varepsilon)| < \beta^m \binom{n^2}{m}^{e(H)}$$

for all $m \geq Cn^{2-1/d_2(H)}$, $n \geq n_0$ and $0 < \varepsilon \leq \varepsilon_0$, where $d_2(H) = \max\{\frac{e(F)-1}{v(F)-2} : F \subseteq H,\ v(F) > 2\}$.

Note that $\binom{n^2}{m}^{e(H)}$ is the number of graphs which are "blow-ups" of H, and it is not hard to see that it is also asymptotically equal to $|\mathcal{S}(H, n, m, \varepsilon)|$, so the conjecture asserts that only an exponentially small fraction β^m of such graphs are H-free. It was shown by Łuczak that $|\mathcal{F}(H, n, m; \varepsilon)| > 0$ for some graphs H, see [12] where Łuczak is quoted.

The conjecture is easily seen to be true for trees. It is also known to be true for cycles [15] and for the complete graphs $H = K_4$ and K_5 on four respectively five vertices [5, 6].

One of the key difficulties in the proof of the KŁR-conjecture is the fact that for $m = o(n^2)$ the size of a neighbourhood of a vertex is on average $o(n)$. The definition of regularity, however, only deals with linear sized subsets and thus regularity seem to be not inherited by subgraphs induced on the neighborhoods of some vertices. Recently we were able to prove that nevertheless in the sparse case a hereditary version holds as well.

Theorem 4 *For all $\beta, \varepsilon', \lambda > 0$ there exist $\varepsilon, C > 0$ such that for all $(\varepsilon, p, \lambda)$-regular graphs $B = (U \dot\cup W, E)$ the following holds. For all $q \geq C(\lambda p)^{-1}$ there exist at most $\beta^q \binom{|U|}{q}$ sets $Q \subseteq U$ such that (Q, W) is not $(\varepsilon', p, \lambda\varepsilon'/32)$-lower regular.*

This lemma readily implies much shorter and elegant proofs of the results known so far. It can also be used to prove the Turán result for $H = K_6$ and, hopefully, more general results in the near future.

References

[1] B. Bollobás. Modern Graph Theory, *Springer* 1998.

[2] R. Diestel. Graph Theory, *Springer* 1997

[3] P. Frankl and V. Rödl. Large triangle-free subgraphs in graphs without K_4. *Graphs and Combinatorics* 2, 1986, pp. 135–144.

[4] Z. Füredi. Random Ramsey graphs for the four-cycle. *Discrete Mathematics* 126, 1994, pp. 407–410.

[5] S. Gerke, H. J. Prömel, T. Schickinger, A. Steger, and A. Taraz. K_4-free subgraphs of random graphs revisited. Submitted for publication, 2002.

[6] S. Gerke, T. Schickinger, and A. Steger. K_5-free subgraphs of random graphs. *Random Structures & Algorithms*, accepted for publication.

[7] P.E. Haxell, Y. Kohayakawa, and T. Łuczak. Turán's extremal problem in random graphs: forbidding even cycles. *Journal of Combinatorial Theory* B 64, 1995, pp. 273–287.

[8] P. E. Haxell, Y. Kohayakawa, and T. Łuczak. Turán's extremal problem in random graphs: forbidding odd cycles.*Combinatorica* 16(1), 1996, pp. 107–122.

[9] S. Janson, T. Łuczak, and A. Rucinski. Random Graphs. *John Wiley & Sons,* 2000.

[10] Y. Kohayakawa. Szemerédi's regularity lemma for sparse graphs. In: Foundations of Computational Mathematics (Berlin, Heidelberg) (F. Cucker and M. Shub, eds.), *Springer-Verlag,* 1997, pp. 216-230.

[11] Y. Kohayakawa, T. Łuczak, and V. Rödl. On K^4-free subgraphs of random graphs. *Combinatorica* 17(2), 1997, pp. 173–213.

[12] Y. Kohayakawa, and V. Rödl. Regular pairs in sparse random graphs. *Random Structures & Algorithms* 22 (4), 2003, pp. 359-434.

[13] Y. Kohayakawa, V. Rödl, and M. Schacht. The Turán Theorem for random graphs. *Combinatorics, Probability and Computing*, accepted for publication.

[14] J. Komlós and M.Simonovits. Szemúredi's regularity lemma and its application in graph theory (in Combinatorics - Paul Ersős is eighty, D. Miklós, V.T. Sós, and T. Szőnyi, eds.) Bolyai Society of Mathematical Studies Vol.2, 1996, pp. 295–352, Budapest.

[15] B. Kreuter. Probabilistic versions of Ramsey's and Turán's theorems. Ph. D. thesis, Humboldt-Universität zu Berlin, 1997, Germany.

[16] T. Szabó and V.H. Vu. Turán's Theorem in sparse random graphs. *Random Structures & Algorithms*, accepted for publication.

[17] E. Szemúredi. Regular partitions of graphs. In Problémes Combinatoires et Théorie des Graphes, Colloques Internationaux CNRS Vol.260, 1978, pp. 399–452.

Solving Extremal Problems Using Stability Theorems

Benjamin Sudakov

(joint work with P. Keevash and in part with N. Alon and J. Balog)

In this talk we discuss a 'stability approach' for solving extremal problems. Roughly speaking, it can be described as follows. In order to show that given configuration is a unique optimum for an extremal problem, we first prove an approximate structure theorem for all constructions whose value is close to the optimum and then use this theorem to show that any imperfection in the structure must lead to a suboptimal configuration. To illustrate this strategy, we use the following results.

- Let $T_k(n)$ be the *Turán graph*, i.e., the complete k partite graph on n vertices with class sizes as equal as possible and denote by $t_k(n)$ the number of edges in $T_k(n)$. Then for $k \geq 2$ and sufficiently large n every graph G on n vertices has at most $2^{t_k(n)}$ distinct 2-edge colorings without a monochromatic clique of size $k+1$. Moreover the equality is only possible if $G = T_k(n)$. This settles a conjecture of Yuster. Our proof is based on Szemerédi's regularity lemma together with some additional tools in Extremal Graph Theory, and provide one of the rare examples of a precise result proved by applying this lemma.
- The Fano plane is a 3-uniform hypergraph with 7 triples on 7 vertices whose edges correspond to the lines of the projective plane over the field with two elements. We show that the maximum number of triples on n vertices not

containing a copy of the Fano plane can be obtain by partitioning vertices into two equal parts and taking all the triples which intersect both of them. This confirms a conjecture of V. Sós from 1976 which was also independently proved by Füredi and Simonovits.

- Let $\mathcal{C}_r^{(2k)}$ be the $2k$-uniform hypergraph obtained by letting $P_1, \cdots, P_r$ be pairwise disjoint sets of size k and taking as edges all sets $P_i \cup P_j$ with $i \neq j$. This can be thought of as the 'k-expansion' of the complete graph K_r: each vertex has been replaced with a set of size k. An example of a hypergraph with vertex set V that does not contain $\mathcal{C}_3^{(2k)}$ can be obtained by partitioning $V = V_1 \cup V_2$ and taking as edges all sets of size $2k$ that intersect each of V_1 and V_2 in an odd number of elements. Let $\mathcal{B}_n^{(2k)}$ denote a hypergraph on n vertices obtained by this construction that has as many edges as possible. We prove a conjecture of Frankl, which states that any hypergraph on n vertices that contains no $\mathcal{C}_3^{(2k)}$ has at most as many edges as $\mathcal{B}_n^{(2k)}$.

  Sidorenko has given an upper bound of $\frac{r-2}{r-1}$ for the Turán density of $\mathcal{C}_r^{(2k)}$ for any r, and a construction establishing a matching lower bound when r is of the form 2^p+1. We also show that when $r = 2^p + 1$, any $\mathcal{C}_r^{(4)}$-free hypergraph of density $\frac{r-2}{r-1} - o(1)$ looks approximately like Sidorenko's construction. On the other hand, when r is not of this form, we show that corresponding constructions do not exist and improve the upper bound on the Turán density of $\mathcal{C}_r^{(4)}$ to $\frac{r-2}{r-1} - c(r)$, where $c(r)$ is a constant depending only on r.

  To prove these results we use the tools from extremal graph theory, linear algebra, the Kruskal-Katona theorem and properties of Krawtchouck polynomials.

All these results were obtained jointly with P. Keevash and the first one was also obtained jointly with N. Alon and J. Balogh.

Canonical Colourings with Many Colours

Anusch Taraz

(joint work with B. Bollobás, Y. Kohayakawa, V. Rödl, M. Schacht)

Canonical colouring theorems state that, roughly spoken, every colouring of a sufficiently large object exhibits a local pattern of a given size that is coloured in a very regular way. From this point of view, partition theorems such as Ramsey's or van der Waerden's theorem deal with the special case of colourings with a bounded numbers of colours and assert that here, the local pattern can be guaranteed to be monochromatic. The topic of this talk, on the contrary, is to determine conditions that ensure local spots which are *rich in colours*. Our objects of interest will be both cliques in hypergraphs as well as arithmetic progressions on the integers.

Let us begin with arithmetic progressions. The classical theorem of van der Waerden states that every colouring of the first n natural numbers with at most t colours must contain a *monochromatic* k-term arithmetic progression, provided that n is sufficiently large compared to t and k. If no restriction on the number of colours is given, then the canonical colouring theorem by Erdős and Graham [2] states that we must find a *monochromatic* k-AP or an *injective* k-AP; i.e. one which uses pairwise distinct colours.

What condition could guarantee the latter of the two outcomes? It is not enough to merely ask for the colouring to use many colours globally, as can be seen by the following simple example. If $\ell = 3^i \cdot r$, where r isn't divisible by 3, then colour the number ℓ with colour i. Obviously this colouring uses an unbounded number of colours, but it is easy to see that not even an injective 3-AP will appear. Thus we need a stronger requirement on the colourfulness.

Theorem 1 *For every $k \in \mathbb{N}$ and for every $\varepsilon > 0$ there exist integers t and n_0 such that for every $n \geq n_0$ every colouring $\gamma : [n] \to \mathbb{N}$ with the property that*

$$\forall T \subseteq [n] \text{ with } |T| \geq (1-\varepsilon)n : \qquad |\gamma(T)| > t$$

must contain an injective k-term arithmetic progression.

The proof of this theorem is in fact quite short, as it can be based on a quantitative version [3] of Szemerédi's famous density theorem for arithmetic progressions. For graphs and hypergraphs such a density result does not hold and therefore the situation becomes more difficult. Here we are considering colourings of $E(K_n^r)$, the edges of the complete r-uniform hypergraph on n vertices. For the sake of a simpler exposition, we only mention the case $r = 3$ here. Given a family of disjoint vertex sets $V_1, \ldots, V_s$, we say that two edges $e, e' \subset V_1 \cup \cdots \cup V_s$ are of the same type if $|e \cap V_j| = |e' \cap V_j|$ for all $j = 1, \ldots, s$. Generalizing the result in [1], the following theorem asserts, roughly spoken, the existence of colourful canonical colourings which may be 1-partite, 2-partite or 3-partite.

Theorem 2 *For every $k \in \mathbb{N}$ and for every $\varepsilon > 0$ there exist integers t and n_0 such that for every $n \geq n_0$ every colouring $\gamma : E(K_n^{(3)}) \to \mathbb{N}$ with the property that*

$$\forall T \subseteq E(K_n^{(3)}) \text{ with } |T| \geq (1-\varepsilon)\binom{n}{3} : \qquad |\gamma(T)| > t$$

must contain one of the following colourful canonical colourings:

- *there exists a set* V_1 *and an index* $i \in \{1,2,3\}$ *such that* $|V_1| = k$ *and so that two edges contained in* V_1 *receive the same colour only if their* i*-th vertices in* V_1 *are identical, or*
- *there exist sets* V_1, V_2 *and indices* $i \in \{1,2,3\}$ *and* $j \in \{1,2\}$ *such that* $|V_1| = |V_2| = k$ *and so that two edges of the same type receive the same colour only if their* i*-th vertices in* V_j *are identical, or*
- *there exist sets* V_1, V_2, V_3 *and indices* $i \in \{1,2,3\}$ *and* $j \in \{1,2,3\}$ *such that* $|V_1| = |V_2| = |V_3| = k$ *and so that two edges of the same type receive the same colour only if their* i*-th vertices in* V_j *are identical.*

As an application of this theorem we consider (ℓ, H)-local colourings. For fixed integer ℓ and hypergraph H, a colouring of $E(K_n^r)$ is said to be (ℓ, H)-local, if every copy of H in K_n^r is coloured with at most ℓ different colours. Obviously, the larger we choose ℓ, the more colours can appear in an (ℓ, H)-local colouring. We address two questions:

- Given H, what is the largest value of ℓ such that the maximum number of colours used by an (ℓ, H)-local colouring is still bounded?
- Given H, what is the largest value of ℓ such that the maximum number of colours used by an (ℓ, H)-local colouring is still *essentially bounded*?

Here the term *essentially bounded* means the following: for every $\varepsilon > 0$, the colouring is such that after the removal of a suitably chosen ε-fraction of the edges, the remaining edges only use a bounded number of colours.

References

[1] B. Bollobás, Y. Kohayakawa, and R. Schelp, *Essentially infinite colourings of graphs,* Journal of the London Mathematical Society, **61** (2000), no. 3, 658–670.

[2] P. Erdős and R. Rado, *Combinatorial theorems on classifications of subsets of a given set,* Proceedings of the London Mathematical Society, **2** (1952), no. 3, 417–439.

[3] P. Frankl, R.L. Graham, and V. Rödl, *Quantitative versions of combinatorial partition theorems,* Jahresber. Deutsch. Math.-Verein. **92** (1990), no. 3, 130–144.

Chromatic Numbers of Triangle-free Graphs and their Complements

Carsten Thomassen

It is easy to see that triangle-free graphs may have large minimum degree. It is also well-known that they may have arbitrarily large chromatic numbers. Can these two phenomena happen simultaneously? Erdős and Simonovits asked in 1973 for which positive real numbers c, there exists a function $f(c)$ such that the following holds: If G is a triangle-free graph with n vertices and minimum degree at least cn, then the chromatic number is at most $f(c)$. (In other words, the chromatic number is independent of the number of vertices of the graph). They proved that $f(c)$ does not exist for $c < 1/3$. I proved a few years ago that $f(c)$ exists for each $c > 1/3$. So only the case $c = 1/3$ remains open. S. Brandt has conjectured that $f(1/3) = 4$.

Hajos' conjecture says that every graph of chromatic number k contains a subdivision of the complete graph on k vertices. The conjecture was disproved by Catlin in 1979 for all k greater than 6. Kühn and Osthus have verified Hajos' conjecture for graphs of girth greater than 100. The conjecture is open for triangle-free graphs. I showed recently that, if a regular triangle-free graph has bipartite edge-index greater than the number of vertices of the graph, then the complement is a counterexample to Hajos' conjecture. Thus, the complements of triangle-free graphs provide a large class of interesting counterexamples, and it is conceivable that some of these might be counterexamples to Hadwiger's conjecture as well. Searching for possible counterexamples, I tried to investigate the bipartite edge-index of triangle-free graphs on a fixed surface, in particular the projective plane. I found no natural graphs with a sufficiently large bipartite edge-index. Instead I found some with a small bipartite edge-index solving two open problems stated in Bollobas' classical monograph "Extremal Graph Theory" from 1978. One of the problems, due to Erdős, involves the smallest possible bipartite edge-index $g(n)$ of a 4-color-critical graph on n vertices. Erdős asked if $g(n)$ tends to infinity as n tends to infinity. I showed that $g(n)$ equals 3 or 4 for infinitely many n.

Dynamic Configuration of Optical Telecommunication Networks

Andreas Tuchscherer

We investigate methods for online call admission and routing and wavelength assignment in optical telecommunication networks. On demand connections are established by lightpaths which are optical channels that operate on one wavelength and can pass several network links without any opto-electronic conversion.

Definition 1 *An* optical network *is a triple* (G, Λ, W), *where*

- $G = (V, E)$ *is a simple and undirected graph,*
- $\Lambda = \{\lambda_1, \ldots, \lambda_k\}$ *is a set of wavelengths, and*
- $W : E \to 2^\Lambda$ *is a map from* E *to the power set of* Λ, *where* $W(e)$ *is the set of wavelengths generally available on edge* e.

A lightpath *in the optical network* (G, Λ, W) *is a pair* (p, λ) *which consists of a path* p *in* G *together with a wavelength* $\lambda \in \Lambda$ *such that* $\lambda \in W(e)$ *for each edge* $e \in E(p)$.

The *Wavelength Division Multiplexing* technique allows for using different wavelengths on one edge simultaneously. However, each wavelength on an edge cannot be used by more than one lightpath at the same time.

Definition 2 (Wavelength conflict constraint) *For each pair of simultaneously routed lightpaths* (p_1, λ_1) *and* (p_2, λ_2) *in an optical network* (G, Λ, W), *we have:*

$$E(p_1) \cap E(p_2) = \emptyset \text{ or } \lambda_1 \neq \lambda_2.$$

A lightpath (p, λ) *is called* free *if it can be realized without violating the wavelength conflict constraint.*

The considered problem can be formulated as follows.

Definition 3 (Dynamic Singleclass Call Admission Problem) *An instance of the* Dynamic Singleclass Call Admission Problem *(*DSCA*) is given by an optical network* (G, Λ, W), *a time horizon* T, *and a sequence of connection requests* $(\sigma_1, \sigma_2, \ldots)$ *with* $\sigma_j = (u_j, v_j, b_j, t_j, d_j, p_j)$, *where*

$$\begin{aligned} u_j, v_j \in V & \quad \textit{are the end nodes,} \\ b_j \in \mathbb{N} & \quad \textit{is the number of required lightpaths,} \\ t_j \in [0, T] & \quad \textit{is the start time,} \\ d_j \in \mathbb{R}_+ & \quad \textit{is the duration,} \\ p_j \in \mathbb{R}_+ & \quad \textit{is the profit.} \end{aligned}$$

The task is to maximize the total profit gained such that valid answers are given to all connection requests. The answer for each σ_j *must be given without knowledge of calls with later start times and specifies whether the request is accepted or rejected. If* σ_j *is accepted, it contributes* p_j *to the total profit but requires that* b_j *lightpaths connecting* u_j *and* v_j *are realized in* (G, Λ, W) *from* t_j *until* $t_j + d_j$. *In doing so, the wavelength conflict constraint must be satisfied all the time.*

Concerning the evaluation of online algorithms for the problem DSCA by competitive analysis, the following negative result can easily be shown.

Theorem 4 ([Tuc03]) *For the problem* DSCA *with $d_j = \infty$ and $p_j = b_j$ for each request σ_j, the competitive ratio of each deterministic competitive algorithm is km, where k denotes the number of wavelengths and m denotes the number of edges in the optical network.*

In the following, we report on the practical approach. The algorithms below are evaluated by simulation. The *greedy algorithms* have originally been proposed in [MA98]. We distinguish between two variants: partial wavelength search (PWS) and total wavelength search (TWS).

PWS: Let $\lambda_{i_1}, \dots, \lambda_{i_k}$ be some order on the set of wavelengths. If there is a free $[u, v]$-lightpath, route a shortest one in wavelength λ, where λ is the first wavelength in the order providing any free $[u, v]$-lightpath.

TWS: Let $\lambda_{i_1}, \dots, \lambda_{i_k}$ be some order on the set of wavelengths. If there is a free $[u, v]$-lightpath, route a shortest one in wavelength λ, where λ is the first wavelength in the order providing a *globally shortest* free $[u, v]$-lightpath.

Sorting the wavelengths in order of decreasing current availability (number of edges where the wavelength can currently be used) turned out to yield the best versions in partial and total wavelength search (see [Tuc03]). We denote the corresponding algorithms by PACK(P) and PACK(T).

The second class of algorithms (*network fitness algorithms*) have been developed at Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB) in a joint project with T-Systems Nova GmbH.

FIT: Let fit : $\mathcal{S} \to \mathbb{R}_+$ be some network fitness function, where $\mathcal{S}$ denotes the set of all possible network states of (G, Λ, W) (a network state corresponds to a configuration of routed lightpaths). If there is a free $[u, v]$-lightpath, route such a lightpath (p, λ) that the resulting state $S + (p, \lambda)$ yields a maximum fitness value.

We consider two network fitness algorithms called *available-lightpaths-reduction* (ALR) and *single-flow-reduction* (SFR). While ALR defines the fitness as the total number of currently free lightpaths, the algorithm SFR defines the fitness as the sum over all pairs of nodes s and t and each wavelength λ of the maximum number of free edge-disjoint $[s, t]$-lightpaths in wavelength λ.

We have investigated by simulation the *blocking probability* (ratio of rejected requests and appeared requests) depending on the traffic load (multiplex factor).

Figure 2 depicts the results for the four presented algorithms in a setting with randomly generated calls. It turns out that the total wavelength search version PACK(T) is superior to the partial wavelength search version PACK(P) and produces solutions with about the same quality as ALR. The network fitness algorithm SFR performs best.

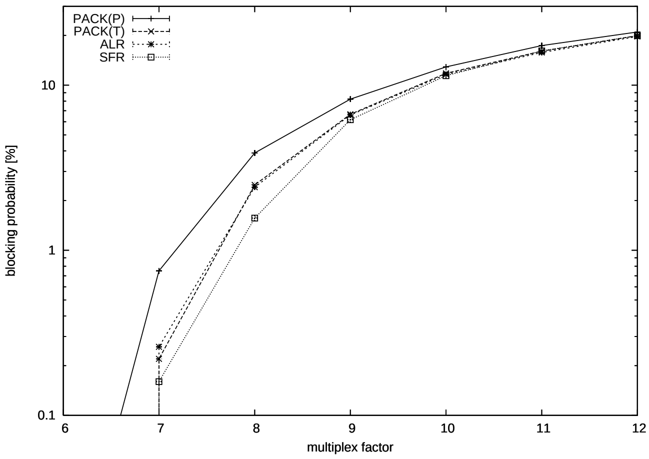


Figure 2: Results of selected algorithms in a 14-nodes network.

References

[MA98] Ahmed Mokhtar and Murat Azizoglu, *Adaptive wavelength routing in all-optical networks*, IEEE/ACM Transactions on Networking **6** (1998), no. 2, 197–206.

[Tuc03] Andreas Tuchscherer, *Dynamical configuration of transparent optical telecommunication networks*, diploma thesis, Technische Universität Berlin, March 2003.

On the Turán Number for the Hexagon

Jacques Verstraëte
(joint work with Zoltan Füredi and Assaf Naor)

One of the fundamental problems in extremal combinatorics is the determination of the maximum number of edges in a graph which contains no $2k$-cycles. The densest constructions of $2k$-cycle-free graphs for certain small values of k arise from the existence of rank two geometries called *generalized k-gons*, first introduced by Tits [5]. These may be defined as rank two geometries whose bipartite incidence graphs are r-regular graphs of diameter k and girth $2k$, where $r > 2$ and $k > 2$, and are known to exist only when k is three, four or six. This fact is an important consequence of a fundamental theorem of Feit and Higman [3]. It is therefore of interest to examine the extremal problem for quadrilaterals, hexagons, and cycles of length ten. In these cases, Lazebnik, Ustimenko and Woldar [4] used the existence of polarities of generalized polygons to construct dense $2k$-cycle-free graphs.

Erdős and Simonovits [2] conjectured the asymptotic optimality of these graphs, by conjecturing that the extremal number for the $2k$-cycle is asymptotic to $\frac{1}{2}n^{1+1/k}$ as n tends to infinity. This was known to hold for quadrilaterals almost fifty years ago, but was recently disproved in [4] for cycles of length ten. The only remaining case allowed by the Feit-Higman theorem is the case of hexagons. In this paper, we refute the Erdős-Simonovits conjecture for hexagons:

Theorem 1 *For infinitely many positive integers n, there are n-vertex hexagon-free graphs of size at least*

$$\frac{3(\sqrt{5}-2)}{(\sqrt{5}-1)^{4/3}}n^{4/3} + O(n) \;\approx\; 0.534n^{4/3}.$$

On the other hand, every n-vertex hexagon-free graph has size at most $\lambda n^{4/3} + O(n)$, where $\lambda \approx 0.627$ is the real root of $16\lambda^3 - 4\lambda^2 + \lambda - 3 = 0$.

The proof of Theorem 1 requires a statement about hexagon-free bipartite graphs, which is interesting in its own right (see de Caen and Szekely [1]).

Theorem 2 *Let m, n be positive integers. Then an m by n bipartite hexagon-free graph has size at most $2^{1/3}(mn)^{2/3} + O(n)$. When $m = 2n$ or $n = 2m$, there are m by n bipartite graphs with $2^{1/3}(mn)^{2/3} + O(n)$ edges.*

References

[1] de Caen, D., Székely, L.A. *The maximum size of* 4- *and* 6-*cycle free bipartite graphs on* m, n *vertices.* Sets, graphs and numbers (Budapest, 1991), 135-142, Colloq. Math. Soc. János Bolyai, 60, Noth-Holland, Amsterdam, 1992.

[2] Erdős, P., Simonovits, M. *Compactness results in extremal graph theory.* Combinatorica 2 (1982), no. 3, 275–288.

[3] Feit, W., Higman, G. The nonexistence of certain generalized polygons. J. Algebra 1 1964 114–131.

[4] Lazebnik, F., Ustimenko, V. A., Woldar, A. J., *Polarities and* $2k$-*cycle-free graphs.* Discrete Math. 197/198 (1999), 503–513.

[5] Tits, J., *Théorème de Bruhat et sous-groupes paraboliques.* C. R. Acad. Sci. Paris 254 (1962), 2910–2912.

Sharp Bounds on Long Arithmetic Progressions in Sumsets

V. H. Vu

(joint work with E. Szemerédi)

One of the main tasks of additive number theory is to examine structural properties of sumsets. For a set A of integers, the sumset $lA = A + \cdots + A$ consists of those numbers which can be represented as a sum of l elements of A. A closely related notion is that of l^*A, which is the collection of numbers which can be represented as a sum of l different elements of A. Among the most well-known results in all mathematics are Vinogradov's theorem which says that $3\mathbb{P}$ ($\mathbb{P}$ is the set of primes) contains all sufficiently large odd number and Waring's conjecture (proved by Hilbert, Hardy and Littlewood, Hua, and many others) which asserts that for any given r, there is a number l such that $l^*\mathbb{N}^r$ ($\mathbb{N}^r$ denotes the set of r^{th} powers) contains all sufficiently large positive integers (see [16] for an excellent exposition concerning these results).

In recent years, a considerable amount of attention has been paid to the study of finite sumsets. For a finite set A, the natural analogue of Vinogadov-Waring results is to show that under proper conditions, a finite set sumset lA (l^*A) contains a long arithmetic progression.

Let us assume that A is a subset of the interval $[n] = \{1, \dots, n\}$, where n is a large positive integer. The concrete problem we would like to talk about is to estimate the length of the longest arithmetic progression in lA (l^*A) as a function of l, n and $|A|$ (we are, of course, talking about the worst set A). This problem was stated explicitly for the sumset lA in a survey of Freiman, but we notice that many results had been proved earlier [1, 11, 12, 5]. We adapt a notation from Freiman's paper and denote by $f(|A|, l, n)$ the minimum length of the longest arithmetic progression in lA, where the minimum is taken over all sets $A \subset [n]$ with $|A|$ elements ($f^*(|A|, l, n)$ is defined similarly).

In this paper, we solve the problem completely for a wide range of l and $|A|$. In fact, our method carries us far beyond our original aim of estimating $f(|A|, l, n)$ and $f^*(|A|, l, n)$. We are able to show that lA and l^*A not only contain large arithmetic progressions, but also large proper generalized arithmetic progressions. Let us state the result for lA.

Theorem 1 *For any fixed positive integer d there are positive constants C and c depending on d such that the following holds. For any positive integers n and l and any set $A \subset [n]$ satisfying $l^d|A| \geq Cn$, lA contains an arithmetic progression of length $cl|A|^{1/d}$.*

Corollary 2 *For any fixed positive integer d there are positive constants C_1, C_2, c_1 and c_2 depending on d and ϵ such that whenever $\frac{C_1 n}{l^d} \leq |A| \leq \frac{C_2 n}{l^{d-1}}$*

$$c_1 l|A|^{1/d} \leq f(|A|, l, n) \leq c_2 l|A|^{1/d}.$$

Theorem 3 *For any fixed positive integer d there are positive constants C and c depending on d such that the following holds. For any positive integers n and l and any set $A \subset [n]$ satisfying $l^d|A| \geq Cn$, lA contains a proper GAP of rank d' and volume at least $cl^{d'}|A|$, for some $d' \leq d$.*

The same results hold for l^*A. However the proofs are much more difficult because of the assumption that the elements in a sum must be different. We can also prove similar results for finite fields.

Our results have some interesting applications. In particular, we settle two forty year old conjectures of Erdős [3] and Folkman [7] (respectively) concerning infinite arithmetic progressions. Let us end this abstract with the statements of these conjectures/theorems. For an infinite sequence of integers A, S_A denotes the collection of partial sums of A.

Theorem 4 *Let $A = a_1 < a_2 < \ldots$ be a sequence of positive integers with density at least $Cn^{1/2}$, where C is a sufficiently large constant. Then S_A contains an infinite arithmetic progression.*

This theorem was conjectured by Folkman in 1966 [7] and was a refined form of an earlier conjecture by Erdős made in 1962 [3] (see also [4] and [10] for more recent discussions).

Theorem 5 *Let $A = a_1 < a_2 < \dots$ be a sequence of positive integers with density at least Cn, where C is a sufficiently large constant. Then S_A contains an infinite arithmetic progression.*

By the density of A, we mean the number of elements of A between 1 and n. In the second theorem, this number may be large than A as we allow repetitions. It is known since the sixties (see [2]) that both statements are sharp, up to the constant C.

Most of the results discussed here appear in [14] and [15]. A related paper is [13], in which an application of different kind is discussed.

References

[1] J. Bourgain, On arithmetic progressions in sums of sets of integers. A tribute to Paul Erds, 105–109, Cambridge Univ. Press, Cambridge, 1990.

[2] J.W.S Cassels, On the representation of integers as the sums of distinct summands taken from a fixed set, *Acta Sci. Math. Szeged* 21 1960 111–124.

[3] P. Erdős, On the representation of large interges as sums of distinct summands taken from a fixed set, *Acta. Arith.* 7 (1962), 345-354.

[4] P. Erdős and R. Graham, **Old and new problems and results in combinatorial number theory**. Monographies de L'Enseignement Mathatique [Monographs of L'Enseignement Mathatique], 28. Universitde Gene, L'Enseignement Mathatique, Geneva, 1980. 128 pp.

[5] G. Freiman, H. Halberstam and I. Ruzsa, Integer sum sets containing long arithmetic progressions, *J. London Math. Soc.* (2) 46 (1992), no. 2, 193–201.

[6] G. Freiman, New analytical results in subset-sum problem. Combinatorics and algorithms (Jerusalem, 1988). *Discrete Math.* 114 (1993), no. 1-3, 205–217.

[7] J. Folkman, On the representation of integers as sums of distinct terms from a fixed sequence, *Canad. J. Math.* 18 1966 643–655.

[8] R. Graham, Complete sequences of polynomial values, *Duke Math. J.* 31 (1964), 275–286.

[9] B. Green, Arithmetic progressions in sumsets, *Geom. Funct. Anal.* 12 (2002), no. 3, 584–597

[10] C. Pomerance and A. Sárközy, Combinatorial number theory, **Handbook of combinatorics**, Vol. 1, 2, 967–1018, Elsevier, Amsterdam, 1995.

[11] I. Ruzsa, Generalized arithmetical progressions and sumsets, *Acta Math. Hungar.* 65 (1994), no. 4, 379–388.

[12] A. Sárközi, Finite addition theorems I, *J. Number Theory,* 32, 1989, 114–130.

[13] E. Szemerédi and V.H. Vu, Long arithmetic progressions in sumbsets and the number of x-sum-free sets, *submitted*. An electronic copy is available at `http://www.math.ucsd.edu/~vanvu/`.

[14] E. Szemerédi and V.H. Vu, Finite and Infinite arithmetic progressions in sumsets, *submitted*. An electronic copy is available at `http://www.math.ucsd.edu/~vanvu/`.

[15] E. Szemerédi and V.H. Vu, Sharp bounds on arithmetic progressions in sumsets, *manuscript*. An electronic copy is available at `http://www.math.ucsd.edu/~vanvu/`.

[16] R. Vaughan, **The Hardy-Littlewood method**. Second edition. Cambridge Tracts in Mathematics, 125. Cambridge University Press, Cambridge, 1997.

On Musin's Proof for the Kissing Number in Dimensions 3 and 4

Günter M. Ziegler

The "kissing number problem" asks for the maximal number of white spheres that can touch a black sphere of the same size in n-dimensional space. The answers in dimensions one, two and three are classical, while the answers in dimensions eight and twenty-four were a big surprise in 1979, based on an extremely elegant method initiated by Philippe Delsarte in the early eighties.

However, despite the fact that in dimension four there is a really special configuration which is conjectured optimal—the shortest vectors in the D_4 lattice, which are also the vertices of a regular 24-cell—it was even *proved* [1] that the bounds given by Delsarte's method aren't good enough to solve the problem in dimension four: This may explain the astonishment even to experts when last fall Oleg Musin announced a solution (currently under review) of the problem, based on a clever modification of Delsarte's method [3, 4].

The purpose of my talk was to outline Musin's new ideas. This started with a short description of the classical approach, due to Delsarte, Goethals & Seidel [2]: If $f(t) = \sum_k c_k G_k^{(n)}(t)$ is a non-negative combination of Gegenbauer polynomials which satisfies $f(t) \leq 0$ in the range $t \in [-1, \frac{1}{2}]$, then $\kappa(n) \leq f(1)/c_0$ is an upper bound for the kissing number in dimension n. Musin's modification is to require the condition $f(t) \leq 0$ only in a range $t \in [t_0, \frac{1}{2}]$ for some fixed $t_0 < -\frac{1}{2}$, while $f(t)$ must be strictly monotonically decreasing in the range $t \in [-1, t_0]$. This leads to an upper bound on $\kappa(n)$ in terms of some non-convex non-linear optimization problems. Musin explains ideas that reduce the dimensions of these optimization problems considerably. Apparently the problems are rather well-behaved, and can be solved numerically.

Their solution not only yields $\kappa(4) = 24$, but it also gives us a systematic and conceptual new proof for the Newton–Gregory problem, $\kappa(3) = 12$, which was first resolved by Schütte and van der Waerden (1953).

References

[1] V. V. Arestov and A. G. Babenko, *Estimates for the maximal value of the angular code distance for* 24 *and* 25 *points on the unit sphere in* $\mathbb{R}^4$, Math. Notes, 68 (2000), pp. 419–435.

[2] P. Delsarte, J. M. Goethals, and J. J. Seidel, *Spherical codes and designs*, Geomeetriae Dedicata, 6 (1977), pp. 363–388.

[3] O. R. Musin, *The kissing number in four dimensions.* Preprint, September 2003, 22 pages; `math.MG/0309430`.

[4] ——, *The problem of the twenty-five spheres*, Russian Math. Surveys, 58 (2003), pp. 794–795.

Reporter: Anusch Taraz

Participants

Prof. Dr. Martin Aigner
aigner@math.fu-berlin.de
Institut für Mathematik II (WE2)
Freie Universität Berlin
Arnimallee 3
D–14195 Berlin

Prof. Dr. Noga Alon
noga@math.tau.ac.il
Department of Mathematics
Sackler Faculty of Exact Sciences
Tel Aviv University
Tel Aviv 69978 – ISRAEL

Prof. Dr. Anders Björner
bjorner@math.kth.se
Dept. of Mathematics
Royal Institute of Technology
S-100 44 Stockholm

Prof. Dr. Aart Blokhuis
aartb@win.tue.nl
Department of Mathematics
Technische Universiteit Eindhoven
Postbus 513
NL-5600 MB Eindhoven

Prof. Dr. Bela Bollobas
Bollobas@msci.memphis.edu
b.bollobas@dpmms.cam.ac.uk
Dept. of Mathematical Sciences
University of Memphis
Memphis, TN 38152 – USA

Prof. Dr. Graham R. Brightwell
graham@tutte.lse.ac.uk
g.r.brightwell@lse.ac.uk
Dept. of Mathematics
London School of Economics
Houghton Street
GB-London WC2A 2AE

Dr. Maria Chudnovsky
mchudnov@Math.Princeton.EDU
Department of Mathematics
Princeton University
Fine Hall
Washington Road
Princeton, NJ 08544-1000 – USA

Prof. Dr. Reinhard Diestel
diestel@math.uni-hamburg.de
Mathematisches Seminar
Universität Hamburg
Bundesstr. 55
D–20146 Hamburg

Prof. Dr. Andras Frank
frank@cs.elte.hu
Department of Operations Research
Eötvös Lorand University
ELTE TTK
Pazmany Peter setany 1/C
H-1117 Budapest

Prof. Dr. Ehud Friedgut
ehudf@math.huji.ac.il
Institute of Mathematics
The Hebrew University
Givat-Ram
91904 Jerusalem – Israel

Prof. Dr. Zoltan Furedi
z-furedi@math.uiuc.edu
furedi@renyi.hu
Department of Mathematics
University of Illinois at
Urbana-Champaign
1409 West Green Street
Urbana IL 61801 – USA

Dr. David Galvin
galvin@microsoft.com
Microsoft Research
8500 148th Ave. NE
Apt. F3020
Redmond WA 98052 – USA

Dr. Stefanie Gerke
gerke@in.tum.de
gerke@inf.ethz.ch
Institut für Theoretische
Informatik, ETH Zentrum
IFW E49.2
Haldeneggsteig 4
CH-8092 Zürich

Prof. Dr. Martin Grötschel
groetschel@zib.de
Konrad-Zuse-Zentrum für
Informationstechnik Berlin (ZIB)
Takustr. 7
D–14195 Berlin

Prof. Dr. Hein van der Holst
host@kam.mff.cuni.cz
hvdholst@math.fu-berlin.de
Laboratoire LEIBNIZ-IMAG
46, Avenue Felix Viallet
F-38031 Grenoble Cedex 1

Prof. Dr. Jeff Kahn
jkahn@math.rutgers.edu
Dept. of Mathematics
Rutgers University
Busch Campus, Hill Center
New Brunswick, NJ 08903 – USA

Dr. Volker Kaibel
kaibel@math.tu-berlin.de
Fakultät II
Institut für Mathematik
MA 6-2
Straße des 17. Juni 136
D–10623 Berlin

Prof. Dr. Gyula O.H. Katona
ohkatona@renyi.hu
Alfred Renyi Mathematical Institute
of the Hungarian Academy of Science
Realtanoda u. 13-15
P.O.Box 127
H-1053 Budapest

Dr. Janos Körner
korner@dsi-uniroma1.it
Dept. of Computer Sciences
Universita "La Sapienza"
Via Salaria 113
I-00198 Roma

Prof. Dr. Alexandr V. Kostochka
sasha@math.nsc.ru
kostochk@math.uiuc.edu
Department of Mathematics
University of Illinois at
Urbana-Champaign
1409 West Green Street
Urbana IL 61801 – USA

Dr. Daniela Kühn
dkuehn@math.fu-berlin.de
Institut für Mathematik II (WE2)
Freie Universität Berlin
Arnimallee 3
D–14195 Berlin

Dr. Monique Laurent
monique@cwi.nl
M.laurent@cwi.nl
Centrum voor Wiskunde en
Informatica
P.O. Box 94079
NL-1090 GB Amsterdam

Dr. Imre Leader
i.leader@dpmms.cam.ac.uk
Centre for Mathematical Sciences
University of Cambridge
Wilberforce Road
GB-Cambridge CB3 OWB

Prof. Dr. Nathan Linial
nati@cs.huji.ac.il
School of Computer Science and Engineering
The Hebrew University
Givat-Ram
91904 Jerusalem –Israel

Prof. Dr. Laszlo Lovasz
lovasz@cs.yale.edu
lovasz@microsoft.com
Microsoft Research
One Microsoft Way
Redmond, WA 98052-6399
USA

Prof. Dr. Jiri Matousek
matousek@kam.ms.mff.cuni.cz
Department of Applied Mathematics
Charles University
Malostranske nam. 25
118 00 Praha 1 – Czech Republic

Prof. Dr. Colin McDiarmid
cmcd@stats.ox.ac.uk
Department of Statistics
University of Oxford
1 South Parks Road
GB-Oxford OX1 3TG

Prof. Dr. Jaroslav Nesetril
nesetril@kam.ms.mff.cuni.cz
Department of Applied Mathematics
Charles University
Malostranske nam. 25
118 00 Praha 1 – Czech Republic

Dr. Deryk Osthus
osthus@informatik.hu-berlin.de
Institut für Informatik
Humboldt-Universität zu Berlin
Unter den Linden 6
D–10099 Berlin

Dr. Oleg Pikhurko
pikhurko@andrew.cmu.edu
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213-3890 – USA

Prof. Dr. Hans Jürgen Prömel
proemel@informatik.hu-berlin.de
Institut für Informatik
Humboldt-Universität zu Berlin
D–10099 Berlin

Dr. Oliver M. Riordan
O.M.Riordan@dpmms.cam.ac.uk
Trinity College
GB-Cambridge CB2 1TQ

Prof. Dr. Vojtech Rödl
rodl@mathcs.emory.edu
Dept. of Mathematics and Computer Science
Emory University
Atlanta, GA 30322 – USA

Prof. Dr. Alexander Schrijver
lex@cwi.nl
lex.schrijver@cwi.nl
CWI
Postbus 94079
NL-1090 GB Amsterdam

Dr. Alex Scott
scott@math.ucl.ac.uk
Department of Mathematics
University College London
Gower Street
GB-London, WC1E 6BT

Dr. Paul Seymour
pds@math.princeton.edu
pseymour@Princeton.EDU
Department of Mathematics
Princeton University – Fine Hall
Washington Road
Princeton, NJ 08544 – USA

Prof. Dr. Miklos Simonovits
miki@renyi.hu
miki@math-inst.hu
Alfred Renyi Mathematical Institute
of the Hungarian Academy of Science
Realtanoda u. 13-15
P.O.Box 127
H-1053 Budapest

Prof. Dr. Vera T. Sos
sos@renyi.hu
Alfred Renyi Mathematical Institute
of the Hungarian Academy of Science
Realtanoda u. 13-15
P.O.Box 127
H-1053 Budapest

Prof. Dr. Angelika Steger
steger@inf.ethz.ch
Institut für Theoretische
Informatik, ETH Zentrum
IFW E49.2
Haldeneggsteig 4
CH-8092 Zürich

Prof. Dr. Benjamin Sudakov
bsudakov@math.princeton.edu
Department of Mathematics
Princeton University
Fine Hall
Washington Road
Princeton, NJ 08544 – USA

Prof. Dr. Endre Szemeredi
szemered@cs.rutgers.edu
szemered@aramis.rutgers.edu
szemered@nyuszik.rutgers
Department of Computer Science
Rutgers University
Hill Center, Busch Campus
New Brunswick, NJ 08903 – USA

Dr. Anuschirawan Taraz
taraz@informatik.hu-berlin.de
Institut für Informatik
HU-Berlin
Unter den Linden 6
D–10099 Berlin

Prof. Dr. Robin Thomas
thomas@math.gatech.edu
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332-0160 – USA

Dr. Andrew Thomason
A.G.Thomason@dpmms.cam.ac.uk
Dept. of Pure Mathematics and
Mathematical Statistics
University of Cambridge
Wilberforce Road
GB-Cambridge CB3 OWB

Prof. Dr. Carsten Thomassen
c.thomassen@mat.dtu.dk
Matematisk Institut
Danmarks Tekniske Universitet
Bygning 303
DK-2800 Lyngby

Dr. Jacques Verstraete
jverstraete@math.uwaterloo.ca
jverstra@pythagoras.math.uwaterloo.ca
Faculty of Mathematics
University of Waterloo
200 University Avenue west
Waterloo ONT N2L 3G1 – Canada

Prof. Dr. Van H. Vu
vanvu@ucsd.edu
Dept. of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0112 – USA

Prof. Dr. Dominic J. A. Welsh
dwelsh@maths.ox.ac.uk
Mathematical Institute
Oxford University
24 - 29, St. Giles
GB-Oxford OX1 3LB

Prof. Dr. Günter M. Ziegler
ziegler@math.tu-berlin.de
Institut für Mathematik
MA 6-2
Technische Universität Berlin
Straße des 17. Juni 136
D–10623 Berlin

Mathematisches Forschungsinstitut Oberwolfach

Report No. 2/2004

Statistics in Finance

Organised by
Richard A. Davis (Ft. Collins)
Claudia Klüppelberg (München)

January 11th – January 17th, 2004

Introduction by the Organisers

The Statistics in Finance Workshop, organized by Richard A. Davis (Ft. Collins) and Claudia Klüppelberg (Technische Universität München), was held January 11-17. This meeting was well attended with over 40 participants with broad geographic representation from Europe, England, Australia, the Far East, and the US. This workshop was a nice blend of researchers with various backgrounds including statistics, probability, and econometrics. Approximately 33 talks, of varying lengths, were delivered during the five days. The talks were given by both leading experts in the field as well as by up and coming stars.

There were several major themes in the various sessions. These included, continuous time models, Levy processes, stochastic volatility models, GARCH models, extreme value theory with applications to financial risk, theory of copulas, and option pricing. This meeting generated a great deal of discussion and often smaller groups of people met in the evenings for expanded and detailed lectures. A number of important research contacts were made which we fully expect to stimulate many new collaborative research projects.

In addition to the excellent scientific program, there were two scheduled social activities. The inclement weather cleared up just in time for the traditional Wednesday afternoon hike to Oberwolfach for coffee and Black Forest Cake. The second activity, which most considered the highlight of the week, was a piano recital performed by Peter Brockwell and Gernot Müller.

For many of the participants, this was their first trip to Oberwolfach, and they came away very impressed from the experience. There was a strong consensus that the “Statistics in Finance Workshop” should become a regular Oberwolfach event.

Workshop on Statistics in Finance

Table of Contents

Abstracts

Continuous Time Stochastic Volatility Modelling and Bipower Variation

Ole. E. Barndorff-Nielsen and Neil Shephard

The theory of semi-martingales and stochastic integration constitutes a powerful and natural background for continuous time modelling of stochastic volatility, as observed in financial time series. However, for the models to make financial sense it is necessary to restrict somewhat from the completely general concept of semi-martingales ($\mathcal{SM}$). Recall that $Y \in \mathcal{SM}$ means that Y is of the form $Y = A + M$ where $A \in \mathcal{FV}$ and $M \in \mathcal{M}_{loc}$. We wish to think of Y as the log price process of a financial asset, with the process A expressing potential rewards and M the risk. For this the decomposition of Y into the sum of A and M should be unique. This is achieved by requiring A to be predictable (in which case Y is said to be a special semi-martingale).

If we further assume that M is continuous, $M \in \mathcal{SM}^c$ (below we comment on alternative possibilities), then this has important consequences. First, if the model Y is to be arbitrage free then A has also to be continuous, i.e. $A \in \mathcal{FV}^c$. Further, by the Dambis-Dubins-Schwarz Theorem, M is then representable as a time changed Brownian motion (BM). So $M = B_{[M]}$ and $B = M_T$ where $[M]$ is the quadratic variation of M and the time-change T is the inverse function of $[M]$. (For this it is necessary that $[M]_t \to \infty$ for $t \to \infty$.) Thirdly, supposing that $[M]$ is absolutely continuous, of the form $[M] = \int_0^t \tau_u \mathrm{d}u$, then, again by the no arbitrage requirement, A must also be absolutely continuous, $A = \int_0^t a_u \mathrm{d}u$. The process $\sigma = \sqrt{\tau}$ expresses the volatility, and it is to be noted that τ, which is termed the variance process, may have jumps. Finally, the absolute continuity together with the time-change representation implies that M can be written as $M_t = \int_0^t \sigma_u \mathrm{d}W_u$ for a BM W.

Thus the choice for Y has been narrowed to the continuous stochastic volatility semi-martingale framework ($\mathcal{SVSM}^c$) where

$$Y_t = \int_0^t a_u \mathrm{d}u + \int_0^t \sigma_u \mathrm{d}W_u. \tag{1}$$

This type of process is sometimes called a Brownian semi-martingale; but, having the financial context in mind, we refer to it as a (continuous time) stochastic volatility process.

For more specific modelling, aimed at representing the important, and widely established, stylised features of financial observational series, choices have to be made of the two ingredients a and σ of (1). A simple example for a, of some definite interest, takes $a_t = \mu + \beta\tau_t$. More generally, one may consider $a_t = g(t, \tau_t)$ for some smooth function g. As to σ, a number of points have to be considered: (i) Should σ be a pure diffusion process (or perhaps a superposition of such processes); or should it be a pure diffusion plus a finite activity (FA) process (finite activity meaning that there are only finitely many jumps in any finite time interval), or should it perhaps be an infinite activity (IA) process (for instance, an inverse Gaussian OU process or a superposition of such processes, or one of the CARMA processes introduced by Brockwell (2001), or one of the Lévy driven long-memory models considered by Anh, Heyde and Leonenko (2002). (ii) Should the model incorporate leverage, in the sense of dependence between σ and W. (iii) Should σ be Markovian. And more specifically, how should the law of τ be chosen so as to capture both the typical 'exbell' shape of log returns and the, generally observed, quasi long range dependence in the log price series. Note that if a is 0, or independent of σ and W, then under (1) the autocorrelations of the returns are necessarily 0, in accordance with the empirical facts.

To account for possible jumps in the price process one possibility is to add an independent FA process to Y. Another is to substitute the $\sigma \bullet W$ term by either $\sigma \bullet L$ or L_T where L denotes a Lévy process and T is a time-change. Note that, except for stable Lévy motions L, these two approaches are not equivalent; each has its advantages and drawbacks. A simpler approach is pure Lévy modelling which replaces Y by L. This already yields significant improvements over the classical Black-Scholes model, but misses, in particular, the key time-wise dependence feature of finance data. A further variant, that does model dependence, is the recently introduced continuous time GARCH model of Klüppelberg, Lindner and Maller (2004). A great amount of interesting work in this area has been carried out in the project led by Ernst Eberlein at Freiburg University. And in monograph form there are now two recent additions to the literature that set out important aspects of the Lévy based methods: Schoutens (2003) and Cont and Tankov (2003).

Our own joint research has fallen within the framework outlined above, and much of this will be described in considerable detail in our forthcoming book Barndorff-Nielsen and Shephard (2005).

The most recent part of this research concerns the new concept of *bipower variation*, that we have introduced and studied in Barndorff-Nielsen and Shephard (2003, 2004). This considers returns over time periods of lengths $\hbar$ and δ, where $n\delta = \hbar$ for some positive integer n and where for concreteness we may think of $\hbar$ as representing a day, with δ corresponding to 5, 10 or 30 minute consecutive intervals during that day. In the simplest 1,1 case, we define the realised bipower variation on the i-th day as the probability limit for $\delta \to 0$ of

$$\{Y_\delta\}_i^{[1,1]} = \sum_{j=2}^{n} |y_{j-1,i}|\,|y_{j,i}|\,;$$

here

$$y_{j,i} = Y_{(i-1)\hbar + j\delta} - Y_{(i-1)\hbar + (j-1)\delta}, \quad j = 1, 2, ..., n.$$

The limit of $\{Y_\delta\}_i^{[1,1]}$ is denoted by $\{Y\}_i^{[1,1]}$. We show that when we add to an SV process a finite activity jump process then, up to proportionality, the probability limit of this object (subject to some weak assumptions) is the quadratic variation of the SV process over the day as $\delta \downarrow 0$. Thus the realised bipower variation process is reasonably robust to jumps.

An asymptotic distribution theory for realised bipower variation can be calculated. Further, the joint distribution of realised bipower variation and the quadratic variation version of this, can be calculated under the assumption that there are no jumps. This allows us to consistently test the hypothesis that the sample path of price processes have jumps.

References

[1] Anh, V.V., Heyde, C.C. and Leonenko, N.N. (2002): Dynamic models of long-memory processes driven by Lévy noise. *J. Appl. Prob.* **39**, 730-747.

[2] Barndorff-Nielsen, O.E. and Shephard, N. (2003): Econometrics of testing for jumps in financial economics using bipower variation. Unpublished paper, Nuffield College, Oxford.

[3] Barndorff-Nielsen, O.E. and Shephard, N. (2004): Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics* **2**, 1-48.

[4] Barndorff-Nielsen, O.E. and Shephard, N. (2005): *Continuous Time Approach to Financial Volatility*. Cambridge University Press (to appear).

[5] Brockwell, P.J. (2001): Lévy driven CARMA processes. *Ann. Inst. Statist. Math.* **53**, 113-124.

[6] Cont, R. and Tankov, P. (2003): *Financial Modelling with Jump Processes*. Chapman and Hall/CRC, London.

[7] Klüppelberg, C., Lindner, A. and Maller, R. (2004): A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour. *J. Appl. Prob.* **41** (3) (to appear).

[8] Schoutens, W. (2003): *Lévy Processes in Finance: Pricing Financial Derivatives*. Wiley, New York.

Fractionally Integrated Continuous Time ARMA Processes

Peter J. Brockwell
(joint work with Tina Marquardt)

Continuous-time models for time series which exhibit both heavy-tailed and long-memory behaviour are of considerable interest, especially for the modelling of financial time series, where such behaviour is frequently observed empirically. A recent paper of Anh, Heyde and Leonenko (2002) develops such models via the Green-function solution of fractional differential equations driven by Lévy processes. A very general class of Gaussian fractionally integrated continuous time models with extensive financial applications has also been introduced by Comte and Renault (1996).

In this paper we consider the class of Lévy-driven continuous-time ARMA (CARMA) processes and the fractionally integrated (FICARMA) processes obtained by fractional integration of the kernel of the CARMA process. For completeness we include a brief account of the derivation of this kernel and indicate its relevance to the stochastic volatility model of Barndorff-Nielsen and Shephard (2001). In the latter paper an Ornstein-Uhlenbeck process driven by a non-decreasing Lévy process was used to model volatility in a stochastic volatility model for log asset prices. The stationary Ornstein-Uhlenbeck process,

$$X(t) = \int_{-\infty}^{t} e^{-c(t-y)} dL(y), \ c > 0,$$

was chosen because it has a non-negative kernel ($g(t) = \exp(-ct)I_{[0,\infty)}(t)$) and consequently, if the driving Lévy process L is non-decreasing, the process X will be non-negative as is necessary if it is to represent volatility. However the use of the Ornstein-Uhlenbeck process restricts the class of volatility autocorrelation functions to functions of the form $\rho(h) = \exp(-ch)$ for some $c > 0$. Barndorff-Nielsen and Shephard suggested extending this class by using linear combinations of independent Ornstein-Uhlenbeck processes with positive coefficients, however the autocorrelation functions are still restricted to be monotone decreasing. If the Ornstein-Uhlenbeck process is replaced by a non-negative Lévy-driven CARMA process, a much larger class of autocorrelations can be modelled, and in particular the monotonicity constraint can be removed (see Brockwell (2003) for further details).

In this paper we derive explicit expressions for the kernel and auto covariance functions of a FICARMA process whose autoregressive polynomial has distinct zeroes. (Corresponding results for multiple zeroes can be obtained by letting distinct roots converge to a common limit.) We also consider the asymptotic behaviour of these functions for large lags. The results are continuous-time analogues of the results of Sowell (1992) for discrete-time fractionally integrated ARMA processes. A comprehensive treatment of the latter processes can be found in the book of Beran (1994). From a second-order point of view, the fractionally integrated CARMA process is a special case of the (Gaussian) fractionally integrated processes defined by Comte and Renault (1996), however the particular form of the kernel of the CARMA process leads to very simple expressions for the kernel and autocovariance functions for the corresponding fractionally integrated process.

If a CARMA(p, q) process is sampled at times $\{0, 1, 2, \ldots\}$, it is well-known that the sampled process is a discrete-time ARMA(p, r) process with $r < p$. It is therefore of interest to compare the behaviour of the fractionally integrated CARMA process sampled at integer times with that of the sampled CARMA process fractionally integrated (in the discrete-time sense). In this paper we make such a comparison for the fractionally integrated Ornstein-Uhlenbeck process.

References

[1] Anh, V.V., Heyde, C.C. and Leonenko, N.N. (2002): Dynamic models of long-memory processes driven by Lévy noise. *J. Appl. Prob.* **39**, 730-747.

[2] Barndorff-Nielsen, O.E. and Shephard, N. (2001): Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B* **63**, 1-42.

[3] Beran, R. (1994): *Statistics for long-memory processes*. Chapman and Hall, New York.

[4] Brockwell, P.J. (2003): Representations of continuous-time ARMA processes. *J. Appl. Prob.* (to appear).

[5] Comte, F. and Renault, E. (1996): Long memory continuous time models. *J. Econometrics* **73**, 101-149.

[6] Sowell, F. (1992): Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *J. Econometrics* **53**, 165-188.

Extremal Behaviour of Fractal Models
Boris Buchmann
(joint work with Claudia Klüppelberg)

A fractional Brownian motion (FBM) is a centred Gaussian process $(B^H{}_t)_{t\in\mathbb{R}}$ with covariance function

$$EB^H{}_tB^H{}_s = \frac{1}{2}\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\}.$$

The parameter $H \in (0,1)$ is the so-called Hurst coefficient. For $H = 1/2$ FBM is the Wiener process, otherwise, both long memory ($H > 1/2$) and short memory ($H < 1/2$) occurs in the increments and FBM is no longer a semi-martingale. FBM has been studied by Kolmogorov in the fourties as a model for turbulence and by Mandelbrot and van Ness (1968) to describe certain aspects in the Nile data. Recently, it has been proposed as tool for financial applications (e.g. Hu and Øksendal (1999), Brody, Syroka and Zervos (2003)). The fractional Ornstein-Uhlenbeck process ($O_t^{H,\gamma,\sigma}$) (FOUP), i.e., the stationary solution of the Langevin equation

$$O_t^{H,\gamma,\sigma} = O_0^{H,\gamma,\sigma} - \gamma\int_0^t O_s^{H,\gamma,\sigma}\,ds + \sigma B_t^H,$$

where $\gamma > 0$, has been studied by Cheridito, Kawaguchi, Maejima (2003). It is a Gaussian process which again exhibits long memory ($H > 1/2$) and short memory ($H < 1/2$). In contrast to the latter authors we have studied the shape of the covariance function near zero. Combining both results allows us to develop the extreme value theory for FOUP based on classical results on Gaussian processes (Pickands (1969), Berman (1971), Leadbetter, Lindgren, Rootzén (1983)). More precisely, we obtain the norming constant $b_T(H,\gamma,\sigma)$ such that

$$\frac{2}{\Gamma(2H{+}1)^{1/2}}\frac{\gamma^H}{\sigma}(\log T)^{1/2}\left\{\max_{0\le t\le T} O_t^{H,\gamma,\sigma} - b_T(H,\gamma,\sigma)\right\} \xrightarrow{d} \quad G,$$

where G is a Gumbel distributed random variable. The extreme value theory can be extended to processes $X_t^{H,\gamma,f} := f(O_t^{H,\gamma})$ where f is a state space transform (SST), i.e., a continuous strictly increasing function. Our concept is related to the work of Davis (1982) where the extreme value theory for diffusions is studied by transforms in time and space. The process $X^{H,\gamma,f}$ to be in the maximum domain of a Gumbel distribution is provided by the following condition on the derivative, namely,

$$\lim_{z\to\infty}\frac{f'(z+a(z))}{f'(z)} = 1$$

for all functions $x \mapsto a(x)$ such that $a(x) = O(x^{-1})$ for $x \to \infty$. Generally, if $f'(x) = r(x)\exp(\kappa x^p)$ for sufficiently large x and some regularly varying function r the condition is satisfied whenever $p < 2$. If f is a SST such that for some constants $C_0 > 0$ and $C_1 \in \mathbb{R}$

$$\log f(x) = C_0 x^2 + C_1 + o(1) \qquad \text{for } x \to \infty,$$

then $X^{H,f,\gamma}$ is found to be in the maximum domain of attraction of a Frechét distribution. The concept of SSTs can be related with a geometric approach to solve integral equations of the type

$$X_t - X_0 = \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dB_s^H.$$

As $(B^H{}_t)$ is not a semi-martingale for $H \neq 1/2$ Itô integration can not be used to define an integral w.r.t FBM. Different approaches have been discussed in the literature (Hu and Øksendal (1999), Mikosch and Norvaiša (2000), Duncan, Hu and Pazik-Duncan (2000), Mazet, Alós and Nualart (2001)). We follow the approach of Zähle (1998, 2001) which works for $H > 1/2$. FBM is then sufficiently Hölder continuous such that the integral w.r.t FBM is well-defined as Riemann-Stieltjes integral whenever the integrand is Hölder of some order strictly larger $1 - H$. By a law of iterated logarithm (Arcones (1995)) FBM takes values in the weighted function space $\tilde{V}_H$ containing all functions f Hölder of at least any order strictly smaller H such that

$$\sup_t \frac{|f(t)|}{1 + |t|^H \sqrt{(\log\log)^+(|t|)}} < \infty\,.$$

Replacing B^H by any possible function $g \in \tilde{V}_H$ we derive a necessary and sufficient condition on μ and σ for existence and uniqueness of a solution in terms of SSTs and the FOUP.

References

[1] Arcones, M.A. (1995): On the law of iterated logarithm for Gaussian processes. *J. Theoret. Probab.* **8** (4), 877-904.

[2] Alòs, E., Mazet, O. and Nualart, D (2000): Stochastic calculus with respect to Gaussian processes. *Ann. Prob.* **29**, 766-801.

[3] Berman, S.M. (1971): Maxima and high level excursions of stationary Gaussian processes. *Trans. Amer. Math. Soc.* **160**, 65-185.

[4] Brody, D., Syroka, J. and Zervos, M. (2003): Dynamical pricing of weather derivatives. *Quantitive Finance* (to appear).

[5] Buchmann, B. and Klüppelberg, C. (2004): Extreme value theory for stochastic processes driven by fractional Brownian motion. (In preparation.)

[6] Cheridito, P., Kawaguchi, H. and Maejima, M. (2003): Fractional Ornstein Uhlenbeck Processes. *Electron. J. Probab.* **8** (3), 1-14.

[7] Davis, R. A. (1982): Maximum and minimum of one-dimensional diffusions. *Stoch. Proc. Appl.* **13** (1), 1-9.

[8] Decreuesefond, L. and Üstünel, A.S. (1999): Stochastic analysis of the fractional Brownian motion. *Potential Anal.* **10** (2), 177–214.

[9] Duncan, T.E., Hu, Y. and Pasik-Duncan, B. (2000): Stochastic calculus for fractional Brownian motion. *SIAM J. Control Optim.* **38**, 582-612.

[10] Hu, Y. and Øksendal, B. (1999): Fractional white noise calculus and applications to finance. *SIAM J. Control Optim.* **38**, 582-612.

[11] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983): *Extremes and related properties of random sequences and processes.* Springer, New York.

[12] Mikosch, T. and Norvaiša, R. (2000): Stochastic integral equation without probability. *Bernoulli* **6** (3), 401-434.

[13] Pickands, J. III (1969): Upcrossing probabilities for stationary Gaussian processes. *Trans. Amer. Math. Soc.* **145**, 51-73.

[14] Pickands, J. III (1969): Asymptotic properties of the maximum in a stationary Gaussian processes. *Trans. Amer. Math. Soc.* **145**, 75-86.

[15] Zähle, M. (1998): Integration with respect to fractal functions and stochastic calculus I. *Probab. Theory Relat. Fields* **111**, 333-374.

[16] Zähle, M. (2001): Integration with respect to fractal functions and stochastic calculus II. *Math. Nachr.* **225**, 154-183.

Structural Models for Credit Risk Migration

Ngai Hang CHAN

1. Abstract

A structural model for credit migration is considered in this paper. The proposed model is firm specific and depends on two parameters: the default distance and credit history. The default distance is the standardized logarithmic asset-to-liability ratio modelled by a Brownian motion and the credit history is modelled by an occupation time variable. By examining the properties of this occupation time variable, the credit performance of a given firm can be analyzed. This model not only allows one to derive a closed-form credit transition probability, but also explains default probability overlaps of different ratings. It can be used to back out the subjective thinking of credit performance of rating agencies.

2. Structural Models

Credit risk management is an important tool in finance, especially in the high-yield bond market and the bank loan market. An essential concern of a financial corporation is changes in credit ratings of companies. Nationally recognized statistical rating organizations (NRSRO), like Standard & Poor's and Moody's Investor Services, classify corporate bond issuers into different credit ratings in order to reflect their credit worthiness. Credit risk managers pay serious attentions to the ratings and transition matrices published by NRSRO. Transition matrices in the form of arrays of migrating probabilities constitute the building block of risk management tools, see for example JP Morgan's Credit Metrics and McKinsey's Credit Portfolio View. The Markov model of Jarrow, Lando and Turnbull (1997) uses the transition matrix to generate the term structure of credit spreads.

Predictions of transition probabilities have been receiving considerable amount of attentions recently. Most of the research make use of historical transition matrices and firm ratings to estimate future transition probabilities, see for example Aderson et. al (1991), Altman and Kao (1992), Kavvathas (2000), and Lando and Skodeberg (2002). Two types of default models are structural approach and reduced-form models. Structural type models suggest that a firm defaults when its asset value drops below its liabilities. KMV corporation implements this structural approach and generates expected default probabilities (EDPs) of firms. Details of the KMV methodology can be found in Crosbe and Bohn (1993). Jarrow and Turnbull (1995) and Duffie and Singleton (1999) propose a second approach called reduced-form models. The time of default in this model is characterized by an exogenously defined intensity process.

This paper proposes a structural model of credit migration. There are at least two reasons to adopt the structural approach. First, it has a solid theoretical basis as it takes into account of the capital structure of a firm. Second, a structural model makes use of the distance-to-default of a firm. The distance-to-default values can be measured from the market, either through KMV or internal models. Finally, the recent acquisition of KMV and Moody's Investor Services provides the market with a possibility of using structural approach to measure credit transition probabilities.

The proposed model is firm specific and depends on distance-to-default and migrating signal duration. Using the idea of Gordy and Heitfield (2001), distance-to-default is mapped into different rating categories by partitioning the distribution of empirical data. The proposed model is able to capture the slow-to-respond features of rating agencies. Such a time-lagged response to new information can be interpreted as an extra rating criterion to reflect the market reputations of a rated company. If the distance-to-default is assumed to follow a symmetric distribution, then migrating probabilities generated from the model can still be skewed on one side. There are several desirable features of the model. Rating agencies can use it to explain changes in firm ratings in relation to current and historical credit performance. The proposed model also allows the overlap of EDPs across different letter grades, and offers a means to reconcile the empirical findings of Kealhofer et. al (1998). Analytical formula for calculating migrating probability can also be obtained.

3. Résumé

Dans cet article, on développe un modèle structurel pour la migration de crédit. On propose un modèle pour chaque entreprise dépendant de deux paramètres: le temps avant la faillite de l'entreprise et l'histoire du crédit. Le temps avant la faillite est défini comme le rapport logarithmique standardisé entre l'actif et la dette, et est modelisé par un mouvement Brownien. L'histoire du crédit est modelisée par une durée variable. En examinant les propriétés de ce temps variable, la performance du credit d'une entreprise donnée peut être analysée. Ce modèle permet non seulement d'obtenir l'expression de la probabilité de transition du crédit, mais il explique aussi les chevauchements des probabilités de faillite pour differents taux. Le modèle peut aussi être utilisé pour reévaluer l'idée subjective donnée par une agence chargée de l'estimation de la performance d'un crédit.

References

[1] Altman, E., Kao, D. (1992): Rating drift in high yield bonds. *Journal of Fixed Income* **2**, 15-20.

[2] Anderson, P., Hanson, L.S., Keilding, N. (1991): Non-and semi parametric estimation of transition probabilities from censored observation of non-homogeneous Markov processes. *Scand. J. Statist.* **22**, 153-167.
[3] Crosbe, P.J., Bohn, J.R. (1993): Modeling default risk. KMV Corporate.
[4] Duffie, D., Singleton, K. (1999): Modeling term structures of defaultable bonds. *Review of Financial Studies* **12**, 687-720.
[5] Gordy, M., Heitfield, E. (2001): Of Moody's and Merton: A structural model of bond rating transitions. Working Paper, Board of Governors of the Federal Reserve System.
[6] Jarrow, R.A., Lando, D., Turnbull, S. (1997): A Markov model for the term structure of credit risk spreads. *Review of Financial Studies* **10**, 481-523.
[7] Jarrow, R.A., Turnbull, S.M. (1995): Pricing derivatives on financial securities subject to credit risk. *J. Finance* **50**, 53-58.
[8] Kavvathas. D (2000): Estimating credit rating transition probabilities for corporate bonds. Working Paper, Department of Economics, University of Chicago.
[9] Kealhofer, S., Kwok, S., Weng, W. (1998): Uses and abuses of bond default rates, KMV Corporation.
[10] Lando, D., Skodeberg, T.M. (2002): Analyzing rating transitions and rating drift with continuous observations. *Journal of Banking and Finance* **26**, 423-444 .

Stochastic Volatility Models for Ordinal Valued Time Series

Claudia Czado
(joint work with Gernot Müller)

Our aim is to model the intraday development of stock prices, in particular the development of the price change process. The price changes have some specific features which we want to be covered by our model. One important feature is that price changes only occur in integer multiples of a certain amount, the so-called tick size. In modelling the price changes we therefore have to take into account that we observe a discrete time series. Also other important features of such time series are covered by the following model:

$$(1) \qquad y_t^{obs} = k \iff y_t \in [c_{k-1}, c_k) \qquad k \in \{1, \ldots, K\}$$

$$(2) \qquad y_t = \boldsymbol{x}_t'\boldsymbol{\beta} + \exp(h_t/2)\varepsilon_t \qquad t \in \{1, \ldots, T\}$$

$$(3) \qquad h_t = \mu + \boldsymbol{z}_t'\boldsymbol{\alpha} + \phi(h_{t-1} - \mu - \boldsymbol{z}_{t-1}'\boldsymbol{\alpha}) + \sigma\eta_t$$

A modified version of the underlying stochastic volatility model (2) and (3) for continuous responses was considered by Chib, Nardari and Shephard (2002). Observed are only the variables y_t^{obs}, which are discretized versions of the latent continuous variables y_t. $\boldsymbol{x}_t$ and $\boldsymbol{z}_t$ are vectors of covariates, ε_t and η_t are assumed to be i.i.d. $N(0, 1)$. We fix c_1 and μ for reasons of identifiability.

For the estimation of the parameters in this model we develop a MCMC algorithm, which is based on the algorithm presented in Chib, Nardari and Shephard (2002) for the underying continuous model. However, standard Gibbs MCMC steps for the additional discretization in Equation (1) lead to bad convergence behaviour of the resulting MCMC iterations. Figure 1 shows the cutpoint chains for simulated data, where the dotted lines indicate true values, when starting values are chosen to be $1.5, 3.0, 4.5, 6.0, 7.5$, respectively.

Therefore we develop additional grouped move (GM) steps to speed up the convergence especially for the chains of the cutpoints c_k. The idea of GM steps is based on a theorem of Liu and Sabatti (2000) which states: If Γ is a locally compact group of transformations defined on the sample space $\boldsymbol{S}$, L its left-Haar measure, $\boldsymbol{w} \in \boldsymbol{S}$ follows a distribution with density π, and $\gamma \in \Gamma$ is drawn from $\pi(\gamma(\boldsymbol{w}))|J_\gamma(\boldsymbol{w})|L(d\gamma)$, with $J_\gamma(\boldsymbol{w}) = \det(\partial\gamma(\boldsymbol{w})/\partial\boldsymbol{w})$, $\partial\gamma(\boldsymbol{w})/\partial\boldsymbol{w}$ the Jacobian matrix, then $\boldsymbol{w}^* = \gamma(\boldsymbol{w})$ also has density π (Liu and Sabatti (2000), Theorem 1).

Commonly π is considered to be the interesting posterior distribution. The difficulty in the choice of a suitable transformation group is to find one where on the one hand the problematic parameters are transformed and on the other hand the distribution
$\pi(\gamma(\boldsymbol{w}))|J_\gamma(\boldsymbol{w})|L(d\gamma)$ allows to draw samples very fast. We apply this theorem only for the conditional distribution of $\boldsymbol{w} := (y_1,\ldots,y_T, c_3,\ldots,c_{K-1}, \beta_0,\ldots,\beta_p)$ given all the observations and all the remaining parameters. This conditional distribution can be computed iteratively. In order to get an easy sampling distribution we now use the scale group, $\Gamma = \{\gamma > 0 : \gamma(\boldsymbol{w}) = (\gamma w_1,\ldots,\gamma w_d)\}$, with $\gamma^{-1}d\gamma$ as left-Haar measure. This finally leads to a Gamma distribution for γ^2. Therefore, after each iteration of our MCMC sampler, we insert the corresponding GM-step, which consists of drawing a γ^2 from the resulting Gamma distribution and update $\boldsymbol{w}$ to $\gamma \cdot \boldsymbol{w}$. As Figure 2 shows, this significantly speeds up the convergence of the algorithm. Here we used the same simulated data as in Figure 1 and the same starting values. By using the GM steps the chains reach the area around the true values within about 50 iterations.

Standard sampler: Extremely slow convergence of cutpoints.

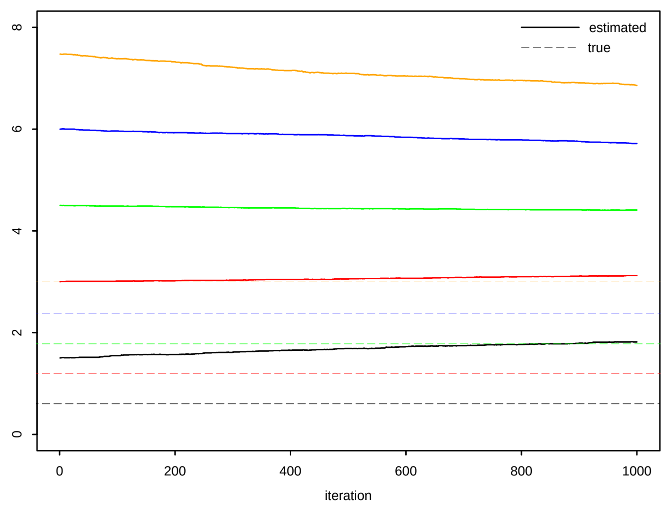


FIGURE 1. First 1000 MCMC iterations for cutpoints produced by standard Gibbs sampler. The dotted lines indicate the true values.

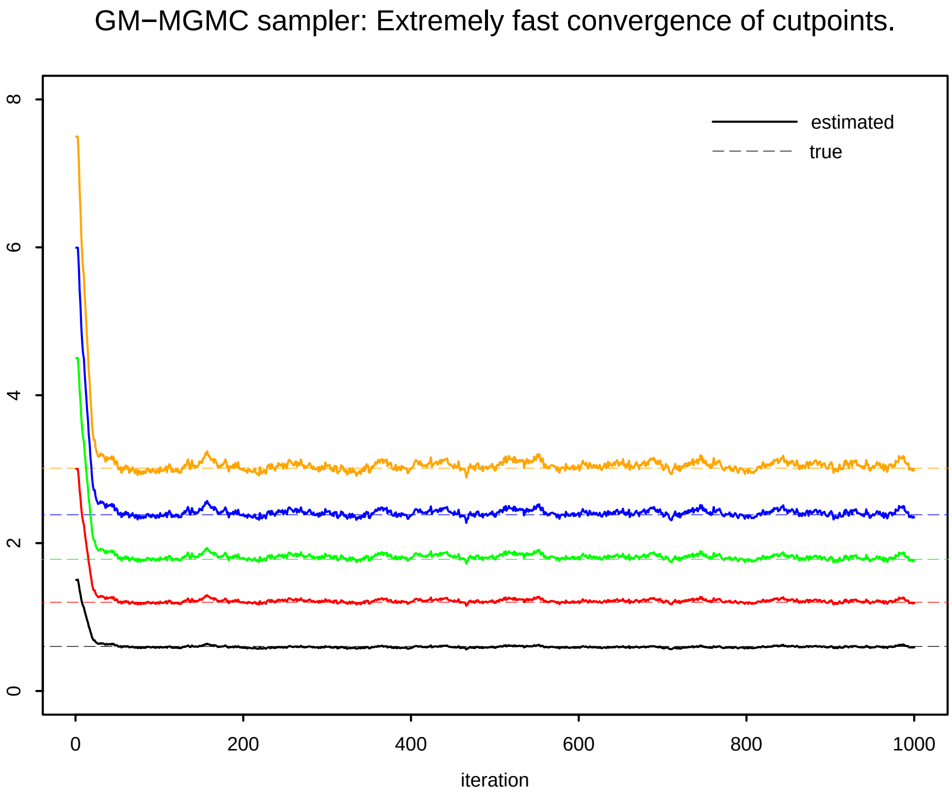


Figure 2. First 1000 iterations of chains for cutpoints produced by GM-MGMC sampler. The dotted lines indicate the true values.

Finally we fit the model to IBM intraday data collected in January 2001. We show that a positive price jump increases the probability that the next price jump will be negative and vice versa. Furthermore, the time between transactions has an impact on the log-volatility in Equation (3): The more time elapses between two subsequent transactions, the higher is the probability for a big price jump (upwards or downwards).

References

[1] Chib, S., Nardari, F. and Shephard, N. (2002): Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics* **108**, 281-316.

[2] Liu, J.S. and C. Sabatti, C. (2000): Generalized Gibbs sampler and multigrid Monte Carlo for Bayesian computation. *Biometrika* **87**, 353-369.

Data Driven Local Coordinates for Linear State Space Systems
M. Deistler
(joint work with T. Ribarits)

The topic of the lecture is embedded in a larger research program at our department concerning identification of ARMA(X) and state space (SS) systems, in particular stressing the multivariate case. The motivation for this program is that in many applications in econometrics and engineering, AR(X) type models are still preferred in modelling linear systems, despite the fact that ARMA(X) and SS models are more flexible. The reasons for this are that parametrization and estimation is much simpler in the AR(X)case, in particular the (maximum likelihood) estimators are of least squares type, they are explicitly given, numerically fast and have no problems of local optima.

In making ARMA(X) and SS systems more competitive, one direction we follow is to look for better parametrizations, in particular for the SS case. Note that ARMA(X) and SS systems represent the same classes of transfer functions, but SS systems in general have larger classes of observationally equivalent systems. The latter fact should be considered as an advantage, because this allows for selection of more suitable representatives.
We consider a state space system

$$x_{t+1} = Ax_t + B\varepsilon_t$$

$$y_t = Cx_t + \varepsilon_t$$

where ε_t is the s–dimensional white noise innovation, x_t is the n–dimensional state and y_t is the s–dimensional output; (A,B,C) are the system matrices.

We consider two approaches: The first approach, *data driven local coordinates (DDLC)*, has been originally introduced by McKelvey and Helmersson. Here (A,B,C) is embedded in $\mathbb{R}^{n^2+2sn}$. For minimal (A,B,C), the equivalence classes are n^2–dimensional manifolds. Commencing from an initial estimator (A,B,C), the orthocomplement to the tangent space to this manifold at (A,B,C) is taken as a parameter space. The likelihood function then is optimized over this parameter space and the procedure is iterated with the new estimate. We analyse the topological and geometrical properties of this parametrization which are relevant for identification. In particular we show that this parametrization is locally homeomorphic, but globally, problems, e.g. of nonidentifiability, arise.

The second approach is *separable least squares DDLC (`slsDDLC`)*. Here first a least squares step is performed in order to concentrate out B. The concentrated likelihood then only depends on (A,C) and for this reduced parameter space, again DDLC is performed. We again analyse the topological and geometrical properties of this parametrization.

Finally, the numerical properties of the maximum likelihood estimation parametrized by the 'classical' echelon form, by DDLC and by `slsDDLC` are investigated in a simulation study. `slsDDLC` is found to be superior to DDLC and both give much better results than echelon forms.

Estimation in Semi-parametric Volatility Models

Feike C. Drost

The availability of large data sets is rapidly growing, especially in finance. In discrete time models, ARMA models with GARCH type errors are quite suitable to pick up the time-varying nature of the first two conditional moments with only a few parameters. However the implications of parametric volatility models for higher order conditional moments, are not reflected in the data. More precisely formulated, the conditional error distribution cannot be described by just a functional form of the conditional volatility and a fixed nonparametric distribution. To avoid this kind of misspecification we use a semi-parametric model where the conditional error distributions are treated as a nuisance parameter. In continuous time models, stochastic volatility models are used to model similar stylized facts. Since the volatility of volatility functions in these models do not affect the first two conditional moments, a nonparametric approach is advised here as well.

Usually, from a practitioners point of view, some finite dimensional parameter is of interest, for example, the mean or median as a measure of location, the Value at Risk as a measure of risk, etc. The question arises how to efficiently estimate such quantities in general semi- and nonparametric models. To study what is best asymptotically, one needs a bound on the asymptotic performance of estimators in the presence of an infinite dimensional nuisance parameter. For the i.i.d. case, a comprehensive account on the present theory along these lines is given in Bickel, Klaassen, Ritov, and Wellner (1993). In financial data, of course, the time dimension also plays an important role. Drost, Klaassen, and Werker (1997) and Koul and Schick (1997) have developed a unified theory for time series models with independently and identically distributed innovations. This covers, for example, semi-parametric ARMA models (Kreiss (1987)). Recent work in applied financial econometrics shows that the assumption of i.i.d. innovations does not hold when using standard semi-parametric time series models, see Engle and Russell (1998).

Based on the first two conditional moments, a popular method to estimate the parameters is the Quasi Maximum Likelihood (QML) approach. This method applies the Maximum Likelihood (ML) procedure to the data as if the conditional distributions are normal. Under some regularity conditions this approach leads to consistent and asymptotically normal estimators, but the efficiency may be quite low.

An alternative to QML, is the Generalized Method of Moments (GMM). Here the conditional moments are used together with a suitably chosen instrument. It is well-known that the QML estimator can be obtained by a suitable choice of the instruments. However, these QML instruments are not optimal since they do not use the possible time varying character of the third and fourth moment of the conditional error distribution. Optimal instruments are easily derived along the lines in Wefelmeyer (1996).

Although the GMM estimator is optimal in the class of estimators based on the first two conditional moments, it is not necessarily the optimal estimator. As in the aforementioned literature, additional information can be gained by estimating the conditional error distribution. In several applications it is known that the conditional error distribution (given the total past) equals the conditional error distribution given some restricted information set $\mathcal{H}_t$. Special cases are:

- i.i.d. errors: choose the restricted information set $\mathcal{H}_t$ to be the trivial sigma field.
- Markov errors: choose the restricted information set $\mathcal{H}_t = \sigma(\varepsilon_t)$, the information set generated by the last error.
- general case: do not put any restrictions on $\mathcal{H}_t$, and choose the restricted information set to be the full information set.

Since the first two conditional moments are of particular interest in financial applications, we present the score functions for QML, GMM, and the Semi-parametric (SP) approach in the following example. The estimator based on the SP score performs best.

Example 1. *Consider the semi-parametric location-scale model,*

$$Y_{t+1} = \mu_t + \sigma_t \varepsilon_{t+1}, \qquad E_{G_t}\varepsilon_{t+1} = 0, E_{G_t}\varepsilon_{t+1}^2 = 1, \qquad G_t \equiv \mathcal{L}(\varepsilon_{t+1}|\mathcal{H}_t),$$

with location-scale score

$$\dot{l}(\varepsilon_{t+1}; G_t) = \left[-\frac{g_t'}{g_t}(\varepsilon_{t+1}), -\frac{1}{2}\left\{1 + \varepsilon_{t+1}\frac{g_t'}{g_t}(\varepsilon_{t+1})\right\}\right]^T.$$

Apart from the model assumptions and some regularity conditions, nothing is known about the conditional error distributions. Put $\gamma_t = E_{G_t}\varepsilon_{t+1}^3, \kappa_t = E_{G_t}\varepsilon_{t+1}^4$, then the score functions of the QML/GMM/SP/ML estimators are given by, respectively,

$$\begin{aligned}
\dot{l}_{t+1}^{QML} &= \left[\frac{\dot{\mu}_t}{\sigma_t}, \frac{\dot{\sigma}_t^2}{\sigma_t^2}\right] \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_{t+1} \\ \varepsilon_{t+1}^2 - 1 \end{bmatrix}, \\
\dot{l}_{t+1}^{GMM} &= \left[\frac{\dot{\mu}_t}{\sigma_t}, \frac{\dot{\sigma}_t^2}{\sigma_t^2}\right] \begin{bmatrix} 1 & \gamma_t \\ \gamma_t & \kappa_t - 1 \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_{t+1} \\ \varepsilon_{t+1}^2 - 1 \end{bmatrix} \equiv \left[\frac{\dot{\mu}_t}{\sigma_t}, \frac{\dot{\sigma}_t^2}{\sigma_t^2}\right] \dot{l}^*(\varepsilon_{t+1}; G_t), \\
\dot{l}_{t+1}^{SP} &= \left\{\left[\frac{\dot{\mu}_t}{\sigma_t}, \frac{\dot{\sigma}_t^2}{\sigma_t^2}\right] - E\left(\left[\frac{\dot{\mu}_t}{\sigma_t}, \frac{\dot{\sigma}_t^2}{\sigma_t^2}\right] \middle| \mathcal{H}_t\right)\right\} \dot{l}(\varepsilon_{t+1}; G_t) \\
&\qquad + E\left(\left[\frac{\dot{\mu}_t}{\sigma_t}, \frac{\dot{\sigma}_t^2}{\sigma_t^2}\right] \middle| \mathcal{H}_t\right) \dot{l}^*(\varepsilon_{t+1}; G_t), \\
\dot{l}_{t+1}^{ML} &= \left[\frac{\dot{\mu}_t}{\sigma_t}, \frac{\dot{\sigma}_t^2}{\sigma_t^2}\right] \dot{l}(\varepsilon_{t+1}; G_t).
\end{aligned}$$

Note that the implied information is strictly increasing: the SP score is closest (in $\mathbb{L}_2$-sense) to the unattainable ML score.

References

[1] Bickel, P.J., Klaassen, C.A.J., Ritov, Y. and Wellner J.A. (1993): *Efficient and Adaptive Statistical Inference for Semi-parametric Models*. John Hopkins University Press, Baltimore.

[2] Drost, F.C., Klaassen, C.A.J. and Werker, B.J.M. (1997): Adaptive estimation in time-series models. *Ann. Statist.* **25**, 786-818.

[3] Engle, R.F. and Russell, J.R. (1998): Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica* **66**, 1127-1162.

[4] Koul, H.L. and Schick, A. (1997): Efficient estimation in nonlinear autoregressive timeseries models. *Bernoulli* **3**, 247-277.

[5] Kreiss, J.-P. (1987): On adaptive estimation in stationary ARMA processes. *Ann. Statist.* **15**, 112-133.

[6] Wefelmeyer, W. (1996): Quasi-likelihood models and optimal inference, *Ann. Statist.* **24**, 405-422.

Modelling Dependence for High–Frequency Data in Finance

Paul Embrechts
(joint work with W. Breymann and A. Dias)

Based on high-frequency data for US$/DM and US$/Yen, stylised facts for extremal dependence in financial data are investigated. Starting with 5' data, through deseasonalisation, data are investigated at the (1 hr, 2 hr, 4 hr, 8 hr, 12 hr, 1 day) frequencies. Dependence is modelled throughout based on the concept of copula. In order to get close to iid bivariate residual vectors, several stochastic models are fitted at the various frequencies. These models include marginal ARMA-GARCH, CCC-GARCH, VECH and DCC-GARCH. The following tests/statistical techniques are performed on the residuals:

- tests for ellipticity
- copula fitting
- dynamic dependence parameters
- comparison of high (low) quantile fitting procedures (leading to the Clayton model)
- spectral measure estimation
- change point analysis.

This work is done jointly with W. Breymann and A. Dias (see [1]) and A. Dias (see [2]). Further references and related work are to be found under www.math.ethz.ch/~embrechts. We also would like to thank Olsen and Associates for providing the data.

References

[1] Breymann, W., Dias, A. and Embrechts P. (2003): Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance* **3** (1), 1-16.

[2] Dias, A. and Embrechts, P.: Dynamic copula models for multivariate high-frequency data in finance. Submitted.

Extremal Behaviour of Continuous-Time Moving Average Processes
Vicky Fasen

We consider a stationary continuous-time moving average (MA) process

$$Y(t) = \int_{-\infty}^{t} f(t-s)\,dL(s) \qquad \text{for } t \geq 0\,,$$

where f is a deterministic kernel function and L is a Lévy process whose increments, represented by $L(1)$, are subexponential and in the domain of attraction of the Gumbel distribution. Examples are Weibull-like distributions with $\alpha \in (0,1)$. The extremal behaviour of subexponential MA processes in the domain of attraction of the Fréchet distribution are well studied Rootzén (1978) and Rosinski and Samorodnitsky (1993). A good overview about subexponential distributions can be found in Embrechts et al. (1997) and about Lévy processes in Sato (1999).

Extremes of $\{Y(t)\}_{t\geq 0}$ are caused by big jumps of the driving Lévy process in combination with large values of the kernel function f. This means that discrete time points $\{t_n\}_{n\in\mathbb{N}}$ chosen properly to incorporate the times where big jumps of the Lévy process and the extremes of the kernel function occur characterise the extremal behaviour of the continuous time process. We restrict ourselves to kernel functions with a finite number of local extremes. Examples for Y include a Weibull-Ornstein-Uhlenbeck process, certain shot noise processes and CARMA processes (Brockwell (2001)).

The extremal behaviour of the discrete-time process $\{Y(t_n)\}_{n\in\mathbb{N}}$ is described by the weak limit of a sequence of marked point processes, i.e.

– by the point processes of exceedances over high thresholds, and
– by marks, which are stochastic processes themselves, and characterize the behaviour of $\{Y(t)\}_{t\geq 0}$, if $Y(t_n)$ exceeds a high threshold.

The limiting distribution of such a sequence of marked point processes is a Poisson process with deterministic marks represented by a scaled version of the kernel function. Further we can compute the normalising constants of the maxima to converge weakly to the Gumbel distribution. The results are similar to the extremal behaviour of discrete MA processes (Davis and Resnick (1988), Rootzén (1986)).

References

[1] Brockwell, P.J. (2001): Lévy-driven CARMA processes. *Ann. Inst. Statist. Math.* **53** (1), 113-123.

[2] Davis, R. and Resnick, S. (1988): Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution. *Stoch. Proc. Appl.* **30**, 41-68.

[3] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997): Modelling Extremal Events for Insurance and Finance. Springer, Berlin.

[4] Rootzén, T. (1978): Extremes of moving averages of stable processes. *Ann. Probab.* **6** (5), 847-869.

[5] Rootzén, H. (1986): Extreme value theory for moving average processes. *Ann. Probab.* **14** (2), 612-652.

[6] Rosinski, J. and Samorodnitsky, G. (1993): Distributions of subadditive functionals of sample paths of infinitely divisible processes. *Ann. Probab.* **21** (2), 996-1014.
[7] Sato, K.I. (1999): *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge.

Nonparametric Value-at-Risk Estimates

Jürgen Franke
(joint work with Mabouba Diagne and Peter Mwita)

We consider a financial time series S_t with returns $R_t = (S_t - S_{t-1})/S_{t-1}$, and we want to estimate the conditional Value-at-Risk, the conditional α-quantile $q_\alpha(r,x) \equiv \mathrm{VaR}(r,x)$ of R_{t+1} given past returns $\vec{R}_t^{(p)} = (R_t, \dots, R_{t-p+1}) = r$ and exogeneous market information $X_t = x \in R^d$ (returns of index or other stock, FX or interest rates etc.), i.e. we have

$$\mathrm{pr}(R_{t+1} \leq q_\alpha(r,x) | \vec{R}_t^{(p)} = r, X_t = x) = \alpha.$$

First, we model the returns as a nonlinear ARX-ARCHX-process

$$R_{t+1} = m(R_t, \dots, R_{t-p+1}, X_t) + \sigma(R_t, \dots, R_{t-p+1}, X_t) Z_{t+1} \tag{1}$$

with i.i.d. innovations Z_t having mean 0 and variance 1 and a known distribution. We can estimate the local trend and volatility functions m, σ nonparametrically by

- kernel estimates or local polynomials for either **lowdimensional** arguments or under restrictions on the functions m, σ (e.g. additive or generalized additive structure)
- neural networks for **highdimensional** arguments

where, for estimating $\sigma^2(s,x)$, we use the residual-based estimator of Fan and Yao (98). Both nonparametric approaches lead to asymptotically normal and, for tuning parameter (bandwidth for local smoothers and number of neurons for neural networks) changing appropriately for increasing sample size, consistent estimates if the time series (R_t, X_t) satisfies some mixing condition (Franke and Diagne, 2002, Franke, Kreiss and Mammen, 2002, Franke et al., 2002, Franke, Neumann and Stockis, 2004). Using estimates for m, σ, we get as a nonparametric VaR-estimate $\hat{q}_\alpha(r,x) = \hat{m}(r,x) + \hat{\sigma}(s,x) q_\alpha^Z$, where q_α^Z denotes the α - quantile of the law of Z_t. An example for a German stock price illustrates the feasibility of the nonparametric approach and the usefulness of incorporating exogeneous information in the calculation of VaR.

Instead of starting from model (1), we can also estimate the quantile function $q_\alpha(r,x)$ directly, either by a nonparametric version of the Koenker-Bassett (1978) regression quantile approach exploiting that

$$q_\alpha(r,x) = \arg\min_{q \in R} E\left(|R_{t+1} - q|_\alpha \middle| \vec{R}_t^{(p)} = r, X_t = x \right)$$

with $|y|_\alpha = (1-\alpha)y^- + \alpha\ y^-$ or by first estimating the conditional distribution function $F(y|r,x) = E(1_{(-\infty,y]}(R_{t+1})|\overset{\rightarrow(p)}{R_t} = r, X_t = x)$ nonparametrically and inverting it. Again, we get consistent nonparametric estimates $\hat{q}_\alpha(r,x)$ based on neural networks (Diagne, 2002) or on local smoothing (Abberger, 1996, Franke and Mwita 2003) if we assume the returns to follow a quantile ARX-model

$$R_{t+1} = q_\alpha(R_t, \ldots, R_{t-p+1}, X_t) + \eta_{t+1}. \tag{2}$$

The innovations η_t may depend on the past $R_n, n < t$, and may have infinite variance, and, in contrast to the VaR estimates based on (1), we do not have to assume their distribution to be known. To get a notion of local variability like volatility which does not require moment assumptions we may specify (2) to the following quantile ARX-ARCHX-model

$$R_{t+1} = q_\alpha(R_t, \ldots, R_{t-p+1}, X_t) + \sigma_\alpha(R_t, \ldots, R_{t-p+1}, X_t)W_{t+1}$$

where the α - scale of R_{t+1} given $\overset{\rightarrow(p)}{R_t} = r, X_t = x$ is the α - quantile of $|R_{t+1} - q_\alpha(r,s)|_\alpha$. The i.i.d. innovations W_t are standardized to have α - quantile 0 and α - scale 1 (compare also Koenker, 1999). Similar as in the familiar model (1), $q_\alpha(r,x)$ and $\sigma_\alpha(r,x)$ may be estimated simultaneously (Mwita, 2003).

References

[1] Abberger, K. (1996): *Nichtparametrische Schätzung bedingter Quantile in Zeitreihen - Mit Anwendung auf Finanzmarktdaten.* Hartung-Gore Verlag, Konstanz.

[2] Diagne, M. (2002): *Financial Risk Management and Portfolio Optimization Using Artificial Neural Networks.* PhD Thesis. TU Kaiserslautern.

[3] Fan, J. and Yao, Q. (1998): Efficient estimation of conditional variance functions in stochastic regression. *Biometrika* **85**, 645-660.

[4] Franke, J., Kreiss, J.-P. and Mammen, E. (2002): Bootstrap of kernel smoothing in nonlinear time series. *Bernoulli* **8**, 1-37.

[5] Franke, J., Kreiss, J.-P., Mammen, E. and Neumann, M. (2002): Properties of the nonparametric autoregressive bootstrap. *J. Time Ser. Anal.* **23**, 555-585.

[6] Franke, J. and Diagne, M. (2002): Estimating market risk with neural networks. Report in Wirtschaftsmathematik 83, University of Kaiserslautern. Submitted.

[7] Franke, J. and Mwita, P. (2003): Nonparametric estimates for conditional quantiles of time series. Report in Wirtschaftsmathematik 87, University of Kaiserslautern. Submitted.

[8] Franke, J., Neumann, M. and Stockis, J.P. (2004): Bootstrapping nonparametric estimates of the volatility function. *J. Econometrics* **118**, 189-218.

[9] Koenker, R. und Basset, G. (1978): Regression quantiles. *Econometrica* **46**, 33-50.

[10] Koenker, R. (1999): Galton, Edgeworth, Frisch and prospects for quantile regression in econometrics. *J. Econometrics* **95**, 347-374.

[11] Mwita, P. (2003): *Semi-parametric Estimation of Conditional Quantiles for Time Series with Applications in Finance.* PhD Thesis, University of Kaiserslautern.

Gibbs Sampling for State Space Modelling of Time Series of Counts
Sylvia Frühwirth-Schnatter
(joint work with Helga Wagner)

For applied statisticians it is not unusual to have to deal with time series of counts. As such data are necessarily non-negative integers, it is often appropriate to assume the observed process y_t follows a Poisson distribution: $y_t \sim \text{Poisson}\,(\lambda_t)$. To capture the effect of exogenous variables z_t, for independent observations a log-linear model could be applied where $\lambda_t = \exp(z_t^{'}\beta)$, with λ_t being the mean of the time series observation y_t given β, and β being a vector of unknown coefficients to be estimated from the data.

To account for the dependency likely to be present in time series data of counts, various extensions of the log-linear model have been suggested which, following Cox (1981), may be classified into parameter-driven and observation-driven models. In an observation driven model the conditional distribution of y_t is specified as a function of the past observations $y_{t-1}, y_{t-2}, \ldots$, see for instance Kaufmann (1987). While observations-driven models are easy to estimate, their theoretical properties can be difficult to derive in comparison to parameter-driven models.

Here we consider parameter-driven models, where the conditional distribution of y_t is allowed to change over time and this change is driven by a latent process. This latent process could be a hidden Markov chain as in Leroux and Puterman (1992), or random effects as in Albert (1992). Smooth changes of the conditional distribution of y_t through state-space models have been considered e.g. in West et al. (1995) and Harvey and Fernandes (1989), whereas a latent stationary autoregressive process has been introduced into the generalized linear model by Zeger (1988).

Estimation of parameter-driven Poisson time series models is known to be a challenging problem. In fact, estimation of these models using maximum likelihood estimation is hampered by the fact that the marginal likelihood, where the latent process is integrated out, is in general not available in closed form. Each evaluation of the likelihood function requires to use some numerical method for solving the necessary high-dimensional integration. One particular useful method in this respect is importance sampling which was applied in Durbin and Koopman (2000) to state space modelling of counts data.

Alternatively, estimation of these models is also feasible within a Bayesian framework using data augmentation as in Tanner and Wong (1987) and Markov chain Monte Carlo (MCMC) methods, as illustrated first by Zeger and Karim (1991). Since this seminal paper, a number of authors have contributed to MCMC estimation of these models. We mention here in particular Shephard and Pitt (1997) and Gamerman (1998) for non-Gaussian time series models based on distributions from the exponential family, and Chib et al. (1998) and Chib and Winkelmann (2001) for more general count data models.

A major difficulties with any of the existing MCMC approaches is that practical implementation requires the use of a Metropolis-Hastings algorithm at least for part of the unknown parameter vector, which in turns make it necessary to define suitable proposal densities in rather high-dimensional parameter spaces. Single-move sampling for this type of models is known to be potentially very inefficient, see e.g. Shepard and Pitt (1997). The main contribution of this article is to show that *straightforward Gibbs sampling of all parameters*, involving only sampling from simple distributions such as multivariate normal, inverse Gamma, exponential and low-dimensional discrete distributions, is feasible for practical Bayesian estimation of most of the parameter-driven models for time series of counts suggested in the literature so far. This rather unexpected result is achieved by introducing two additional sequences of latent variables through data augmentation. One of these sequences are the unobserved inter arrival times of the events under investigation. The introduction of this first sequence eliminates the non-linearity of the observation equation whereas the non-normality of the error term remains which follows a log exponential distribution. As the mean of the exponential distribution is equal to 1, this distribution is independent of any parameter and may be approximated by a mixture of normal distribution in a similar way as in Kim et al. (1998) who used a mixture approximation to the density of a log χ^2-distribution in the context of stochastic volatility models. By introducing the component indicator as a second sequence of missing data, the resulting model may be thought of a partially Gaussian state space model as in Shephard (1994). This is particularly useful for state space models for Poisson time series, as multi-move-sampling of the whole state process through forward-filtering backward sampling as in Frühwirt-Schnatter (1994), Carter and Kohn (1994) and de Jong and Shephard (1995) is now possible.

References

[1] Albert, J.H. (1992): A bayesian analysis of a poisson random-effects model for home run hitters. *Amer. Statist.* **46**, 246-253.

[2] Carter, C.K. and R. Kohn, R. (1994): On Gibbs sampling for state space models. *Biometrika* **81**, 541-553.

[3] Chib, S., Greenberg, E. and Winkelmann, R. (1998): Posterior simulation and Bayes factors in panel count data models. *J. Econometrics* **86**, 33-54.

[4] Chib, S. and Winkelmann, R. (2001): Markov Chain Monte Carlo analysis of correlated count data. *J. Bus. Econom. Statist.* **19**, 428-435.

[5] Cox, D.R. (1981): Statistical analysis of time series: Some recent developments. *Scand. J. Statist.* **8**, 93-108.

[6] de Jong, P. and Shephard, N. (1995): The simulation smoother for time series models. *Biometrika* **82**, 339-350.

[7] Durbin, J. and Koopman, S.J. (2000): Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives.*J. R. Stat. Soc. Ser. B* **62** (1), 3-56.

[8] Frühwirth-Schnatter, S. (1994): Data augmentation and dynamic linear models. *J. Time Ser. Anal.* **15**, 183-202.

[9] Gamerman, D. (1998): Markov chain Monte Carlo for dynamic generalized linear models. *Biometrika* **85**, 215-227.

[10] Harvey, A. C. and Fernandes, C. (1989): Time series models for count or qualitative observations. *J. Bus. Econom. Statist.* **7**, 407-417.

[11] Kaufmann, H. (1987): Regression models for nonstationary categorical time series: Asymptotic estimation theory. *Ann. Statist.* **15**, 79-98.

[12] Kim, S., Shephard, N. and Chib, S. (1998): Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies* **65**, 361-393.
[13] Leroux, B.G. and Puterman, M.L. (1992): Maximum-penalized-likelihood estimation for independent and Markov-dependent mixture models. *Biometrics* **48**, 545-558.
[14] Shephard, N. (1994): Partial non-Gaussian state space. *Biometrika* **81**, 115-131.
[15] Shephard, N. and Pitt, M.K. (1997): Likelihood analysis of non-Gaussian measurement time series. *Biometrika* **84**, 653-667.
[16] Tanner, M.A. and Wong, W.H. (1987): The calculation of posterior distributions by data augmentation, *J. Amer. Statist. Assoc.* **82**, 528-540.
[17] West, M., Harrison, Jeff, P., Migon, H.S. (1985): Dynamic generalized linear models and Bayesian forecasting. *J. Amer. Statist. Assoc.* **80**, 73-83.
[18] Zeger, S.L. (1988): A regression model for time series of counts. *Biometrika* **75**, 621-629.
[19] Zeger, S.L. and Karim, M.R. (1991): Generalized linear models with random effects: a Gibbs sampling approach. *J. Amer. Statist. Assoc.* **86**, 79-86.

Estimation and Change Point Detection with a Hidden Markov Model in Finance

X. Guo

Consider a probability space $(\Omega, \mathcal{F}, P)$ and $t \in [0, T]$ for some $T > 0$. Suppose that $\alpha(t)$ is a finite-state continuous time Markov process with state space $M = \{z^1, \ldots, z^m\}$ and generator $Q = (q^{ij}) \in R^{m \times m}$.

Assume that the Markov process $\alpha(t)$ is observed with the process $y(t)$ such that

$$(1) \qquad \begin{cases} dy(t) = & \mu_{\alpha(t)} dt + \sigma_{\alpha(t)} dw(t), \\ y(0) = & 0 \text{ w.p. } 1, \end{cases}$$

where $w(\cdot)$ is a standard one-dimensional Brownian motion independent of $\alpha(t)$, and the drift μ and diffusion σ take different values when $\alpha(t)$ is in different states.

Given Eq. (1), we are primarily interested in the parameter estimation problem that is motivated by checking the validation of this Markov modulated (or regime switching) model in the financial time series data. The critical issue is the identification of the Markov chain $\alpha(\cdot)$

The problem is trivial when σ's are all distinct and the observation is continuous: the quadratic variation of Ito's calculus will easily reveal the state of $\alpha(t)$. Therefore, we are mostly interested in two cases: (A) when σ's are independent of $\alpha(t)$ and the observation is continuous, and (B) when the observation is discrete. In a joint work with G. Yin [2], we address (A) under the (more general) framework of Wonham filters; In a joint work with D. Chan (included in the summary report [3]), we address (B) and propose a statistical estimation method for applying this regime switching model to analyze financial time series data. Here, we suggest a notion of "regime shift" and a detection method based on a case study of AT&T stock price.

The optimality of Wonham filter is a direct corollary of a result of independent interest concerning the relationship between choices of error functions and the optimality of conditional expectations. This is joint work with A. Banerjee and H. Wang [1].

1. Wonham filters

Assume that the Markov process $\alpha(t)$ is observed with the process $y(t)$ such that

$$\left\{ \begin{array}{rl} dy(t) = & \alpha(t)dt + \sigma(t)dw(t), \\ y(0) = & 0 \ \text{ w.p. } 1, \end{array} \right. \tag{2}$$

as in Eq. (1) where $\sigma(\cdot) : R \mapsto R$, is a continuously differentiable function with $\sigma(t) \geq c$ for all $t \in [0, T]$ and some $c > 0$.

In this framework, suppose we assume that the values of the states $z^1, \cdots, z^m$ and the generator Q are *known* a priori and fixed. Then, a classical result states that the posterior probability $p(t) = (p^1(t), \ldots, p^m(t)) \in R^{1\times m}$, with $p^i(t) = P(\alpha(t) = z^i | y(s),\ 0 \leq s \leq t), p^i(0) = p^i_0, (i = 1, \ldots, m)$ satisfies the following system of stochastic differential equations:

$$\begin{array}{rl} dp^i(t) = & \sum_{j=1}^m p^j(t) q^{ji} dt - \sigma^{-2}(t)\overline{\alpha}(t)[z^i - \overline{\alpha}(t)]p^i(t)dt \\ & +\sigma^{-2}(t)[z^i - \overline{\alpha}(t)]p^i(t)dy(t),\ i = 1, \ldots, m. \end{array} \tag{3}$$

Here, $\overline{\alpha}(t) = \langle z, p(t)\rangle, z = (z^1, \ldots, z^m)'$, and v' denotes the transpose of v. This is known as the *Wonham filter*, which is the first finite dimensional filter for non-Gaussian processes. It is known to be optimal under the mean square error.

1.1. Optimality of conditional expectation as BLFs.

We first show the optimality of Wonham filter under a general class of loss functions known as *Bregman loss functions* (BLFs) (including L^2-loss functions). This is a direct corollary of our study [1], where we provide necessary and sufficient conditions for general loss functions under which the conditional expectation is the unique optimal predictor.

Theorem 1 (Optimality Property)**.** *Let $\phi : R^d \mapsto R$ be a strictly convex, differentiable function. Let $(\Omega, \mathcal{F}, P)$ be an arbitrary probability space and $\mathcal{G}$ a sub-σ-algebra of $\mathcal{F}$. Let X be any $\mathcal{F}$-measurable random variable taking values in R^d for which both $E[X]$ and $E[\phi(X)]$ are finite. Then*

$$\arg\min_{Y\in\mathcal{G}} E[D_\phi(X, Y)] = E[X|\mathcal{G}].$$

Theorem 2 (Exhaustiveness of BLFs)**.** *Let $F : R \times R \mapsto R$ be a non-negative function such that $F(x, x) = 0, \forall x \in R$. Assume that F and F_x are both continuous functions. If for all random variables X, $E[X|\mathcal{G}]$ is the unique minimizer for $E[F(X, Y)]$ over random variables $Y \in \mathcal{G}$, i.e., $\arg\min_{Y\in\mathcal{G}} E[F(X, Y)] = E[X|\mathcal{G}]$, then $F(x, y) = D_\phi(x, y)$ for some strictly convex, differentiable function $\phi : R \mapsto R$.*

Here the BLF $D_\phi : R^d \times R^d \mapsto R$ is defined as $D_\phi(x, y) = \phi(x) - \phi(y) - \langle x - y, \nabla\phi(y)\rangle$, for any (strictly) convex and differentiable function $\phi : R^d \mapsto R$.

For further properties of BLFs and corresponding exhaustiveness results for higher dimensions, see [1].

1.2. **Wonham filter with random parameters.** Now, assume that z^i's (or (q^{ij})) are not available, and that only their noisy/corrupted measurements/observations/distributional information are at our disposal. We propose approximated (suboptimal) filters and prove their (exponential rate) of convergence to the desired Wonham filter under simple ergodic conditions.

For instance, if we assume that a sequence of observations of the form $\widehat{z}_n = (\widehat{z}_n^1, \ldots, \widehat{z}_n^m)' \in R^{m\times 1}$ such that $E\widehat{z}_n = z$ can be obtained, then by defining $\overline{z}_n = \frac{1}{n}\sum_{j=1}^n \widehat{z}_j$, we can construct a sequence of approximations $p_n(t)$ by

(4)
$$\begin{cases} dp_n(t) &= p_n(t)Qdt - \sigma^{-2}(t)\overline{\alpha}_n(t)p_n(t)A_n(t)dt + \sigma^{-2}(t)p_n(t)A_n(t)dy(t), \\ p_n(0) &= p_0, \end{cases}$$

where $\overline{\alpha}_n(t) = \langle p_n(t), \overline{z}_n\rangle$, $A_n(t) = \mathrm{diag}(\overline{z}_n^1 - \overline{\alpha}_n(t), \ldots, \overline{z}_n^m - \overline{\alpha}_n(t))$.

Let $e_n(t) = p_n(t) - p(t)$. Now, if we assume that $\{\widehat{z}_n\}$ is a stationary ergodic sequence with $E\widehat{z}_n = z$, uniformly bounded, and that the sequence $\{\widehat{z}_n\}$ is independent of $\alpha(\cdot)$ and the Brownian motion $w(\cdot)$, then we have:

Theorem 3. *As* $n \to \infty$, $\sup_{0\le t\le T} E|e_n(t)|^2 \to 0$.

Theorem 4.

$$\sup_{0\le t\le T} E|e_n^\kappa(t)|^2 = \begin{cases} o(1), & 0<\kappa<1/2, \\ O(1), & \kappa = 1/2, \end{cases} \quad \text{as } n\to\infty. \tag{5}$$

Theorem 5. (i) *For any positive integer* $\ell > 1$,

$$\sup_{0\le t\le T} E|e_n^\kappa|^{2\ell} = \begin{cases} o(1), & 0<\kappa<1/2, \\ O(1), & \kappa = 1/2, \end{cases} \quad \text{as } n\to\infty. \tag{6}$$

(ii) *As* $n\to\infty$, $\sup_{0\le t\le T} E\exp(|e_n^{1/2}(t)|) = O(1)$.

Similar results are obtained for the error bound estimates in the case when the generator Q is not known a priori. For more details, see [2].

2. Statistical estimation and change point detection in financial time series data

Given discrete feature of financial time series data, a natural statistical problem is the estimation of the states of the Markov chain $\alpha(t)$ when the stock price is observed at discrete time intervals $t = 1, 2, \ldots, n$, i.e.,

$$y_t \;=\; \mu_{\alpha(t)} + \sigma_{\alpha(t)}e_t, \quad e_t \sim \mathrm{N}(0,1), \tag{7}$$

In statistical literature, a model of the above form falls under the umbrella of a more generic class of models called *hidden Markov models* (HMMs). Within a Bayesian framework, we propose a recursive approach for parameter estimation, together with model selection strategies.

A case study of AT&T stock price data indicates that in the financial markets, a given pattern change is more gradual and takes time before its pattern is more sustainable. In this regard, the regime switching model captures this feature well;

our recursive algorithm can be a promising tool in identifying this type of regime change.

For more details of the estimation procedure and on the pros and cons of regime switching models, together with related research problems, see [3].

References

[1] Banerjee, A., Guo, X. and Wang., H. (2003): On the optimality of conditional expectation as a Bregman predictor. http://www.orie.cornell.edu/∼xinguo/papers.html.

[2] Guo, X. and Yin., G. (2003): Wonham filter with random parameters: (exponential) rate of convergence and error bounds. http://www.orie.cornell.edu/∼xinguo/papers.html.

[3] Guo, X.: A regime switching model: estimation, empirical evidence and change point detection. *Proceedings of SIAM-AMS-IMA Research Conference in Mathematical Finance* (to appear). http://www.orie.cornell.edu/∼xinguo/papers.html.

Estimation in Discretely Observed Diffusions: Two Examples of Using Small Δ-Optimality

M. Jacobsen
(joint work with M.L. Østerdal)

Consider a d-dimensional diffusion,

$$dX_t = b_\theta(X_t)\,dt + \sigma_\theta(X_t)\,dB_t$$

driven by a standard d-dimensional Brownian motion and with b_θ a d-dimensional drift function and σ_θ a $d \times d$-matrix valued diffusion function, where both b_θ and $C_\theta := \sigma_\theta\sigma_\theta^T$ are allowed to depend on an unknown p-dimensional parameter $\theta \in \Theta$. It is assumed that for all $\theta \in \Theta$, X has an invariant distribution μ_θ and is ergodic and suitably 'nice'. The task is then to estimate θ based on the observation of $X_{t_1}, \ldots, X_{t_n}$ where $0 < t_1 < \cdots < t_n$. With the likelihood function typically untractable, this may be done using *unbiased estimating* functions $g_{t,\theta}(x,y) = \left(g^k_{t,\theta}(x,y)\right)_{1\le k\le p}$ where the $g_{t,\theta}$ are given in an explicit analytic form and unbiasedness means that (the μ_θ signifying that X_0 has distribution μ_θ)

$$\mathbb{E}_{\mu_\theta} g^k_{t,\theta'}(X_0, X_t) = 0 \quad \text{iff} \quad \theta = \theta'.$$

The estimator $\hat\theta_n$ for θ is now found by solving the equations $G^k_n(\theta) = 0\ (1 \le k \le p)$, where

$$G^k_n(\theta) = \sum_{i=1}^n g^k_{\Delta_i,\theta}\left(X_{t_{i-1}}, X_{t_i}\right),$$

writing $\Delta_i = t_i - t_{i-1}$. If $t_i = i\Delta$ for some $\Delta > 0$, it often holds that if θ is the true parameter value, then $\sqrt{n}\left(\hat\theta_n - \theta\right)$ converges in distribution as $n \to \infty$ to a Gaussian limit $N(0, var_{\Delta,\theta}(g))$. Good choices for $g_{t,\theta}$ are obtained by minimising the asymptotic covariance matrix $var_{\Delta,\theta}(g)$ in a suitable sense.

Some basic examples of unbiased estimating functions are (i) the *simple estimating functions*, see Kessler (2000),

$$g_{t,\theta}^k(x,y) = A_\theta h_\theta^k(x),$$

with $A_\theta = \sum_{i=1}^d b_\theta^i \partial_{x_i} h + \frac{1}{2}\sum_{i,j=1}^d C_\theta^{ij} \partial_{x_i x_j}^2 h$ the infinitesimal generator for X, and (ii) the *martingale estimating functions* introduced by Bibby and Sørensen (1995),

$$g_{t,\theta}^k(x,y) = \sum_{\ell=1}^r \phi_\theta^{k\ell}(x)\left(f_\theta^\ell(y) - \pi_{t,\theta} f_\theta^\ell(x)\right).$$

where $\pi_{t,\theta} f_\theta^\ell(x) = \mathbb{E}_\theta\left[f_\theta^\ell(X_t)\,|X_0 = x\right]$ is known explicitly. The number r is the dimension of the *base* $\left(f_\theta^\ell\right)$ for the estimating functions.

The concept of *small Δ-optimality* (*SΔ-O*), Jacobsen (2001, 2002), aims at minimising $var_{\Delta,\theta}(g)$ as $\Delta \to 0$: (I) if C_θ does not depend on θ, typically

$$var_{\Delta,\theta}(g) = \Delta^{-1} v_{-1,\theta}(g) + O(1)$$

and there is a universal lower bound for $v_{-1,\theta}$ and g is SΔ-O if it achieves this lower bound. With g simple, this is possible only if X is reversible (automatic for $d=1$); for g a martingale estimating function, SΔ-O may be obtained using a base of dimension $r=d$. (II) By contrast, if C_θ depends on all the p parameters,

$$var_{\Delta,\theta}(g) = \Delta^{-1} v_{-1,\theta}(g) + v_{0,\theta}(g) + O(\Delta)$$

and g is SΔ-O provided $v_{-1,\theta}(g) = 0$ (!) and $v_{0,\theta}(g)$ attains its universal lower bound. In this case (II) it is not possible to find simple g that are SΔ-O and martingale estimating functions that are SΔ-O require a base of dimension $d(d+3)/2$, e.g. $f^\ell(x)$ of the form x_i for $1 \le i \le d$ and $x_i x_j$ for $1 \le i \le j \le d$.

The purpose of the present study is to find SΔ-O estimation functions that combine 'simple' with 'martingale' estimating functions,

$$g_{t,\theta}^k(x,y) = tA_\theta h_\theta^k(x) + \sum_{i=1}^d \phi_\theta^{ki}(x)\left(y_i - \pi_{t,\theta} x_i\right) \tag{1}$$

for models where the first order conditional moments are known explicitly, i.e. typically b_θ an affine function of x. Such g *are SΔ-O provided there are functions Φ_θ^k such that*

$$\begin{aligned} h_\theta^k &= \Phi_\theta^k, \quad \phi_\theta^{ki} = \partial_{x_i}\Phi_\theta^k, \\ \partial_{x_i,x_j}^2 \Phi_\theta^k &= \sum_{i',j'=1}^d \left(\partial_{\theta_k} C_\theta^{i'j'}\right)\left(C_\theta^{i'i}\right)^{(-1)}\left(C_\theta^{j'j}\right)^{(-1)}. \end{aligned} \tag{2}$$

(Warning: for $d \ge 2$, a special structure for C and its inverse is of course required for Φ_θ^k that satisfy the last condition to exist at all!)

SΔ-O estimating functions of the form (1) are simpler in structure and may be easier to find than the pure martingale estimating functions needed for models of type (II). To illustrate this, two examples are considered:

Example 1. *A model suggested in the finance literature as a generalization of the Cox-Ingersoll-Ross process: let* $d = 1$, $p = 4$, *with* $b(x) = a + bx$, $\sigma(x) = \sigma x^\gamma$. *Here* $\pi_{t,\theta}x$ *is known but not* $\pi_{t,\theta}x^2$, *which makes it difficult to obtain* $S\Delta$-O *when estimating* σ^2 *and* γ. *But* $\left(g^1_\theta, g^2_\theta\right)$ *of the form (1) with the* h^k_θ *and* ϕ^{ki}_θ *as in (2) is* $S\Delta$-O *provided*

$$\partial_x \Phi^1_\theta(x) = x^{1-2\gamma}, \quad \partial_x \Phi^1_\theta(x) = x^{1-2\gamma}\left((1-2\gamma)\log x - 1\right).$$

Whether this works in practice, is currently being tested! For estimating a *and* b *also (type (I) model), one may combine with a* $S\Delta$-O *martingale estimating function with base* $f^1(x) = x$ *of dimension* 1.

Example 2. *Let* $d \geq 2$ *and consider the* d*-dimensional Ornstein-Uhlenbeck process with* $b(x) = bx$, $C(x) \equiv C$. *Here* $\theta = (b, C)$ *where* $b \in \mathbb{R}^{d\times d}$ *while* $C \in \mathbb{R}^{d\times d}$ *is positive definite. The transition function and therefore the likelihood function is known explicitly, but for* t_i *that are not equidistant becomes most unpleasant to maximize. Again, for estimating* C, *one may use (2) to find* $g^{i_0 j_0}$ *of the form (1) that are* $S\Delta$-O, *viz.*

$$\Phi^{i_0 j_0}_\theta(x) = \sum_{1\leq i\leq j\leq d} x_i x_j \left[C^{(-1)}_{i_0 i} C^{(-1)}_{j_0 j} + C^{(-1)}_{i_0 j} C^{(-1)}_{j_0 i}\right].$$

For estimating b, *combine with the* $S\Delta$-O *martingale estimating function for type (I) models with base* $f^i(x) = x_i$ *of dimension* d*: this still gives quite an unpleasant set of equations for estimating the* b_{ij}, *but it is certainly simpler than the likelihood equations.*

References

[1] Bibby, B.M., Sørensen, M. (1995): Martingale estimating functions for discretely observed diffuions. *Bernoulli* **1**, 17-39.
[2] Jacobsen, M. (2001): Discretely observed diffusions: classes of estimating functions and small Δ-optimality. *Scand. J. Statist.* **28**, 123-149.
[3] Jacobsen, M. (2002): Optimality and small Δ-optimality of martingale estimating functions. *Bernoulli* **8**, 643-668.
[4] Kessler, M. (2000): Simple and explicit estimating functions for a discretely observed diffusion process. *Scand. J. Statist.* **27**, 65-82.

Lévy Copulas for General Lévy Processes

Jan Kallsen
(joint work with Peter Tankov)

Copulas constitute a popular tool to model the dependence of multivariate random variables e.g. in financial and actuarial applications. By virtue of Sklar's theorem, the dependence structure can be considered completely separately from the marginal laws. Various parametric families of Archimedean copulas allow for flexible and parsimoneous modelling (cf. e.g. Nelsen 1999).

In a continuous-time setup, Lévy processes are often applied successfully in order to describe in particular univariate data in finance and insurance. Parametric multivariate Lévy models, however, are scarce and typically very limited as far as the dependence between the components is concerned.

This suggests to transfer the notion of copulas to Lévy processes. In order to obtain a time-independent concept one works with the Lévy-Khinchine triplet. Since the correlation structure of the Brownian motion part is completely determined by the covariance matrix, it remains to consider the Lévy measure.

Tankov (2003) introduced a notion of copulas on the level of Lévy measures for multivariate processes with only positive jumps. In Kallsen and Tankov (2004) this concept of *Lévy copulas* is generalized to general Lévy processes X. Similarly as for random vectors, they are defined as tail integrals of measures with uniform marginals. An analogue of Sklar's theorem states that the Lévy measure can be recovered from the Lévy copula and the marginal Lévy measures. Conversely, any Lévy copula and any univariate Lévy measures can be combined to yield a Lévy measure. Archimedean Lévy copulas as e.g. the Clayton family are defined similarly as in the case of random vectors.

Finally, two limit theorems are discussed which show how to obtain the Lévy copula and also the Gaussian copula corresponding to the Brownian motion part of X as a limit of properly rescaled copulas of the random vectors X_t for $t \to 0$. The proof of these results relies on a characterization of weak convergence in terms of copula convergence by Lindner and Szimayer (2004).

References

[1] Kallsen, J. and Tankov, P. (2004): Lévy copulas for general Lévy processes. Preprint.

[2] Lindner, A. and Szimayer, A. (2004): A limit theorem for copulas. Preprint.

[3] Nelsen, R. (1999): *An Introduction to Copulas*. Springer, New York.

[4] Tankov, P. (2003): Dependence structure of spectrally positive Lévy processes. Preprint.

Forecasting Daily Variability of the S&P 100 Stock Index Using Historical, Realised and Implied Volatility Measurements

Siem Jan Koopman

(joint work with Borus Jungbacker and Eugenie Hol)

ABSTRACT

The increasing availability of financial market data at intraday frequencies has not only led to the development of improved volatility measurements but has also inspired research into their potential value as an information source for volatility forecasting. In this paper we explore the forecasting value of historical volatility (extracted from daily return series), of implied volatility (extracted from option pricing data) and of realised volatility (computed as the sum of squared high frequency returns within a day). First we consider unobserved components and long

memory models for realised volatility which is regarded as an accurate estimator of volatility. The predictive abilities of realised volatility models are compared with those of stochastic volatility models and generalised autoregressive conditional heteroskedasticity models for daily return series. These historical volatility models are extended to include realised and implied volatility measures as explanatory variables for volatility. The main focus is on forecasting the daily variability of the Standard & Poor's 100 stock index series for which trading data (tick by tick) of almost seven years is analysed. The forecast assessment is based on the hypothesis of whether a forecast model is outperformed by alternative models. In particular, we will use superior predictive ability tests to investigate the relative forecast performances of some models. Since volatilities are not observed, realised volatility is taken as a proxy for actual volatility and is used for computing the forecast error. A stationary bootstrap procedure is required for computing the test statistic and its p-value. The empirical results show convincingly that realised volatility models produce far more accurate volatility forecasts compared to models based on daily returns. Long memory models seem to provide the most accurate forecasts.

Description of paper

Modelling and forecasting volatility in financial markets has gained much interest in the financial and economic literature. The seminal paper of Engle (1982) has started the development of a large number of so-called historical volatility models in which a time-varying volatility process is extracted from financial returns data. Most volatility models can be regarded as variants of the generalised autoregressive conditional heteroskedasticity (GARCH) models of Bollerslev (1986), see Bollerslev et al. (1994) for a review. A rival class of volatility models is associated with the stochastic volatility (SV) model, see Taylor (1986) and Harvey et al. (1994). The overviews presented in Shephard (1996) and Ghysels et al. (1996) provide an excellent introduction to historical volatility models. A more recent review of volatility models together with an assessment of their forecasting performances is given by Poon and Granger (2003).

Both GARCH and SV models are regularly used for the analysis of daily, weekly and monthly returns. ¿From a theoretical perspective these models can also be applied to returns data measured at higher frequencies (intraday). However, it is learned from empirical studies that these models can not accomodate all information in high frequency returns. The initial work of Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (2001) show that realised volatility (a daily volatility measure) as computed by the cumulative sum of squared intraday returns is less subject to measurement error and therefore less noisy. This empirical fact is supported by the theory that the measurement noise contained in daily squared returns prevents the observation of the volatility process while it is reduced as the sampling frequency of the return series from which volatility is calculated is increased, see Andersen, Bollerslev, Diebold and Labys (2001) and Barndorff-Nielsen and Shephard (2001, 2002). These results also justify the earlier work of French et al. (1987), amongst others. Andersen and Bollerslev (1998) show that daily forecasts of exchange rates based on GARCH models, when evaluated against realised volatility, are far more accurate than had been previously assumed. These findings were subsequently confirmed with regards to stock index data by Blair et al. (2001) and Hansen and Lunde (2003) who examined the predictive accuracy of volatility forecasts based on GARCH models.

Volatility can be extracted from returns data but it can also be derived from option pricing data in combination with an option pricing model. Early empirical studies have indicated that implied volatility, when compared with historical standard deviations, can be regarded as a good predictor of future volatility. Implied volatility is often referred to as the market's volatility forecast and is said to be forward looking as opposed to historical based methods which are by definition backward looking. Recent study by Blair et al. (2001) shows that accurate volatility forecasts for returns on stock indices are often based on implied volatility. Moreover, their research strongly suggests that daily returns contain little or no incremental information about future volatility.

In this paper we investigate the potential gains of different measures of volatility and different ways of modelling these data for the purpose of volatility forecasting. For example, it is suggested to incorporate realised volatility as an explanatory variable in the variance equation of a daily GARCH model. They found a considerable improvement in the forecasting performance in this way. Another possible explanatory variable for volatility is implied volatility. We will explore this option further by incorporating such explanatory variables in both GARCH and SV models.

Realised volatility can also be modelled directly which is reminiscent of the methods adopted for monthly volatility in a number of earlier studies. The forecasting performance of realised volatility models has been studied, amongst others, by Andersen, Bollerslev, Diebold and Ebens (2001) and Barndorff-Nielsen and Shephard (2004). In the first paper, it is stressed that long memory features are present in the logarithms of realised volatility and that the autoregressive fractionally integrated moving average (ARFIMA-RV) model is effective in empirical modelling. The second paper builds on Barndorff-Nielsen and Shephard (2002) where volatility is represented as a continuous time series process, the sum of independent Lévy driven Ornstein-Uhlenbeck (OU) processes. This approach forms the basis of an unobserved components (UC-RV) model for realised volatility that consists of independent ARMA components with restricted parameters.

The empirical investigation is for the Standard & Poor's 100 (S&P 100) stock index series over the period 6 January 1997 to 14 November 2003 with 1725 trading days. Opening and closure prices for all trading days in the sample are available in this period together with all price quotes within the days (tick by tick). Further we have obtained the S&P 100 implied volatility index from the Chicago Board Options Exchange Market volatility index (VIX) which is known to be a highly liquid options market. The forecasting performance of various volatility models for the last 525 days of the data set is the focus of the empirical study. We compare the forecasts of ARFIMA-RV, UC-RV, SV and GARCH volatility models;

the latter two models are considered with and without explanatory variables. The forecasts are generated by a rolling-window of 1200 observations through the last 525 daily observations. Forecast comparison is based on four different loss functions including the mean squared error and the mean absolute error statistics. The fact that a particular loss criterion is smallest for a particular model does not provide any information about its forecast superiority in other samples of the data set and in future samples of the data. The results in White (2000) and the important refinements in Hansen (2001) constitute a framework that constructs a formal test for superior prediction ability (SPA) of a benchmark or base model relative to a set of rival models. Since volatility can never be observed, realised volatility is taken as a proxy for actual volatility and used for determining the forecast error. This may introduce inconsistencies in the ranking of forecast models but it is argued that the occurrence of such inconsistencies are unlikely in our study. The method of computing the SPA test statistic and its p-value requires bootstrap samples obtained by, for example, the stationary bootstrap procedure of Politis and Romano (1994). The construction of the test and some details of implementation are discussed.

The findings of this extensive empirical study are presented by reporting a selection of the most interesting results. The maximum likelihood estimates for the coefficients of the considered models are reported for the full sample. Although these estimates are not used for forecasting since all models are re-estimated for each rolling window sample (starting from 17 October 2001), the reported estimation results provide insights about the S&P 100 data set and the effectiveness of models to capture volatility information from the data. A selection of the forecasting results is also presented but most attention is paid to the SPA results. It has become clear that the realised volatility models are overwhelmingly superior and therefore making comparisons between, say, GARCH and ARFIMA-RV is not useful. We therefore concentrate on the comparison of models within the two classes of realised volatility models and historical volatility models. It will be concluded that both the ARFIMA-RV and the SV model with realised volatility as the explanatory variable are superior within their classes for the forecasting of S&P 100 volatility. To get some insight in how forecasts evolve over time in our study, in Figure 1 we present one-step ahead forecasts for the S&P 100 volatility between 9 September 2002 and 18 November 2002 (51 trading days).

References

[1] Andersen, T.G. and Bollerslev, T. (1998): Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *Internat. Econom. Rev.* **39**, 885–905.

[2] Andersen, T.G., Bollerslev, T., Diebold, F.X., and Ebens, H. (2001): The distribution of realized stock return volatility. *Journal of Financial Economics* **61**, 43–76.

[3] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. (2001): The distribution of exchange rate volatility. *J. Amer. Statist. Assoc.* **96**, 42–55.

[4] Barndorff-Nielsen, O.E. and Shephard, N. (2001): Non-Gaussian OU based models and some of their uses in financial economics (with discussion). *J. R. Stat. Soc. Ser. B* **63**, 167–241.

[5] Barndorff-Nielsen, O.E. and Shephard, N. (2002): Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *J. R. Stat. Soc. Ser. B* **64**, 253–280.

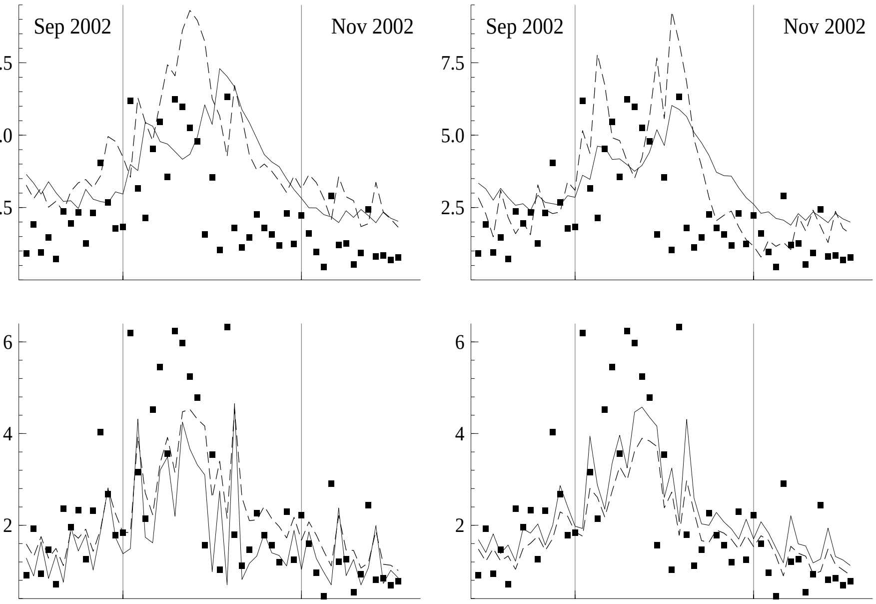


FIGURE 1. Realised volatility (as dots) and one-day ahead volatility forecasts from (i) GARCH (solid) and GARCH with RV (dashed), (ii) SV (solid) and SV with RV (dashed), (iii) UC-RV1 (solid) and UC-RV2 (dashed) and (iv) ARFIMA-RV (solid) and log ARFIMA-RV (dashed) models for the period between 9 September 2002 and 18 November 2002 (day 225 to 275).

[6] Barndorff-Nielsen, O.E. and Shephard, N. (2004): Measuring and forecasting financial variability using realised variance. In: Harvey, A.C., Koopman, S.J. and Shephard, N. (Eds.), *State Space and Unobserved Components Models: Theory and Applications*, Cambridge University Press, Cambridge.

[7] Beckers, S. (1981): Standard deviations implied in options prices as predictors of future stock price variability. *Journal of Banking and Finance* **5**, 363–381.

[8] Blair, B.J., Poon, S.H. and Taylor, S.J. (2001): Forecasting S&P 100 volatility: the incremental information content of implied volatilities and high frequency returns. *J. Econometrics* **105**, 5–26.

[9] Bollerslev, T. (1986): Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**, 307–327.

[10] Bollerslev, T., Engle, R.F. and Nelson, D.B. (1994): ARCH models. In: Engle, R.F. and McFadden, D.L. (Eds.), *Handbook of Econometrics*, Volume 4, pp. 2959–3038, Elsevier Science, Amsterdam.

[11] Chiras, D.P. and Manaster, S. (1978): The information content of options prices and a test of market efficiency. *Journal of Financial Economics* **6**, 213–234.

[12] Christensen, B.J. and Prabhala, N.R. (1998): The relation between implied and realized volatility. *Journal of Financial Economics* **50**, 125–150.

[13] Engle, R.F. (1982): Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom Inflation. *Econometrica* **50**, 987–1007.

[14] Fleming, J. (1998): The quality of market volatility forecast implied by S&P 100 index option prices. *Journal of Empirical Finance* **5**, 317–345.

[15] French, K.R., Schwert, G.W. and Stambaugh, R.F. (1987): Expected stock returns and volatility. *Journal of Financial Economics* **19**, 3–29.

[16] Ghysels, E., Harvey, A.C. and Renault, E. (1996): Stochastic Volatility. In: Maddala, G.S. and Rao, C.R. (Eds.), *Handbook of Statistics*, Volume 14, Statistical Methods in Finance, pp. 119–191, North-Holland, Amsterdam.

[17] Giot, P. (2003). The information content of implied volatility in agricultural commodity markets. *Journal of Futures Markets* **23**, 441–454.

[18] Hansen, P.R. (2001): A test for superior predictive ability. Discussion Paper, Brown University Working Paper 2001-06.

[19] Hansen, P.R. and Lunde, A. (2003): A forecast comparison of volatility models: does anything beat a GARCH(1,1)? Discussion Paper, Brown University Working Paper.

[20] Harvey, A.C., Ruiz, E. and Shephard, N. (1994): Multivariate Stochastic Variance models. *Rev. Econom. Stud.* **61**, 247–264.

[21] Latané, H.A. and Rendleman, R.J. (1976): Standard deviations of stock price ratios implied in option prices. *J. Finance* **31**, 369–381.

[22] Martens, M. (2001): Forecasting daily exchange rate volatility using intraday returns. *Journal of International Money and Finance* **20**, 1–23.

[23] Politis, D.N. and Romano, J.P. (1994): The stationary bootstrap. *J. Amer. Statist. Assoc.* **89**, 1303–1313.

[24] Poon, S.H. Granger, C. (2003): Forecasting volatility in financial markets: a review. *Journal of Economic Literature* **41**, 478–539.

[25] Poon, S.H. and Taylor, S.J. (1992): Stock returns and volatility: an empirical study of the UK stock market. *Journal of Banking and Finance* **16**, 37–59.

[26] Shephard, N. (1996): Statistical aspects of ARCH and Stochastic Volatility. In: Cox, D.R., Hinkley, D.V. and Barndorff-Nielsen, O.E. (Eds.), *Time Series Models in Econometrics, Finance and Other Fields*, Number 65 in Monographs on Statistics and Applied Probability, pp. 1–67, Chapman and Hall, London.

[27] Taylor, S.J. (1986): *Modelling financial time series.* Wiley, Chichester.

[28] White, H. (2000): A reality check for data snooping. *Econometrica* **68**, 1097–1126.

Leroux's method for General Hidden Markov Models and Stochastic Volatility Models

Catherine Larédo

(joint work with Valentine Genon-Catalot)

Parametric inference for Hidden Markov Models (H.M.M.) has been widely investigated, especially in the last decade. The observed process (Z_n) is modelled via an unobserved Markov chain (U_n). When studying the statistical properties of H.M.M.s, a difficulty arises since the exact likelihood cannot be explicitly calculated. As a consequence, many authors have studied approximations by means of numerical and simulation techniques (see for instance Del Moral et al., 2001; Pitt and Shephard, 1999; Durbin and Koopman, 1997).

The theoretical study of the exact maximum likelihood has been investigated for finite state space (see Leroux, 1992; Bickel and Ritov, 1996; Bickel et al., 1998) and for compact state space (see Jensen and Petersen, 1999; Douc and Matias, 2001). In previous papers (Genon-Catalot et al., 1998, 1999, 2000, 2003), we have investigated some statistical properties of discretely observed Stochastic Volatility models (S.V.). When the sampling interval is fixed, stochastic volatility models are H.M.M.s, for which the hidden chain has non-compact state space.

We extend here a method of Leroux (1992) to study the likelihood and related contrast processes for general hidden Markov models. We define the entropy associated to these models and characterize the limit of the loglikelihood and related processes, under specific assumptions.

Generic examples of such processes are obtained setting $Z_n = G(U_n, \epsilon_n)$, where $G : \mathcal{U} \times \mathbb{R}^l \to \mathbb{R}$ is a known function, (U_n) is a strictly stationary Markov chain on $\mathcal{U}$, and (ϵ_n) a sequence of i.i.d random variables on $\mathbb{R}^l$, independent of (U_n) with known density. These methods are applied to the Kalman filter ($G(u, v) = u + v$ and (U_n) is AR(1)), to stochastic volatility models ($G(u, v) = \sqrt{u} \times v$ and (U_n) a Markov chain in $\mathbb{R}^2$), and to the multiplicative explicit filter proposed by Genon-Catalot and Kessler (2004).

This research was supported in part by Dynstoch European Network.

References

[1] Bickel, P.J. and Ritov, V. (1996): Inference in hidden Markov models. I. Local asymptotic normality in the stationary case. *Bernoulli* **2** (3), 199-228.

[2] Bickel, P.J., Ritov, Y. and Ryden, T. (1998): Asymptotic normality of the maximum likelihood estimator for general hidden Markov models. *Ann. Statist.* **26**, 1614-1635.

[3] Del Moral, P., Jacod, J. and Protter, Ph. (2001): The Monte-Carlo method for filtering with discrete time observations. *Probab. Theory Related Fields* **120** (3), 346-368.

[4] Douc, P. and Matias, L. (2001): Asymptotics of the maximum likelihood estimator for general hidden Markov models. *Bernoulli* **7** (3), 381-420.

[5] Durbin, J. and Koopman, S.J. (1997): Monte-Carlo maximum likelihood estimation for non-Gaussian state space models. *Biometrika* **84** (3), 669-684.

[6] Genon-Catalot, V., Jeantheau, T. and Larédo, C. (1998): Limit theorems for discretely observed stochastic volatility models. *Bernoulli* **4** (3), 283-303.

[7] Genon-Catalot, V., Jeantheau, T. and Larédo, C. (1999): Parameter estimation for discretely observed stochastic volatility models. *Bernoulli* **5** (5), 855-872.

[8] Genon-Catalot, V., Jeantheau, T. and Larédo, C. (2000): Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli* **6** (6), 1051-1079.

[9] Genon-Catalot, V., Jeantheau, T. and Larédo, C. (2003): Conditional likelihood estimators for hidden Markov models and stochastic volatility models. *Scand. J. Statist.* **30** (2), 297-316.

[10] Genon-Catalot, V. and Kessler, M. (2004): Random scale perturbation of an AR(1) and its properties as a non-linear explicit filter. *Bernoulli* (to appear).

[11] Jensen, J.L. and Petersen, N.V. (1999): Asymptotic normality of the maximum likelihood estimator in state space models. *Ann. Statist.* **27** (2), 514-535.

[12] Leroux, B.G. (1992): Maximum likelihood estimation for hidden Markov models. *Stoch. Proc. Appl.* **40**, 127-143.

[13] Pitt, M.K. and Shephard, N. (1999): Filtering via simulation: auxiliary particle filters. *J. Amer. Statist. Assoc.* **94** (446), 590-599.

A Continuous Time GARCH(1,1) Process

Alexander Lindner
(joint work with Claudia Klüppelberg and Ross Maller)

1. Introduction

Discrete time GARCH(1,1) models are commonly used to model financial time series like asset prices and exchange rates. They capture many of the so-called stylized features such as heavy tails and uncorrelatedness without being independent. The latter is e.g. manifested in the nonzero autocorrelation of the squared sequence. Various attempts have been made to capture these features in a *continuous time* model such as diffusion approximations (see e.g. Duan (1996) or Nelson (1990)) and other stochastic volatility models, as e.g. in Anh et al. (2002) or Barndorff-Nielsen and Shephard (2001). These models have in common that they are driven by two random processes. Here, we propose a continuous time GARCH(1,1) model with only one source of randomness, capturing the stylized features by the dependence structure alone. The talk is based on results of Klüppelberg et al. (2004).

2. From discrete to continuous GARCH

The discrete time GARCH(1,1) process is given by

$$Y_n = \sigma_n \varepsilon_n,\ n \in \mathbb{N}_0,\ \text{where } \sigma_n^2 = \beta + \lambda Y_{n-1}^2 + \delta \sigma_{n-1}^2$$

with constants $\beta, \delta > 0$, $\lambda \geq 0$ and an iid sequence $(\varepsilon_n)_{n\in\mathbb{N}_0}$, independent of σ_0^2. Then σ_n can be written as

$$\begin{aligned} \sigma_n^2 &= \beta \sum_{i=0}^{n-1} \prod_{j=i+1}^{n-1} (\delta + \lambda \varepsilon_j^2) + \sigma_0^2 \prod_{j=0}^{n-1} (\delta + \lambda \varepsilon_j^2) \\ (1) = \quad &\beta \int_0^n \exp\left(\sum_{j=\lfloor s \rfloor + 1}^{n-1} \log(\delta + \lambda \varepsilon_j^2) \right) ds + \sigma_0^2 \exp\left(\sum_{j=0}^{n-1} \log(\delta + \lambda \varepsilon_j^2) \right),\ n \in \mathbb{N}. \end{aligned}$$

This suggests, in continuous time, to replace the noise variables ε_n by the increments $\Delta L_t = L_t - L_{t-}$ of a Lévy process $(L_t)_{t\geq 0}$. Keep $\beta, \delta > 0$, $\lambda \geq 0$, and define the process $(X_t)_{t\geq 0}$ by

$$X_t = -t \log \delta - \sum_{0<s\leq t} \log(1 + \frac{\lambda}{\delta}(\Delta L_s)^2), \quad t \geq 0.$$

Then, in analogy with (1), for a finite random variable $\sigma_0 \geq 0$, independent of $(L_t)_{t\geq 0}$, define the left-continuous *volatility process* $(\sigma_t)_{t\geq 0}$ by

$$\sigma_t^2 = \left(\beta \int_0^t e^{X_s} ds + \sigma_0^2 \right) e^{-X_t}, \quad t \geq 0, \tag{2}$$

and the *continuous time GARCH* (*"COGARCH"*) process $(G_t)_{t\geq 0}$ as the cádlág process satisfying

$$dG_t = \sigma_t \, dL_t, \quad t \geq 0, \quad G_0 = 0. \tag{3}$$

Thus G jumps as the same times as L does, and has jumps of size $\Delta G_t = \sigma_t \Delta L_t$.

3. Properties of the model

In this section we give some of the properties of model (2), (3). First, we note that $(X_t)_{t\geq 0}$ defines a spectrally negative Lévy process of bounded variation with drift $\gamma_{X,0} = -\log \delta$. For the volatility process, it holds:

Proposition 1. *The process $(\sigma_t^2)_{t\geq 0}$ satisfies the stochastic differential equation*

$$d\sigma_{t+}^2 = \beta dt + \sigma_t^2 e^{X_{t-}} d(e^{-X_t}), \quad t > 0,$$

and we have

$$\sigma_t^2 = \beta t + \log \delta \int_0^t \sigma_s^2 ds + (\lambda/\delta) \sum_{0<s<t} \sigma_s^2 (\Delta L_s)^2 + \sigma_0^2, \quad t \geq 0.$$

Denote by Π_L the Lévy measure of $(L_t)_{t\geq 0}$, and assume that it is nonzero. Then we can give necessary and sufficient conditions for strict stationarity of the volatility process $(\sigma_t^2)_{t\geq 0}$.

Theorem 2. *The volatility process $(\sigma_t^2)_{t\geq 0}$ is a time homogeneous Markov process. The random variable σ_0 can be chosen such that $(\sigma_t^2)_{t\geq 0}$ is strictly stationary, if and only if*

$$\int_{-\infty}^{\infty} \log(1 + \frac{\lambda}{\delta} y^2) \Pi_L(dy) < -\log \delta.$$

In that case, for any $k \in \mathbb{N}$, σ_t^2 has finite k'th moment if and only if $EL_1^{2k} < \infty$ and

$$\Psi(k) := k \log \delta + \int_{-\infty}^{\infty} \left((1 + \frac{\lambda}{\delta} y^2)^k - 1 \right) \Pi_L(dy) < 0.$$

If $EL_1^4 < \infty$ and $\Psi(2) < 0$, then the autocovariance function of σ_t^2 decreases exponentially with the lag.

Using Theorem 2, it can be shown that for any Lévy process $(L_t)_{t\geq 0}$ the stationary version of the volatility process $(\sigma_t^2)_{t\geq 0}$ has certain infinite moments. In that sense, the volatility process is heavy tailed. For the COGARCH process $(G_t)_{t\geq 0}$ itself, we have:

Theorem 3. *Assume $(\sigma_t^2)_{t\geq 0}$ is the strictly stationary volatility process. Then the integrated GARCH(1,1) process $(G_t)_{t\geq 0}$ has stationary increments. Assume further that $EL_1^8 < \infty$ and $\Psi(4) < 0$, that $(L_t)_{t\geq 0}$ is a quadratic pure jump process* *(i.e. has no Gaussian component) and that* $EL_1 = 0$, $\int_{\mathbb{R}} y^3 \Pi_L(dy) = 0$. *Let* $r > 0$ *be fixed. Then there exists a positive constant* C_r *such that for any* $t \geq 0$ *and* $h \geq r$*:*

$$\begin{aligned} \mathrm{Cov}(G_{t+r} - G_t, G_{t+r+h} - G_{t+h}) &= 0, \\ \mathrm{Cov}((G_{t+r} - G_t)^2, (G_{t+r+h} - G_{t+h})^2) &= C_r e^{h\Psi(1)}. \end{aligned}$$

Theorem 3 shows, in analogy with the discrete time GARCH model, that the increments of $(G_t)_{t\geq 0}$ are uncorrelated, but that their squares are not.

References

[1] Anh, V.V., Heyde, C.C. and Leonenko, N.N. (2002): Dynamic models of long–memory processes driven by Lévy noise. *J. Appl. Prob.* **39**, 730-747.

[2] Barndorff-Nielsen, O.E. and Shephard, N. (2001): Non-Gaussian Ornstein–Uhlenbeckbased models and some of their uses in financial economics (with discussion). *J. R. Stat. Soc. Ser. B* **63**, 1-42.

[3] Duan, J.C. (1996): A unified theory of option pricing under stochastic volatility – from GARCH to diffusion. Working paper, available at http://www.rotman.utoronto.ca/∼jcduan/.

[4] Klüppelberg, C., Lindner, A. and Maller, R. (2004): A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour. *J. Appl. Prob.* **41** (3) (to appear).

[5] Nelson, D.B. (1990): ARCH models as diffusion approximations. *J. Econometrics* **45**, 7-38.

The Large-Sample Distribution of the Sharpe Ratio

R. A. Maller

1. Introduction and Summary

In the Markowitz efficient portfolio paradigm, we maximise the expected return on a portfolio of assets for a given level of "risk", as measured by the standard deviation of the portfolio return. Among the set of portfolios derived in this way, we can select the one which has the *maximum* return to risk tradeoff, as measured by the ratio of expected return (excess over the risk-free rate) to standard deviation of return, that is, the portfolio with maximum *Sharpe ratio.* This portfolio has desirable optimality properties and is important both for purposes of allocation of resources and for the performance evaluation of portfolios.

Given sample estimates of the mean vector and covariance matrix of the excess returns which are asymptotically normally distributed, we might expect to get asymptotic normality of the maximised Sharpe ratio. But because the maximisation procedure means that we are not dealing with just a simple ratio of mean to standard deviation, this is not true in general, though it is in some cases. We are able to give a complete description of the large-sample behaviour of the Sharpe ratio for a wide class of portfolios, and (when there are restrictions on short-selling), a partial solution which still covers some useful situations – but we merely summarise the results here. (For details see Maller, 2004.)

Although not always asymptotically normal, the Sharpe ratio is in the domain of attraction of the normal in the cases we study, so the usual kinds of statistical analyses which are applied to the Sharpe ratio are valid, at least in large samples.

2. Background – the Markowitz paradigm

We are given a d–vector $\tilde{\mu}$ of expected asset returns and an associated $d \times d$ positive definite covariance matrix Σ. The excess returns are:

$$\mu = \tilde{\mu} - ri,$$

where r is the risk-free rate and i is a d-vector each of whose elements is 1. The optimisation problem is to choose a d–vector x of asset weights such that the portfolio standard deviation

$$\sigma_p = \sqrt{x^T \Sigma x}$$

is minimised for a specified expected return,

$$\mu_p = x^T \mu$$

(or, equivalently, μ_p is maximised for a specified level of risk, σ_p.) ¿From pairs (μ_p, σ_p) constructed in this way we can trace out an efficient frontier, representing portfolios whose return/risk tradeoff is optimal in the mean–variance sense.

The vector x will be further restricted to a class C, say, which must include $\{i^T x = 1\}$ (the "total allocation constraint"). We only consider C of the form

$$C_A = \mathbb{R}^d \cap \left\{x : i^T x = 1\right\},$$

or

$$C_+ = \mathbb{R}^d \cap \left\{x : x \geq 0, i^T x = 1\right\}.$$

In C_A, the components of x may be negative – short sales of assets are allowed. In C_+, the components of x are non-negative – short sales of assets are not allowed.

The *Sharpe ratio* (SR) of a portfolio (or a single asset) is its expected (excess) return divided by its standard deviation. We ask for the portfolio with the *maximum SR* along the efficient frontier. This is the portfolio with the highest return/risk tradeoff achievable from the assets: the *optimal risky portfolio*.

The basics of the optimisation problem have been well understood since the seminal work of Markowitz (1952, 1991). When short sales are allowed and

$$i^T \Sigma^{-1} \mu > 0, \tag{1}$$

the optimal risky portfolio is located at the point of tangency of a line from the origin (since we have excess returns) to the efficient frontier. But when

$$i^T \Sigma^{-1} \mu < 0 \tag{2}$$

following this method gives a portfolio with the *minimum* SR. Maller and Turkington (2002) showed how to find the portfolio with the maximum SR achievable in this case. The case in (2) is not by any means pathological.

Suppose first that $C = C_A$, so we wish to maximise the function

$$f(x) = \frac{x^T\mu}{\sqrt{x^T\Sigma x}}$$

for variations in x, under the sole constraint that $i^T x = 1$. An easy analysis gives

$$\sup_{i^T x=1} |f(x)| \le \sqrt{\mu^T\Sigma^{-1}\mu},$$

and, supposing now that (1) holds, we get

$$\sup_{i^T x=1} f(x) = +\sqrt{\mu^T\Sigma^{-1}\mu}.$$

This is achieved for the allocation

$$x_{\max} = \frac{\Sigma^{-1}\mu}{i^T\Sigma^{-1}\mu}.$$

This is a textbook solution (e.g., Elton and Gruber, 1995).

By contrast, when (2) holds, the maximum of $f(x)$ occurs at infinite values of x, having value (Maller and Turkington, 2002)

$$+\sqrt{\mu^T\Sigma^{-1}\mu - (i^T\Sigma^{-1}\mu)^2/i^T\Sigma^{-1}i}. \tag{3}$$

The term under the square root sign of (3) is non-negative, and is zero if and only if μ is proportional to i, i.e., if the excess returns of all N assets are equal. Thus we can expect to achieve a positive SR regardless of the value of $i^T\Sigma^{-1}\mu$.

3. Sample Statistics

In practise we will have estimates

$$\widehat{\mu}_n = (\widehat{\mu}_{n1}, \dots, \widehat{\mu}_{nd}),$$

of the mean (excess) returns calculated from a sample of size n, and an estimate $\widehat{\Sigma}_n = (\widehat{\sigma}_{nij})$ of a positive definite matrix. We carry out a Markowitz (1952) optimal allocation of funds among the securities. For our analysis $\widehat{\Sigma}_n$ need not be related to the covariance matrix of the returns, though in practise it usually is. (We keep $d \ge 2$ from now on.)

The sample Sharpe ratio is defined as

$$\widehat{SR}_n = \sup_{x\in C}\left(\frac{x^T\widehat{\mu}_n}{\sqrt{x^T\widehat{\Sigma}_n x}}\right).$$

Note that we maximise the ratio with regard to sign, as advocated, eg. by Sharpe (1994), rather than taking the absolute value or square, as is occasionally done. The statistic $\widehat{SR}_n$ provides one way of summarising the risk/return tradeoff of the optimal portfolio. Comparisons between portfolios can be made by comparing their Sharpe ratios. So it's natural to ask how the precision of estimation of $\widehat{\mu}_n$ and $\widehat{\Sigma}_n$ is transferred to $\widehat{SR}_n$.

It is relatively easy to show that the estimator is consistent for any choice of C if based on consistent estimators of μ and Σ. (This is not quite trivial to prove since the supremum can occur at infinite values of x. But we omit details here.) We might further guess that asymptotic normality of $\widehat{\mu}_n$ and $\widehat{\Sigma}_n$, i.e., assuming

$$\sqrt{n}\left(\widehat{\mu}_n - \mu, \mathrm{vech}(\widehat{\Sigma}_n - \Sigma)\right) \xrightarrow{D} N(0, \zeta),$$

where ζ is a positive definite matrix, will imply asymptotic normality of $\widehat{SR}_n$. This is not the case in general, even for the class C_A, though it is sometimes. Specifically, when $C = C_A = \mathbb{R}^d \cap \left\{i^T x = 1\right\}$, and $\mu \neq 0$, then we can show that, as $n \to \infty$,

$$\sqrt{n}(\widehat{SR}_n - SR) \xrightarrow{D} N(0, \sigma^2_{C_A})$$

for some $\sigma^2_{C_A} > 1$ (depending on μ, Σ, and ζ), where SR is the population Sharpe ratio. When $\mu = 0$ the limit of $\sqrt{n}(\widehat{SR}_n - SR)$ can be explicitly worked out for C_A, and is not normal (in fact it is a non-negative random variable and depends on the unknown Σ). Heuristically, what happens when $\mu = 0$ is that the sample estimate $\widehat{\mu}$ can oscillate around zero, alternately bringing into play the situations in (1) and (2).

Finally, when $C = C_+$, and further assuming that Σ is diagonal, the limit of $\sqrt{n}(\widehat{SR}_n - SR)$ can again be worked out, and again is not normal in all situations. Especially, the case $\mu = 0$ leads to non-normality, but so do some other values of μ, in the C_+ case.

Although not always asymptotically normal, the Sharpe ratio is in the domain of attraction of the normal in the cases we study. Details of these results, together with some practical implications of the analyses, are in Maller (2004).

References

[1] Elton, E.J. and Gruber, M.J. (1995): *Modern Portfolio Theory and Investment Analysis*, 5th Ed., Wiley, New York.

[2] Maller, R.A. (2004): The Large-Sample Distribution of the Sharpe Ratio. Preprint.

[3] Maller, R.A. and Turkington, D.A. (2002): New light on the portfolio allocation problem. *Math. Methods Oper. Res.* **56** (3), 501-511.

[4] Markowitz, H. (1952): Portfolio Selection. *J. Finance* **7**, 77-91.

[5] Markowitz, H. (1991): *Portfolio Selection: Efficient Diversification of Investment.* Blackwell, Cambridge, Mass.

[6] Sharpe, W.F. (1994): The Sharpe Ratio. *The Journal of Portfolio Management* (Fall), 49-58.

The t Copula and Related Copulas

Alexander J. McNeil
(joint work with Stefano Demarta)

The t copula (see for example Embrechts et al. (2001) or Fang and Fang (2002)) can be thought of as representing the dependence structure implicit in a multivariate t distribution. It is a model which has received much recent attention, particularly in the context of modelling multivariate financial return data (for example daily relative or logarithmic price changes on a number stocks). A number of recent papers such as Mashal and Zeevi (2002) and Breymann et al. (2003) have shown that the empirical fit of the t copula is generally superior to that of the so-called Gaussian copula, the dependence structure of the multivariate normal distribution. One reason for this is the ability of the t copula to capture better the phenomenon of dependent extreme values, which is often observed in financial return data.

The objective of this talk is to bring together what is known about the t copula, particularly with regard to its extremal properties, to present some extensions of the t copula, and to describe copulas that are related to the t copula through extreme value theory.

The two new extensions of the t copula are known respectively as the skewed (or asymmetric) t copula and the grouped t copula. Both are constructed by generalising the Gaussian mixture construction of the multivariate t distribution. The skewed t copula is obtained as the copula of a mean-variance mixture of multivariate normals using an inverse gamma mixing distribution, and is a member of the family of generalised hyperbolic copulas. The grouped t copula is the copula of a distribution that is obtained by mixing different subvectors of a Gaussian vector with different inverse-gamma distributed mixing variables, all of which are perfectly positively dependent. Both copulas are interesting for applied work as they suggest ways of incoporating more heterogeneity into the modelling of tail-dependent risks.

The two new copulas arising from extreme value theory are known as the t extreme value (t-EV) copula and the t lower tail limit copula. The former is the limiting copula of componentwise maxima of t distributed random vectors; the latter is the limiting copula of bivariate observations from a t distribution that are conditioned to lie below some joint threshold that is progressively lowered. Both these copulas may be approximated for practical purposes by simpler, better-known copulas, these being the Gumbel and Clayton copulas respectively. They are thus of more theoretical than practical interest.

The finding that the Clayton copula may successfully approximate the t lower tail copula provides some support for the empirical finding by Breymann et al. (2003) that bivariate exchange rate return data are consistent with a t copula as overall model and a Clayton copula for the most extreme negative returns at many different sampling frequencies.

References

[1] Breymann, W., Dias, A. and Embrechts, P. (2003): Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance* **3**, 1-14.

[2] Embrechts, P., McNeil, A. and Straumann, D. (2001): Correlation and dependency in risk management: properties and pitfalls. In: *Risk Management: Value at Risk and Beyond*, M. Dempster and H. Moffatt, eds., Cambridge University Press.

[3] Fang, H. and Fang, K. (2002): The meta-elliptical distributions with given marginals. *J. Multivariate Anal.* **82**, 1-16.

[4] Mashal, R. and Zeevi, A. (2002): Beyond correlation: extreme co–movements between financial assets. Unpublished, Columbia University.

Stable Limits for GARCH Parameter Estimation

Thomas Mikosch
(joint work with Daniel Straumann)

This talk is based on joint work with Daniel Straumann (ETH Zurich); see [9].

We consider a GARCH(p,q) (generalized autoregressive conditionally heteroscedactic process of order (p,q)) given by the equations

$$X_t = \sigma_t \, Z_t \,, \quad X_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j \, X_{t-j}^2 + \sum_{k=1}^{q} \beta_k \, \sigma_{t-k}^2 \,, \quad t \in \mathbb{Z} \,, \tag{1}$$

for non-negative coefficients α_j and β_k. This process is one of the standard models for returns of speculative prices. It is a well-known empirical fact that returns are heavy-tailed. The GARCH model allows for modeling those tails either by heavy tails of the σ- or Z-processes.

Regular variation and stochastic recurrence equations

A theoretical means to describe heavy tails in the univariate and multivariate cases is *regular variaton*: a random vector $\mathbf{X} \in \mathbb{R}^d$ and its distribution are *regularly varying with index* $\alpha \geq 0$ if there exists $\mathbf{\Theta} \in \mathbb{S}^{d-1}$ such that for any $t > 0$, $S \subset \mathbb{S}^{d-1}$ with $P(\mathbf{\Theta} \in \partial S) = 0$,

$$\lim_{x \to \infty} \frac{P\left(|\mathbf{X}| > tx \,, \widetilde{\mathbf{X}} \in S\right)}{P(|\mathbf{X}| > x)} = t^{-\alpha} \, P(\mathbf{\Theta} \in S) \,,$$

where $\widetilde{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$. The limiting distribution $P_{\mathbf{\Theta}}$ is the *spectral measure* of $\mathbf{X}$.

The notion of multivariate regular variation is a very natural one. It is used as necessary and sufficient domain of attraction condition for partial sums of iid random vectors with infinite variance stable weak limits ([12]) and for componentwise maxima of iid random vectors ([11]). Moreover, under mild conditions on the sequence of iid non-negative random vectors $((\mathbf{A}_i \,, \mathbf{B}_i))$, the stationary solution $(\mathbf{X}_t)$ to the stochastic recurrence equation

$$\mathbf{X}_t = \mathbf{A}_t \, \mathbf{X}_{t-1} + \mathbf{B}_t \,, \quad t \in \mathbb{Z} \,. \tag{2}$$

is regularly varying in the sense that

$$P\left((\widetilde{\mathbf{x}}\,,\mathbf{X}) > x\right) \sim c(\widetilde{\mathbf{x}})\,x^{-\alpha}\,,\quad x\to\infty,,\quad \widetilde{\mathbf{x}}\in\mathbb{S}^{d-1}\,, \tag{3}$$

for some $\alpha > 0$ ([7]). It is not difficult to verify that the vector

$$\mathbf{X}_t = \left(\sigma_{t+1}^2,\ldots,\sigma_{t-q+2}^2, X_t^2,\ldots,X_{t-p+2}^2\right)'\,,$$

which is constructed from the GARCH($p.q$) process (1) satisfies (2) and, hence, (3) applies. See [8] for a review on GARCH models, regular variation and stochastic recurrence equations.

Gaussian maximum likelihood estimation with heavy-tailed innovations

Gaussian maximum likelihood for the GARCH parameters α_i and β_j is based on the maximization of the log-likelihood function of a sample $X_1,\ldots,X_n$ (assuming the Z_t's iid standard normal)

$$L_n(\theta) = -\frac{1}{n}\sum_{t=1}^n [\log(\sigma_t^2(\theta)) + \frac{\sigma_t^2(\theta_0)Z_t^2}{\sigma_t^2(\theta)}]\,,$$

with respect to the GARCH parameter θ, where θ_0 is the true parameter of the GARCH model, underlying the observations, and $\widehat{\theta}_n$ is the resulting *Gaussian maximum likelihood estimator*. Taylor expansion of $L_n'(\widehat{\theta}_n)$ at θ_0 yields

$$\widehat{\theta}_n - \theta_0 = -(L_n''(\theta_n))^{-1}\,L_n'(\theta_0)\,,$$

for some θ_n with $|\theta_0-\theta_n| \le |\theta_0-\widehat{\theta}_n|$. By the ergodic theorem, $L_n''(\theta_n)\to\mathbf{B}_0$ a.s. for some deterministic matrix $\mathbf{B}_0$, and therefore weak limit theory for $\widehat{\theta}_n$ reduces to

$$L_n'(\theta_0) = \frac{1}{n}\sum_{t=1}^n \frac{(\sigma_t^2(\theta_0))'}{\sigma_t^2(\theta_0)}\,(Z_t^2-1) = \frac{1}{n}\sum_{t=1}^n \mathbf{G}_t\,Y_t\,.$$

If $EZ_1^4 < \infty$ the CLT for stationary ergodic martingale differences ([3]) gives asymptotic normality for $\widehat{\theta}_n$

$$\sqrt{n}\,(\widehat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}(\mathbf{0}, -E(Z_1^4-1)\,\mathbf{B}_0^{-1})\,. \tag{4}$$

This was proved in [2]. An interesting observation as regards (4) is that the GARCH structure is not essential for the limit theorem (4): as long as $\mathbf{G}_t = (\sigma_t^2(\theta_0))'/\sigma_t^2(\theta_0)$ is stationary ergodic and predictable, and $E|\mathbf{G}_0Y_0|^2 < \infty$ the CLT applies. In the GARCH context it is remarkable, that $\mathbf{G}_0$ has finite moments of all orders (see [2]) and therefore the regular variation of the X_t's (see the previous section) is not essential for the asymptotic theory of $\widehat{\theta}_n$, even if $\mathrm{var}(X_0)=\infty$.

Recently, [6] have extended (4) to the case when $EZ_1^4=\infty$. Assuming that Z_1 is regularly varying with index $\alpha\in(2,4)$, they show that infinite variance stable limits appear in (4). In the paper [9] it is shown that such limits appear for general models $X_t=\sigma_tZ_t$, if (σ_t) is predictable, stationary ergodic, β-mixing with geometric rate, (Z_t) is an iid sequence and regularly varying with index $\alpha\in(2,4)$

and if $E|\mathbf{G}_0|^{(\alpha/2)+\delta} < \infty$ for some $\delta > 0$. The latter conditions are satisfied for the GARCH model.

The results of [9] are based on an analogue to the CLT for stationary ergodic sequences in the case of infinite variance summands $\mathbf{X}_t$ which was proved in [4, 5]. Assuming that $(\mathbf{X}_t)$ satisfies a mild mixing condition (milder than strong mixing) and that its finite-dimensional distributions are regularly varying with index $\kappa \in (0,2)$, $a_n^{-1}\sum_{t=1}^n \mathbf{X}_t$ (suitably centered) weakly converges to a stable limit, where $P(|\mathbf{X}_0| > a_n) \sim n^{-1}$.

In particular, it applies to summands of the form $\mathbf{X}_t = \mathbf{G}_t Y_t$ for vector-valued predictable stationary ergodic $\mathbf{G}_t$ with $E|\mathbf{G}_0|^{\kappa+\delta} < \infty$, some $\delta > 0$, and regularly varying Z_t with index $\alpha = 2\kappa \in (2,4)$. Indeed, then regular variation of the finite-dimensional distributions is conveniently verified. If one has a particular structure such as GARCH, the verification of the β-mixing condition for $\mathbf{G}_t = ((\sigma_t^2(\theta_0))'/\sigma_t^2(\theta_0))$ can be derived from β-mixing for $((\sigma_t^2)', \sigma_t^2)$. In the GARCH case, this condition can be verified by applying a result of [10] on mixing properties of solutions to stochastic recurrence equations (2): then $(\sigma_t^2, (\sigma_t^2)')$ can be embedded in such a stochastic recurrence equation.

The CLT for the GARCH Gaussian maximum likelihood estimator of θ_0 when $EZ_1^4 < \infty$ has $\sqrt{n}$-rates of convergence. This is in contrast to the case when Z_1 is regularly varying with index $\alpha \in (2,4)$, where the rate of convergence is of the order $n^{1-2/\alpha}$. This means that slow rates of convergence and unusually wide confidence bands for the parameter estimators appear.

References

[1] Basrak, B., Davis, R.A. and Mikosch. T. (2002): Regular variation of GARCH processes. *Stoch. Proc. Appl.* **99**, 95-116 .

[2] Berkes, I., Horváth, L. and Kokoszka, P. (2003): GARCH processes: structure and estimation. *Bernoulli* **9**, 201-228.

[3] Billingsley, P. (1968): *Convergence of Probability Measures.* Wiley, New York.

[4] Davis, R.A. and Hsing, T. (1995): Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.* **23**, 879-917.

[5] Davis, R.A. and Mikosch, T. (1998): The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Ann. Statist.* **26**, 2049-2080.

[6] Hall, P.G. and Yao, G. (2003): *Econometrica* **71**, 285-317.

[7] Kesten, H. (1973): Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207-248.

[8] Mikosch, T. (2003): Modeling dependence and tails of financial time series. In: Finkenstädt, B. and Rootén, H. (Eds.) *Extreme Values in Finance, Telecommunications, and the Environment*, pp. 185-286. Chapman and Hall, Boca Raton.

[9] Mikosch, T., Straumann, D. (2003): Stable limits of martingale transforms with application to the estimation of GARCH parameters. Technical report. Available under www.math.ku.dk/~mikosch.

[10] Mokkadem, A. (1990): Propriétés de mélange des processus autorégressifs polynomiaux. *Ann. Inst. H. Poincaré Probab. Statist.* **26**, 219-260.

[11] Resnick, S.I. (1987): *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

[12] Rvačeva, E. L. (1962): On domains of attraction of multi-dimensional distributions. *Select. Transl. Math. Statist. and Probability*, American Mathematical Society, Providence, R.I. **2**, 183-205.

The Effects of Random and Discrete Sampling when Estimating Continuous-Time Diffusions

Per Mykland
(joint work with Yacine Aït-Sahalia)

Diffusion models, and their extensions such as jump-diffusions and Markov models driven by Lévy processes, are essential tools for much of theoretical asset pricing. Estimating these models from discrete time observations has become in recent years an active area of research in econometrics and statistics. Beyond the choice of inference strategy, an important debate in this area concerns the question of what sampling scheme to use, if a choice is available, and in any event what to do with the sampling times. The most straightforward thing to do, in accordance with the usual low-frequency data collection procedures in finance, is to view the sampling as occurring at fixed discrete time intervals, such as a day or a week. In many circumstances, however, this is not realistic. In fact, all transaction-level data are available at irregularly and randomly spaced times.

Not only are the data randomly spaced in time, but whenever a theoretical model is spelled out in continuous time, its estimation necessarily relies on discretely sampled data. By now, there is a good understanding in the literature of the implications of sampling discreteness, and how to design estimation methods that correct for it. The objective in this work is to understand the additional effect that the randomness of the sampling intervals might have when estimating a continuous-time model with discrete data. Specifically, we seek to disentangle the effect of the sampling randomness from the effect of the sampling discreteness, and to compare their relative magnitudes. We also examine the effect of simply ignoring the sampling randomness. We achieve this by comparing the properties of three likelihood-based estimators, which make different use of the observations on the state process and the times at which these observations have been recorded. We design these estimators in such a way that each one of them is subject to a specific subset of the effects we wish to measure. As a result, the differences in their properties allow us to zero in and isolate these different effects.

Our main conclusion is that the loss from not observing, or not using, the sampling intervals, will be at least as great, and often substantially greater, than the loss due to the fact that the data are discrete rather than continuous. While correcting for the latter effect has been the main focus of the literature in recent years, our results suggest however that empirical researchers using randomly spaced data should pay as much attention, if not more, to sampling randomness as they do to sampling discreteness.

The second paper develops tools for analyzing similar problems in the context of non-likelihood inference (estimating or moment equations), and studies specifically the effect of using approximations such as the Euler scheme.

A further contribution of the work is the development of a set of tools that allows these calculations to be performed in closed form.

References

[1] Aït-Sahalia, Y. and Mykland, P.A. (2003): The effects of random and discrete sampling when estimating continuous-time diffusions. *Econometrica* **71**, 483-549.

[2] Aït-Sahalia, Y. and Mykland, P.A. (2004?): Estimators of diffusions with discrete observations: a general theory. Technical report 539, Dept. of Statistics, The University of Chicago (to appear in *Ann. Statist.*).

Tail Behaviour of the Stationary Distribution of a Random Coefficient Autoregressive Model

Serguei Pergamenchtchikov
(joint work with Claudia Klüppelberg)

We consider the following autoregressive process with ARCH errors.

$$x_n = a_1 x_{n-1} + \cdots + a_q x_{n-q} + \sqrt{1+\sigma_1^2 x_{n-1}^2 + \cdots + \sigma_q^2 x_{n-q}^2}\,\varepsilon_n\,, \quad n \in \mathbb{N}\,, \tag{1}$$

where (ε_n) are i.i.d. $\mathcal{N}(0,1)$. We are interested in the existence of a stationary version of the process $(x_n)_{n\in\mathbb{N}}$, represented by a random variable (rv) x_∞ and its properties. We investigate the tail behaviour

$$\mathbf{P}(x_\infty > t) \quad \text{as} \quad t \to \infty\,. \tag{2}$$

This is, in particular, the first step for an investigation of the extremal behaviour of the corresponding stationary process. For $q = 1$ the model (1) was investigated in Borkovec and Klüppelberg [3] by direct analytic methods. For the general case $q > 1$ it is not possible to apply this approach since in this case the model (1) is a non-linear equation with respect to x_n. One can, however, show (see Lemma 2.7 in [12]) that this model is in distribution equivalent to a random coefficient autoregressive process

$$y_n = \alpha_{1n} y_{n-1} + \cdots + \alpha_{qn} y_{n-q} + \xi_n\,, \quad n \in \mathbb{N}\,, \tag{3}$$

where the independent coefficient sequences $(\alpha_{in}\,,\, n \ge 1)$ are i.i.d. and $\alpha_{in} \sim \mathcal{N}(a_i, \sigma_i^2)$ for each $1 \le i \le q$. Moreover the noise variables $(\xi_n)_{n\in\mathbb{N}}$ are an i.i.d. $\mathcal{N}(0,1)$ sequence independent of $(\alpha_{in}\,,\, n \ge 1)_{1\le i\le q}$. Consequently, the problem (2) is equivalent to the investigation of the tail behaviour of a stationary version of the process (3) represented by a random variable y_∞.

To obtain the asymptotic behaviour of the tail of y_∞ we embed $(y_n)_{n\in\mathbb{N}}$ into a multivariate set-up.

Set $Y_n = (y_n, \ldots, y_{n-q+1})'$. Then the multivariate process (Y_n) can be considered in the much wider context of random recurrence equations of the type

$$Y_n = A_n Y_{n-1} + \zeta_n\,, \quad n \in \mathbb{N}\,, \tag{4}$$

where $\zeta_n = (\xi_n, 0, \ldots, 0)'$ and

$$A_n = \begin{pmatrix} \alpha_{1n} & \cdots & \alpha_{qn} \\ I_{q-1} & & 0 \end{pmatrix}, \quad n \in \mathbb{N}\,, \tag{5}$$

where I_{q-1} denotes the identity matrix of order $q-1$.

Such equations play an important role in many applications as e.g. in queueing; see Brandt, Franken and Lisek [4] and in financial time series; see Engle [7]. See also Diaconis and Freedman [5] for an interesting review article with a wealth of examples.

If the Markov process defined in (4) has a stationary distribution and Y has this stationary distribution, then certain results are known on the tail behaviour of Y. In the one-dimensional case ($q = 1$) Goldie [8] has derived the tail behaviour of Y in a very elegant way by a renewal type argument: the tail decreases like a power-law. For the multivariate model for the matrix A_n with positive elements Kesten [9] shows that for each non-zero vector $x \in \mathbb{R}^q$ there exists some $\lambda > 0$ such that $\lim_{t\to\infty} t^\lambda\, \mathbf{P}(x'Y > t) < \infty$.

However, our model (4) does not satisfy the positivity condition on the matrices A_n. Consequently, we derived a new limiting theorem for the model (4) with the matrix of special form (5) in the spirit of Kesten's results. The following is our main result.

Theorem 1. *We assume that the eigenvalues of the matrix* $\mathbf{E}\, A_1 \otimes A_1$ *have moduli less than one and* $a_q^2 + \sigma_q^2 > 0$. *Then the stationary distribution* Y *of the process (4) satisfies*

$$\lim_{t\to\infty} t^\lambda \mathbf{P}(x'Y > t) = h(x)\,, \quad x \in S = \{z \in \mathbb{R}^q \,:\, |z| = 1\}\,.$$

The function $h(\cdot)$ *is strictly positive and continuous on* S *and the parameter* λ *is given as the unique positive solution of*

$$\kappa(\lambda) = 1\,, \tag{6}$$

where for some probability measure ν *on* S

$$\kappa(\lambda) := \lim_{n\to\infty} \left(\mathbf{E}|A_1 \cdots A_n|^\lambda\right)^{1/n} = \int_S \mathbf{E}\,|x'A_1|^\lambda\, \nu(\underset{\cdot}{x})\,,$$

and the solution of (6) satisfies $\lambda > 2$.

References

[1] Basrak, B., Davis, R.D. and Mikosch, T. (2002): Regular variation of GARCH processes. *Stoch. Proc. Appl.* **99**, 95-115.

[2] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989): *Regular Variation.* Revised paperback edition. Cambridge University Press, Cambridge.

[3] Borkovec, M. and Klüppelberg, C. (2001): The tail of the stationary distribution of an autoregressive process with ARCH(1) errors. *Ann. Appl. Probab.* **11**, 1220-1241.

[4] Brandt, A., Franken, P. and Lisek, B. (1990): *Stationary Stochastic Models.* Wiley, Chichester.

[5] Diaconis, P. and Freedman, D. (1999): Iterated random functions. *SIAM Review* **41**, 45-76.

[6] Diebolt, J. and Guegan, D. (1993): Tail behaviour of the stationary density of general nonlinear autoregressive processes of order 1. *J. Appl. Prob.* **30**, 315-329.

[7] Engle, R.F. (1995) *ARCH. Selected Readings.* Oxford University Press, Oxford.

[8] Goldie, C.M. (1991): Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1**, 126-166.

[9] Kesten, H. (1973): Random difference equations and renewal theory for products of random matrixes. *Acta Math.* **131**, 207-248.

[10] Kesten, H. (1974): Renewal theory for functional of a Markov chain with general state space. *Ann. Probab.* **2**, 355-386.

[11] Klüppelberg, C. and Pergamenchtchikov, S. (2003): Renewal theory for functionals of a Markov chain with compact state space. *Ann. Probab.* **31** (4), 2270 - 2300

[12] Klüppelberg, C. and Pergamenchtchikov, S. (2004): The tail of the stationary distribution of a random coefficient AR(q) model. *Ann. Applied Probab.* (to appear). Available under www.ma.tum.de/stat/.

Multivariate Extremes, Max-Stable Processes and Financial Risk

Richard L. Smith

1. Introduction

Extreme value theory has become increasingly applied in mathematical finance, especially in conjunction with "Value at Risk" calculations (Embrechts *et al.* 1997, Finkenstädt and Rootzén, 2003). Univariate extreme value theory is fairly well understood by now, with extensive development of the threshold approach, which is based on the Generalized Pareto distribution (GPD) fitted to exceedances over some high threshold (Davison and Smith 1991, Coles 2001). These methods are applicable to financial time series provided some account is taken of volatility. One approach to that is due to McNeil and Frey (1999), who proposed fitting a GARCH model to financial time series with residuals from an unknown distribution, whose tail was analyzed using threshold methods.

However, there has been relatively less work on dependence in the extremes, by which we mean both temporal dependence in a single time series, and cross-dependence between time series. Multivariate extreme value theory and its generalization, the theory of max-stable processes, are natural candidates to model the joint extremal behaviour of several financial time series. This is the subject of the present paper.

2. Multivariate Extreme Value Theory

Suppose $\mathbf{Y}_i = (Y_{i1}, ..., Y_{iD}),\ i = 1, 2, ...$ is an i.i.d. sequence of D-dimensional random vectors. For each $d \in \{1, ..., D\}$, let $M_{nd} = \max\{Y_{id},\ 1 \le i \le n\}$.

If normalizing constants a_{nd}, b_{nd} and a D-dimensional distribution function G exist such that as $n \to \infty$,

$$\Pr\left\{\frac{M_{nd} - b_{nd}}{a_{nd}} \le x_d,\ 1 \le d \le D\right\} \to G(x_1, ..., x_D)$$

then G is called a *multivariate extreme value distribution*.

There are various representations of multivariate extreme value distributions due to Pickands, de Haan and Resnick, Deheuvels, etc. (Resnick (1987) has a comprehensive account) but these are too general to be directly applicable to statistics. Some authors (e.g. Tawn, Coles) have used parametric subfamilies while others (e.g. de Haan) used nonparametric approaches, but it is not easy to apply any of the existing methods to series in very high dimensions. This motivates an alternative approach.

3. Max-Stable Processes

Max-stable processes are the infinite-dimensional generalization of multivariate extreme value distributions. They are a natural framework within which to study extremal properties of multivariate time series.

Suppose $\{Y_{id},\ i = 0, \pm 1, \pm 2,\ d = 1, ..., D\}$ is a D-dimensional time series with discrete time index i. Without loss of generality, we may assume $\Pr\{Y_{id} \le y\} = e^{-1/y}$ for $0 < y < \infty$ (the unit Fréchet distribution). In practice, this would be achieved only above a given threshold, by first fitting a univariate threshold model to the marginal distributions.

The process is *max-stable* if for any $n \ge 1,\ N \ge 1,\ y_{id} \ge 0$ for $i = 1, ..., n,\ d = 1, ..., D$,

$$\Pr{}^N \left\{Y_{id} \le N y_{id},\ 1 \le i \le n,\ 1 \le d \le D\right\} = \Pr\left\{Y_{id} \le y_{id},\ 1 \le i \le n,\ 1 \le d \le D\right\}.$$

A subclass of max-stable consists of *multivariate maxima of moving maxima* (M4 for short) defined by

$$Y_{id} = \max_{\ell=1}^{\infty} \max_{k=-\infty}^{\infty} a_{\ell,k,d} Z_{\ell,i-k},$$

where $Z_{\ell,i}$ are independent unit Fréchet for all ℓ, i; $a_{\ell,k,d} \ge 0$; and $\sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} a_{\ell,k,d} = 1, \quad d = 1, ..., D$. For this process,

$$\Pr\left\{Y_{id} \le y_{id},\ i = 1, ..., n,\ d = 1, ..., D\right\} = \exp\left(- \sum_{\ell=1}^{\infty} \sum_{m=-\infty}^{\infty} \max_{k=1-m}^{n-m} \max_{d=1}^{D} \frac{a_{\ell,k,d}}{y_{m+k,d}} \right).$$

Smith and Weissman (1996), generalizing Deheuvels (1983), showed that subject to some non-degeneracy conditions, any max-stable process may be approximated arbitrarily closely by an M4 processes.

Statistically, however, these processes are hard to estimate, because of the presence of "signature patterns" of the form

$$Y_{id} = a_{\ell^*, i-m^*, d} Z_{\ell^*, m^*},\ i = 1, ..., n,\ d = 1, ..., D,$$

which arise when a single very large value Z_{ℓ^*,m^*} dominates all its neighbors. If these relations hold, it is possible to derive very precise estimates of the coefficients (Zhang and Smith, 2004a) but this approach is not robust against even tiny deviations from the model. For this reason, it is not a practical approach with real data.

Some alternative estimation strategies include

- **(a):** estimation based on the empirical distribution function (Hall, Peng and Yao, 2002; Zhang and Smith, 2004b);
- **(b):** assuming the observed process is of the form $X_{id} = Y_{id} + \epsilon_{id}$ with Y an M4 process and $\{\epsilon_{id}\}$ random noise; it may then be possible to filter out the noise by Monte Carlo methods. In ongoing PhD research, Francisco Chamú of the University of North Carolina has been exploring this approach;

(c): a more *ad hoc* method in which observed signature patterns are grouped into clusters and the coefficients of the M4 process inferred from the cluster centers (Smith 2003).

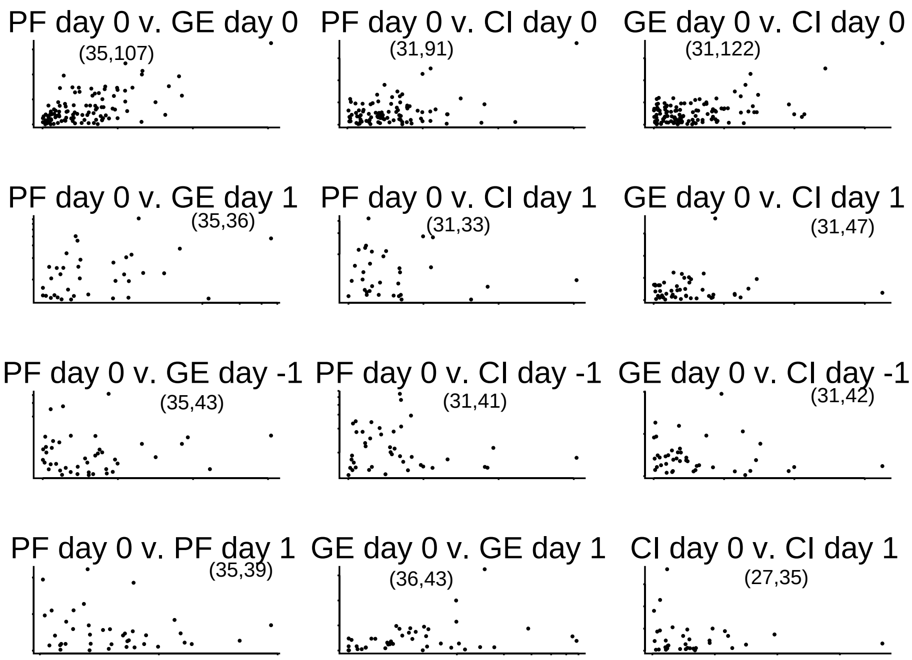


FIGURE 1. Scatterplots of standardized exceedances on unit Fréchet scale. The three stocks are Pfizer (PF), General Electric (GE) and Citibank (CI); plotted are the values on the current day (day 0) versus current day, following day (day 1) and previous day (day –1). The two numbers displayed on each plot are the actual number of joint exceedances (second number), and the expected value of the number of joint exceedances if the two variables were independent (first number). It can be seen that for all three "day 0 v.day 0" plots, there is substantial dependence between the extremes of the two series. For plots of day 0 against day 1 or day –1, however, the evidence for dependence is much less clear-cut.

4. APPLICATION TO FINANCIAL TIME SERIES

We consider 20 years of financial returns data from three stocks (Pfizer, GE and Citibank). For each series, the GARCH(1,1) model is used to estimate volatility, and the process standardized by dividing returns by the estimated volatility. They are then transformed to unit Fréchet margins (above a threshold) by applying univariate extreme value technology. Pairwise correlation plots of the Fréchet standardized exceedances (Fig. 1) show substantial dependence among the three series, which we model using an M4 process. In this application, we estimated coefficients $a_{\ell,k,d}$ which are assumed non-zero for $\ell = 1, 2, ..., 25$ and $k = -2, -1, 0, 1, 2$.

Finally, a cross-validation exercise shows that the model provides good estimation of some simple functionals of the joint extremes. One possible functional is the probability that, over a window of 10 trading days (a typical time window in Value at Risk calculations), at least one of the three series crosses a given target value. Fig. 2 shows both an empirical crossing rate and a cross-validated model-generated crossing rate (Smith, 2003) as the target value increases; the agreement is excellent until the very highest target values, where neither method can be expected to give accurate results.

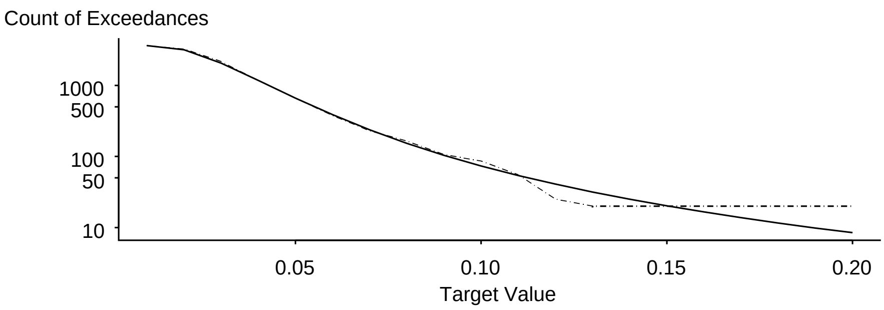


Figure 2. Plot of estimated number of expected exceedances of a given target value for the maximumof the daily returns over all 3 stocks over a 10-day window. Solid curve: cross-validated model-based estimate. Dashed curve: empirical value.

5. Acknowledgements

Portions of this work are collaborative with Ishay Weissman (Technion), Zhengjun Zhang (Washington University, St.Louis) and Francisco Chamú (University of North Carolina). Financial support is acknowledged from NSF grants DMS-0084375 and DMS-9971980.

References

[1] Coles, S.G. (2001): *An Introduction to Statistical Modeling of Extreme Values.* Springer-Verlag, New York.

[2] Davison, A.C. and Smith, R.L. (1990): Models for exceedances over high thresholds (with discussion). *J.R. Statist. Soc. Ser. B* **52**, 393-442.

[3] Deheuvels, P. (1983): Point processes and multivariate extreme values. *J. Multivar. Anal.* **13**, 257-272.
[4] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997): *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, New York.
[5] Finkenstädt, B. and H. Rootzén, H. (eds.) (2003): *Extreme Values in Finance, Telecommunications and the Environment*. Chapman and Hall/CRC Press, London.
[6] Hall, P., Peng, L. and Yao, Q. (2002): Moving-maximum models for extrema of time series. *J. Stat. Plann. Inference* **103**, 51-63.
[7] McNeil, A. and Frey, R. (1999): Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *J. Empirical Finance* **7**, 271-300.
[8] Resnick, S. (1987): *Extreme Values, Point Processes and Regular Variation*. Springer Verlag, New York.
[9] Smith, R.L. (2003): Statistics of extremes, with applications in environment, insurance and finance. Chapter 1 of Finkenstädt and Rootzén, 1-78.
[10] Smith, R.L. and Weissman, I. (1996): Characterization and estimation of the multivariate extremal index. http://www.stat.unc.edu/postscript/rs/extremal.pdf (under revision.)
[11] Zhang, Z. and Smith, R.L. (2004a): The behaviour of multivariate maxima of moving maxima processes. Tentatively accepted for *J. Appl. Prob.*
[12] Zhang, Z. and Smith, R.L. (2004b): On the estimation and application of max-stable processes. Submitted for publication.

A Flexible Class of Stochastic Volatility Models of the Diffusion-Type

Michael Sørensen

Stochastic volatility models of the type

$$dX_t = (\alpha + \beta V_t)dt + \sqrt{V_t}dW_t,$$

where V_t is a suitable positive stochastic process, are widely used in finance to model the logarithm of the price of an asset. Several possible specifications of the process V have been proposed. Barndorff-Nielsen and Shephard (2001) proposed to model V as an Ornstein-Uhlenbeck process driven by a Lévy process or a sum of such processes, which is a very flexible class of models. Here we present a class of models with a similar flexibility where the volatility process is a sum of mean-reverting processes driven by Wiener processes. The results given here are based on the paper Bibby, Skovgaard and Sørensen (2003).

Let f be a given continuous, bounded, and strictly positive probability density on $(0, \infty)$ that is zero when $x \leq 0$ and has finite variance. Define a function v by

$$v(x) = \frac{2\theta \int_l^x (\mu - y) f(y) dy}{f(x)}, \quad x > 0,$$

where μ denotes the expectation of f. It is not difficult to see that $v(x) > 0$ for $x > 0$. The stochastic differential equation

$$dV_t = -\theta(V_t - \mu)dt + \sqrt{v(V_t)}dW_t, \quad t \geq 0,$$

where W denotes a standard Wiener process, has a unique weak solution that is ergodic with invariant density f. If $V_0 \sim f$, V is stationary, and its autocorrelation function is given by

$$\operatorname{Corr}(V_{s+t}, V_s) = e^{-\theta t}, \quad s, t \geq 0.$$

The diffusion V is the only ergodic mean-reverting diffusion with invariant density f.

The function v can be found explicitly for a number of standard distributions on $(0, \infty)$. For the gamma-distribution with shape-parameter α and scale-parameter β

$$v(x) = 2\theta\beta x.$$

In this case the volatility model is the square root process, and the corresponding volatility model is the Heston (1993) model. For the inverse Gaussian distribution with density

$$f(x) = \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} x^{-\frac{3}{2}} \exp\left(-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right), \quad x > 0,$$

we find that

$$v(x) = 4\theta\mu\sqrt{\frac{2\pi}{\lambda}} e^{\lambda/\mu} x^{3/2} \exp\left(\frac{\lambda}{2\mu^2}x + \frac{\lambda}{2}x^{-1}\right) \Phi\left(-\sqrt{\frac{\lambda}{x}} - \sqrt{\frac{\lambda x}{\mu^2}}\right),$$

where Φ is the standard normal distribution function.

Usually, the correlation function $e^{-\theta t}$ is too simple to fit the autocorrelation of the volatility observed in financial time series. Therefore the following construction is useful. Let f be a strictly positive, infinitely divisible probability density on $(0, \infty)$ that is zero when $x \leq 0$, and let $C(t)$ denote the characteristic function of f. Suppose the positive real numbers φ_i, $i = 1, \ldots, m$, satisfy that $\varphi_1 + \cdots + \varphi_m = 1$. Then the functions $C(t)^{\varphi_i}$, $i = 1, \ldots, m$, are characteristic functions too. Assume that the corresponding density functions f_i, $i = 1, \ldots, m$, satisfy the conditions imposed on f earlier, and define

$$v_i(x) = \frac{2\theta_i}{f_i(x)} \int_0^x (\varphi_i\mu - y) f_i(y) dy.$$

Then the process

$$V_t = V_t^{(1)} + \cdots + V_t^{(m)},$$

where

$$dV_t^{(i)} = -\theta_i \left(V_t^{(i)} - \varphi_i\mu\right) dt + \sqrt{v_i\left(V_t^{(i)}\right)} dB_t^{(i)},$$

with $B^{(1)}, \ldots, B^{(m)}$ denoting independent standard Wiener processes, has marginal density f, provided that $V_0^{(i)} \sim f_i$, $i = 1, \ldots, m$. The autocorrelation function of V is given by

$$\operatorname{Corr}(V_{s+t}, V_s) = \varphi_1 \exp(-\theta_1 u) + \cdots + \varphi_m \exp(-\theta_m u).$$

For the gamma-distribution with shape-parameter α and scale-parameter β

$$v_i(x) = 2\beta\theta_i x,$$

and for the inverse Gaussian distribution

$$v_i(x) = 4\theta_i\mu\sqrt{\frac{2\pi}{\lambda}}e^{\varphi_i\lambda/\mu}x^{3/2}\exp\left(\frac{\lambda}{2\mu^2}x + \frac{\varphi_i^2\lambda}{2}x^{-1}\right)\Phi\left(-\varphi_i\sqrt{\frac{\lambda}{x}} - \sqrt{\frac{\lambda x}{\mu^2}}\right).$$

For distributions where the f_i cannot be found explicitly, an approximation to v_i can be found in Bibby, Skovgaard and Sørensen (2003).

References

[1] Barndorff-Nielsen, O.E. and Shephard, N. (2001): Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial econometrics (with discussion). *J. R. Stat. Soc. Ser. B* **63**, 167-241.

[2] Bibby, B.M., Skovgaard, I.M. and Sørensen, M. (2003): Diffusion-type models with given marginals and autocorrelation function. Preprint No. 5, Department of Applied Mathematics and Statistics, University of Copenhagen.

[3] Heston, S.L. (1993): A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* **6**, 327-343.

Adaptive Estimation for a Varying Coefficient GARCH Model

Vladimir Spokoiny
(joint work with Jörg Polzehl)

Financial time series are often modelled by parametric ARCH or GARCH models under the assumption of stationarity. This approach is not flexible enough to incorporate models with structural breaks and time varying parameters. This paper presents a unified approach for modeling non (local) stationary time series including change point and smooth transition models. The procedure is based on the *Adaptive Weights* idea from Polzehl and Spokoiny (2000, 2002, 2003). The paper discusses important theoretical properties of the method and illustrates its numerical performance by mean of simulated examples and applications to real data.

References

[1] Polzehl, J. and Spokoiny, V. (2000): Adaptive weights smoothing with applications to image segmentation. *J. R. Stat. Soc. Ser. B* **62**, 335-354.

[2] Polzehl, J. and Spokoiny, V. (2002): Local likelihood modeling by adaptive weights smoothing. Preprint 787. WIAS 2002. Available under http://www.wias-berlin.de/publications/preprints/787.

[3] Polzehl, J. and Spokoiny, V. (2003): Varying coefficient regression modeling by adaptive weights smoothing. Preprint 818, WIAS 2003.

Is GARCH(1,1) as good a model as the Nobel prize accolades would imply

Cătălin Stărică

1. Abstract

This paper investigates the relevance of the stationary, conditional, parametric ARCH modeling paradigm as embodied by the GARCH(1,1) process to describing and forecasting the dynamics of returns of the Standard & Poors 500 (S&P 500) stock market index.

A detailed analysis of the series of S&P 500 returns featured in Section 3.2 of the Advanced Information note on the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel reveals that during the period under discussion, there were no (statistically significant) differences between GARCH(1,1) modeling and a simple non-stationary, non-parametric regression approach to next-day volatility forecasting.

A second finding is that the GARCH(1,1) model severely over-estimated the unconditional variance of returns during the period under study. For example, the annualized implied GARCH(1,1) unconditional standard deviation of the sample is 35% while the sample standard deviation estimate is a mere 19%. Over-estimation of the unconditional variance leads to poor volatility forecasts during the period under discussion with the MSE of GARCH(1,1) 1-year ahead volatility more than 4 times bigger than the MSE of a forecast based on historical volatility.

We test and reject the hypothesis that a GARCH(1,1) process is the true data generating process of the longer sample of returns of the S&P 500 stock market index between March 4, 1957 and October 9, 2003. We investigate then the alternative use of the GARCH(1,1) process as a local, stationary approximation of the data and find that the GARCH(1,1) model fails during significantly long periods to provide a good local description to the time series of returns on the S&P 500 and Dow Jones Industrial Average indexes.

Since the estimated coefficients of the GARCH model change significantly through time, it is not clear how the GARCH(1,1) model can be used for volatility forecasting over longer horizons. A comparison between the GARCH(1,1) volatility forecasts and a simple approach based on historical volatility questions the relevance of the GARCH(1,1) dynamics for longer horizon volatility forecasting for both the S&P 500 and Dow Jones Industrial Average indexes.

2. Figures

Figure 1 displays the estimated $\alpha_1 + \beta_1$ under the assumption of non-stationary data. The Garch(1,1) model has been initially estimated on the first 2000 observations of the sample corresponding roughly to the period 1957-1964, then re-estimated every 50 observations on a sample containing 2000 past observations.

The graph shows that the IGARCH effect significantly[1] affects the GARCH(1,1) models (estimated on a sample that ends) during the period 1997-2003[2]. This fact at its turn, is likely to cause the explosion of the estimated unconditional variance of the GARCH(1,1) processes fitted on samples that end during this period.

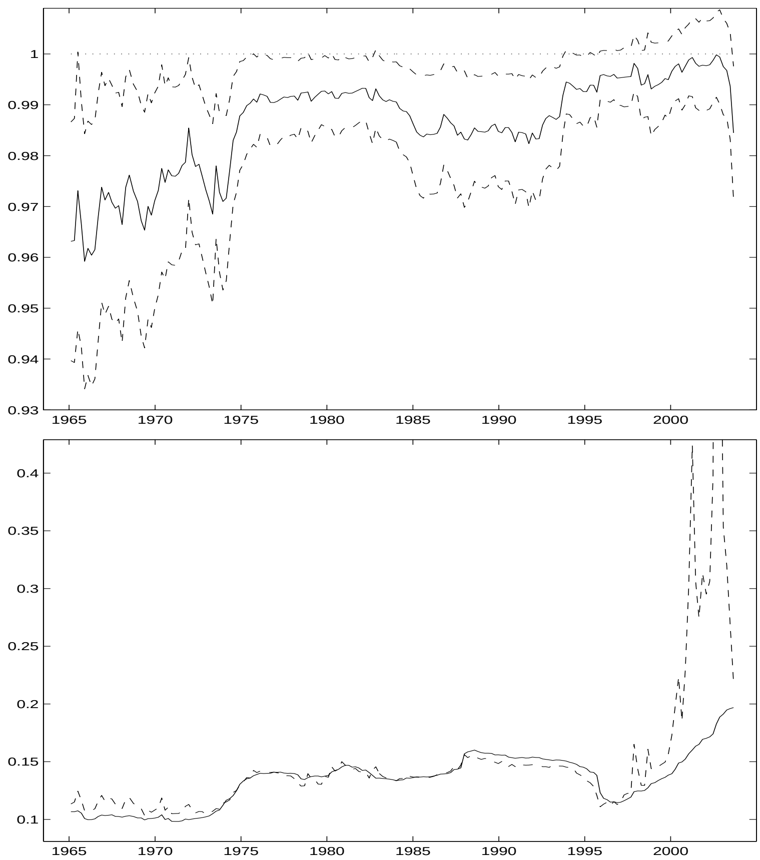


FIGURE 1. Top: Estimated $\alpha_1 + \beta_1$. Bottom: Estimated GARCH (1,1) sd (dotted line) together with sample sd (both estimates are annualized) (full line) for the S&P 500 log-returns. The time mark corresponds to the end of the sub-sample that yields the two standard deviation estimates. While most of the time the two curves in the bottom graph are remarkably close to each other, the GARCH(1,1) variance seems to explode towards the end of the sample.

[1]The point estimate is close to 1 and, more importantly, 1 belong to the 95% confidence interval.

[2]During the interval 1994-1996, the value 1 is the upper bound of the confidence interval.

To see that indeed this is the case, let us take look at the bottom graph of the same Figure 1 where the GARCH(1,1) unconditional sd (broken line) and the corresponding sample sd (full line) are displayed. The GARCH(1,1) unconditional sd is obtained from the values of the parameters estimated on a window of size 2000 moving through the data. The graph shows a good agreement between the two estimates at all times except during the period when the IGARCH effect becomes strongly statistically significant, i.e. samples that end in the interval 1997-2003[3],[4].

The bottom graph in Figure 1 show that the GARCH(1,1) model fails to provide a local stationary approximation to the time series of returns on the S&P 500 during significantly long periods.

An explanation for the strong IGARCH effect in the second half of the 90's can be the sharp change in the unconditional variance (see Mikosch and Starica [1]). There it is proved, both theoretically and empirically, that sharp changes in the unconditional variance can cause the IGARCH effect. Figure 2 displays non-parametric estimates of the unconditional sd together with the 95% confidence intervals[5] for the S&P 500 returns (top) and the Dow Jones industrial index returns (bottom). The two graphs show a pronounced increase of the volatility from around 5% in 1993-1994 to three times as much (around 15%) in the period 2000-2003.

References

[1] Mikosch, T. and Stărică, C. (2004): Non-stationarities in financial time series, the long-range dependence and IGARCH effects. *Rev. Econom. Statist.* (to appear). Available under www.math.ku.dk/∼mikosch.

[2] Mikosch, T. and Stărică, C. (2003) Stock market risk-return inference. An unconditional non-parametric approach. Technical Report. Available under www.math.ku.dk/∼mikosch.

Pricing of Contingent Claims When Prices Are Perturbed: An Elementary Example for Discussion

J. Michael Steele

The basic aim of this talk was to suggest consideration of a class of models that one may view as perturbations of another (unobserved) price processes which is either well-understood or which may be blessed with some special theoretical

[3]The analysis was also performed with smaller sample sizes of 1500, 1250 and 1000. As expected, the confidence intervals in Figures 1 get wider and hence less meaningful. However, for every sample sized mentioned, there is always a period between 1997 and 2003 where the unconditional variance of the estimated model explodes. Estimation based on samples smaller than 1000 observations is infeasible as it produces extremely unstable coefficients and renders problematic the use of any asymptotic result.

[4]Contrast this finding with the statement on page 16 of the Advanced Information note: "Condition $\alpha_1+\beta_1 < 1$ is necessary and sufficient for the first-order GARCH process to be weakly stationary, and the estimated model (on the short S&P 500, n.n.) satisfies this condition."

[5]The method used to obtain the estimates is that of kernel smoothing in the framework of non-parametric regression with non-random equi-distant design points. For more details on the performance of this method on financial data see Mikosch and Starica [2].

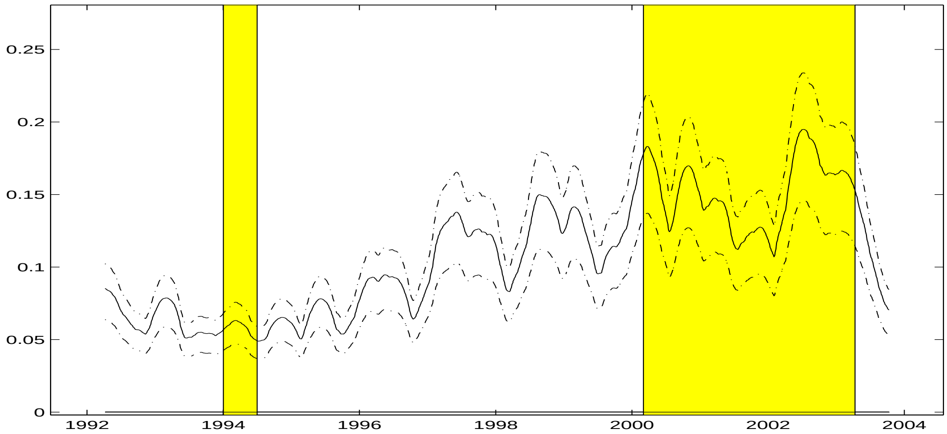


FIGURE 2. Estimated unconditional standard deviation (annualized) with 95% confidence intervals for the S&P 500 returns. The shaded areas correspond to bear market periods.

appeal. This discussion is part of a larger program which hopes to explain more fully the costs and benefits of relying on misspecified models.

The example used in the talk was simply the Black-Scholes model where the observed log prices are perturbed by a mean-zero mean-reverting process, and, for specificity, we took the perturbing process to be an independent Ornstein-Uhlenbeck process. Formally, we considered processes S_t and O_t which satisfy

$$dS_t = \mu dt + \sigma dW_t \quad \text{and} \quad dO_t = -\alpha O_t dt + \epsilon d\widetilde{W}_t$$

where the process $(W_t, \widetilde{W}_t)$ is an uncorrelated Brownian motion in $\mathbb{R}^2$, and we then considered a price process $\{P_t\}$ which is specified by setting

$$P_t = P_0 \exp(Y_t) \quad \text{and} \quad Y_t = S_t + O_t. \tag{1}$$

One reason to consider this model is that it contains as special cases both the Black-Scholes model and the model of Lo and Wang (1995). Like the Lo and Wang (1995) model, $\{P_t\}$ exhibits aspects of predictability, but here it also captures additional elements of economic reality. Specifically, we view $P_t^T \equiv \exp S_t$ as a "true" (but unobserved) price process, and we posit that market forces will drive the observed price P_t back to P_t^T after any random deviations from P_t^T. The model offers a practical compromise between a theoretically appealing model, and one which manifests some modest predictability.

On interesting feature of $\{P_t\}$ is that it is not a Markov process, so *a priori* one might not expect that the PDE methods for pricing contingent claims would apply. Nevertheless, in the case of European call options, easy calculations and ancient recipes quickly bring one to an almost exact replicate of the Black-Scholes PDE — only the volatility parameter is changed. Moreover, this heuristic derivation turns out to be theoretically justified; risk-neutral pricing theory leads one to the same valuation formula.

A further instructive feature of the price process $\{P_t\}$ is its relation to the notion of viability which Bick (1990) introduced to addresses the consistency of a price process with a certain economic equilibrium. He and Leland (1993) later developed a PDE based criterion for viability, and, although it is not strictly applicable here, one can check that the process $\{P_t\}$ does not pass the He-Leland test (where, with eyes closed, we pretend for a moment that $\{P_t\}$ is Markovian!). It remains to be seen if $\{P_t\}$ is viable in the more general framework of Decamps and Lazrak (2000), this also seems doubtful. Nevertheless, the practical motivation underlying consideration of the process $\{P_t\}$ remains in tact; it is, after all, a perturbation of a process $\{P_t^T\}$ that passes anyone's test of viability.

References

[1] Bick, A. (1990): On viable diffusion price processes of the market portfolio. *J. Finance* **45**, 673-689.

[2] Decamps, J.P. and Lazrak, A. (2000): A martingale characterization of equilibrium asset prices processes. *Econom. Theory* **15**, 207-213.

[3] He, H. and Leland, H. (1993): On equilibrium asset price processes. *Review of Financial Studies* **6**, 593-617.

[4] Lo, A.W. and Wang, J. (1995): implementing option pricing models when asset returns are predictable. *J. Finance* **50**, 87-129.

Quasi-Maximum Likelihood Estimation and Conditional Heteroskedastic Time Series

Daniel Straumann

By exploiting the techniques of stochastic recurrence equations, we develop a general and unifying limit theory for the maximum likelihood estimator (MLE) and quasi maximum likelihood estimator (QMLE) in a certain parametric class of conditionally heteroscedastic processes, which contains widely used financial time series models: (asymmetric) GARCH(1,1) and EGARCH. Our approach generalizes and clarifies work of Lumsdaine (1996) and Berkes et al. (2003). We furthermore discuss the issue of misspecification in the MLE and the behaviour of the QMLE in the presence of a heavy-tailed noise distribution. This complements work by Newey and Steigerwald (1997) and Hall and Yao (2003).

A Multinomial Approximation of American Option Prices in a Lévy Process Model

Alex Szimayer

(joint work with Ross A. Maller and David H. Soloman)

This paper examines the pricing of American options in models where the stock price follows an exponential Lévy process. We propose a multinomial model approximating the stock price process which can be viewed as a generalisation of the binomial model of Cox et al. (1979) adapted from Brownian motion to the broader class of Lévy processes. Under mild conditions, it is proved that American option prices obtained under the multinomial model converge to the corresponding prices under the continuous time Lévy process model. Further, explicit schemes are given for the jump diffusion model, the variance gamma model.

The Distribution of the LR Test for a Nonlinear Latent Variable Model of Equity Returns

Mark Van De Vyver

(joint work with Ross A. Maller)

1. Abstract

This paper is devoted to deriving, under quite general conditions, the distribution of a likelihood ratio statistic for testing whether several versions of a generalized autoregressive conditional heteroscedasticity (GARCH) model are superior to a general random walk model, in depicting the true (unknown) data generating process for the natural log of an equity price or their continuously compounded returns. This is the statistic which the one sided LM test statistic approximates, and provides a first check as to whether GARCH effects are in fact present in the data. The application of these results is illustrated using equity market data (contained in the full paper).

2. Introduction

This paper extends the results of [7] and provides new results in the subject of latent variable model specification testing in the discrete time, continuous state setting. Specifically, we consider models that naturally arise in the context of financial modelling, and derive the distribution of the deviance statistic (negative two times the quasi log–likelihood ratio), which is of use in testing for the reduction of the alternative model to a more parsimonious null model. The deviance statistic is that which the more commonly used Lagrange Multiplier (LM) statistic approximates ([5]). Initially the alternative model is specified quite generally, and the parametrization we consider includes the nonlinear GARCH model. The NGARCH-M model we consider is suitable for option pricing, and belongs to a class for which [9] have developed an accurate and parsimonious option pricing algorithm, capable of pricing American and exotic options. While less general, the null model is of special interest. This is a random walk with innovations that are independently and identically distributed. When the innovations have the normal distribution this model converges to the geometric Brownian motion model that lies behind the celebrated Black-Scholes and Merton (BSM) option pricing model ([1] and [8]). The following is the model of an asset price series, which incorporates volatility clustering as well as an asymmetric correlation between returns and volatility innovations:

$$\begin{aligned} X_i &= \phi X_{i-1} + \left(\mu - \tfrac{1}{2}\sigma_i^2 + \lambda\sigma_i\right) + \varepsilon_i\sigma_i \\ \sigma_i &= \sqrt{\omega + \alpha\sigma_{i-1}^2\left(\varepsilon_{i-1} - c\right)^2 + \beta\sigma_{i-1}^2}. \end{aligned} \tag{1}$$

Here X_i is the natural logarithm of the stock price, μ is a drift parameter that frequently is interpreted as the risk free rate; ϕ is an autoregression coefficient (and we are particularly interested in testing the hypothesis that ϕ is unity); λ is another drift term interpreted as the market price of risk; ω is the instantaneous variance of a Gaussian random walk when there are in fact no ARCH or GARCH effects present; α and β are, respectively, the ARCH and GARCH terms that feedback the effects of past observations into the variance equation; c is a generic parameter in the variance equation which permits an asymmetric correlation between the returns and volatility process (see [6] and [4]). We refer to the first two equations as the price (or, when $\phi = 1$, return) equation and volatility equation. We wish to test two null hypotheses of interest, against different alternatives. Both null models will have a common feature $\phi = 1$, and $\alpha = \beta = 0$. Then $\sigma_i = \sigma = \omega$ does not depend on i and we can write (1) as:

$$X_i = X_{i-1} + \mu - \frac{1}{2}\omega + \lambda\sqrt{\omega} + \varepsilon_i\sqrt{\omega}, \quad i = 1, 2, \ldots, n \tag{2}$$

Notice that the parameter c is not present under the null (the term containing c disappears from (1) when $\alpha = 0$), and two other parameters, μ and λ, combine into a single drift parameter and thus cannot be uniquely determined under the null. To reflect this we introduce a drift parameter, ψ, in a simple reparameterization, as $\psi = \mu + \lambda\sqrt{\omega} - \frac{1}{2}\omega$. Applying this to (1) introduces an additional parameter that disappears under the null hypothesis. We handle these parameters using the methods of [2] and [3].

3. Results

3.1. **Testing for GARCH effects alone.**

Theorem 1. *Assume X_i satisfies (1) and that $\phi = 1$, for i.i.d ε_i with expectation 0, variance 1 and finite third and fourth moments, μ_3 and μ_4. Suppose that the null hypothesis holds; $\phi = 1$, and $\alpha = \beta = 0$. When evaluating the null and alternate models, the deviance statistic, d_n, has asymptotic distribution:*

$$d_n^{(2)}(\tau) \xrightarrow{D} kN^2(0,1)I\left(N \geqslant 0\right),$$

where N is standard normal and k is

$$k = 1 + \frac{\mu_4 + 2\mu_3\big(\lambda_0 - \sqrt{\omega_0}\big) - 3}{2 + \big(\lambda_0 - \sqrt{\omega_0}\big)^2}. \tag{3}$$

$I\big(Z \geqslant 0\big)$ is 1 if $Z \geqslant 0$ and 0 otherwise.

3.2. **Testing for GARCH effects and a unit root.**

Theorem 2. *Assume X_i satisfies (1) for i.i.d ε_i with expectation 0, variance 1 and finite third and fourth moments, μ_3 and μ_4. Suppose that the null hypothesis holds; $\phi = 1$, and $\alpha = \beta = 0$. When evaluating the null and alternate models, the deviance statistic, d_n, has asymptotic distribution:*

$$d_n^{(1)}(\tau) \xrightarrow{D} N_1^2 + N_2^2 I\big(N_2 \geqslant 0\big).$$

where N_i are normal random variables with some variance-covariance given in the full paper.

4. Empirical Application

For each company in the S&P 500 index as at 8/8/2003, we select those which have at least 1000 observations, leaving 481 firms. Using individual company returns over this period we fit (1) and (2), and calculate the robust deviance as set out in Theorem 1. The first box plot summarizes the lower range of values of the robust deviances calculated using the 95% CI of the moments of the QML estimated resdiuals. The second and third box plots summarize the range of values of the robust deviance, and it's upper range of values. The last box plot shows the two sided lagrange multiplier test statistic's empirical values. In each plot the box spans the 25%-75% quantiles, the whiskers cover 3/2 of the interquantile range from the edges of the box, individual points represent outlying observations and the dashed line indicates the median value. The two horizontal lines indicate the one-sided chi-square critical values, with one degree of freedom, at the 5% and 1% levels of significance. The upper and lower robust deviance values reflect the range of values the deviance may take, using the 95% CI surrounding the null model residuals' moments, calculated using QML estimated residuals.

References

[1] Black, F. and Scholes, M. (1973): The pricing of options and corporate liabilities. *Journal of Political Economy* **81** (3), 637-54.

[2] Davies, R.B. (1977): Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* **64** (2), 247-254.

[3] Davies, R.B. (1987): Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* **74** (1), 33-43.

[4] Duan, J.-C. (1995): The GARCH option pricing model. *Math. Finance* **5**, 13-32.

[5] Engle, R. F.(1983): Estimates of the variance of U.S. inflation based upon the ARCH model. *Journal of Money, Credit and Banking* **15** (3), 286-301.

[6] Engle, R.F. and Ng, V. K. (1993): Measuring and testing the impact of news on volatility. *J. Finance* **48** (5), 1749-78.

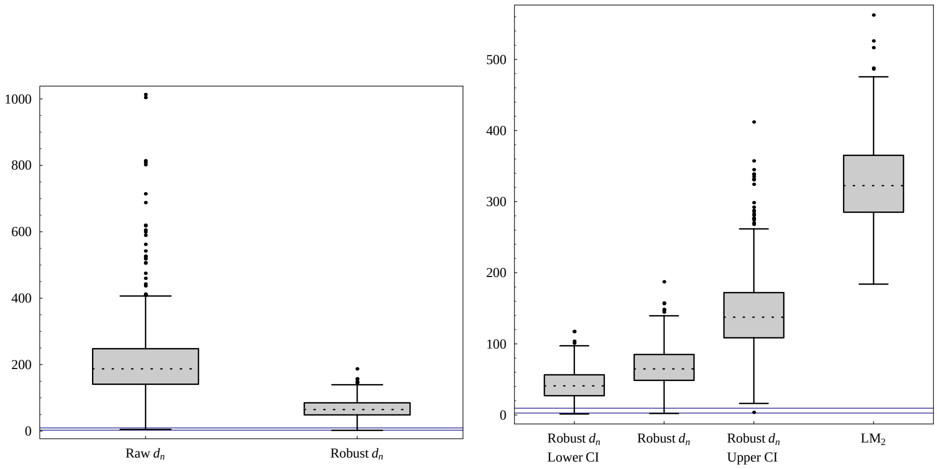


FIGURE 1. GARCH vs. random walk hypothesis raw and robust deviances of 481 companies in the S&P 500 index: 18/8/1999 to 8/8/2003

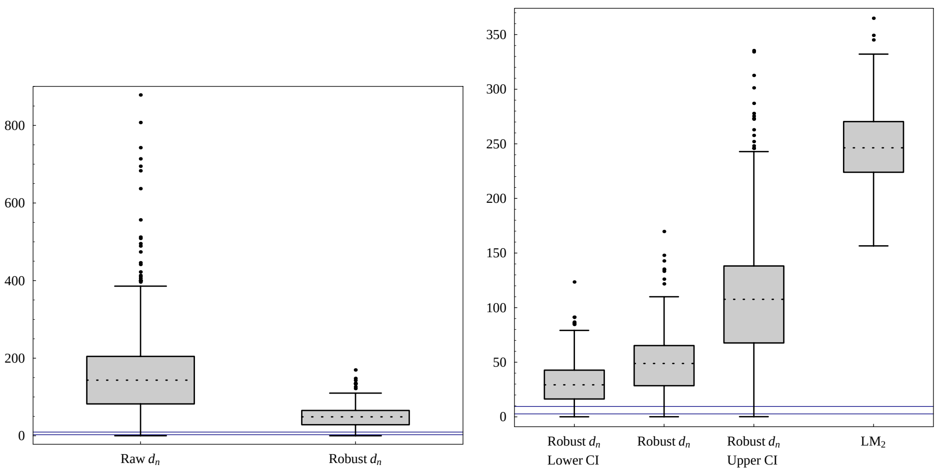


FIGURE 2. ARCH vs. random walk hypothesis raw and robust deviances of 481 companies in the S&P 500 index: 18/8/1999 to 8/8/2003

[7] Klüppelberg, C., Maller, R., Van De Vyver, M. and Wee, D. (2002): Testing for Reduction to Random Walk in Autoregressive Conditional Heteroscedasticity Models. *Econom. J.* **5** (2), 387-416.

[8] Merton, R.C. (1973): Theory of rational option pricing. *The Bell Journal of Economics* **4** (1), 141-83.

[9] Ritchken, P. and Trevor, R. (1999): pricing options under generalized GARCH and stochastic volatility processes. *J. Finance* **54**, 377-402.

Option Pricing and Statistics Inference for GARCH Models and Diffusions

Yazhen Wang

Stock market modeling has two types of approaches. One is continuous-time modeling that assumes a stock price to change with time continuously and obey a continuous-time stochastic process. Historically continuous-time models based on stochastic differential equations have been developed in financial economics, and modern finance theory is much based on the continuous-time modeling. However, in reality all data are recorded only at discrete intervals. Unknown parameters in the continuous-time models need to be estimated and tested from the observed discrete-time data. Due to the difficulty in statistical inference for the continuous time model based on the discrete data, the validity of the continuous-time modeling is not straightforward to check. Another approach is discrete-time modeling of available discrete data. Successful discrete-time models are the autoregressive conditionally heteroscedastic (ARCH) models. These discrete-time models often provide parsimonious representations for the observed discrete-time data, and their statistical inference is relatively much easier. The weak convergence of the discrete-time ARCH model to continuous-time diffusion established first by D. Nelson in early 1990 has generated a general belief that the ARCH model and diffusions are more or less equivalent.

This talk presents asymptotic equivalence of the Garch, discrete stochastic volatility (SV), and diffusion models with respect to option pricing, implied volatility, and statistical inferences based on option data (or implied volatility). As discrete observation intervals shrink to zero, the GARCH and SV models weakly converge to a bivariate diffusion. First we prove that the GARCH option price converges to diffusion price at the speed near to the square root of the observation interval length. Second we show that under the three models, the prices of a European option and their corresponding implied volatilites are equal up to the order near to the square root of the observation interval length, and asymptotically option based statistical inferences under the three models are statistically equivalent. This shows that asymptotically the three models are equivalent in all aspects regarding to option pricing, implied volatility and statistical inference for option data. It presents a sharp contrast with nonequivalence of the GARCH and its diffusion limit regarding to statistical inferences for historical time series price data.

Valuation of American Options via Basis Functions
Samuel Po-Shing Wong
(joint work with Tze Leung Lai)

Using the methodology of pricing and hedging American options proposed by AitSahlia and Lai (2001), we apply the idea of neuro-dynamic programming to develop

(1) nonparametric pricing formulas for actively traded American options, and
(2) simulation-based optimization strategies for complex over-the-counter options, whose optimal stopping problems are prohibitively difficult to solve numerically by standard backward induction algorithms because of the curse of dimensionality.

An important issue in this approach is the choice of basis functions, for which some guidelines and their underlying theory are provided.

This paper is going to be published by *IEEE Transactions in Automatic Control* in 2004.

References

[1] AitSahlia, F. and Lai, T.L. (2001): Exercise boundaries and efficient approximations to American option prices and hedge parameters. *J. Computational Finance* **4**, 85-103.

Approximating Volatilities by Asymmetric Power GARCH Functions
Qiwei Yao
(joint work with Jeremy Penzer and Mingjin Wang)

Let $\{X_t\}$ be a strictly stationary process defined by the volatility model

$$X_t = \sigma_t \varepsilon_t, \tag{1}$$

where $\{\varepsilon_t\}$ is a sequence of independent random variables with mean 0, $\sigma_t \geq 0$ is $\mathcal{F}_{t-1}$-measurable, and $\mathcal{F}_{t-1}$ is the σ-algebra generated by $\{X_{t-k}, k \geq 1\}$. Furthermore, ε_t is independent of $\mathcal{F}_{t-1}$. The conventional ARCH/GARCH formulation assumes that the conditional standard deviation σ_t is of the form

$$\begin{aligned} \sigma_t^2 &= \operatorname{Var}(X_t|\mathcal{F}_{t-1}) = E(X_t^2|X_{t-1}, X_{t-2}, \cdots) \\ &= c + \sum_{i=1}^{p} b_i X_{t-i}^2 + \sum_{j=1}^{q} a_j \sigma_{t-j}^2. \end{aligned} \tag{2}$$

where $c > 0$ and b_i, a_j are non-negative. The above model also implies $\operatorname{Var}(\varepsilon_t) = 1$. Under the condition $\sum_j a_j < 1$, (2) admits the representation

$$\sigma_t^2 = E(X_t^2|X_{t-1}^2, X_{t-2}^2, \cdots) = d_0 + \sum_{j=1}^{\infty} d_j X_{t-j}^2, \tag{3}$$

where $d_i \geq 0$ are some constants. This suggests that σ_t^2 is the autoregressive function of X_t^2 on its lagged values $X_{t-1}^2, X_{t-2}^2, \cdots$.

On the other hand, there exists the abundance of empirical evidence indicating that for some financial returns the autocorrelation of the squared returns $\{X_t^2\}$, although significant, is often not as strong as, for example the autocorrelation of the absolute returns $\{|X_t|\}$. See, for example, Granger et al (1999) and Rydberg (2000) and the references within. Therefore, instead of modelling the conditional second moments as in (3), Ding, Granger and Engle (1993) proposed to model the conditional γ-th absolute moment of X_t given $\mathcal{F}_{t-1}$ by an asymmetric power GARCH formula, with $\gamma \in (0, 2]$ determined by the data; see (4) below.

In this paper, we do not impose any explicit form on σ_t which is merely assumed to be $\mathcal{F}_{t-1}$-measurable. Instead we seek for an index $\gamma \in (0, 2]$ such that a GARCH-like model for $\{|X_t|^\gamma\}$ provides the *best* approximation for σ_t^γ. More specifically, we approximate σ_t^γ by an asymmetric GARCH function

$$\begin{aligned}\xi_{t,\gamma} &\equiv c + \sum_{i=1}^{p} b_i\{|X_{t-i}| - d_i X_{t-i}\}^\gamma + \sum_{j=1}^{q} a_j \xi_{t-j,\gamma} \\ &= c + \sum_{i=1}^{p} b_i |X_{t-i}|^\gamma \{1 - d_i \operatorname{sgn}(\varepsilon_{t-i})\}^\gamma + \sum_{j=1}^{q} a_j \xi_{t-j,\gamma},\end{aligned} \tag{4}$$

for any $\gamma \in (0, 2]$, where the parameters c, b_i, a_j are non-negative, and $d_i \in (-1, 1)$. We then choose the γ such that the approximation is *optimum* in certain sense. Equation (4) admits a unique strictly stationary solution

$$\xi_{t,\gamma} = \frac{c}{1 - \sum_{j=1}^{q} a_j} + \sum_{i=1}^{p} b_i |X_{t-i}|^\gamma \{1 - d_i \operatorname{sgn}(\varepsilon_{t-i})\}^\gamma \tag{5}$$

$$+ \sum_{i=1}^{p} b_i \sum_{k=1}^{\infty} \sum_{j_1=1}^{q} \cdots \sum_{j_k=1}^{q} a_{j_1} \cdots a_{j_k} |X_{t-i-j_1-\cdots-j_k}|^\gamma \{1 - d_i \operatorname{sgn}(\varepsilon_{t-i-j_1-\cdots-j_k})\}^\gamma \tag{6}$$

with $E(\xi_{t,\gamma}) < \infty$, provided that $\{X_t\}$ is strictly stationary with $E|X_t|^\gamma < \infty$, and $\boldsymbol{\theta} \equiv (c, b_1, \cdots, b_p, a_1, \cdots, a_q, d_1, \cdots, d_p)^\tau \in \Theta$, where

$$\Theta = \Big\{(c, b_1, ..., b_p, a_1, ..., a_q, d_1, ..., d_p) \,\Big|\, c, b_i, a_j > 0,\ d_i \in [-1+\delta_0, 1-\delta_0],\ \sum_{j=1}^{q} a_j < 1\Big\}, \tag{7}$$

where $\delta_0 > 0$ is a small constant. We restrict d_i in a closed interval contained in $(-1, 1)$ for some technical convenience.

First we consider how to estimate $\boldsymbol{\theta}$, for a given γ. To make σ_t uniquely defined in (1), we always assume that the median of $|\varepsilon_t|$ is equal to 1, unless specified otherwise. Now $\log(|\varepsilon_t|) = \log(|X_t|) - \gamma^{-1}\log(\sigma_t^\gamma)$ are i.i.d. with median 0. Therefore it holds that

$$\sigma_t^\gamma = \arg\min_{a>0} E\big\{\big|\log|X_t| - \frac{1}{\gamma}\log a\big| \,\big|\, \mathcal{F}_{t-1}\big\}.$$

This leads to an L_1 estimator

$$\widehat{\boldsymbol{\theta}}_1 \;\equiv\; \widehat{\boldsymbol{\theta}}_1^{(\gamma)} = \arg\min_{\boldsymbol{\theta}} \sum_{t=\nu}^{n} \Big| \log|X_t| - \frac{1}{\gamma}\log\{\xi_{t,\gamma}(\boldsymbol{\theta})\}\Big|.$$

An *approximate* (conditional) Gaussian MLEs may also be entertained based on an additional assumption that ε_t in (1) are independent $N(0,1)$ random variables. This condition implies a different parametrisation since now the median of $|\varepsilon_t|$ is not 1. Note σ_t defined in (1) differs under the two parameterisations by a constant independent of t. This impacts on the parameters in $\xi_{t,\gamma}$ as follows; c and all b_i differ by a common constant under the two parametrisation while d_i and a_j remain unchanged. The resulting estimator is

$$\widehat{\boldsymbol{\theta}}_2 \equiv \widehat{\boldsymbol{\theta}}_2^{(\gamma)} = \arg\min_{\boldsymbol{\theta}} \sum_{t=\nu}^{n} \big[X_t^2/\{\xi_{t,\gamma}(\boldsymbol{\theta})\}^{2/\gamma} + 2\gamma^{-1}\log\{\xi_{t,\gamma}(\boldsymbol{\theta})\}\big]. \tag{8}$$

We note that the method is based on approximating σ_t by $\xi_{t,\gamma}^{1/\gamma}$.

Now we consider the problem of estimating the power index γ. Since our goal is to estimate volatility function σ_t, a good estimation should ensure the residuals $\widehat{\varepsilon}_t = X_t/\widehat{\sigma}_t$ behave like an i.i.d. sequence, or, contain little information on $\mathcal{F}_{t-1}$, where $\widehat{\sigma}_t$ denotes an estimator for σ_t. Let $\widehat{\boldsymbol{\theta}}^{(\gamma)}$ be a reasonable estimator for the parameter $\boldsymbol{\theta} \equiv \boldsymbol{\theta}_\gamma$ of $\xi_{t,\gamma}$. Define residuals

$$\widehat{\varepsilon}_t^{(\gamma)} = X_t/\{\xi_{t,\gamma}(\widehat{\boldsymbol{\theta}}^{(\gamma)})\}^{1/\gamma}, \quad t = \nu, \cdots, n. \tag{9}$$

If $\widehat{\varepsilon}_t^{(\gamma)}$ is a good estimator for ε_t, $E\{\widehat{\varepsilon}_t^{(\gamma)} I(X_{t-j} \le x)\} \approx 0$ for any $j \ge 1$ and x. This suggests to choose $\widehat{\gamma} \in [u_0, 2]$ which minimises

$$R(\gamma) \equiv \sum_{j=1}^{k} \sup_x \frac{1}{n}\Big|\sum_{t=\nu}^{n} \widehat{\varepsilon}_t^{(\gamma)} I(X_{t-j} \le x)\Big|, \tag{10}$$

where $k \ge 1$ is an integer, $u_0 > 0$ is a small constant. We restrict $\widehat{\gamma}$ to be bounded away from 0 for technical convenience. The statistics of this type have been used for model checking by, for example, Stute (1997) and Koul and Stute (1999). In practice, we may use either the least absolute deviations estimator $\widehat{\boldsymbol{\theta}}_1^{(\gamma)}$ or the Gaussian MLE $\widehat{\boldsymbol{\theta}}_2^{(\gamma)}$ as $\widehat{\boldsymbol{\theta}}^{(\gamma)}$ in (9), and we may also standardise $\widehat{\varepsilon}_t^{(\gamma)}$ such that the sample mean and variance are, respectively, 0 and 1.

Under some regularity conditions, we have establish the asymptotic normality for the estimators $\widehat{\boldsymbol{\theta}}_1$ and $\widehat{\boldsymbol{\theta}}_2$, and the weak consistency for the estimator $\widehat{\gamma}$. The method has also been illustrated with four sets of financial return data. It is interesting to see that the estimated power index $\widehat{\gamma}$ is often around 1 for those real data sets, leading to better estimation for the volatility function σ_t in comparison with a conventional GARCH fitting.

A Tale of Two Time Scales:
Determining Integrated Volatility with Noisy High-Frequency Data

Lan Zhang
(joint work with Per A. Mykland and Yacine Aït-Sahalia)

In the analysis of high frequency financial data, a major problem concerns the nonparametric determination of the volatility of an asset return process. A common practice is to estimate volatility from the sum of the frequently-sampled squared returns. Though this approach is justified under the assumption of a continuous stochastic model in an idealized world, it meets the challenge from market microstructure in real applications. We argue that this customary way of estimating volatility is flawed in that it overlooks observation error. The usual mechanism for dealing with the problem is to throw away some data, by sampling less frequently or constructing "time-aggregated" returns from the underlying high frequency asset prices. We propose here a statistically sounder device. Our device is model-free, it takes advantage of the rich sources in tick-by-tick data, and to a great extend it corrects the effect of the microstructure noise on volatility estimation. In the course of constructing our estimator, it becomes clear why and where the "usual" volatility estimator fails when the returns are sampled at high frequency.

Our interest lies in using high frequency intraday data to estimate the integrated volatility over some time periods. To fix the ideas, let $\{S_t\}$ denote the price process of a security, and suppose the log-return process $\{X_t\}$, where $X_t = \log S_t$, follows an Itô process

$$X_t = \mu_t dt + \sigma_t dB_t \tag{1}$$

where B_t is a standard Brownian motion. Typically, σ_t^2, the instantaneous variance (or diffusion coefficient) of the return process $\{X_t\}$, will be stochastic. The parameter of interest is the integrated (cumulative) volatility over one or successive time periods, $\int_0^{T_1} \sigma_t^2 dt$, $\int_{T_1}^{T_2} \sigma_t^2 dt$, ... A natural way to estimate the cumulative volatility over, say, a single time interval from 0 to T, is to use the sum of squared incremental returns,

$$\sum_{t_i} (X_{t_{i+1}} - X_{t_i})^2 \approx \int_0^T \sigma_t^2 dt, \tag{2}$$

where the X_{t_i}'s are all the observations of the return process in $[0, T]$. The estimator $\sum_{t_i} (X_{t_{i+1}} - X_{t_i})^2$ is commonly used and generally called "realized volatility" or "realized variance." For a sample of the recent literature in integrated volatility, see Hull and White (1987), Jacod and Protter (1998), Gallant et al. (1999), Chernov and Ghysels (2000), Gloter (2000), Andersen et. al. (2001), Barndorff-Nielsen and Shephard (2001), Mykland and Zhang (2002) and others.

Under model (1), the approximation in (2) is justified by theoretical results in stochastic processes which state that

$$\text{plim} \sum_{t_i} (X_{t_{i+1}} - X_{t_i})^2 = \int_0^T \sigma_t^2 dt, \tag{3}$$

as the sampling frequency increases. In other words, the estimation error of the realized volatility diminishes. According to (3), realized volatility computed from the highest frequency data ought to provide the best possible estimate for $\int_0^T \sigma_t^2 dt$ the integrated volatility.

However, this is not the general viewpoint from the finance literature. It is generally held there that the returns process X_t should not be sampled too often, regardless of the fact that the asset prices can often be observed with extremely high frequency, such as several times per second. It has been found empirically that the estimator is not robust when the sampling interval is quite small. Issues including bigger bias in the estimate and non-robustness to changes in sampling interval have been reported (see e.g., Brown (1990), Campell et al. (1997), Bai et al. (2000)). The main explanation for this phenomenon is a vast array of issues collectively known as market microstructure, such as, but not limited to, the existence of the bid-ask spread: see Aït-Sahalia et al. (2003) for a description of these phenomena and their grounding in the vast theoretical literature describing the frictions inherent in the trading process. When prices are sampled at finer intervals, microstructure issues become more pronounced. It is then suggested that the bias induced by market microstructure effects makes the most finely sampled data unusable for the calculation, and many authors prefer to sample over longer time horizons to obtain reasonable estimates. The length of the typical choices in the literature is ad hoc and ranges from 5 to 30 minutes for exchange rate data, for instance.

This approach to handling the data poses a conundrum from the statistical point of view. We argue here that sampling over longer horizon merely reduces the impact of microstructure, rather than quantifying and correcting its effect for volatility estimation. And it goes against the grain to throw away data. On the other hand, market microstructure may pose so many problems that subsampling is the only way out.

In this paper we analyze the trade-offs involved in the choice of sampling frequency and develop a method to estimate integrated volatility in such a way as to lessen this conflict. Our contention in the following is that the contamination due to market microstructure is, to first order, the same as what statisticians usually call "observation error". We shall incorporate the observation error into the estimating procedure for integrated volatility. In other words, we shall suppose that the return process as observed at the sampling times is of the form

$$Y_{t_i} = X_{t_i} + \epsilon_{t_i}. \tag{4}$$

Here X_t is a latent true, or efficient, return process, and the ϵ_{t_i}'s are independent noise around the true return. A similar structure was used in the parametric context where $\sigma_t = \sigma$ is constant by Aït-Sahalia et al. (2003). In that paper, due to the parametric nature of volatility, we proposed likelihood-based corrections for market microstructure.

We show in the paper that, if the data have a structure of the form (4), ignoring microstructure noise would have a devastating effect on the use of the realized volatility. Instead of (2), one gets

$$\sum_{t_i, t_{i+1} \in [0,T]} (Y_{t_{i+1}} - Y_{t_i})^2 = 2nVar(\epsilon) + O_p(n^{1/2}) \tag{5}$$

where the errors ϵ_{t_i}'s are i.i.d. with mean 0, and n is the number of sampling intervals over $[0, T]$. As we will show, ignoring market microstructure noise in the context of stochastic volatility leads to an even more dangerous situation than when σ is constant and $T \to \infty$. The results from equation (5) suggest that the realized volatility does not estimate the true integrated volatility, but rather the variance of the contamination noise. In fact, we will show that the true integrated volatility, which is $O_p(1)$, is even dwarfed by the magnitude of the asymptotically Gaussian $O_p(n^{1/2})$ term in (5).

Of course, the model (4) may also not be correct. When made the basis of inference, it could still occur that one does not wish to sample as frequently as the data would permit. It may, however, make it possible to use substantially larger amounts of data than what would be possible under (2).

In seeking to create an inference procedure under measurement error, we have sought to draw some lessons from the empirical practice that one should not use all the data, while at the same time not violating basic statistical principles. Our approach is built on separating the observations into multiple "grids". We found that the best results can be obtained by combining the usual ("single grid") realized volatility with the multiple grid based device. This gives an estimator which is approximately unbiased. We have also shown how to assess the (random) variance of this estimator, and how to balance the effect in (5) and an effect due to the sampling frequencies.

The theory, including asymptotic distributions, is developed mainly in the context of finding the integrated volatility over one time period; at the end, we extend this to multiple periods. Also, in the case where the noise can be taken to be almost negligible, we provide a way of optimizing the sampling frequency if one wishes to use the classical "realized volatility" or its multi-grid extension.

One important message of the paper: Any time one has an impulse to sample sparsely, one can always do better with a multi-grid method. No matter what the model is, no matter what quantity is being estimated.

References

[1] Aït-Sahalia, Y., Mykland, P.A. and Zhang, L. (2003): How often to sample a continuous-time process in the presence of market microstructure noise. Technical Report, Princeton University.

[2] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. (2001): The distribution of exchange rate realized Volatility. *J. Amer. Statist. Assoc.* **96**, 42-55.

[3] Bai, X., Russell, J.R. and Tiao G.C. (2000): Beyond Merton's utopia I: effects of non-normality and dependence on the precision of variance estimates using high frequency financial data. Technical Report, University of Chicago.

[4] Barndorff-Nielsen, O.E. and Shephard, N. (2001): Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B* **63**, 167-241.

[5] Brown, S.J. (1990): Estimating volatility. In: *Financial Options: From Theory to Practice*. Eds: S. Figlewski, W.L. Silber and M.G. Subrahmanyam, pp. 516-537. Business One-Irwin, Homewood, IL,

[6] Campbell, J.Y., Lo, A.W. and MacKinlay, A.C. (1997): *The Econometrics of Financial Markets*. Princeton University Press. Princeton, NJ.

[7] Chernov, M. and Ghysels, E. (2000): A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *Journal of Financial Economics* **57**, 407-458.

[8] Gallant, A.R., Chien-Te Hsu and Tauchen, G.T. (1999): Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *Rev. Econom. Statist.* **81**, 617-631.

[9] A. Gloter (2000): *Estimation des Paramètres d'une Diffusion Cachée*. PhD Thesis Université de Marne-la-Vallée.

[10] Hull, J. and White, A. (1987): The pricing of options on assets with stochastic volatilities. *J. Finance* **42**, 281-300.

[11] J. Jacod and P. Protter (1998): Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.* **26**, 267-307.

[12] Mykland, P.A. and Zhang, L. (2002): ANOVA for diffusions. Technical Report, University of Chicago, Department of Statistics.

Reporter: Alexander Lindner

Participants

Prof. Dr. Ole E. Barndorff-Nielsen
oebn@imf.au.dk
Dept. of Mathematical Sciences
Aarhus University
Ny Munkegade
DK-8000 Aarhus C

Prof. Dr. Peter J. Brockwell
pjbrock@stat.colostate.edu
Department of Statistics
Colorado State University
Fort Collins CO 80523-1877 – USA

Dr. Boris Buchmann
bbuch@mathematik.tu-muenchen.de
bbuch@ma.tum.de
Zentrum Mathematik
Technische Universität München
Boltzmannstr. 3
D–85747 Garching bei München

Prof. Dr. Ngai Hang Chan
nhchan@sta.cuhk.edu.hk
Department of Statistics
Chinese University of Hong Kong
Risk Management Science Program
Shatin N.T. – Hong Kong

Prof. Dr. Claudia Czado
cczado@ma.tum.de
Lehrstuhl für Mathematische
Statistik
Technische Universität München
D–85747 Garching bei München

Prof. Dr. Rainer Dahlhaus
dahlhaus@statlab.uni-heidelberg.de
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294
D–69120 Heidelberg

Prof. Dr. Richard A. Davis
rdavis@stat.colostate.edu
Department of Statistics
Colorado State University
Fort Collins CO 80523-1877 – USA

Prof. Dr. Manfred Deistler
Manfred.Deistler@tuwien.ac.at
Institut für Ökonometrie und
Operations Research
Technische Universität
Argentinierstraße 8
A-1040 Wien

Dr. Feike C. Drost
F.C.Drost@TilburgUniversity.nl
Department of Econometrics
Tilburg University
P. O. Box 90153
NL-5000 LE Tilburg

Prof. Dr. Paul Embrechts
embrechts@math.ethz.ch
Departement Mathematik
ETH-Zentrum
Rämistr. 101
CH-8092 Zürich

Vicky Fasen
fasen@ma.tum.de
Zentrum Mathematik
TU München
Boltzmannstr. 3
D–85748 Garching bei München

Prof. Dr. Jürgen Franke
franke@mathematik.uni-kl.de
Fachbereich Mathematik
Universität Kaiserslautern
Erwin-Schrödinger-Straße
D–67653 Kaiserslautern

Prof. Dr. Rüdiger Frey
frey@mathematik.uni-leipzig.de
Mathematisches Institut
Universität Leipzig
Augustusplatz 10/11
D–04109 Leipzig

Prof. Dr. Sylvia Frühwirth-Schnatter
sfruehwi@wv-wien.ac.at
Sylvia.Fruehwirth-schnatter@jku.at
Institut für Angewandte Statistik
Johannes Kepler Universität
A-4040 Linz

Prof. Dr. Xin Guo
xinguo@orie.cornell.edu
School of Operations Research and
Industrial Engineering
Cornell University
Rhodes Hall
Ithaca, NY 14853 – USA

Dipl.-Math. Ralph Högn
hoegn@ma.tum.de
Zentrum Mathematik
Technische Universität München
Boltzmannstr. 3
D–85747 Garching bei München

Prof. Dr. Martin Jacobsen
martin@math.ku.dk
Department of Applied Mathematics
and Statistics
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen

Dr. Jan Kallsen
kallsen@stochastik.uni-freiburg.de
Kallsen@ma.tum.de
HVB-Institute for Mathematical
Finance
Munich University of Technology
Boltzmannstr. 3
D–85747 Garching

Prof. Dr. Claudia Klüppelberg
cklu@ma.tum.de
cklu@mathematik.tu-muenchen.de
Zentrum Mathematik
TU München
Boltzmannstr. 3
D–85748 Garching bei München

Prof. Dr. Siem Jan Koopman
s.j.koopman@feweb.vu.nl
Department of Econometrics
Vrije University
De Boelelaan 1105
NL-1081 HV Amsterdam

Prof. Dr. Jens-Peter Kreiß
j.kreiss@tu-bs.de
Institut für Mathematische
Stochastik der TU Braunschweig
Pockelsstr. 14
D–38106 Braunschweig

Prof. Dr. Catherine Laredo
catherine.laredo@jouy.inra.fr
Laboratoire de Biometrie
Departement de Mathematiques
I.N.R.A.
Domaine de Velvert
F-78352 Jouy-en-Josas Cedex

Dr. Alexander Lindner
lindner@ma.tum.de
Lehrstuhl für Mathematische
Statistik
Technische Universität München
D–85747 Garching bei München

Prof. Dr. Ross A. Maller
Ross.Maller@anu.edu.au
School of Finance and Applied
Statistics
Australian National University
Acton 0200 – Australia

Prof. Dr. Alexander McNeil
mcneil@math.ethz.ch
Departement Mathematik
ETH-Zentrum
Rämistr. 101
CH-8092 Zürich

Prof. Dr. Thomas Mikosch
mikosch@math.ku.dk
Laboratory of Actuarial Mathematics
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen

Dipl.-Math. Gernot Müller
mueller@ma.tum.de
Zentrum Mathematik
Technische Universität München
Boltzmannstr. 3
D–85747 Garching bei München

Prof. Dr. Per Mykland
mykland@galton.uchicago.edu
Department of Statistics
The University of Chicago
5734 University Avenue
Chicago, IL 60637 – USA

Prof. Dr. Michael H. Neumann
mi.neumann@tu-bs.de
Institut für Mathematische
Stochastik der TU Braunschweig
Pockelsstr. 14
D–38106 Braunschweig

Prof. Dr. Serguei Pergamenchtchikov
Serge.Pergamenchtchikov@univ-rouen.fr
Laboratoire de Mathematiques
Raphael Salem
UMR CNRS 6085, UFR Sciences
Univerisite de Rouen
F-76821 Mont Saint Aignan Cedex

Mark Podolskij
podolski@cityweb.de
Fakultät für Mathematik
Ruhr-Universität Bochum
D–44780 Bochum

Dr. Martin Schlather
Martin.Schlather@uni-bayreuth.de
schlather@cu.lu
Laboratoire de Mathematique
Universite du Luxembourg
162 A, avenue de la Faiencerie
L-1511 Luxembourg

Prof. Dr. Neil Shephard
neil.shephard@nuffield.ox.ac.uk
Nuffield College
Oxford University
GB-Oxford OX1 1NF

Prof. Dr. Richard L. Smith
rls@email.unc.edu
Statistics Department
University of North Carolina
Chapel Hill, NC 27599-3260 – USA

Prof. Dr. Michael Sorensen
michael@math.ku.dk
Department of Applied Mathematics
and Statistics
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen

Prof. Dr. Vladimir Spokoiny
spokoiny@wias-berlin.de
WIAS
Mohrenstr. 39
D–10117 Berlin

Prof. Dr. Catalin Starica
starica@math.chalmers.se
Department of Mathematics
Chalmers University of Technology
and University of Göteborg
Eklandag. 86
S-412 96 Göteborg

Prof. Dr. J.Michael Steele
steele@wharton.upenn.edu
Department of Statistics
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6302 – USA

Dr. Daniel Straumann
strauman@math.ethz.ch
Mathematik Departement
RiskLab, H6F 42.3
ETH-Zentrum
Rämistr. 101
CH-8092 Zürich

Dr. Alex Szimayer
aszimaye@ecel.uwa.edu.au
The University of Western Australia
Accounting & Finance
UWA Business School
35 Stirling Highway
Crawley WA 6009 – Australia

Mark Van de Vyver
mvdv@spamcop.net
School of Business (H69)
The University of Sydney
Cnr Rose and Codrington Ave.
Sydney 2006 – Australia

Prof. Dr. Yazhen Wang
yzwang@stat.uconn.edu
Department of Statistics
University of Connecticut
Box U-4120
Storrs, CT 06269-3120 – USA

Prof. Dr. Samuel P. Wong
imsam@ust.hk
Department of Information and
Systems Management
School of Business and Management
Clear Water Bay
Kowloon – Hong Kong

Prof. Dr. Qiwei Yao
q.yao@lse.ac.uk
Department of Statistics
London School of Economics
Houghton Street
GB-London,WC2A 2AE

Prof. Dr. Lan Zhang
lzhang@stat.cmu.edu
Dept. of Statistics
Carnegie Mellon University
Pittsburgh, PA 15213 – USA

Mathematisches Forschungsinstitut Oberwolfach

Report No. 3/2004

Numerical Methods for Instationary Control Problems

Organised by
Karl Kunisch (Graz)
Angela Kunoth (Bonn)
Rolf Rannacher (Heidelberg)

January 18th – January 24th, 2004

Introduction by the Organisers

The topic for the current Mini–Workshop organized by Karl Kunisch (Univ. Graz), Angela Kunoth (Univ. Bonn) and Rolf Rannacher (Univ. Heidelberg) emerged from the Oberwolfach Workshop "Numerical Techniques for Optimization Problems with PDE Constraints" which was held February 16–22, 2003. It was realized that numerically solving control problems which are constrained by time–dependent nonlinear PDEs (Partial Differential Equations) are particularly challenging with respect to the complexity of the problem.

Mathematically one has to minimize a functional under PDE constraints and possibly additional constraints on the state and the control. Standard discretizations on uniform grids in space and time will only yield solutions where the inherent structures of the problem (nonlinearity, constraints) are not sufficiently captured. Certain optimization problems for large coupled systems of partial differential equations are currently not yet numerically treatable or do not satisfy the time constraints required in practice. Overcoming this barrier can only be achieved by designing new mathematically founded algorithmic approaches. The road towards this goal leads to many interesting problems in optimization, linear algebra, numerics, analysis, and approximation theory.

The conference had 21 participants which represented continuous optimization, numerical analysis and scientific computing. 18 talks were given. The 10 longer, overview–style talks were on optimization with PDEs, focussing on

- Modelling and Global Optimization
- Snapshot Selection
- Treatment of State Constraints

- ODE Techniques for PDEs
- Automatic Differentiation
- Adaptive Finite Elements
- Parameter Estimation
- Adaptive Wavelets.

These talks were intended to bridge the gap between the different research fields Optimization and Numerics. They were complemented by 8 shorter talks on more specialized research topics, ranging from efficiency indices for optimization over iterative methods for the coupled systems and multigrid acceleration to modelling issues in optimization, and data compression by means of proper orthogonal decompositions and central Voronoi tesselations.

Different modern approaches to overcome the complexity issues in numerical simulations for PDE–constrained optimization have been presented and discussed. One of the approaches is to employ fast iterative solvers like multigrid on uniform grids. The methodology which conceptually provides the largest potential is to introduce adaptivity. This drastically reduces complexity but depending on the context may require solving an additional problem. Wavelet approaches particularly allow to resolve each of the variables separately and in addition provide a built–in preconditioning. Yet another approach uses compressed information in order to efficiently solve the primal–dual system.

All these issues are currently very active research areas. The extensive discussions held during this workshop have produced a number of new ideas and connections. It was agreed upon that a mere black–box–style matching of efficient PDE codes with optimization tool boxes would on one hand remain much below its potential and on the other hand not help overcoming complexity barriers. Some concepts from automatic differentiation seem to carry over to adaptive methods. Even combining adaptivity with proper orthogonal decompositions may be a very promising direction. The many new ideas discussed during this workshop will have to be further elaborated in future.

Mini–Workshop on Numerical Methods for Instationary Control Problems

Table of Contents

Abstracts

Adaptivity for Optimization of Time-dependent Partial Differential Equations

Roland Becker

In this talk, we derive a posteriori error estimators for optimal control problems. The error estimates are general and apply to Galerkin discretizations of optimization problems governed by partial differential equations. In the first part of the talk we describe some algorithmic aspects of a resulting adaptive algorithm. We focus on the case of time-dependent partial differential equations, where we want to adapt the step size and the dynamically varying meshes for space discretization. This estimate can be used to derive automatic adaptation of $(h, p) \times (k, r)$ method where h denotes the spatial mesh, p the distribution of polynomial degree in the spatial mesh h, k the temporal mesh, and r the polynomial degree in time. In order to simplify the situation we focus on the case where only the mesh h is to be adapted dynamically and the control is frozen, since this seems to be the difficult part. The goal of our adaptive algorithm is to find a method which has computing time linear with respect to the overall number of unknowns and storage requirements proportional to the temporal mean of the employed meshes. In order to achieve this goal, we have to use a divide-and-conquer algorithm as the checkpointing/windowing algorithm known from automatic differentiation and optimal control. The essential additional difficulty in our context is the fact that we need information about the co-state (solving a backwards-in-time equation) when computing forward. This is due to the structure of the error estimator.

The second part of the talk describes a posteriori error estimates for optimization problems. We consider the following general case: we are interest in computing an interest function $I(q, u)$ which depends on both control q and state variable u. The interest functional is independent of the optimization problem which determines q and u. By specialization we obtain estimators for:

- error in the cost functional [2, 3]
- error in a functional of the controls [4]
- error in an independent functional of the state [5]
- norm of the error of controls [1]

Beside the first estimator, the others require the solution of an additional problem involving the adjoint of the linearized state operator. The right-hand side of this problem depends on the special context.

In the last part of the talk, we show that the information given by the additional adjoint problem might be used for further purposes. We employ the well-known concept of condition numbers in order to produce answers to the following questions: which parameters have been most important, which measurements have been most important in computing the solution?

References

[1] R. Becker, Adaptive finite element methods for optimal control of partial differential equations: Norms of the error, Technical report, LMA, Université de Pau et du Pays de l'Adour, 2004.

[2] R. Becker, H. Kapp, R. Rannacher, Adaptive finite element methods for optimal control of partial differential equations: Basic concepts, SIAM J. Contr. Optim., 39(1):113–132, 2000.

[3] R. Becker, R. Rannacher, An optimal control approach to a-posteriori error estimation, Acta Numerica 2001, pp. 1–101, Cambridge University Press.

[4] R. Becker, B. Vexler, A posteriori error estimation for finite element discretizations of parameter identification problems, to appear in Numer. Math.

[5] R. Becker, B. Vexler, Numerical sensitivity analysis with adaptive finite elements: Applications to CFD problems, submitted.

Fast Integral Equation Solvers and Applications to Problems with Dynamic Interfaces

George Biros

(joint work with L. Ying and D. Zorin (New York University))

Boundary integral formulations have been extensively used for the analysis and numerical solution of elliptic partial differential equations. However, with the exception of problems in inverse scattering, there has been limited work in boundary integral equation formulations for optimal control and optimal design problems. There are several reasons for that: prohibitive complexity of efficient implementations for non-Laplacian kernels, difficulty with distributed force terms, and restriction to problems with piecewise constant coefficients.

Recent developments however, indicate integral equation formulations might have impact to a larger class of problems. We present two new algorithms and two examples that illustrate the efficiency of the new methods: (1) A fast solver for Stokes and Navier-Stokes equations, (2) a new kernel-independent fast-multipole method for kernels related to constant coefficient elliptic PDEs, (3) a three-dimensional rigid body-fluid interaction problem, and (4) a prototypical shape optimization problem of a Dirichlet interior Stokes problem.

1. Fast Solvers for Stokes and Navier-Stokes Equations. Our motivation in designing this method is the design of efficient algorithms for flows with dynamic interfaces. Solvers for such problems should be built on algorithms that do not require expensive preprocessing, like unstructured mesh generation, since the interface is moved to a new location at each time step. The main features of the method are: It can be applied to arbitrary piecewise smooth geometries; It does not require mesh generation; It can solve problems with distributed forces; It is second-order accurate and readily generalizes to arbitrary order of accuracy. If an optimal boundary integral equation solver is used, the method has $\mathcal{O}(N)$ complexity.

Our method is based on potential theory for linear second-order elliptic operators. Using an indirect integral formulation, the solution of a Dirichlet problem can be written as the sum of a double layer potential and a *Newton potential* (the domain convolution of the Green's function with the distributed force). Under such a scheme the evaluation of the solution must consist of three steps: (1) computation of the Newton potential, (2) solution of a boundary integral equation to compute a double layer potential that satisfies the boundary conditions, and (3) evaluation of a double layer potential everywhere in the domain. Details can be found in [2] for the Stokes equations and [1] for the unsteady Navier-Stokes problem.

2. Kernel-independent Fast Multipole Method. The main feature of the new method is its black-box application to several different non-oscillatory elliptic kernels. Our algorithm has the structure of the adaptive FMM algorithm [3] but requires only kernel evaluations without sacrificing accuracy and efficiency compared to the original algorithm. The crucial element of our approach is to replace the analytic expansions and translations with *equivalent density* representations. These representations are computed by solving local exterior and interior problems on circles (2D), spheres or cubes (3D) using integral equation formulations. We have demonstrated the efficiency of our method in both 2D and 3D for many kernels: the single and double layer potentials of the Laplacian, the modified Laplacian, the Navier, the Stokes, and their modified variants. Our method has $\mathcal{O}(N \log N)$ asymptotic complexity, whereas for reasonable assumptions on the initial particle distribution the complexity becomes $\mathcal{O}(N)$. Like analytic FMM, our method works well for nonuniform particle distributions. Details can be found in [4]. We have also developed an MPI-based parallel version of the method, and have performed systematic scalability tests. Overall we have achieved very good iso-granular and fixed-size scalability on up to 3000 processors. A detailed discussion can be found in [5].

3. Fluid-structure Interaction Formulation. We have developed algorithms to simulate the interaction of rigid bodies of arbitrary geometry with Stokesian fluids, ignoring inertial terms in the fluid and using an integral formulation for the equations which describe the motion of the dynamics. These equations are a set of integrodifferential equations the interaction between the fluid and a rigid object and consist of the linear and angular momentum conservation for the rigid body and the Stokes equations for the fluid. The coupling is induced by the requirement of non-slip condition and force balance across the interface. These results are work in progress. For the fluid-structure interaction runs we are currently working on convergence studies, and we have not performed systematic scalability analysis. However, the main cost in these simulations is the solution of the underlying integral equations. To this end we are working on efficient preconditioning schemes.

4. Shape Optimization for Stokes-constrained Systems. We present a 2D shape optimization problem for the interior Dirichlet problem of a Stokesian fluid. The flow is represented using a second kind integral equation formulation. The objective function is of tracking type. The boundary is represented using a periodic B-spline and the optimization variables are the control points. Adjoints

are used to compute shape sensitivities and the problem is solved using a reduced space quasi-Newton method globalized by trust-region. The first derivatives of the adjoint and forward problem involve hypersingular kernels (Hilbert transforms) which are approximated using an odd-even Nyström integration scheme. The results, although very preliminary, are very encouraging since only a small number of quasi-Newton iterations are sufficient for convergence.

References

[1] G. Biros, L. Ying, and D. Zorin, The embedded boundary integral method for the unsteady incompressible Navier-Stokes equations, Technical Report TR2003-838, Courant Institute, New York University, 2002.

[2] G. Biros, L. Ying, and D. Zorin, A fast solver for the Stokes equations with distributed forces in complex geometries, Journal of Computational Physics, 193(1):317–348, 2003.

[3] L. Greengard, The Rapid Evaluation of Potential Fields in Particle Systems, MIT Press, Cambridge, MA, 1988.

[4] G. Biros, L. Ying, and D. Zorin, A kernel-independent adaptive fast multipole method in two and three dimensions, Journal of Computational Physics, to appear.

[5] G. Biros, L. Ying, D. Zorin, and H. Langston, A new parallel kernel-independent fast multiple algorithm, in: Proceedings of SC03, The SCxy Conference series, Phoenix, Arizona, November 2003, ACM/IEEE.

On Space-time Multigrid Solution of Unsteady Optimal Control Problems

Alfio Borzì
(joint work with K. Kunisch and R. Griesse)

The development and investigation of space-time multigrid schemes for unsteady reaction diffusion optimal control problems are reported. We focus on the control of the time evolution of chemical and biological processes characterized by non-monotone nonlinearities. For benchmarking our algorithms, we propose two models:

The solid fuel ignition model

$$-\partial_t y + \delta e^y + \Delta y = u, \quad \delta > 0,$$

results in a singular optimal control problem which cannot be solved by any method based on a free evolution of the state variable.

The lambda-omega system is given by

$$\frac{\partial}{\partial t}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} \lambda(y_1,y_2) & -\omega(y_1,y_2) \\ \omega(y_1,y_2) & \lambda(y_1,y_2) \end{bmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \sigma\,\Delta\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

where

$$\lambda(y_1,y_2) = 1 - (y_1^2 + y_2^2) \quad \text{and} \quad \omega(y_1,y_2) = -\beta(y_1^2 + y_2^2).$$

The functional form of λ and ω was proposed in [4] to model chemical turbulence. The evolution into a chaotic state of the $\lambda - \omega$ system can also be observed from a principal component analysis via proper orthogonal decomposition of its snapshots. As the system becomes less ordered, the energy content becomes more and more evenly distributed among the eigenmodes.

In the first model case, the control, represented by u, is applied to avoid blow-up or to optimize the combustion process [1]. In the second case the control (u_1, u_2) is applied to drive the system from a turbulent to a regular state [2]. The optimality systems characterizing the optimal control solution are solved by space-time multigrid schemes with typical multigrid efficiency and robustness with respect to the choice of the optimization parameters. These features are obtained by developing appropriate collective smoothing schemes.

Using two-grid Fourier analysis, sharp estimates of convergence factors are obtained for linear model problems [3]. Results of numerical experiments demonstrate that these estimates remain sharp also for the nonlinear cases considered here.

References

[1] A. Borzì, K. Kunisch, A multigrid method for optimal control of time-dependent reaction diffusion processes, In K.H. Hoffmann et al. (eds.), Fast Solution of Discretized Optimization Problems, Internat. Ser. on Numerical Mathematics, Vol. 138, Birkhäuser, 2001.

[2] A. Borzì, R. Griesse, Distributed optimal control of lambda-omega systems, submitted to SICON.

[3] A. Borzì, Multigrid methods for parabolic distributed optimal control problems, J. Comp. Appl. Math, 157 (2003), pp. 365–382.

[4] Y. Kuramoto, S. Koga, Turbulized rotating chemical waves, Prog. Theor. Phys., 66(3) (1981), pp. 1081–1085.

Algorithms for Optimization of Time-dependent PDE Systems: Can We Realize the Same Efficiencies as in the Steady Case?

Omar Ghattas

(joint work with G. Biros (University of Pennsylvania), V. Akcelik, J. Bielak, and I. Epanomeritakis(Carnegie Mellon University))

The answer to the question posed in the title depends of course on the type of problem being solved. We begin by recalling some of our earlier work on fast solvers for optimization problems that are governed by PDEs [1,2,3,4,5]. The method we developed, which we refer to as *Lagrange-Newton-Krylov-Schur (LNKS)*, solves the full optimality system consisting of state, adjoint, and control equations using an inexact preconditioned Newton-QMR method. The preconditioner is a block factorization that emulates a reduced quasi-Newton SQP method: it approximates the reduced Hessian via suitably-initialized limited memory BFGS updates while discarding other second derivative terms, and replaces the exact state and adjoint solves with application of appropriate preconditioners, e.g. additive Schwarz or multigrid. If sufficient descent cannot be obtained with a line search, then we drop down to the reduced space and take a quasi-Newton step. Experiments with this method on some problems of optimal control of three-dimensional steady Navier-Stokes flows via boundary suction/injection demonstrate high parallel scalability, mesh-independence of Newton iterations, mesh-independence of Krylov iterations (provided an optimal state preconditioner is available), and solution to optimality in four times the cost of a flow solution, for a problem with over 600,000 state and 9000 control variables. This small constant multiple of the state solve cost is due to iterating in the full space, which folds the iterations (linear and nonlinear) needed to converge the flow into those required for optimization. LNKS is most effective when the state equations are difficult to solve, requiring many iterations.

We next discuss an inverse parameter estimation problem governed by earthquake modeling via the elastic wave equation. The problem is to find the distribution of elastic parameters of large sedimentary basins such as Los Angeles, from surface observations of past earthquakes [6,7]. The forward problem alone requires terascale computing: our typical earthquake simulations resolve up to 1 Hz ground motion frequencies, involve 100 million grid points and 40,000 times steps, and require several hours of run time on the 3000 processor AlphaCluster supercomputer in Pittsburgh. The inverse problem is formulated via output least squares, regularized by a total variation (TV) functional. TV eliminates oscillatory components of the material properties, while preserving discontinuities at material interfaces. The solver is an inexact Newton-CG method in the reduced space, with the same preconditioning as in LNKS. However, because we iterate in the reduced space, an exact forward and adjoint solve are required at each CG iteration. Mesh independence of Newton and CG iterations is observed, and the number of inner and outer iterations required for convergence is similar to those observed for the flow control problem. However, the difference here is the requirement for exact solution of the forward and adjoint wave equation. For the largest inverse problem we solved, involving 17 million inversion parameters, the product of inner and outer iterations is such that 800 total forward and adjoint wave propagations are required. The essentially renders the inverse problem intractable for our goal of reconstructing the structure of the LA Basin to a 1 Hz frequency resolution.

We conclude the talk with a somewhat pessimistic discussion of several opportunities for speeding up the convergence of the earthquake inversion problem. Additional processors won't help, since the granularity of the computation is low to begin with. We could switch from a reduced space solver to a full space LNKS method, but there is nothing to be gained since the explicit forward solver offers little opportunity for approximation with a simpler solve. A coarser mesh misses the finest wavelengths, and a longer time step loses the shorter periods, both of which contribute importantly to the surface response. Implicit methods are not useful for wave propagation problems in which the system is responding in all of its resolvable scales. Adding processors in the time direction is not helpful, since information propagates at finite speed. Reduced order modeling in the state space is likely unproductive, since the system is responding at all scales (indeed the mesh was designed to just resolve the finest scales of interest). Similarly, reduced order modeling in the parameter space faces the problem of trying to construct a response surface in a very high dimensional (e.g. 10 million) space. Some improvements in the linear preconditioner can probably be made by exploiting the compact and differential structure of the reduced Hessian, but since we have typically just 20 CG iterations per Newton iteration, the reduction in iterations must be balanced against the cost of construction of the preconditioners. Similar statements can be made about improvements in the nonlinear solver, e.g. through nonlinear preconditioning. We conclude that the inverse earthquake modeling problem to frequencies of engineering interest remains a major challenge, when measured against the cost of the forward simulation.

References

[1] G. Biros, O. Ghattas, Parallel domain decomposition methods for optimal control of viscous incompressible flows, Proceedings of Parallel CFD '99, Williamsburg, Virginia, USA, May 1999.

[2] G. Biros, O. Ghattas, Parallel Newton-Krylov methods for PDE-constrained optimization, Proceedings of SC1999, IEEE/ACM, Portland, OR, USA, November 1999.

[3] G. Biros, Parallel Lagrange-Newton-Krylov-Schur Methods for PDE-constrained Optimization, with Application to Optimal Control of Viscous Flows, Ph.D. Thesis, Carnegie Mellon University, September 2000.

[4] G. Biros and O. Ghattas, Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part I: The Krylov-Schur solver, to appear in SIAM J. on Scientific Computing.

[5] G. Biros and O. Ghattas, Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part II: The Lagrange-Newton solver, and its application to optimal control of steady viscous flows, to appear in SIAM J. on Scientific Computing.

[6] V. Akcelik, G. Biros, and O. Ghattas, Parallel multiscale Gauss-Newton-Krylov methods for inverse wave propagation, Proceedings of SC2002, Baltimore, MD, USA, November 2002.

[7] V. Akcelik, J. Bielak, G. Biros, I. Epanomeritakis, A. Fernandez, O. Ghattas, E. Kim, J. Lopez, D. O'Hallaron, T. Tu, and J. Urbanic, High-resolution forward and inverse earthquake modeling on terascale computers, Proceedings of SC2003, IEEE/ACM, Phoenix, AZ, USA, November 2003.

From Algorithmic Differentiation to Automated Design

Andreas Griewank

To facilitate the transition from simulation to optimization we suppose that we are provided (only) with an iterative solver for some state equation and a procedure for evaluating an objective function. This is a realistic scenario in aerodynamics, where the state equation is some discretized variant of the Navier Stokes equation governing the flow around a wing, and the objective may be the drag, which is to be minimized by varying the design of the wing. Since the state space may have very high dimensionality we prefer not to modify the given solver but merely assume that it is contractive as an iterative map. However, using algorithmic differentiation we may derive from it, in an automated fashion, a fixed point solver for the corresponding adjoint equation and the computation of an approximate reduced gradient. This methodology had been introduced by Griewank and Faure under the name of piggy-back differentiation [1].

In this talk we show first that while the two fixed point iterations converge with the same asymptotic contractivity rate, the accuracy of succesive approximate solutions to the adjoint equation lags behind that for the underlying state space iterates. More specifically, the ratio between the residuals of the two equations at the k-th, coupled iteration grows linearly with k [2].

Secondly, we examine the choice of a matrix for preconditioning the approximate reduced gradient in a simultaneous update of the design variables. As it turns out the seemingly natural choice of the reduced Hessian is not the best choice, but may lead to divergence. Instead we find that local convergence can be assured by projecting the Lagrangian Hessian onto another subspace, at least when the full Hessian is positive definite [3]. These theoretical observation are confirmed numerically on a 2D test problem provided by Volker Schulz. Either of the two projected Hessians can also be evaluated by automatic differentiation, so that we obtain a methodology for optimal design in a fairly automated fashion. Practical validations on Euler and Navier Stokes codes for 3D and 2D flows are under way.

References

[1] A. Griewank, C. Faure, Reduced functions, gradients and hessians from fixed point iteration for state equations, Numer. Algor., 30 (2002), no. 2, 113–139.

[2] A. Griewank, C. Faure, Linear and quadratic delay of reduced gradients and Hessians obtained from fixed point iterations for state equations, in preparation.

[3] A. Griewank, The correct projected Hessian for preconditioning the approximate reduced gradient in piggy-back optimization, in preparation.

Design of Experiments and Snapshot Generation in Reduced-order Modeling

Max Gunzburger

Reduced-order modeling strategies such as proper orthogonal decomposition (POD) methods are developed from a set of snapshots. The reduced-order model cannot contain more information than that contained in the snapshot set. Thus, the generation of snapshots is crucial to the success of reduced-order models. The generation of snapshots is an exercise in the design of experiments, i.e., how does one choose the values of the parameters used to generate the snapshot simulations or the time instants at which one evaluates the snapshots? We discuss the use of design of experiment-based strategies for parameter selection for snapshot generation. Issues that arise in selecting a method for sampling points in parameter space are considered, and the relative merits of different methods (e.g., quasi-Monte Carlo sequences, Latin hypercube sampling, centroidal Voronoi tessellation, etc.) are discussed. Several notions of uniformity for point sets are compared, as are their effect on deciding which sampling methods are best for specific applications. Also, the role of known information about the parameters and how to incorporate this information into the point sampling process for snapshot generation are considered.

A major consideration of the talk is a new point sampling strategy that we have developed that is based on centroidal Voronoi tessellations (CVT's). These are special Voronoi tessellations for which the generators of the Voronoi tessellation are also the centers of mass, with respect to a given density function, of the associated Voronoi cells. CVT's have many uses in many applications; in particular, CVT's can be used for generating very high-quality point sets in regions and on surfaces. Using several volumetric measures of uniformity, CVT point samples are shown to be more "uniform" that those obtained by existing strategies. On the other hand, CVT point sets do not have good properties when projected onto lower dimensional surfaces, e.g., the faces of a hypercube. Such a property is desirable in some applications such as high-dimensional integration. For the latter application, one can define "Latinized" CVT point sets that possess good projections. For design of experiment applications relevant to snapshot generation, both CVT and Latinized CVT point sets are superior to existing points sampling methods. Incidentally, CVT strategies also offer an alternative to POD as a means for defining a reduced basis from a set of snapshots.

Various aspects of the talk represent joint work with John Burkardt, Hoa Nguyen, Janet Peterson, and Yuki Saka (Florida State University) and Hyung-Chun Lee (Ajou University, Korea).

References

[1] J. Burkardt, M. Gunzburger, and H.-C. Lee, Centroidal Voronoi tessellation-based reduced-order modeling of complex systems, to appear in SIAM J. on Scientific Computing.

[2] J. Burkardt, M. Gunzburger, H. Nguyen, J. Peterson, and Y. Saka,Centroidal Voronoi tessellation points sampling I: Uniform sampling in hypercubes, in preparation.

[3] J. Burkardt, M. Gunzburger, H. Nguyen, J. Peterson, and Y. Saka, Centroidal Voronoi tessellation point sampling II: Nonuniform, anisotropic sampling in general regions, in preparation.

Adjoint-based Adaptive Time-stepping for Partial Differential Equations using Proper Orthogonal Decomposition

Vincent Heuveline

(joint work with M. Hinze (TU Dresden))

We present an effective adjoint-based a-posteriori goal-oriented error control mechanism [1] for time integration of partial differential equations. The sensitivity information is obtained from the adjoint of a reduced order model of the full partial differential equations [2] while the reduced order model is adapted during the error estimation process. Several numerical examples illustrate the performance of the method.

References

[1] R. Becker, R. Rannacher, An optimal control approach to a posteriori error estimation in finite element methods, In Acta Numerica 2001, Cambridge University Press, pp. 1–101.

[2] M. Hinze, K. Kunisch, Three control methods for time-dependent fluid flow, Flow, Turbulence and Combustion 65, 273-298 (2000).

Model Reduction (and Spacetime Multigrid) in Control of Time-dependent PDEs

Michael Hinze
(joint work with K. Afanasiev (ZIB Berlin) and G. Büttner (TU Berlin))

We present an effective control method for mathematical models governed by systems of nonlinear time-dependent partial differential equations. It takes account of the fact that control inputs may alter the regime of the underlying physical process. The method in an adaptive manner constructs a hierarchy of appropriate low dimensional approximations to the mathematical model which then are used as subsidiary condition in the related optimization problem. We discuss different possibilities to construct low dimensional systems and the related modes (eigenfunctions of stationary problem, eigenfunctions of the linearized model and snapshot form of proper orthogonal decomposition).

As numerical example we present tracking-type control of the incompressible Navier-Stokes system as mathematical model for periodic flow around a cylinder. The numerical results of the adaptive approaches for different modes are compared. Furthermore they are compared to the result of the optimal control approach applied to the full Navier-Stokes system. It turns out that the quality of the controls obtained from the suboptimal approaches compares to that obtained by optimal control, and the computational costs for the optimal approach are at least one order of magnitude larger. More specifically, for the numerical example considered we obtain

$$\text{Runtime(Optimization)} = \text{(6-8) x Runtime(pdesolve)},$$

see [1].

In the second part of the talk we present preliminary results for multigrid in spacetime following the integral equation approach of Hackbusch. It turns out that for the numerical solution of linear-quadratic control problems we achieve

$$\text{Runtime(Optimization)} = \text{(7-8) x Runtime(pdesolve)},$$

for the numerical solution of a nonlinear problem with inexact Newton methods and spacetime multigrid for the Newton defect system we achieve

$$\text{Runtime(Optimization)} = \text{20 x Runtime(pdesolve)},$$

see [2].

References

[1] K. Afanasiev, M. Hinze, Adaptive control of a wake flow using proper orthogonal decomposition, Lecture Notes in Pure and Applied Mathematics 216 (2001).
[2] G. Büttner, Ein Mehrgitterverfahren zur optimalen Steuerung parabolischer Probleme, Dissertation TU Berlin (2004).

Adaptive Finite Elements for Inverse Scattering
Claes Johnson
(joint work with L. Beilina)

We apply an adaptive hybrid FEM/FDM method for an inverse scattering problem for the time-dependent acoustic wave equation in 2D and 3D where we seek to reconstruct an unknown sound velocity $c(x)$ from measured wave-reflection data. Typically this corresponds to identifying an unknown object [scatterer] in a surrounding homogeneous medium.

We use an optimal control approach where we seek a sound velocity $c[x]$ which minimizes the difference between computed and measured output data in a discrete L_2 norm. We solve the optimization problem by a quasi-Newton method where in each step we compute the gradient by solving a forward [direct] and an backward [adjoint] wave propagation problem.

To compute the backward and forward wave propagation problems we use an adaptive hybrid finite element/finite difference method where we exploit the flexibility of mesh refinement and adaption of the finite element method in a domain covering the object and the efficiency of a structured mesh finite difference method in the surrounding homogeneous domain. The hybrid scheme can be viewed as a finite element scheme on a partially unstructured mesh which gives a stable coupling of the two methods.

We use an adaptive mesh refinement algorithm to improve the accuracy of the reconstruction and speed up the convergence of the quasi-Newton method.

References

[1] L. Beilina, C. Johnson, Adaptive hybrid FEM/FDM methods for inverse scattering problems, Preprint, Chalmers Finite Element Center, 2003.

Structural Aspects for Numerical Methods in Optimal Control of Evolution Equations
Karl Kunisch

Motivated by optimal control problems in fluid dynamics describing vortex or drag reduction I remark in this survey talk on some aspects in open as well as closed loop numerical optimal control.

In the first part the difference between the optimization based approach and methods focused on solving the optimality system is explained. Sequential quadratic programming methods are compared to Newtons method. In the former the linearized state equation in the latter the nonlinear state equations are solved, resulting in primal feasibility in case of Newtons method [HK2]. In view of time-stepping techniques which are typically used to integrate the dynamical system, the numerical cost between the solution of the linearized equation and inexact solutions to the nonlinear equation can be low. This suggests to favor the Newton over the SQP method for optimal control of evolution problems. — Finally a new cost functional for vortex suppression relying on local phase plane analysis is proposed [SK].

The second part is devoted to feedback control. For linear quadratic problems feedback control is completely characterized by means of an operator Riccati equation. In the nonlinear case the feedback control relies on the viscosity solution of a Hamilton-Jacobi-Bellman equation, which is numerically unfeasible unless the dimension of the (discretized) state-space is unreasonably small. For optimal control of fluid flow this requires the use of approximation strategies (beyond state space discretization). Here we explain techniques which rely on model reduction based on proper orthogonal decomposition (POD) combined with numerical solutions of the HJB-equation for infinite as well as finite horizon problems for the reduced problem [KV1, KV2, KX]. We also describe receding horizon techniques which rely on a time domain splitting strategy. To analytically justify their use we consider the stabilization problem of steady states. If a control Ljapunov functional is used as terminal penalty then the receding horizon control successfully drives the trajectory to the steady state [IK1, IK2].

References

[HK1] M. Hinze, K. Kunisch, Three control methods for time-dependent fluid flow, Journal of Flow, Control and Combustion, 65 (2000), 273–298.

[HK2] M. Hinze, K. Kunisch, Second order methods for optimal control of time-dependent fluid flow, SIAM J. Control and Optimization, 40 (2001), 925–946.

[IK1] K. Ito, K. Kunisch, On asymptotic properties of receding horizon optimal control, SIAM J. on Control and Optimization, 40 (2001), 1455–1472.

[IK2] K. Ito, K. Kunisch, Receding horizon control for infinite dimensional systems, ESAIM, Control, Optimization and Calculus of Variations, 8 (2002), 741-760.

[IK3] K. Ito, K. Kunisch, Semi-smooth Newton methods for state-constrained optimal control problems, to appear in Systems and Control Letters.

[KM] K. Kunisch, X. Marduel, Sub-optimal control of transient non-isothermal viscoelastic fluid flows, Physics of Fluids, 13(2001), 2478–2491.

[KV1] K. Kunisch, S. Volkwein, Galerkin proper orthogonal decomposition methods for parabolic systems, Numerische Mathematik, 90 (2001), 117–148.

[KV2] K. Kunisch, S. Volkwein, Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics, to appear in SIAM J. on Numerical Analysis.

[KX] K. Kunisch, L. Xie, Suboptimal feeback control of flow separation by POD Model Reduction, Submitted.

[SK] Y. Spasov, K. Kunisch, Dynamical system based optimal control of incompressible fluids: Boundary control, Submitted.

Adaptive Wavelets for Optimal Control Problems

Angela Kunoth

For the fast numerical solution of control problems governed by partial differential equations, an adaptive algorithm based on wavelets is proposed. The framework allows for linear elliptic and parabolic PDEs in full weak space–time formulation as constraints as well as for problems with distributed or with Neumann or Dirichlet boundary control.

A quadratic cost functional which may involve fractional Sobolev norms of the state and the control is to be minimized subject to linear constraints in weak form. Placing the problem into the framework of (biorthogonal) wavelets allows us to represent the functional and the constraints in terms of ℓ_2–norms of wavelet expansion coefficients and constraints in form of an ℓ_2 automorphism. The resulting first order necessary conditions are then derived as a (still infinite) system in ℓ_2.

Applying the machinery developed in [CDD], we propose an adaptive method for the coupled system for the state, adjoint and control variables. An essential ingredient is that the scheme can be interpreted as an inexact gradient descent method, where in each iteration step the primal and the adjoint system needs to be solved up to a prescribed accuracy. In particular, the method resolves each of the involved variables separately, without having to resort to a common underlying grid.

The convergence analysis of the algorithm is crucially based on the fact that the wavelet framework allows us to step by step break down an ideal iteration on the infinite system to computable quantities. Thus, the method captures all relevant features from the original control problem with respect to infinite function spaces and specifically resolves any type of singularity coming from the data or the domain. Consequently, the approximate solutions generated by the adaptive algorithm can be shown to converge (in the energy norm) to the exact solution triple (state, adjoint state, control) for any prescribed accuracy.

Moreover, it is shown that the adaptive algorithm is asymptotically optimal, that is, the convergence rate achieved for computing the solution up to a desired target tolerance is asymptotically the same as the wavelet–best N–term approximation of the solution, and the total computational work is proportional to the number of computational unknowns.

References

[CDD] A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet methods for elliptic operator equations — Convergence rates, Math. Comp., 70 (2001), 27–75.

[DK] W. Dahmen, A. Kunoth, Adaptive wavelet methods for linear–quadratic elliptic control problems: Convergence Rates, Preprint #46, SFB 611, Universität Bonn, December 2002, submitted for publication, revised, November 2003.

[GK] M. Gunzburger, A. Kunoth, Adaptive wavelets for optimal control of linear parabolic PDE's, in preparation.

[K] A. Kunoth, Adaptive wavelet schemes for an elliptic control problem with Dirichlet boundary control, Preprint #109, SFB 611, Universität Bonn, November 2003, to appear in Numer. Algor.

A Paradigm for Adaptivity and Optimal Control
Rolf Rannacher

We present a general approach to “goal-oriented” error estimation and mesh adaptation in the context of optimal control problems of the abstract form

$$J(u,q) = \min!, \qquad A(u,q)(\cdot) = 0,$$

where the equation constraint is a PDE. The starting point is the first-order optimality condition, the so-called KKT system,

$$\nabla_{u,q,z} L(u,q,z) = 0, \qquad L(u,q,z) := J(u,q) - A(u,q)(z),$$

obtained by the classical Euler-Lagrange approach. This set of nonlinear and strongly coupled PDEs for the primal variable (state) u, the control variable q (control), and the dual variable z (adjoint) is approximated by a Galerkin finite element method. The topic of this talk is a strategy for the a posterori construction of finite element meshes which are most economical for the optimization process. Exploiting the particular structure of the KKT system an error representation is derived for the discretization with respect to the cost functional in terms of the residuals of the computed solution and a remainder term which is cubic in the approximation errors,

$$J(u,q) - J(u_h,q_h) = \tfrac{1}{2}\nabla_{u,q,z} L(u_h,q_h,z_h)(z-\psi_h, q_h-\chi_h, u-\phi_h) + R_h^{(3)}(e^u, e^q, e^z).$$

The remainder term is usually neglected. The computational evaluation of the residual term yields “weighted” a posteriori error estimates of the form

$$J(u,q) - J(u_h,q_h) \approx \sum\nolimits_{K\in T_h} \{\rho_K^u \omega_K^z + \rho_K^q \omega_K^q + \rho_h^z \omega_K^u\},$$

which can guide the mesh adaptation process. In these estimates “primal” residuals and “dual” weights as well as “dual” residuals and “primal” weights are crosswise multiplied, while in the traditional approach these quantities are added. Therefore, this approach of mesh adaptation is called “Dual Weighted Residual (DWR) Method”. The resulting meshes for discretizing the KKT system are most economical since only that information is represented which is really needed for the optimization process. When an optimal control q_h^{opt} is determined for this sparse model, using for instance a Newton-type iteration, a more accurate primal solution $\tilde{u}_h^{\mathrm{opt}}$ can be generated in a post-processing step by solving the state equation once again on a finer mesh. The performance of the DWR method is demonstrated for a simple diffusion-reaction problem with Neumann-boundary control. The very same approach can also be used in the context of time discretization and eigenvalue problems, e.g., in the control of the Navier-Stokes equations for drag minimization and for the stability analysis of the resulting optimal states. For further information and other classes of applications, we refer to the literature listed below.

References

[1] R. Becker, H. Kapp, R. Rannacher, Adaptive finite element methods for optimal control of partial differential equations: Basic concepts, SIAM J. Control Optim. 39, 113–132 (2000).

[2] R. Becker, R. Rannacher, An optimal control approach to error estimation and mesh adaptation in finite element methods, Acta Numerica 2001, pp. 1–101, Cambridge University Press.

[3] R. Becker, Mesh adaptation for stationary flow control, J. Math. Fluid Mech. 3, 317–341 (2001).

[4] R. Becker, V. Heuveline, R. Rannacher, An optimal control approach to adaptivity in computational fluid mechanics, Int. J. Numer. Meth. Fluids. 40, 105–120 (2002).

[5] W. Bangerth, R. Rannacher, Adaptive Finite Element Methods for Differential Equations, Lectures in Mathematics, ETH Zürich, Birkhäuser, Basel-Boston-Berlin, 2003.

Modelling and Globalizing Applied to Optimal Control Algorithms

Ekkehard W. Sachs

Various models in algorithms for optimization have led to different strategies of globalizing optimization methods. The trust region concept is a general globalization strategy which can be applied to various model functions.

For an unconstrained optimization problem the trust region strategy defines at each iteration point u_k a nonlinear model function $m_k(u_k + \cdot)$ which is minimized over a ball of radius Δ. The radius is updated according to the accuracy of the value of the model function in comparison to the value of the true function at the candidate at the next iteration point. The trust region is updated in a similar fashion. These strategies together with a condition on a sufficient descent for a function value yield a global convergence statement.

The linear model leads to the well known Armijo step size rule, whereas the quadratic model yields a smooth transition from steepest descent to the Newton step. It is less known that the trust region method can also be applied to nonlinear models as pointed out by Toint [4].

The training of neural networks is one area where these techniques have been applied successfully for large scale problems.

Another application is reduced order modelling, in particular, proper orthogonal decomposition for the optimal control of Navier-Stokes equation. Afanasiev and Hinze [1] also use adaptive techniques to adjust the POD model during the iteration. In Arian, Fahl, Sachs [2, 3] the trust region approach is utilized to control the adaption of the POD model. In this case the differential equation is replaced by a reduced order model of smaller size. In this case the original function $f(y(\cdot))$ is replaced by the model function $m_k(u_k + \cdot) = f(y_{POD,u_k}(\cdot))$.

Numerical examples are given for an example of the control of a driven cavity flow problem.

References

[1] K. Afanasiev, M. Hinze, Adaptive control of a wake flow using proper orthogonal decomposition, Lecture Notes in Pure and Applied Mathematics 216 (2001).

[2] E. Arian, M. Fahl, and E.W. Sachs, Trust-region proper orthogonal decomposition for flow control, ICASE Report 2000-25, ICASE, NASA Langley Research Center, Hampton, 2000.
[3] E. Arian, M. Fahl and E.W. Sachs, Managing POD Models by optimization methods, IEEE CDC Conference Proceedings, Las Vegas, 2002.
[4] P.L. Toint, Global convergence of a class of trust region methods for nonconvex minimization in Hilbert space, IMA J. Num. Anal., 8 (2) 231-252 (1988).

ODE Concepts for PDE Optimization

Volker H. Schulz
(joint work with S. Hazra (U Trier))

In this talk, we exploit algorithmic concepts from the ODE world for the solution of optimization problems with PDE constraints. In particular, methodological results for two specific application problems are reported.

In the first application problem, we solve parameter identification problems for instationary multiphase flow in the subsurface. Multiple shooting in combination with a reduced Gauss-Newton approach, due to Bock and Schloeder (1981-), yields an efficient and robust algorithm with low storage requirements (cf. [1] and several subsequent papers).

In the second application problem, we study shape optimization for the design of parts of the surface of airplanes under drag optimization. First results regarding a new collaborative project together with DLR Braunschweig, Airbus Germany, EADS and others, which has started recently, are reported. The algorithmic workhorse is a generalization of reduced SQP techniques to continuous reduced SQP techniques in a pseudo-timestepping framework.

Both applications show that one can profit from knowledge of ODE concepts for optimization problems also in a PDE context.

References

[1] S. Hazra, V. Schulz, Numerical identification in multiphase flow through porous media, Comput. Vis. Sci. 5:107-113 (2002).
[2] S. Hazra, V. Schulz, Simultaneous pseudo-timestepping for PDE-model based optimization problems, Report 02-23, U Trier, 2002, submitted to BIT.
[3] S. Hazra, V. Schulz, Continuous reduced SQP methods for airwing design, in preparation.

On Optimal Control Problems with Pointwise State Constraints

Fredi Tröltzsch

We consider the parabolic optimal control problem with pointwise constraints on the control and the state,

$$\min J(y,u) = \frac{1}{2}\,\|y(T) - y_d\|^2_{L^2(\omega)} + \frac{\nu}{2}\,\|u\|^2_{L^2(Q)}$$

subject to

$$y_t - \Delta y + d(y) = u \qquad \text{in } Q = \Omega \times (0,T)$$

with homogeneous Neumann boundary condition and initial condition $y(0) = y_0$. Moreover, pointwise constraints are imposed on the control $u \in L^\infty(Q)$ and the state $y \in Y := W(0,T) \cap C(\bar{Q})$,

$$\begin{aligned} 0 &\le u(x,t) \le b(x,t) \\ 0 &\le c(x,t) + \gamma\, y(x,t), \end{aligned}$$

$(x,t) \in Q$. The following data are given: A bounded domain $\Omega \subset \mathbb{R}^N$ with $C^{1,1}$-boundary Γ, functions y_d, $y_0 \in C(\bar{\Omega})$, b, $c \in C(\bar{Q}), d \in C^{2,1}(\mathbb{R})$ with $d'(y) \ge 0$, and $\gamma > 0$, $\nu > 0$. For all $u \in L^\infty(Q)$, a unique state $y = y(u)$ exists in Y. Let $\bar{u}$ be locally optimal in the sense of $L^\infty(Q)$ and define $\bar{y} = y(\bar{u})$. Then the associated first order necessary optimality conditions can be formulated on using the Lagrange functional

$$\mathcal{L}(y,u,p,\mu) = J(y,u) - \int\limits_Q [(y_t + d(y) - u)p + \nabla y \cdot \nabla p]\, dxdt - \int\limits_Q (y+c)\, d\mu(x,t),$$

where p stands for the adjoint state and $\mu \in C(\bar{Q})^\star$ is the Lagrange multiplier associated with the state constraints. If a constraint-qualification is satisfied at $(\bar{y}, \bar{u})$, then a non-negative Borel measure μ exists such that the first order necessary optimality conditions

$$\begin{aligned} \tfrac{\partial \mathcal{L}}{\partial y}(\bar{y},\bar{u},p,\mu)\, y &= 0 \quad \forall\, y \in W(0,T) \text{ with } y(0) = 0 \\ \tfrac{\partial \mathcal{L}}{\partial u}(\bar{y},\bar{u},p,\mu)(u-\bar{u}) &\ge 0 \quad \forall u \in L^\infty(\Omega) \text{ with } 0 \le u \le b \\ \int\limits_Q (y+c)\, d\mu &= 0 \end{aligned}$$

are satisfied. This follows from results [2] and [5]. Assume conversely that a pair $(\bar{y}, \bar{u})$ is given that satisfies all constraints and the first-order necessary conditions. One might expect that the following standard condition is sufficient for local optimality of $\bar{u}$:
(SSC) There is $\delta > 0$ such that

$$\mathcal{L}''(\bar{y},\bar{u},p,\mu)(y,u)^2 = \|y(T)\|^2_{L^2(\Omega)} + \nu \|u\|^2_{L^2(Q)} - \int\limits_Q d''(\bar{y}) p\, y^2\, dxdt \ge \delta\, \|u\|^2_{L^2(Q)}$$

holds for all pairs $(y,u) \in Y \times L^\infty(Q)$ satisfying the state equation linearized at $(\bar{y}, \bar{u})$.

Then $\bar{u}$ is expected to be locally optimal with respect to the L^2-topology. This holds true, if the mapping $u \mapsto y(u)$ is continuous from L^2 to $C(\bar{Q})$, i.e. for $N = 1$ in our parabolic example with distributed control. For the case of Neumann boundary control, this is not true.

If, however, the state-constraints are deleted, then (SSC) is sufficient for local optimality in the L^∞-topology for all N.

Unfortunately, we have not been able to prove the same result with pointwise state constraints. Local optimality can still be shown in the sense of L^∞ for $N = 2$. For $N > 2$ we cannot prove that (SSC) is really sufficient for local optimality. The obstacle is the need to estimate $\mathcal{L}''(\bar{y},\bar{u},p,\mu)(y,u)^2$ against the L^2-norm of u. Due to the presence of the measure μ in the right-hand side of the adjoint equation, p is not bounded in the case of state-constraints. Therefore, the estimation of the third quantity in the expression for $\mathcal{L}''$ causes troubles.

The situation is even worse, if (SSC) is weakened by considering also strongly active state-constraints. Then only for $N = 1$ the local optimality can be shown in the case of distributed control, while boundary controls cannot be handled at all. We refer to [6]. The results are slightly better for elliptic problems. L^2-optimality can be shown for $N \leq 3$ and distributed control (since $H^2(\Omega) \subset C(\bar{\Omega})$) and $N = 2$ for Neumann boundary control ($H^{3/2-\varepsilon}(\Omega) \subset C(\bar{\Omega})$), [3]. To overcome these difficulties, a Lavrentiev type regularization is suggested. Consider the mixed-pointwise control-state constraints

$$-\varepsilon\, u(x,t) \leq c(x,t) + \gamma\, y(x,t).$$

In this case, the existence of an associated regular Lagrange multiplier $\mu \in L^\infty(Q)$ can be shown, [1]. Moreover, this concept is useful for numerical approximations. For $\varepsilon \downarrow 0$, the associated optimal control converges to $\bar{u}$. This is demonstrated for linear-quadratic elliptic problems with $N = 2$, [4].

References

[1] M. Bergounioux, F. Tröltzsch, Optimal control of semilinear parabolic equations with state-constraints of bottleneck type, ESAIM, Control, Optimisation and Calculus of Variations, 4 (1999), 595–608.

[2] E. Casas, Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations, SIAM J. Control Optimization 35 (1997), 1297–1327.

[3] E. Casas, F. Tröltzsch, and A. Unger, Second order sufficient optimality conditions for some state-constrained control problems of semilinear elliptic equations, SIAM J. Control and Optimization, 38 (2000), 1369–1391.

[4] C. Meyer, A. Rösch, and F. Tröltzsch, Optimal control problems of PDEs with regularized pointwise state constraints, Tech. Report 14-2003, Inst. of Math., TU Berlin. Submitted.

[5] J.-P. Raymond, H. Zidani, Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations, Applied Mathematics and Optimization 39 (1999), 143–177.

[6] J.-P. Raymond and F. Tröltzsch, Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints, Discrete and Continuous Dynamical Systems, 6 (2000), 431–450.

Optimal Control of Nonlinear Parabolic Systems by Second Order Methods

Stefan Volkwein

In the talk three different optimal control problems for nonlinear parabolic systems are considered.

The first example is concerned with optimal boundary control of an instationary reaction-diffusion system in three spatial dimensions. This problem involves a coupled nonlinear system of parabolic differential equations with bilateral as well as integral control constraints. We include the integral constraint in the cost by a penalty term whereas the bilateral control constraints are handled explicitly. A primal-dual active set strategy is utilized to compute optimal solutions numerically. The algorithm is compared to a semi-smooth Newton method.

As a second example an optimal control boundary problem for the Stefan problem is considered. Here, an inexact Newton method is applied with quasi-Newton approximations for the Hessian. To ensure positivity of the Hessian, a line search based on the Wolfe-Powell conditions is utilized.

Finally, laser surface hardening of steel is formulated in terms of an optimal control problem, where the state equations are a semilinear heat equation and an ordinary differential equation, which describes the evolution of the high temperature phase. To avoid the melting of the steel we have to impose state constraints for the temperature. Including the state constraints into the cost functional by a penalty term, a globalized SQP method with a reduced Hessian is applied to solve the control problem numerically. To ensure the convergence of the algorithm a numerically inexpensive globalization strategy is used.

References

[1] R. Griesse, S. Volkwein, Analysis for optimal boundary control for a three-dimensional reaction-diffusion system, Technical Report No. 286, Special Research Center F 003 Optimization and Control, Project area Continuous Optimization and Control, University of Graz & Technical University of Graz, Dezember 2003.

[2] M. Hintermüller, K. Ito, and K. Kunisch, The primal-dual active set strategy as a semismooth Newton method, SIAM J. Optimization, 13:865-888, 2003.

[3] K. Ito, K. Kunisch, The primal-dual active set method for nonlinear problems with bilaterally constraints, Technical Report No. 214, Special Research Center F 003 Optimization and Control, Project area Continuous Optimization and Control, University of Graz & Technical University of Graz, 2001, to appear in SIAM J. Control and Optimization.

[4] M. Ulbrich, Semismooth Newton methods for operator equations in function spaces, SIAM J. Optimization, 13:805-842, 2003.

[5] S. Volkwein, A globalized SQP method for the optimal control of laser surface hardening, Technical Report No. 272, Special Research Center F 003 Optimization and Control, Project area Continuous Optimization and Control, University of Graz & Technical University of Graz, June 2003, submitted.

Interior–Point Methods for Optimization Problems with PDE Constraints

Beate Winkelmann

(joint work with R.E. Bank and Ph.E. Gill)

Methods are proposed for the numerical solution of optimal control problems with partial differential equation (PDE) constraints and inequality constraints on the control variable. The general form of the problem is:

$$\underset{y,u}{\text{minimize}}\ \rho(y,u) = \int_{\Omega} p_1(y, \nabla y, u)\, dx + \int_{\partial\Omega} p_2(y, \nabla y, u)\, ds$$

$$\begin{aligned} \text{subject to }\ &\langle E(y,u), v\rangle = 0, \quad \forall v \in H,\\ &\underline{u} \le u(x) \le \overline{u}, \end{aligned}$$

where $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2\}$, y is the state variable, u is the control variable, H is an appropriate function space, and $\langle E(y,u), v\rangle = 0$ is the weak form of a partial differential equation in divergence form. State and control variables are discretized using an adaptive finite-element approach. Algorithms for optimization and PDEs are combined to solve a discretized optimization problem over a sequence of adaptive meshes.

An interior-point method is used for the optimization part of the algorithm. The two main types of interior-point method are *primal methods* and *primal-dual methods*. As the names suggest, primal methods iterate over the primal variables only, while primal-dual methods iterate over the primal and dual variables simultaneously. However, both methods approach the solution by following a continuous path that approaches the solution from the interior of the set of admissible solutions. For a primal method the path is the trajectory of solutions of a sequence of equality-constrained problems parameterized by a scalar μ. Primal-dual methods define the path as the trajectory of points satisfying the perturbed first-order optimality conditions for the constrained problem.

Primal-dual methods are usually preferred for general constrained optimization because of their rapid convergence near the solution. However, in the PDE context, the use of a primal-dual method requires the adaptive discretization of both the primal and dual variables, which can lead to serious numerical difficulties if the dual variables are not sufficiently smooth. It is shown that these difficulties may be avoided if the primal method is implemented using an extrapolation scheme for the parameter μ.

Regardless of the particular choice of interior-point method, the linear systems to be solved at each iteration are large, symmetric and have PDE-like structure. These systems also become increasingly ill-conditioned as the solution is approached. In order to handle the size and sparsity of these systems, a preconditioned Krylov-space method is used. The choice of preconditioner is crucial for the performance of the optimization since the cost of solving the linear system dominates the overall cost of the computation. A good preconditioner lowers the cost of solving the linear systems significantly. The aim is to incorporate as much information as possible into the preconditioner without dramatically increasing the cost of the computation. To this end, the preconditioner has the same block structure as the original system and existing multilevel PDE preconditioners are used for some of the blocks. In particular, an algebraic multigrid preconditioner with

ILU smoothing is used for the PDE constraint blocks, and a symmetric Gauss-Seidel preconditioner is used for the control block and the full linear system. The preconditioner is fully parallel.

The PDE part of the algorithm uses adaptive mesh refinement based on an *a posteriori* hierarchical basis error estimator for the state variables. The path-following parameter μ and the PDE parameters are chosen to allow the discretization error and optimization error to go to zero at the same rate. An error estimator based on the state variable allows the mesh to be adaptively refined and unrefined without the additional cost of solving the adjoint equation.

These ideas are illustrated in the context of the elliptic finite-element PDE package PLTMG. Numerical results are presented for a particular optimal control problem involving an elliptic PDE constraint. The adaptive refinement algorithm requires the solution of an optimization problem with up to 22.5M variables. This problem was solved on a 256 processor Beowulf cluster in approximately 8 minutes.

References

[1] R.E. Bank, M. Holst, A new paradigm for parallel adaptive mesh refinement, SIAM J. Sci. and Statist. Comput., 22(4):1411–1443, 2000.

[2] R.E. Bank, PLTMG: A Software Package for Solving Elliptic Partial Differential Equations, Users' Guide 8.0, Software, Environments and Tools, Vol. 5, SIAM, Philadelphia, 1998.

[3] R. E. Bank, P. E. Gill, R. F. Marcia, Interior methods for a class of elliptic variational inequalities, In L.T. Biegler et al. (eds.), Large-scale PDE-constrained Optimization, vol. 30, Lecture Notes in Computational Science and Engineering, 218–235, Berlin, Heidelberg and New York, 2003. Springer-Verlag.

[4] A. Forsgren, P.E. Gill, M.H. Wright, Interior methods for nonlinear optimization, SIAM Rev., 44:525–597, 2002.

[5] B. Winkelmann, Interior-point Methods and their Application to Optimal Control with PDE Constraints, University of California San Diego, 2004, Ph.D. Thesis, in preparation.

Reporter: Angela Kunoth

Participants

Dr. Roland Becker
roland.becker@univ-pau.fr
roland.becker@iwr.uni-heidelberg.de
Laboratoire de Mathematiques
Appliquees
Universite de Pau et des Pays de l'Adour
BP 1155
F-64013 Pau Cedex

Dr. George Biros
biros@seas.upenn.edu
Dept. of Computer and Information
Science, University of Pennsylvania
220 Towne Building
220 S. 33rd. St.
Philadelphia PA 19104-6315 – USA

Prof. Dr. Hans Georg Bock
scicom@iwr.uni-heidelberg.de
Bock@IWR.Uni-Heidelberg.De
Interdisziplinäres Zentrum
für Wissenschaftliches Rechnen
Universität Heidelberg
Im Neuenheimer Feld 368
D–69120 Heidelberg

Dr. Alfio Borzi
alfio.borzi@uni-graz.at
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstr. 36
A-8010 Graz

Carsten Burstedde
burstedde@iam.uni-bonn.de
Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 6
D–53115 Bonn

Prof. Omar Ghattas
oghattas.@cs.cmu.edu
Carnegie Mellon University
Department of Civil and
Environmental Engineering
Pittsburgh PA 15213-3890 – USA

Prof. Dr. Andreas Griewank
griewank@math.hu-berlin.de
Institut für Mathematik
Humboldt-Universität Berlin
Unter den Linden 6
D–10099 Berlin

Prof. Dr. Max D. Gunzburger
gunzburg@csit.fsu.edu
CSIT/DSL 400
School of Computational Science and
Information Technology
Florida State University
Tallahassee FL 32306-4120 – USA

Dr. Vincent Heuveline
vincent.heuveline@iwr.uni-heidelberg.de
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294
D–69120 Heidelberg

Prof. Dr. Michael Hinze
hinze@math.tu-dresden.de
TU Dresden
Institut für Numerische Mathematik
Willersbau C 312
D–01062 Dresden

Prof. Dr. Claes Johnson
claes@math.chalmers.se
Department of Mathematics
Chalmers University of Technology
S-412 96 Göteborg

Prof. Dr. Karl Kunisch
karl.kunisch@uni-graz.at
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstr. 36
A-8010 Graz

Prof. Dr. Angela Kunoth
kunoth@iam.uni-bonn.de
Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 6
53115 Bonn

Prof. Dr. Günter Leugering
leugering@am.uni-erlangen.de
Lehrst. f. Angewandte Mathematik II
Universität Erlangen-Nürnberg
Martensstr. 3
D–91058 Erlangen

Daniel Oeltz
oeltz@iam.uni-bonn.de
Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 6
D–53115 Bonn

Prof. Dr. Rolf Rannacher
rannacher@iwr.uni-heidelberg.de
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294
D–69120 Heidelberg

Prof. Dr. Ekkehard Sachs
sachs@uni-trier.de
Abteilung Mathematik
Fachbereich IV
Universität Trier
D–54286 Trier

Prof. Dr. Volker Schulz
Volker.Schulz@uni-trier.de
Fachbereich IV - Mathematik
Numerik - Optimierung, partielle
Differentialgleichungen
Universität Trier
D–54286 Trier

Prof. Dr. Fredi Tröltzsch
troeltzsch@math.tu-berlin.de
Institut für Mathematik
Technische Universität Berlin
Sekr. MA 4-5
Strasse des 17. Juni 136
D–10623 Berlin

Doz.Dr. Stefan Volkwein
stefan.volkwein@uni-graz.at
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstr. 36
A-8010 Graz

Beate Winkelmann
bmwinkelmann@ucsd.edu
Dept. of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0112 – USA

Mathematisches Forschungsinstitut Oberwolfach

Report No. 4/2004

Mini-Workshop: Multiscale Modelling in Epitaxial Growth

Organised by
Axel Voigt (Bonn)

January 18th – January 24th, 2004

Introduction by the Organisers

Thin film epitaxy is a modern technology of growing single crystals that inherit atomic structures from substrates. Various mathematical models and numerical algorithms are proposed to be used for describing epitaxial growth processes. Due to the underlying multiscale phenomena, which range from the interaction of single atoms at steps up to an engineering scale, on which the transport of material to the surface in a MBE (molecular beam epitaxy) furnace needs to be described, the models can be distinguished by the relevant length scales they are living on

(a) discrete atomic models: Individual atoms are the basic degrees of freedom and single hoppings to neighbouring lattice sites are simulated by kinetic Monte Carlo methods. A n example are the so-called Solid-on-Solid models.
(b) discrete-continuous models: The atomic distance in the growth direction is discrete, but the atomic distance in the lateral direction is coarse grained. The steps are assumed to be smooth curves and serve as free boundaries for an adatom diffusion equation on terraces. These models are known as Burton-Cabrera-Frank models.
(c) continuous models: The atomic processes at steps are neglected , the overall surface is assumed to be smooth and phenomenological equations describe directly the height of the growing film. An example is the Villain equation.

The main goal in modelling epitaxial growth is to bridge the gap between these different models and to describe growth process on a continuous scale by incorporating atomic effects. The focus of this workshop was to bring together materials scientists, theoretical physicists and applied mathematicians to exchange ideas on the three different regimes (a),(b) and (c). The mini-workshop consisted of three introducing lectures, one for each approach and several lectures which focus on connections both in an analytical and numerical fashion. The contributions ranged from quantum-chemistry, molecular dynamics and kinetic Monte Carlo to step flow and continuum models. Several multiscale approaches have been considered to combine at least two of these models.

Besides the mathematical aspects of modelling epitaxial growth also the connection to experimental results was dealt with in order to drive the recent theoretical developments into a direction which is relevant for a large variety of industrial applications. The mini-workshop was also used to give young researchers the opportunity to be introduced into such an actual interdisciplinary field.

Multiscale Modelling in Epitaxial Growth

Table of Contents

Abstracts

Fronttracking for Epitaxial Growth by a Cellular Automaton Algorithm

Rainer Backofen
(joint work with Frank Haußer, Axel Voigt)

The description of steps is a crucial part in the numerical modelling of step flow or island growth. There are a lot of different methods, e.g. level sets, parametric finite elements or phasefield methods. Each of them have their advantages and disadvantages. But for all this methods faceting of the steps or strong anisotropic properties are numerically hard to treat[1].
In cases of strong anisotropic growth laws Gandin and Rappaz [1] introduced a cellular automata (CA) algorithm for the description of grain growth in metallurgical solidification.We present first steps to adapt this algorithm to epitaxial island growth.

In the presence of strong anisotropies the growth of an island is limited by its slowest growing directions. In these directions the island form facets. A complete representation of such an island is given by the normal directions of the facets and their distances from a central point. A set of facet normals together with a distance to a point defines an evolution element. In order to take into account local effects such as island impingement or spatially varying growth velocities, a local description of the island is needed. Thus the step or island boundary is approximated by a set of local evolution elements defined at points near the step. In figure 1 the evolution algorithm is shown.

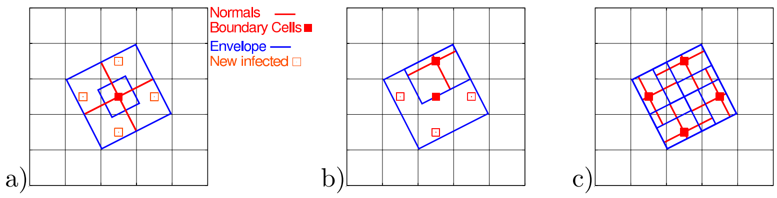


FIGURE 1. a) An island is nucleated in a cell and successively grows until adjacent cells are inside the envelope. b) Adjacent cells are infected, that is, the island is represented by a evolution element defined at the new infected cell. c) The whole island is now locally defined by evolution elements defined in cells nearby

The CA uses a regular structured grid. Every grid cell has a state, ξ_i : terrace, no terrace or boundary. In the boundary cells are additionally evolution elements defined to track the step.

[1] at least compared to a isotropic case

The driving force for the step evolution is derived by a standard Burton-Cabrera-Frank (BCF) type model [1]. The adatom density ρ_i at the terrace Ω_i of atomic height is described by the diffusion equation

$$\rho_i - D\Delta\rho_i \quad = \quad F - \tau^{-1}\rho_i \quad \text{in} \quad \Omega_i(t) \tag{1}$$

where F and $\tau^{-1}\rho_i$ model the flux onto the surface and evaporation. The steps (terrace boundaries) Γ_i are free boundaries with normal velocity v_i governed by the adatom fluxes toward the steps and edge diffusion.

$$v_i \quad = \quad -D\nabla\rho_i \cdot \vec{n_i} - \rho v_i + D\nabla\rho_{i-1} \cdot \vec{n_i} + \nu\partial_{ss}\kappa_i \tag{2}$$

Until now we use thermodynamic boundary conditions at the steps

$$\rho_i \quad = \rho_i \quad = \rho_{eq}(1 + \mu\kappa_i) \quad \text{on } \Omega_i \tag{3}$$

To solve this set of equations, we use a operator splitting approach, as in Bänsch et al. [2]. The step evolution is modelled with the CA, which is coupled to the FEM algorithm for adatom diffusion, see figure 2.

As a first test case a circular island is treated. To approximate an isotropic situation evolution elements with 90 facets are used, see figure 3 a),b).

The instabilities of the island growth is triggered by the approximation of the step by a polygon. For a slightly tilted evolution element with five facets the instabilities are clearly caused by the prescribed anisotropy of the growth algorithm,see figure 3 c).

The next major step will be to connect the anisotropy of the growth algorithm to physical situations.

Another important issue is the effective derivation of a smooth and nearly equidistant polygon from the CA description of the island. Since the implementation of topological changes and incorporation of faceting is very natural, the algorithm seems to be worth further considerations.

References

[1] Ch.-A. Gandin and M. Rappaz, *A coupled finite element-cellular automaton model for the prediction of dendritic grain structures in solidification processes*, Acta Metal. and Mat., 1994, 42.

[2] Bänsch, E. and Haußer, F. and Voigt, A., *Finite element method for epitaxial growth with thermodynamic boundary conditions*,SIAM J. Sci. Comput.,2003 (submitted)

[3] Burton, W.K. and Cabrera, N. and Frank, F.C., *The growth of crystals and the equilibrium structure of their surfaces*,Phil. Trans. Roy. Soc. London Ser. A, 1951, 243(866)

Lattice Gas Models and Kinetic Monte Carlo Simulation of Epitaxial Crystal Growth

Michael Biehl

A brief introduction is given to the Kinetic Monte Carlo (KMC) simulation of epitaxial crystal growth. Molecular Beam Epitaxy (MBE) serves as a particularly clear-cut prototype situation, but many of the aspects discussed here would carry

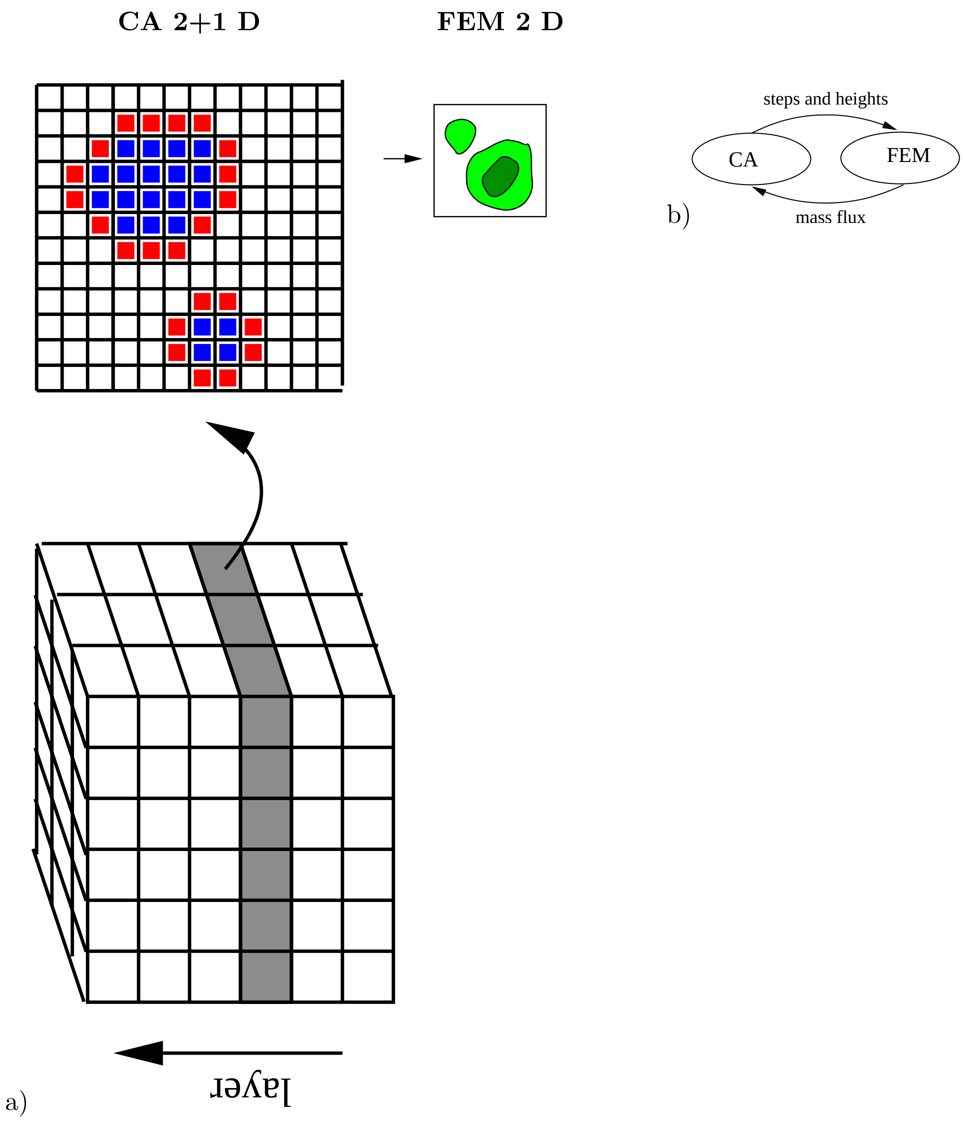


FIGURE 2. a) For each atomic layer a CA grid is defined. In each layer the islands of the corresponding height are defined. The steps are then transferred as polygons to the FEM calculation of adatom diffusion, the FEM algorithm then calculates the growth velocity of the steps, b).

over to other techniques. MBE has become a standard experimental setup for the production of high quality crystals, such as thin magnetic films or nano-scale

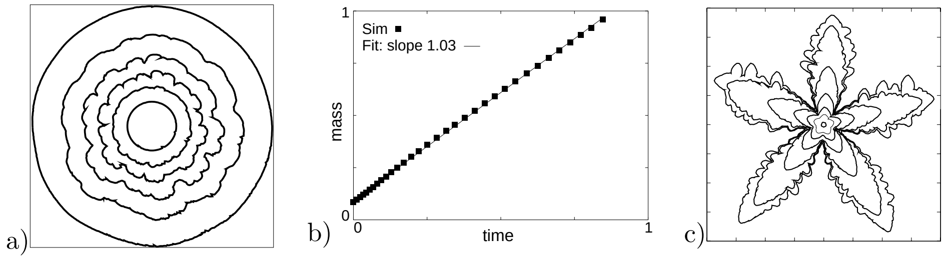


FIGURE 3. a) Envelope of the island. Curvature effects are not taken into account, so the growth is instable for small islands. The instability is driven by fluctuations in the definition of the steps. b) Overall mass conservation is 3%. c) Five sided evolution element. The instability is triggered by the anisotropy of the evolution element. (mass conservation $\approx 10\%$) .

semiconductor structures. At the same time it provides a framework in which to develop theoretical and computational concepts for the description of growth and more general non-equilibrium processes.

Different approaches to the modelling and simulation of MBE and similar growth techniques have been applied. They range from the full microscopic quantum mechanics treatment of the dynamics to the coarse grained description in terms of, for instance, stochastic differential equations. Here, the focus will be on discrete models such as lattice gas and Solid-On-Solid (SOS) models and the corresponding Kinetic Monte Carlo techniques. Various levels of simplification or sophistication have been employed in this context, depending on the precise goal of the investigation.

This contribution is far from giving an exhaustive review of the field. It is intended to provide a brief discussion of the basic concepts of KMC simulations and their strengths and limitations in the modelling of crystal growth processes. The following example books and review articles give a detailed and more complete overview of, both, the physics of epitaxial growth and the KMC method. They also provide plenty of further references.

A.-L. Barabasy and H.E. Stanley, *Fractal concepts in surface growth* Cambridge University Press, Cambridge (UK) 1995.

A. Pimpinelli and J. Villain, *Physics of Crystal Growth* Cambridge University Press, 1998.

M.E.J. Newman and G.T. Barkema, *Monte Carlo methods in statistical physics* Clarendon Press, Oxford 1999.

P. Politi, G. Grenet, A. Marty, A. Ponchet, and J. Villain, *Instabilities in crystal growth by atomic or molecular beams*, *Physics Reports* **324** (2000) 271.

M. Kotrla, N.I. Pananicolaou, D.D. Vvedensky, and L.T. Wille (eds.), Proc. of the NATO Advanced Research Worshop on *Atomistic Aspects of Epitaxial Growth* Kluwer, Dordrecht 2001.

T. Michely and J. Krug, *Islands, mounds and atoms*
Springer, Heidelberg 2004.

Off-lattice Kinetic Monte Carlo Simulation of strained hetero-epitaxial growth

Michael Biehl

An off-lattice, continuous space Kinetic Monte Carlo algorithm is introduced and discussed [1, 2, 3, 4], which allows to study various phenomena known from strained, hetero-epitaxial crystal growth [5, 6].

As a starting point, we study a simplifying, 1+1 dimensional model with Lennard-Jones interactions. It exhibits, for instance, the appearance of misfit dislocations at a characteristic layer thickness [6, 3].

The focus of this talk is on the appearance of strain induced multilayer islands or *dots* upon a persisting wetting layer, i.e. the so-called Stranski-Krastanow growth mode [5, 7, 8, 9]. The transition from monolayer to multilayer islands occurs at a critical film thickness. Its dependence on the model parameters (lattice misfit, growth rate, and temperature) is investigated quantitatively. We find that for sufficiently large deposition rates the properties of the mounds is governed by the lattice mismatch only [8, 9].

The method is also applied in the context of surface alloy formation of immiscible metals on appropriate substrates. Two competing mechanisms for the emergence of nano-scale stripe structures are investigated [10].

References

[1] M. Schroeder and D.E. Wolf, *Diffusion on strained surfaces* , Surf. Sci. **375** (1997) 129.
[2] A. Schindler, *Theoretical aspects of growth on one and two dimensional strained surfaces*, Ph.D. thesis, Gerhard-Mercator-Universität Duisburg, Duisburg 1999.
[3] F. Much, M. Ahr, M. Biehl, and W. Kinzel, *Kinetic Monte Carlo simulations of dislocations in heteroepitaxial growth*, Europhys. Lett. **56** (2001) 791.
[4] F. Much, M. Ahr, M. Biehl, and W. Kinzel, *A Kinetic Monte Carlo method for the simulation of heteroepitaxial growth*, Comp. Phys. Comm. **147** (2002) 107.
[5] A. Pimpinelli and J. Villain, *Physics of Crystal Growth*, Cambridge University Press, Cambridge (UK) 1998.
[6] P. Politi, G. Grenet, A. Marty, A. Ponchet, and J. Villain, sl Instabilities in crystal growth by atomic or molecular beams, Phys. Rep. bf 324 (2000) 271.
[7] B. Joyce, P. Kelires, A. Naumovets, and D.D. Vvedensky (eds.), Proceedings of the NATO Advanced Research Workshop on *Quantum Dots: Fundamentals, Applications, and Frontiers*, Kluwer, Dordrecht, in press.

[8] F. Much and M. Biehl, *Simulation of wetting-layer and island formation in heteroepitaxial growth*, Europhys. Lett. **63** (2003) 14.
[9] M. Biehl and F. Much, *Off-lattice Kinetic Monte Carlo simulations of Stranski-Krastanov-like growth*, to appear in [7].
[10] M. Kotrla, T. Volkmann, F. Much, and M. Biehl, *Mechanisms of formation of self assembled nanostructures in heteroepitaxy*, in: *Proceedings of NANO '03, Brno/Cz*, in press.

A Multiscale Approach to the Modelling of Chemical Vapor Deposition

Carlo Cavallotti

Computational material science of thin solid films has undergone great advancements in the last years. Significant progress has been made not only in the prediction and description of the surface and bulk properties of the materials, but also, from an engineering point of view, in the comprehension of the influence that the operating conditions of the growth process have on the desired material properties. An approach that has recently proved successful in the description of the thin film deposition processes is the multiscale modeling approach. It is based on the fact that growth of materials with well controlled morphological and compositional properties is a processes complicated by chemical and physical phenomena that occurs on time and length scales that can differ even by several orders of magnitude. I present a multiscale approach that has been developed to investigate the Chemical Vapor Deposition of epitaxial thin films at different time and length scales. The multiscale approach here outlined is designed to investigate the influence that gas phase and surface reactions have on the morphological and compositional evolution of thin solid films deposited by chemical vapor deposition. Atomic scale energetic and kinetic parameters, when not available from the literature, are estimated by means of quantum chemistry computations. The local gas phase composition, fluid dynamic and thermal fields are evaluated by integration of mass, energy and momentum equations at the reactor scale using kinetic and thermodynamic data calculated with quantum chemistry. The morphology of the film is finally investigated using 3 dimensional Kinetic Monte Carlo, which inputs are the gas phase fluxes calculated at the reactor scale and the kinetic parameters determined at the atomic scale. The calculation of kinetic parameters for CVD processes by means of quantum chemistry is usually performed by means of density functional theory (DFT). DFT calculations can be essentially of two different types, depending on the choice of the basis set between plane waves and Gaussian basis functions. While the first type of calculations has the advantage of treating more correctly systems with delocalized electrons, such as metals, the second offers the possibility to systematically increase the dimension of the basis set used for the calculations, and thus describe in higher detail the electronic density distribution. Since our analysis is focused mainly on semiconductors, in which electrons are usually localized within covalent bonds and atomic orbitals, we choose to perform our atomic scale calculations using gaussian basis functions with gradient corrected functionals, such as B3LYP [1]. Surface processes are studied using clusters of different dimension to represent the surface structure. We choose to investigate the morphology evolution of the thin solid films with 3 Dimensional Kinetic Monte Carlo, that has the advantage over other mesoscale models to require as inputs kinetic constants or diffusion parameters that can be directly calculated by means of quantum chemistry. Our implementation of KMC follows the theory outlined by Weinberg [2], with direct tracking of real time and a rejection free choice of the random transition. The starting conditions of the KMC simulation are the surface structure at time 0, the surface temperature and the fluxes of gas phase species towards the surface. The output of a KMC simulation consists in the detailed surface morphology of the film after the deposition of a certain amount of layers. It is thus possible to determine the growth regime of the film, be it 3 dimensional, terrace step flow or 2 dimensional. The reactor scale modeling of CVD processes can nowadays be considered as a mature field. Several commercial CFD codes dedicated to CVD are in fact available and have been tested in many different occasions against experimental data. However, being the focus of our research the integrated multiscale modeling of the CVD process, we still rely on the use of our codes when the intent is that of linking together consistently KMC and reactor scale models [3]. The multiscale approach here proposed was used to investigate the epitaxial CVD of Si and ZnSe and the selective Metal Organic CVD of AlGaAs and InP [4, 5]. The results of the calculations were compared with experimental data with the aim of improving our understanding of the growth process.

References

[1] M. J. Frisch et al., G98, Revision A.6; Gaussian, Inc., Pittsburgh, PA, 1998.
[2] K. A. Fichthorn, W. H. Weinberg, J. Chem. Phys. **95** (1991) 1090.
[3] C. Cavallotti, V. Gupta, C. Sieber, and K.F. Jensen, Phys. Chem. Chem. Phys. **5** (2003) 2818.
[4] C.Cavallotti, D. Moscatelli, and S. Carr, J. Phys. Chem A **108** (2004) 1214.
[5] C. Cavallotti, M. Nemirovskaya, and K.F. Jensen, J. Crystal Growth **248** (2003) 411.

Configurational continuum modelling of crystal surface evolution

Navot Israeli
(joint work with Daniel Kandel)

We propose a novel approach to continuum modelling of dynamics of crystal surfaces. Our model follows the evolution of an ensemble of step configurations, which are consistent with the macroscopic surface profile. Contrary to the usual approach where the continuum limit is achieved when typical surface features consist of many steps, our continuum limit is approached when the number of step configurations of the ensemble is very large. The model is capable of handling singular surface structures such as corners and facets and has a clear computational advantage over discrete models.

Thermal decay and Ostwald ripening in homoepitaxy

Frank Haußer
(joint work with Axel Voigt)

The thermal relaxation of isolated (single layer) homoepitaxial islands and craters and of isolated nanomounds is simulated using a 2+1 dimensional step flow model. Numerical simulations based on adaptive finite elements are used to study decay rates of these structures in the diffusion limited and attachment-detachment limited regime under the influence of anisotropic effects.

Configurational continuum modelling of crystal surface evolution

Navot Israeli
(joint work with Daniel Kandel)

The behavior of classical physical systems is typically described in terms of equations of motion for discrete microscopic objects (e.g. atoms). The dynamics of the microscopic objects is usually very erratic and complex. Nevertheless, in many cases a smooth behavior emerges when the system is observed on macroscopic length and time scales (e.g. in fluid flow through a pipe). A fundamental problem in physics is to understand the emergence of the smooth macroscopic behavior of a system starting from its microscopic description. A useful way to address this problem is to construct a continuum, coarse-grained model, which treats the dynamics of the macroscopic, smoothly varying, degrees of freedom rather than the microscopic ones. The derivation of continuum models from the microscopic dynamics is far from trivial. In most cases it is done in a phenomenological manner by introducing various uncontrolled approximations.

In this work we address the above problem in the context of the dynamics of crystal surfaces. The evolution of crystal surfaces below the roughening transition proceeds by the motion of discrete atomic steps which are separated by high symmetry orientation terraces. One can model step motion by solving the diffusion problem of adatoms on the terraces with appropriate boundary conditions at step edges. This approach was introduced long ago by Burton, Cabrera and Frank [1], and was further developed by other authors [2]. The resulting models are capable of describing surface evolution on the mesoscopic scale with significant success [3, 4]. However, such models pose a serious challenge for numerical computations, and can be solved only for small systems.

Several attempts were made to construct continuum models for stepped surfaces [5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17], in order to understand their large scale properties. The general idea behind these attempts is that step flow can be treated continuously in regions where every morphological surface feature is composed of many steps. If we label surface steps by the index n, the continuum limit in these models is obtained by taking n to be continuous. The outcome of these attempts are partial differential equations for surface evolution. Such continuum models are fairly successful in describing the evolution of smooth surfaces with very simple morphologies. However, they suffer from fundamental drawbacks, which do not allow generalizations to more complex and realistic situations.

The most severe drawback is that below the roughening temperature, crystal surfaces have singularities in the form of corners and macroscopic facets. The latter are a manifestation of the cusp singularity of the surface free energy at high symmetry crystal orientations. The assumption that every surface feature is composed of many steps clearly breaks down on macroscopic facets where there are no steps at all. Thus, existing continuum models fail conceptually near singular regions. Several authors have tried to overcome this problem by solving a continuum model only in the non-singular parts of the surface and then carefully match the boundary conditions at the singular points or lines[11, 12, 13, 14]. In most cases however it is not at all clear how these matching conditions can be derived. Another approach is to round the surface free energy cusp [15, 16, 17], replacing true facets by relatively flat but analytic regions. This method implicitly assumes that the surface free energy derived for non singular orientations determines the dynamics on facets as well. This assumption is often found to be false because steps near facet edges obey different dynamics than steps in the sloping parts of the surface[12, 13].

In this work we propose a conceptually new definition of the continuum limit, which we term Configurational-Continuum[18]. Configurational-Continuum allows construction of continuum models, which are free of all the limitations of standard continuum models discussed above. It provides a rigorous way of deriving the continuum model directly from the discrete step equations of motion. Like other continuum models, Configurational-Continuum has a clear computational advantage over the discrete step model due to the small number of discretization points it requires for the description of smooth surface regions in a numerical scheme.

Our key observation in deriving Configurational-Continuum is that a continuous surface height profile defines an *ensemble* of microscopic step configurations which are all consistent with the continuous profile. The continuous profile in this picture evolves as the upper envelope of the ensemble with each configuration obeying the microscopic step dynamics. We derive the envelope equation of motion in the continuum limit when the number of configurations in the ensemble is very large. In contrast to the situation in standard continuum models, this Configurational-Continuum limit is exact.

References

[1] W. K. Burton, N. Cabrera and F. C. Frank, Philos. Trans. R. Soc. London, Ser. A **243**, 299 (1951).

[2] For general reviews see H.-C. Jeong and E. D. Williams, Surf. Sci. Rep. **34**, 175 (1999); E. D. Williams, Surf. Sci. **299/300**, 502 (1994).

[3] E. S. Fu, M. D. Johnson, D.-J. Liu, J. D. Weeks and E. D. Williams, Phys. Rev. Lett. **77**, 1091 (1996).

[4] S. Tanaka, N. C. Bartelt, C. C. Umbach, R. M. Tromp and J. M. Blakely, Phys. Rev. Lett. **78**, 3342 (1997).

[5] W. W. Mullins, J. Appl. Phys. **28**, 333 (1957).

[6] M. Ozdemir and A. Zangwill, Phys. Rev. B **42**, 5013 (1990).
[7] P. Nozières, J. Phys. I France **48**, 1605 (1987).
[8] F. Lançon and J. Villain, Phys. Rev. Lett. **64**, 293 (1990).
[9] M. Uwaha, J. Phys. Soc. Jap. **57**, 1681 (1987).
[10] A. Chame, S. Rousset, H. P. Bonzel and J. Villain, Bul. Chem. Commun. **29**, 398 (1996/1997).
[11] J. Hager and H. Spohn, Surf. Sci. **324**, 365 (1995).
[12] N. Israeli and D. Kandel, Phys. Rev. B **62** 13707 (2000).
[13] N. Israeli and D. Kandel, Phys. Rev. Lett. **80** 3300 (1998); Phys. Rev. B **60** 5946 (1999).
[14] N. Israeli, H.-C. Jeong, D. Kandel and J. D. Weeks, Phys. Rev. B **61** 5698 (2000).
[15] H. P. Bonzel, E. Preuss and B. Steffen, Appl. Phys. A **35**, 1 (1984).
[16] H. P. Bonzel and E. Preuss, Surf. Sci. **336**, 209 (1995).
[17] M. V. Ramana Murty, Phys. Rev. B **62**, 17004 (2000).
[18] N. Israeli and D. Kandel, Phys. Rev. Lett. **88**, 116103-1 (2002).

Universality classes for step bunching?

Joachim Krug
(joint work with V. Tonchev, S. Stoyanov)

In a remarkable recent paper [1], Pimpinelli and coworkers proposed a classification of step bunching instabilities in terms of scaling exponents characterizing the shape of the bunches and the time evolution of their size. The scaling exponents α, β and γ are defined through the relations

$$N \sim W^{\alpha}, \quad l_{\min} \sim N^{-\gamma}, \quad N \sim L \sim t^{\beta} \tag{4}$$

between the number of steps N in a bunch, the width W of the bunch, the minimal terrace size $l_{\min}$ in the bunch and the spacing L between bunches. The second and third of theses scaling relations have been observed experimentally in electromigration-induced step bunching on surfaces vicinal to Si(111) [2, 3].

Pimpinelli et al. derive the scaling exponents by dimensional analysis of a continuum height equation of the generic form

$$\frac{\partial h}{\partial t} = -\frac{\partial}{\partial x}\left[Bm^{\rho} + K\frac{\partial^2}{\partial x^2}m^n\right], \tag{5}$$

where $m = \partial h/\partial x > 0$ is the slope of the surface, assumed to be positive, the exponent ρ characterizes the instability mechanism and n is the exponent of the step-step interactions (usually $n = 2$). In this talk I address two questions: First, how can equations of the form (5) be derived from the underlying step dynamics? Second, are the scaling exponents correctly given by the dimensional analysis employed in [1]?

With regard to the first question, we show that an equation of the form (5) with $\rho = -1$ follows whenever the destabilizing part of the step dynamics can be assumed to be *linear* in the step spacings [4]. Examples for such kind of linear step dynamics include growth with strong inverse Ehrlich-Schwoebel barriers [5], as well as sublimation with conventional Ehrlich-Schwoebel barriers [6], and surface electromigration [7], provided attachment to the steps is slow. The latter implies an additional mobility factor $1/m$ in front of the coefficient K of the stabilizing term in (5) [8, 9]. In a well-defined sense, these three problems therefore belong to the same universality class. The corresponding prediction of the theory of [1] for the scaling exponents is

$$\alpha = 1 + \frac{2}{n}, \quad \gamma = \frac{2}{2+n}, \quad \beta = \frac{1}{2}. \tag{6}$$

To address the second question, we have analyzed the stationary solutions of (5), which are characterized by the condition of constant current. For the universality class of interest it reads

$$J = \frac{B}{m} + \frac{K}{m}\frac{d^2}{dx^2}m^n \equiv J_0. \tag{7}$$

Interpreting m^n as a particle coordinate, this is Newton's equation for motion in a one-dimensional potential. The bunch shape corresponds to a trajectory starting at and returning to $m = 0$.

In analyzing this problem, it is important to realize that the mean current J_0 is not an adjustable integration constant; instead, it is forced by the microscopic boundary conditions to remain at the value $J_0 = B/m_0$ that it would have on the initial undisturbed vicinal surface of slope m_0 [6, 7]. For large bunches (large slopes) this implies that the mean current J_0 much exceeds the destabilizing part B/m inside the bunch, which is therefore irrelevant for the shape of the bunch. The latter is instead determined by the balance between the stabilizing step-step interaction term and the mean current. This problem was first analyzed by Nozières [8]. It gives rise to a bunch profile with the characteristic Pokrovsky-Talapov singularity

$$h(x) - h(x_0) \sim (x - x_0)^{3/2} \tag{8}$$

near the edges x_0 of the bunch. The scaling exponent for the minimal terrace size (i.e., the maximum value of the slope) turns out to be $\gamma = 2/(n+1)$, in contradiction to (6). This is because the relevant part of the current is in fact independent of the slope, so that the dimensional analysis should be carried out with $\rho = 0$ rather than with $\rho = -1$.

Numerical integration of the step dynamical equations shows good agreement with the expression for the minimal terrace size $l_{\min}$ derived through the above analysis, with regard to the scaling exponent γ as well as with regard to the prefactor in the scaling law [6, 10]. Similar agreement is found for the size l_1 of the *first* terrace in the bunch. On the other hand, the numerics indicates that the exponents α and β, which describe, in a sense, the *global* properties of bunches, are correctly given by the expressions (6) derived by dimensional analysis assuming $\rho = -1$. In particular, the scaling relation $\gamma = 1 - 1/\alpha$ suggested by trivial geometric considerations seems to be violated. We conjecture that this is related to the distinctly asymmetric bunch shape, which is not captured by the (manifestly symmetric) solutions of (7). In particular, the scaling of the size of the last terrace in the bunch (which is in fact hard to unambiguously identify) is completely different from that of the first terrace. Qualitatively, the asymmetry in the bunch shape is related to the drift of bunches and the exchange of steps between bunches. Further work is needed to clarify to what extent these phenomena, and thus, the overall scaling of the bunch morphology, can be captured by continuum height equations.

The talk is based on joint work with Vesselin Tonchev and Stoyan Stoyanov.

References

[1] A. Pimpinelli, V. Tonchev, A. Videcoq, M. Vladimirova: Phys. Rev. Lett. **88**, 206103 (2002)
[2] K. Fujita, M. Ichikawa, S.S. Stoyanov: Phys. Rev. B **60**, 16006 (1999)
[3] Y.-N. Yang, E.S. Fu, E.D. Williams: Surf. Sci. **356**, 101 (1996)
[4] J. Krug, in: *Dynamics of Fluctuating Interfaces and Related Phenomena*, ed. by D. Kim, H. Park, B. Kahng (World Scientific, Singapore 1997) pp. 95–113
[5] M. Sato, M. Uwaha: Surf. Sci. **493**, 494 (2001)
[6] J. Krug, V. Tonchev, S. Stoyanov, A. Pimpinelli (in preparation)
[7] D.-J. Liu, J.D. Weeks: Phys. Rev. B **57**, 14891 (1998)
[8] P. Nozières: J. Physique **48**, 1605 (1987)
[9] D.-J. Liu, E.S. Fu, M.D. Johnson, J.D. Weeks, E.D. Williams: J. Vac. Sci. Technol. B **14**, 2799 (1996)
[10] V. Tonchev, unpublished data

Surface Electromigration of single Islands

Philipp Kuhn
(joint work with Joachim Krug)

Surface electromigration is the biased diffusion of adatoms in the presence of an electric field. In order to understand the influence of this effect on the morphology of the surface we investigate the motion of a single island on a flat surface. We utilise a continuum approach where the island edge is treated as a continuous curve which evolves due to the competition between capillary forces and the electromigration force. We present an exact solution for the case without capillarity, and show numerical evidence for an oscillatory instability induced by crystal anisotropy in the step edge mobility.

Upper bounds on coarsening rates

Felix Otto
(joint work with Robert Kohn)

We consider two standard models of surface-energy-driven coarsening: a constant-mobility Cahn-Hilliard equation, whose large-time behaviour corresponds to Mullins-Sekerka dynamics; and a degenerate-mobility Cahn-Hilliard equation, whose large-time behaviour corresponds to motion by surface diffusion. Arguments based on scaling suggest that the typical length scale should behave as $l(t) \approx t^{1/3}$ in the first case and $l(t) \approx t^{1/4}$ in the second. We prove a weak, one-sided version of this assertion – showing, roughly speaking, that no solution can coarsen faster than the expected rates. Our results constrains the behaviour in a time-averaged sense rather than pointwise in time, and it constrains not the physical length scale but rather the perimeter per unit volume.

Discretization and numerical tests of a diffuse-interface model with Ehrlich-Schwoebel barrier
Patrick Penzler, Tobias Rump
(joint work with Felix Otto)

We consider a step–flow model for epitaxial growth, as proposed by Burton, Cabrera and Frank. This type of model is discrete in the growth direction but continuous in the lateral directions. The effect of the Ehrlich–Schwoebel barrier, which limits the attachment rate of adatoms to a step from an upper terrace, is included. Mathematically, this model is a 2+1–dimensional dynamic free boundary problem for the steps. In [Nonlinearity 17, 477(2004)] a diffuse–interface approximation which reproduces an arbitrary Ehrlich–Schwoebel barrier has been proposed. It is a version of the Cahn–Hilliard equation with variable mobility. In this talk, we propose a discretization for this diffuse–interface approximation. Our approach is guided by the fact that the diffuse–interface approximation has a conserved quantity and a Liapunov functional. We obtain an implicit finite volume discretization of symmetric structure. We test the discretization by comparison with the matched asymptotic analysis. We also test the diffuse–interface approximation itself by comparison with theoretically known features of the original free boundary problem. More precisely, we investigate quantitatively the phenomena of step–bunching and the Bales–Zangwill instability.

Phase field models for step flow growth
Olivier Pierre-Louis

The relation between phase field and discontinuous models for crystal steps is analyzed. - Different formulations of the kinetic boundary conditions of the discontinuous model are first presented. We show that: (i) step transparency, usually interpreted as the possibility for adatoms to jump through steps, may be seen as a modification of the equilibrium concentration engendered by step motion. (ii) The interface definition (i.e. the position of the dividing line) intervenes in the expression of the kinetic coefficients only in the case of fast attachment kinetics. (iii) We also identify the thermodynamically consist ent reference state in the kinetic boundary conditions. - Asymptotic expansions of the phase field models in the limit where the interface width is small, lead to various discontinuous models: (1) A phase field model with one global concentration field and variable mobility is shown to lead to a discontinuous model with fast step kinetics. (2) A phase field model with one concentration field per terrace allows one to recover arbitrary step kinetics (i.e. arbitrary strong Ehrlich-Schwoebel effect and step transparency). - Quantitative agreement is found in the linear and nonlinear regimes, between the numerical solution of the phase field models and the analytical solution of the discontinuous model.

Phase-field models for epitaxial growth
Andreas Rätz
(joint work with Axel Voigt)

Different phase field models are proposed as an approximation of classical sharp interface models of Burton-Cabrera-Frank type. The motion of island boundaries of discrete atomic layers is determined by the time evolution of an introduced phase-field variable. In order to describe attachment-detachment kinetics in epitaxial growth a reduced mobility is applied for the modelling of the asymmetry in the kinetic boundary conditions, while an increased mobility is used for the approximation of edge diffusion along the free boundary. We apply matched asymptotic expansion to determine the asymptotic limit of vanishing interfacial thickness and show the reduction to classical sharp interface models. Furthermore an adaptive finite element discretization and numerical results are shown.

Continuum ("height") models for surface growth, an overview
Martin Rost

In crystal growth models the surface is often represented by a height field h(x,t). Its dynamics can be derived on heuristic grounds yielding equations of the form ht(x,t) =3D =85, where the growth velocity above the substrate point x and at time t depends on the present surface configuration h. This talk attempts to give an introductory overview on the use of continuum height field dynamics for crystal growth focussing on three key issues: (i) thermodynamic and kinetic basis for its derivation, (ii) symmetries and conserved quantities, also in connection to analogous approaches in other fields, and (iii) typical applications and results, also linking it to more detailed crystal growth models.

Quasicontiuum Monte Carlo: A computational method for surface growth calculations
Peter Smereka
(joint work with Jason Devita, Giovanni Russo, Len Sander)

Epitaxial growth on surfaces is of central importance both for applications and as a very interesting example of statistical processes out of equilibrium. This growth process is commonly modeled by Kinetic Monte Carlo (KMC) and continuum models. In KMC each adatom is represented individually; therefore, it automatically includes internal noise processes. However, when there are many adatoms, (e.g. close to equilibrium) these simulations slow down considerably. A deterministic continuum model which represents the adatoms as a continuous fluid does not have this problem, and should be much faster. There has been considerable work in the development of such models for epitaxial growth (see [1] references therein). In some cases they have been quite successful, but in other cases they have failed to reproduce the structures seen in experiment. One reason for such problems is that deterministic continuum models neglect important fluctuations. In this talk we present a method of dealing with some fluctuations without giving up the advantages of a continuum treatment. We call this approach Quasi-Continuum Monte Carlo (QCMC). The most important use of this method will be in cases where fluctuations are important, but which would be difficult to treat with KMC because of the presence of a large number of adatoms.

The first version of our QCMC algorithm goes as follows: we treat the islands on the surface as crystals containing discrete atoms which occupy the sites of a lattice. To illustrate the method we use a square lattice. On the other hand, the adatoms are treated as a continuum whose density, ρ, is governed by:

$$\partial_t \rho = D\nabla^2 \rho + F. \tag{9}$$

In practice, we solve this equation numerically on a discrete grid which is commensurate with the crystal lattice. On the surface of the island, we include boundary conditions that model both attachment and detachment processes. We then compute the velocity of the interface as one would have for a continuum model. However, in QCMC we interpret this in a way that includes fluctuations. For example, when attachment is the only process we compute the total flux onto the island boundary and when the total flux exceeds one atom, then one or more adatoms are attached to the boundary at random with the probability proportional to the normal speed. For more details see [4]. We have shown that this method agrees quite well with KMC and retains the advantages of a continuum method. This approach is similar to the dielectric breakdown model[3] which can be considered a generalization of diffusion limited aggregation[2]. Our method has been extended to multi-layer growth including nucleation in [5].

References

[1] S. Chen, B. Merriman, M. Kang, R. E. Caflisch, C. Ratsch, L. Cheng, M. Gyure, R. P. Fedkiw, C. R. Anderson, and S. J. Osher, *Level set method for thin film epitaxial growth,* Journal of Computational Physics **167**, 475-500 (2000).

[2] T. A. Witten and L. M. Sander, *Diffusion-limited aggregation, a kinetic critical phenomenon,* Physical Review Letters **47**, 1400-1403 (1981).

[3] L. Niemeyer, L. Pietronero, and H. J. Wiesmann, *Fractal dimension of dielectric breakdown,* Physical Review Letters **52**, 1033-1036 (1984).

[4] G. Russo, L.M. Sander, and P. Smereka, *Quasicontinuum Monte Carlo: A method for surface growth simulations* Physical Review B, accepted for publication, January 2004.

[5] J. Devita, L.M. Sander, and P. Smereka, in preparation.

Semi-implicit level set methods for curvature and surface diffusion motion

Peter Smereka

We introduce semi-implicit methods for evolving interfaces by mean curvature flow and surface diffusion using level set methods.

Regularized anisotropic curve shortening flow

Axel Voigt
(joint work with Frank Haußer)

Realistic interfacial energy densities are often non convex, which results in backward parabolic behaviour of the corresponding anisotropic curve shortening flow, thereby inducing phenomena such as the formation of corners and facets. Adding a term being quadratic in the curvature to the interfacial energy yields a regularized evolution equation for the interface, which is fourth order parabolic. Using a semi-implicit time discretization, we present a variational formulation of this equation, which allows the use of linear finite elements. The resulting linear system is shown to be uniquely solvable. We also present numerical examples. The described algorithm can also be used to solve Willmore flow.

A general finite element framework for Burton-Cabrera-Frank equations

Axel Voigt
(joint work with Eberhard Bänsch, Frank Haußer)

An adaptive finite element method is presented for step flow models in homoepitaxial growth. Such problems consist of an adatom (adsorbed atom) diffusion equation on each terrace; boundary condition on steps between the terraces including thermodynamic or kinetic conditions; and a normal velocity law for the motion of the steps, which is determined by a two-sided flux, together with edge-diffusion. Mathematically speaking it is a 2+1 model and it is solved using independent meshes, a two-dimensional mesh for the adatom diffusion and a one-dimensional mesh for the boundary evolution. The diffusion equation is discretized using linear composite finite elements in space and an implicit scheme in time in the case of attachment limited growth (kinetic boundary conditions). For diffusion limited growth (thermodynamic boundary conditions) a penalty method is applied. The evolution of the steps included surface diffusion, curvature flow and forcing terms. Its governing equation is solved by a semi-implicit front-tracking method using linear parametric finite elements. Simple adaptive techniques are employed in solving the adatom diffusion equation as well as the boundary motion problem.

Levelset formulation for fourth order geometric evolution problems

Ulrich Weikard
(joint work with Ulrich Clarenz, Frank Haußer, Axel Voigt)

A level set formulation of anisotropic surface diffusion is derived using the gradient flow perspective. Starting from single embedded surfaces and the corresponding gradient flow, the metric is generalized to sets of level set surfaces using the identification of normal velocities and variations of the level set function in time via the level set equation. The approach in particular allows to identify the natural dependent quantities of the derived variational formulation. Furthermore, spatial and temporal discretization are discussed and some numerical simulations in two and three dimensions are presented.

Reporter: Axel Voigt

Participants

Prof. Dr. Karsten Albe
albe@tu-darmstadt.de
FB Material- und Geowissenschaften
Technische Universität Darmstadt
Petersenstr. 23
D–64289 Darmstadt

Rainer Backofen
backofen@caesar.de
Crystal Growth Group
Research Center CAESAR
Ludwig-Erhard-Allee 2
D–53175 Bonn

Prof. Dr. Eberhard Bänsch
baensch@math.uni-bremen.de
baensch@wias-berlin.de
Weierstraß-Institut für
Angewandte Analysis und Stochastik
im Forschungsverbund Berlin e.V.
Mohrenstr. 39
D–10117 Berlin

Dr. Michael Biehl
biehl@physik.uni-wuerzburg.de
biehl@cs.rug.nl
Institute for Mathematics and
Computing Science
P.O.Box 800
NL-9700 AV Groningen

Prof. Dr. Carlo Cavallotti
carlo.cavallotti@polimi.it
Dept. Chimica, Materiali e
Ingegneria Chimica 'G. Natta'
Politecnico di Milano
Via Mancinelli 7
I-20131 Milano

Dr. Frank Haußer
hausser@caesar.de
Crystal Growth Group
Research Center CAESAR
Ludwig-Erhard-Allee 2
D–53175 Bonn

Dr. Navot Israeli
navot.israeli@weizmann.ac.il
Dept. of Physics of Complex Systems
The Weizmann Institute of Science
76100 Rehovot – Israel

Prof. Dr. Joachim Krug
krug@thp.uni-koeln.de
Institut für Theoretische Physik
Universität Köln
Zülpicher Str. 77
D–50937 Köln

Philipp Kuhn
philipp@theo-phys.uni-essen.de
Fachbereich Physik
Universität Duisburg-Essen
Standort Essen
D–45117 Essen

Prof. Dr. Felix Otto
otto@iam.uni-bonn.de
Otto@rieman.iam.uni-bonn.de
Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 10
D–53115 Bonn

Patrick Penzler
penzler@iam.uni-bonn.de
Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 6
D–53115 Bonn

Prof. Dr. Olivier Pierre-Louis
olivier.pierre-louis@ujf-grenoble.fr
Laboratoire de Spectrometrie
Physique
Universite Joseph Fourier (CNRS)
Grenoble 1, B.P. 87
F-38402 Saint-Martin d'Heres Cedex

Andreas Rätz
raetz@caesar.de
Crystal Growth Group
Research Center CAESAR
Ludwig-Erhard-Allee 2
D–53175 Bonn

Dr. Martin Rost
martin.rost@uni-bonn.de
Institut für zelluläre & molekulare
Botanik (IZMB), Theoret. Biologie
Universität Bonn
Kirschallee 1
D–53115 Bonn

Dipl.Math. Tobias Rump
rump@iam.uni-bonn.de
Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 6
D–53115 Bonn

Prof. Dr. Peter Smereka
psmereka@umich.edu
Dept. of Mathematics
The University of Michigan
2074 East Hall
525 E. University Ave.
Ann Arbor, MI 48109-1109 – USA

Dr. Axel Voigt
voigt@caesar.de
Crystal Growth Group
Research Center CAESAR
Ludwig-Erhard-Allee 2
D–53175 Bonn

Ulrich Weikard
weikard@math.uni-duisburg.de
Numerische Mathematik und
Wissenschaftliches Rechnen
Gerhard-Mercator-Universität
Lotharstr. 65
D–47048 Duisburg

MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

Report No. 5/2004

Wave Motion

Organised by
Adrian Constantin (Lund)
Joachim Escher (Hannover)

January 25th – January 31th, 2004

Introduction by the Organisers

The workshop **Wave Motion** that took place in the period January 25–31, 2004 was devoted to the study of nonlinear wave phenomena. The modelling of waves leads to a variety of difficult mathematical issues, involving several domains of mathematics: partial differential equations, harmonic analysis, dynamical systems, topological degree theory.

The progam of the workshop consisted in 18 talks, presented by international specialists in nonlinear waves coming from England, France, Germany, Japan, Norway, Sweden, Switzerland, U.S.A., and by three discussion sessions on the topics "Open Problems in PDEs", "Stability Phenomena in the Theory of Nonlinear Waves", and "Geodesic Flows and Fluid Mechanics". Moreover, several doctoral and post-doctoral fellows participated in the workshop and did benefit from the unique academic atmosphere at the Oberwolfach Institute.

The proceedings of the workshop "Wave Motion" will appear as a special issue of the *Journal of Nonlinear Mathematical Physics*.

Workshop on Wave Motion

Table of Contents

Abstracts

Nonlinear water waves and spatial dynamics

MARK D. GROVES

The water-wave problem is the study of the three-dimensional irrotational flow of a perfect fluid bounded below by a rigid horizontal bottom $\{y = 0\}$ and above by a free surface $\{y = h + \eta(x, z, t)\}$ subject to the forces of gravity and surface tension. This remarkable problem, first formulated in terms of a potential function ϕ by Euler (Figure 1), has become a paradigm for most modern methods in nonlinear functional analysis and nonlinear dispersive wave theory. Its mathematical study has historically called upon many different approaches (iteration methods, bifurcation theory, complex variable methods, PDE methods, the calculus of variations, positive operator theory, topological degree theory, KAM theory, symplectic geometry, . . .). In this talk I would like to illustrate the role of the water-wave problem as a paradigm in the theory of Hamiltonian systems and conservative pattern-formation problems.

$$
\begin{aligned}
&\phi_{zz} + \phi_{yy} + \phi_{zz} = 0, && 0 < y < h + \eta,\\
&\phi_y = 0, && y = 0,\\
&\phi_y = \eta_t + \eta_x\phi_x + \eta_z\phi_z, && y = h + \eta,\\
&\phi_t = -\frac{1}{2}(\phi_x^2 + \phi_y^2 + \phi_z^2) - g\eta && \\
&\qquad + \sigma\left[\frac{\eta_x}{\sqrt{1 + \eta_x^2 + \eta_z^2}}\right]_x + \sigma\left[\frac{\eta_z}{\sqrt{1 + \eta_x^2 + \eta_z^2}}\right]_z, && y = h + \eta
\end{aligned}
$$

FIGURE 1. Euler (1707–1783), who first formulated the water-wave problem (left)

Travelling water waves are solutions of the water-wave problem which are stationary in a uniformly translating reference frame, so that $\eta(x, z, t) = \eta(\xi, z)$, where $\xi = x - ct$. The resulting time-independent problem can be approached using the method of *spatial dynamics*, which was devised by K. Kirchgässner specifically with water waves in mind and has now found applications in a huge range of other problems (reaction-diffusion equations, spiral waves, mathematical biology, . . .). The idea is to formulate a stationary problem as an evolutionary equation in which an unbounded *spatial* coordinate plays the role of the time-like variable. In the travelling water-wave problem one can take any horizontal direction $X = \sin\theta_2\,\xi - \cos\theta_2\,z$ as the time-like variable and formulate the equations as an evolutionary equation

$$
u_X = Lu + Nu, \qquad u \in \mathcal{X}; \tag{1}
$$

the infinite-dimensional phase space $\mathcal{X}$ is constructed to contain functions which are, for example, $2\pi/\nu$-periodic in a second, different horizontal direction $Z = \sin\theta_1\,\xi - \cos\theta_1\,z$.

The evolutionary equation (1) is found by performing a Legendre transform upon the classical variational principle

$$\delta \int \int_0^{2\pi} \left\{ \int_0^{h+\eta} \Big(-\sin\theta_2\,\phi_X - \nu\sin\theta_1\,\phi_Z + \frac{1}{2}\left(\phi_X^2 + \phi_y^2 + \nu^2\phi_Z^2 + 2\nu\cos(\theta_1-\theta_2)\phi_X\phi_Z\right)\Big)\,dy \right.$$
$$\left. + \frac{1}{2}g\eta^2 + \sigma\left(\sqrt{1+\eta_X^2+\nu^2\eta_Z^2+2\nu\cos(\theta_1-\theta_2)\eta_X\eta_Z} - 1\right)\right\}\,dZ\,dX = 0$$

for the desired wave motions. In many cases equation (1) can be treated using an invariant-manifold theory due to A. Mielke, which was again developed with this problem in mind, but is now used in a wide variety of problems (elasticity, solid mechanics, ...). This theory shows that all small, bounded solutions lie on a finite-dimensional invariant manifold and thus reduces the water-wave problem to a locally equivalent finite-dimensional Hamiltonian system; the dimension and character of this reduced system depend upon the values of the physical parameters (gravity g, surface tension σ, wave speed c, water depth h).

Two-dimensional i.e. z-independent travelling waves lend themselves naturally to an application of the spatial dynamics method with $X = \xi$. B. Buffoni, M. D. Groves & J. F. Toland showed that in a certain parameter regime the invariant manifold is four dimensional and controlled by the Hamiltonian equation

$$u'''' + Pu'' + u - u^2 = 0, \qquad P \in (-2, -2+\epsilon).$$

Amazingly, this equation turns up in many, seemingly unrelated problems in applied science, for example in nonlinear elasticity, nonlinear optics and now nonlinear water waves. One of its most interesting features is that it exhibits *chaotic behaviour:* there is a Smale-horseshoe structure in its solution set. As a consequence, it has infinitely many *homoclinic solutions*, that is solutions which decay to zero as the time-like variable tends to infinity. The corresponding solutions of the water-wave problem are called *solitary waves* and decay to the undisturbed state of the water as $\xi \to \pm\infty$. This result shows that there are infinitely many of them; they are waves of depression with $2, 3, 4, \ldots$ large troughs separated by $2, 3, \ldots$ small oscillations, and their oscillatory tails decay exponentially to zero. Two waves from this family are sketched in Figure 2.

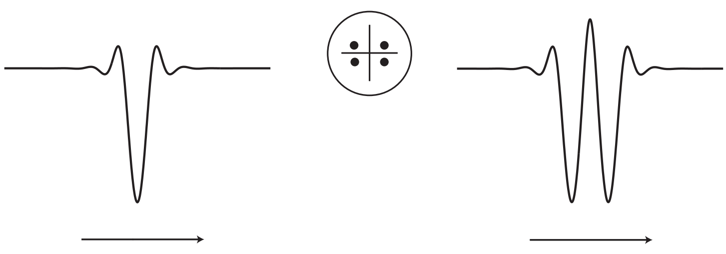


FIGURE 2. Two of the multi-troughed solitary waves found by B. Buffoni, M. D. Groves & J. F. Toland on a four-dimensional invariant manifold with the depicted eigenvalue structure.

The study of two-dimensional solitary waves was continued by G. Iooss & K. Kirchgässner and B. Buffoni & M. D. Groves, who noticed that there are parameter values for which the invariant manifold is four-dimensional and a *Hamiltonian-Hopf bifurcation* takes place (two nonsemisimple imaginary eigenvalues become complex as a parameter is varied). Hamiltonian-Hopf bifurcations are well-known to researchers in the field of celestial mechanics, where they occur in the restricted three-body problem for the planar motion of a light body orbiting the centre of mess of two heavy bodies; the Hamiltonian-Hopf bifurcation occurs for a certain value of the mass ratio of the two heavy particles (Routh's ratio). Iooss & Kirchgässner used the Birkhoff normal form to show that Hamiltonian-Hopf bifurcations generate homoclinic solutions which take the form of periodic wave trains modulated by exponentially decaying envelopes (Figure 3). Buffoni & Groves showed that there are in fact infinitely many such solutions which resemble multiple copies of Iooss & Kirchgässner's solutions; their proof is based upon modern methods from the calculus of variations (mountain-pass arguments and the concentration-compactness principle) and the topological degree. These results are not restricted to the water-wave problem in which they emerge; they provide dramatic new solutions to the three-body problem and indeed Hamiltonian-Hopf bifurcations have been detected in a range of situations (Taylor-Couette flows, nonlinear elasticity,. . .).

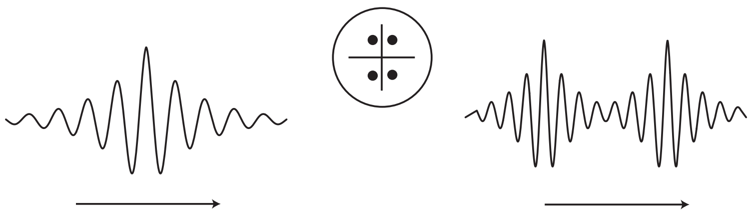


FIGURE 3. Two of the multi-packet solitary waves found by B. Buffoni & M. D. Groves on a four-dimensional invariant manifold with the depicted eigenvalue structure.

M. D. Groves & M. Haragus have recently classified all the possible bifurcation scenarios for *three-dimensional* travelling waves using the spatial dynamics method. In particular, they compiled a catalogue of three-dimensional waves which have solitary-wave or *generalised solitary-wave* profiles in a distinguished horizontal direction (the time-like direction); these profiles decay respectively to zero and to a periodic wavetrain at large distances. Some of the waves are rather exotic, as Figure 4 shows.

Groves & Haragus also examined doubly periodic travelling waves using spatial dynamics. Periodicity in the Z-direction is built into the method, so that doubly periodic waves are found as solutions of the reduced Hamiltonian system which are periodic in the time-like direction X. Such solutions are found using the classical Lyapunov centre theorem, and depending upon the physical parameters one encounters all possible cases: nonresonant eigenvalues, semisimple eigenvalue resonances, nonsemisimple eigenvalue resonances and equal or opposite Krein signatures! Doubly periodic surface waves (Figure 5) and periodic motion of heavenly bodies (the n-body problem in celestial mechanics) are, according to the above observations, two aspects of the same mathematical theory, namely finite-dimensional Hamiltonian systems and the Lyapunov centre theorem.

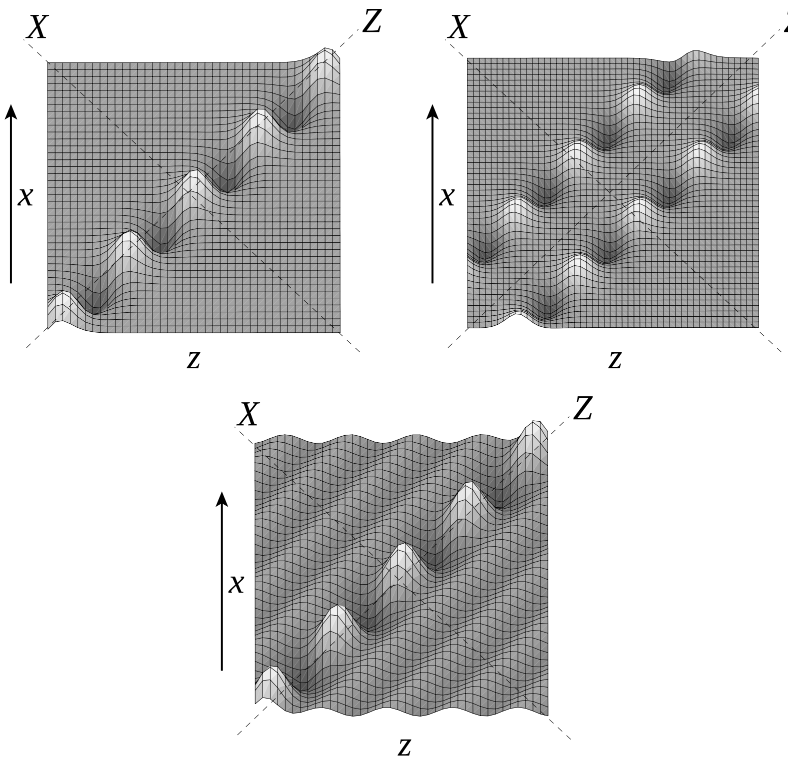


FIGURE 4. Three examples from the catalogue of three-dimensional travelling waves compiled by M. D. Groves & M. Haragus. These waves have the profile of (a) a one-pulse solitary wave, (b) a two-pulse solitary wave and (c) a generalised solitary-wave in one distinguished spatial direction (X) and are periodic in another (Z); they move with constant speed and without change of shape in the x direction (arrowed).

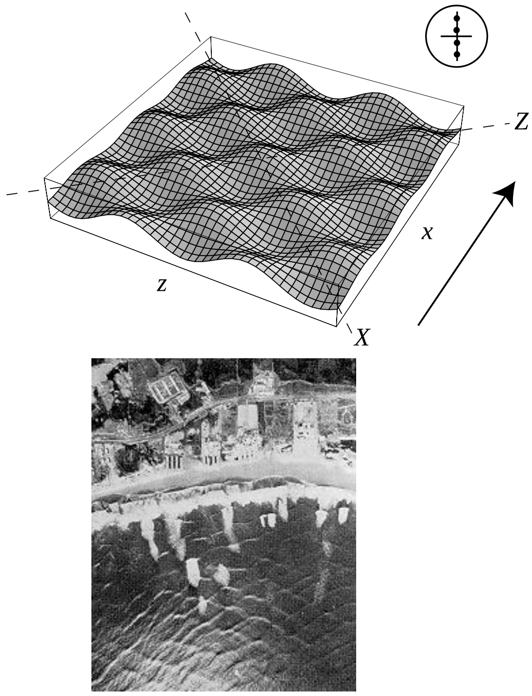


FIGURE 5. The doubly periodic wave on the left is constructed using the Lyapunov centre theorem on a four-dimensional invariant manifold. Hexagonal doubly periodic waves are often seen in nature, as this ariel photograph on the right shows; they can be explained mathematically by this procedure.

Corner defects in almost planar interface propagation

MARIANA HARAGUS
(joint work with Arnd Scheel)

We study existence and stability of almost planar interfaces in reaction-diffusion systems. Almost planar here refers to the angle of the interface at each point, relative to a fixed planar interface. Most of the interfaces that we construct are planar at infinity, with possibly different orientations at $+\infty$ and $-\infty$ in an arclength parameterization. We refer to all these types of interfaces as *corner defects*. According to their angles at $\pm\infty$ we distinguish between interior corners, exterior corners, steps and holes.

We construct corner defects as perturbations of a planar interface. Assumptions are solely on the existence of a primary planar travelling-wave solution and spectral properties of the linearization at the planar wave. All interfaces that we construct are stationary or time-periodic patterns in an appropriately comoving frame. The corner typically but not always points in the direction opposite to the direction of propagation. In addition, we give stability results which show that “open” classes of initial conditions actually converge to the corner-shaped interfaces we constructed before. The results are stated for reaction-diffusion systems but the method is sufficiently general to cover different situations, as well. In particular we do not rely on monotonicity arguments or comparison principles such that we can naturally include the case of interfaces separating patterned states from spatially homogeneous states.

The method we use is based on the (essentially one-dimensional) dynamical systems approach to the existence of bounded solutions to elliptic equations in cylinders introduced by Kirchgässner. The main idea is to consider an elliptic equation, posed on the (x, y)-plane in a neighborhood of an x-independent wave $q_*(y)$ as a dynamical system in the x-variable and rely on dynamical systems tools such as center-manifold reduction and bifurcation theory to construct bounded solutions to the elliptic equation in a neighborhood of the original wave. Nontrivial, that is non-equilibrium, x-“dynamics” then correspond to nontrivial x-profiles. In the present work, we extend these ideas, incorporating the shift of the y-profile $q_*(y)$ into the reduced dynamics. We then respect this affine action of the symmetry group in the construction and parameterization of the center-manifold such that the reduced equations take a skew-product form. The reduced ordinary differential equation can be viewed as the travelling-wave equation to a viscous conservation law or variants of the Kuramoto-Sivashinsky equation.

Invariance principle and the inverse problems for the periodic Camassa-Holm equation

EVGENI KOROTYAEV

Consider the nonlinear mapping $F : H_1 \to H$ given by

$$F(y) = y' + u_1(y) + u_2(Jy) - u_0(y, y'), \quad Jy = \int_0^x y(s)ds, \quad y \in H_1$$

where the Hilbert spaces $H = \{q : q \in L^2_R(\mathbb{T}),\ \int_0^1 q(x)dx = 0\}$ and $H_1 = \{y, y' \in H\}$. Here the functions u_1, u_2 are real analytic and $u_2'(y) \le 0$: the constant u_0 is such that $F(y) \in H$. We prove that the map F is a real analytic isomorphism. Furthermore, a priori two-sided estimates of norms of $F(y), y$ are obtained. We apply these results to the inverse problems for the Schrödinger operators $S = -d^2dx^2 + F(y)$ in $L^2(\mathbb{R})$ with a 1-periodic potential $F(y)$. The inverse problem for the operator $S_0 S = -d^2dx^2 + p$, (p is periodic) is the well known fact. Thus we have the factorization, which yields the solution the inverse problem for S, with "variable" y. We call these result invariance principle since the results about the inverse problems depend on only the large conditions on the functions u_1, u_2.

In the second part we use this result to study the Camassa-Holm equation. We consider the periodic weighted operator $Ty = -\rho^{-2}(\rho^2 y')' + 14\rho^{-4}$ in $L^2(\mathbb{R}, \rho^2 dx)$ where ρ is a 1-periodic positive function satisfying $q = \rho'/\rho \in L^2(0, 1)$. The spectrum of T consists of intervals separated by gaps. Using the Liouville transform we get the Schrödinger operators $-d^2dx^2 + F$, where $F = q' + q^2 + u_2(Jq) - u_0$, i.e.,in this case $u_1(q) = q^2,\ \ u_2(Jq) = 14e^{-4Jq}$. Firstly, we construct the Marchenko-Ostrovski mapping $q \to$

$h(q)$ and solve the corresponding inverse problem. For our approach it is essential that the mapping h has the factorization $h(q) = h^0(F(q))$, where $q \to F(q)$ is a certain nonlinear mapping and $V \to h^0(F)$ is the Marchenko-Ostrovski mapping for the Hill operator. Secondly, we solve the inverse problem for the gap length mapping and we obtain the trace formula for T.

Noncommutative Deformation of Solitons

OLAF LECHTENFELD

A noncommutative (Moyal) deformation of a function space over $\mathbb{R}^{2n}$ is achieved formally by subjecting the coordinate functions $(x^\mu)_{\mu=1,\ldots,2n}$ to the (Heisenberg) algebra

$$x^\mu \star x^\nu - x^\nu \star x^\mu = \mathrm{i}\theta^{\mu\nu},$$

where '$\star$' denotes the deformed product and $(\theta^{\mu\nu})$ is a constant antisymmetric matrix. This induces an associative product on $C^\infty(\mathbb{R}^{2n})$ involving arbitrary powers of a bidifferential operator. A convenient realization of the deformed function algebra trades the coordinate dependence (and the deformed product) for operator valuedness (and the standard compositional product), the operators acting on an auxiliary Fock space such as $L^2(\mathbb{R}^n)$. The model is quantum mechanics.

The generalization of integrable differential equations, for instance the sine-Gordon equation, to the noncommutative setup is ambiguous, but a distinguished choice arises from demanding the existence of a noncommutative Lax pair. The rewriting of the nonlinear differential equation as the compatibility condition of a linear system does not notice the noncommutative deformation since one is already dealing with matrix-valued, i.e. noncommuting, objects. Likewise, established methods for generating solutions to the linear system, e.g. the dressing method, can be deformed painlessly.

We demonstrate this stategy for the example of the U(m) principal chiral (or nonlinear sigma-)

model in two dimensions which allows for solitonic solutions to its equation of motion. The dressing method reduces the task to solving an eigenvalue problem for a linear differential operator which in the noncommutative situation is realized by a simple linear operator in $L^2(\mathbb{R})$. After picking a basis in this Fock space, the noncommutative deformation formally amounts to replacing $n \times n$ matrices by semi-infinite ones. It is therefore not surprising to find not only smooth deformations of the commutative solitonic solutions but also a class of new solitons (even for the U(1) case!) which display a singular commutative ($\theta \to 0$) limit.

Like in the commutative case, all known integrable equations in 3, 2, and 1 dimensions descend from the four-dimensional self-dual Yang-Mills equations by dimensional reduction and specialization. The noncommutative deformation, however, is not compatible with picking any subgroup of U(m). Nevertheless, the method presented here can be applied to any integrable system, such as NLS, KdV, KP, Burgers, Boussinesq, sine-Gordon, or Camassa-Holm. The resulting noncommutative equations will be nonlocal (featuring an infinite number of derivatives) but of a controlled kind; their solutions will have many properties in common with the standard solitons.

REFERENCES

[1] M.F. Atiyah, R.S. Ward, Commun. Math. Phys. **55** (1977) 117;
M.F. Atiyah, N.J. Hitchin, I.M. Singer, Proc. Roy. Soc. Lond. **A362** (1978) 425.
[2] V.E. Zakharov, A.B. Shabat, Funct. Anal. Appl. **13** (1979) 166.
[3] L.D. Faddeev, L.A. Takhtajan, *Hamiltonian methods in the theory of solitons*, Springer 1987.
[4] M.R. Douglas, N.A. Nekrasov, Rev. Mod. Phys. **73** (2002) 977.
[5] A. Konechny, A. Schwarz, Phys. Rept. **360** (2002) 353.
[6] O. Lechtenfeld, A.D. Popov, JHEP **0111** (2001) 040 [hep-th/0106213]
[7] O. Lechtenfeld, *Noncommutative instantons and solitons* [hep-th/0401158].

Renormalized entropy solutions for quasilinear anisotropic degenerate parabolic equations

KENNETH HVISTENDAHL KARLSEN
(joint work with Mostafa Bendahmane)

We consider the Cauchy problem for quasilinear anisotropic degenerate parabolic equations with L^1 data. This convection–diffusion type problem is of the form

$$\partial_t u + \mathrm{div} f(u) = \nabla \cdot (a(u)\nabla u) + F, \qquad u(0,x) = u_0(x), \tag{1}$$

where $(t,x) \in (0,T) \times \mathbf{R}^d$; $T > 0$ is fixed; div and ∇ are with respect to $x \in \mathbf{R}^d$; and $u = u(t,x)$ is the scalar unknown function that is sought. The (initial and source) data $u_0(x)$ and $F(t,x)$ satisfy

$$u_0 \in L^1(\mathbf{R}^d), \qquad F \in L^1((0,T) \times \mathbf{R}^d). \tag{2}$$

The diffusion function $a(u) = (a_{ij}(u))$ is a symmetric $d \times d$ matrix of the form

$$a(u) = \sigma(u)\sigma(u)^\top \geq 0, \qquad \sigma \in \left(L^\infty_{\mathrm{loc}}(\mathbf{R})\right)^{d\times K}, \qquad 1 \leq K \leq d, \tag{3}$$

and hence has entries

$$a_{ij}(u) = \sum_{k=1}^{K} \sigma_{ik}(u)\sigma_{jk}(u), \qquad i,j = 1,\dots,d.$$

The inequality in (3) means that for all $u \in \mathbf{R}$

$$\sum_{i,j=1}^{d} a_{ij}(u)\lambda_i\lambda_j \geq 0, \qquad \forall \lambda = (\lambda_1,\dots,\lambda_d) \in \mathbf{R}^d.$$

Finally, the convection flux $f(u)$ is a vector–valued function that satisfies

$$f(u) = (f_1(u),\dots,f_d(u)) \in \left(\mathrm{Lip}_{\mathrm{loc}}(\mathbf{R})\right)^d. \tag{4}$$

It is well known that (1) possesses discontinuous solutions and that weak solutions are not uniquely determined by their initial data (the scalar conservation law is a special case of (1)). Hence (1) must be interpreted in the sense of entropy solutions [15, 19, 20]. In recent years the isotropic diffusion case, for example the equation

$$(5) \qquad \partial_t u + \mathrm{div} f(u) = \Delta A(u), \quad A(u) = \int_0^u a(\xi)\, d\xi, \quad 0 \le a \in L^\infty_{\mathrm{loc}}(\mathbf{R}),$$

has received much attention, at least when the data are regular enough (say $L^1 \cap L^\infty$) to ensure $\nabla A(u) \in L^2$. Various existence results for entropy solutions of (5) (and (1)) can be derived from the work by Vol'pert and Hudjaev [20]. Some general uniqueness results for entropy solutions have been proved in the one-dimensional context by Wu and Yin [21] and Bénilan and Touré [2]. In the multi-dimensional context a general uniqueness result is more recent and was proved by Carrillo [6, 5] using Kružkov's doubling of variables device. Various extensions of his result can be found in [4, 12, 13, 14, 16, 17, 18], see also [7] for a different approach. Explicit "continuous dependence on the nonlinearities" estimates were proved in [10]. In the literature just cited it is essential that the solutions u possess the regularity $\nabla A(u) \in L^2$. This excludes the possibility of imposing general L^1 data, since it is well known that in this case one cannot expect that much integrability.

The general anisotropic diffusion case (1) is more delicate and was successfully solved only recently by Chen and Perthame [9]. Chen and Perthame introduced the notion of kinetic solutions and provided a well posedness theory for (1) with L^1 data. Using their kinetic framework, explicit continuous dependence and error estimates for $L^1 \cap L^\infty$ entropy solutions were obtained in [8]. With the only assumption that the data belong to L^1, we cannot expect a solution of (1) to be more than L^1. Hence it is in general impossible to make distributional sense to (1) (or its entropy formulation). In addition, as already mentioned above, we cannot expect $\sqrt{a(u)}\nabla u$ to be square-integrable, which seems to be an essential condition for uniqueness. Both these problems were elegantly dealt with in [9] using the kinetic approach.

The purpose of the present paper is to offer an alternative "pure" L^1 well posedness theory for (1) based on a notion of renormalized entropy solutions and the classical Kružkov method [15]. The notion of renormalized solutions was introduced by DiPerna and Lions in the context of Boltzmann equations [11]. This notion (and a similar one called entropy solutions) was then adapted to nonlinear elliptic and parabolic equations with L^1 (or measure) data by various authors. We refer to [3] for some recent results in this context and a list of relevant references. Bénilan, Carrillo, and Wittbold [1] introduced a notion of renormalized Kružkov entropy solutions for scalar conservation laws with L^1 data and proved the existence and uniqueness of such solutions. Their theory generalizes the Kružkov well posedness theory for L^∞ entropy solutions [15].

Motivated by [1, 3] and [9], we introduce herein a notion of renormalized entropy solutions for (1) and prove its well posedness. Let us illustrate our notion of an L^1 solution on the isotropic diffusion equation (5) with initial data $u|_{t=0} = u_0 \in L^1$. To this end, let $T_l : \mathbf{R} \to \mathbf{R}$ denote the truncation function at height $l > 0$ and let $\zeta(z) = \int_0^z \sqrt{a(\xi)}\, d\xi$. A renormalized entropy solution of (5) is a function $u \in L^\infty(0, T; L^1(\mathbf{R}^d))$ such that (i) $\nabla\zeta(T_l(u))$ is square–integrable on $(0, T) \times \mathbf{R}^d$ for any $l > 0$; (ii) for any convex C^2 entropy-entropy flux triple (η, q, r), with η' bounded and $q' = \eta' f'$, $r' = \eta' a$, there exists for any $l > 0$ a nonnegative bounded Radon measure μ_l on $(0, T) \times \mathbf{R}^d$, whose total mass tends to zero as $l \uparrow \infty$, such that

$$\begin{aligned}(6)\qquad &\partial_t \eta(T_l(u)) + \mathrm{div} q(T_l(u)) - \Delta r(T_l(u))\\ &\qquad \le -\eta''(T_l(u)) \left|\nabla \zeta(T_l(u))\right|^2 + \mu_l(t,x) \quad \text{in } \mathcal{D}'((0,T)\times \mathbf{R}^d).\end{aligned}$$

Roughly speaking, (6) expresses the entropy condition satisfied by the truncated function $T_l(u)$. Of course, if u is bounded by M, choosing $l > M$ in (6) yields the usual entropy formulation for u, i.e., a bounded renormalized entropy solution is an entropy solution. However, in contrast to the usual entropy formulation, (6) makes sense also when u is merely L^1 and possibly unbounded. Intuitively the measure μ_l should be supported on $\{|u| = l\}$ and carry information about the behavior of the "energy" on the set where $|u|$ is large. The requirement is that the energy should be small for large values of $|u|$, that is, the total mass of the renormalization measure μ_l should vanish as $l \uparrow \infty$. This is essential for proving uniqueness of a renormalized entropy solution. Being explicit, the existence proof reveals that $\mu_l((0,T)\times \mathbf{R}^d) \le \int_{\{|u_0|>l\}} |u_0| \, dx \to 0$ as $l \uparrow \infty$.

We prove existence of a renormalized entropy solution to (1) using an approximation procedure based on artificial viscosity [20] and bounded data. We derive a priori estimates and pass to the limit in the approximations.

Uniqueness of renormalized entropy solutions is proved by adapting the doubling of variables device due to Kružkov [15]. In the first order case, the uniqueness proof of Kružkov depends crucially on the fact that

$$\nabla_x \Phi(x-y) + \nabla_y \Phi(x-y) = 0, \qquad \Phi \text{ smooth function on } \mathbf{R}^d,$$

which allows for a cancellation of certain singular terms. The proof herein for the second order case relies in addition crucially on the following identity involving the Hessian matrices of $\Phi(x-y)$:

$$\nabla_{xx}\Phi(x-y) + 2\nabla_{xy}\Phi(x-y) + \nabla_{yy}\Phi(x-y) = 0,$$

which, when used together with the parabolic dissipation terms (like the one found in (6)), allows for a cancellation of certain singular terms involving the second order operator in (1). Compared to [9], our uniqueness proof is new even in the case of bounded entropy solutions.

References

[1] P. Bénilan, J. Carrillo, and P. Wittbold. Renormalized entropy solutions of scalar conservation laws. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 29(2):313–327, 2000.

[2] P. Bénilan and H. Touré. Sur l'équation générale $u_t = a(\cdot, u, \phi(\cdot, u)_x)_x + v$ dans L^1. II. Le problème d'évolution. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 12(6):727–761, 1995.

[3] D. Blanchard, F. Murat, and H. Redwane. Existence and uniqueness of a renormalized solution for a fairly general class of nonlinear parabolic problems. *J. Differential Equations*, 177(2):331–374, 2001.

[4] R. Bürger, S. Evje, and K. H. Karlsen. On strongly degenerate convection-diffusion problems modeling sedimentation-consolidation processes. *J. Math. Anal. Appl.*, 247(2):517–556, 2000.

[5] J. Carrillo. On the uniqueness of the solution of the evolution dam problem. *Nonlinear Anal.*, 22(5):573–607, 1994.

[6] J. Carrillo. Entropy solutions for nonlinear degenerate problems. *Arch. Rational Mech. Anal.*, 147(4):269–361, 1999.

[7] G.-Q. Chen and E. DiBenedetto. Stability of entropy solutions to the Cauchy problem for a class of nonlinear hyperbolic-parabolic equations. *SIAM J. Math. Anal.*, 33(4):751–762 (electronic), 2001.

[8] G.-Q. Chen and K. H. Karlsen. L^1 framework for continuous dependence and error estimates for quasilinear degenerate parabolic equations. *Trans. Amer. Math. Soc.* To appear.

[9] G.-Q. Chen and B. Perthame. Well-posedness for non-isotropic degenerate parabolic-hyperbolic equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 20(4):645–668, 2003.

[10] B. Cockburn and G. Gripenberg. Continuous dependence on the nonlinearities of solutions of degenerate parabolic equations. *J. Differential Equations*, 151(2):231–251, 1999.

[11] R. J. DiPerna and P.-L. Lions. On the Cauchy problem for Boltzmann equations: global existence and weak stability. *Ann. of Math. (2)*, 130(2):321–366, 1989.

[12] R. Eymard, T. Gallouët, R. Herbin, and A. Michel. Convergence of a finite volume scheme for nonlinear degenerate parabolic equations. *Numer. Math.*, 92(1):41–82, 2002.

[13] K. H. Karlsen and M. Ohlberger. A note on the uniqueness of entropy solutions of nonlinear degenerate parabolic equations. *J. Math. Anal. Appl.*, 275(1):439–458, 2002.

[14] K. H. Karlsen and N. H. Risebro. On the uniqueness and stability of entropy solutions of nonlinear degenerate parabolic equations with rough coefficients. *Discrete Contin. Dyn. Syst.*, 9(5):1081–1104, 2003.

[15] S. N. Kružkov. First order quasi-linear equations in several independent variables. *Math. USSR Sbornik*, 10(2):217–243, 1970.

[16] C. Mascia, A. Porretta, and A. Terracina. Nonhomogeneous dirichlet problems for degenerate parabolic-hyperbolic equations. *Arch. Ration. Mech. Anal.*, 163:87–124, 2002.

[17] A. Michel and J. Vovelle. Entropy formulation for parabolic degenerate equations with general Dirichlet boundary conditions and application to the convergence of FV methods. Preprint, 2002.

[18] É. Rouvre and G. Gagneux. Solution forte entropique de lois scalaires hyperboliques-paraboliques dégénérées. *C. R. Acad. Sci. Paris Sér. I Math.*, 329(7):599–602, 1999.

[19] A. I. Vol'pert. The spaces BV and quasi-linear equations. *Math. USSR Sbornik*, 2(2):225–267, 1967.

[20] A. I. Vol'pert and S. I. Hudjaev. Cauchy's problem for degenerate second order quasilinear parabolic equations. *Math. USSR Sbornik*, 7(3):365–387, 1969.

[21] Z. Wu and J. Yin. Some properties of functions in BV_x and their applications to the uniqueness of solutions for degenerate quasilinear parabolic equations. *Northeastern Math. J.*, 5(4):395–422, 1989.

Perturbations of gradient flow / real Ginzburg-Landau systems

JOHN NORBURY

Suppose we are given the (smooth or C^1) functions $k(x), f(x), g(x)$ which are strictly positive in the closure of the bounded connected domain Ω, and we are given the constants $\epsilon > 0$ and $\alpha \in (0,1)$. Then we consider the parabolic system of partial differential equations (henceforth PDEs)

$$(P) \qquad \left\{ \begin{array}{l} \epsilon \frac{\partial u}{\partial t} = \frac{\epsilon^2}{2}\mathrm{div}(k\nabla u) + f^2 u[g^2 - u^2 - \alpha v^2] = -\epsilon \frac{\delta E_\epsilon}{\delta u}, \\ \epsilon \frac{\partial v}{\partial t} = \frac{\epsilon^2}{2}\mathrm{div}(k\nabla v) + f^2 v[g^2 - v^2 - \alpha u^2] = -\epsilon \frac{\delta E_\epsilon}{\delta v}, \end{array} \right.$$

for $x \in \Omega$, where

$$E_\epsilon(u,v) \quad = \quad \int_\Omega \frac{\epsilon}{4}(|\nabla u|^2 + |\nabla v|^2) + \frac{1}{\epsilon} W(x,u,v)dx$$

for $u, v \in H^1(\Omega)$, with u, v satisfying homogeneous Neumann boundary conditions $\frac{\partial u}{\partial n} = \frac{\partial v}{\partial n} = 0$ on $\partial\Omega$ (which is sufficiently smooth, say C^2, for the derivatives to exist). We are interested in the longtime behaviour of the solutions $u(.,t), v(.,t)$ of such systems, when ϵ is small, and in particular the solutions with changing signs in Ω. Thus we define

$$\mathcal{A}_1 = \mathcal{A}_1(\hat{u}) = \{x : \hat{u}(x) = g(x)/\sqrt{1+\alpha}\},\ \mathcal{A}_2 = \mathcal{A}_2(\hat{v}) = \{x : \hat{v}(x) = g(x)/\sqrt{1+\alpha}\}$$

and describe the curves $\mathcal{C}_i := \partial\mathcal{A}_i$ (i=1,2) as the nodal curves of the solution. These curves $\mathcal{C}_i$ define the pattern of the solution because they separate Ω into subdomains $\mathcal{A}_i(\epsilon), \Omega\setminus\overline{\mathcal{A}_i(\epsilon)} =: \mathcal{A}_i^c(\epsilon)$ such that, if $\mathcal{A}_i(\epsilon) \to \mathcal{A}_i(0)$ as $\epsilon \to 0$, then $\frac{u}{g\sqrt{1+\alpha}}, \frac{v}{g\sqrt{1+\alpha}} \to \pm 1$ for $x \in \mathcal{A}_i(0)$ or for $x \in \mathcal{A}_i(0)^c := \Omega \setminus \overline{\mathcal{A}_i(0)}$. See Figures 6 and 7. The

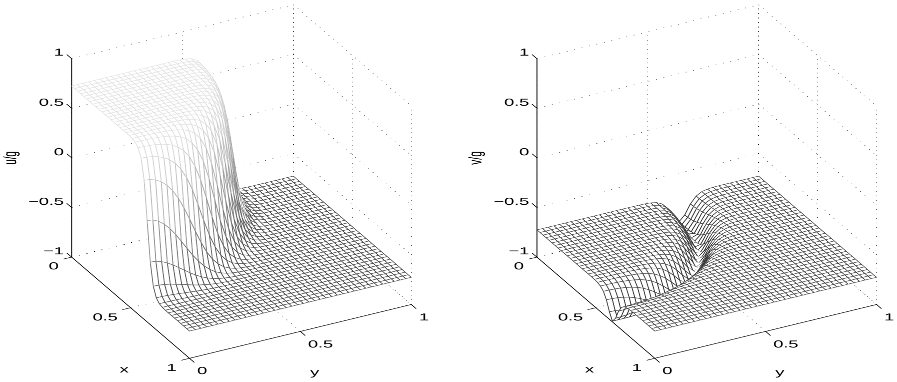


FIGURE 6. Steady solutions to problem (P), with $\Omega = (0,1) \times (0,1)$, $\alpha = 0.9$, $\epsilon = \frac{1}{200}$, $f \equiv k \equiv 1$, and $g(x,y) = \begin{cases} 1 - 3\cosh(\pi r)^2 \exp(-\frac{1}{1-r})\exp(-\frac{1}{r}) & \text{if } r := \sqrt{x^2+y^2} \leq 1; \\ 1 & \text{otherwise.} \end{cases}$
(Steady limit of a numerical simulation of the time-dependent gradient system showing that these local minimisers act as stable attractors for a wide range of initial data with appropriate sign changes.)

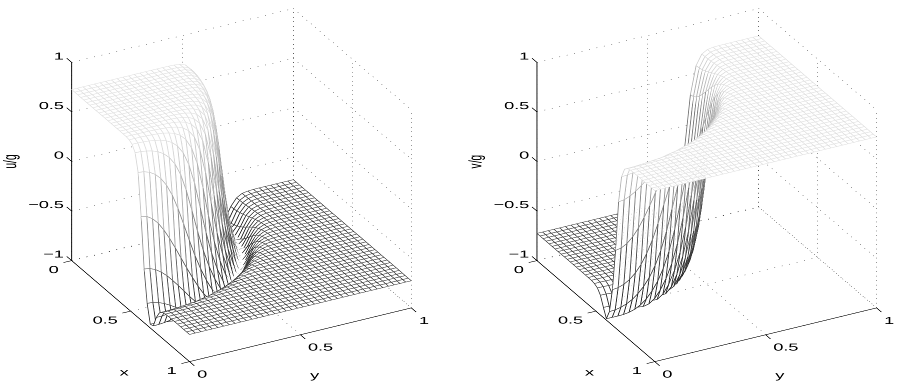


FIGURE 7. Another (stable) solution (with two interfaces) to the same problem as in Figure 6.

talk described the Γ-convergence limit $E_0(u,v)$ of the functional $E_\epsilon(u,v)$ extended to

$u, v \in L^1(\Omega)$ by defining $E_\epsilon(u, v) = \infty$ for $u, v \in L^1(\Omega) \setminus H^1(\Omega)$. In fact, as a simple one-dimensional example for the steady solutions of the following nonlinearly forced heat equation in one space dimension with boundary conditions $u_x(-1) = 0 = u_x(1)$, shows, $u(x)$ tends to a bounded discontinuous limit as $\epsilon \to 0$, and so for appropriate initial data (which of necessity must change sign in Ω), the corresponding time dependent problem

$$\epsilon u_t = \frac{\epsilon^2}{2} u_{xx} + u[g^2 - u^2]$$

will have attracting steady states which become discontinuous in Ω in the limit $\epsilon \to 0$. Hence this example (embedded using the appropriate symmetrical domain Ω) shows that in general the Γ-limit $E_0(u, v)$ cannot be defined on elements u, v in the (Banach) space $H^1(\Omega)$, but may be bounded in the Banach space of functions of bounded variation $BV(\Omega)$, where $\|w\|_{BV} := \|w\|_{L^1(\Omega)} + \int_\Omega |Dw|$

$$\text{and} \qquad \int_\Omega |Dw| := \sup \left\{ \int_\Omega w \operatorname{div} \phi \, dx \ : \phi \in C_0^1(\Omega, \mathbb{R}^2), |\phi| \leq 1 \right\} < \infty.$$

In fact we can show that the $\epsilon \to 0$ limits of solutions u, v of our problem always lie in the subspace $SBV(\Omega)$ (where $H^1(\Omega) \subset SBV(\Omega) \subset BV(\Omega) \subset L^1(\Omega)$), the subspace of special functions of bounded variation that possess no Cantor part in the measure valued derivative (in other words, the singular measure in the generalised derivative always consists of a bounded jump $u_-(x) < u_+(x)$ for $x \in J(u)$, the jump set of Hausdorff dimension one in Ω; note that this jump set has a generalised normal ν and belongs to a generalised curve in the geometric measure theory sense, see [1], [2]).

The key result of Girardet and Norbury [3], that

$$E_0(u, v) = K(\alpha) \left\{ \int_\Omega \sqrt{k} f g^3 |D\chi_{\mathcal{A}_1}| + \int_\Omega \sqrt{k} f g^3 |D\chi_{\mathcal{A}_2}| \right\}$$

for $u, v \in SBV(\Omega)$, is interpreted as an equation for extremal geodesics $\mathcal{C}(0) \subset \Omega$ that may act as stable attractors for our time dependent dynamical system problem when $\mathcal{C}(0)$ are isolated local minimisers of $E_0(u, v)$ in $SBV(\Omega) \times SBV(\Omega)$. These (extremal) geodesics for the domain Ω are calculated from the equation

$$\kappa(x) = -\frac{\partial}{\partial n} \ln h(x)$$

for $x \in \mathcal{C}(0) \subset \Omega$, where $\kappa(x)$ is the (Gaussian) mean curvature of $\mathcal{C}(0)$ at the point x and $\frac{\partial}{\partial n}$ is the normal derivative to $\mathcal{C}(0)$ (given the direction of an increasing arc length parameter as the direction of the tangent, and with the normal then making the usual sense to the tangent), and where the metric $h(x) := \sqrt{k(x)} f(x) g(x)^3$ for $x \in \Omega$, is smooth $(C^1(\bar{\Omega}))$ and strictly positive in $\bar{\Omega}$. Either a geodesic intersects $\partial\Omega$, or otherwise extremal geodesics are periodic, and (at least) C^1 in the arc length parameter (note that geodesics may intersect, including self-intersections). New features of problem (P) that were described include:

(a) "Ridges" in one solution component exist which act as indicators or markers of interfaces and their nodal curves in the other solution component (these ridges are proportional in height to α and persist as $\epsilon \to 0$ in the component (say u) that keeps the same sign when the other component (say v) changes sign, where both the centrelines of such a ridge and such an interface converge to $\mathcal{C}(0)$ as $\epsilon \to 0$).

(b) Stable "double interface" solutions exist whenever a stable single interface solution exists (the double interface is of the type where there is a single interface in each component, and the distance between the single interfaces vanishes as $\epsilon \to 0$ but does not remain uniformly $\mathcal{O}(\epsilon)$).

These new features appear in the solutions of the coupled ordinary differential equation problem

$$0 = \frac{1}{2}u_{xx} + u[1 - u^2 - \alpha v^2], \qquad\qquad 0 = \frac{1}{2}v_{xx} + v[1 - v^2 - \alpha u^2],$$

for $-\infty < x < \infty$, where $u_x, v_x \to 0$ as $|x| \to \infty$; here $u \to \pm(1+\alpha)^{-\frac{1}{2}}$ as $x \to \pm\infty$, while either $v \to (1+\alpha)^{-\frac{1}{2}}$ as $x \to \pm\infty$ in the single interface case, or $v \to \mp(1+\alpha)^{-\frac{1}{2}}$ as $x \to \pm\infty$ in the double interface case (see Girardet and Norbury [4] where this ordinary differential equation problem is further analysed). This problem models the key behaviour in Problem (P) near interfaces when x is a variable measuring stretched (in ϵ) distance normal to the interface, and the u, v are scaled by g.

REFERENCES

[1] L. AMBROSIO, N. FUSCO AND D. PALLARA, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, Oxford University Press, 2000.
[2] I. FONSECA AND S. MÜLLER, *Relaxation of quasiconvex functionals in $BV(\Omega, \mathbb{R}^p)$ for integrands $f(x, u, \nabla u)$*, Arch. Rational Mech. Anal., **123**, 1993, pp 1–49.
[3] C. GIRARDET AND J. NORBURY, *Patterns for a four well gradient elliptic system*, submitted.
[4] C. GIRARDET AND J. NORBURY, *New solutions for perturbed steady Ginzburg-Landau equations in $\mathbb{R}^1$*, submitted.

Well-Posedness of Free Surface Problems In 2D Fluids

DAVID AMBROSE

The main subject of this talk is my recent proof of a long-standing conjecture in fluid dynamics: that the motion of a vortex sheet subject to surface tension is well-posed (for a short time). The method employed was strongly related to numerical methods developed by T. Hou, J. Lowengrub, and M. Shelley [4]. In particular, both analysis and computation of the problem become possible when the problem is reformulated using natural variables and convenient parameterizations. This will be described in more detail below. Even more recently, Nader Masmoudi and I have extended the analytical method to provide a new proof of the well-posedness of two-dimensional water waves without surface tension.

The proofs of well-posedness for both the water wave and for the vortex sheet with surface tension allow a wide class of initial data to be used. The position of the interface and the initial vorticity (which is concentrated on the interface) are taken in Sobolev spaces, and the interface may be of multiple heights. Also, there is no restriction on the size of the initial data; the only requirement is that a natural non-self-intersection condition must be met.

The vortex sheet is the interface between two incompressible, irrotational, inviscid fluids flowing past each other. It is well known that without surface tension, the vortex sheet is ill-posed. It had long been believed that when surface tension is accounted for in the evolution equations, the initial value problem would become well-posed.

At each time, the sheet can be viewed as a curve in the complex plane. The curve, $z,$ is parameterized by a spatial variable, $\alpha,$ and by time, $t.$ The classical vortex sheet evolves according to the Birkhoff-Rott integral,

$$z_t^*(\alpha,t) = \frac{1}{2\pi i}\mathrm{PV}\int_{-\infty}^{\infty} \frac{\gamma(\alpha')}{z(\alpha,t)-z(\alpha',t)}\,d\alpha'. \tag{1}$$

The $*$ denotes complex conjugation; γ is the vortex sheet strength. Notice that γ is not a function of time. This problem has been studied for many years and has been found to be ill-posed. In particular, it exhibits the well-known Kelvin-Helmholtz instability: in the linearization of the evolution equations about equilibrium, Fourier modes with high wave numbers grow without bound. Equation (1) neglects the effect of surface tension at the interface. Surface tension is a restoring force, and when surface tension is accounted for in the equations of motion, Fourier modes of high wave number remain bounded in the linearization. Taking this further, Beale, Hou, and Lowengrub demonstrated that even far from equilibrium, surface tension makes the linearized equations well-posed [5]. For these reasons, it had been conjectured that surface tension makes the full problem well-posed.

The HLS formulation has two important components: first, they compute dependent variables which are naturally related to the surface tension. In particular, surface tension enters the evolution equations in the form $\gamma_t = \frac{1}{\mathrm{We}}\kappa_\alpha,$ where κ is the curvature of the vortex sheet and We is the Weber number. The Weber number is a dimensionless parameter that is inversely proportional to the surface tension; the case without surface tension corresponds to $\mathrm{We} = \infty.$ (Recall that without surface tension, $\gamma_t = 0.$) To simplify this curvature term in the evolution equations, Hou, Lowengrub, and Shelley described the curve by its tangent angle and arclength rather than by the Cartesian variable, $z.$ The notation s will be used for arclength and θ will be used for the tangent angle the curve forms with the horizontal. The strength of this choice of variables lies in the close relationship between curvature and the tangent angle, $\kappa = \theta_\alpha/s_\alpha.$

Second, HLS added an extra tangential velocity to the evolution equation for the vortex sheet. This does not change the shape of the vortex sheet; rather, it only reparameterizes it. The most convenient choice will make the curve always be parameterized by arclength. Since the evolution equation for s_α is

$$s_{\alpha t} = T_\alpha - \theta_\alpha U, \tag{2}$$

where T and U are the tangential and normal velocities of z, T can be chosen to essentially eliminate s_α from the system. That is, setting $s_{\alpha t}$ equal to a function of time yields an equation for $T.$ This essentially reduces the problem by one dependent variable. Rather than evolving $s_\alpha(\alpha, t)$, it is only necessary to keep track of $L(t)$, the length of one period of the sheet.

The choice of tangential velocity also changes the evolution equation for γ. These new terms introduced in the γ_t equation are of a lower order than the term which comes from surface tension.

The main tool in the proof of well-posedness of vortex sheets with surface tension is an energy estimate. This energy estimate is performed after much rewriting of the evolution equations. The equations can be stated as

$$\theta_t = \frac{2\pi^2}{L^2} H(\gamma_\alpha) + P, \tag{3}$$

$$\gamma_t = \frac{2\pi}{L\mathrm{We}} \theta_{\alpha\alpha} + \frac{2\pi^2}{L^2} \gamma H(\gamma\theta_\alpha) + \widetilde{Q}. \tag{4}$$

The Hilbert transform is denoted by H. Since L is a function of t only, notice that the evolution equations are effectively semilinear. After performing the energy estimate, standard methods can be applied to prove well-posedness. The energy functional is related the H^s Sobolev norm of θ and the $H^{s-1/2}$ Sobolev norm of γ.

Also, the analysis described above proves well-posedness in the case where the upper and lower fluids have different densities. With a density difference, the γ_t equation has additional terms, although none of them are significant. (It is worth noting that in the two-density case, the equation for γ_t is actually an integral equation for γ_t; it was proven in [2] that this integral equation can be solved.) A particular case of the two-density problem is when the upper fluid has density equal to zero; this is the water wave. Thus, the work described above establishes well-posedness of the two-dimensional water wave with surface tension.

Without surface tension, well-posedness of the full water wave problem was demonstrated by Wu in [6]. The proof, however, requires significant use of complex analysis (the Riemann mapping theorem in particular). In [3], Nader Masmoudi and I have given a simpler proof of well-posedness for the two-dimensional irrotational water wave. The method resembles the method of [1], but there are important differences. In particular, the variable γ is insufficient in the water wave case. An appropriate new variable is δ, the difference between the Lagrangian tangential velocity and the special tangential velocity: $\delta = \mathbf{W} \cdot \hat{\mathbf{t}} + \frac{\gamma}{2s_\alpha} - T$. Frequently the derivative δ_α is more useful than δ itself. The evolution equation for δ_α can be written

$$\delta_{\alpha t} = -c\theta_\alpha + \psi, \tag{5}$$

where ψ is a collection of terms which are easy to deal with when performing energy estimates. Here, $c = c(\alpha, t)$ is defined by $c = -\nabla p \cdot \hat{\mathbf{n}}$. A necessary condition for well-posedness is that $c(\alpha, t) > 0$. This is a generalization of a condition of G. I. Taylor.

In terms of δ instead of γ, the evolution equation for θ can be written

$$\theta_t = \frac{2\pi}{L} H(\delta_\alpha) + \phi. \tag{6}$$

Here, ϕ is a collection of terms which can be handled routinely in the energy estimates.

Energy estimates for the system (5), (6) can be made with $\delta_\alpha \in H^{s-1/2}$ and $\theta \in H^s$, for s large enough. Similar estimates hold when surface tension is accounted for. It is then found that solutions to the water wave problem without surface tension are the limit of solutions to the water wave problem with surface tension as surface tension goes to zero.

REFERENCES

[1] D. AMBROSE, *Well-posedness of vortex sheets with surface tension*, SIAM J. Math. Anal., **35**, 2003, pp 211–244.

[2] G. BAKER AND D. MEIRON AND S. ORSZAG, *Generalized vortex methods for free-surface flow problems*, J. Fluid Mech., **123**, 1982, pp 477–501.

[3] D. AMBROSE AND N. MASMOUDI, *The zero surface tension limit of 2D water waves*, 2003, Submitted.

[4] T. HOU AND J. LOWENGRUB AND M. SHELLEY, *Removing the stiffness from interfacial flows with surface tension*, J. Comput. Phys., **114**, 1994, pp 312–338.

[5] J. T. BEALE AND T. HOU AND J. LOWENGRUB, *Growth rates for the linearized motion of fluid interfaces away from equilibrium*, Comm. Pure Appl. Math., **46**, 1993, pp 1269–1301.

[6] S. WU, *Well-posedness in Sobolev spaces of the full water wave problem in 2-D*, Invent. Math., **130**, 1997, pp 39–72.

Stability of Periodic Peakons

JONATAN LENELLS

Abstract. The peakons are peaked traveling wave solutions of a nonlinear integrable equation modeling shallow water waves. We give a simple proof of their stability.

AMS SUBJECT CLASSIFICATION (2000): 35Q35, 37K45.
KEYWORDS: Water waves, Peakons, Stability.

INTRODUCTION

The Camassa-Holm equation

$$u_t - u_{txx} + 3uu_x = 2u_x u_{xx} + uu_{xxx}, \tag{0.1}$$

arises as a model for the unidirectional propagation of shallow water waves over a flat bottom, $u(x,t)$ representing the water's free surface in non-dimensional variables. We are concerned with periodic solutions of (0.1), i.e. $u : \mathbb{S} \times [0,T) \to \mathbb{R}$ where $\mathbb{S}$ denotes the unit circle and $T > 0$ is the maximal existence time of the solution. Equation (0.1) was first obtained [6] as an abstract bi-Hamiltonian equation with infinitely many conservation laws and was subsequently derived from physical principles [2]. Equation (0.1) is a re-expression of the geodesic flow in the group of compressible diffeomorphisms of the circle [7], just like the Euler equation is an expression of the geodesic flow in the group of incompressible diffeomorphisms of the torus [1]. This geometric interpretation leads to a proof that equation (0.1) satisfies the Least Action Principle [3]: a state of the system is transformed to another nearby state through a uniquely determined flow that minimizes the energy. For a large class of initial data, equation (0.1) is an infinite-dimensional completely integrable Hamiltonian system: by means of an isospectral problem one can

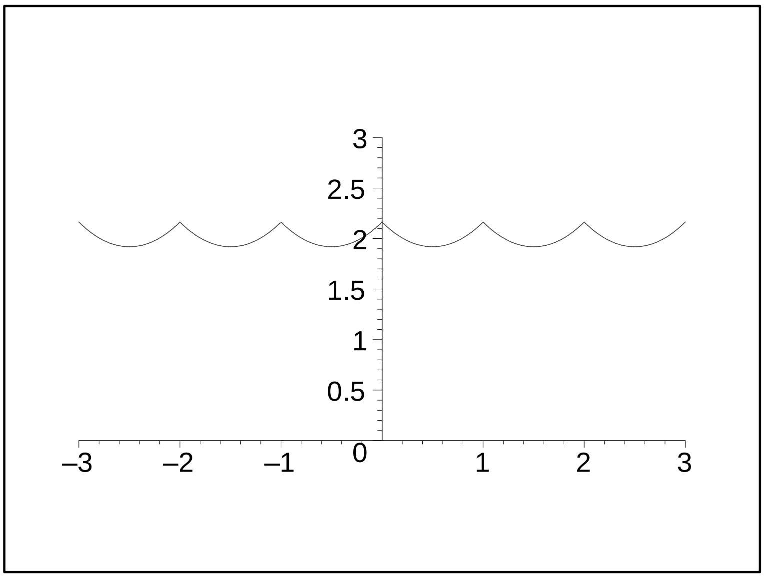


FIGURE 8. The peakon for $c = \frac{\cosh(1/2)}{\sinh(1/2)}$.

convert the equation into an infinite sequence of linear ordinary differential equations which can be trivially integrated [4].

Equation (0.1) has the periodic traveling solution

$$u(x,t) = \frac{c\varphi(x-ct)}{M_\varphi}, \quad c \in \mathbb{R},$$

where $\varphi(x)$ is given for $x \in [0,1]$ by

$$\varphi(x) = \frac{\cosh(1/2 - x)}{\sinh(1/2)}$$

and extends periodically to the real line, and

$$M_\varphi = \max_{x \in S}\{\varphi(x)\} = \frac{\cosh(1/2)}{\sinh(1/2)}.$$

Because of their shape (they are smooth except for a peak at their crest, see Figure 1) these solutions are called (periodic) peakons. Note that the height of the peakon is equal to its speed. Equation (0.1) can be rewritten in conservation form as

$$u_t + \frac{1}{2}\left(u^2 + \varphi * [u^2 + \frac{1}{2}u_x^2]\right)_x = 0. \tag{0.2}$$

This is the exact meaning in which the peakons are solutions.

Numerical simulations suggest that the sizes and velocities of the peakons do not change as a result of collision so that these patterns are expected to be stable. Moreover, for the peakons to be physically observable it is necessary that their shape remains approximately the same as time evolves. Therefore the stability of the peakons is of great interest. We prove the following:

Theorem *The periodic peakons are stable.*

Outline of Proof. Equation (0.1) has the conservation laws

$$H_0[u] = \int_{\mathbb{S}} u dx, \quad H_1[u] = \frac{1}{2}\int_{\mathbb{S}} (u^2 + u_x^2) dx, \quad H_2[u] = \frac{1}{2}\int_{\mathbb{S}} (u^3 + uu_x^2) dx. \tag{0.3}$$

To each solution $u(x,t)$ we associate a function $F_u(M,m)$ of two real variables (M,m) depending only on the three conservation laws H_0, H_1, H_2. Since H_0, H_1, H_2 are conserved quantities, F_u does not depend on time. If we let $M_{u(t)} = \max_{x\in S}\{u(x,t)\}$ and $m_{u(t)} = \min_{x\in S}\{u(x,t)\}$ be the maximum, respectively the minimum of u at the time t, it turns out that

$$F_u(M_{u(t)}, m_{u(t)}) \geq 0, \quad t \in [0,T). \tag{0.4}$$

Moreover, for the peakon we have $F_\varphi(M,m) \leq 0$ with equality only at the point (M_φ, m_φ) (see Figure 2). If u is a solution starting close to φ, the conserved quantities $H_i[u]$ are close to $H_i[\varphi]$, $i = 0,1,2$, and hence F_u is a small perturbation of F_φ. Therefore, the set where $F_u \geq 0$ is contained in a small neighborhood of (M_φ, m_φ). We conclude by (0.4) that $(M_{u(t)}, m_{u(t)})$ stays close to (M_φ, m_φ) for all times. The proof is completed by showing that if the maximum of u stays close to the maximum of the peakon, then the shape of the whole wave remains close to that of the peakon. □

The proof is inspired by [5] where the case of peaked solitary waves of (0.1) is considered. The approach taken here is similar but there are differences. For example, in the case of the peaked solitary waves [5] a polynomial in M plays the role of our function $F_u(M,m)$. The variable m enters in the periodic case because of non-zero boundary conditions.

REFERENCES

[1] V. Arnold, Sur la géometrie différentielle des groupes de Lie de dimension infinie et ses application à l'hydrodynamique des fluides parfaits, *Ann. Inst. Fourier (Grenoble)* **16** (1966), 319–361.

[2] R. Camassa and D. Holm, An integrable shallow water equation with peaked solitons, *Phys. Rev. Lett.* **71** (1993), 1661–1664.

[3] A. Constantin and B. Kolev, On the geometric approach to the motion of inertial mechanical systems, *J. Phys. A* **35** (2002), R51–R79.

[4] A. Constantin and H. P. McKean, A shallow water equation on the circle, *Comm. Pure Appl. Math.* **52** (1999), 949–982.

[5] A. Constantin and W. Strauss, Stability of peakons, *Comm. Pure Appl. Math.* **53** (2000), 603–610.

[6] A. Fokas and B. Fuchssteiner, Symplectic structures, their Bäcklund transformation and hereditary symmetries, *Physica D* **4** (1981), 47–66.

[7] G. Misiolek, A shallow water equation as a geodesic flow on the Bott-Virasoro group, *J. Geom. Phys.* **24** (1998), 203–208.

Well-posedness of KdV on $H^{-1}(\mathbb{T})$

THOMAS KAPPELER
(joint work with Peter Topalov)

Let us consider the Initial Value Problem (IVP) for the Korteweg-deVries equation on the circle

$$
\begin{aligned}
v_t &= -v_{xxx} + 6vv_x \qquad t \in \mathbb{R},\ x \in \mathbb{T} = \mathbb{R}/\mathbb{Z} \\
v\big|_{t=0} &= q \in H^\alpha(\mathbb{T}).
\end{aligned}
$$

This problem has been studied extensively. In particular it is known that for $q \in C^\infty(\mathbb{T})$, the (IVP) admits a unique solution $\mathcal{S}(t,q)$ which exists for all times (see [BS]). Our aim is to solve the (IVP) for very rough initial data such as distributions in the Sobolev space $H^{-1}(\mathbb{T})$.

We say that a continuous curve $\gamma : [T_1, T_2] \to H^\alpha(\mathbb{T})$ with $T_1 < 0 < T_2, \gamma(0) = q$ and $\alpha \in \mathbb{R}$ is a solution of (IVP) if for any $T_1 < t < T_2$ and for any sequence $(q_k)_{k\geq 1} \subseteq C^\infty(\mathbb{T})$ with $q = \lim_{k\to\infty} q_k$ in $H^\alpha(\mathbb{T})$, the solutions $\mathcal{S}(\cdot, q_k)$ have the property that $\gamma(t) = \lim_{k\to\infty} \mathcal{S}(t, q_k)$ in $H^\alpha(\mathbb{T})$. It then follows from the definition of a solution of (IVP) that it is unique whenever it exists. If the solution of (IVP) exists, we denote it by $\mathcal{S}(t,q)$.

The above (IVP) is said to be globally [uniformly] C^0-wellposed on $H^\alpha(\mathbb{T})$ if for any $q \in H^\alpha(\mathbb{T})$ the solution $\mathcal{S}(t,q)$ exists globally in time and the solution map $\mathcal{S}$ is continuous [uniformly continuous on bounded sets] as a map $\mathcal{S} : H^\alpha(\mathbb{T}) \to C^0(\mathbb{R}, H^\alpha(\mathbb{T}))$.

Theorem 1. ([KT1]) KdV is globally C^0-wellposed on $H^\alpha(\mathbb{T})$ for any $-1 \leq \alpha \leq 0$.

Remarks: (1) Theorem 1 improves in particular on earlier results of [Bou1], [Bou2], [KPV], [CKSTT]. Using earlier results, it is proved in [CKSTT] that KdV is globally uniformly C^0-wellposed on $H_0^\alpha(\mathbb{T})$ for any $\alpha \geq -1/2$.

(2) In [CCT] it is shown that KdV is *not* uniformly C^0-wellposed on $H_0^\alpha(\mathbb{T})$ for $-2 < \alpha < -1/2$ where $H_0^\alpha(\mathbb{T}) = \{q \in H^\alpha(\mathbb{T}) \mid \int_{\mathbb{T}} q = 0\}$. See also [Bou2].

The following theorem states that well known features [MT] of solutions of (IVP) for smooth initial data continue to hold for rough initial data.

Theorem 2. ([KT1]) For any $q \in H^\alpha(\mathbb{T})$ with $-1 \leq \alpha \leq 0$, the solution of (IVP) has the following properties:

(i) the orbit $t \mapsto \mathcal{S}(t,q)$ is relatively compact.
(ii) $t \mapsto \mathcal{S}(t,q)$ is almost periodic.

Theorem 1 and Theorem 2 can be applied to obtain corresponding results for the IVP of the modified KdV (mKdV)

$$u_t = -u_{xxx} + 6u^2 u_x \qquad t \in \mathbb{R},\ x \in \mathbb{T}$$

$$u\Big|_{t=0} = r \in H^\alpha(\mathbb{T}).$$

Theorem 3. ([KT2]) mKdV is globally C^0-wellposed on $H^\alpha(\mathbb{T})$ for $0 \le \alpha \le 1$.

Remarks: (1) Theorem 3 improves on earlier results of [Bou1], [KPV], [CKSTT]. Using earlier results it is proved in [CKSTT] that mKdV is globally uniformly C^0-wellposed on $H^\alpha(\mathbb{T})$ for any $\alpha \ge 1/2$.

(2) In [CCT] it is shown that mKdV is *not* uniformly C^0-wellposed on $H_0^\alpha(\mathbb{T})$ for $-1 < \alpha < 1/2$. See also [Bou2].

Besides Theorem 1, the main ingredient of the proof of Theorem 3 is the following result on the Miura map, $B : L^2(\mathbb{T}) \to H^{-1}(\mathbb{T}), r \mapsto r_x + r^2$, first introduced by Miura [Mi] and proved to be a Bäcklund transformation, mapping solutions of mKdV to solutions of KdV.

Theorem 4. ([KT2])

- (i) For any $\alpha \ge 0$, the Miura map $B : H^\alpha(\mathbb{T}) \to H^{\alpha-1}(\mathbb{T})$ is a global fold.
- (ii) Restricted to $H_0^\alpha(\mathbb{T})$, B is a real analytic isomorphism onto the real analytic submanifold $H_0^{\alpha-1}(\mathbb{T}) := \{q \in H^{\alpha-1}(\mathbb{T}) \,|\, \lambda_0(q) = 0\}$ where $\lambda_0(q)$ denotes the lowest eigenvalue in the periodic spectrum of the operator $-d^2/dx^2 + q$.

Remark: Theorem 4 is based on earlier results on the Riccati map [KT3] which used as one of the ingredients estimates on the gaps of the periodic spectrum of impedance operators of [Kor1]. Some of the results in [KT3] have been obtained independently by [Kor2].

The main ingredient in the proof of Theorem 1 is a result on the normal form of the Korteweg-deVries equation considered as an integrable Hamiltonian system. To formulate it, introduce the following model spaces $(\alpha \in \mathbb{R})$

$$h^\alpha := \{(x_k, y_k)_{k\ge1} \,\big|\, x_k, y_k \in \mathbb{R};\ \sum_{k\ge1} k^{2\alpha}(x_k^2 + y_k^2) < \infty\}$$

with the standard Poisson bracket where $\{x_k, y_k\} = 1 = -\{y_k, x_k\}$ and all other brackets between the coordinate functions vanish.

On the space $H_0^\alpha(\mathbb{T}) := \{q = \sum_{k\ne0} \hat{q}_k e^{2\pi ikx} \,\big|\, q \in H^\alpha(\mathbb{T})\}$ we consider the Poisson bracket introduced by Gardner and, independently, by Faddeev and Zakharov

$$\{F, G\} = \int_{\mathbb{T}} \frac{\partial F}{\partial q(x)} \frac{d}{dx} \frac{\partial G}{\partial q(x)} dx.$$

Theorem 5. ([KP], [KMT]) There exists a real analytic diffeomorphism $\Omega : H_0^{-1}(\mathbb{T}) \to h^{-1/2}$ so that

- (i) Ω preserves the Poisson bracket;

(ii) for any $-1 \leq \alpha \leq 0$, the restriction Ω_α of Ω to $H_0^\alpha(\mathbb{T})$ is a real analytic isomorphism, $\Omega_\alpha : H_0^\alpha(\mathbb{T}) \to h^{\alpha+1/2}$;

(iii) on $H_0^1(\mathbb{T})$, the KdV Hamiltonian $\mathcal{H}(q) = \int_{\mathbb{T}}(\frac{1}{2}q_x^2 + q^3)dx$, when expressed in the new coordinates $(x_k, y_k)_{k\geq 1}$, is a real analytic function of the actions $I_k := (x_k^2 + y_k^2)/2$ $(k \geq 1)$ alone.

Remark: In [KP] it is shown that $\Omega_0 : L_0^2 \to h^{1/2}$ is a real analytic isomorphism with properties (i) and (iii). Moreover it is proved that for any $\alpha \in \mathbb{N}$, the restriction Ω_α of Ω to $H_0^\alpha(\mathbb{T})$ is a real analytic isomorphism, $\Omega_\alpha : H_0^\alpha(\mathbb{T}) \to h^{\alpha+1/2}$. This result has been extended in [KMT] as formulated in Theorem 5.

REFERENCES

[BS] J.-L. Bona, R. Smith, *The initial-value problem for the Korteweg-deVries equation*. Phil. Trans. Roy. Soc. London, Series A, Math. and Phys. Sciences, **278**(1975), p. 555-601.

[Bou1] J. Bourgain, *Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations, part II: KdV-equation*. GAFA, **3**(1993), p. 209-262.

[Bou2] J. Bourgain, *Periodic Korteweg-de Vries equation with measures as initial data*. Sel. Math., **3**(1997), p. 115-159.

[CCT] M. Christ, J. Colliander, T. Tao, *Asymptotics, frequency modulation, and low regularity ill-posedness for canonical defocusing equations*. Preprint.

[CKSTT] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Sharp global well-posedness for KdV and modified KdV on* $\mathbb{R}$ *and* $\mathbb{T}$. arXiv:math.AP / 0110045 v2.

[KMT] T. Kappeler, C. Möhr, P. Topalov, *Birkhoff coordinates for KdV on phase spaces of distributions*. Preprint Series, Institute of Mathematics, University of Zurich, 2003.

[KT1] T. Kappeler, P. Topalov, *Global well-posedness of KdV in* $H^{-1}(\mathbb{T}, \mathbb{R})$. Preprint Series, Institute of Mathematics, University of Zurich, 2003.

[KT2] T. Kappeler, P. Topalov, *Global well-posedness of mKdV in* $H^{-1}(\mathbb{T}, \mathbb{R})$. Preprint Series, Institute of Mathematics, University of Zurich, 2003.

[KT3] T. Kappeler, P. Topalov, *Riccati representation for elements in* $H^{-1}(\mathbb{T}^1)$ *and its applications*. Preprint Series, Institute of Mathematics, University of Zurich, 2002; abridged version in Pliska Stud. Math., **15**(2003).

[KP] T. Kappeler, J. Pöschel, *KdV & KAM*. Ergebnisse Math. u. Grenz- gebiete, Springer Verlag, 2003.

[KPV] C. Kenig, P. Ponce, L. Vega, *A bilinear estimate with applications to the KdV equations*. J. Amer. Math. Soc., **9**(1996), p. 573-603.

[Kor1] E. Korotyaev, *Periodic weighted operators*. SFB-288 preprint 388 (1999) and J. of Differential Equations **189**(2003), p 461-486.

[Kor2] E. Korotyaev, *Characterization of spectrum for Schrödinger operator with periodic distributions*. Preprint, 2002; and Int. Math. Res. Not. 37(2003), p 2019 - 2031.

[MT] H. McKean, E. Trubowitz, *Hill's operator and hyperelliptic function theory in the presence of infinitely many branch points*. CPAM, **24**(1976), p. 143-226.

[Mi] R. Miura, *Korteweg-deVries equation and generalizations. I. A remarkable explicit nonlinear transformation*. J. Math. Phys., **9**(1968), p. 1202-1204.

Exact Periodic water waves with Vorticity

WALTER A. STRAUSS

(joint work with Adrian Constantin)

The work presented here is an application of global continuation methods to a classical problem in fluid mechanics. The main tools are Leray-Schauder degree, bifurcation theory, and estimates for elliptic PDEs.

We consider the most classical kind of water wave, traveling at a constant speed $c > 0$. We assume that it is two-dimensional and horizontally periodic with a period L. The water is treated as incompressible and inviscid. We denote the horizontal variable by x and the vertical variable by y. The bottom is assumed to be flat. The surface S has average height $y = 0$. Let S be given by the equation $y = \eta(x - ct)$. Gravity acts on the water with gravitational constant g. The air pressure is assumed to be a constant P_{atm} and it is assumed that there is no surface tension. Let (u, v) denote the velocity, ψ denote the stream function and ω the vorticity. Then $\omega = \gamma(\psi)$ for some function γ.

The special case of irrotational flow, when $\omega = 0$, has been studied much more than the general case because then ψ is a harmonic function and the techniques of complex analysis are readily available. However, we want to focus on the case of general vorticity.

It follows from the preceding equations that the relative mass flux

$$p_0 = \int_{-d}^{\eta(x)} \{u(x, y) - c\} dy$$

is independent of x. We will be looking for waves with $u < c$ and therefore $p_0 < 0$.

Theorem 1 (Main Theorem)**.** *Let $L > 0, c > 0, p_0 < 0$. Let these quantities satisfy Condition A given below. Then there exist traveling waves (u, v, η) of period L, flux p_0 and speed c, with $u < c$, which are symmetric around each crest and trough. In fact there exists a connected set $\mathcal{C}$ of such waves in the space $C2 \times C2 \times C3$ such that*

(i) $\mathcal{C}$ contains a trivial flow with $\eta \equiv 0$ (that is, a flat surface), and

(ii) $\mathcal{C}$ contains waves for which $\max u \nearrow c$ (that is, stagnation).

Condition A. *Let $\Gamma'(p) = \gamma(-p)$ with $\Gamma(0) = 0$. Denote $a(p) = \sqrt{\lambda + 2\Gamma(p)}$ defined for $\lambda > -2\Gamma_{min}$. Condition A requires that for some λ the Sturm-Liouville problem*

$$-(a3M_p)_p = \mu a M, \quad M(p_0) = 0, \quad a^3 M_p(0) = gM(0)$$

has an eigenvalue $\mu \leq -1$.

This condition is necessary for the validity of the theorem. We will construct the continuum $\mathcal{C}$ by bifurcation from the curve of trivial solutions. Condition A is required for the existence of a local bifurcation curve.

The proof is based in part on the following ingredients: (1) a transformation due to Choquet-Bruhat that fixes the free boundary, (2) a local bifurcation argument using the Crandall-Rabinowitz theorem, (3) a global bifurcation argument of Rabinowitz type using the Healey-Simpson degree, (4) a nodal characterization of the solutions using the Hopf and Serrin maximum principles, and (5) regularity theorems of Schauder type due to Lieberman and Trudinger for fully nonlinear elliptic problems.

On Stokes's extreme wave

LUDWIG EDWARD FRAENKEL

Stokes conjectured in 1880 that (in the absence of surface tension and viscosity) the 'highest' gravity wave on water

(i) is distinguished by sharp crests of included angle $2\pi/3$;
(ii) has a profile (by which we mean the free upper boundary of the water) that is convex between successive crests.

Part (i) of this conjecture was proved in 1982 in two quite different ways. In England, Amick, Fraenkel and Toland used the integral equation of Nekrasov and real-variable methods for functions of one variable. In Novosibirsk, Plotnikov used complex-variable methods and an extension of a certain function beyond its domain in the plane of the complex potential. (This was an inspired sharpening for a particular case of a general construction due to H. Lewy.)
Part (ii) of the conjecture has been proved recently by Plotnikov and Toland (who have been collaborators since 1997). The proof uses complex-variable methods of the kind initiated by Plotnikov; it is a tour de force, but far from simple.

The present talk describes two unsuccessful attempts to obtain a relatively simple existence proof for the extreme wave by means of the Nekrasov equation, in the hope that such a proof might yield both parts of the Stokes conjecture more or less directly.

For periodic gravity waves on water of infinite depth, the Nekrasov equation is

$$\phi(s) = (T_\nu \phi)(s) := \frac{1}{3}\int_0^\pi K(s,t)\,\frac{\sin\phi(t)}{\nu + \int_0^t \sin\phi}\,dt, \quad 0 < s \leq \pi,$$

where

$$K(s,t) := \frac{1}{\pi}\log\frac{\tan\frac{1}{2}s + \tan\frac{1}{2}t}{|\tan\frac{1}{2}s - \tan\frac{1}{2}t|},$$

and where $\tan\phi(s) := Y'(x) \geq 0$ is the slope of half a wave-length of the free boundary $\{(x, Y(x))\,|\,x \in \mathbb{R}\}$. The points $s = \pi$ and $s = 0$ correspond to a trough and a crest, respectively. The parameter $\nu \in \left[0, \frac{1}{3}\right)$ and is such that $\frac{1}{3} - \nu$ is small for waves of small amplitude, while $\nu = 0$ for the extreme (or 'highest') wave. In 1978, Toland proved the existence of a suitable solution for $\nu = 0$ by considering a sequence $\left(\phi_{\nu(n)}\right)_{n=1}^\infty$ of solutions for which $\nu(n) \downarrow 0$ as $n \to \infty$.

My first attempt involves the sequence $(\phi_n)_{n=0}^\infty$ of functions defined by $\sin\phi_0(s) := \frac{1}{2}\cos\frac{s}{2}$ $(0 \leq s \leq \pi)$ and $\phi_{n+1} := T_0\phi_n$. The functions ϕ_1 and ϕ_2 are known exactly; ϕ_3 and ϕ_4 are described by formulae which result partly from fitting a trigonometric series to numerical values of the smooth part of $\sin\phi_n(t)/\int_0^t \sin\phi_n$ for $n = 2$ and 3. Graphs of ϕ_0 to ϕ_4 suggest rapid convergence. The leading four terms of $\phi_n(s)$ for $s \downarrow 0$ are known for every n and form a part of the formulae for ϕ_0 to ϕ_4.
However, I have failed to prove convergence.

In the equation $\phi = T_0\phi$ for the extreme wave, let

$$(\mathcal{N}\phi)(s) := \frac{\sin\phi(s)}{\int_0^s \sin\phi}$$

and

$$(\mathcal{K}f)(s) := \int_0^\pi K(s,t)\ f(t)\ dt \qquad (0 < s \le \pi),$$

so that $T_0\phi \equiv \frac{1}{3}\mathcal{K} \circ \mathcal{N}\phi.$ Perhaps the main difficulty of the problem is that $\mathcal{N}\phi$ is not a monotonic function of ϕ (under the usual partial ordering of continuous functions on $(0,\pi])$. However, the inverse not only exists but is an increasing function of $\mathcal{N}\phi$; in fact,

$$\sin\psi(s) = \sin\frac{s}{2}\ \mathcal{N}\psi(s)\ \ \exp\int_0^s \left\{\mathcal{N}\psi(t) - \frac{1}{2}\cot\frac{t}{2}\right\} dt$$

whenever $\mathcal{N}\psi(t) \sim 1/t$ as $t \downarrow 0$, $\mathcal{N}\psi \in C(0,\pi]$ and $\mathcal{N}\psi(s) \ge 0$.

Accordingly, my second attempt has been to pursue $f := \mathcal{N}\phi$, rather than ϕ, by means of the new equation $f = Af$, where

$$\begin{aligned}(Af)(s): \quad &= \quad \frac{\sin(\frac{1}{3}\mathcal{K}f)(s)}{\sin\frac{s}{2}} \quad \exp\textstyle\int_0^s\{f_0 - f\}, \qquad f_0(t) := \mathcal{N}\phi_0(t) = \dfrac{1}{2}\cot\dfrac{t}{2},\\ &= \quad \frac{\sin(T_0\phi)(s)}{\int_0^s \sin\phi} \qquad \text{if} \quad \phi := \mathcal{N}^{-1}f.\end{aligned}$$

An encouraging property of this equation is that its linearization about f_0 is solvable, as follows. If we set $f = f_0 + h$, then, formally,

$$f = Af \Longleftrightarrow h - A'(f_0)h = Af_0 - f_0 + O(h^2).$$

Theorem. *The equation*

$$h - A'(f_0)h = g \quad \textit{in} \quad L_2 := L_2(0,\pi)$$

has a unique solution satisfying

$$\| h \|_{L_2} \le \left(\frac{7}{9} - \frac{3}{4}\log\frac{4}{3}\right)^{-1} \| g \|_{L_2}.$$

Uniqueness issues on permanent progressive water-waves

HISASHI OKAMOTO
(joint work with Kenta Kobayashi)

Abstract. We consider two-dimensional water-waves of permanent shape with constant propagation speed. Two theorems concerning the uniqueness of certain solutions are reported. Uniqueness of Crapper's pure capillary waves is proved under a positivity assumption. The proof is based on the theory of positive operators. Also proved is the uniqueness of the gravity waves of mode one. This is done by a combination of new inequalities and numerical verification algorithm.

Keywords. Crapper's wave, gravity waves, uniqueness, positivity, the Perron-Frobenius theory, verified numerics.

SUMMARY

We consider progressive waves of permanent shape on 2D irrotational flow of incompressible inviscid fluid. For the sake of simplicity, we consider only those fluid flows whose depth are infinite. We show that, under a positivity assumption, the pure capillary waves of Crapper are unique. Also, the positive gravity waves are shown to be unique.

Specifically we consider a solution θ of

$$q\frac{\mathrm{d}\theta}{\mathrm{d}\sigma} = -\sinh(H\theta) \qquad (-\pi \leq \sigma \leq \pi) \tag{1}$$

such that θ is 2π-periodic and satisfies $\theta(-\sigma) = -\theta(\sigma)$. For its meaning, see [5]. In 1957, G.D. Crapper found a family of solutions of (1), which are written, in our context, as follows:

$$q = \frac{1+A2}{1-A2}, \qquad \theta(\sigma) = -2\arctan\left(\frac{2A\sin\sigma}{1-A2}\right).$$

A natural question would be: Does the differential equation (1) has a solution other than Crapper's waves?

Our results is:

Theorem 1. *Suppose that a solution of* (1) *satisfies* $\theta(-\sigma) = -\theta(\sigma)$ *and either the following* **A1** *or* **A2***.*

A1: $0 \leq \theta(\sigma) \leq \pi$ *everywhere in* $0 \leq \sigma \leq \pi$*;*
A2: $\frac{\mathrm{d}\tau}{\mathrm{d}\sigma}(\sigma) \geq 0$ *everywhere in* $0 \leq \sigma \leq \pi$*.*

Then it is one of Crapper's solutions of mode one.

The proof of this theorem depends crucially on [7]. See [6].

We now move on to a uniqueness theorem on the gravity waves, which has been recently obtained by the second author. Now the assumption is that the surface tension is neglected and only the gravity acts. In this case the solutions are obtained by solving the following integral equation, called Nekrasov's equation:

$$\theta(\sigma) = \frac{1}{3\pi}\int_0^\pi \log\left|\frac{\sin\frac{\sigma+s}{2}}{\sin\frac{\sigma-s}{2}}\right| \frac{\mu\sin\theta(s)}{1+\mu\int_0^s \sin(\theta(u))\mathrm{d}u}\mathrm{d}s. \tag{2}$$

Here μ is a new parameter related to the gravity acceleration.

The equation (2) has a rather long history but the structure of the solutions had long been unclear except for those solutions of small amplitude. See [5]. The first satisfactory answer was given by [3] as in the following form:

Theorem 2 (Keady & Norbury, '78)**.** *For all* $3 < \mu < \infty$*, there exists at least one non-trivial solution satisfying* $0 \leq \theta \leq \pi/2$*.*

It is known that there exist solutions which change sign in $0 \leq \sigma \leq \pi$. It is also known that secondary bifurcations exist along such solutions. Therefore uniqueness does not hold among solutions of different signs. However, no secondary bifurcation seems to exist along a positive solutions, and we expect uniqueness for positive solutions. The second author proved in [4] the following

Theorem 3. *For all $3 < \mu \leq 40.0$, there exists at most one non-trivial solution satisfying $0 \leq \theta \leq \pi$.*

The proof in [4] uses the validated numerics or "interval analysis", which gives us exact (i.e., including round-off error) bound for numerical computations.

REFERENCES

[1] B. Chen and P.G. Saffman, Numerical evidence for the existence of new types of gravity waves of permanent form in deep water, Stud. Appl. Math., vol. 62 (1980), pp. 1–21.
[2] G.D. Crapper, J. Fluid Mech., vol. 2 (1957), pp. 532–540.
[3] G. Keady and J. Norbury, On the existence theory for irrotational water waves, Math. Proc. Camb. Phil. Soc., vol. 83 (1978), pp. 137–157.
[4] K. Kobayashi, Numerical verification of the global uniqueness of a positive solution for Nekrasov's equation, to appear in Japan J. Indust. Appl. Math.
[5] H. Okamoto and M. Shōji, The Mathematical Theory of Permanent Progressive Water-Waves, World Scientific, (2001).
[6] H. Okamoto, Uniqueness of Crapper's pure capillary waves of permanent shape, submitted to J. Math. Sci., University of Tokyo,
[7] J.F. Toland, The Peierls-Nabarro and Benjamin-Ono equations, J. Func. Anal., vol. 145, (1997), 136–150.

On the spectral problem associated with the Camassa-Holm equation

CHRISTER BENNEWITZ

INTRODUCTION

Associated with the Camassa-Holm (CH) equation (see [1])

$$u_t - u_{txx} + 3uu_x + 2\varkappa u_x = 2u_x u_{xx} + uu_{xxx},$$

where $x \in \mathbb{R}$, $t \geq 0$ and $\varkappa$ is a parameter, there is the spectral problem

$$-y'' + \tfrac{1}{4}y = \lambda(\varkappa + w(\cdot, t))y, \tag{1}$$

where $w = u - u_{xx}$ and t is viewed as a parameter. There are complications in copying the scattering-inverse scattering approach for the KdV equation to this situation. In particular, an interesting feature of the CH equation is the presence of wave breaking. It is known, however, that this can only occur if $\varkappa + w(x, 0)$ is not of one sign. Standard spectral theory, however, considers (1) in an L^2-space with weight $\varkappa + w$, which is only possible if the weight is of one sign.

One may instead use $H^1(\mathbb{R})$ as the Hilbert space for (1), provided with a slightly modified scalar product

$$\langle y, z \rangle = \int_{\mathbb{R}} (y'\overline{z'} + \tfrac{1}{4}y\overline{z}), \qquad \|y\| = \sqrt{\langle y, y \rangle}.$$

A simple scaling argument shows that one need only consider the cases $\varkappa = 0$ and $\varkappa = 1$. Consider now the case $\varkappa = 1$. There is then a scattering theory for (1), with standard decay assumptions on w. It was proved in [2] [3] that all eigen-values and the transmission coefficient are conserved quantities under the CH flow, and that the reflection coefficient and the normalization constants for the eigenfunctions evolve in a simple, explicit way. Unfortunately, no inverse scattering theory is available unless $1 + w \geq 0$. In this case one may transform (1) to a standard Schrödinger equation, and use the inverse scattering theory then available. This was carried out by Constantin [2] and Lenells [3].

On the other hand, there is a complete spectral theory for (1), and some inverse spectral theory. This is of some use in the case $\varkappa = 0$, as we shall see.

SPECTRAL THEORY

We sketch a general spectral theory for equations of the form

$$-(py')' + qy = \lambda wy \quad \text{in } [0, b), \tag{2}$$

where $p \geq 0$, $q \geq 0$ and $1/p$, q and w are all in $L^1_{\text{loc}}[0, b)$. For simplicity also assume $\operatorname{supp} w = [0, b)$. We study the equation in the completion $\mathcal{H}$ of $C^1_0(0, b)$ with respect to the norm-square $\|y\|^2 = \int_0^b (p|y'|^2 + q|y|^2)$. Let $\varphi(x, \lambda)$ be the solution of (2) with initial data $\varphi(0, \lambda) = 0$, $p\varphi'(0, \lambda) = 1$. There then exists a uniquely determined positive measure $d\rho$ on $\mathbb{R}$, called the *spectral measure*, some of the properties of which are as follows. Let L^2_ρ be the Hilbert space of functions $\hat{y}$ measurable $(d\rho)$ and such that $\int |\hat{y}|^2 \, d\rho < \infty$. Given $y \in \mathcal{H}$ the integral $\hat{y}(t) = \int_0^b (py'\varphi'(\cdot, t) + qy\varphi(\cdot, t))$ converges in L^2_ρ and gives a unitary map $\mathcal{F} : \mathcal{H} \ni y \mapsto \hat{y} \in L^2_\rho$, the *generalized Fourier transform* for (2). The spectrum of the operator corresponding to (2) is $\operatorname{supp} d\rho$, eigenvalues corresponding to point-masses in the measure. We have the following inverse spectral theorem.

Theorem 0.1. Suppose the interval $[0, b)$ and the coefficients p and q are given. Then the spectral measure determines the coefficient w uniquely.

If the spectrum is discrete with eigenvalues λ_n and we define the normalization constants $c_n = \|\varphi(\cdot, \lambda_n)\|^{-2}$, then $d\rho = \sum c_n \delta_{\lambda_n}$. So, in this case knowing the spectral measure is equivalent to knowing all eigenvalues and normalization constants.

For a brief indication of the proof of the theorem, assume that two coefficients w and $\tilde{w}$ give the same spectral measure, and let $\mathcal{U} = \tilde{\mathcal{F}}^{-1}\mathcal{F}$, $\mathcal{F}$ and $\tilde{\mathcal{F}}$ being the generalized Fourier transforms associated with w and $\tilde{w}$ respectively. Then $\mathcal{U}$ is a unitary operator on $\mathcal{H}$, and we are done if we can prove that it is the identity. It is not hard to see that this follows if we can prove that $\mathcal{U}$ preserves supports. To see that it does, one may use a generalization of the classical Paley-Wiener theorem, valid for the generalized Fourier transforms used here.

APPLICATION TO THE CAMASSA-HOLM EQUATION

We consider (1) with $\varkappa = 0$ on $(-\infty, \infty)$, where w is locally integrable. We may transform this problem using a *Liouville transform*, introducing new independent and dependent variables $\xi(x) = e^{-x}$ and $\tilde{y}(\xi) = e^{-x/2}y(x)$. If $y \in \mathcal{H}$ we obtain

$$\int_{\mathbb{R}} (|y'|^2 + \tfrac{1}{4}|y|^2) = \int_0^\infty |\tilde{y}'|^2.$$

The equation (1) is transformed to $-\tilde{y}'' = \lambda\tilde{w}\tilde{y}$ where $\tilde{w}(\xi) = e^{2x}w(x)$. The spectral theory sketched above applies if $\tilde{w}$ is integrable near 0. This translates into the requirement that $e^x w(x)$ is integrable near $+\infty$.

Assuming $(1 + |x|)w(x) \in L^1(\mathbb{R})$ and $\varkappa = 0$, the spectrum of (1) is discrete and the equation has a solution $f_+(x, \lambda)$ asymptotic to $e^{-x/2}$ at $+\infty$ for any λ. It is easy to see that f_+ transforms to φ, and f_+ will be in $\mathcal{H}$ precisely if $\lambda = \lambda_n$ is an eigen-value. Define the corresponding normalization constant

$$c_n = \Big(\int_{\mathbb{R}} (|f'_+(\cdot, \lambda_n)|^2 + \tfrac{1}{4}|f_+(\cdot, \lambda_n)|^2) \Big)^{-1}. \tag{3}$$

Eigen-values are still conserved under the CH flow, and the normalization constants evolve according to $c_n(t) = c_n(0)\exp(-t/4\lambda_n)$. Clearly $c_n = \|\varphi(\cdot, \lambda_n)\|^{-2}$, so Theorem 0.1 gives the following theorem, which is at least a step in the direction of a scattering-inverse scattering approach for CH.

Theorem 0.2. Assume that $\varkappa = 0$ and $(1 + e^x)w(x) \in L^1(\mathbb{R})$. Then (1) has discrete spectrum and the eigenvalues and normalization constants (3) determine w uniquely.

References

[1] Camassa, R. and Holm, D.: *An integrable shallow water equation with peaked solitons.* Phys. Rev. Lett. **71** (1993), 1661–1664.

[2] Constantin, A.: *On the scattering problem for the Camassa-Holm equation.* Proc. Roy. Soc. London A (2001), 953–970.

[3] Lenells, J.: *The scattering approach for the Camassa-Holm equation.* J. Nonlinear Math. Phys. Vol. 9 Nr. 4 (Nov 2002).

Algebro-Geometric Solutions of the KdV and Camassa-Holm equation

Helge Holden

(joint work with Fritz Gesztesy)

The KdV hierarchy

To construct the KdV hierarchy, one assumes u to be a smooth function on $\mathbb{R}$ (or meromorphic in $\mathbb{C}$) in the stationary context or a smooth function on $\mathbb{R}^2$ in the time-dependent case, and one introduces the recursion relation for some functions f_ℓ of u by

$$f_0 = 1, \quad f_{\ell,x} = -(1/4)f_{\ell-1,xxx} + uf_{\ell-1,x} + (1/2)u_x f_{\ell-1}, \quad \ell \in \mathbb{N}. \tag{1}$$

Given the recursively defined sequence $\{f_\ell\}_{\ell\in\mathbb{N}_0}$ (whose elements turn out to be differential polynomials with respect to u, defined up to certain integration constants) one defines the Lax pair of the KdV hierarchy by

$$L = -\frac{d^2}{dx^2} + u, \tag{2}$$

$$P_{2n+1} = \sum_{\ell=0}^{n} \Big(f_{n-\ell} \frac{d}{dx} - \frac{1}{2} f_{n-\ell,x} \Big) L^{\ell}. \tag{3}$$

The commutator of P_{2n+1} and L then reads

$$[P_{2n+1}, L] = 2f_{n+1,x}, \tag{4}$$

using the recursion (1). Introducing a deformation (time) parameter $t_n \in \mathbb{R}$, $n \in \mathbb{N}_0$ into u, the *KdV hierarchy* of nonlinear evolution equations is then defined by imposing the *Lax commutator relations*

$$\frac{d}{dt_n} L - [P_{2n+1}, L] = 0, \tag{5}$$

for each $n \in \mathbb{N}_0$. By (4), the latter are equivalent to the collection of evolution equations

$$\mathrm{KdV}_n(u) = u_{t_n} - 2f_{n+1,x}(u) = 0, \quad n \in \mathbb{N}_0. \tag{6}$$

Explicitly,

$$\begin{aligned} \mathrm{KdV}_0(u) &= u_{t_0} - u_x = 0, \\ \mathrm{KdV}_1(u) &= u_{t_1} + \tfrac{1}{4} u_{xxx} - \tfrac{3}{2} u u_x - c_1 u_x = 0, \\ \mathrm{KdV}_2(u) &= u_{t_2} - \tfrac{1}{16} u_{xxxxx} + \tfrac{5}{8} u u_{xxx} + \tfrac{5}{4} u_x u_{xx} - \tfrac{15}{8} u^2 u_x \\ &\qquad + c_1 \big(\tfrac{1}{4} u_{xxx} - \tfrac{3}{2} u u_x \big) - c_2 u_x = 0, \quad \text{etc.,} \end{aligned} \tag{7}$$

represent the first few equations of the time-dependent KdV hierarchy.

We construct a special class of explicitly defined solutions given by the Its–Matveev formula

$$u(x, t_n) = \Lambda_0 - 2\partial_x^2 \ln(\theta(\underline{A} + \underline{B}x + \underline{C}_r t_n)), \tag{8}$$

Here $\Lambda_0, \underline{A}, \underline{B}, \underline{C}_r$ are all constants, and θ is Riemann's theta function. Observe that the argument in the theta-function is linear both in space and time.

The Camassa–Holm hierarchy

The Camassa–Holm (CH) equation reads

$$4u_t - u_{xxt} - 2u u_{xxx} - 4u_x u_{xx} + 24 u u_x = 0, \quad (x, t) \in \mathbb{R}^2 \tag{9}$$

(chosing a scaling of x, t that's convenient for our purpose), with u representing the fluid velocity in x-direction. Actually, (9) represents the limiting case $\kappa \to 0$ of the general Camassa–Holm equation,

$$4v_t - v_{xxt} - 2v v_{xxx} - 4v_x v_{xx} + 24 v v_x + 4\kappa v_x = 0, \quad \kappa \in \mathbb{R}, \ (x, t) \in \mathbb{R}^2. \tag{10}$$

However, in our formalism the general Camassa–Holm equation (10) just represents a linear combination of the first two equations in the CH hierarchy and hence we consider without loss of generality (9) as the first nontrivial element of the Camassa–Holm hierarchy. Alternatively, one can transform

$$v(x, t) \mapsto u(x, t) = v(x - (\kappa/2)t, t) + (\kappa/4) \tag{11}$$

and thereby reduce (10) to (9).

We start by formulating the basic polynomial setup. One defines $\{f_\ell\}_{\ell\in\mathbb{N}_0}$ recursively by

$$f_0 = 1,$$
$$f_{\ell,x} = -2\mathcal{G}\big(2(4u - u_{xx})f_{\ell-1,x} + (4u_x - u_{xxx})f_{\ell-1}\big), \quad \ell \in \mathbb{N}, \tag{12}$$

where $\mathcal{G}$ is given by

$$\mathcal{G}\colon L^\infty(\mathbb{R}) \to L^\infty(\mathbb{R}), \quad (\mathcal{G}v)(x) = \frac{1}{4}\int_{\mathbb{R}} dy\, e^{-2|x-y|}v(y), \quad x \in \mathbb{R},\ v \in L^\infty(\mathbb{R}). \tag{13}$$

One observes that $\mathcal{G}$ is the resolvent of minus the one-dimensional Laplacian at energy parameter equal to -4, that is,

$$\mathcal{G} = \left(-\frac{d^2}{dx^2} + 4\right)^{-1}. \tag{14}$$

The first coefficient reads

$$f_1 = -2u + c_1, \tag{15}$$

where c_1 is an integration constant. Subsequent coefficients are nonlocal with respect to u. At each level a new integration constant, denoted by c_ℓ, is introduced. Moreover, we introduce coefficients $\{g_\ell\}_{\ell\in\mathbb{N}_0}$ and $\{h_\ell\}_{\ell\in\mathbb{N}_0}$ by

$$g_\ell = f_\ell + \frac{1}{2}f_{\ell,x}, \quad \ell \in \mathbb{N}_0, \tag{16}$$
$$h_\ell = (4u - u_{xx})f_\ell - g_{\ell+1,x}, \quad \ell \in \mathbb{N}_0. \tag{17}$$

Explicitly, one computes

$$\begin{aligned}
f_0 &= 1,\\
f_1 &= -2u + c_1,\\
f_2 &= 2u^2 + 2\mathcal{G}\big(u_x^2 + 8u^2\big) + c_1(-2u) + c_2,\\
g_0 &= 1,\\
g_1 &= -2u - u_x + c_1,\\
g_2 &= 2u^2 + 2uu_x + 2\mathcal{G}\big(u_x^2 + u_xu_{xx} + 8uu_x + 8u^2\big)\\
&\quad + c_1(-2u - u_x) + c_2,\\
h_0 &= 4u + 2u_x,\\
h_1 &= -2u_x^2 - 4uu_x - 8u^2\\
&\quad - 2\mathcal{G}\big(u_xu_{xxx} + u_{xx}^2 + 2u_xu_{xx} + 8uu_{xx} + 8u_x^2 + 16uu_x\big)\\
&\quad + c_1(4u + 2u_x), \text{ etc.}
\end{aligned} \tag{18}$$

Next one introduces the 2×2 matrix U by

$$U(z,x) = \begin{pmatrix} -1 & 1 \\ z^{-1}(4u(x) - u_{xx}(x)) & 1 \end{pmatrix}, \quad x \in \mathbb{R}, \tag{19}$$

and for each $n \in \mathbb{N}_0$ the following 2×2 matrix V_n by

$$V_n(z,x) = \begin{pmatrix} -G_n(z,x) & F_n(z,x) \\ z^{-1}H_n(z,x) & G_n(z,x) \end{pmatrix}, \quad n \in \mathbb{N}_0,\ z \in \mathbb{C} \setminus \{0\},\ x \in \mathbb{R}, \tag{20}$$

assuming F_n, G_n, and H_n to be polynomials of degree n with respect to z and C^∞ in x. Postulating the zero-curvature condition

$$-V_{n,x}(z,x) + [U(z,x), V_n(z,x)] = 0, \tag{21}$$

one finds

$$F_{n,x}(z,x) = 2G_n(z,x) - 2F_n(z,x), \tag{22}$$

$$zG_{n,x}(z,x) = (4u(x) - u_{xx}(x))F_n(z,x) - H_n(z,x), \tag{23}$$

$$H_{n,x}(z,x) = 2H_n(z,x) - 2(4u(x) - u_{xx}(x))G_n(z,x). \tag{24}$$

From (22)–(24) one infers that

$$\frac{d}{dx}\det(V_n(z,x)) = -\frac{1}{z}\frac{d}{dx}\Big(zG_n(z,x)^2 + F_n(z,x)H_n(z,x)\Big) = 0, \tag{25}$$

and hence

$$zG_n(z,x)^2 + F_n(z,x)H_n(z,x) = Q_{2n+1}(z), \tag{26}$$

where the polynomial Q_{2n+1} of degree $2n+1$ is x-independent. Actually it turns out that it is more convenient to define

$$R_{2n+2}(z) = zQ_{2n+1}(z) = \prod_{m=0}^{2n+1}(z - E_m), \quad E_0 = 0,\ E_1, \dots, E_{2n+1} \in \mathbb{C} \tag{27}$$

so that (26) becomes

$$z^2G_n(z,x)^2 + zF_n(z,x)H_n(z,x) = R_{2n+2}(z). \tag{28}$$

Next one makes the ansatz that F_n, H_n, and G_n are polynomials of degree n, related to the coefficients f_ℓ, h_ℓ, and g_ℓ by

$$F_n(z,x) = \sum_{\ell=0}^{n} f_{n-\ell}(x)z^\ell, \tag{29}$$

$$G_n(z,x) = \sum_{\ell=0}^{n} g_{n-\ell}(x)z^\ell, \tag{30}$$

$$H_n(z,x) = \sum_{\ell=0}^{n} h_{n-\ell}(x)z^\ell. \tag{31}$$

Insertion of (29)–(31) into (22)–(24) then yields the recursion relations (12)–(13) and (16) for f_ℓ and g_ℓ for $\ell = 0, \dots, n$. For fixed $n \in \mathbb{N}$ we obtain the recursion (17) for h_ℓ for $\ell = 0, \dots, n-1$ and

$$h_n = (4u - u_{xx})f_n. \tag{32}$$

(When $n=0$ one directly gets $h_0=(4u-u_{xx})$.) Moreover, taking $z=0$ in (28) yields

$$f_n(x)h_n(x)=-\prod_{m=1}^{2n+1}E_m. \tag{33}$$

In addition, one finds

$$h_{n,x}(x)-2h_n(x)+2(4u(x)-u_{xx}(x))g_n(x)=0, \quad n\in\mathbb{N}_0. \tag{34}$$

Using the relations (16) and (32) permits one to write (34) as

$$\text{s-CH}_n(u)=(u_{xxx}-4u_x)f_n-2(4u-u_{xx})f_{n,x}=0, \quad n\in\mathbb{N}_0. \tag{35}$$

Varying $n\in\mathbb{N}_0$ in (35) then defines the stationary CH hierarchy. We record the first few equations explicitly,

$$\begin{aligned}
&\text{s-CH}_0(u)=u_{xxx}-4u_x=0,\\
&\text{s-CH}_1(u)=-2uu_{xxx}-4u_xu_{xx}+24uu_x+c_1(u_{xxx}-4u_x)=0,\\
&\text{s-CH}_2(u)=2u^2u_{xxx}-8uu_xu_{xx}-40u^2u_x+2(u_{xxx}-4u_x)\mathcal{G}\big(u_x^2+8u^2\big)\\
&\qquad -8(4u-u_{xx})\mathcal{G}\big(u_xu_{xx}+8uu_x\big)\\
&\qquad +c_1(-2uu_{xxx}-4u_xu_{xx}+24uu_x)+c_2(u_{xxx}-4u_x)=0, \text{ etc.}
\end{aligned} \tag{36}$$

Next, we turn to the time-dependent CH hierarchy. Introducing a deformation parameter $t_n\in\mathbb{R}$ into u (replacing $u(x)$ by $u(x,t_n)$), the definitions (19), (20), and (29)–(31) of U, V_n, and F_n, G_n, and H_n, respectively, still apply. The corresponding zero-curvature relation reads

$$U_{t_n}(z,x,t_n)-V_{n,x}(z,x,t_n)+[U(z,x,t_n),V_n(z,x,t_n)]=0, \quad n\in\mathbb{N}_0, \tag{37}$$

which results in the following set of equations

$$\begin{aligned}
&4u_{t_n}(x,t_n)-u_{xxt_n}(x,t_n)-H_{n,x}(z,x,t_n)+2H_n(z,x,t_n)\\
&-2(4u(x,t_n)-u_{xx}(x,t_n))G_n(z,x,t_n)=0,
\end{aligned} \tag{38}$$

$$F_{n,x}(z,x,t_n)=2G_n(z,x,t_n)-2F_n(z,x,t_n), \tag{39}$$

$$zG_{n,x}(z,x,t_n)=(4u(x,t_n)-u_{xx}(x,t_n))F_n(z,x,t_n)-H_n(z,x,t_n). \tag{40}$$

Inserting the polynomial expressions for F_n, H_n, and G_n into (39) and (40), respectively, first yields recursion relations (12) and (16) for f_ℓ and g_ℓ for $\ell=0,\dots,n$. For fixed $n\in\mathbb{N}$ we obtain from (38) the recursion (17) for h_ℓ for $\ell=0,\dots,n-1$ and

$$h_n=(4u-u_{xx})f_n. \tag{41}$$

(When $n=0$ one directly gets $h_0=(4u-u_{xx})$.) In addition, one finds

$$\begin{aligned}
&4u_{t_n}(x,t_n)-u_{xxt_n}(x,t_n)-h_{n,x}(x,t_n)+2h_n(x,t_n)\\
&-2(4u(x,t_n)-u_{xx}(x,t_n))g_n(x,t_n)=0, \quad n\in\mathbb{N}_0.
\end{aligned} \tag{42}$$

Using relations (16) and (41) permits one to write (42) as

$$\mathrm{CH}_n(u)=4u_{t_n}-u_{xxt_n}+(u_{xxx}-4u_x)f_n-2(4u-u_{xx})f_{n,x}=0, \quad n\in\mathbb{N}_0. \tag{43}$$

Varying $n \in \mathbb{N}_0$ in (43) then defines the time-dependent CH hierarchy. We record the first few equations explicitly,

$$\begin{aligned}
&\mathrm{CH}_0(u) = 4u_{t_0} - u_{xxt_0} + u_{xxx} - 4u_x = 0,\\
&\mathrm{CH}_1(u) = 4u_{t_1} - u_{xxt_1} - 2uu_{xxx} - 4u_x u_{xx} + 24uu_x + c_1(u_{xxx} - 4u_x) = 0,\\
&\mathrm{CH}_2(u) = 4u_{t_2} - u_{xxt_2} + 2u^2 u_{xxx} - 8uu_x u_{xx} - 40u^2 u_x\\
&\qquad + 2(u_{xxx} - 4u_x)\mathcal{G}\big(u_x^2 + 8u^2\big) - 8(4u - u_{xx})\mathcal{G}\big(u_x u_{xx} + 8uu_x\big)\\
&\qquad + c_1(-2uu_{xxx} - 4u_x u_{xx} + 24uu_x) + c_2(u_{xxx} - 4u_x) = 0, \text{ etc.}
\end{aligned} \tag{44}$$

We show the analogue of the Its–Matveev formula for the CH hierarchy. Here we find

$$u(x,t_n) = A + \sum_{j=1}^{n} U_j \frac{\partial}{\partial w_j} \ln\left(\frac{\theta\big(\underline{z}(P_{\infty_+}, \underline{\hat{\mu}}(x,t_r)) + \underline{w}\big)}{\theta\big(\underline{z}(P_{\infty_-}, \underline{\hat{\mu}}(x,t_r)) + \underline{w}\big)} \right)\Bigg|_{\underline{w}=0}. \tag{45}$$

Here $(U_1, \ldots, U_n)$ is a constant, and

$$\underline{\hat{z}}(P, \underline{Q}) = B(P) + \underline{\hat{\alpha}}(\mathcal{D}_{\underline{Q}}), \tag{46}$$

where $B(P)$ is a constant, and $\underline{\hat{\alpha}}$ is the Abel map, and $\mathcal{D}_{\underline{Q}}$ is a divisor at $\underline{Q}$. Finally, $\underline{\hat{\mu}}$ is the set of solutions of the Dubrovin equations. All constants can be explicitly computed in terms of quantities of a hyperelliptic curve. Unfortunately, the argument inside the theta function is not linear in the space and time variable.

Extensive background information and complete details as well as references to the early literature can be found in [1].

References

[1] F. Gesztesy and H. Holden. *Soliton Equations and Their Algebro-Geometric Solutions. Vol. I:* $(1+1)$-*Dimensional Continuous Models*. Cambridge Studies in Advanced Mathematics, Cambridge Univ. Press, Cambridge, 2003.

Lie Groups and Mechanics: an introduction

Boris Kolev

Euler equation of a rigid body

In classical mechanics, a *material system* (Σ) in the ambient space $\mathbb{R}^3$ is described by a *positive measure* μ on $\mathbb{R}^3$ with compact support. This measure is called the *mass distribution* of (Σ).

In the *Lagrangian formalism of Mechanics*, a *motion* of a material system is described by a smooth path φ^t of *embeddings* of the *reference state* $\Sigma = Supp(\mu)$ in the ambient space. A material system (Σ) is *rigid* if each map φ is the restriction to Σ of an isometry g of the Euclidean space $\mathbb{R}^3$.

In what follows, we are going to study the motions of a rigid body (Σ) such that $\Sigma = Supp(\mu)$ spans the 3 space. In that case, the manifold of all possible configurations of (Σ) is completely described by the group $\mathcal{D}_3$ of orientation-preserving isometries of $\mathbb{R}^3$.

Although the physically meaningful rigid body mechanics is in dimension 3, we will not use this peculiarity in order to distinguish easier the main underlying concepts. Hence, in what follows, we will study the motion of an n-dimensional rigid body. To avoid heavy computations, we will restrain our study to motions of a rigid body having a fixed point. In these circumstances, the configuration space reduces to the group $SO(n)$ of isometries which fix a point.

The Lie algebra $\mathfrak{so}(n)$ of $SO(n)$ is the space of all skew-symmetric $n \times n$ matrices[1]. There is a canonical inner product, the so-called *Killing form*

$$\langle \Omega_1, \Omega_2 \rangle = -\frac{1}{2}\ \mathrm{tr}(\Omega_1 \Omega_2)$$

which permit us to identify $\mathfrak{so}(n)$ with its dual space $\mathfrak{so}(n)^*$. For x and y in $\mathbb{R}^n$, we define

$$L^*(x,y)(\Omega) = (\Omega\, x) \cdot y, \quad \Omega \in \mathfrak{so}(n)$$

which is skew-symmetric in x, y and thus defines a linear map

$$L^* : \bigwedge^2 \mathbb{R}^n \to \mathfrak{so}(n)^* \ .$$

This map is injective and defines therefore an isomorphism between $\mathfrak{so}(n)^*$ and $\bigwedge^2 \mathbb{R}^n$, which have the same dimension. We let $L(x,y)$ be the corresponding element of $\mathfrak{so}(n)$ (using the Killing form).

The location of a point a of the body Σ is described by the column vector r of its coordinates in the frame $\Re_0$. At time t, this point occupies a new position $r(t) = g(t)r$, where $g(t)$ is an element of the group $SO(3)$ and its velocity is given by $\mathbf{v}(a,t) = \dot{g}(t)\, r$. The kinetic energy K of the body Σ at time t is defined by

$$K(t) = \frac{1}{2}\int_\Sigma \|\mathbf{v}(a,t)\|^2 \ d\mu = \frac{1}{2}\int_\Sigma \|\dot{g}\, r\|^2 \ d\mu = \frac{1}{2}\int_\Sigma \|\Omega r\|^2 \ d\mu \tag{1}$$

where $\Omega = g^{-1}\,\dot{g}$ lies in the Lie algebra $\mathfrak{so}(n)$.

Lemma 1. *We have $K = -\frac{1}{2}\ \mathrm{tr}(\Omega J \Omega)$, where J is the symmetric matrix with entries*

$$J_{ij} = \int_\Sigma x_i x_j \ d\mu \ .$$

The kinetic energy K is therefore a positive quadratic form on the Lie algebra $\mathfrak{so}(n)$. To K, a linear operator $A : \mathfrak{so}(n) \to \mathfrak{so}(n)$, called the *inertia tensor* or the *inertia operator*, is associated by means of the relation

$$K = \frac{1}{2}\langle A(\Omega), \Omega \rangle\, , \qquad \Omega \in \mathfrak{so}(n).$$

More precisely, this operator is given by

$$A(\Omega) = J\Omega + \Omega J = \int_\Sigma \left(\Omega\, r r^t + r r^t \Omega\right) \, d\mu \ . \tag{2}$$

[1]In dimension 3, we generally identify the Lie algebra $\mathfrak{so}(3)$ with $\mathbb{R}^3$ endowed with the Lie bracket given by the cross product $\omega_1 \times \omega_2$.

The *angular momentum* of the rigid body is defined by the following 2-vector

$$\mathcal{M}(t) = \int_\Sigma (gr) \wedge (\dot{g}r)\; d\mu \;.$$

Lemma 2. *We have* $L(\mathcal{M}) = gA(\Omega)g^{-1}$.

If there are no external actions on the body, the spatial angular momentum is a constant of the motion,

$$\frac{d\mathcal{M}}{dt} = 0\;. \tag{3}$$

Coupled with the relation $L(\mathcal{M}) = gA(\Omega)g^{-1}$, we deduce that

$$A(\dot{\Omega}) = A(\Omega)\Omega - \Omega A(\Omega) \tag{4}$$

which is the generalization in n dimensions of the traditional *Euler equation*. Notice that if we let $M = A(\Omega)$, this equation can be rewritten as

$$\dot{M} = [M, \Omega\,]\,. \tag{5}$$

GENERAL ARNOLD-EULER EQUATION

A *Riemannian* or *pseudo-Riemannian* metric on a Lie group G is left invariant if it is preserved under every left shift L_g, that is,

$$\langle X_g, Y_g\,\rangle_g = \langle L_h\,X_g, L_h\,Y_g\,\rangle_{hg}\,, \qquad g,h \in G.$$

A left-invariant metric is uniquely defined by its restriction to the tangent space to the group at the unity and hence by a quadratic form on the Lie algebra of the group, $\mathfrak{g}$. To such a quadratic form on $\mathfrak{g}$, correspond a symmetric operator $A : \mathfrak{g} \to \mathfrak{g}^*$ defined by

$$\langle \xi, \omega\,\rangle = (A\xi, \omega\,) = (A\omega, \xi\,)\,, \qquad \xi, \omega \in \mathfrak{g}\,.$$

The operator A is called the *inertia operator*. It can be extended to a left-invariant tensor $A_g : T_gG \to T_gG^*$ defined by $A_g = L^*_{g^{-1}}AL_{g^{-1}}$.

The geodesics of the metric are defined as extremals of the *Lagrangian*

$$\mathcal{L}(g) = \int K\left(g(t), \dot{g}(t)\right)\, dt \tag{6}$$

where

$$K(X) = \frac{1}{2}\,\langle X_g, X_g\,\rangle_g = \frac{1}{2}\,(A_g\,X_g, X_g\,)_g$$

is called the *kinetic energy* or *energy functional*.

If $g(t)$ is a geodesic, the velocity $\dot{g}(t)$ can be translated to the identity via left or right shifts and we obtain two elements of the Lie algebra $\mathfrak{g}$,

$$\omega_L = L_{g^{-1}}\dot{g}, \qquad \omega_R = R_{g^{-1}}\dot{g},$$

called the *left angular velocity*, respectively the *right angular velocity*. Letting $m = A_g\,\dot{g} \in T_gG^*$, we define the *left angular momentum* m_L and the *right angular momentum* m_R by

$$m_L = L^*_g m \in \mathfrak{g}^*, \qquad m_R = R^*_g m \in \mathfrak{g}^*.$$

Between these four elements, we have the relations

$$\omega_R = Ad_g\ \omega_L, \quad m_R = Ad_g^* m_L, \quad m_L = A\,\omega_L.$$

The invariance of the energy with respect to left translations leads to the existence of a *momentum map* $\mu : TG \to \mathfrak{g}^*$ defined by

$$\mu((g,\dot g))(\xi) = \frac{\partial K}{\partial \dot g}\, Z_\xi = \langle \dot g, R_g\,\xi\,\rangle_g = (m, R_g\,\xi\,) = \left(R_g^*\,m, \xi\,\right) = m_R(\xi),$$

where Z_ξ is the right-invariant vector field generated by $\xi \in \mathfrak{g}$. According to Noether's theorem, this map is constant along a geodesic, that is

$$\frac{dm_R}{dt} = 0.$$

As we did in the special case of the group $SO(n)$, using the relation $m_R = Ad_g^*\,m_L$ and computing the time derivative, we obtain the *Arnold-Euler equation*

$$\frac{dm_L}{dt} = ad^*_{\omega_L}\,m_L. \tag{7}$$

Using $\omega_L = A^{-1}\,m_L$ and the bilinear operator B defined by

$$\langle [a,b\,]\,,c\,\rangle = \langle B(c,a), b\,\rangle\,, \quad a,b,c \in \mathfrak{g},$$

equation (7) can be rewritten as an evolution equation on the Lie algebra

$$\frac{d\omega_L}{dt} = B(\omega_L, \omega_L)\,.$$

Well-posedness results for the generalized Benjamin-Ono equation with arbitrary large initial data

Luc Molinet

(joint work with Francis Ribaud)

Abstract. We prove new local well-posedness results for the generalized Benjamin-Ono equation (GBO) $\partial_t u + \mathcal{H}\partial_x^2 u + u^k \partial_x u = 0$, $k \geq 2$. By combining a gauge transformation with dispersive estimates we establish the local well-posedness of (GBO) in $H^s(\mathbb{R})$ for $s \geq 1/2$ if $k \geq 5$, $s > 1/2$ if $k = 2, 4$ and $s \geq 3/4$ if $k = 3$. Moreover we prove that in all these cases the flow map is locally Lipschitz on $H^s(\mathbb{R})$.

Presention of the problem

This work is devoted to the study of the local well-posedness problem for the generalized Benjamin-Ono equation

$$\text{(GBO)} \qquad \begin{cases} \partial_t u + \mathcal{H}\partial_x^2 u \pm u^k \partial_x u = 0\ ,\ (t,x) \in \mathbb{R}\times\mathbb{R}\ , \\ u(0,x) = u_0(x)\ , \end{cases}$$

where $\mathcal{H}$ is the Hilbert transform defined by

$$\mathcal{H}(f)(x) = -i \int_{-\infty}^{+\infty} e^{ix\xi} \operatorname{sgn}(\xi) \hat{f}(\xi) \, d\xi \ ,$$

and $k \geq 2$ is an integer.

The Benjamin-Ono equation (k=1) arises as a model for long internal gravity waves in deep stratified fluids, see [2], and have been studied in a large amount of works. When $k \geq 2$, (GBO) is an infinite dimensional Hamiltonian system (for $k = 1$ it is even formally completely integrable) and possesses the following invariant quantities :

$$I(u) = \int_{-\infty}^{+\infty} u(t,x) \, dx \,, \quad M(u) = \int_{-\infty}^{+\infty} u^2(t,x) \, dx \,,$$

and

$$E(u) = \int_{-\infty}^{+\infty} \Big(\frac{1}{2} |D_x^{1/2} u(t,x)|^2 \mp \frac{1}{(k+1)(k+2)} u(t,x)^{k+2} \Big) \, dx \ \ (energy) \ \ .$$

One of the challenging problem about this family of equations is probably to establish a well-posedness result in the energy space $H^{1/2}(\mathbb{R})$.

Recall that the Cauchy problem for the Benjamin-Ono equation (k=1) has been shown to be locally well-posed in $H^s(\mathbb{R})$ for $s \geq 3$ in [19], $s > 3/2$ in [8], [1] and later on for $s \geq 3/2$ in [18]. These results have been extended to global ones by using conservation laws. Recently, by establishing dispersive estimates for the non homogeneous linearized equation, H. Koch and N. Tzvetkov [14] and then C. Kenig and K. Koenig [9] have improved these local well-posedness results in $H^s(\mathbb{R})$ to respectively $s > 5/4$ and $s > 9/8$. More recently, using a gauge transformation and standard dispersive estimates, T. Tao [20] has gone down to $H^1(\mathbb{R})$. It is worth noticing that all these results have been obtained by compactness methods. Moreover, it has been proved in [17] that, for all $s \in \mathbb{R}$, the flow-map $u_0 \mapsto u(t)$ is not of class C^2 at the origin in $H^s(\mathbb{R})$ which implies that it is not possible to obtain well-posedness results in $H^s(\mathbb{R})$ for the Benjamin-Ono equation by contraction methods. In this direction, H. Koch and N. Tzvetkov [15] have recently proved that this flow-map is even not locally uniformly continuous in $H^s(\mathbb{R})$.

Now, concerning the case $k \geq 2$, the local well-posedness of (GBO) is also known in $H^s(\mathbb{R})$ for $s > 3/2$, see [8], [1], [13]. Recently, using the approach developped in [14], C. Kenig and K. Koenig [9] have shown the local well-posedness of (GBO) in $H^1(\mathbb{R})$ for $k = 2$ (note that only the continuity of the flow-map is established). Unfortunately, this approach does not seem to permit to go below $H^1(\mathbb{R})$ due to the weakness of the smoothing effect of the associated free evolution. On the other hand, in the context of small initial data, C. Kenig, G. Ponce and L. Vega [13] have proved local well-posedness results for (GBO) in $H^s(\mathbb{R})$ by a Picard iterative sheme on the integral equation. This denotes of course a strong difference with the case $k = 1$ and implies the real analiticity of the flow-map in a neighborhood of the origin. Very recently, these results have been improved by the authors in [16] where it is proven that, for small initial data, (GBO) is locally well-posed in $H^s(I\!R)$ as soon as

$$\begin{cases} s > 1/2 \ \text{ if } \ k = 2\,, \\ s > 1/3 \ \text{ if } \ k = 3\,, \\ s > s_k \ \text{ if } \ k \geq 4\,, \end{cases}$$

and globally well-posed as soon as

$$\begin{cases} s \geq 1/2 \ \text{ if } \ k = 3\,, \\ s > s_k \ \text{ if } \ k \geq 4\,, \end{cases}$$

where $s_k = 1/2 - 1/k$ is the critical scaling Sobolev index. Moreover these results are almost sharp for $k \neq 3$: in [16] we prove that for $k = 2$ and $k \geq 4$, the flow map is not respectively of class C^3 below $H^{1/2}(I\!R)$ and of class C^{k+1} below $H^{s_k}(I\!R)$ at the origin. Note that the above results imply the global well-posedness in the energy space $H^{1/2}(I\!R)$ for small initial data when $k \geq 3$.
It is worth recalling that the dispersion of the free evolution $V(t)$ of (GBO) is just sufficient to recover the lost derivative in the nonlinear term but does not seem to permit to get a contraction factor for T small when estimating the operator

$$\mathcal{G} \ : \ u \mapsto V(t)u_0 - \int_0^t V(t-t')\partial_x(u^{k+1}(t'))\,dt'$$

in the appropriate resolution space. This explains the smallness assumption on the initial data in [13] and [16]. In this sense (GBO) seems to be a limit case for the balance between dispersion and derivative nonlinearity of order one.

In this work we improve the existing local well-posedness results in the case of arbitrary large initial data. As mentioned above, the aim is to reach the energy space $H^{1/2}(I\!R)$. This will be achieved for $k \geq 5$. More precisely, we prove that (GBO) is locally well-posed in $H^s(I\!R)$ as soon as

$$\begin{cases} s > 1/2 \ \text{ if } \ k = 2, 4\,, \\ s \geq 3/4 \ \text{ if } \ k = 3\,, \\ s \geq 1/2 \ \text{ if } \ k \geq 5\,. \end{cases}$$

Moreover we show that in all these cases, in a sharp contrast with the case $k = 1$, the flow-map is locally Lipschitz. This has to be anderstood as a stability result for (GBO) when $k \geq 2$.
To establish our results, inspired by the recent work [20], we introduce a gauge transform w of u a smooth solution of (GBO) and derive a dispersive equation satisfied by w. Using dispersive estimates we will be able to get a positive power of T in front of the Duhamel part when estimating w in our resolution space. Next, rewriting (GBO) with the help of w, we obtain the desired estimate on the solution u. Our results follow then by regularizing the initial data and passing to the limit on smooth solutions to (GBO).

1. MAIN RESULTS

Let us state our main result.

Theorem 1. *For any* $u_0 \in H^s(\mathbb{R})$ *with*

$$\begin{cases} s > 1/2 \ \ \textit{if} \ \ k = 2, 4\,, \\ s \geq 3/4 \ \ \textit{if} \ \ k = 3\,, \\ s \geq 1/2 \ \ \textit{if} \ \ k \geq 5\,, \end{cases}$$

there exists $T = T^s_k(\|u_0\|_{H^s}) > 0$ *with* $T^s_k(\alpha) \nearrow \infty$ *as* $\alpha \searrow 0$*, and a unique solution* u *to (GBO) satisfying*

$$u \in C([0,T]; H^s(\mathbb{R})) \cap X^s_T\ .$$

Moreover, for the class of s *defined above, the flow-map is Lipschitz on every bounded set of* $H^s(\mathbb{R})$.

Remark 1.1. Actually we prove that $T = T_k(\|u_0\|_{H^{s(k)}})$ where

$$s(k) = \begin{cases} 1/2+ \ \ \text{if} \ \ k = 2, 4\,, \\ 3/4 \ \ \text{if} \ \ k = 3\,, \\ 1/2 \ \ \text{if} \ \ k \geq 5\,. \end{cases}$$

Remark 1.2. Theorem 1 yields a global existence result for the solutions to the following (GBO) equation

$$\partial_t u + \mathcal{H}\partial_x^2 u - u^k \partial_x u = 0\ , \tag{1}$$

where k is an odd integer greater or equal to 5. Indeed, the energy is then given by

$$E(u) = \frac{1}{2}\int_{-\infty}^{+\infty} |D_x^{1/2} u(t,x)|^2\, dx + \frac{1}{(k+1)(k+2)}\int_{-\infty}^{+\infty} u(t,x)^{k+2}\, dx$$

and thus, for $k \geq 5$ odd, Theorem 1.1 combining with the conservation of $E(u)$ leads to the global well-posedness of (1) in $H^s(\mathbb{R})$, $s \geq 1/2$. Note that for $k \geq 2$ with the reverse sign in front of the nonlinear term, numerical simulations suggest that blow-up in finite time can occur for large initial data [4].

Remark 1.3. Following the approach developped in this work with some additional technical points, one can certainly improve the results of Theorem 1 to $s > s_k = 1/2 - 1/k$ at least for $k \geq 5$ large enough. This would be in some sense optimal since it is shown in [3] that the flow-map cannot be uniformly continuous in $H^s(\mathbb{R})$ for $s = s_k$.

References

[1] L. Abdelouhab, J. Bona, M. Felland, and J.C. Saut, Nonlocal models for nonlinear, dispersive waves, *Phys. D* **40** *(1989), 360–392.*

[2] T.B. Benjamin, Internal waves of permanent form in fluids of great depth, *J. Fluid Mech.* **29** *(1967), 559–592.*

[3] H.A. Biagioni and F. Linares, Ill-posedness for the derivative Schrödinger and generalized Benjamin-Ono equations, *Trans. Amer. Math. Soc.* **353** *(2001), 3649-3659.*

[4] J. Bona and H. Kalisch, Singularity formation in the generalized Benjamin-Ono equation, *To appear in Disc. Cont. Dyn. Systems-Ser. B.*

[5] M Christ and A. Kiselev, Maximal functions associated to filtrations, *J. Funct. Anal.* **179** *(2001), 406-425.*

[6] J. Ginibre and G. Velo, Smoothing properties and existence of solutions for the generalized Benjamin-Ono equation, *J. Differential Equations* **93** *(1991), 150–232.*

[7] N. HAYASHI AND T. OZAWA, Remarks on Schrödinger equations in one space dimension, *Diff. Int. Equ.* **7** *(2) (1994), 453-461.*

[8] J.R. IORIO, On the Cauchy problem for the Benjamin-Ono equation, *Comm. Partial Differential Equations* **11** *(10) (1986), 1031-1081.*

[9] C.E. KENIG AND K. KOENIG, On the local well-posedness of the Benjamin-Ono and modified Benjamin-Ono equations, *Preprint (2003).*

[10] C.E. KENIG, G. PONCE AND L. VEGA, Well-posedness of the initial value problem for the Korteweg-de Vries equation, *J. Amer. Math. Soc.* **4** *(2) (1991), 323-347.*

[11] C.E. KENIG, G. PONCE AND L. VEGA, Oscillatory integrals and regularity of dispersives equations, *Indiana Univ. Mat. J.* **40** *(1) (1991), 32-69.*

[12] C.E. KENIG, G. PONCE AND , L. VEGA, Well-posedness and scattering results for the generalized Korteweg-de Vries equation via contraction principle. Comm. and Pure Pure and Appl. Math. **1993**, *46*, 527-620.

[13] C.E. KENIG, G. PONCE AND L. VEGA, On the Generalized Benjamin-Ono equations, *Trans. Amer. Math. Soc.* **342** *(1994), 155-172.*

[14] H. KOCH AND N. TZVETKOV, On the local well-posedness of the Benamin-Ono equation in $H^s(\mathbb{R})$, *IMRN* **26** *(2003), 1449-1464.*

[15] H. KOCH AND N. TZVETKOV, Nonlinear wave interactions for the Benamin-Ono equation, *Preprint (2002).*

[16] L. MOLINET AND F. RIBAUD, Well-posedness results for the Generalized Benjamin-Ono equation with small initial data, *J. Math. Pures Appl.* **83** *(2004), 277-311.*

[17] L. MOLINET, J.C. SAUT AND N. TZVETKOV, Ill-posedness issues for the Benjamin-Ono and related equations, *SIAM J. Math. Anal.* **33** *(4) (2001), 982-988.*

[18] G. PONCE, On the global well-posedness of the Benjamin-Ono equation, *Differential Integral Equations* **4** *(3) (1991), 527-542.*

[19] J.-C. SAUT, Sur quelques generalisation de l'équation de Korteweg-de Vries, *J. Math. Pures Appl.* **58** *(1979), 21-61.*

[20] T. TAO, Global well-posedness of the Benjamin-Ono equation in $H^1(\mathbb{R})$, *Preprint (2003).*

About the "loss of regularity" for hyperbolic problems

MICHAEL REISSIG

Mathematics Subject Classification 2000: 35L15, 35L80, 35S30, 35B45.

In this lecture we will study hyperbolic problems with quite different goals from the first point of view. It turns out that these problems have common features which are described in the following table:

strictly hyperbolic theory with non-Lipschitz continuous coefficients	$\longleftrightarrow$	$L_p - L_q$ decay estimates for wave equations with bounded time-dependent coefficients
$\updownarrow$		$\updownarrow$
weakly hyperbolic theory	$\longleftrightarrow$	$L_p - L_q$ decay estimates for wave equations with increasing time-dependent coefficients

The question for $L_p - L_q$ decay estimates is related with the question for global in time small data solutions for the Cauchy problem for nonlinear wave equations like

$$u_{tt} - \triangle u = f(u_t, \nabla u, \nabla u_t, \nabla^2 u), \;\; u(0,x) = \varepsilon\phi(x), \; u_t(0,x) = \varepsilon\psi(x).$$

The goal is to prove under suitable assumptions that for all $\varepsilon \in (0, \varepsilon_0(\phi,\psi)]$ there exists a global (in time) small data solution. One of the key tools is the so-called Strichartz' decay estimate

$$E(u)(t)|_{L_q} \le C(1+t)^{-\frac{n-1}{2}(\frac{1}{p}-\frac{1}{q})} E(u)(0)|_{W_p^{N_p}}$$

on the conjugate line $2 \le q \le \infty, \;\; 1/p + 1/q = 1$ for solutions of the Cauchy problem for classical wave equations, where $N_p > n\left(\frac{1}{p} - \frac{1}{q}\right)$. Generalizing such type of estimates (with $-\frac{n}{2}$ instead of $-\frac{n-1}{2}$ in the decay rate) to Klein-Gordon equations or damped wave equations (with an additional term $-\frac{1}{2}$ in the decay rate of the latter case coming from the dissipation itself) one can show the global existence of small data solutions for

$$u_{tt} - \triangle u + m^2 u = f(u_t, \nabla u, \nabla u_t, \nabla^2 u), \;\; u(0,x) = \varepsilon\phi(x), \; u_t(0,x) = \varepsilon\psi(x), \; m > 0;$$
$$u_{tt} - \triangle u + u_t = f(u_t, \nabla u, \nabla u_t, \nabla^2 u), \;\; u(0,x) = \varepsilon\phi(x), \; u_t(0,x) = \varepsilon\psi(x).$$

In general one can find such $L_p - L_q$ decay estimates for solutions of partial differential equations (or systems) with constant coefficients.

For this reason the author asked if one can generalize such estimates to solutions for wave equations with time dependent coefficients like

$$u_{tt} - a(t) \triangle u + m(t)u + b(t)u_t = 0.$$

Here one can use the WKB-method and construct explicit representations of solutions. The dependence of coefficients on spatial variables brings essential difficulties, e.g. the global existence (in time) of phase functions in the FIO-representations.
The above model is to general, one should assume some more structure of the coefficients.

1.case: wave equations with weak dissipation
The model under consideration is

$$u_{tt} - \triangle u + b(t)u_t = 0.$$

Under the main assumptions $b' < 0, \lim_{t\to\infty} b(t) = 0$ we have a complete picture from wave to damped wave equations which reads in the following form (we only describe the decay rates):

- $b \in L_1(\mathbb{R}^+)$: scattering results with the free wave equation,
- $b(t) \sim (t \log t)^{-1}$ for large t: hyperbolic decay rate $-\frac{n-1}{2}(\frac{1}{p} - \frac{1}{q})$ and a term coming from the dissipation itself, such dissipations are *not effective*,
- $b(t) = \mu(1+t)^{-1}$, $\mu > 0$: critical case $\mu = 2$ gives the best $L_2 - L_2$ decay rate $(1+t)^{-1}$, here the decay rate changes from the hyperbolic one (small μ) to the parabolic one $-\frac{n}{2}(\frac{1}{p} - \frac{1}{q})$ (large μ),
- $b(t) \sim t^{-1} \log t$ for large t: parabolic decay rate $-\frac{n}{2}(\frac{1}{p} - \frac{1}{q})$ and a term coming from the dissipation itself, such dissipations are *effective*.

Question: What are the relations in the 3-d case between the influence of dissipation (effective or non effective) on $L_p - L_q$ decay estimates and assumptions to the asymptotic behavior of the nonlinearity $f = f(u_t, \nabla u, \nabla u_t, \nabla^2 u)$ in 0?

2.case: general model There exist several difficulties:
- There exists an interplay between oscillating behavior and increasing behavior of coefficients.
- An interplay between $a = a(t)$ and $m(t)$ decides if the *mass term is effective*. In such a case it should be included into the phase function. This gives difficulties to develop a stationary phase method.
- An interplay between $a = a(t)$ and $b(t)$ decides if the *dissipation term is effective*.

One can prove the following results:
1. *Let us consider the model problem*

$$u_{tt} - \exp(2t^\alpha)(2 + \sin t)^2 \triangle u = 0, \;\; u(0, x) = \phi(x), \;\; u_t(0, x) = \psi(x).$$

Then the following Strichartz' type estimate holds with some regularity $W_p^{N_p}$*:*

$$E(u)(t)|_{L_q} \leq C\Big(1 + \int_0^t \exp \tau^\alpha d\tau\Big)^{s_0 - \frac{n-1}{2}(\frac{1}{p} - \frac{1}{q})} E(u)(0)|_{W_p^{N_p}},$$

where s_0=ε *sufficiently small,* s_0 *is a positive constant,* $s_0 = \infty$ *(no* $L_p - L_q$ *decay estimate) if* $\alpha > \frac{1}{2}$*,* $\alpha = \frac{1}{2}$*,* $\alpha < \frac{1}{2}$ *respectively.*
2. A mass term can have an improving influence (less increasing behavior is necessary) as the next result shows.

Let us consider the model problem

$$u_{tt} - (1 + t)^2(2 + \sin t)^2(\triangle u - u) = 0, \;\; u(0, x) = \phi(x), \;\; u_t(0, x) = \psi(x).$$

Then the following Strichartz' type estimate holds with some regularity $W_p^{N_p}$*:*

$$E(u)(t)|_{L_q} \leq C(1 + t^2)^{s_0 - \frac{n}{2}(\frac{1}{p} - \frac{1}{q})} E(u)(0)|_{W_p^{N_p}},$$

where s_0 *is a positive constant.*

Are there some relations to other hyperbolic problems? **Yes!** There exist relations to *weakly hyperbolic problems* or to *strictly hyperbolic problems with non-Lipschitz coefficients*. Let us demonstrate this connection by the following result:

Let us consider the strictly hyperbolic Cauchy problem

$$u_{tt} - \sum_{k,l=1}^{n} a_{kl}(t, x) u_{x_k x_l} = f(t, x), \;\; u(0, x) = \phi(x), u_t(0, x) = \psi(x),$$

in the strip $\mathbb{R}^n \times [0,T]$. *The non-Lipschitz behavior of coefficients is described by the following conditions for all multi-indices* β *and all* $p \in \mathbb{N}$*:*

$$|D_t^p D_x^\beta a_{kl}(t,x)| \le C_{p\beta}\Big(\frac{1}{t}\Big(\log\frac{1}{t}\Big)\Big)^\gamma.$$

Then for large s *the energy inequality*

$$E(u)(t)|_{H^{s-s_0}} \le C_s E(u)(0)|_{H^s}$$

holds with $s_0 = 0$, $s_0{=}\varepsilon$ *sufficiently small,* s_0 *is a positive constant,* $s_0 = \infty$ *(no* C^∞ *well-posedness) if* $\gamma = 0$, $\gamma \in (0,1)$, $\gamma = 1$, $\gamma > 1$ *respectively. Moreover, there exists a parametrix in the cases* $\gamma \in [0,1]$.

Remarks:

- There exists in all those hyperbolic problems a connection between the oscillating behavior of coefficients and the "loss of regularity" (for $L_p - L_q$ decay estimates this means how the decay rate differs from the classical decay rates for the wave, Klein-Gordon, or damped wave operator). An optimal classification of oscillations can be given for all problems.
- The construction of parametrix in form of Fourierintegral operators is closely related to the construction of representation of solutions by Fourier multipliers to derive $L_p - L_q$ decay estimates.
- A careful division of the phase space into zones, a symbolic calculus for non-standard symbol classes, hierarchies of symbols, ellipticity, the construction of phase functions and amplitudes in FIO-representations, and a suitable perfect diagonalization procedure, form the main tools for the construction of parametrix.
- Counterexamples ($s_0 = \infty$ in the above results) are proved by the application of Floquet's theory.

Genesis of Solitons Arising from Individual Flows of the Camassa-Holm Hierarchy

Enrique Loubet

The present work offers a detailed account of the large time development of the velocity profile v run by a single "individual" Hamiltonian flow of the Camassa-Holm (CH) hierarchy, the Hamiltonian employed being the invariant $\mathrm{H} = 1/\lambda$, where λ is *any* of the bound states of the associated spectral problem: $(\frac{1}{4} - D^2)(f) = \lambda m f$, with "mass" potential $m \equiv v - v''$. The flow may be expressed as in $\partial m/\partial t = [mD + Dm](f^2) = 1/(2\lambda)D(1-D^2)(f^2)$, or more simply, as $\partial v/\partial t = 1/(2\lambda)D(f^2)$. Unlike the formation of the soliton train that is produced by Korteweg-de Vries (KdV) $\partial V/\partial t = 3V\partial V/\partial X - \frac{1}{2}\partial^3 V/\partial X^3$, which accounts, except for the reflectionless potential V, only for the part of the total energy ascribed to the bound states of the associated spectral problem $(-D^2+V)(F) = \lambda F$, the deficiency being carried by the evanescent radiation corresponding to the continuous spectrum; for *summable* m, CH has *only* bound states $\lambda_n\ ,\ n \in \mathbb{Z} - \{0\}$, each of which characterizes the speed=amplitude of the associated individual soliton $\mathcal{S}_n(t,x) \equiv 1/(2\lambda_n)e^{-|x-t/(2\lambda_n)|}$. They embody respectively an energy $\frac{1}{2}\int[(\mathcal{S}_n')^2 + \mathcal{S}_n^2] = 1/(4\lambda_n^2)$, and all these individual pieces add up to the whole: $\mathrm{H}_{\mathrm{CH}} \equiv \frac{1}{2}\int m\mathrm{G}[m] = \sum 1/(4\lambda_n^2)$ where $\mathrm{G} \equiv (1-D^2)^{-1}$, so here nothing is lost. And indeed, the present investigation confirms this:

Let m be summable and odd, having the signature of x, and consider the individual flow based upon $\mathrm{H} = 1/\lambda$ with $\lambda > 0$. With the help of a private "Lagrangian" scale determined by $\overline{x}^\bullet = -f^2(t,\overline{x})$ and $\overline{x}(0,x) = x$; the updated velocity profile $(e^{t\mathbb{X}_\mathrm{H}}v(0,\cdot))(\overline{x}) \equiv v(t,\overline{x}(t,x))$ is found to shape itself like the soliton

$$\mathcal{S}_\lambda(t,x) = 1/(2\lambda)e^{-|x-t/(2\lambda)|}$$

escaping to $-\infty$ as $t \uparrow +\infty$, leaving behind a "residual" $v(+\infty,\overline{x}(+\infty,x))$ having the same spectrum as the one attatched to the initial $v(0,x)$ except that λ is excised. Doubtless, the map induced by the large time asymptotics $v(0,x) \mapsto v(+\infty,\overline{x}(+\infty,x))$, is some counterpart of the standard Darboux transformation for removing/adding the bottom bound state for KdV, with the difference now that you need not proceed in such orderly fashion. I did not succeed in casting such correspondence in the form of an "addition" as in [1], but it should be closely connected to that circle of ideas.

References

[1] McKean, *Fredholm determinants and the Camassa-Holm hierarchy,* Comm. Pure Appl. Math., **56**, (2003), no.5, 638–680.

Reporter: Carlheinz Kneisel

Participants

Dr. David Ambrose
david.ambrose@cims.nyu.edu
Courant Institute of
Mathematical Sciences
New York University
251, Mercer Street
New York, NY 10012-1110 – USA

Prof. Dr. Christer Bennewitz
Christer.Bennewitz@math.lu.se
Dept. of Mathematics
University of Lund
Box 118
S-221 00 Lund

Prof. Dr. Adrian Constantin
Adrian.Constantin@math.lu.se
Dept. of Mathematics
University of Lund
Box 118
S-221 00 Lund

Prof. Dr. Joachim Escher
escher@ifam.uni-hannover.de
Institut für Angewandte Mathematik
Universität Hannover
Welfengarten 1
D–30167 Hannover

Prof. Dr. Ludwig Edward Fraenkel
L.E.Fraenkel@bath.ac.uk
lef@maths.bath.ac.uk
School of Mathematical Sciences
University of Bath
Claverton Down
GB-Bath Somerset BA2 7AY

Dr. Mark D. Groves
M.D.Groves@lboro.ac.uk
Dept. of Mathematical Sciences
Loughborough University
Loughborough
GB-Leicestershire LE11 3TU

Dr. Mariana Haragus
haragus@math.univ-fcomte.fr
Department de Mathematiques
UFR de Sciences et Techniques
Universite de Franche-Comte
16, route de Gray
F-25000 Besancon

Prof. Dr. Helge Holden
holden@math.ntnu.no
Dept. of Mathematical Sciences
Norwegian University of Science
and Technology
A. Getz vei 1
N-7491 Trondheim

Prof. Dr. Henrik Kalisch
Henrik.Kalisch@math.lu.se
Department of Mathematics
Lund Institute of Technology
P.O. Box 118
S-22100 Lund

Prof. Dr. Thomas Kappeler
tk@math.unizh.ch
Institut für Mathematik
Universität Zürich
Winterthurerstr. 190
CH-8057 Zürich

Dr. Kenneth Karlsen
kennethk@mi.uib.no
Matematisk Institutt
Johannes Brunsgate 12
5008 Bergen – Norway

Carlheinz Kneisel
kneisel@ifam.uni-hannover.de
Institut für Angewandte Mathematik
Universität Hannover
Welfengarten 1
D–30167 Hannover

Prof. Dr. Boris Kolev
kolev@cmi.univ-mrs.fr
CMI Marseille
Place des Aires
F-84160 Marseille

Prof. Dr. Evgeni Korotyaev
evgeni@math.uni-potsdam.de
ek@mathematik.hu-berlin.de
Institut für Reine Mathematik
Fachbereich Mathematik
Humboldt-Universität Berlin
D–10099 Berlin

Prof. Dr. Olaf Lechtenfeld
lechtenf@itp.uni-hannover.de
Institute for Theoretical Physics
University of Hannover
Appelstr. 2
D–30167 Hannover

Prof. Dr. Jonatan Lenells
jonatan@maths.lth.se
Department of Mathematics
Lund Institute of Technology
P.O. Box 118
S-22100 Lund

Dr. Enrique Loubet
eloubet@math.unizh.ch
Institut für Mathematik
Universität Zürich
Winterthurerstr. 190
CH-8057 Zürich

Prof. Dr. Nader Masmoudi
masmoudi@cims.nyu.edu
Courant Institute of Math. Sciences
New York University
251, Mercer Street
New York NY 10012-1185 – USA

Prof. Dr. Luc Molinet
molinet@math.univ-paris13.fr
Laboratoire Analyse, Geometrie et
Applications, UMR CNRS 7539
Institut Galilee, Univ. Paris 13
Avenue JB Clement
F-93430 Villetaneuse

Prof. Dr. John Norbury
john.norbury@lincoln.ox.ac.uk
Mathematical Institute
Oxford University
24 - 29, St. Giles
GB-Oxford OX1 3LB

Prof. Dr. Hisashi Okamoto
okamoto@kurims.kyoto-u.ac.jp
Research Institute for
Mathematical Sciences
Kyoto University
Kitashirakawa, Sakyo-ku
Kyoto 606-8502 – Japan

Prof. Dr. Michael Reissig
reissig@math.tu-freiberg.de
Fakultät für Mathematik und
Informatik; Technische Universität
Bergakademie Freiberg
Agricolastr. 1
D–09599 Freiberg

Prof. Dr. Nils Henrik Risebro
nilshr@math.uio.no
16, rue Mathias Tresch
L-2626 Luxembourg

Prof. Dr. Walter A. Strauss
wstrauss@math.brown.edu
Dept. of Mathematics
Brown University
Box 1917
Providence, RI 02912 – USA

Dr. Petar J. Topalov
topalov@math.bas.bg
Institut für Mathematik
Universität Zürich
Winterthurerstr. 190
CH-8057 Zürich

Prof. Dr. Erik Wahlen
ewahlen@df.lth.se
Dept. of Mathematics
University of Lund
Box 118
S-221 00 Lund

Mathematisches Forschungsinstitut Oberwolfach

Report No. 6/2004

Finite and Infinite Dimensional Complex Geometry and Representation Theory

Organised by
Alan Huckleberry (Bochum)
Karl-Hermann Neeb (Darmstadt)
Joseph A. Wolf (Berkeley)

February 1st – February 7th, 2004

Introduction by the Organisers

As the theme of the conference indicates, one of the organizers' main goals was to put together a group of participants with a wide range of interests in and around the complex geometric side of the representation theory of Lie groups and algebras. It is their belief that a hybrid approach to representation theory, in particular interaction between complex geometers and harmonic analysts standing on a strong foundation of finite- and infinite-dimensional Lie theory, will open up new avenues of thought and lead to progress in a number of areas.

Since the previous Oberwolfach conference (Represention Theory and Complex Analysis, April 2000), there has been quite a positive development toward these goals. A number of breakthroughs were achieved, and of course these were reported at this year's conference. More than half of the 47 participants (from 15 countries) are now working in some middle ground between complex geometry and representation theory. Furthermore, it was clear from the discussions both after the talks and in the evenings that most participants now understand each other's language well enough to discuss high level research projects.

A basic new component, infinite-dimensional complex geometry and related representation theory, was added this year. This quickly developing subject is already attracting wide attention. A goal for the future is to better integrate this with the more classical finite-dimensional theory.

One consequence of the broad range of backgrounds of the participants is that, without prodding by the organizers, virtually all speakers gave quite comprehensive introductions to their subjects before going into their most recent results. This was greatly appreciated by all!

Instead of attempting to summarize these talks we will let the following abstracts speak for themselves.

Workshop on Finite and Infinite Dimensional Complex Geometry and Representation Theory

Table of Contents

Abstracts

Visible actions on complex manifolds and multiplicity-one theorems

Toshiyuki Kobayashi

Multiplicity-free representations appear in various contexts such as Fourier transforms, Taylor series expansions, the Peter-Weyl theorem, branching laws for $GL_n \downarrow GL_{n+1}$, Clebsch-Gordan formula, Pieri's law, $GL_m - GL_n$ duality, the Plancherel formula for Riemannian symmetric spaces G/K, etc.

The aim of this talk is to report a simple principle based on complex geometry that explains the multiplicity-free property of various representations as above and more.

Suppose $\mathcal{V} \to D$ is an H-equivariant holomorphic vector bundle. Then, a representation of the group H is naturally defined on the Fréchet space $\mathcal{O}(D, \mathcal{V})$ of holomorphic sections. One asks:

"When does $\mathcal{O}(D, \mathcal{V})$ become multiplicity-free?"

We present a sufficient condition which comprises of a 'balance' of the base space D and fibers $\mathcal{V}_x$. To be more precise, let $P \to D$ be an H-equivariant principle K-bundle, $\mu : K \to GL_{\mathbb{C}}(V)$ a finite dimensional unitary representation, and $\mathcal{V} \simeq P \times_K V$. Suppose we are given automorphisms of Lie groups H and K, and a diffeomorphism of P, for which we use the same letter σ, satisfying the following two conditions:

$$\sigma(hpk) = \sigma(h)\sigma(p)\sigma(k)(h \in H; p \in P; k \in K),$$

The induced action of σ on $D(\simeq P/K)$ is anti-holomorphic.

For a subset B in P^σ, we define the following σ-stable subgroup

$$M := \{k \in K : bk \in Hb for any b \in B\}.$$

Theorem. *Assume that there exist σ and a subset B of P^σ satisfying the following three conditions:*

a) *HBK contains an interior point of P.*
b) *The restriction $\mu|_M$ decomposes as a multiplicity-free sum of irreducible representations of M.*

We shall write the decomposition as $\mu|_M \simeq \bigoplus_i \nu^{(i)}$.

c1) *$\mu \circ \sigma$ is isomorphic to μ^* (the contragredient representation of μ) as representations of K.*
c2) *$\nu^{(i)} \circ \sigma$ is isomorphic to $(\nu^{(i)})^*$ as representations of M for every i.*

Then, for any (abstract) unitary representation π of H which can be realized as a subrepresentation of $\mathcal{O}(D, \mathcal{V})$, π is multiplicity-free as an H-module.

Loosely speaking, our theorem asserts that the multiplicity-free property propagates from the smaller group M acting on fibers (see (b)) to the larger group H acting on holomorphic sections under a suitable condition (see (a)) on the H-action on the complex manifold D.

In light of the geometric condition (a) given in Theorem, we introduce the following notion:

Definition. The action of a Lie group H on a connected complex manifold D is **visible** if there exists a totally real submanifold N which meets generic H-orbit on D and satisfies

$$J(T_x N) \subset T_x(H \cdot x) \quad \text{for all } x \in N.$$

Example. 1) The natural action of $\mathbb{T}^n$ on the projective space $\mathbb{P}^{n-1}\mathbb{C}$ is visible.

2) The natural action of the direct product group $U(n_1) \times U(n_2) \times U(n_3)$ on the Grassmann variety $Gr_p(\mathbb{C}^n)(n = n_1 + n_2 + n_3 = p + q)$ is visible if $\min(n_1 + 1, n_2 + 1, n_3 + 1, p, q) \leq 2$.

3) Let G be a compact Lie group, and $G_{\mathbb{C}}$ its complexification. Then the action of $G \times G$ on $G_{\mathbb{C}}$ is visible.

4) Let $\mathcal{N}$ be a nilpotent orbit of $GL(n, \mathbb{C})$ corresponding to a partition $2^p 1^{n-2p}$. Then the action of $U(n)$ on $\mathcal{N}$ is visible for any p.

5) Let G/K be a Riemannian symmetric space of the non-compact type, and Ω its crown in $G_{\mathbb{C}}/K_{\mathbb{C}}$. Then the action of G on Ω is visible.

The above examples lead us to various kinds of multiplicity free representations. For example, (1) gives rise to the multiplicity-free property of the restriction $GL_n \downarrow GL_{n-1}$ as well as the Pieri rule for tensor product representations; (2) does to the list of all multiplicity-free tensor product representations of GL_n, which Stembridge found by a completely different method based on combinatorial argument; (3) does to the multiplicity-free property of the Peter-Weyl theorem of $L^2(G)$; (4) does to spherical nilpotent orbits whose complete list was recently given by Panyushev.

References

[1] T. Kobayashi, Multiplicity-free branching laws for unitary highest weight modules, Proceedings of the Symposium on Representation Theory held at Saga, Kyushu 1997 (ed. K. Mimachi) (1997), 9–17.

[2] T. Kobayashi, Geometry of multiplicity-free representations of $GL(n)$, visible actions on flag varieties, and triunity, preprint.

[3] T. Kobayashi, in preparation.

A new look at the Maslov index

Bent Ørsted

The Maslov index is an invariant that appears several places in mathematics; roughly speaking it encodes qualitive aspects of solutions to certain variational problems - this includes asymptotic solutions to partial differential equations and flows of Hamiltonian systems. It also appears in the study of Lagrangian subspaces of a fixed symplectic vector space, where it gives an integer invariant for each triple of such subspaces. In this lecture we give several new ways of looking at the Maslov index, generalizing to the setting of bounded symmetric domains and defining a Maslov index for transversal triples of points in the Shilov boundary. This is done by integrating the canonical Kähler form over geodesic triangles in the domain and taking a limit to the boundary. We also extend to the infinite-dimensional situation and define a Maslov map from transversal triples on an appropriate Shilov boundary to the first homotopy group of the stabilizer of a base point in the domain. A crucial identity is shown in the context of Jordan triple systems, which gives a good algebraic framework for the infinite-dimensional case of such generalized flag manifolds and their invariants. This represents joint work, partly in progress, with J.-L. Clerc, K-H. Neeb, and W. Bertram.

References

[1] Jean-Louis Clerc and Bent Ørsted: The Maslov index revisited, in *Transformation Groups* Vol. 6, pp. 303-320, 2001.

[2] Jean-Louis Clerc and Bent Ørsted: The Gromov norm of the Kaehler class and the Maslov index, in *Asian J. Math.* Vol. 7, No. 2, pp. 269-296, June 2003.

Analysis on the crown of a Riemannian symmetric space

Jacques Faraut

The crown of a Riemannian symmetric space $\mathcal{X} = G/K$ of non-compact type is a domain $\mathcal{D}$ in its complexification $\mathcal{X}_{\mathbb{C}} = G_{\mathbb{C}}/K_{\mathbb{C}}$, which has been intoduced by Akhiezer and Gindikin [1990]. It is also called the Akhiezer-Gindikin domain. It is interesting from various points of view: Riemannian geometry, complex geometry, analysis. From the analytic viewpoint it has the following remarkable property: All eigenfunctions of the invariant differential operators have a holomorphic extension to the crown $\mathcal{D}$, and the domain $\mathcal{D}$ is maximal for this property.

Consider the Cartan decomposition of $\mathfrak{g} = Lie(G)$, $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$, and let $\mathfrak{a} \subset \mathfrak{p}$ be a Cartan subspace. Define

$$\omega = \{H \in \mathfrak{a} \mid \forall\alpha \in \Delta(\mathfrak{g}, \mathfrak{a}),\ |\alpha(H)| < \frac{\pi}{2}\}.$$

The crown can be described as

$$\mathcal{D} = G \exp i\omega \cdot o \quad (o = eK_{\mathbb{C}}).$$

On the other hand consider an Iwasawa decomposition $\mathcal{X} = NA \cdot o$, and define

$$\Xi = \bigcap_{k \in K} kN_{\mathbb{C}}A_{\mathbb{C}} \cdot o.$$

Theorem

The crown $\mathcal{D}$ is equal to the connected component Ξ_0 of Ξ which contains $\mathcal{X}$.

The inclusion $\mathcal{D} \subset \Xi_0$ has been proved by Krötz and Stanton for classical groups G [2001], and by Huckleberry in general [2002]. The reverse inclusion $\Xi_0 \subset \mathcal{D}$ has been proved by Barchini [2003].

If the symmetric space $\mathcal{X}$ is Hermitian, then $\mathcal{D} = \mathcal{X} \times \bar{\mathcal{X}}$ ([Huckleberry,2002], [Burns-Halverscheid-Hind,2003]). Let $\mathrm{Aut}(\mathcal{D})$ be the group of all holomorphic automorphisms of the crown of $\mathcal{D}$. In all cases $G \subset \mathrm{Aut}(\mathcal{D})$. In case of equality one says that $\mathcal{D}$ is rigid. Then $\mathcal{D}$ is either rigid or Hermitian ([Burns-Halverscheid-Hind,2003]).

Corollary
Every eigenfunction of all invariant differential operators has a holomorphic extension to the crown $\mathcal{D}$, and $\mathcal{D}$ is maximal for this property.

Such a joint eigenfunction f has a Poisson integral representation over the maximal boundary B of $\mathcal{X}$:

$$f(x) = \int_B P_\lambda(x,b)dT(b) \quad (\lambda \in \mathfrak{a}_{\mathbb{C}}^*),$$

where T is an analytic functional on $B = K/M$ (M is the centralizer of A in K). The Poisson kernel $P_\lambda(x,b)$ is related to the Iwasawa decomposition as follows. If $x = n \exp H \cdot o$ ($n \in N$, $H \in \mathfrak{a}$) one writes $H = \mathcal{A}(x)$. Then

$$P_\lambda(x,b) = e^{\langle \rho-\lambda, \mathcal{A}(k^{-1}x)\rangle} \quad (b = kM).$$

By [Clerc,1988],

$$e^{\langle \lambda, \mathcal{A}(x)\rangle} = \prod_{j=1}^{\ell} \psi_j(x)^{\lambda_j},$$

where ψ_j is a holomorphic function on $\mathcal{X}_{\mathbb{C}}$ which does not vanish on $N_{\mathbb{C}}A_{\mathbb{C}} \cdot o$. Since the crown $\mathcal{D}$ is simply connected, if follows that the function $x \mapsto P_\lambda(x,b)$ has a holomorphic extension to $\mathcal{D}$.

On the other hand, for any point z_0 on the boundary of the crown $\mathcal{D}$, one can find $\lambda \in \mathfrak{a}_{\mathbb{C}}^*$ and $b \in B$ such that the function $z \mapsto P_\lambda(z,b)$ has a singularity at z_0.

References

Akhiezer, D.N.; Gindikin, S.G. (1990). On Stein extensions of real symmetric spaces, *Math. Ann.*, 286, 1–12.

Barchini, L. (2003). Stein exensions of real symmetric spaces and the geometry of the flag manifold, *Math. Ann.*, 326, 331–346.

Burns, D.; Halverscheid, S.; Hind, R. (2003). The geometry of Grauert tubes and complexification of symmetric spaces, *Duke Math. J.*, 118, 465–491.

Clerc, J.-L. (1988). Fonctions sphériques des espaces symétriques compacts, *Trans. A. M. S.*, 306, 421–431.

Huckleberry, A. (2002), On certain domains in cycle spaces of flag manifolds, *Math. Ann.*, 323, 797–810.

Krötz, B.; Stanton, R.J. (2001), Holomorphic extension of representations: (I) Automorphic functions, *Preprint.*

Hardy spaces for the most continuous spectrum
Bernhard Krötz

We report on joint work with Simon Gindikin and Gestur Ólafsson (cf. [GKÓ02]).

Holomorphic extensions and boundary value maps have been valuable tools to solve problems in representation theory and harmonic analysis on *real* symmetric spaces. Two of the best known constructions are Hardy spaces with their boundary value maps and Cauchy-Szegö-kernels, and Fock space constructions with their corresponding Segal-Barmann transform. It is in this flavour that we establish a correspondence between eigenfunctions on a Riemannian symmetric spaces $X = G/K$ and a non-compactly causal (NCC) symmetric spaces $Y = G/H$ in this talk. In particular we, via analytic continuation, relate a *spherical function* ϕ_λ on G/K to a *holomorphic* H-invariant distribution on G/H.

Let us explain our results in more detail. On the geometric level we construct a certain minimal G-invariant Stein domain $\Xi_H \subseteq X_{\mathbb{C}} = G_{\mathbb{C}}/K_{\mathbb{C}}$ with the following properties: The Riemannian symmetric space X is embedded into Ξ_H as a totally real submanifold and the affine non-compactly causal space Y is isomorphic to the distinguished (Shilov) boundary of Ξ_H.

The minimal tube Ξ_H is a subdomain of the complex crown $\Xi \subseteq X_{\mathbb{C}}$ of X – an object first introduced in [AG90] which became subject of intense study over the last few years. A consequence is that all $\mathbb{D}(X)$-eigenfunctions on X extend holomorphically to Ξ_H [KS01b]. Another key fact is that $\mathbb{D}(X) \simeq \mathbb{D}(Y)$. Thus by taking limits on the boundary Y we obtain a realization of the $\mathbb{D}(X)$-eigenfunctions on X as $\mathbb{D}(Y)$-eigenfunctions on Y. Conversely, eigenfunctions on Y which holomorphically extend to Ξ_H yield by restriction eigenfunctions on X.

It seems to us that the above mentioned transition between eigenfunctions on X and Y is most efficiently described using the techniques from representation theory. To fix the notation let $(\pi, \mathcal{H})$ denote an admissible Hilbert representation of G with finite length. We write $\mathcal{H}^K$ for the space of K-fixed vectors and $(\mathcal{H}^{-\infty})^H$ for the space of H-fixed distribution vectors of π. Using the method of analytic continuation of representations as developed in [KS01a] we establish a bijection

$$\mathcal{H}^K \to (\mathcal{H}^{-\infty})^H_{\mathrm{hol}}, \quad v_K \mapsto v_H$$

where $(\mathcal{H}^{-\infty})^H_{\mathrm{hol}} \subseteq (\mathcal{H}^{-\infty})^H$ denotes the subspace characterized through the property that associated matrix coefficients on Y extend holomorphically to Ξ_H.

We give an application of our theory towards the geometric realization of the most-continuous spectrum $L^2(Y)_{\mathrm{mc}}$ of $L^2(Y)$. First progress in this direction was achieved in [GKÓ01]. There, for the cases where $\Xi = \Xi_H$, we defined a Hardy space $\mathcal{H}^2(\Xi)$ on Ξ and showed that there is an isometric boundary value mapping realizing $\mathcal{H}^2(\Xi)$ as a multiplicity one subspace of $L^2(Y)_{\mathrm{mc}}$ of full spectrum. It was an open problem how to define Hardy spaces for general NCC symmetric spaces

Y and to determine the Plancherel measure explicitely. We solve this problem by giving a spectral definition of the Hardy space, i.e., we take the conjectured Plancherel measure and define a Hilbert space of holomorphic functions $\mathcal{H}^2(\Xi_H)$ on Ξ_H. The identification of $\mathcal{H}^2(\Xi_H)$ as a Hardy space then follows by establishing an isometric boundary value mapping $b : \mathcal{H}^2(\Xi_H) \hookrightarrow L^2(G/H)_{\mathrm{mc}}$. In particular we achieve a geometric realization of a multiplicity free subspace of $L^2(Y)_{\mathrm{mc}}$ in holomorphic functions.

References

[AG90] Akhiezer, D. N., and S. G. Gindikin, *On Stein extensions of real symmetric spaces*, Math. Ann. **286**, 1–12, 1990.

[GKÓ01] Gindikin, S., B. Krötz and G. Ólafsson, *Hardy spaces for non-compactly causal symmetric spaces and the most continuous spectrum*, Math. Ann. **327**, 25–66 (2003).

[GKÓ02] — , *Holomorphic H-spherical distribution vectors in principal series representations*, preprint.

[KS01a] Krötz, B., and R.J. Stanton, *Holomorphic extension of representations: (I) automorphic functions*, Ann. Math., to appear.

[KS01b] Krötz, B., and R.J. Stanton, *Holomorphic extensions of representations: (II) geometry and harmonic analysis*, preprint.

Structure of $gl(\infty)$

Ivan Dimitrov

Let U and V be two (infinite dimensional) complex vector spaces with a non–degenerate pairing $\langle \circ, \circ \rangle : U \times V \to \mathbb{C}$. Consider the Lie algebra $\mathfrak{g} := U \otimes V$. When both U and V are countable dimensional, $\mathfrak{g}$ is isomorphic to the Lie algebra $gl(\infty)$ of finitary infinite matrices, see [M]. A maximal locally solvable subalgebra of $\mathfrak{g}$ is called a *Borel subalgebra of* $\mathfrak{g}$. In this talk we describe the Borel subalgebras of $\mathfrak{g}$ and discuss their relation with maximal toral subalgebras of $\mathfrak{g}$.

In order to describe the Borel subalgebras of $\mathfrak{g}$ we need the notion of a generalized flag in U introduced in [DP]. A chain $\mathcal{F} = \{F'_\alpha, F''_\alpha\}_{\alpha \in A}$ of subspaces of U is a *generalized flag in* U if F'_α is the immediate predecessor of F''_α and $U \backslash \{0\} = \cup_\alpha F''_\alpha \backslash F'_\alpha$. (Here we allow $F'_\alpha = F''_\beta$.) For any chain $\mathcal{C}$ of subspaces of U, there is a canonical generalized flag $\mathit{fl}(\mathcal{C})$ associated with $\mathcal{C}$. The pairing between U and V defines the closure operation on subspaces of U given by $\bar{H} := H^{\perp\perp}$. This operation was first introduced and studied by Mackey in his thesis, see also [M]. For any generalized flag $\mathcal{F}$ in U we define the closure $\bar{\mathcal{F}}$ of $\mathcal{F}$ as $\mathit{fl}(\mathcal{F}^{\perp\perp})$, where $\mathcal{F}^{\perp\perp}$ denotes the chain in U consisting of the closures of all subspaces in $\mathcal{F}$. $\mathcal{F}$ is a *closed generalized flag in* U if $\bar{\mathcal{F}} = \mathcal{F}$, and $\mathcal{F}$ is a *strongly closed generalized flag in* U if $\mathcal{F}^{\perp\perp} = \mathcal{F}$. Clearly, any strongly closed generalized flag in U is closed. $\mathcal{F}$ is closed if and only if $\bar{F''_\alpha} = F''_\alpha$ and $\bar{F'_\alpha}$ equals either F'_α or F''_α. For any generalized flag $\mathcal{F}$ in U the subalgebra of $\mathfrak{g}$ which stabilizes $\mathcal{F}$ is $\mathrm{St}_{\mathcal{F}} = \sum_{\alpha} F''_{\alpha} \otimes (F'_{\alpha})^{\perp}$. The following theorem describes the Borel subalgebras of $\mathfrak{g}$.

Theorem 1. *The map $\mathcal{F} \mapsto \mathrm{St}_{\mathcal{F}}$ establishes a bijection between maximal closed generalized flags in U and Borel subalgebras of $\mathfrak{g}$.*

This theorem provides a rather explicit description of all Borel subalgebras of $\mathfrak{g}$. The results are most interesting in the case when both U and V are countable dimensional, i.e. $\mathfrak{g} \simeq gl(\infty)$. In this case we can represent $\mathfrak{g}$ as the direct limit $\varinjlim \mathfrak{g}_n$, where $\mathfrak{g}_n \simeq gl(n)$. It is clear that choosing a direct system of Borel subalgebras $\mathfrak{b}_n$ of $\mathfrak{g}_n$, the limit subalgebra $\mathfrak{b} := \varinjlim \mathfrak{b}_n$ is necessarily a Borel subalgebra of $\mathfrak{g}$. The converse, however, is not true. In fact we have the following theorem.

Theorem 2. *A Borel algebra $\mathfrak{b}$ of $\mathfrak{g}$ is the direct limit of Borel algebras $\mathfrak{b}_n$ of $\mathfrak{g}_n$ for some (but not every) direct system $\mathfrak{g} = \varinjlim \mathfrak{g}_n$, such that $\mathfrak{g}_n \simeq gl(n)$, if and only if the maximal closed generalized flags corresponding to $\mathfrak{b}$ both in U and in V are strongly closed.*

Finally, we consider the relationship between maximal toral subalgebras of $\mathfrak{g}$ and Borel subalgebras of $\mathfrak{g}$. We prove that, for any $\mathfrak{b} \subset \mathfrak{g}$, there exists a maximal toral subalgebra $\mathfrak{t} \subset \mathfrak{b}$ which is the compliment (as a vector space) of the locally nilpotent radial of $\mathfrak{b}$, i.e. $\mathfrak{b} = \mathfrak{t} \oplus \mathfrak{n}$, where $\mathfrak{n}$ is the locally nilpotent radical of $\mathfrak{b}$. Furthermore, we establish another criterion for $\mathfrak{b} = \varinjlim \mathfrak{b}_n$ as in Theorem 2. To state it we need to recall the definition of a splitting maximal toral subalgebra of $\mathfrak{g}$. A maximal toral subalgebra $\mathfrak{t}$ of $\mathfrak{g}$ is called *splitting* if it acts locally finitely on $\mathfrak{g}$, equivalently, if $\mathfrak{g}$ admits a root decomposition with respect to $\mathfrak{t}$. (For more details on maximal toral subalgebras of $\mathfrak{g}$ see [NP].) We then prove that the conditions of Theorem 2 are equivalent to the requirement that $\mathfrak{b}$ contains a splitting maximal toral subalgebra of $\mathfrak{g}$.

The talk is based on a joint work with Ivan Penkov.

References

[DP] Ivan Dimitrov and Ivan Penkov, Ind–varieties of generalized flags as homogeneous spaces for classical ind–groups, preprint, 2003, 17 pp.

[M] George Mackey, On infinite–dimensional linear spaces, Trans. Amer. Math. Soc. 57 (1945), 155–207.

[NP] Karl-Hermann Neeb and Ivan Penkov, Cartan subalgebra of $gl(\infty)$, Cand. Math. Bull. 46 (2003), 597–616.

Dolbeault cohomology of a loop space

László Lempert

(joint work with Ning Zhang (Riverside))

Loop spaces LM of compact complex manifolds M promise to have rich analytic cohomology theories, and it is expected that sheaf and Dolbeault cohomology groups of LM will shed new light on the complex geometry and analysis of M itself. This idea first occurs in [W], in the context of the infinite dimensional Dirac operator, and then in [HBJ] that touches upon Dolbeault groups of loop spaces; but in all this both works stay heuristic. Our goal here is to present rigorous results concerning the $H^{0,1}$ Dolbeault group of the first interesting loop space, that of the Riemann sphere $\mathbb{P}_1$. One noteworthy fact that emerges from this research is that analytic cohomology of loop spaces, unlike topological cohomology, is very sensitive to the regularity of loops admitted in the space. Another concerns local functionals, a notion from theoretical physics. Roughly, if M is a manifold, a local functional on a space of loops $x\colon S^1 \to M$ is one of form

$$f(x) = \int_{S^1} \Phi(t, x(t), \dot{x}(t), \ddot{x}(t), \dots)dt,$$

where Φ is a function on $S^1\times$ an appropriate jet bundle of M. It turns out that all cohomology classes in $H^{0,1}(L\mathbb{P}_1)$ are given by local functionals. Nonlocal cohomology classes exist only perturbatively, i.e., in a neighborhood of constant loops in $L\mathbb{P}_1$; but none of them extends to the whole of $L\mathbb{P}_1$.

We fix a smoothness class C^k, $k = 1, 2, \dots, \infty$, or Sobolev $W^{k,p}$, $k = 1, 2, \dots, 1 \le p < \infty$. If M is a finite dimensional complex manifold, consider the space $LM = L_k M$ resp. $L_{k,p}M$ of maps $S^1 = \mathbb{R}/\mathbb{Z} \to M$ of the given regularity. These spaces are complex manifolds modeled on a Banach space, except for $L_\infty M$, which is modeled on a Fréchet space. We shall focus on the loop space(s) $L\mathbb{P}_1$. As on any complex manifold, one can consider the space $C^\infty_{r,q}(L\mathbb{P}_1)$ of smooth (r, q) forms, the operators $\bar\partial : C^\infty_{r,q}(L\mathbb{P}_1) \to C^\infty_{r,q+1}(L\mathbb{P}_1)$, and the associated Dolbeault groups $H^{r,q}(L\mathbb{P}_1)$; for all this, see e.g. [L1,2]. On the other hand, let $\mathfrak{F}$ be the space of holomorphic functions $F\colon \mathbb{C} \times L\mathbb{C} \to \mathbb{C}$ that have the following properties:

(1) $F(\zeta/\lambda, \lambda^2 y) = O(\lambda^2)$, as $\mathbb{C} \ni \lambda \to 0$;
(2) $F(\zeta, x + y) = F(\zeta, x) + F(\zeta, y)$, if $\operatorname{supp} x \cap \operatorname{supp} y = \emptyset$;
(3) $F(\zeta, y + \text{const}) = F(\zeta, y)$.

As we shall see, the additivity property (2) implies $F(\zeta, y)$ is local in y.

Theorem 1. $H^{0,1}(L\mathbb{P}_1) \approx \mathbb{C} \oplus \mathfrak{F}$.

In the case of $L_\infty\mathbb{P}_1$, examples of $F \in \mathfrak{F}$ are

$$F(\zeta, y) = \zeta^\nu \left\langle \Phi, \prod_{j=0}^{m} y^{(d_j)} \right\rangle, \tag{1}$$

where Φ is a distribution on $S^1, y^{(d)}$ denotes d'th derivative, each $d_j \geq d_0 = 1$, and $0 \leq \nu \leq 2m$. A general function in $\mathfrak{F}$ can be approximated by linear combinations of functions of form (1).

This brings us to the issue of topology on $H^{0,1}(L\mathbb{P}_1)$ and on $\mathfrak{F}$. On any, possibly infinite dimensional complex manifold X the space $C^\infty_{r,q}(X)$ can be given the compact–C^∞ topology as follows. First, the compact–open topology on $C^\infty_{0,0}(X) = C^\infty(X)$ is generated by C^0 seminorms $\|f\|_K = \sup_K |f|$ for all $K \subset X$ compact. The family of C^ν seminorms is defined inductively: each $C^{\nu-1}$ seminorm $\| \quad \|$ on $C^\infty(TX)$ induces a C^ν seminorm $\|f\|' = \|df\|$ on $C^\infty(X)$. The collection of all C^ν seminorms, $\nu = 0, 1, \ldots$, defines the compact–C^∞ topology on $C^\infty(X)$. The compact–C^∞ topology on a general $C^\infty_{r,q}(X)$ is induced by the embedding $C^\infty_{r,q}(X) \subset C^\infty(\overset{r+q}{\oplus} TX)$. With this topology $C^\infty_{r,q}(X)$ is a separated locally convex vector space, complete if X is first countable. The quotient space $H^{r,q}(X)$ inherits a locally convex topology, not necessarily separated. We note that on the subspace $\mathcal{O}(X) \subset C^\infty(X)$ of holomorphic functions the compact–C^∞ topology restricts to the compact–open topology. The isomorphism in Theorem 1 is topological; it is also equivariant with respect to the obvious actions of the group of C^k diffeomorphisms of S^1.

There is another group, the group $G \approx \mathrm{PSL}(2, \mathbb{C})$ of holomorphic automorphisms of $\mathbb{P}_1$, whose holomorphic action on $L\mathbb{P}_1$ (by post–composition) and on $H^{0,1}(L\mathbb{P}_1)$ will be of greater concern to us. Theorems 2, 3, 4 below will describe the structure of $H^{0,1}(L\mathbb{P}_1)$ as a G–module. Recall that any irreducible (always holomorphic) G–module is isomorphic, for some $n = 0, 1, \ldots$, to the space $\mathfrak{K}_n$ of holomorphic differentials $\psi(\zeta)(d\zeta)^{-n}$ of order $-n$ on $\mathbb{P}_1$; here ψ is a polynomial, deg $\psi \leq 2n$, and G acts by pullback. The n'th isotypical subspace of a G–module V is the sum of all irreducible submodules isomorphic to $\mathfrak{K}_n$. In particular, the 0'th isotypical subspace is the space V^G of fixed vectors.

Theorem 2. *If $n \geq 1$, the n'th isotypical subspace of $H^{0,1}(L_\infty\mathbb{P}_1)$ is isomorphic to the space $\mathfrak{F}^n$ spanned by functions of form (0.1), with $m = n$.*

The fixed subspace of $H^{0,1}(L\mathbb{P}_1)$ can be described more explicitly, for any loop space:

Theorem 3. *The space $H^{0,1}(L\mathbb{P}_1)^G$ is isomorphic to $C^{k-1}(S^1)^*$ resp. $W^{k-1,p}(S^1)^*$, if the dual spaces are endowed with the compact–open topology.*

The isomorphisms in Theorem 3 are not Diff S^1 equivariant. To remedy this, one is led to introduce the spaces $C^l_r(S^1)$ resp. $W^{l,p}_r(S^1)$ of differentials $y(t)(dt)^r$ of order r on S^1, of the corresponding regularity; $L^p_r = W^{0,p}_r$. Then $H^{0,1}(L\mathbb{P}_1)^G$ will be Diff S^1 equivariantly isomorphic to $C^{k-1}_1(S^1)^*$, resp. $W^{k-1,p}_1(S^1)^*$.

For low regularity loop spaces one can very concretely represent all of $H^{0,1}(L\mathbb{P}_1)$:

Theorem 4. *(a) If $1 \leq p < 2$, all of $H^{0,1}(L_{1,p}\mathbb{P}_1)$ is fixed by G, hence it is isomorphic to $L^{p'}(S^1)$, with $p' = p/(p-1)$.*

(b) If $1 \le p < \infty$ then $H^{0,1}(L_{1,p}\mathbb{P}_1)$ is isomorphic to

$$\bigoplus_{0\le n\le p-1} \mathfrak{K}_n \otimes L^{p/(n+1)}_{n+1}(S^1)^* \approx \bigoplus_{0\le n\le p-1} \mathfrak{K}_n \otimes L^{p_n}_{-n}(S^1), \qquad p_n = \frac{p}{p-1-n},$$

and so it is the sum of its first $[p]$ isotypical subspaces. Indeed, the isomorphisms above are $G \times \mathrm{Diff}\, S^1$ equivariant, G, resp. Diff S^1 acting on one of the factors $\mathfrak{K}_n$, L^q_r naturally, and trivially on the other.

Again, the dual spaces are endowed with the compact–open topology.

To finish this write up, here is a list of relevant literature:

References

[BD] T. Bröcker, T. tom Dieck, *Representations of compact Lie groups,* Springer, New York, 1985.

[D] M.M. Day, *The space L^p with $0 < p < 1$,* Bull. Amer. Math. Soc. **46** (1940), 816-823.

[H] L. Hörmander, *The analysis of linear partial differential operators I,* Springer, Berlin, 1983.

[HBJ] F. Hirzebruch, T. Berger, R. Jung, *Manifolds and modular forms,* Viehweg, Braunschweig, 1992.

[K] J. Kurzweil, *On approximations in real Banach spaces,* Studia Math. **14** (1954), 214–231.

[L1] L. Lempert, *The Dolbeault complex in infinite dimensions, I,* J. Amer. Math. Soc. **11** (1998), 485–520.

[L2] ______, *Holomorphic functions on (generalised) loop spaces,* Proc. Royal Irish Acad. (to appear).

[MZ] J. Millson, B. Zombro, *A Kähler structure on the moduli space of isometric maps of a circle into Euclidean space,* Invent. Math. **123** (1996), 35–59.

[P] R. Palais, *Foundations of global nonlinear analysis,* Benjamin, Inc., New York–Amsterdam, 1968.

[W] E. Witten *The index of the Dirac operator in loop space,* Elliptic curves and modular forms in algebraic topology, Lecture Notes in Math., vol. 1326, Springer, Berlin, 1988, pp. 161–181.

[Z] N. Zhang, *Holomorphic line bundles on the loop space of the Riemann sphere,* J. Diff. Geom. (to appear).

Infinite-Dimensional Homogeneous Spaces and Operator Ideals

Daniel Beltiţă

The existence of invariant Kähler structures on homogeneous spaces of certain Lie groups turns out to be a phenomenon that is not confined to finite dimensions. Our research concerns this phenomenon in the case of some classes of infinite-dimensional Lie groups associated with ideals of compact operators on Hilbert spaces.

More specifically, we have introduced in the paper [Be03] a notion of admissible pair of operator ideals $(\mathfrak{I}_0, \mathfrak{I}_1)$ and have used it to construct Kähler homogeneous spaces of Banach-Lie groups naturally associated with such pairs. One special instance of admissible pair is a pair of Schatten ideals $(\mathfrak{S}_p, \mathfrak{S}_q)$, where $2 \le p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. More generally, certain dual pairs of Lorentz ideals also turn out to be admissible.

Here is the precise definition of an admissible pair:

Definition. Let $\mathcal{H}$ be a complex Hilbert space and $\mathcal{B}(\mathcal{H})$ the Banach algebra of all bounded linear operators on $\mathcal{H}$. An *admissible pair* of ideals of $\mathcal{B}(\mathcal{H})$ is a pair $(\mathfrak{I}_0, \mathfrak{I}_1)$ of two-sided ideals of $\mathcal{B}(\mathcal{H})$ satisfying the following conditions:

(a) The ideal $\mathfrak{I}_0$ is equipped with a norm $\|\cdot\|_{\mathfrak{I}_0}$ making it into a reflexive separable Banach space satisfying

$$\|T\| \le \|T\|_{\mathfrak{I}_0} = \|T^*\|_{\mathfrak{I}_0} \text{ and } \|ATB\|_{\mathfrak{I}_0} \le \|A\| \cdot \|T\|_{\mathfrak{I}_0} \cdot \|B\|$$

whenever $A, B \in \mathcal{B}(\mathcal{H})$ and $T \in \mathfrak{I}_0$.

(b) We have $\mathfrak{I}_1 \cdot \mathfrak{I}_0 \subseteq \mathfrak{S}_1(\mathcal{H})$ and the bilinear functional

$$\mathfrak{I}_1 \times \mathfrak{I}_0 \to \mathbb{C}, \quad (K, T) \mapsto \mathrm{Trace}\,(KT)$$

induces a vector space isomorphism of $\mathfrak{I}_1$ onto the topological dual of the Banach space $(\mathfrak{I}_0, \|\cdot\|_{\mathfrak{I}_0})$, where $\mathfrak{S}_1(\mathcal{H})$ denotes the trace class on $\mathcal{H}$.

(c) We have $\mathfrak{I}_1 \subseteq \mathfrak{I}_0$.

Using the notion of admissible pair, one can construct infinite-dimensional Kähler manifolds as described in the following theorem. In this statement, for any operator ideal $\mathfrak{I}$ we denote by $\mathfrak{u}_{\mathfrak{I}} = \{T \in \mathfrak{I} \mid T^* = -T\}$ the Lie algebra of skew-adjoint operators in $\mathfrak{I}$, and we also denote by $\mathrm{U}_{\mathfrak{I}} = \{T \in \mathrm{id}_{\mathcal{H}} + \mathfrak{I} \mid T^*T = TT^* = \mathrm{id}_{\mathcal{H}}\}$ the group of all unitary operators in $\mathrm{id}_{\mathcal{H}} + \mathfrak{I}$.

Theorem. *Let $(\mathfrak{I}_0, \mathfrak{I}_1)$ be an admissible pair of ideals of $\mathcal{B}(\mathcal{H})$ and A a self-adjoint element of $\mathcal{B}(\mathcal{H})$. Consider the following objects:*

- $\mathrm{U}_{\mathfrak{I}_0,\mathfrak{I}_1}(A) = \{T \in \mathrm{U}_{\mathfrak{I}_0} \mid T^*AT \in A + \mathfrak{I}_1\} = \{T \in \mathrm{U}_{\mathfrak{I}_0} \mid [A,T] \in \mathfrak{I}_1\}$,
- $H_{\mathfrak{I}_0,A} = \{T \in \mathrm{U}_{\mathfrak{I}_0} \mid T^*AT = A\}$,
- $\mathfrak{u}_{\mathfrak{I}_0,\mathfrak{I}_1}(A) = \{T \in \mathfrak{u}_{\mathfrak{I}_0} \mid [A,T] \in \mathfrak{I}_1\}$,
- $\omega\colon \mathfrak{u}_{\mathfrak{I}_0,\mathfrak{I}_1}(A) \times \mathfrak{u}_{\mathfrak{I}_0,\mathfrak{I}_1}(A) \to \mathbb{R}, \quad \omega(T_1, T_2) = \mathrm{Trace}\,(\mathrm{i}[A, T_1]T_2)$.

Then the following assertions hold.

(a) *The group $\mathrm{U}_{\mathfrak{I}_0,\mathfrak{I}_1}(A)$ has a natural structure of connected real Banach-Lie group with the Lie algebra $\mathfrak{u}_{\mathfrak{I}_0,\mathfrak{I}_1}(A)$, and the bilinear functional ω is a continuous 2-cocycle of the real Banach-Lie algebra $\mathfrak{u}_{\mathfrak{I}_0,\mathfrak{I}_1}(A)$. Furthermore, $H_{\mathfrak{I}_0,A}$ is a Banach-Lie subgroup of $\mathrm{U}_{\mathfrak{I}_0,\mathfrak{I}_1}(A)$ whose Lie algebra equals $\{T \in \mathfrak{u}_{\mathfrak{I}_0,\mathfrak{I}_1}(A) \mid \omega(T,\cdot) \equiv 0\}$.*

(b) *The 2-cocycle ω induces a $\mathrm{U}_{\mathfrak{I}_0,\mathfrak{I}_1}(A)$-invariant weakly symplectic form Ω on the homogeneous space $\mathrm{U}_{\mathfrak{I}_0,\mathfrak{I}_1}(A)/H_{\mathfrak{I}_0,A}$.*

(c) *If the spectrum of the operator A is finite, then there exists a* $\mathrm{U}_{\mathfrak{I}_0,\mathfrak{I}_1}(A)$*-invariant complex structure making the weakly symplectic homogeneous space*
$(\mathrm{U}_{\mathfrak{I}_0,\mathfrak{I}_1}(A)/H_{\mathfrak{I}_0,A},\Omega)$ *into a weakly Kähler homogeneous space.*

We now outline the method used in [Be03] to construct the aforementioned Kähler structures. The main point is that we actually study Banach-Lie groups associated with admissible pairs and with certain n-tuples of self-adjoint operators. We use the joint functional calculus of those n-tuples (which is a special instance of the Weyl functional calculus) to construct Kähler polarizations in the complexified Lie algebras of the Lie groups under consideration. In fact, the polarizations arise as spectral subspaces corresponding to certain subsets of the joint spectrum of the corresponding n-tuple. A remarkable point of this approach is that it actually holds in a quite general setting. E.g., besides the homogeneous spaces of groups associated with operator ideals, that approach leads to complex structures on the flag manifolds associated with arbitrary associative unital Banach algebras.

We mention that certain special instances of the complex homogeneous spaces constructed by the above described method were already shown to play a significant role in representation theory of certain Hilbert-Lie groups associated with the Hilbert-Schmidt ideal (see e.g., [Bo80], [Ca85], [Ne00], [Ne02]). From this point of view, it is interesting to investigate the role played by the new classes of complex homogeneous spaces in the representation theory of more general Banach-Lie groups. On the other hand, it would be important to understand whether the specific properties of the operator ideals correspond to any particular phenomena in the complex geometry of the corresponding homogeneous spaces (compare also [Up85]).

References

[Be03] D. Beltiţă, *Complex homogeneous spaces of pseudo-restricted groups*, Math. Research Letters 10(2003), no. 4, 459–467.

[Bo80] R.P. Boyer, *Representation theory of the Hilbert-Lie group* $\mathrm{U}(\mathfrak{H})_2$*,* Duke Math. J. 47(1980), no. 2, 325-344.

[Ca85] A.L. Carey, *Some homogeneous spaces and representations of the Hilbert Lie group* $\mathcal{U}(H)_2$*.* Rev. Roumaine Math. Pures Appl. 30(1985), no. 7, 505–520.

[Ne00] K.-H. Neeb, *Infinite-dimensional groups and their representations.* (Lectures at the European School in Group Theory, SDU-Odense Univ., August 2000), Preprint Technische Universität Darmstadt, no. 2206 (2002), April.

[Ne02] K.-H. Neeb, *Highest weight representations and infinite-dimensional Kähler manifolds,* in: Recent advances in Lie theory (Vigo, 2000), vol. 25, Res. Exp. Math., Heldermann, Lemgo, 2002, pp. 367-392.

[Up85] H. Upmeier, *Symmetric Banach Manifolds and Jordan* C^**-algebras.* North-Holland Mathematics Studies, 104. Notas de Matemática , 96. North-Holland Publishing Co., Amsterdam, 1985.

Realizing Lie Groups as Automorphism Groups of Complex Manifolds

Jörg Winkelmann

Let X be a hyperbolic (in the sense of Kobayashi) complex connected manifold. Then the group of all holomorphic automorphisms of X (endowed with the compact-open topology) is a finite-dimensional real Lie group with countably many connected components. This raises the question whether conversely every such Lie group can be realized as a the full automorphism group of a hyperbolic complex manifold.

We prove that this is true if the group is connected or discrete.

Theorem 1. *Let G be a (finite-dimensional) real connected Lie group or a countable discrete group.*

Then there exists a Stein hyperbolic connected complex manifold X such that G is isomorphic to the group of all automorphisms (i.e. biholomorphic selfmaps) of X.

The first step in this direction was the result for compact Lie groups. Saerens and Zame ([5]), and independently Bedford and Dadok ([1]) proved that, given a compact real Lie group K there always exists a strictly pseudoconvex bounded domain $D \subset \mathbb{C}^n$ such that $Aut(D) \simeq K$. By the theorem of Wong-Rosay (which states that every strictly pseudoconvex bounded domain with non-compact automorphism group is isomorphic to the ball) it is clear that an arbitrary non-compact real Lie group can not be realized as the automorphism of a strictly pseudoconvex bounded domain in $\mathbb{C}^n$. However, as proved in [8], for any connected real Lie group G there does exist a complex manifold X on which G acts effectively. Moreover, X can be chosen in such a way that it enjoys several of the key properties of strictly pseudoconvex bounded domains. Namely, X can be chosen such that it is both Stein and hyperbolic in the sense of Kobayashi.

In [10] we verified that one can rule out additional automorphisms, i.e. it is possible to achieve $Aut(X) \simeq G$. The precise result is the following:

Theorem 2. *Let G be a connected real Lie group. Then there exists a Stein, complete hyperbolic complex manifold X on which G acts effectively, freely, properly and with totally real orbits such that* $\mathrm{Aut}_{\mathcal{O}}(X) \simeq G$.

The idea is to follow the strategy of Saerens and Zame: Construct the desired manifold as an open subset of a larger Stein manifold in such a way that the given group acts on this open subset. Ensure that every automorphism of this open subset can be extended to the boundary, then modify the boundary in such a way that this CR-hypersurface simply has no automorphisms other than those from the given group. The latter can be done using the fact that a CR-hypersurface (unlike a complex manifold) does have local invariants. A principal difficulty in this approach is to obtain an extension of automorphisms of the open subset to the boundary. If one is concerned only with compact Lie groups, then one can work with a strictly pseudoconvex bounded domain D. For such a domain it is evident that for every automorphism ϕ of D there exists a sequence $x_n \in D$ such that both x_n and $\phi(x_n)$ converge to a strictly pseudoconvex point in the boundary. This is the starting point for the extension of the automorphism ϕ to the boundary ∂D.

Now, our goal is to obtain a result for arbitrary connected Lie groups, which are not necessarily compact.

This lack of compactness assumption creates some difficulties.

There are two main problems: First, an arbitrary non-compact Lie group is not necessarily linear. For instance, the universal cover of $SL_2(\mathbb{R})$ cannot be embedded into a linear group. Second, as already mentioned, the theorem of Wong-Rosay implies that in general a non-compact Lie group can not be realized as the full automorphism group of a strictly pseudoconvex bounded domain with smooth boundary. Thus we have to work with domains which are not bounded or where the boundary is not everywhere smooth. The trouble is that it is therefore no longer clear that for every automorphism ϕ there exists a sequence x_n in the domain such that both x_n and $\phi(x_n)$ converge to a nice point in the boundary.

In [7] a result similar to ours is claimed for certain Lie groups with a rather sketchy outline of a possible proof.

The first of the aforementioned two problems is dealt with by *assuming* the group G to be linear while the second problem is simply ignored. Since the second problem is in fact a serious obstacle, the proof sketched in [7] can not be regarded as complete.

We proceed in the following way: To deal with the first problem, we note that every Lie algebra is linear by the theorem of Ado. Therefore, in a certain sense, every Lie group is linear up to coverings and the first problem can be attacked by working carefully with coverings.

For the second problem, we use bounded domains whose boundaries are smooth outside an exceptional set E which is small in a certain sense. Exploiting this smallness we prove that for every automorphism ϕ there must exist a sequence x_n such that both x_n and $\phi(x_n)$ converge to a boundary point outside the "bad set" E.

Once this has been verified, we can prove (using arguments similar to those used in[5], [1]) that ϕ extends as holomorphic map near $\lim(x_n)$, and use the theory of Chern-Moser-invariants to deduce that ϕ was in fact given by left multiplication with an element of G.

For discrete groups the following statement is proved in [9]:

Theorem 3. *Let G be a countable discrete group. Then there exists a non-compact Riemann surface X, hyperbolic in the sense of Kobayashi, such that G is isomorphic to the automorphism group of X.*

References

[1] Bedford, E.; Dadok, J.: Bounded domains with prescribed group of automorphisms. *Comm. Math. Helv.* **62** , 561–572 (1987)

[2] Borel, A.: Semisimple groups and Riemannian symmetric spaces. Texts and Readings in Mathematics **16**, Hindustan Book Agency. (1998)

[3] Diederich, K.; Fornaess, J.: Proper Holomorphic Mappings between real-analytic pseudoconvex domains in $\mathbb{C}^n$. *Math. Ann.* **282**, 681-700 (1988)
[4] Forstneric, F.; Rosay, J.P.: Localization of the Kobayashi Metric and the boundary Countinuity of Proper Holomorphic Mappings. *Math. Ann.* **279** , 239–252 (1987)
[5] Saerens, R.; Zame, W.R.: The Isometry Groups of Manifolds and the Automorphism Groups of Domains. *Trans. A.M.S.* **301**, no. 1, 413-429 (1987)
[6] Stein, K.: Überlagerungen holomorph–vollständiger komplexer Räume. *Arch. Math.* **VIII** , 354–361 (1956)
[7] Tumanov, A.E.; Shabat, G.B.: Realization of linear Lie Groups by Biholomorphic Automorphisms *Funct. Anal. Appl.* **24**, 255–257 (1991)
[8] Winkelmann, J.: Invariant Hyperbolic Stein Domains. *manu. math.* **79**, 329–334 (1993)
[9] Winkelmann, J.: Realizing Countable Groups as Automorphism Groups of Riemann Surfaces. *documenta math.*, (2002)
[10] Winkelmann, J.: Realizing Connected Lie Groups as Automorphism Groups of Complex Manifolds. *Commentarii Mathematici Helvetici* , (2004)

Principal Series Representations and Dirac Operators

Roger Zierau

Kostant defined a remarkable invariant differential operator in [7] which he called the cubic Dirac operator. Given a connected semisimple Lie group, a closed reductive subgroup and a homogeneous vector bundle $\mathcal{E} \to G/H$ of finite rank, the cubic Dirac operator is a G-invariant differential operator on sections:

$$D : C^\infty(G/H, \mathcal{E} \otimes \mathcal{S}) \to C^\infty(G/H, \mathcal{E} \otimes \mathcal{S}). \tag{1}$$

Here S is the spin representation of $\mathfrak{h}$. In this lecture we discuss joint work with Salah Mehdi in which the kernel of D is studied when H is noncompact and $\operatorname{rank}(\mathfrak{g}) = \operatorname{rank}(\mathfrak{h})$. The main result is an integral formula for certain solutions of $Df = 0$. In particular, the kernel is nonzero and certain interesting representations occur.

The cubic Dirac operator is defined as follows. There is an orthogonal decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{q}$ with respect to the Killing form of $\mathfrak{g}$ (however, we need to assume the Killing form on $\mathfrak{h}$ is nondegenerate). Then $\mathfrak{q}$ is equipped with a nondegenerate (possibly indefinite) symmetric form. Thus, one may build a corresponding Clifford algebra and spin representation of $\mathfrak{so}(\mathfrak{q})$. Since $ad : \mathfrak{h} \to \mathfrak{so}(\mathfrak{q})$ we obtain the representation $\sigma \circ \mathrm{ad}$ of $\mathfrak{h}$, which we call the spin representation of $\mathfrak{h}$. In (1) we require only that E is a representation of $\mathfrak{h}$ so that $E \otimes S$ integrates to a representation of H. Then $\mathcal{E} \otimes \mathcal{S} \to G/H$ is the corresponding homogeneous vector bundle. Now choose a basis $\{X_j\}$ of $\mathfrak{q}$ so that $\langle X_j , X_k \rangle_{\mathfrak{q}} = \epsilon_j \delta_{jk}$, with $\epsilon_j = \pm 1$. Let $c \in Cl(\mathfrak{q})$ correspond to the alternating 3-form $\langle X , [Y, Z] \rangle_{\mathfrak{q}}$ on $\mathfrak{q}$. The cubic Dirac operator of (1) is defined by

$$D = \sum_j \epsilon_j r(X_j) \otimes 1 \otimes \gamma(X_j) + 1 \otimes 1 \otimes \gamma(c). \tag{2}$$

Here γ denotes Clifford multiplication and $r(X)$ is the right action of $X \in \mathfrak{g}$ on functions.

There are several well-known cases where such an operator has been studied. Most notably, when H is a maximal compact subgroup of G, then $c = 0$ and D is the 'usual' Dirac operator arising from an invariant connection. In this case, the kernel of D (on L_2-sections) is a relative discrete series representation and all relative discrete series representations of G occur this way. See [11], [1] and [12]. Another case is when G is compact. Then, in [8] and [9] the kernel of D is seen to be an irreducible G-representation. This is a generalization of the Bott-Borel-Weil Theorem. A remarkable property of D which relates D to infinitesimal character is contained in [6].

Now let us turn to a noncompact group G and noncompact reductive subgroup H. Let E and S be as above. Our goal is to study the kernel of D and our approach is to find a G-intertwining map from a principal series representation of G into $Ker(D)$.

We briefly describe the construction. Let θ be a Cartan involution of $\mathfrak{g}$ which stabilizes $\mathfrak{h}$ and let $\mathfrak{g} = \mathfrak{k} + \mathfrak{s}$ be the corresponding Cartan decomposition of $\mathfrak{g}$. The principal series consists of representations induced from representations of real parabolic subgroups of G. Our subgroup H determines a parabolic subgroup as follows. Choose a maximal abelian subalgebra $\mathfrak{a}$ in $\mathfrak{h} \cap \mathfrak{s}$. Then $\mathfrak{a}$ determines a parabolic $P = MAN$ (up to a choice of N). Note that it is important here that $\mathfrak{g}_{\mathbf{C}}$ and $\mathfrak{h}_{\mathbf{C}}$ have the same ranks. It follows that $P \cap H = (M \cap H)A(N \cap H)$ is a minimal parabolic subgroup of H. In particular $H \cap K \cdot eP = H \cdot eP$ is a closed H-orbit in G/P.

Lemma 3. *Each relative discrete series representation of M occurs in the kernel of*
$$D_{M/M\cap H} : \mathbf{C}^\infty(M/M \cap H, \mathcal{F} \otimes S) \to \mathbf{C}^\infty(M/M \cap H, \mathcal{F} \otimes S)$$
for some homogeneous bundle $\mathcal{F} \to M/M \cap H$. Note that, with our choice of P, $M \cap H$ is compact.

This Lemma is of course related to the results on the discrete series mentioned above. However, here we are not concerned with the L_2 statement; by relative discrete series here we mean a representation infinitesimally equivalent to a relative discrete series representation.

For a representation W of P we write $C^\infty(G/P, \mathcal{W})$ for the induced representation (the smooth principal series representation).

Lemma 4. *For any smooth representation W of P, given some nonzero $t \in Hom_{P\cap H}(W \otimes \mathbf{C}_{\rho+2\rho_{\mathfrak{h}}}, E \otimes S)$ there is a nonzero G-intertwining map*
$$P_t : C^\infty(G/P, \mathcal{W} \otimes \mathbf{C}_{\rho_{\mathfrak{h}}}) \to \mathbf{C}^\infty(G/H, \mathcal{E} \otimes S)$$
$$(P_t\phi)(g) = \int_{H\cap K} \ell \cdot (\phi(g\ell))\, d\ell.$$

Therefore, we need to find a W and t so that the image of P_t lies in the kernel of D. This is accomplished by finding a relative discrete series representation W of M so that, when realizing W as $Ker(D_{M/M\cap H})$ as in Lemma 3, t is evaluation at $e \in M$ and the following holds.

Theorem 5. *When the highest weight μ of E is sufficiently regular, the image of P_t lies in $Ker(D)$.*

Remark 6. Note the analogy between our construction and that of the Poisson integral. The Poisson integral is a formula giving harmonic functions on the unit disk in $\mathbf{C}$. In fact, the generalization of this is the Poisson transform (see, for example, [5, Ch. II, Section 4.1]) producing joint eigenfunctions of the G-invariant differential operators on the riemannian symmetric space G/K. One notes that the Poisson transform is an integral over the boundary G/P of G/K and the formula comes from an analogue of Lemma 4 with S and E replaced by the trivial representation. In our setting, $H \cap K \cdot eP = H \cdot eP \subset G/P$. We may therefore say that integration over 'a piece of the boundary' of G/H gives solutions to the Dirac equation $Df = 0$.

Remark 7. The results discussed here may be viewed as a generalization of [3], [2] and [4] in the following sense. If G/H is a measurable open orbit in a flag variety (i.e., an elliptic coadjoint orbit), then $D = \overline{\partial} + \overline{\partial}^*$. In this case, the operator initially studied in [3] coincides with the intertwining operator P_t above.

Remark 8. The principal series representations are fairly well understood. Thus, certain representatins occurring in $Ker(D)$ can be identified via the Langlands classification. Furthermore, the growth of harmonic spinors of the form $P_t\phi$ can be studied by considering properties of ϕ and using techniques of Harish-Chandra.

References

[1] M. Atiyah and W. Schmid, *A geometric construction of the discrete series for semisimple Lie groups*, Inv. Math. **42** (1977), pp. 1-62.

[2] L. Barchini, Szegö mappings, harmonic forms and Dolbeault cohomology, J. Funct. Anal. **118** (1993), pp. 351–406.

[3] L. Barchini, A. W. Knapp and R. Zierau, Intertwining operators into Dolbeault cohomology representations, J. Funct. Anal. **107** (1992), pp. 302–341.

[4] R. W. Donley, Intertwining operators into cohomology representations for semisimple Lie groups, J. Funct. Anal. **151** (1997), pp. 138-165.

[5] S. Helgason, Groups and Geometric analysis, 1984, Academic Press, New York.

[6] J.-S. Huang and P. Pandžić, *Dirac cohomology, unitary representations and a proof of a conjecture of Vogan*, J. Amer. Math. Soc. **15** (2002), no. 1, pp. 185-202.

[7] B. Kostant, *A cubic Dirac operator and the emergence of Euler number multiplets of representations for equal rank subgroups*, Duke Math. J. **100** (1999), no. 3, pp. 447-501.

[8] B. Kostant, *A generalization of the Bott-Borel-Weil theorem and Euler number multiplets of representations*, Conference Mosh Flato 1999 (Dijon). Lett. Math. Phys. **52** (2000), no. 1, pp. 61-78.

[9] G. D. Landweber,
Harmonic spinors on homogeneous spaces, Representation Theory **4** (2000), pp. 466-473.

[10] S. Mehdi and R. Zierau, *Principal series representations and harmonic spinors*, preprint.

[11] R. Parthasarathy,
Dirac operator and discrete series, Ann. of Math. **96** (1972), pp. 1-30.

[12] J. A. Wolf,
Partially harmonic spinors and representations of reductive Lie groups, J. of Functional Analysis **15**, no. 2, (1974), pp. 117-154.

Theta lifting of unitary lowest weight representations and their associated cycles

Kyo Nishiyama

We consider a reductive dual pair (G, G') in the stable range with G' the smaller member and of Hermitian symmetric type. Namely, the following three kinds of dual pairs will be treated.

	the pair (G, G')	stable range condition
Case $\mathbb{R}$:	$(O(p,q), Sp(2n,\mathbb{R}))$	$2n < \min(p,q)$
Case $\mathbb{C}$:	$(U(p,q), U(m,n))$	$m+n \leq \min(p,q)$
Case $\mathbb{H}$:	$(Sp(p,q), O^*(2n))$	$n \leq \min(p,q)$

We study the theta lifting of a unitary lowest weight representation π' of G', which may be singular. The main result is an explicit determination of the associated cycle of the lifted representation $\theta(\pi')$. More precisely, we prove that

$$\theta(\mathcal{AC}(\pi')) = \mathcal{AC}(\theta(\pi')),$$

where θ(associated cycle) means the theta lifting of nilpotent orbits in the stable range. We also obtained a K-type formula for $\theta(\pi')$ in terms of the branching coefficient of classical groups; the associated nilpotent orbit is realized as a quotient of a minimal nilpotent orbit of a lager group. The K-type formula is not new though, since $\theta(\pi')$ is a derived functor module. However, our K-type formula is not a variant of Blattner's one, and we believe ours has some advantage.

Also, we have given a brief survey on the associated cycles of the unitary lowest weight representations in the terminology of classical invariant theory ([1]). This idea is crucial for the investigation of the theta lifting of the lowest weight representations explained above.

The talk is based on the joint research ([2], [3], [4]) with Chen-bo Zhu (National University of Singapore) and Hiroyuki Ochiai (Nagoya University).

References

[1] Kyo Nishiyama, Hiroyuki Ochiai, Kenji Taniguchi, Hiroshi Yamashita and Shohei Kato, *Nilpotent orbits, associated cycles and Whittaker models for highest weight representations. Astérisque* **273** (2001).
[2] Kyo Nishiyama and Chen-bo Zhu, Theta lifting of holomorphic discrete series, (The case of $U(p,q) \times U(n,n)$). *Trans. AMS.* **353** (2001), 3327-3345.
[3] Kyo Nishiyama and Chen-bo Zhu, Theta lifting of unitary lowest weight modules and their associate cycles. To appear in *Duke Math. Jour.*
[4] Kyo Nishiyama Hiroyuki Ochiai, and Chen-bo Zhu. Theta lifting of nilpotent orbits for symmetric pairs. **math.RT/0312453**, 2003.

Quantum Chaos and Cohomology of Arithmetic Groups

Joachim Hilgert
(joint work with A. Deitmar)

Our work [1] is motivated by the following problem: given a classical system (symplectic manifold plus Hamiltonian function) and a quantization of this system (Hilbert space plus a self adjoint operator), can one detect from the quantum system whether the classical system shows chaotic behavior (e.g. ergodic or hyperbolic behavior)? For the modular surface and its geodesic flow (so that a suitable quantization is given by the corresponding L^2-space together with the Laplace-Beltrami operator Δ) Lewis and Zagier [2] have constructed a natural correspondence between Maass cusp forms (which are eigenfunctions of Δ) and holomorphic functions $\psi\colon \mathbb{C} \setminus \mathbb{R}^- \to \mathbb{C}$ satisfying a three term functional equation (called the Lewis equation) which has a natural interpretation in terms of the classical system. So far one has this correspondence only for this surfaces, but it is expected that it can be extended to coverings or even more general locally symmetric spaces of finite volume.

The Lewis equation admits a cohomological interpretation which suggests a starting point for generalizations. On the other hand Maass cusp forms can be defined in terms of representation theory and correspond to Γ-invariant vectors in principal series representations π_s of $\mathrm{PSL}(2,\mathbb{R})$, which leads to an interpretation of the dimension of the space of Maass cusp forms as multiplicities $N_\Gamma(\pi_s)$ of π_s in $L^2(\Gamma\backslash G)$.

Our main theorem is the following multiplicity formula for split semisimple Lie groups with arithmetic torsion free subgroups: If π is any irreducible unitary principal series representation and r, d the rank, respectively the dimension of the non-compact Riemannian symmetric space associated with G, then

$$N_\Gamma(\pi) = \sum_{j\geq 0} (-1)^{j+r} \binom{j}{r} \dim H^{N-j}_{\text{cusp}}(\Gamma, \pi^\omega).$$

In order to prove this formula establish a functorial isomorphism

$$H^j(\mathfrak{g}, KF\hat{\otimes}V^{\max} \to \mathrm{Ext}^j_{\mathfrak{g},K}(\tilde{V}, F)$$

for Harish-Chandra modules V (then $V^{\max}$ is the maximal globalization and $\tilde{V}$ is the dual Harisch-Chandra module) and smooth G-representations F, as well as a Poincaré duality

$$H^j_{\text{cusp}}(\Gamma, V^{\max}) \cong H^{N-j}_{\text{cusp}}(\Gamma, \tilde{V}^{\min}),$$

where $\tilde{V}^{\min}$ is the minimal globalization of $\tilde{V}$. As a corollary we derive

Theorem: Let Γ be a Fuchsian group of finite covolume and $s \in \mathbb{R}$. Then $N_\Gamma(\pi_s) = \dim H^1_{\text{cusp}}(\Gamma, \pi_s^\omega)$, where π_s^ω is the G-module of analytic vectors in the representation space of π_s.

References

[1] A. Deitmar, J. Hilgert: Cohomology of arithmetic groups with infinite dimensional coefficient spaces. Preprint arXiv: math.NT/0311526

[2] J.Lewis and D. Zagier: Period functions for Maass wave forms. I Annals of Mathematics, II. Series Vol. 153, 191-258 (2001)

[3] J. Hilgert , D. Mayer, H. Movasati : $\Gamma_0(n)$ and the Hecke operators for the period functions for $PSL(2, Z)$. Preprint arXiv: math.NT/0303251

Global deformations of the Virasoro algebra

Alice Fialowski

This talk is based on a joint work with Martin Schlichenmaier (see [4]).

Introduction. Deformation is one of the tools to study a specific object, by deforming it into some families of "similar" structure objects. Another question related to deformation: Can we equip the set of nonequivalent deformations with the structure of a topological or maybe geometric space? In other words, does there exist a moduli space for these structures? If so, then for a fixed object its deformations should reflect the local structure of the moduli space at the point corresponding to this object.

There is a lot of confusion in the literature in the notion of a deformation. Several different (inequivalent) approaches exist. May aim now is to clarify the difference between deformations of geometric origin and so-called formal deformations. Formal deformation theory has the advantage of using cohomology. It is also complete in the sense that under some natural cohomology assumptions there exists a versal formal deformation which induces all other deformations. Formal deformations are deformations with a complete local algebra base. A deformation with a commutative (non-local) algebra base gives a much richer picture of deformation families, depending on the augmentation of the base algebra. If we identify the base of deformation – which is a commutative algebra of functions – with a smooth manifold, an augmentation corresponds to choosing a point on the manifold. So choosing different points should in general lead to different deformation situations. I will show in the case of the Witt and Virasoro algebra that – in the case of infinite dimensional Lie algebras – there is no tight relation between global formal deformations.

1. Deformations. Let $\mathcal{L}$ be a Lie algebra.

i) Intuitively: One-parameter family $\mathcal{L}_t$ of Lie algebras with bracket $\mu_t = \mu_0 + t\phi_1 + t^2\phi_2 + \dots$.

ii) Global deformations: Consider a deformation $\mathcal{L}_t$ not as a family of Lie algebras, but as a Lie algebra over the algebra $\mathbb{K}[[t]]$. Call it the *base* of the deformation. The natural generalization is to allow more parameters, or to take in general a commutative algebra A over $\mathbb{K}$ with identity as base of a deformation. Take such an A over $\mathbb{K}$ of char 0 with an augmentation $\varepsilon : A \to \mathbb{K}$ and $m = \operatorname{Ker}\varepsilon$ maximal ideal.

Definition. A *global deformation* λ of $\mathcal{L}$ *with base* (A, m) is a Lie A-algebra structure on $A\otimes_{\mathbb{K}}\mathcal{L}$ with $[\ ,\ ]_\lambda$ such that $\varepsilon \otimes \mathrm{id} : A \otimes \mathcal{L} \to \mathbb{K} \otimes \mathcal{L} = \mathcal{L}$ is a Lie algebra homomorphism.

A deformation is called *trivial* if $A\otimes_{\mathbb{K}}\mathcal{L}$ carries the trivially extended Lie structure, i.e. $[1 \otimes x, 1 \otimes y]_\lambda = 1 \otimes [x, y]$. Two deformations of a Lie algebra $\mathcal{L}$ with the same base A are called *equivalent* if there exists a Lie algebra isomorphism between the two copies of $A\otimes\mathcal{L}$ with the two Lie algebra structures, compatible with $\varepsilon \otimes \mathrm{id}$. We say that the deformation is *local* if A is a local $\mathbb{K}$-algebra with unique maximal ideal $m_A = \operatorname{Ker}\varepsilon$. In case that in addition, $m_A^2 > 0$, the deformation is called *infinitesimal*.

iii) We call a deformation *formal*, if its base is a complete local algebra (with a unique maximal ideal) (see [1]).

Proposition (see [3])**.** *If* $\dim \mathrm{H}^2(\mathcal{L}, \mathcal{L}) < \infty$, *there exists a universal infinitesimal deformation* $\eta_{\mathcal{L}}$ *of* $\mathcal{L}$ *with base* $B = \mathbb{K} \oplus \mathrm{H}^2(\mathcal{L}, \mathcal{L})'$.

This means that for any infinitesimal deformation λ of the Lie algebra $\mathcal{L}$ with finite-dimensional (local) algebra base A there exists a unique homomorphism $\phi : \mathbb{K} \oplus \mathrm{H}^2(\mathcal{L}, \mathcal{L})' \to A$ such that λ is equivalent to the push-out $\phi_*\eta_{\mathcal{L}}$.

Definition ([1])**.** A formal deformation η of $\mathcal{L}$ parametrized by a complete local algebra B is called *versal* if for any deformation λ, parametrized by (A, m_A), there exists $f : B \to A$ morphism such that the push-out

1) $f_*\eta$ is equivalent to λ.

2) If A satisfies ${m_A}^2 = 0$, then f is unique.

Theorem. *Assume* $\mathrm{H}^2(\mathcal{L}, \mathcal{L})$ *is finite dimensional.*

a) ([1]) *There exists a versal formal deformation of* $\mathcal{L}$.

b) ([3]) *The base of the versal deformation is formally embedded into* $\mathrm{H}^2(\mathcal{L}, \mathcal{L})$, *i.e. it can be described in* $\mathrm{H}^2(\mathcal{L}, \mathcal{L})$ *by a finite system of formal equation.*

Corollary. $\mathrm{H}^2(\mathcal{L}, \mathcal{L}) = \{0\}$ *implies that* $\mathcal{L}$ *is formally rigid.*

Theorem ([2])**.** *The Witt and Virasoro algebra is formally rigid.*

2. Krichever–Novikov algebras. They are generalizations of the Virasoro and all its related algebras. Let M be a compact Riemann surface of genus g, or a smooth projective curve over $\mathbb{C}$. Let $I = \{P\}$ and $O = \{Q\}$ be distinct points ("marked points") on the curve. Denote $A = I \cup O$ as a set. Denote by $\mathcal{L}$ the Lie algebra consisting of those meromorphic sections of the holomorphic tangent line bundle which are holomorphic outside of A, equipped with the Lie bracket of vector field. Call them Krichever–Novikov algebras. For the Riemann sphere ($g = 0$) with quasi-global coordinate z, $I = \{0\}$, $O = \{\infty\}$, the introduced algebra is the Witt algebra. The Witt and Virasoro algebras are graded, but these Krichever–Novikov algebras are only almost graded, as was observed by Krichever–Novikov in the two-point case [5] and generalized by Schlichenmaier [6].

We consider the genus one case, i.e., the case of one-dimensional complex tori, or, equivalently the elliptic curve case. Consider now two marked points. One marking we always put to $\infty = (0 : 1 : 0)$, and the other one to the affine coordinate $(e, 0)$. Set

$$B := \{(e_1, e_2, e_3) \in \mathbb{C}^3 \mid e_1 + e_2 + e_3 = 0,\ e_i \neq e_j \text{ for } i \neq j\}.$$

In $B \times \mathbb{P}^2$ we consider the family of elliptic curves $\mathcal{E}$ over B defined via $Y^2Z = 4(X - e_1Z)(X - e_2Z)(X - e_3Z)$. Consider the complex lines in $\mathbb{C}^2$:

$$D_s := \{(e_1, e_2) \in \mathbb{C}^2 \mid e_2 = s \cdot e_1\},\ s \in \mathbb{C}, \quad D_\infty := \{(0, e_2) \in \mathbb{C}^2\}.$$

Then B is isomorphic to $\mathbb{C}^2 \setminus (D_1 \cup D_{-\frac{1}{2}} \cup D_{-2})$.

Theorem ([7])**.** *For any elliptic curve $E_{(e_1,e_2)}$ over $(e_1, e_2) \in \mathbb{C}^2 \setminus (D_1 \cup D_{-1/2} \cup D_{-2})$ the Lie algebra $\mathcal{L}^{(e_1,e_2)}$ of vector fields on $E_{e_1,e_2)}$ has a basis $\{V_n, n \in \mathbb{Z}\}$ such that the Lie algebra structure is given as*

$$(*) \qquad [V_n, V_m] = \begin{cases} (m-n)V_{n+m}, & n, m \text{ odd},\\ (m-n)(V_{n+m} + 3e_1V_{n+m-2} & \\ \quad +(e_1 - e_2)(e_1 - e_3)V_{n+m-4}), & n, m \text{ even},\\ (m-n)V_{n+m} + (m-n-1)3e_1V_{n+m-2} & \\ \quad +(m-n-2)(e_1 - e_2)(e_1 - e_3)V_{n+m-4}, & n \text{ odd, } m \text{ even}. \end{cases}$$

These algebras make sense also for the points $(e_1, e_2) \in D_1 \cup D_{-\frac{1}{2}} \cup D_{-2}$. Altogether this defines a 2-dimensional family of Lie algebras parametrized over $\mathbb{C}^2$. In particular, for $(e_1, e_2) = 0$ we get the Witt algebra.

Now consider the family of algebras obtained by taking as base variety the line D_s (for an s). We get that for fixed s in all cases the algebras will be isomorphic above every point in D_s as long as we are not above $(0, 0)$.

Proposition. *For $(e_1, e_2) \neq (0, 0)$ the algebras $\mathcal{L}^{(e_1,e_2)}$ are* not *isomorphic to $\mathcal{W}$.*

In particular, we obtain a family of algebras over the base D_s, which is always the affine line. In this family, the algebra over the point $(0, 0)$ is the Witt algebra and the isomorphy type above all other points will be the same but different from this special Witt element. We obtain the following

Theorem. *For every $s \in \mathbb{C} \cup \{\infty\}$ the families of Lie algebras defined by $(*)$ define global deformations $\mathcal{W}_t^{(s)}$ of $\mathcal{W}$ over the affine line $\mathbb{C}[t]$. Here t corresponds to the parameter e_1 and e_2 respectively. The Lie algebra over $t = 0$ corresponds* *always to the Witt algebra, the algebras above $t \neq 0$ belong (if s is fixed) to the same isomorphy class, but are not isomorphic to $\mathcal{W}$.*

Remark. It is easy to incorporate a central term defined by a local cocycle and easy to show that the centrally extended algebras have the same properties.

References

[1] Fialowski, A.: An example of formal deformations of Lie algebras. In: *Proceedings of NATO Conference on Deformation Theory of Algebras and Applications, Il Ciocco, Italy, 1986*, pp. 375–401, Kluwer, Dordrecht, 1988

[2] Fialowski, A.: Deformations of some infinite-dimensional Lie algebras. J. Math. Phys. **31**, 1340–1343 (1990)

[3] Fialowski, A. and Fuchs, D.: Construction of miniversal deformations of Lie algebras. J. Funct. Anal. **161**, 76–110 (1999)

[4] Fialowski, A. and Schlichenmaier, M.: Global deformations of the Witt algebras of Krichever–Novikov type. Comm. in Contemp. Math. **5**, 921–945 (2003)

[5] Krichever, I. M. and Novikov, S. P.: Algebras of Virasoro type, Riemann surfaces and structures of the theory of solitons, Funktional Anal. i Prilozhen. **21**, 46–63 (1987)

[6] Schlichenmaier, M.: Krichever–Novikov algebras for more than two points. Lett. Math. Phys. **19**, 151–165 (1990)

[7] Schlichenmaier, M.: Degenerations of generalized Krichever–Novikov algebras on tori. Jour. Math. Phys. **34**, 3809–3824 (1993)

Direct limits of Lie groups

Helge Glöckner

1. Existing methods. Let $G_1 \subseteq G_2 \subseteq \cdots$ be an ascending sequence of finite-dimensional real Lie groups, such that the inclusion maps are smooth homomorphisms. Then $G := \bigcup_{n\in\mathbb{N}} G_n$ is a group in a natural way, and it becomes a topological group when equipped with the final topology with respect to the inclusion maps $G_n \to G$, the so-called *DL-topology* ([1], [11]). Provided certain technical conditions are satisfied (ensuring in particular that $\exp_G := \varinjlim \exp_{G_n} : \varinjlim L(G_n) \to \varinjlim G_n = G$ is a local homeomorphism at 0), the map $\exp_G$ and its translates can be used as charts which make G a (usually infinite-dimensional) Lie group (see [9] and subsequent work by the same authors). It is also known that every Lie subalgebra of $\mathfrak{gl}_\infty(\mathbb{R}) := \varinjlim \mathfrak{gl}_n(\mathbb{R})$ integrates to a subgroup of $\mathrm{GL}_\infty(\mathbb{R}) := \varinjlim \mathrm{GL}_n(\mathbb{R})$ [6]; this facilitates an alternative construction of a Lie group structure on various direct limits of linear Lie groups. However, neither of these methods is general enough to tackle arbitrary direct limits of Lie groups. In particular, examples show that $\exp_G$ need not be injective on any 0-neighbourhood [1], whence a general construction of a Lie group structure on $G = \bigcup_n G_n$ cannot make use of $\exp_G$.

2. A new construction principle. In [1], a Lie group structure on $G = \bigcup_n G_n$ was constructed in the case where the inclusion maps are embeddings (strict direct systems). Later, the strictness condition could be removed [2]. In [2], direct limits of Lie groups are discussed as special cases of direct limits of direct sequences

$M_1 \subseteq M_2 \subseteq \cdots$ of finite-dimensional smooth manifolds and injective immersions. To make $M := \bigcup_n M_n$ a smooth manifold, the idea is to start with a chart ϕ_1 of some M_n (say M_1) and then to use tubular neighbourhoods to extend ϕ_n already defined (or its restriction to a slightly smaller open set) to a chart of M_{n+1}. Then $\varinjlim \phi_n$ is a chart for M. It can be shown that M is smoothly paracompact [2]. Furthermore (see [2]), the direct limit groups G are regular Lie groups in the sense of convenient differential calculus [6] (this is easy) and also regular Lie groups in Milnor's sense [8] (this is much harder to prove). If all manifolds (or Lie groups) and all bonding maps are real or complex analytic, then the direct limit manifolds constructed in [2] are real analytic in the sense of convenient differential calculus, resp., complex analytic.

3. Lie theory for direct limit groups. Despite the fact that $\exp_G$ need not be well-behaved, all of the basic constructions of finite-dimensional Lie theory can be pushed to the case of direct limit groups $G = \bigcup_n G_n$. Thus, subgroups and Hausdorff quotient groups of G are Lie groups, a universal complexification $G_{\mathbb{C}}$ exists, subalgebras of $L(G)$ integrate to analytic subgroups, and Lie algebra homomorphisms integrate to smooth homomorphisms in the expected way. Furthermore, every locally finite real or complex Lie algebra of countable dimension is enlargible, i.e., it is the Lie algebra of a regular Lie group [2]. Such Lie algebras have been studied by Bahturin, Baranov, Benkart, Dimitrov, Neeb, Penkov, Strade, Stumme, and Zalesskii. If $H \subseteq G$ is a closed subgroup, then H is a smooth submanifold of G, and in fact a conveniently real analytic (c^ω-) submanifold, under mild additional conditions [2]. Furthermore, the homogeneous space G/H can be given a c^ω-manifold structure which makes $\pi\colon G \to G/H$ a smooth principal bundle (and a c^ω-principal bundle under additional conditions), [2]. Similar results are available for complex Lie groups [2]. Special cases of complexifications and homogeneous spaces have already been used in [10], in the context of a Bott-Borel-Weil theorem for direct limit groups.

4. Direct limits of infinite-dimensional Lie groups. The situation becomes more complicated if the G_n's are infinite-dimensional Lie groups. Let us assume that a direct limit $\phi := \varinjlim \phi_n$ of compatible charts is defined on some open (or c^∞-open) subset of the locally convex direct limit $\varinjlim L(G_n)$. Provided $\varinjlim L(G_n)$ is regular (viz. it is Hausdorff, and each bounded subset is contained and bounded in some $L(G_n)$), then it is straightforward to make $G = \bigcup_n G_n$ a (possibly not smoothly Hausdorff) Lie group in the sense of convenient differential calculus [4], whose group multiplication however need not be continuous (cf. [11]). All Lie groups of relevance are Lie groups in a stronger sense (as in Milnor [8]), based on a notion of smooth maps which are, in particular, continuous (Keller's C^∞_c-maps). Pathological examples show that, even if ϕ is a global chart, it need not make $G = \bigcup_n G_n$ a Milnor-Lie group, [4]. But what happens for the examples encountered in practice?

5. Discussion of the main examples. Given a σ-compact smooth manifold M of finite dimension, the group $\mathrm{Diff}_c(M)$ of compactly supported smooth diffeomorphisms of M is a Lie group in Milnor's sense (see [7] or [5], where also regularity of $\mathrm{Diff}_c(M)$ in Milnor's sense is proved in detail). It is a union $\mathrm{Diff}_c(M) = \bigcup_K \mathrm{Diff}_K(M)$ of the Fréchet-Lie groups $\mathrm{Diff}_K(M)$ of diffeomorphisms supported in a given compact subset $K \subseteq M$. Because the DL-topology does not make $\mathrm{Diff}_c(M)$ a topological group [11], the DL-topology is strictly finer than the topology on the Lie group $\mathrm{Diff}_c(M)$. Hence, there exists a discontinuous map on $\mathrm{Diff}_c(M)$ which is continuous on $\mathrm{Diff}_K(M)$ for each K. There even exists a discontinuous map on $\mathrm{Diff}_K(M)$ which is smooth on each $\mathrm{Diff}_K(M)$, whence $\mathrm{Diff}_c(M) \neq \varinjlim \mathrm{Diff}_K(M)$ as a smooth manifold [4]. However, homomorphisms on $\mathrm{Diff}_c(M)$ are smooth (resp., continuous) if and only if they are so on each $\mathrm{Diff}_K(M)$, [4]. The situation is similar for test function groups $C^\infty_c(M,G)$ with values in a Lie group G. Thus $\mathrm{Diff}_c(M) = \varinjlim \mathrm{Diff}_K(M)$ and $C^\infty_c(M,G) = \varinjlim C^\infty_K(M,G)$ holds or does not hold, in the following categories (see [4]):

	$C^\infty_c(M,G)$	$\mathrm{Diff}_c(M)$
Lie groups	yes	yes
topological groups	yes	yes
smooth manifolds	no	no
topological spaces	no	no

6. Smooth homomorphisms vs. continuous homomorphisms. The continuity and smoothness questions just analyzed are related to the general (open) problem (due to Milnor) whether every continuous homomorphism between infinite-dimensional Lie groups is smooth. Some progress concerning this problem has been made recently: Every *Hölder* continuous homomorphism between Milnor-Lie groups is smooth [3], and Lip^0-homomorphisms between Lie groups in the sense of convenient differential calculus are smooth in the convenient sense (the author, work in progress).

References

[1] Glöckner, H., *Direct limit Lie groups and manifolds*, J. Math. Kyoto Univ. **43** (2003), 1–26.
[2] ——, *Fundamentals of direct limit Lie theory*, preprint, arXiv:math, March 2004.
[3] ——, *Hölder continuous homomorphisms between infinite-dimensional Lie groups are smooth*, manuscript, will be posted in arXiv:math in March 2004.
[4] ——, *Direct limits of Lie groups compared to direct limits in related categories*, in preparation.
[5] ——, *Patched locally convex spaces, almost local mappings, and the diffeomorphism groups of non-compact manifolds*, in preparation.
[6] Kriegl, A. and P. W. Michor, "The Convenient Setting of Global Analysis," Math. Surveys and Monographs **53**, AMS, Providence, 1997.
[7] Michor, P. W., "Manifolds of Differentiable Mappings," Shiva, 1980.
[8] Milnor, J., *Remarks on infinite dimensional Lie groups*, in: B. DeWitt and R. Stora (eds.), Relativity, Groups and Topology II, North-Holland, 1983.
[9] Natarajan, L., E. Rodríguez-Carrington and J. A. Wolf, *Differentiable structure for direct limit groups*, Lett. Math. Phys. **23** (1991), 99–109.
[10] ——, *The Bott-Borel-Weil Theorem for direct limit groups*, Trans. Amer. Math. Soc. **353** (2001), 4583–4622.

[11] Tatsuuma, N., H. Shimomura and T. Hirai, *On group topologies and unitary representations of inductive limits of topological groups and the case of the group of diffeomorphisms*, J. Math. Kyoto Univ. **38** (1998), 551–578.

Flag manifolds and cycles

Gregor Fels

Let G be a complex semisimple Lie group and $Q \subset G$ a parabolic subgroup. Let $S \subset G$ be a (connected) real form of G. Let $\mathfrak{s} \subset \mathfrak{g} = \mathfrak{s}^{\mathbb{C}}$ denote the corresponding Lie algebras. Fix a Cartan decomposition $\mathfrak{s} = \mathfrak{k} \oplus \mathfrak{p}$ and let $\mathfrak{g} = \mathfrak{k}^{\mathbb{C}} \oplus \mathfrak{p}^{\mathbb{C}}$ be its complexification. Finally, let $K \subset S$ denote the corresponding maximal compact subgroup and let $K^{\mathbb{C}} \subset G$ be its complexification. In order to avoid some awkward case by case distinctions we assume that G is simple. All the result below can be easily generalized for semisimple G.

Let $X := G/Q$ be a flag manifold. The orbit structure of the canonical action $S \times X \to X$ by left translations is well understood, see [Wo1]. Since there are only finitely many Sfrm-e–orbits in X we conclude that open orbits exist. Any open Sfrm-e–orbit is called a flag domain.
Every flag domains $D = S{\cdot}x$ contains a unique compact $K^{\mathbb{C}}$frm-e–orbit C_D. Such orbit has the property $C_D = K^{\mathbb{C}}{\cdot}x = K{\cdot}x$. This is a special case of a more general fact: There is a natural duality between the Sfrm-e–orbits and the $K^{\mathbb{C}}$frm-e–orbits in X, and an Sfrm-e–orbit $\mathbf{s}$ and an $K^{\mathbb{C}}$frm-e-orbit $\mathbf{k}$ are said to be dual if the intersection $\mathbf{s} \cap \mathbf{k}$ is a single Kfrm-e–orbit. see [Mat], [MUV], [BrLo].
Every compact $K^{\mathbb{C}}$frm-e–orbit C_D defines a cycle $1{\cdot}C_D$ in X i.e., a point in the Barlet cycle space. The Barlet cycle space $\mathfrak{C}(X)$ provides a universal family parameterizing all cycles in X. The construction of the Barlet space $\mathfrak{C}(Z)$ can be given for an arbitrary complex space Z, see [Bar] for the details. From the point of view of group actions, a natural family can be defined as follows ([WeWo]). For a given compact $K^{\mathbb{C}}$frm-e–orbit $C = C_D$, consider $\widetilde{\mathcal{M}}_D := \{g \in G \mid gC \subset D\}$.
Notice that the stabilizer $G_{[C]}$ of C acts freely and properly on the right on this set, and that the quotient

$$\mathcal{M}_D := (\widetilde{\mathcal{M}}_D/G_{[C]})^{\circ}$$

can be identified with a domain in the complex homogenous space $G/G_{[C]}$. Observe that this space parameterizes the (connected component) of the family of submanifolds of D which are obtained by moving the base manifold C by elements of $g \in G$ such that $g(C) \subset D$. We refer to such $\mathcal{M}_D$ as the Wolf parameter space.
The analysis of the quotient $G/G_{[C]}$ shows that the following cases occur:

• $G/G_{[C]} = \{\mathrm{pt}\}$ in the rare case when a non-compact real form S acts transitively on X
• $G/G_{[C]}$ is a compact Hermitian space Y. This happens only if S is of Hermitian type and certain cycles $C_D \subset G/Q$

• $G/G_{[C]}$ is the affine symmetric space $G/N_G(\mathfrak{k}^{\mathbb{C}})$.

Our first main result is the description of the Wolf parameter spaces $\mathcal{M}_D$ for all S, X and the corresponding flag domains D. In the particular case when S is of Hermitian type, the structure of $\mathcal{M}_D$ was determined in [WZ]: In this case $\mathcal{M}_D \cong \Delta$ or $\mathcal{M}_D \cong \Delta \times \overline{\Delta}$ where Δ denotes the bounded symmetric domain such that $\mathrm{Aut}^\circ(\Delta) = S/Z(S)$.

We deal only with the case where S is not of Hermitian case. Let $H := N_G(\mathfrak{k}^{\mathbb{C}}) = G_{[C_D]}$. It turns out that

Theorem 1. *Let a (non-Hermitian) real form S be fixed. For arbitrary X and flag domain $D \subset X$ all domains $\mathcal{M}_{D_C} \subset G/H$ coincide.* The domains $\mathcal{M}_{D_C}$ can

be also described in a more explicite way: Fix a maximal Abelian subalgebra $\mathfrak{a} \subset \mathfrak{p}$

and an Iwasawa decomposition $\mathfrak{s} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$. Fix a Borel subgroup $B \subset G$ such that $\mathfrak{b} \supset \mathfrak{a} \oplus \mathfrak{n}$. It should be noted that $B\cdot[H]$ is open in G/H and its complement consists of $\dim \mathfrak{a}$ irreducible Bfrm-e–stable hypersurfaces: $G/H \setminus B\cdot[H] = \mathcal{H}_1 \cup \cdots \cup \mathcal{H}_r$. For any Bfrm-e–stable hypersurface $\mathcal{H}$ define the set

$$\Omega_{\mathcal{H}} := \big(G/H \setminus \bigcup_{s \in S} s\mathcal{H}\big)^\circ = \big(G/H \setminus \bigcup_{k \in K} k\mathcal{H}\big)^\circ$$

where $(\cdots)^\circ$ denote the connected component containing $[H]$. This set is open and is called the <u>hypersurface domain</u>, associated with $\mathcal{H}$.

Theorem 2. *For an arbitrary but fixed (non-Hermitian) S, any flag domain $D \subset X$ and arbitrary Bfrm-e–stable divisor $\mathcal{H} \subset G/H$ we have*

$$\mathcal{M}_D = \Omega_{\mathcal{H}} = S\cdot \exp i\omega_{AG}\cdot[H]$$

where $\omega_{AG} := \{X \in \mathfrak{a} \mid |\lambda(X)| < \pi/2$ for all $\lambda \in \Phi(\mathfrak{a})\}$. Here, $\Phi(\mathfrak{a})$ denotes the restricted root system of $\mathfrak{s}$ with respect to $\mathfrak{a}$.
All above domains are Kobayashi hyperbolic.

See [FeHu], [HuWo].

Remark. The open set $S\cdot \exp i\omega_{AG}\cdot[H] \subset G/H$ is also called the Akhiezer-Gindikin domain, see [AG]. Note that H is disconnected and $K^{\mathbb{C}} = H^\circ$.
The covering map $G/K^{\mathbb{C}} \to G/H$ maps biholomorphically $S\cdot \exp i\omega_{AG}\cdot[K^{\mathbb{C}}]$ onto $S\cdot \exp i\omega_{AG}\cdot[H]$. Furthermore, being interested in local properties of various cycle spaces, we do not need to distinguish between $H = G_{[C]}$ and $K^{\mathbb{C}} = H^\circ$.

As already mentioned, instead of moving the reference cycles C_D by elements of a given transformation group G one can also consider the universal family of cycles, i.e., the component of the Barlet cycle space $\mathfrak{C}(X)$ which contains C_D. Such a universal family depends only on the complex geometry of the ambient space and the embedding $C_D \hookrightarrow X$. A compact $K^{\mathbb{C}}$frm-e–orbit C can be now considered as a point $[C] \in \mathfrak{C}(D) = \mathfrak{C}$, and $\mathcal{M}_D$ is just a subset of $\mathfrak{C}$. Therefore one naturally asks if $\mathcal{M}_D = \mathfrak{C}(D)$ or if $\mathcal{M}_D$ is at least open in $\mathfrak{C}$.

In order to “see” cycles in the proximity of the given C it is necessary to compute the full Zariski tangent space $T_{[C]}\mathfrak{C}$ at a point $[C]$. In general, the Barlet spaces $\mathfrak{C}$ are singular and in principle, the point $[C] \in \mathfrak{C}$ might be singular. Note that we have the canonical subspace $T_{[C]}(G{\cdot}[C])$ of $T_{[C]}\mathfrak{C}$, consisting of vectors tangent to the orbit $G{\cdot}[C]$.

Our first result here is that for *certain* real forms $S \subset G$ the tangent spaces to $\mathfrak{C}$ computed at all compact $K^{\mathbb{C}}$frm-e–orbits C and for all G-flags X the spaces $T_o(G{\cdot}[C])$ and $T_{[C]}\mathfrak{C}$ coincide. In particular, $\mathcal{M}_D$ is open in $\mathfrak{C}(X)$.

On the other hand, there are real forms S and flag manifolds X in which there are situations which do not arise by moving the base cycle by elements of $\mathrm{Aut}(X)$: There exist real forms and compact orbits $C \subset X$ (we give a precise list below) such that $\dim T_{[C]}\mathfrak{C} > \dim G/G_{[C]}$. In such a case we compute in detail the isotropy representation $K^{\mathbb{C}} \times T_{[C]}\mathfrak{C} \to T_{[C]}\mathfrak{C}$. It is actually quite difficult to obtain precise quantitative results of this type, and a substantial part of our work consists in developing effective methods for computing certain cohomology groups which are necessary for our purposes.

The calculations are carried out mostly for the full flag $X = G/B$. It should be noted that in this case $\mathrm{Aut}^{\circ}(X) = G/Z(G)$.

Theorem 3. *In all cases the Barlet space $\mathfrak{C}(G/B)$ is smooth at C_D.*

Note that for G simple and $S \subset G$ a real form, all $\mathfrak{k}^{\mathbb{C}}$frm-e–modules $\mathfrak{p}^{\mathbb{C}}$ in the complexified Cartan decomposition of $\mathfrak{s}$ are irreducible if S is not of Hermitian type and sum of two irreducible submodules $\mathfrak{p}^{\mathbb{C}} = (\mathfrak{p}^{\mathbb{C}})^{+} \oplus (\mathfrak{p}^{\mathbb{C}})^{-}$ if S is of Hermitian type. Further, for every complex group H of the classical type A–D let $H \hookrightarrow \mathrm{GL}(V^{H}_{\mathrm{std}})$ denote the standard representation. It turns out that the isotropy groups in all cases listed below are of classical type.

Theorem 4.

i) *For all real forms $\mathfrak{s}$ listed below, there exist compact $K^{\mathbb{C}}$frm-e–orbits $C \subset G/B = X$, such that the Zariski tangent space $T_{[C]}\mathfrak{C}(X)$ is bigger than $T_{[C]}(G/G_{[C]})$. $T_{[C]}\mathfrak{C} = T_{[C]}(G{\cdot}[C])$. The real forms listed below are also the only ones with this property:*
   (1) $\mathfrak{so}(2p, 2q+1)$ *for* $p \geq 2$,
   (2) $\mathfrak{so}(2p+1, 2q+1)$, *for* $p, q \geq 1$
   (3) $\mathfrak{sp}_n(\mathbb{R})$ *for* $n \geq 3$,
   (4) G_2
   (5) $\mathfrak{sl}_3(\mathbb{R})$.

ii) *At the same time, for all real forms except $\mathfrak{sl}_3(\mathbb{R})$ there exist compact $K^{\mathbb{C}}$frm-e–orbits $C' \subset G/B$, such that $T_{[C']}\mathfrak{C}(X) = T_{[C']}(G{\cdot}[C'])$.*

iii) *For those compact $K^{\mathbb{C}}$frm-e–orbits $C \subset G/B$ with the property as in i) the tangent space $T_{[C]}\mathfrak{C}(X)$ has the following decomposition as a $K^{\mathbb{C}}$frm-e–module:*

$\mathfrak{s} = \mathfrak{so}(2p, 2q+1)$	$T_{[C]}\mathfrak{C} = \mathfrak{p}^{\mathbb{C}} \oplus V_{\mathrm{std}}^{\mathrm{SO}_{2p}}$
$\mathfrak{s} = \mathfrak{so}(2p+1, 2q+1)$	$T_{[C]}\mathfrak{C} = \mathfrak{p}^{\mathbb{C}} \oplus V_{\mathrm{std}}^{\mathrm{SO}_{2p+1}}$ *or* $T_{[C]}\mathfrak{C} = \mathfrak{p}^{\mathbb{C}} \oplus V_{\mathrm{std}}^{\mathrm{SO}_{2q+1}}$
$\mathfrak{s} = \mathfrak{sp}_n(\mathbb{R})$	$T_{[C]}\mathfrak{C} = (\mathfrak{p}^{\mathbb{C}})^+ \oplus \bigwedge^2 V_{\mathrm{std}}^{\mathrm{GL}_n}$ *or* $T_{[C]}\mathfrak{C} = (\mathfrak{p}^{\mathbb{C}})^- \oplus \bigwedge^2 (V_{\mathrm{std}}^{\mathrm{GL}_n})^*$
$\mathfrak{s} = \mathbf{G}_2$	$T_{[C]}\mathfrak{C} = \mathfrak{p}^{\mathbb{C}} \oplus V_{\mathrm{std}}^{\mathrm{SO}_4}$
$\mathfrak{s} = \mathfrak{sl}_3(\mathbb{R})$	$T_{[C]}\mathfrak{C} = \mathfrak{p}^{\mathbb{C}} \oplus V_{\mathrm{std}}^{\mathrm{SO}_3}$

See [Fe] for the proofs and further details.

References

[1] [AG] D. N. Akhiezer, S. G. Gindikin: On Stein extensions of real symmetric spaces. Math. Ann. 286 (1990), no. 1-3, 1–12.

[2] D. Barlet: Espace analytique réduit des cycles analytiques complexes compacts d'un espace analytique complexe de dimension finie. Fonctions de plusieurs variables complexes, II (Sm. Franois Norguet, 1974–1975), pp. 1–158. Lecture Notes in Math., Vol. 482, Springer, Berlin, 1975.

[3] R. Bremigan, J. Lorch: Orbit duality for flag manifolds. Manuscripta Math. **109** (2002), no. 2, 233–261.

[4] G. Fels: On complex analytic cycle spaces of flag domain. Habilitationsschrift 2004.

[5] G. Fels, A. T. Huckleberry: Characterization of cycle domains via Kobayashi hyperbolicity. Erscheint 2004 in Bulletin de la Société Mathématique de France.

[6] A. T. Huckleberry, J. A. Wolf: Schubert varieties and cycle spaces. math.AG/0204033

[7] T. Matsuki: Orbits on affine symmetric spaces under the action of parabolic subgroups. Hiroshima Math. J. **12** (1982), no. 2, 307–320.

[8] I. Mirković, T. Uzawa, K. Vilonen: Matsuki correspondence for sheaves. Invent. Math. **109** (1992), no. 2, 231–245.

[9] R. O. Wells, Jr.; J. A. Wolf: Poincar series and automorphic cohomology on flag domains. Ann. of Math. (2) **105** (1977), no. 3, 397–448.

[10] J. A. Wolf: The action of a real semisimple group on a complex flag manifold. I. Orbit structure and holomorphic arc components. Bull. Amer. Math. Soc., **75** (1969) 1121–1237.

[11] J. A. Wolf: The Stein condition for cycle spaces of open orbits on complex flag manifolds. Ann. of Math. (2) **136** (1992), no. 3, 541–555.

[12] J. A. Wolf, R. Zierau: Linear cycle spaces in flag domains. Math. Ann. **316** (2000), no. 3, 529–545.

Berezin transform on root systems of type BC

Genkai Zhang

In the present talk we present our recent result on Berezin transform on root systems with general multiplicities. The Berezn transform on symmetric domains arises when one studies the branching of holomorphic representation on a Hermitian symmetric space G/K of a semisimple Lie group G under a symmetric subgroup H with the corresponding symmetric space H/L being a real form of G/K. More precisely, consider the restriction map R of a scalar holomorphic discrete series $\mathcal{H}_\nu$ (and its analytic continuation) realized as a Hilbert space of holomorphic functions on G/K to the real form H/L. The map $B_\nu = RR^*$ on $L^2(H/L)$ is then the Berezin transform. It is H-invariant, and is bounded on $L^2(H/L)$ for larger parameter of ν. The spectral symbol of B_ν has been computed by Unterberger-Upmeier [3], Zhang [6] [5] van Dijk and Pevsner [1] and have found several applications [4]. In the present work we consider a general root system of type BC with general positive multiplicity. The Berezin transform can be defined as an integral operator whose kernel is defined by a series. We find the spectral symbol of the Berezin transform and find a Bernstein-Sato type formula for the integral kernel. The precise results are summarized below.

Let $\mathfrak{a} = \mathbb{R}^r$ be an Euclidean space with inner product $(\cdot,\cdot)$ and let $R \subset \mathfrak{a}^*$ be a root system of type BC. We fix an orthogonal basis γ_j, $j = 1, \dots, r$ of $\mathfrak{a}^*$ so that $R = \{\frac{\gamma_j}{2}; j = 1, \cdots, r\} \cup \{\gamma_j; j = 1, \cdots, r\} \cup \{\frac{\gamma_j \pm \gamma_k}{2}; j \neq k = 1, \cdots, r\}$ and let $k = (k_1, k_2, k_3)$ be the root multiplicity with k_1, k_2 and k_3 the multiplicities of the respective subsets of R. We assume that $k_1 + k_2 > 0$ and $k_3 > 0$.

Let $\{\xi_j\}$ be the basis of $\mathfrak{a}$ dual to $\frac{\gamma_j}{2}$, $j = 1, \dots, r$, i.e., $\frac{\gamma_j}{2}(\xi_k) = \delta_{jk}$. A function $f(x)$ on $\mathfrak{a}^{\mathbb{C}}$ will be identified with $f(x_1, \cdots, x_r)$, for $x = x_1\xi_1 + \cdots + x_r\xi_r$. Let $D_j = D_{\xi_j}$ be be the Cherednik operators and let ϕ_λ be Heckman-Opdam the spherical function. Consider the function

$$f_\nu(x) = \prod_{j=1}^{r} \cosh(x_j)^{-2\nu}.$$

The integral kernel $B(x, y)$ of the Berezin transform is given by $B(x, 0) = f_\nu(x)$ and by an infinite series with using the Jack symmetric polynomials. Its spectral symbol is determined by the integral

$$b_\nu(\lambda) = \widetilde{f_\nu}(\lambda) = \int_{\mathfrak{a}} f_\nu(x)\phi_\lambda(x)d\mu_\Sigma(x)$$

where $d\mu_\Sigma(x)$ is the invariant measure for the root system Σ (which corresponds to the radial A-part of the Riemannian measure in the case of symmetric space $H/L = LA \cdot 0$).

We prove first a Bernstein-Sato type formula using the Cherednik operator.

Theorem 1. *(Bernstein-Sato type formula) The following formula holds*

$$\prod_{j=1}^{r}(D_j^2 - (-\nu/2 + \rho_1)^2))f_\nu = \prod_{j=1}^{r}(-\nu/2 + k_3(j-1))(1 + \nu/2 - k_2 - k_3(r-j))f_{\nu+1}$$

In proving the theorem we also find some interesting commutation relation for the Hecke algebra elements and multiplication operators by polynomials of e_j^x.

We can then find the spectral symbol.

Theorem 2. *The spherical transform of f_ν is given by*

$$b_\nu(\lambda) = c_\delta \prod_{j=1}^{r} \prod_{\varepsilon=\pm 1} \Gamma(\nu - \frac{p-1}{2} + \varepsilon\lambda_j)$$

The result has also some applications to orthogonal polynomials, the details will appear later.

References

[1] G. van Dijk and M. Pevzner, *Berezin kernels of tube domains.* J. Funct. Anal. **181** (2001), no. 2, 189–208.
[2] E. Opdam, *Harmonic analysis for certain representations of graded Hecke algebras*, Acta Math. **175** (1995), no. 1, 75–121.
[3] A. Unterberger and H. Upmeier, *The Berezin transform and invariant differential operators*, Comm. Math. Phys. **164** (1994), 563–597.
[4] G. Zhang, *Branching coefficients of holomorphic representations and Segal-Bargmann transform.* J. Funct. Anal. **195** (2002), no. 2, 306–349.
[5] G. Zhang, *Berezin transform on real bounded symmetric domains.* Trans. Amer. Math. Soc. **353** (2001), no. 9, 3769–3787
[6] G. Zhang, *Berezin transform on line bundles over bounded symmetric domains.* J. Lie Theory 10 (2000), no. 1, 111-126.

General Differential Calculus and General Lie Theory

Wolfgang Bertram

In joint work with H. Glöckner and K.-H. Neeb [1], a simple and at the same time very general approach to differential calculus, manifolds and Lie groups is proposed which not only works in arbitrary dimension over the real and complex numbers, but more generally for arbitrary topological modules over (commutative) base rings k having a dense group of invertible elements (in particular, over all non-discrete topological fields). All notions and results from differential geometry and Lie theory that are essentially algebraic in nature continue to make sense in this general framework – one may call these parts of the theory "general differential geometry" and "general Lie theory".

In our talk we present a basic result of this theory which in a way provides a rigorous justification of the use of "infinitesimals" in differential geometry (cf. [3]): if M is a manifold over k, then the tangent bundle TM is, in a natural way, a manifold over the ring of dual numbers $k[\epsilon] = k \oplus \epsilon k \cong k[x]/(x^2)$ (with relation $\epsilon^2 = 0$), and tangent maps are smooth over $k[\epsilon]$; thus the tangent functor really is a functor of scalar extension from k to dual numbers over k. It immediately follows that the iterated tangent bundles $T^n M$ are manifolds over the ring $T^n k := k[\epsilon_1, \ldots, \epsilon_n]$ and that the "jet bundles" $J^n M = (T^n M)^{\Sigma_n}$ (the subbundle fixed under the canonical action of the permutation group Σ_n on $T^n M$) are manifolds over the ring $J^n k := (T^n k)^{\Sigma_n}$. Likewise, if G is a Lie group over k, then $T^n G$ is a Lie group over $T^n k$ and $J^n G$ is a Lie group over $J^n k$. Another approach to infinitesimals has been proposed by A. Weil in 1953 and lead to various concepts such as the "Weil-functors" defined in the book "Natural Operations in Differential Geometry" by I. Kolář, P. Michor and J. Soválk (Springer-Verlag 1993) or the theory of "smooth toposes" and "synthetic differential geometry" (see the book

"Models for Smooth Infinitesimal Analysis" by I. Moerdijk and G. Reyes, Springer-Verlag 1991); our result may be seen as an alternative and much simpler approach to these objects.

Finally, we give a short overview over problems and further topics in the context of general Lie theory. In general, it is not possible to integrate differential equations in our general context (this is known to be so already in the p-adic case or in the locally convex real case), and so most problems take the form of "integration problems". For instance, for a general Lie group over k, there is no exponential map, but pushing the theory of connections somewhat further than usual one can define a certain bundle isomorphism on the level of higher order tangent bundles which serves to replace the missing exponential map (work in progress, cf. [3]). Then one may ask whether there is also an analog of the Campbell-Hausdorff formula. This seems to be indeed the case, but the precise form of this formula is unknown at present (note that the characteristic of k is arbitrary). The ultimate integration problem in Lie theory would be to find an analog of "Lie's third theorem" in our general context, i.e. to find necessary and sufficient conditions for a Lie algebra to be "integrable" to a Lie group. This problem can also be posed for symmetric spaces and Lie triple systems. Remarkably enough, for *Jordan algebraic structures* the integration problem can be solved (cf. [2], [4]): there is a functor assigning to every Jordan-structure over k (-algebra, -triple system or -pair, satisfying some natural continuity condition) a geometry which is smooth over k. This is possible since "Jordan geometries" tend to be algebraic, whereas "Lie geometries" only tend to be analytic.

References

(all available on my homepage: http://www.iecn.u-nancy.fr/˜bertram/)

[1] Differential Calculus, manifolds and Lie groups over arbitrary infinite fields (with H. Gloeckner and K.-H. Neeb), Expo. Math., to appear, see also arXiv math.GM/0303300

[2] Projective completions of Jordan pairs. Part I: The generalized projective geometry of a Lie algebra (with K.-H. Neeb), J. of Algebra, to appear, see also arXiv math.RA/0306272

[3] Differential Geometry over General Base Fields and Rings. Part I: First and Second Order Geometry. Preprint, Nancy 2003 (Part II in preparation).

[4] Projective completions of Jordan pairs. Part II: Manifold structures and symmetric spaces (with K.-H. Neeb). Preprint, Nancy – Darmstadt 2004, see also arXiv math. GR/0401236

Cohomology of holomorphic vector fields on a punctured Riemann surface

Friedrich Wagemann

Let Σ denote a compact Riemann surface of genur g and $\Sigma_r = \Sigma \setminus \{p_1, \ldots, p_r\}$ a punctured Riemann surface, punctured in $r \geq 1$ distinct points.
Let $Hol(\Sigma_r)$ denote the Lie algebra of holomorphic vector fields on Σ_r. It is a topological Lie algebra with respect to the topology of uniform convergence on compact sets in Σ_r. The underlying topological space is Fréchet.
The goal of this survey is an Ext-description of the continuous cohomology of $Hol(\Sigma_r)$, i.e. to describe it in terms of (topologically split) exact sequences of $Hol(\Sigma_r)$-modules.
In a first section, we recalled the setting of continuous cohomology of a Fréchet Lie algebra $\mathfrak{g}$ [1]. The Ext-description, which is standard for ordinary cohomology by work of Yoneda, is more difficult here as there is no standard category of topological $\mathfrak{g}$-modules which posesses enough projectives and injectives. Nevertheless, $H^2(\mathfrak{g}, \mathbb{C})$ classifies central extensions which are topologically split (i.e. split as sequences of topological vector spaces). Our first theorem [6] is that the standard map associating to a (topologically split) crossed module its continuous 3-cocycle induces a bijection of the set of crossed modules of $\mathfrak{g}$ with V to $H^3(\mathfrak{g}, V)$ in case there is a topologically split exact sequence $0 \to V \to W \to U \to 0$ such that $H^3(\mathfrak{g}, W) = 0$.
In a second section, we recalled N. Kawazumi's theorem [2] on the continuous cohomology of $Hol(\Sigma_r)$. It states that $H^*(Hol(\Sigma_r), \mathbb{C})$ is isomorphic to the singular cohomology of the space $Map(\Sigma_r, S^3)$ of continuous maps from Σ_r to the 3-sphere S^3, equipped as a topological space with the compact-open topology. The latter cohomology algebra is a graded commutative Hopf algebra in N generators of degree 2 (where N equals the dimension of $H^1(\Sigma_r)$) and one generator of degree 3, a kind of Godbillon-Vey generator. We generalized Kawazumi's work to n dimensional complex manifolds [4], and showed also that one can obtain from it the continuous cohomology of the topological subspace of meromorphic vector fields [3] (i.e. those holomorphic vector fields on Σ_r which have at most poles in $p_1, \ldots, p_r$) which play an important rôle in Krichever-Novikov's approch to string theory.
In a third section, we showed in our main theorem how to construct a crossed module representing the mentioned Godbillon-Vey type generator [5]. The corresponding 4-term exact sequence is constructed by splicing together the short exact de Rham sequence of holomorphic differential forms on the universal cover $\widetilde{\Sigma_r}$ of Σ_r, say

$$0 \to \mathbb{C} \to \Omega^0(\widetilde{\Sigma_r}) \to \Omega^1(\widetilde{\Sigma_r}) \to 0,$$

and an abelian extension of $Hol(\Sigma_r)$ by $\Omega^1(\widetilde{\Sigma_r})$ by a certain 2-cocycle.

References

[1] D. B. Fuchs, Cohomology of Infinite Dimensional Lie Algebras, Consultant Bureau, New York (1987)

[2] Nariya Kawazumi, *On the Complex Analytic Gelfand–Fuks Cohomology of open Riemann Surfaces* Ann. Inst. Fourier, Grenoble **43**, 3 (1993) 655–712

[3] Friedrich Wagemann, *Some Remarks on Krichever-Novikov Algebras*, Lett. Math. Phys. **47** (1999) 173–177

[4] Friedrich Wagemann, *Differential Graded Cohomology and Lie Algebras of Holomorphic Vector Fields*, Commun. Math. Phys. **208** (1999) 521–540
[5] F. Wagemann, A crossed module representing the Godbillon-Vey cocycle, Lett. Math. Phys. **51** (2000) 293–299
[6] F. Wagemann, On Lie algebra crossed modules, preprint.

On the holomorphic structure of G-orbits in compact hermitian symmetric spaces

Wilhelm Kaup

In this lecture we give a survey on the results of the submitted paper [4]. Let us start with a complex Banach space E of dimension n (that is $\mathbb{C}^n$ with a fixed norm $\| \, \|$). The open unit ball $D \subset E$ is called a *bounded symmetric domain* if the group $G := \mathrm{Aut}(D)$ of all biholomorphic automorphisms of D acts transitively on D (this is not an essential restriction to the usual more abstract definition, see e.g. [2]). Then it is well known that G is a semi-simple Lie group and that the isotropy subgroup $K\colon = \{g \in G : g(0) = 0\}$ is a maximal compact subgroup coinciding with the group of all linear isometries of the complex Banach space E. The compact dual Z of D in the sense of symmetric hermitian spaces is a compact homogeneous complex manifold containing E as open dense subset in such a way that $G \cong \{g \in \mathrm{Aut}(Z) : g(D) = D\}$ (Z is the Riemann sphere in case $E = \mathbb{C}$ and D the open unit disk). In this sense G also acts on Z by biholomorphic transformations and has only finitely many orbits there (one of which is the domain $D \subset Z$, another one is the Shilov boundary of D, the unique closed G-orbit in Z).

The G-orbits in Z as homogeneous spaces and the holomorphic arc components of their closures have been described explicitly in [5]. Here we are interested in the Cauchy-Riemann structure of G-orbits (which for open orbits is just the usual holomorphic structure as complex manifold). For fixed orbit $M\colon = G(a)$, $a \in Z$, let us briefly recapitulate its CR-structure (take [1] as general reference for arbitrary CR-manifolds): For every $c \in M$ the tangent space T_cM is canonically contained in the tangent space T_cZ, which is a complex vector space in a natural way. Clearly, $H_cM\colon = T_cM \cap i\,T_cM$ (called the *holomorphic tangent space at* $c \in M$) is the biggest complex subspace contained in T_cM. The CR-structure on M is given by the complex vector bundle $HM \subset TM$. In particular, a smooth function $f\colon M \to \mathbb{C}$ (or more generally with values in another CR-manifold) is called CR if it satisfies the tangential Cauchy-Riemann partial differential equations in the sense that the differential $df(c)\colon T_cM \to \mathbb{C}$ is complex linear on every holomorphic tangent space H_cM, $c \in M$. Here we are interested in the holomorphic extendibility of CR-functions, the explicit determination of CR-automorphism groups and the CR-equivalence problem for G-orbits.

For simplicity and without essential loss of generality we restrict to the case where the bounded symmetric domain D is *irreducible*, that is, not a direct product of lower dimensional bounded symmetric domains. Then, if D has rank r, there exist precisely $\binom{r+2}{2}$ G-orbits in Z, which can be indexed in a canonical way as $M_{p,q}$ with integers $p,q \geq 0$ satisfying $p+q \leq r$ (compare the special example below). There are precisely $r+1$ open orbits (those with $p+q=r$) and also $r+1$ orbits (those with $q=0$) contained in the closure $\overline{D}$ of D. In case D is of tube type, the Shilov boundary $M_{0,0}$ of D is totally real in Z, and there is a biholomorphic transformation ι of Z with period 2, mapping every $M_{p,q}$ onto $M_{q,p}$, thus giving a real-analytic CR-equivalence between $M_{p,q}$ and $M_{q,p}$. It's the existence of this transformation ι that is responsible for some extra phenomena in the tube case. For a presentation of our results therefore assume in the following that the irreducible bounded symmetric domain D is not of tube type: Then, if $M = M_{p,0}$ (that is, $M \subset \overline{D}$), every continuous CR-function f on M has a unique continuous extension $\hat{f}$ to the linear convex hull $\hat{M} = \bigcup_{k\geq p} M_{k,0}$ of M that is holomorphic on the domain $D = M_{r,0}$, and $\hat{M}$ is maximal in Z with respect to this extension property. For every other orbit $M = M_{p,q}$, $q > 0$, every continuous CR-function on M is constant and every continuous CR-function on $M \cap E$ has a unique holomorphic extension to E, implying that then every infinitesimal CR-transformation of M extends to a holomorphic vector field on Z. This can be used to show for every G-orbit M in Z that the group $\mathrm{Aut}(M)$ of all CR-automorphisms of M is just the group G and also that the G-orbits in Z are pairwise CR-inequivalent. The proofs use extensively the Jordan algebraic description of bounded symmetric domains as well as the CR-extension results for K-orbits obtained in [3].

For the announced example fix integers $s \geq r \geq 1$ in the following and denote by $E\colon = \mathbb{C}^{s\times r}$ the Banach space of all complex $s \times r$–matrices, where $\|z\|$ is the operator norm of the matrix z, considered as a linear operator $\mathbb{C}^r \to \mathbb{C}^s$. Then the open unit ball $D \subset E$ is an irreducible bounded symmetric domain of rank r, and D is of tube type if and only if $s = r$. The subgroup $\mathrm{SU}(s,r) \subset \mathrm{SL}(s+r,\mathbb{C})$ acts by linear fractional transformations transitively on D in the following way: Write every $g \in \mathrm{SU}(s,r)$ in block form $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $a \in \mathbb{C}^{s\times s}$, $b \in \mathbb{C}^{s\times r}$, $c \in \mathbb{C}^{r\times s}$, $d \in \mathbb{C}^{r\times r}$ and put $g(z) := (az+b)(cz+d)^{-1}$ for all $z \in D$. Then the connected identity component of $G = \mathrm{Aut}(D)$ consists of all transformations obtained this way. The compact dual Z of D is the Grassmann manifold $\mathbb{G}_{r,s}$ of all r-planes in $\mathbb{C}^r \times \mathbb{C}^s$, in which E is embedded by identifying every matrix $z \in E$ with its graph $\{(x, zx) : x \in \mathbb{C}^r\} \in \mathbb{G}_{r,s}$. For every $z = (z_{jk}) \in E$ let $z^*\colon = (\overline{z_{jk}}) \in \mathbb{C}^{r\times s}$ and $\mathbb{1}$ the unit matrix in $\mathbb{C}^{r\times r}$. Then, if the hermitian matrix $\mathbb{1} - z^*z \in \mathbb{C}^{r\times r}$ has type (p,q) (meaning p positive and q negative eigenvalues), we have $G(z) = M_{p,q}$. In particular,

$$D = M_{r,0} = \{z \in E : \mathbb{1} - z^*z \text{ positive definite}\}$$

and

$$M_{0,0} = \{z \in E : z^*z = \mathbb{1}\}$$

is the Shilov boundary of D. In the tube case, i.e. $r = s$, the involutory transformation ι leaves $\mathrm{GL}(r,\mathbb{C}) \subset Z$ invariant and satisfies $\iota(z) = z^{-1}$ there.

References

[1] Boggess, A.: *CR Manifolds and the Tangential Cauchy-Riemann Complex*. Studies in Advanced Mathematics. CRC Press. Boca Raton Ann Arbor Boston London 1991.

[2] Helgason, S.: *Differential Geometry and Symmetric Spaces*. Academic Press 1962.

[3] Kaup, W., Zaitsev, D.: On the CR-structure of compact group orbits associated with bounded symmetric domains. *Inventiones math.* **153**, 45-104 (2003).

[4] Kaup, W.: On the holomorphic structure of G-orbits in compact hermitian symmetric spaces. To appear. See also `http://www.mathematik.uni-tuebingen.de/~kaup/`.

[5] Wolf, A.J.: *Fine Structure of Hermitian Symmetric Spaces*. Symmetric spaces (Short Courses, Washington Univ., St. Louis, Mo., 1969–1970), pp. 271–357. Pure and App. Math., Vol. 8, Dekker, New York, 1972.

Deformation quantization of Kähler manifolds

Martin Schlichenmaier

In this talk I presented results on the Berezin-Toeplitz deformation quantization for compact quantizable Kähler manifolds. Some of them were obtained jointly with M. Bordemann and E. Meinrenken. Some of them jointly with A. Karabegov.

Let (M, ω) be a Kähler manifold and $(C^\infty(M), \cdot)$ the associative and commutative algebra of C^∞-functions under the pointwise product. This algebra is endowed with a Poisson structure via $\{f, g\} := \omega(H_f, H_g)$, with H_f the Hamiltonian vector field defined by $\omega(H_f, .) = df(.)$. A formal deformation quantization or a star product is an associative product $\star$ on the vector space of formal power series $C^\infty(M)[[\nu]]$, which is ν-adically continuous and fulfills

$$(1)\quad f \star g = f \cdot g \bmod \nu, \qquad (2)\quad \frac{1}{\nu}(f \star g - g \star f) = \mathrm{i}\{f, g\} \bmod \nu.$$

In particular,

$$f \star g = \sum_{j=0}^{\infty} \nu^j C_j(f, g),$$

with bilinear maps $C_j : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$. A star product is called a differential star product if the C_j are bidifferential operators. Usually one assumes also $f \star 1 = 1 \star f = f$. Two star products $\star$ and $\star'$ (for the same Poisson structure) are called equivalent if there is an isomorphism of algebras B (i.e. $B(f) \star' B(g) = B(f \star g)$) such that the formal sum $B = \sum_{j=0}^{\infty} \nu^j B_j$ starts with $B_0 = id$. A differential star product is called a star product with the property of "separation of variables" (in the terminology of Karabegov) or of Wick-type (in the terminology of Bordemann-Waldmann) if in the first argument of C_j only holomorphic and in the second argument only anti-holomorphic derivatives appear. In joint work with A. Karabegov I showed that the Berezin-Toeplitz (BT) deformation quantization is a differential star product with the separation of variables property [KS].

The approach presented here works for arbitrary compact and quantizable Kählermanifolds. A Kähler manifold is called quantizable if there exists a holomorphic hermitian line bundle over M: (L, h, ∇), (∇ is the compatible connection) such that $curv_{L,\nabla} = -\mathrm{i}\omega$. Important examples of such quantizable Kähler manifolds are given by the projective space with the hyperplane section bundle, projective submanifolds, abelian varieties, moduli spaces of flat $su(N)$-connections on a Riemann surface, moduli spaces of stable algebraic vector bundles of rank N and degree d over an algebraic curve, etc.

The metric h on L extends to $h^{(m)}$ on $L^m := L^{\otimes m}$. By integrating it against the Liouville form it defines a scalar product on the space of C^∞ sections. Inside the L^2 completion there is the finite-dimensional subspace $\Gamma_{hol}^{(m)}$ of holomorphic sections. Let $\Pi^{(m)}$ be the projection onto this subspace. The BT quantum operators associated to a function f on M are defined as $(T_f^{(m)})_{m\in\mathbb{N}}$ with

$$T_f^{(m)} : \Gamma_{hol}^{(m)} \to \Gamma_{hol}^{(m)}; \quad s \mapsto T_f^{(m)}(s) = \Pi^{(m)}(f \cdot s) \ .$$

Theorem 1. [BMS].
(a) $\lim_{m\to\infty} ||T_f^{(m)}|| = ||f||_{sup}$.
(b) $||m\mathrm{i}[T_f^{(m)}, T_g^{(m)}] - T_{\{f,g\}}^{(m)}|| = O(1/m)$,
(c) $||T_f^{(m)} \cdot T_g^{(m)} - T_{f\cdot g}^{(m)}|| = O(1/m)$.

Theorem 2. [Bia], [BMS], [CMF]. *There exists a unique star product* $f \star_{BT} g = \sum_{k=0}^\infty \nu^k C_k(f, g)$, *such that*

$$T_f^{(m)} \cdot T_g^{(m)} \sim \sum_{k=0}^{\infty} \left(\frac{1}{m}\right)^k T_{C_k(f,g)}^{(m)}, \quad m \to \infty.$$

This star product is called the Berezin-Toeplitz star product.

Theorem 3. [KS]. *The BT-star product is a differential star product with the separation of variables property. It has as Karabegov classifying form* $\tilde{\omega}_{BT} = -\frac{1}{\lambda}\omega + \omega_{can}$ *and as Fedosov-Deligne class* $c(\star_{BT}) = \frac{1}{\mathrm{i}}\left(\frac{1}{\lambda}[\omega] - \frac{\epsilon}{2}\right)$.

Here λ is a formal variable, ω_{can} is the curvature form of the canonical holomorphic line bundle and ϵ is the canonical class. Furthermore, it should be recalled that all star products with the separation of variables property are uniquely given by their (formal) Karabegov form, and all differential star products up to equivalence given by their (formal) Fedosov-Deligne class. As an important tool in the proof of the last theorem the Berezin transform $I^{(m)} : C^\infty(M) \to C^\infty(M)$ was used. With the help of the (suitably generalized) Berezin covariant symbol map σ it can be described as $I^{(m)}(f) = \sigma(T_f^{(m)})$. In [KS] it was shown that it has a complete asymptotic expansion in $1/m$ which starts with $f(x) + (1/m)\Delta f + \ldots$

References

[BMS] Martin Bordemann, Eckhard Meinrenken, Martin Schlichenmaier, *Toeplitz quantization of Kähler manifolds and* $gl(N)$, $N \to \infty$ *limits*, Communications in Mathematical Physics **165**(1994), 281-296.

[KS] Alexander Karabegov, Martin Schlichenmaier, *Identification of Berezin-Toeplitz deformation quantization*, J. reine angew. Math. **540**(2001), 49–76.

[CMF] Martin Schlichenmaier, *Deformation quantization of compact Kähler manifolds by Berezin-Toeplitz quantization*, Conference Moshé Flato 1999 (September 1999, Dijon, France) (eds. G. Dito, and D. Sternheimer), Kluwer, 2000, Vol. 2, pp. 289–306, math.QA/9910137.

[Bia] Martin Schlichenmaier, *Berezin-Toeplitz quantization of compact Kähler manifolds*, in: Quantization, Coherent States and Poisson Structures, Proc. XIV'th Workshop on Geometric Methods in Physics, Białowieza, Poland, 9-15 July 1995, (eds. A. Strasburger, S.T. Ali, J.-P. Antoine, J.-P. Gazeau , A. Odzijewicz), Polish Scientific Publisher PWN 1998, pp.101-115, q-alg/9601016.

The generalized Cayley map from an algebraic group to its Lie algebra
Peter W. Michor

This talk is mainly based on the paper [4].

Let $\pi : G \to \mathrm{End}(V)$ be an infinitesimally faithful complex representation of a connected Lie group G. Consider $(A, B) \mapsto \mathrm{tr}(AB)$ on $\mathrm{End}(V)$ and suppose that it is non-degenerate on the linear subspace $\pi'(\mathfrak{g}) \subseteq \mathrm{End}(V)$. Then the orthogonal projection $\mathrm{pr}_\pi : \mathrm{End}(V) \to \pi'(\mathfrak{g})$ is defined:

$$\begin{tikzcd} G \arrow[r, "\text{representation } \pi"] \arrow[d, "\Phi_\pi \mid \text{Caley map}"'] & \mathrm{End}(V) \arrow[d, "\mathrm{pr}_\pi"', "\text{orthoproj.}"] \\ \mathfrak{g} \arrow[r, tail, "\text{infinites. repr. } \pi'"] & \pi'(\mathfrak{g}) \end{tikzcd} \qquad \Psi_\pi(g) = \Psi(g) := \det(d\Phi(g))$$

The Cayley mapping Φ has the following simple properties:

(1) $\Phi(bxb^{-1}) = \mathrm{Ad}_b(\Phi(x))$.
(2) We have $\Phi(g) \in \mathrm{Cent}(\mathfrak{g}^g) \subset Z_{\mathfrak{g}}(\mathfrak{g}^g)$.
(3) $d\Phi(e) : \mathfrak{g} \to \mathfrak{g}$ is the identity mapping.
(4) $H \subset G$ be a Cartan subgroup with Cartan algebra $\mathfrak{h} \subset \mathfrak{g}$. Then $\Phi(H) \subset \mathfrak{h}$.
(5) For the character $\chi_\pi(g) = \mathrm{tr}(\pi(g))$ of π we have $d\chi_\pi(g)(T_e(\mu_g)X) = \mathrm{tr}(\pi'(\Phi_\pi(g))\pi'(X))$

Further results are:

• Let $\pi : G \to \mathrm{Aut}(V)$ be a representation admitting a Cayley mapping. Let $H = (\bigcap_{a \in A} G^a)_o = (G^A)_o \subseteq G$ be a subgroup which is the connected centralizer of a subset $A \subseteq G$ and suppose that H is itself reductive. Then $\pi|H : H \to \mathrm{End}(V)$ admits a Cayley mapping and $\Phi_\pi|H = \Phi_{\pi|H} : H \to \mathfrak{h}$.

• Let G be a semisimple real or complex Lie group, let $\pi : G \to \mathrm{Aut}(V)$ be an infinitesimally effective representation. Let $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k$ be the decomposition into the simple ideals $\mathfrak{g}_i$. Let $G_1, \ldots, G_k$ be the corresponding connected subgroups of G. Then $\Phi_\pi|G_i = \Phi_{\pi|G_i} \quad$ for $i = 1, \ldots, k$.

• G a simple Lie group, for direct sum and tensor product representations

$$\Phi_{\pi_1 \oplus \pi_2}(g) = \frac{j_{\pi_1}}{j_{\pi_1 \oplus \pi_2}} \Phi_{\pi_1}(g) + \frac{j_{\pi_2}}{j_{\pi_1 \oplus \pi_2}} \Phi_{\pi_2}(g) \in \mathfrak{g}.$$

$$\Phi_{\pi_1 \otimes \pi_2}(g) = \frac{j_{\pi_1} \chi_{\pi_2}(g)}{j_{\pi_1 \otimes \pi_2}} \Phi_{\pi_1}(g) + \frac{\chi_{\pi_1}(g) j_{\pi_2}}{j_{\pi_1 \otimes \pi_2}} \Phi_{\pi_2}(g) \in \mathfrak{g}.$$

Results for algebraic groups. Now let G be a reductive complex algebraic group and π a rational representation. We have $A(\mathfrak{g}) = A(\mathfrak{g})^G \otimes \mathrm{Harm}(\mathfrak{g})$, where $\mathrm{Harm}(\mathfrak{g})$ is the space of all regular functions killed by all invariant differential operators with constant coefficients. We define $\mathrm{Harm}_\pi(G) := \Phi_\pi^*(\mathrm{Harm}(\mathfrak{g}))$. It is a G-module.

• For the localization at Ψ we have $A(G)_\Psi = A(G)^G_\Psi \otimes \mathrm{Harm}_\pi(G)$. Moreover, we have $A(G) = A(G)^G \otimes \mathrm{Harm}_\pi(G)$ if and only if $\Phi : G \to \mathfrak{g}$ maps regular orbits in G to regular orbits in $\mathfrak{g}$.
• If $\Phi(e) = 0 \in \mathfrak{g}$ then for the G-equivariant extension of the rational function fields $\Phi^* : Q(\mathfrak{g}) \to Q(G)$ the degrees satisfy $[Q(G) : Q(\mathfrak{g})] = [Q(G)^G : Q(\mathfrak{g})^G]$.
• Let $a \in G$ be regular. Assume that $d\Phi(a)$ is invertible. Then Φ restricts to an isomorphism $\Phi : \overline{\mathrm{Conj}_G(a)} \to \overline{\mathrm{Ad}_G(\Phi(a))}$ of affine varieties.
• Let $a \in G$. Then for the semisimple parts we have $\Phi(a_s) = \Phi(a)_s$ and $\Phi(a) = \Phi(a_s) + \Phi(a)_n \in \mathfrak{g}^a$ is the Jordan decomposition.
• Let G be a connected reductive complex algebraic group and let $\Phi : G \to \mathfrak{g}$ be the Cayley mapping of a rational representation with $\Phi(e) = 0$. Then $\Phi : G_{\mathrm{pos}} \to \mathfrak{g}_{\mathrm{real}}$ is bijective and a fiber respecting isomorphism of real algebraic varieties, where G_{pos} is the set of all $a \in G$ whose semisimple part has positive eigenvalues, and $\mathfrak{g}_{\mathrm{real}}$ is the set of all $X \in \mathfrak{g}$ whose semisimple part has only real eigenvalues.

Relation to the classical Cayley mapping. Let $T : \mathrm{Spin}(n, \mathbb{C}) \to SO(n, \mathbb{C})$ be the double cover. We consider the spin representation $\mathrm{Spin} : \mathrm{Spin}(n, \mathbb{C}) \to \mathrm{Aut}(S_n)$.
• There is a choice of the sign of the square root so that $\chi(g) := \sqrt{\det(1 + T(g))}$ satisfies

$$\Phi_{\mathrm{Spin}}(g) = -\frac{2}{2^{n/2}}\,\chi(g)\,\Gamma(T(g)) \in \mathfrak{so}(n, \mathbb{C}).$$

for all $g \in \mathrm{Spin}(n, \mathbb{C})$. Moreover, $\chi \in A(\mathrm{Spin}(n, \mathbb{C}))$ and we have for the rational function fields

$$Q(\mathrm{Spin}(n))^{\mathrm{Spin}(n)} = Q(\mathfrak{so}(n, \mathbb{C}))^{\mathrm{Spin}(n)}[\chi],$$
$$Q(\mathrm{Spin}(n)) = Q(\mathfrak{so}(n, \mathbb{C}))[\chi].$$

Thus the generalized Cayley mapping $\Phi_{\mathrm{Spin}} : \mathrm{Spin}(n, \mathbb{C}) \to \mathfrak{so}(n, \mathbb{C})$ factors to the classical Cayley transform $\Gamma : SO(n, \mathbb{C})^* \to \mathrm{Lie}\,\mathrm{Spin}(n, \mathbb{C})^{(*)}$, up to multiplication by a function, via the natural identifications.

Relation to Poisson structures. For a representation π of a Lie group G we can try to pull back the Poisson structure on $\mathfrak{g}^*$ via the derivative of the character $d\chi_\pi : G \to \mathfrak{g}^*$. This pullback is a rational Poisson structure on G which in fact is an integrable Dirac structure in the sense of [1], [2], [3]. Let us explain this a little:

Let M be a smooth manifold of dimension m. A *Dirac structure* on M is a vector subbundle $D \subset TM \times_M T^*M$ with the following two properties:

(1) Each fiber D_x is maximally isotropic with respect to the metric of signature (m, m) on $TM \times_M T^*M$ given by $\langle (X, \alpha), (X', \alpha')\rangle_+ = \alpha(X') + \alpha'(X)$. So D is of fiber dimension m.
(2) The space of sections of D is closed under the non-skew-symmetric version of the Courant-bracket $[(X, \alpha), (X', \alpha')] = ([X, X'], \mathcal{L}_X\alpha' - i_{X'}d\alpha)$.

Natural examples of Dirac structures are the following: Symplectic structures ω on M, where $D = D^\omega = \{(X, \omega(X)) : X \in TM\}$ is just the graph of $\omega : TM \to T^*M$; these are precisely the Dirac structures D with $TM \cap D = \{0\}$. Poisson structures

P on M where $D = D^P = \{(P(\alpha), \alpha) : \alpha \in T^*M\}$ is the graph of $P : T^*M \to TM$; these are precisely the Dirac structures D which are transversal to T^*M.

Given a Dirac structure D on M we consider its range $R(D) = \mathrm{pr}_{TM}(D) = \{X \in TM : (X, \alpha) \in D \text{ for some } \alpha \in T^*M\}$. There is a skew symmetric 2-form Θ_D on $R(D)$ which is given by $\Theta_D(X, X') = \alpha(X')$ where $\alpha \in T^*M$ is such that $(X, \alpha) \in D$. The range $R(D)$ is an integrable distribution of non-constant rank in the sense of Stefan and Sussmann, see [5], so M is foliated into maximal integral submanifolds L of $R(D)$ of varying dimension, which are all initial submanifolds. The form Θ_D induces a closed 2-form on each leaf L and (L, Θ_D) is thus a presymplectic manifold (Θ_D might be degenerate on L). If the Dirac structure corresponds to a Poisson structure then the (L, Θ_D) are exactly the symplectic leaves of the Poisson structure.

The main advantage of Dirac structures is that one can apply arbitrary push forwards and pull backs to them. So if $f : N \to M$ is a smooth mapping and D_M is a Dirac structure on M then the pull back is defined by $f^*D_M = \{(X, f^*\alpha) \in TN \times_N T^*N : (Tf.X, \alpha) \in D_M\}$. Likewise the push forward of a Dirac structure D_N on N is given by $f_*D_N = \{(Tf.X, \alpha) \in TM \times_M T^*M : (X, f^*\alpha) \in D_N\}$.

References

[1] Courant, T. *Dirac manifolds*, Trans. AMS **319** (1990), 631–661.

[2] Bursztyn, H.; Radko, O., *Gauge equivalence of Dirac structures*, Ann. Inst. Fourier **53** (2003), 309–337.

[3] Bursztyn, H.; Crainic, M.; Weinstein, A.; Zhu, C., *Integration of twisted Dirac brackets*, (2003).

[4] Kolár, Ivan; Slovák, Jan; Michor, Peter W.: *Natural operations in differential geometry.* Springer-Verlag, Berlin, Heidelberg, New York, (1993)

[5] Kostant, Bert; Michor, Peter W., *The generalized Cayley map from an algebraic group to its Lie algebra*, The orbit method in geometry and physics: In honor of A. A. Kirillov. (Duval, Guieu, Ovsienko, eds.), Progress in Mathematics 213, Birkhäuser, Boston, 2003, pp. 259–296.

Θ-hypergeometric functions and shift operators

Angela Pasquale

The noncompactly causal (NCC) symmetric spaces are a small but nice class of pseudo-Riemannian symmetric spaces. The interest in these spaces was raised by the studies on the global structure of the space-time (see for instance [5]). In 1994, Faraut, Hilgert and Olafsson [1] could exploit the geometry of these spaces to extend to them the theory of spherical functions, which Harish-Chandra had developed in the late 50s on the Riemannian symmetric spaces of noncompact type [4]. As in the Riemannian case, the spherical functions on a NCC symmetric space G/H are the (suitably normalized) smooth H-invariant joint eigenfunctions of the commutative algebra of G-invariant differential operators on G/H. However, due to the non-compactness of H, they turn out to be much less regular than those of Harish-Chandra: they are only defined on an open submanifold of G/H; they are meromorphic (not entire) in the spectral parameter; they can be described by integral formulas only for certain values of the spectral parameters. Many of the difficulties encountered when studying the spherical functions on NNC symmetric spaces can be overcome by working in the more general setting of Θ-spherical functions.

The Θ-hypergeometric functions are special functions associated with root systems that generalize the spherical functions on both the NCC and the Riemannian symmetric spaces. Their definition has been suggested by Olafsson's expansion formula [7] for the spherical functions on a NCC symmetric space G/H. This formula shows that the restriction of the spherical functions of G/H to a specific Weyl chamber of Cartan subgroup is a certain linear combination of Harish-Chandra series for the Riemannian dual G/K. In their theory of hypergeometric functions associated with root systems [3, 2, 10, 6, 11], Heckman and Opdam developed very powerful methods for studying this kind of linear combinations without relying on a Riemannian structure. It is then quite natural to to try to extend Heckman-Opdam's definitions and methods to enclose also the spherical functions on NCC symmetric spaces. The big family of special functions originating from this extension gives precisely the Θ-hypergeometric functions. They are constructed from a triple $(\mathfrak{a}, \Sigma, m)$, where $\mathfrak{a}$ is a Euclidean symmetric space, Σ is a root system in the dual $\mathfrak{a}^*$ of $\mathfrak{a}$, and m is a multiplicity functions on Σ. As the hypergeometric functions associated with root systems, the Θ-hypergeometric functions are joint eigenfunctions of the hypergeometric system of Heckman and Opdam. The parameter Θ designates a subset of a fixed fundamental system Π of positive simple roots in Σ. The different choices of Θ lead to a lattice of special functions associated with the given root system. At the top of the lattice, that is for $\Theta = \Pi$, we find the hypergeometric functions of Heckman and Opdam; at the bottom, that is for $\Theta = \emptyset$, (certain multiples of) the Harish-Chandra series. In the middle appear many new special functions. For "geometric" triples $(\mathfrak{a}, \Sigma, m)$, the Θ-hypergeometric functions corresponding to $\Theta = \Pi$ yield Harish-Chandra's spherical functions, whereas the spherical functions on NCC symmetric spaces arise from some of the new special functions in the central part of the lattice. This unified framework allows us, for instance, to derive information on the spherical functions on NCC symmetric spaces from those of the spherical functions of the Riemannian dual.

A particularly nice situation occurs for even multiplicity functions on reduced root systems. Geometrically, this situation corresponds to Riemannian symmetric spaces G/K with the property that all Cartan subalgebras in the Lie algebra $\mathfrak{g}$ of G are conjugate. The simplest example is when $\mathfrak{g}$ admits a complex structure, in which case all multiplicities are equal to 2. The analysis of Θ-hypergeometric functions with even multiplicities is simplified by the use of Opdam's shift operators (see e.g. [6]). By modifying one of these operators, it is possible to obtain a Weyl-group-invariant differential operator with regular coefficients yielding Θ-hypergeometric functions with even multiplicities from averages of exponential functions. In particular, this provides new formulas for the spherical functions on

Riemannian symmetric spaces with even multiplicities of both noncompact and compact type. The study of of the Θ-hypergeometric functions in even multiplicities and their associated harmonic analysis is a joint work with Gestur Ólafsson [9].

References

[1] J. Faraut, J. Hilgert, and G. Ólafsson, *Spherical functions on ordered symmetric spaces*, Ann. Inst. Fourier **44** (1994), 927–966.

[2] G. J. Heckman, *Root systems and hypergeometric functions.* II, Composition Math. **64** (1987), no. 3, 353–373.

[3] G. J. Heckman and E. M. Opdam, *Root systems and hypergeometric functions.* I, Compositio Math. **64** (1987), no. 3, 329–352.

[4] Harish-Chandra, *Spherical functions on a semisimple Lie group.* I, II. Amer. J. Math. **80** (1958), 241–310; 553–613.

[5] S. W. Hawking and G. F. R. Ellis, *The large scale structure of space-time.* Cambridge Monographs on Mathematical Physics, No. 1. Cambridge University Press, London-New York, 1973.

[6] G. J. Heckman and H. Schlichtkrull, *Harmonic analysis and special functions on symmetric spaces*, Perspectives in Mathematics, 16. Academic Press, Inc., San Diego, CA, 1994.

[7] G. Ólafsson, *Spherical functions and spherical Laplace transform on ordered symmetric spaces*, Preprint. Available at `http://www.math.lsu.edu/preprint`, 1997.

[8] G. Ólafsson and A. Pasquale, *On the meromorphic extension of the spherical functions on noncompactly causal symmetric spaces*, J. Funct. Anal. **181** (2001), no. 2, 346–401.

[9] G. Ólafsson and A. Pasquale, *A Paley-Wiener theorem for the Θ-hypergeometric functions: the even multiplicity case.* Preprint 2003. To appear on J. Math. Pures Appl.

[10] E.M. Opdam, *Root systems and hypergeometric functions.* III, IV Compositio Math. **67** (1988), 21–49; 191–209.

[11] E.M. Opdam, *Harmonic analysis for certain representations of graded Hecke algebras.* Acta math. **175** (1995), 75–121.

[12] A. Pasquale, *A theory of Θ-spherical functions*, Habilitationsschrift, Technische Universität Clausthal, 2002.

[13] A. Pasquale, *Asymptotic analysis of Θ-hypergeometric functions*, DOI: 10.1007/s00222-003-0349-9. To appear on Invent. Math. (2004).

Maximal adapted complexifications of Riemannian homogeneous spaces

Andrea Iannuzzi
(joint work with Stefan Halverscheid)

For a Riemannian real-analytic manifold M one can construct canonical complexifications by defining the *adapted* complex structure on a domain of the tangent bundle TM, as shown by Guillemin-Stenzel and Lempert-Szoeke ([GS], [LS]). This uniquely determines the complexification in a neighborhood of M, which is identified with the zero section in TM, however in general there are questions about existence and unicity of a maximal domain Ω_{max} on which the adapted complex structure exists. If Ω_{max} is understood, by functoriality of the definition it may be regarded as an invariant of the metric, i.e., isometric manifolds have biholomorphic maximal domains. For instance examples are given by symmetric spaces of non-compact type ([BHH]), compact normal Riemannian Homogeneous spaces ([Sz2]), compact symmetric spaces ([Sz1]) and spaces obtained by Kählerian reduction of these ([A]). Note that in the mentioned cases maximal domains turn out to be Stein.

Let us consider a Riemannian homogeneous space $M = G/K$, with G a Lie group of isometries and K compact. It is reasonable to assume that $\dim_{\mathbb{C}} G^{\mathbb{C}} = \dim_{\mathbb{R}} G$, where $G^{\mathbb{C}}$ is the universal complexification of G. Then $K^{\mathbb{C}}$ acts on $G^{\mathbb{C}}$, the left action on M induces a natural G-action on TM and under certain extensibility assumptions on the geodesic flow of M one obtains a real-analytic and G-equivariant map $P : TM \to G^{\mathbb{C}}/K^{\mathbb{C}}$ such that

the connected component of the non-singular locus of DP containing M is the unique maximal domain on which the adapted complex structure exists.

This applies to the case of generalized Heisenberg groups and naturally reductive Riemannian homogeneous spaces, among which one finds all isotropy irreducible homogeneous spaces classified by J. Wolf [W].

As an application it is shown that for all generalized Heisenberg groups such maximal domain is neither holomorphically separable, nor holomorphically convex. We are not aware of previous non-Stein examples. Moreover allready in the case of the 3-dimensional Heisenberg group one notices mixed signature Ricci curvature, suggesting an influence of curvature properties of M on the complex geometry of the maximal adapted complexification. Some recent new examples give a different light to such point of view.

References

[A] Aguilar, R. *Symplectic reduction and the homogeneous complex Monge-Ampère equation* Ann. Global Anal. Geom. 19 (2001), no. 4, 327–353

[AG] Akhiezer, D.-N.; Gindikin, S.-G. *On Stein extension of real symmetric spaces* Math. Ann. 286 (1990), 1–12

[BHH] Burns, D.; Halverscheid, S.; Hind, R. *The Geometry of Grauert Tubes and Complexification of Symmetric Spaces* Preprint, 2001, math.CV/0109186

[BTV] Berndt, J.; Tricerri, F.; Vanhecke, L. *Generalized Heisenberg Groups and Damek-Ricci Harmonic Spaces* LNM 1598, Springer-Verlag 1995

[GS] Guillemin, V.; Stenzel, M. *Grauert tubes and the homogeneous Monge-Ampère equation* J. Differential Geom. 34 (1991), no. 2, 561–570 (first part); J. Differential Geom. 35 (1992), no. 3, 627–641 (second part)

[Ha] Halverscheid, S. *On Maximal Domains of Definition of Adapted Complex Structures for Symmetric Spaces of Non-compact Type* Thesis, Ruhr-Universität Bochum, 2001

[LS] Lempert, L.; Szőke, R. *Global solutions of the homogeneous complex Monge-Ampère equation and complex structures on the tangent bundles of Riemannian manifolds* Math. Ann. 290 (1991), 689–712

[PW] Patrizio, G.; Wong, P.-M. *Stein manifolds with compact symmetric center* Math. Ann. 289 (1991), no. 4, 355–382

[Sz1] Szőke, R. *Complex structures on tangent bundles of Riemannian manifolds* Math. Ann. 291 (1991), 409–428

[Sz2] Szőke, R. *Adapted complex structures and Riemannian homogeneous spaces* Complex analysis and applications (Warsaw, 1997) Ann. Polon. Math. 70 (1998), 215–220.

[W] Wolf, J. *The geometry and structure of isotropy irreducible homogeneous spaces* Acta Math. 120 (1968), 59–148.

Reporter: Patrick Schützdeller

Participants

Prof. Dr. Dmitry N. Akhiezer
akhiezer@mccme.ru
akhiezer@cplx.ruhr.uni-bochum.de
Institute for Information Transmission Problems
Russian Academy of Sciences
19 Bol.Karetny per,
101447 Moscow GSP-4 – Russia

Prof. Dr. Leticia Barchini
leticia@math.okstate.edu
Dept. of Mathematics
Oklahoma State University
401 Math Science
Stillwater, OK 74078-1058 – USA

Prof. Dr. Daniel Barlet
barlet@iecn.u-nancy.fr
Departement de Mathematiques
Universite de Nancy I
Boite Postale 239
F-54506 Vandoeuvre les Nancy Cedex

Prof. Dr. Daniel Beltita
Daniel.Beltita@imar.ro
Institute of Mathematics
"Simion Stoilow"
of the Romanian Academy
P.O. Box 1-764
014700 Bucharest – Romania

Prof. Dr. Wolfgang Bertram
bertram@iecn.u-nancy.fr
Departement de Mathematiques
Universite de Nancy I
Boite Postale 239
F-54506 Vandoeuvre les Nancy Cedex

Dr. Harald Biller
biller@mathematik.tu-darmstadt.de
Fachbereich Mathematik
TU Darmstadt
Schloßgartenstr. 7
D–64289 Darmstadt

Prof. Dr. Ralph Bremigan
bremigan@math.bsu.edu
Dept. of Mathematical Sciences
Ball State University
Muncie, IN 47306-0490
USA

Prof. Dr. Jean-Louis Clerc
clerc@iecn.u-nancy.fr
Institut Elie Cartan
-Mathematiques-
Universite Henri Poincare, Nancy I
Boite Postale 239
F-54506 Vandoeuvre les Nancy Cedex

Prof. Dr. Ivan Dimitrov
dimitrov@mast.queensu.ca
Department of Mathematics
Queen's University
Jeffery Hall
99 University Avenue
Kingston ONT K7L 3N6 – Canada

Prof. Dr. Alexander Dvorsky
dvorsky@math.miami.edu
Dept. of Mathematics and Computer Science
University of Miami
P.O. Box 249085
Coral Gables, FL 33124-4250 – USA

Prof. Dr. Jacques Faraut
faraut@math.jussieu.fr
Institut de Mathematiques
Analyse Algebrique
Universite Pierre et Marie Curie
4, place Jussieu, Case 247
F-75252 Paris Cedex 5

Gregor Fels
gfels@uni-tuebingen.de
Mathematisches Institut
Universität Tübingen
D–72074 Tübingen

Prof. Dr. Alice Fialowski
fialowsk@cs.elte.hu
Department of Analysis
ELTE TTK
Pazmany Peter setany 1/c
1117 Budapest – Hungary

Dr. Laura Geatti
geatti@mat.uniroma2.it
Dipartimento di Matematica
II. Universita di Roma
Via della Ricerca Scientifica
I-00133 Roma

Dr. Helge Glöckner
gloeckner@mathematik.tu-darmstadt.de
Fachbereich Mathematik
TU Darmstadt
Schloßgartenstr. 7
D–64289 Darmstadt

Dr. Anna Gori
gori@math.unifi.it
Dipartimento Matematica "U.Dini"
Universita degli Studi
Viale Morgagni, 67/A
I-50134 Firenze

Prof. Dr. Laurent Guieu
guieu@math.univ-montp2.fr
guieu@darboux.math.univ-montp2.fr
Departement de Mathematiques
Universite Montpellier II
Place Eugene Bataillon
F-34095 Montpellier Cedex 5

Prof. Dr. Peter Heinzner
Heinzner@cplx.ruhr-uni-bochum.de
Fakultät für Mathematik
Ruhr-Universität Bochum
Universitätsstr. 150
D–44801 Bochum

Prof. Dr. Joachim Hilgert
hilgert@math.tu-clausthal.de
Institut für Mathematik
Technische Universität Clausthal
Erzstr. 1
D–38678 Clausthal-Zellerfeld

Prof. Dr. Jaehyun Hong
jhhong@math.berkeley.edu
Department of Mathematics
University of California
Berkeley, CA 94720-3840 – USA

Prof. Dr.Dr.h.c. Alan T. Huckleberry
ahuck@cplx.ruhr-uni-bochum.de
Fakultät für Mathematik
Ruhr-Universität Bochum
D–44780 Bochum

Dr. Andrea Iannuzzi
iannuzzi@mat.uniroma2.it
Dipartimento di Matematica
II. Universita di Roma
Via della Ricerca Scientifica
I-00133 Roma

Prof. Dr. Wilhelm Kaup
kaup@uni-tuebingen.de
Mathematisches Institut
Universität Tübingen
Auf der Morgenstelle 10
D–72076 Tübingen

Prof. Dr. Toshiyuki Kobayashi
toshi@kurims.kyoto-u.ac.jp
Research Institute for
Mathematical Sciences
Kyoto University
Kitashirakawa, Sakyo-ku
Kyoto 606-8502 – Japan

Bernhard Krötz
kroetz@math.uoregon.edu
kroetz@darkwing.uoregon.edu
Dept. of Mathematics
University of Oregon
Eugene, OR 97403-1222 – USA

Prof. Dr. Laszlo Lempert
lempert@math.purdue.edu
Dept. of Mathematics
Purdue University
West Lafayette, IN 47907-1395 – USA

Prof. Dr. Joshua A. Leslie
jleslie@howard.edu
joshuales1@aol.com
Dept. of Mathematics
Howard University
Washington, DC 20059 – USA

Prof. Dr. Peter W. Michor
peter.michor@esi.ac.at
Fakultät für Mathematik
Universität Wien
Nordbergstr. 15
A-1090 Wien

Prof. Dr. Karl-Hermann Neeb
neeb@mathematik.tu-darmstadt.de
Fachbereich Mathematik
TU Darmstadt
Schloßgartenstr. 7
D–64289 Darmstadt

Prof. Dr. Kyo Nishiyama
kyo@math.kyoto-u.ac.jp
Division of Mathematics
Graduate School of Science
Kyoto University
Kyoto 606-8502 – Japan

Ben Ntatin
ntatin@cplx.ruhr-uni-bochum.de
Fakultät für Mathematik
Ruhr-Universität Bochum
D–44780 Bochum

Prof. Dr. Arkadiy L. Onishchik
aonishch@aha.ru
onishch.@univ.uniyar.ac.ru
Department of Mathematics
Yaroslavl' State University
Sovjetskaya ul. 14
Yaroslavl 150000 – Russia

Prof. Dr. Bent Orsted
orsted@imada.ou.dk
orsted@imada.ou.dk.bitnet
orsted@imada.sdu.dk
Matematisk Institut
Odense Universitet
Campusvej 55
DK-5230 Odense M

Prof. Dr. Angela Pasquale
pasquale@poncelet.univ-metz.fr
Departement et Laboratoire de
Mathematiques, Universite de Metz
ISGMP, Batiment A
Ile Du Saulcy
F-57045 Metz

Prof. Dr. Ivan Penkov
penkov@math.ucr.edu
ivanpenkov@yahoo.com
Dept. of Mathematics
University of California
Riverside, CA 92521-0135 – USA

Prof. Dr. Martin Schlichenmaier
schlichenmaier@cu.lu
Laboratoire de Mathematique
Universite du Luxembourg
162 A, avenue de la Faiencerie
L-1511 Luxembourg

Patrick Schützdeller
patrick@cplx.ruhr-uni-bochum.de
Fakultät für Mathematik
Ruhr-Universität Bochum
Gebäude NA4
D–44780 Bochum

Prof. Dr. Andrew Sinton
sinton@math.berkeley.edu
Department of Mathematics
University of California
Berkeley, CA 94720-3840 – USA

Prof. Dr. Harald Upmeier
upmeier@mathematik.uni-marburg.de
Fachbereich Mathematik
Universität Marburg
D–35032 Marburg

Prof. Dr. Cornelia Vizman
vizman@math.uvt.ro
Mathematisches Institut
West University of Timisoara
Bul.V.Parvan n.4
1900 Timisoara – Romania

Dr. Friedrich Wagemann
wagemann@math.univ-nantes.fr
Laboratoire de Mathematique
Universite de Nantes
2 rue de la Houssiniere
F-44322 Nantes Cedex 03

Prof. Dr. Jörg Winkelmann
jwinkel@member.ams.org
jw@cplx.ruhr-uni-bochum.de
Institut Elie Cartan
-Mathematiques-
Universite Henri Poincare, Nancy I
Boite Postale 239
F-54506 Vandoeuvre les Nancy Cedex

Prof. Dr. Joseph Albert Wolf
jawolf@math.berkely.edu
Department of Mathematics
University of California
Berkeley, CA 94720-3840 – USA

Prof. Dr. Tilmann Wurzbacher
wurzbacher@poncelet.univ-metz.fr
Laboratoire de Mathematiques
Universite de Metz et C.N.R.S.
Ile du Saulcy
F-57045 Metz Cedex 01

Dr. Dmitri Zaitsev
zaitsev@maths.tcd.ie
Dept. of Mathematics
Trinity College
University of Dublin
Dublin 2 – Ireland

Prof. Dr. Genkai Zhang
genkai@math.chalmers.se
Department of Mathematics
Chalmers University of Technology
S-412 96 Göteborg

Prof. Dr. Roger Zierau
zierau@math.okstate.edu
zierau@littlewood.math.okstate.edu
Dept. of Mathematics
Oklahoma State University
401 Math Science
Stillwater, OK 74078-1058 – USA

Mathematisches Forschungsinstitut Oberwolfach

Report No. 7/2004

Funktionentheorie

Organised by
Walter Bergweiler (Kiel)
Stephan Ruscheweyh (Würzburg)
Ed Saff (Vanderbilt)

February 8th – February 14th, 2004

Introduction by the Organisers

The present conference was organized by Walter Bergweiler (Kiel), Stephan Ruscheweyh (Würzburg) and Ed Saff (Vanderbilt).

The 24 talks gave an overview of recent results and current trends in function theory.

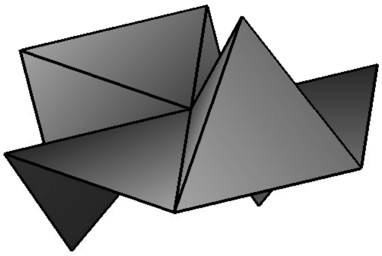
Workshop on Funktionentheorie

Table of Contents

Abstracts

Curvature flow in conformal mapping
Kenneth Stephenson
(joint work with Charles Collins and Tobin Driscoll)

In joint work with Charles Collins (Tennessee) and Tobin Driscoll (Delaware), the author investigates the conformal mapping of a non-planar Riemann surface to a rectangle in the plane. The methods involve circle packing, and the discussion centres on a simple prototype problem: A Riemann surface $\mathcal{S}$ is created as a nonplanar cone space by pasting 10 equilateral triangles together in a specified pattern. Four vertices on the boundary are designated as "corners". It is well known classically that there is a conformal map $F : \mathcal{S} \longrightarrow \mathcal{R}$ mapping S to a plane rectangle $\mathcal{R}$ with corners going to corners, as suggested in Figure 1.



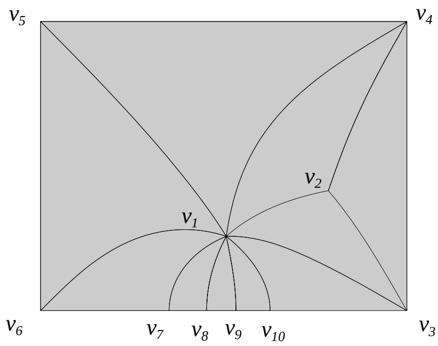


FIGURE 1. Conformally mapping an equilateral surface to a rectangle

Circle packing provides a means for numerically approximating F. A sequence of ever finer *insitu* circle packings Q_n are created in $\mathcal{S}$ based on its equilateral structure and a "repacking" computation then lays out circle packings P_n in the plane having the same combinatorics but with carriers that form rectangles $\mathcal{R}_n$. For each n the associated map $f_n : Q_n \longrightarrow P_n$ is defined as a "discrete conformal map". It has been established by Phil Bowers and the author that as n grows, appropriately normalized rectangles $\mathcal{R}_n$ converge to $\mathcal{R}$ and the discrete conformal maps f_n converge uniformly on compact subsets of $\mathcal{S}$ to F. (See [1] and for background, [3, 4].) The circle packing on the left in Figure 2 is P_6; the images of the 10 faces of $\mathcal{S}$ here are very close to their correct conformal shapes.

In studying this mapping, the authors parametrized the flattening process, both classical and discrete, in a natural way to obtain a continuous family of surfaces stretching from $\mathcal{S}$ to $\mathcal{R}$. One can observe experimentally the "flow" of radius adjustments as the circle packings are computed from one discrete surface to the next; that flow reflects the movement of "curvature" at the circle centres during

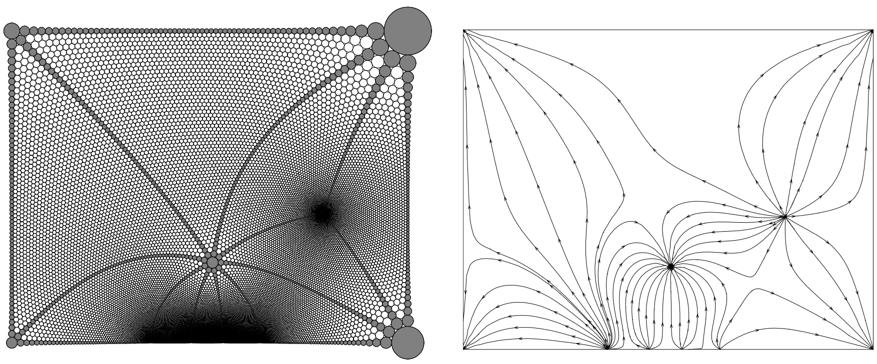


FIGURE 2. The packing P_6 and the associated "flow" field

the adjustment process. The surprise came in our observation that this flow was essentially independent of the parametrization stage. In other words, from the beginning to the end of the parametrization the circles seemed to move in accordance with an unchanging prescription about how to coordinate their size adjustments. On the right in Figure 2 is one of these simulated flow fields.

This field ultimately describes the flow of cone angle (curvature) among the ten cone points of $\mathcal{S}$ during the flattening process. The authors looked for a classical parallel and obtained it via a modification of the Schwarz-Christoffel (SC) method [2]. That modification introduces interior cone points and cuts to allow mapping to a non-planar surface. The experimental flows are nearly exact copies of the gradient field $\nabla \log |\Phi'(z)|$, where Φ' is the derivative of the mapping function generated by our modified SC method (and then lifted to $\mathcal{R}$). This raises a number of questions about the classical interpretation and the possible uses for this "curvature" flow.

References

[1] Philip L. Bowers and Kenneth Stephenson, *Uniformizing dessins and Belyĭ maps via circle packing*, Amer. Math. Soc., to appear.

[2] Tobin A. Driscoll, *A MATLAB toolbox for Schwarz-Christoffel mapping*, 1996.

[3] Kenneth Stephenson, *Approximation of conformal structures via circle packing*, Computational Methods and Function Theory 1997, Proceedings of the Third CMFT Conference (N. Papamichael, St. Ruscheweyh, and E. B. Saff, eds.), vol. 11, World Scientific, 1999, pp. 551–582.

[4] Kenneth Stephenson, *Circle packing and discrete analytic function theory*, Handbook of Complex Analysis, Vol. 1: Geometric Function Theory (R. Kühnau, ed.), Elsevier, 2002.

Julia polynomials and the Szegő kernel method
Igor Pritsker

Let G be a Jordan domain bounded by a rectifiable curve L of length l. The Smirnov space of analytic functions $E_2(G)$ is defined by the product $\langle f, g\rangle = \frac{1}{l}\int_L f(z)\overline{g(z)}|dz|$ (see [2], [3] and [10]). Consider the associated contour orthonormal polynomials $\{p_n(z)\}_{n=0}^{\infty}$. If G is a Smirnov domain, then polynomials are dense in $E_2(G)$ [2]. In this case, the Szegő kernel is given by $K(z,\zeta) = \sum_{k=0}^{\infty} \overline{p_k(\zeta)} p_k(z) = \frac{l}{2\pi}\sqrt{\varphi'(z)\overline{\varphi'(\zeta)}},\ z,\zeta \in G$, where φ is the conformal map of G onto the unit disk, normalized by $\varphi(\zeta) = 0,\ \varphi'(\zeta) > 0$ [11]. Julia polynomials approximate φ, with a construction resembling Bieberbach polynomials in the Bergman kernel method,

$$J_{2n+1}(z) = \frac{2\pi}{l}\int_{\zeta}^{z}\left(\sum_{k=0}^{n}\overline{p_k(\zeta)}p_k(t)\right)^2 dt \,\Big/ \sum_{k=0}^{n}|p_k(\zeta)|^2, \quad n \in \mathbb{N}.$$

The uniform convergence of Bieberbach polynomials has been extensively studied, but methods based on the Szegő kernel did not receive a comprehensive attention. It is not difficult to see that J_{2n+1} converge to φ locally uniformly in G. We show in [9] that J_{2n+1} converge to φ uniformly on the closure of any Smirnov domain. This class contains all Ahlfors-regular domains [8], allowing arbitrary (even zero) angles at the boundary. For the piecewise analytic domains, we also give the estimate

$$\|\varphi - J_{2n+1}\|_{L_\infty(\overline{G})} \le C(G)\ n^{-\frac{\lambda}{4-2\lambda}}, \quad n \in \mathbb{N}, \tag{1}$$

where $\lambda\pi,\ 0 < \lambda < 2$, is the smallest exterior angle at the boundary of G. The rate of convergence for J_{2n+1} on compact subsets of G is essentially squared comparing to (1). These results have standard applications to the rate of decay for the contour orthogonal polynomials inside the domain, and to the rate of locally uniform convergence of Fourier series.

The approximating polynomials of this kind were first introduced via an extremal problem by Keldysh and Lavrentiev (cf. [5], [6] and [7]), who developed the ideas of Julia [4]. Set $\|f\|_p = \left(\int_L |f(z)|^p|dz|\right)^{1/p}$ for $f \in E_p(G),\ 0 < p < \infty$, where $E_p(G)$ is the Smirnov space [2]. Let $Q_{n,p}$ be a polynomial minimizing $\|P_n\|_p$ among all polynomials P_n such that $P_n(\zeta) = 1$. Julia [4] showed that the corresponding extremal problem in the class of all $E_p(G)$ functions is solved by $(\phi')^{1/p}$, where ϕ is the conformal map of G onto a disk $\{z : |z| < R\}$, normalized by $\phi(\zeta) = 0$ and $\phi'(\zeta) = 1$. Keldysh and Lavrentiev [7] proved that $Q_{n,p}$ converge to $(\phi')^{1/p}$ locally uniformly in G if and only if G is a Smirnov domain. Thus the polynomials $J_{n,p}(z) := \int_\zeta^z Q_{n,p}^p(t)dt$ provide an approximation to $\phi(z)$. If $p = 2$ then $J_{n,2}$ differ from J_{2n+1} just by a constant factor. This case was studied by Ahlfors [1], Warschawski [12] and Gaier [3]. Again, the locally uniform convergence of $J_{n,p}$ to ϕ in Smirnov domains is immediate for any $p \in (0,\infty)$. We prove the uniform convergence on $\overline{G}$ in arbitrary Smirnov domains, and give the convergence rates generalizing (1) for piecewise analytic domains.

References

[1] L. V. Ahlfors, *Two numerical methods in conformal mapping*, Experiments in the computation of conformal maps, pp. 45-52. National Bureau of Standards Applied Mathematics Series, No. 42. U. S. Government Printing Office, Washington, D.C., 1955.

[2] P. L. Duren, Theory of H^p Spaces, Dover, New York, 2000.

[3] D. Gaier, Konstruktive Methoden der konformen Abbildung, Springer-Verlag, Berlin, 1964.

[4] G. Julia, Lecons sur la représentation conforme des aires simplement connexes, Paris, 1931.

[5] M. V. Keldysh, *On a class of extremal polynomials*, Dokl. Akad. Nauk SSSR **4** (1936), 163-166. (Russian)

[6] M. V. Keldysh and M. A. Lavrentiev, *On the theory of conformal mappings*, Dokl. Akad. Nauk SSSR **1** (1935), 85-87. (Russian)

[7] M. V. Keldysh and M. A. Lavrentiev, *Sur la représentation conforme des domaines limités par des courbes rectifiables*, Ann. Sci. École Norm. Sup. **54** (1937), 1-38.

[8] Ch. Pommerenke, Boundary Behaviour of Conformal Maps, Springer-Verlag, Berlin, 1992.

[9] I. E. Pritsker, *Approximation of conformal mapping via the Szegő kernel method*, Comp. Methods and Function Theory **3** (2003), 79-94.

[10] V. I. Smirnov and N. A. Lebedev, Functions of a Complex Variable: Constructive Theory, MIT Press, Cambridge, 1968.

[11] G. Szegő, Orthogonal Polynomials, Amer. Math. Soc., Providence, 1975.

[12] S. E. Warschawski, *Recent results in numerical methods of conformal mapping*, in "Proceedings of Symposia in Applied Mathematics. Vol. VI. Numerical Analysis," McGraw-Hill Book Company, Inc., New York, 1956, pp. 219-250.

Conformal Pseudo–metrics and a free boundary value problem for analytic functions

Daniela Kraus

The starting point is the following free boundary value problem for analytic functions f which are defined on a domain $G \subset \mathbb{C}$ and map into the unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$.

Problem 1 *Let $z_1, \ldots, z_n$ be finitely many points in a bounded simply connected domain $G \subset \mathbb{C}$ and let $\phi : \partial G \to (0, \infty)$ be a continuous function. Show that there exists a holomorphic function $f : G \to \mathbb{D}$ with critical points z_j (counted with multiplicities) and no others such that*

$$\lim_{z \to \xi} \frac{|f'(z)|}{1 - |f(z)|^2} = \phi(\xi)$$

for all $\xi \in \partial G$.

If $G = \mathbb{D}$, $\phi \equiv 1$, Problem 1 was solved by Kühnau [5] in case of one critical point, which is sufficiently close to the origin, and for more than one critical point by Fournier and Ruscheweyh [2]. The method employed by Kühnau, Fournier and Ruscheweyh easily extends to more general domains G, say bounded by a Dini–smooth Jordan curve, but does not work for arbitrary bounded simply connected domains.

We present a completely new approach to Problem 1, which shows that this boundary value problem is not an isolated question in complex analysis, but is intimately connected to a number of basic (open) problems in conformal geometry and non–linear PDE. To solve Problem 1 for arbitrary bounded simply connected domains we divide it into the following two parts.

In a first step we construct a conformal metric in a bounded regular domain $G \subset \mathbb{C}$ with prescribed non–positive Gaussian curvature $\kappa(z)$ and prescribed singularities by solving the first boundary value problem for the Gaussian curvature equation $\Delta u = -\kappa(z)e^{2u}$ in G with prescribed singularities and continuous boundary data. More precisely, we have

Theorem 1 *Let $G \subset \mathbb{C}$ be a bounded and regular domain, let $z_1, z_2, \ldots, z_n \in G$ be finitely many distinct points and let $\alpha_1, \ldots, \alpha_n \in (0, \infty)$. Let $\phi : \partial G \to (0, \infty)$ be a continuous function and $\kappa : G \to (-\infty, 0]$ a bounded and locally Hölder continuous function with exponent α, $0 < \alpha \leq 1$. Then there exists a unique pseudo–metric $\lambda : G \to [0, \infty)$ of curvature $\kappa(z)$ in $G \backslash \{z_1, z_2, \ldots, z_n\}$ with zeros of orders α_j at z_j and no others such that λ is continuous on $\overline{G}$ with $\lambda(z) = \phi(z)$ for $z \in \partial G$.*

Theorem 1 is related to the Berger–Nirenberg problem in Riemannian geometry, that is, the question which functions on a surface R can arise as the Gaussian curvature of a Riemannian metric on R. The special case, where $\kappa(z) \equiv -4$ and the domain G is bounded by finitely many analytic Jordan curves was treated by Heins [4].

In a second step we show every conformal pseudo–metric on a simply connected domain $G \subseteq \mathbb{C}$ with constant negative Gaussian curvature and isolated zeros of integer order is the pullback of the hyperbolic metric on $\mathbb{D}$ under an analytic map $f : G \to \mathbb{D}$:

Theorem 2 *Let $E = \{z_1, z_2, \ldots\}$ be a discrete set in a simply connected domain $G \subseteq \mathbb{C}$, let $\alpha_1, \alpha_2, \ldots$ be positive integers, and let $\lambda : G \to [0, \infty)$ be a pseudo–metric of constant curvature $\kappa = -4$ in $G \backslash E$ with zeros of orders α_j at z_j and no others. Then λ is the pullback of the hyperbolic metric under a holomorphic function $f : G \to \mathbb{D}$, i.e.*

$$\lambda(z) = \frac{|f'(z)|}{1 - |f(z)|^2}, \quad z \in G.$$

If $g : G \to \mathbb{D}$ is another holomorphic function such that

$$\lambda(z) = \frac{|g'(z)|}{1 - |g(z)|^2}, \quad z \in G,$$

then $g = T \circ f$, where T is a conformal automorphism of the unit disk $\mathbb{D}$.

This extends a theorem of Liouville [6] which deals with the case that the pseudo–metric has no zeros at all.
Theorem 1 and Theorem 2 together allow in particular a quick and complete solution of Problem 1.

References

[1] Agranovsky, M. L., Bandman, T. M., *Remarks on a Conjecture of Ruscheweyh*, Complex Variables (1996), **31**, 249–258.

[2] Fournier, R., Ruscheweyh, St., *Free boundary value problems for analytic functions in the closed unit disk*, Proc. Amer. Math. Soc. (1999), **127** no. 11, 3287–3294.

[3] Fournier, R., Ruscheweyh, St., *A generalization of the Schwarz–Carathéodory reflection principle and spaces of pseudo–metrics*, Math. Proc. Camb. Phil. Soc. (2001), **130**, 353–364.

[4] Heins, M., *On a class of conformal metrics*, Nagoya Math. J. (1962), **21**, 1–60.

[5] Khnau, R., *Lngentreue Randverzerrung bei analytischer Abbildung in hyperbolischer und sphrischer Geometrie*, Mitt. Math. Sem. Giessen (1997), **229**, 45–53.

[6] Liouville, J., *Sur l'équation aux différences partielles* $\frac{d^2 \log \lambda}{du dv} \pm \frac{\lambda}{2a^2} = 0$, J. de Math. (1853), **16**, 71–72.

Critical points of discrete potentials in space

J.K. Langley

(joint work with J. Rossi)

The following was conjectured in [1]: let $z_k \in \mathbb{C}, a_k > 0$,

$$z_k \to \infty, \quad \sum_{z_k \neq 0} \left| \frac{a_k}{z_k} \right| < \infty, \quad f(z) = \sum_{k=1}^{\infty} \frac{a_k}{z - z_k}. \tag{1}$$

Then f has infinitely many zeros.

The zeros of f correspond to equilibrium points of the electrostatic field generated by wires carrying charge density $a_k/2$, perpendicular to the plane at z_k. The conjecture is known to be true in two contrasting cases: (i) if the total charge $\sum_{k=1}^{\infty} a_k$ is finite (or, more generally, if $\sum_{|z_k| \leq r} a_k = o(\sqrt{r})$ as $r \to \infty$) [1]; (ii) if $\inf\{a_k\} > 0$ [2].

For point charges in space, the following was proved in [1]. Let $x_k \in \mathbb{R}^3$, with

$$x_k \to \infty, \quad \sum_{x_k \neq 0} \frac{a_k}{|x_k|} < \infty, \quad u(x) = \sum_{k=1}^{\infty} \frac{a_k}{|x - x_k|}. \tag{2}$$

If $\inf\{a_k\} > 0$ then u has infinitely many critical points in $\mathbb{R}^3$.

In this case the critical points of u are equilibrium points of the electrostatic field generated by charges a_k at x_k. Langley and Rossi [5] have recently shown that instead of the condition $\inf\{a_k\} > 0$ it suffices that the x_k have finite exponent of convergence, which follows at once from (2) if $\inf\{a_k\} > 0$. The Cartan lemma [3, p.366] is used to prove that there exist spheres $|x| = r_n \to \infty$ on which the maximum of $u(x)$ tends to 0, following which the method of [1] is applied.

The talk concludes with some results from [4] concerning zeros of $f(z)$ when the a_k are complex in (1). A number of methods are applied, including quasiconformal surgery.

References

[1] J. Clunie, A. Eremenko and J. Rossi, On equilibrium points of logarithmic and Newtonian potentials, J. London Math. Soc. (2) 47 (1993), 309-320.

[2] A. Eremenko, J.K. Langley and J. Rossi, On the zeros of meromorphic functions of the form $\sum_{k=1}^{\infty} \frac{a_k}{z-z_k}$, J. d'Analyse Math. 62 (1994), 271-286.

[3] W.K. Hayman, Subharmonic functions Vol. 2, Academic Press, London, 1989.

[4] J.K. Langley and J. Rossi, Meromorphic functions of the form $f(z) = \sum_{n=1}^{\infty} a_n/(z - z_n)$, Rev Mat. Iberoamericana 20 (2004), 285-314.

[5] J.K. Langley and J. Rossi, Critical points of certain discrete potentials, preprint.

Efficient Discretization of Green Energy and Grunsky-Type Development of Functions Univalent in an Annulus

Marcus Stiemer

Let Γ be an analytic Jordan curve in the complex plane. In 1970, K. Menke introduced an extremal point system on Γ and applied it to approximate the logarithmic capacity of Γ and the conformal mapping Φ from the outer domain of the unit circle onto the outer domain of Γ with $\Phi(z) = dz + O(1)$, $z \to \infty$, $d > 0$ geometrically fast [2, 3, 4, 5, 6]. D. Gaier introduced the notation *Menke points* for systems of this type. In contrast to Fekete-points, which possess a worse distribution on analytic Jordan curves [10, 11], Menke-points consist of two sets of points that alternate on the curve Γ. An extension to the hyperbolic situation (see below) has been developed in [9].

Let now $F \subset \widehat{\mathbb{C}}$ be a set with connected complement Ω, such that the Green function $G(z, \zeta)$ in Ω with pole in $\zeta \in \Omega$ exists. Moreover, let Γ be an analytic Jordan curve in Ω with $E = \overline{\operatorname{Int}\Gamma}$.

The purpose of this work is to develop a Menke-type discretization for the measure of minimal Green energy on Γ with respect to Ω and to prove that this discretization provides a geometrically fast converging approximation to minimal Green energy.

Particularly for the hyperbolic situation, $F = \widehat{\mathbb{C}} \setminus \mathbb{D}$, $\Omega = \mathbb{D}$, we prove that Menke-points approximate the images of rotated roots of unity under the conformal mapping Φ from $\{1 < |z| < e^{1/C(E,F)}\}$ onto $\mathcal{R} = \mathbb{D} \setminus E$ with $\Phi(e^{1/C(E,F)}) = 1$ geometrically fast. Here, $C(E, F)$ denotes the capacity of the condenser (E, F). Thus, hyperbolic Menke points possess a better distribution on analytic Jordan curves than points of Fekete-type, which are called Tsuji-points in the hyperbolic situation [7, 8]. The latter has only been shown under additional assumptions so far.

The key to the presented proof is to utilize the connection between Green energy and the coefficients of the logarithmic development of functions univalent in an annulus. In particular, an extension of the Grunsky inequaltities to functions univalent in an annulus due to R. Kühnau is applied [1].

Finally, a pointwise geometrically fast approximation to the Green potential in $\mathcal{R} = \mathbb{D} \setminus E$ is derived and several numerical examples are presented.

References

[1] R. Kühnau. Koeffizientenbedingungen für schlicht abbildende Laurentsche Reihen. *Bull. Acad. Polon. Sciences, Sér. math., astr., phys.*, 20:7–10, 1972.

[2] K. Menke. *Extremalpunkte und konforme Abbildung.* PhD thesis, Technische Universität Berlin, 1970.

[3] K. Menke. Extremalpunkte und konforme Abbildung. *Math. Ann.*, 195:292–308, 1972.

[4] K. Menke. Bestimmung von Näherungen für die konforme Abbildung mit Hilfe von stationären Punktsystemen. *Num. Math.*, 22:111–117, 1974.

[5] K. Menke. Zur Approximation des transfiniten Durchmessers bei bis auf Ecken analytischen geschlossenen Jordankurven. *Israel Journ. Math.*, 17:136–141, 1974.

[6] K. Menke. Über die Verteilung von gewissen Punktsystemen mit Extremaleigenschaften. *Journal für die reine und angewandte Mathematik*, 283/284:421–435, 1976.

[7] K. Menke. On the distribution of Tsuji points. *Mathematische Zeitschrift*, 190:439–446, 1985.

[8] K. Menke. Tsuji points and conformal mapping. *Annales Polonici Mathematici*, XLVI:183–187, 1985.

[9] K. Menke. Point systems with extremal properties and conformal mapping. *Numerische Mathematik*, 54:125–143, 1988.

[10] C. Pommerenke. Über die Verteilung der Fekete-Punkte. *Math. Ann.*, 168:111–127, 1969.

[11] C. Pommerenke. Über die Verteilung der Fekete-Punkte II. *Math. Ann.*, 179:212–218, 1969.

Boundary Interpolation in the Theory of Nonlinear Riemann-Hilbert Problems

Gunther Semmler

We study Riemann-Hilbert problems for a holomorphic function w in the unit disc $\mathbb{D}$ with the boundary condition

$$w(t) \in M_t \tag{1}$$

for all $t \in \mathbb{T}$. The restriction manifold

$$M := \bigcup_{t \in \mathbb{T}} \{t\} \times M_t$$

is supposed to be smooth so that the existence of solutions that are continuous on the closed unit disc is secured by well-known theorems. Given k points $z_1, \ldots, z_k$ in the unit disc, there is exactly one solution of the boundary value problem (1) satisfying the side conditions

$$w(z_j) = w_j,\ j = 1, \ldots, k \qquad w(t_0) = w_0 \in M_{t_0}$$

The ambition of our research is to replace these conditions solely by interpolation points on M, i.e. we require

$$w(t_j) = w_j, \qquad j = 0, \ldots, k \tag{2}$$

where $t_j \in \mathbb{T}$ and $w_j \in M_{t_j}$ are given. As a generalization of a result by Ruscheweyh and Jones for Blaschke products, we show that the interpolation problem (2) has a solution with winding number at most k about M. This raises the question to determine a solution of (2) with minimal winding number about M. For three interpolation points we define the notion of counterclockwise turning around M with respect to the holomorphic parametrization, which allows to finally solve this problem. For more than three interpolation points, the situation is more involved. It turns out that we can distinguish three classes of problems which will be called rigid, fragil, and flexible. Problems in these classes have different properties concerning uniqueness and stability of solutions.

It is remarkable that also for finite Blaschke products (which solve the most simple Riemann-Hilbert problem where $M_t = \mathbb{T}$), no solvability criterium for (2) is known. In order to find at least an algorithmic approach we transformed this problem to an interpolation problem for a rational funtion on the real line, the numerator and denumerator polynomial of which have the interlacing property.

References

[1] Semmler, G. and E. Wegert: Nonlinear Riemann-Hilbert problems and Boundary Interpolation. To appear in Comp. Meth. Func. Theory

[2] Wegert, E. Nonlinear Boundary Value Problems for Holomorphic Functions and Singular Integral Equations. Akademie Verlag Berlin, 1992.

Restriction operators on Bergman space

Mihai Putinar

(joint work with B. Gustafsson and H.S. Shapiro)

Let Ω be a bounded planar domain and let $A^2(\Omega)$ be the associated Bergman space (of analytic square integrable functions). For a positive measure μ, compactly supported by Ω we consider the restriction operator:

$$R : A^2(\Omega) \longrightarrow L^2(\mu), \; Rf = f|_{\,\mathrm{supp}\mu}.$$

It is a trace class operator, whose modulus square R^*R has a complete system of eigenvectors $f_k \in A^2(\Omega)$, corresponding to a descending sequence of eigenvalues λ_k (after putting aside the null vectors). The typical eigenvalue problem for f_k can be written as an integral equation:

$$\lambda_k f_k(z) = \int K(z,w) f_k(w) d\mu(w).$$

This shows that each function f_k analytically extends across the boundary of Ω.

The system of functions f_k is doubly orthogonal with respect to the two inner products:

$$\lambda_k \langle f_k, f_m \rangle_{2,\Omega} = \delta_{km} \langle f_k, f_m \rangle_\mu.$$

Such doubly orthogonal systems have appeared a long time ago in function theory and approximation theory. Most of the references below illustrate such instances.

We are interested in qualitative properties of the eigenfunctions f_k. A central result in this direction is the following.

Theorem. *Let Ω be a bounded domain with smooth boundary, such that its Green function of the bi-Laplacian(associated to an arbitrary point of the boundary) is non-negative. Let $H(z,w)$ denote the reproducing kernel for all harmonic, square integrable functions in Ω, and assume that the positivity set: $P = \{z \in \Omega;\ H(z,w) > 0,\ w \in \partial\Omega\}$ is non-empty.*

Suppose that the positive measure μ is supported by a compact subset of P. Then each eigenfunction f_k does not vanish on the boundary of Ω and it possesses exactly k zeros in Ω.

For instance, if $\Omega = \mathbf{D}$ is the unit disk, then the conditions of the theorem are met for the set $P = \{z; |z| < \sqrt{2} - 1\}$. The analogous theorem for restrictions from the Hardy space was discovered by Fisher and Micchelli [7] and it played an important role in best approximation results and estimates on n-widths.

The proof of the theorem is based on potential theoretic techniques, starting from the observation that each eigenfunction f_k satisfies the balayage identity:

$$\lambda_k \int_\Omega |f_k(z)|^2 u(z) dArea(z) = \int |f_k(z)|^2 u(z) d\mu(z),$$

valid for an arbitrary harmonic function u, defined on a neighborhood of the closure of Ω.

This particular framework of doubly orthogonal systems can be used to estimate the growth of the contractive divisors in the Bergman space, best approximation in the $L^2(\mu)$ norm with control of the $L^2(\Omega, dArea)$ norm or exact identification of the inner measure μ from the matricial elements of the restriction operator.

This is a report on results published in:

B. Gustafsson, M. Putinar and H.S.Shapiro *Restriction operators, balayage and doubly orthogonal systems of analytic functions*, J. Funct. Analysis **199**(2003), 332-378.

References

[1] M.E.Andersson, *An inverse problem connected to doubly orthogonal sequences in Bergman space*, Math. Proc. Camb. Philo. Soc. **128**(2000), 535-538.

[2] S.Bergman, *The Kernel Function and Conformal Mapping*, Amer. Math. Soc., Providence, R.I., 1970.

[3] V.G.Cherednichenko, *Inverse Logarithmic Potential Problem*, VSP Publishers, Utrecht, The Netherlands, 1996.

[4] Ph.Davis, *An application of doubly orthogonal functions to a problem of approximation in two regions*, Trans. Amer. Math. Soc. **72**(1952), 104-137.

[5] P.Duren, D.Khavinson, H.S.Shapiro, C.Sundberg, *Contractive zero-divisors in Bergman space*, Pacific J. Math. **157**(1993), 37-56.

[6] S.Fisher, *Function Theory on Planar Domains*, Wiley Interscience, 1983.

[7] S.D. Fisher, C.A. Micchelli, *The n width of sets of analytic functions*, Duke Math. J. **47**(1980), 789-801.

[8] B. Gustafsson, *Existence of weak backward solutions to a generalized Hele-Shaw flow moving boundary problem*, Nonlinear Analysis **9**(1985), 203-215.
[9] B.Gustafsson, *On quadrature domains and an inverse problem in potential theory*, J. d'Analyse Math. **55**(1990), 172-216.
[10] B. Gustafsson, M.Sakai, H.S.Shapiro, *On domains in which harmonic functions satisfy generalized mean value properties*, Potential Analysis **7**(1997), 467-484.
[11] I.F. Krasichkov, *Systems of functions with double orthogonality property* (in Russian), Mat. Zametki **4**(1968), 551-557.
[12] V.L. Oleinik, *Estimates of the widths of compact sets of analytic functions in L^p with a weight* (in Russian), Vestnik Leningrad Univ. Math. **8**(1980), 219-224.
[13] O. Parfenov, *Asymptotics of the singular numbers of the embedding operators for certain classes of analytic functions* (in Russian), Mat. Sbornik **4**(1981), 632-641.
[14] J.R.Partington, *Interpolation, Identification, and Sampling*, Oxford University Press, New York, 1997.
[15] M. Putinar, H.S.Shapiro, *The Friedrichs operator of a planar domain*, Operator Theory: Advances and Applications **113**(2000), 303-330; part II, Operator Theory: Advances and Applications (2001), to appear.
[16] K.Seip, *Reproducing formulas and double orthogonality in Bargmann and Bergman spaces*, Preprint **4**(1989), Department of Mathematics, University of Trondheim, 27 pp.
[17] H.S.Shapiro, *Stefan Bergman's theory of doubly-orthogonal functions. An operator theoretic approach*, Proc. Royal Irish Acad. **79:6**(1979), 49-58.
[18] H.S.Shapiro, *Reconstructing a function from its values on a subset of its domain- a Hilbert space approach*, J. Approx. Theory **46**(1986), 385- 402.
[19] H.S.Shapiro, *The Schwarz Function and its Generalization to Higher Dimensions*, Wiley-Interscience, New York, 1992.
[20] A.A.Shlapunov, N.N.Tarkhanov, *Bases with double orthogonality in the Cauchy problem for systems with injective symbols*, Proc. London Math. Soc. **71**(1995), 1-52.
[21] G. Szegö, *A problem concerning orthogonal polynomials*, Trans. Amer. Math. Soc. **37**(1935), 196-206.
[22] G. Szegö, *On some Hermitian forms associated with two given curves of the complex plane*, Trans. Amer. Math. Soc. **40**(1936), 450-461.
[23] J.L.Walsh, *Note on the orthogonality of Tchebycheff polynomials on confocal ellipses*, Bull. Amer. Math. Soc. **40**(1934), 84-88.
[24] J.L.Walsh, G.M.Merriman, *Note on the simultaneous orthogonality of harmonic polynomials on several curves*, Duke Math. J. **3**(1937), 279-288.

Measurable dynamics of transcendental entire functions on their Julia sets

Jan-Martin Hemke

One of the main ideas in complex dynamics is to divide the plane into the Fatou set of points, where the iterates behave stable, i.e. where they form a normal family, and its complement, the Julia set. By definition the dynamics in the Fatou set is easier and understood very well. We are interested in the dynamics of meromorphic functions on their Julia set and study it in terms of the Lebesgue measure. In [1] H. Bock proved, that for any non-constant meromorphic function, which is defined on the whole complex plane, one of the two following cases holds:

(1) The Julia set is the entire plane and almost every orbit is dense in the sphere $\hat{\mathbb{C}}$;

(2) almost every forward-orbit in the Julia set accumulates only in the post-singular set.

Here the post-singular set denotes the closure of the union of the forward-orbits of all singularities of the inverse function, which are the critical and asymptotic values. This result is a generalization of similar results for rational functions obtained by M. Lyubich [8] and C. McMullen [10].
It is natural to ask for a given function, which case holds. Since a non-empty Fatou set always implies (ii), one can restrict to the cases, in which the Julia set consists of the whole complex plane. If the Julia set is not the entire plane, and thus (ii) holds, it would still be interesting to know if the Julia set has positive measure, since otherwise the statement (ii) would be trivial.
In the paper mentioned H. Bock gives sufficient conditions for (i): If f is entire and the set of singularities of the inverse function is finite, all of these are pre-periodic but not periodic, then (i) is satisfied. Thus the function $f(z) = 2\pi i \exp(z)$ is an example for this first case, in which the post-singular set consists of the only asymptotic value zero and its image $2\pi i$. Other conditions concerning this case are given by L. Keen and J. Kotus [4]. Conversely it was already shown in 1984 independently by M. Rees [6] and M. Lyubich [7] that the function $f(z) = \exp(z)$ is an example for (ii). Here the post-singular set consists of the the closure of the forward-orbit of the only asymptotic value zero, which tends to infinity on the real axis. This result was generalized in [11] to functions $f_\lambda(z) = \lambda \exp(z)$, if $f_\lambda^n(0)$ tends to infinity sufficiently fast. M. Urbanski and A. Zdunik [3] even showed, that the Hausdorff-dimension of the remaining set is smaller than 2.
The difference between the dynamics of $\exp(z)$ and $2\pi i \exp(z)$ is caused by the different behavior of the asymptotic value zero under iteration. We consider functions of the type $f(z) = \int_0^z P(t)\exp(Q(t))dt + c$, with polynomials P and Q and $c \in \mathbb{C}$, such that Q is not constant and P not zero. Counting multiplicity these functions have exacty $\deg(Q)$ asymptotic values and $\deg(P)$ critical points and may even be characterized as those entire functions with this property. In the extremal case that all singularities of the inverse are pre-periodic but not periodic, the theorem of H. Bock implies (i). We consider the other extreme and may neglect the critical values but have to specify the speed of escape. We assume that every asymptotic values s escapes exponentially fast, i.e. that $|f^n(s)| \geq \exp(|f^{n-1}(s)|^\delta)$ for some $\delta > 0$ and almost all $n \in \mathbb{N}$. Then we can prove that the Julia set has positive measure and that (ii) is satisfied. If the degree of Q is at least three, using an argument introduced by H. Schubert in [13], we obtain that the measure of the Fatou set is even finite.

References

[1] Bock, H., Über das Iterationsverhalten meromorpher Funktionen auf der Juliamenge, Dissertation, Aachener Beiträge zur Mathematik, 23, 1998
[2] Brolin, H, Invariant sets under iteration of rational functions, Ark. Math. 6, 103-144, 1965
[3] Urbanski, M. and Zdunik, A., Geometry and ergodic theory of non-hyperbolic exponential maps, Preprint, 2003

[4] Keen, L and Kotus, J, Ergodicity of some classes of meromorphic functions, Ann. Acad. Sci. Fenn., 24, 133-145, 1999

[5] Eremenko, A. and Lyubich, M., Dynamical Properties of some classes of entire Functions , Ann. Inst. Fourier, 42, 889-1019, 1992

[6] Rees, M., The Exponetial Map is Not Recurrent, Math. Z., 191, 593-598, 1986

[7] Lyubich, M., The measurable dynamics of the exponential, Sibirsk. Mat. Zh., 28, 111-127, 1987

[8] Lyubich, M., On typical behaviour of the trajectories of a rational mapping on the sphere, Soviet Math. Dokl., 27, 22-25, 1983

[9] McMullen, C., Area and Hausdorff dimension of Juliasets of entire functions, J. Amer. Math. Soc., 300, 329-342, 1987

[10] McMullen, C. T., Complex Dynamics and Renormalization, Princeton University Press, 1994

[11] Hemke, J-M., Typische Orbits der Exponentialfamilie, Diplomarbeit, CAU-Kiel, 2002

[12] Rempe, L., Dynamics of Exponential Maps, Dissertation, CAU-Kiel, 2003

[13] Schubert, H., Über das Maß der Fatoumenge trigonometrischer Funktionen, Diplomarbeit, CAU-Kiel, 2003

On Periodic Rays of Certain Entire Functions

Lasse Rempe

A well-known theorem of Douady and Hubbard [M, Theorem 18.10] states that *periodic dynamic rays* of polynomials always have a periodic landing point. This result forms the basis of the combinatorial methods which have been an essential ingredient in the success story of polynomial dynamics since the early studies of the Mandelbrot set [DH].

In this talk, we will consider the analogous question for periodic rays of transcendental entire functions. For our purposes, a *periodic dynamic ray* of an entire function $f : \mathbb{C} \to \mathbb{C}$ is a maximal curve

$$\gamma : (t_0, \infty) \to I(f) := \{z \in \mathbb{C} : |f^n(z)| \to \infty\}$$

which satisfies $f^n(\gamma(t)) = \gamma(t+1)$ for some $n \geq 1$ and all $t > t_0$. (Here $t_0 \in [-\infty, \infty)$.) As usual, we say that γ *lands* at a point $z_0 \in \mathbb{C}h$ if $\lim_{t \to t_0} \gamma(t) = z_0$.

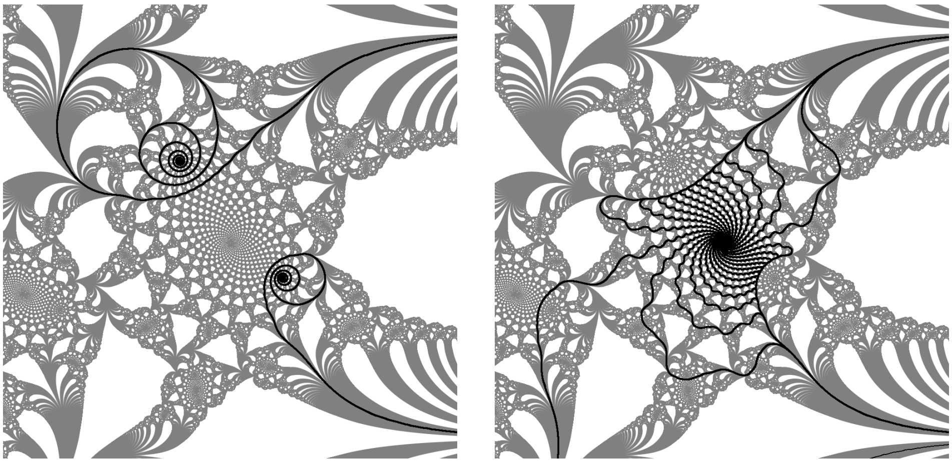
For the family of *exponential maps*[1]

$$E_\kappa k : z \mapsto \exp(z) + \kappa,$$

landing behavior of periodic rays has recently been used to great advantage by Schleicher (see e.g. [S2, RS]). However, it was previously not known whether periodic rays of exponential maps always land. We can now answer this question.

Theorem 1 (Periodic rays land [R1]**)** *Every periodic ray of every exponential map lands.*

[1]This family forms the simplest parameter space of transcendental entire functions, as exponential maps are the only such functions with only one singular value. Also, the exponential family can be considered to be the limit of the families of *unicritical polynomials*, $z \mapsto z^d + c$ [BDG], which are by far the best-understood polynomial families.



(a) (b)

FIGURE 3. Periodic rays for $z \mapsto \exp(z) + 1.0038 + 2.8999i$. (a) shows two rays forming a period 2-cycle; (b) shows a cycle of 25 rays landing at a common fixed point.

The proof of Douady and Hubbard's landing theorem for polynomials uses a hyperbolic contraction principle, and this argument can be carried over to several situations in which there is some form of expansion along the ray. However, it is conceivable that a periodic ray γ might accumulate on a singular value, whose orbit again accumulates everywhere on γ. In such a situation, a proof by hyperbolic contraction would be impossible. Thus, in order to apply this method to maps with large postsingular sets, it seems that one must *a priori* show that the given ray does not accumulate on singular values. The problem is that it can be very difficult to control the accumulation behavior of these rays; even for many tame exponential maps, there are many (nonperiodic) dynamic rays with complicated accumulation behavior [DJ, R2].

Our proof of Theorem 1 circumvents these difficulties by using a theorem of Schleicher [S1] on landing properties of *parameter* rays.[2] However, there is little hope for this method to generalise to higher-dimensional parameter spaces. For example, we currently know of no argument which would prove the analogue of Theorem for *cosine maps*,

$$z \mapsto a \exp(z) + b \exp(-z),$$

[2]Thus, we are reversing Douady's famous principle: we plough in the parameter plane to harvest in the dynamical plane.

where $a, b \in \mathbb{C}$. (Many results for the exponential family are known to generalise to this two-dimensional family; in particular, there is a complete classification of escaping points in terms of dynamic rays [Ro].)

On the other hand, we were able to show that the above problem is indeed the only obstruction for a large set of functions in the class

$$\mathcal{B} := \{f : \mathbb{C} \to \mathbb{C} \text{ entire}; \operatorname{sing}(f^{-1}) \text{ is bounded}\}.$$

Theorem 2 (Periodic rays with nonsingular accumulation sets [R3]**)** *Let f be either*

- *a cosine map $z \mapsto a\exp(z) + b\exp(-z)$ or*
- *a function $f \in \mathcal{B}$ all of whose singular values lie in the Julia set.*

If γ is a fixed dynamic ray of f which has no accumulation points in $\overline{\operatorname{sing}(f^{-1})}$), then γ lands.

References

[BDG] Clara Bodelón, Robert L. Devaney, Michael Hayes, Gareth Roberts, Lisa R. Goldberg, and John H. Hubbard, *http://math.bu.edu/people/bob/papers/hairs-2.ps* Dynamical convergence of polynomials to the exponential, J. Differ. Equations Appl. **6** (2000), no. 3, 275–307.

[DJ] Robert L. Devaney and Xavier Jarque, *http://math.bu.edu/people/bob/papers/inde.ps* Indecomposable continua in exponential dynamics, Conform. Geom. Dyn. **6** (2002), 1–12.

[DH] Adrien Douady and John Hubbard, *Etude dynamique des polynômes complexes*, Prépublications mathémathiques d'Orsay (1984 / 1985), no. 2/4.

[M] John Milnor, *Dynamics in one complex variable*, Friedr. Vieweg & Sohn, Braunschweig, 1999, http://www.arXiv.org/abs/math.DS/9201272arXiv:math.DS/9201272.

[R1] Lasse Rempe, *A landing theorem for periodic rays of exponential maps*, Manuscript, 2003, http://www.arXiv.org/abs/math.DS/0307371arXiv:math.DS/0307371, submitted for publication.

[R2] Lasse Rempe, *On exponential maps with accessible singular value*, Manuscript, 2003.

[R3] Lasse Rempe, *Siegel disks and periodic rays of entire functions*, Manuscript, 2003.

[RS] Lasse Rempe and Dierk Schleicher, *Bifurcations in the space of exponential maps*, Manuscript, 2003, http://www.arXiv.org/abs/math.DS/0311480arXiv:math.DS/0311480, to appear in the Stony Brook IMS preprint series.

[Ro] Günter Rottenfußer, *Entkommende Punkte einer Familie ganzer transzendenter Funktionen*, Diplomarbeit, TU München, 2002.

[S1] Dierk Schleicher, *On the dynamics of iterated exponential maps*, Habilitation thesis, TU München, 1999.

[S2] Dierk Schleicher, *Attracting dynamics of exponential maps*, Ann. Acad. Sci. Fenn. Math. **28** (2003), 3–34.

[SZ1] Dierk Schleicher and Johannes Zimmer, *Escaping points of exponential maps*, J. London Math. Soc. (2) **67** (2003), no. 2, 380–400.

[SZ2] Dierk Schleicher and Johannes Zimmer, *Periodic points and dynamic rays of exponential maps*, Ann. Acad. Sci. Fenn. Math. **28** (2003), 327–354.

On the zeros of the solutions of a functional equation

Walter Hayman

We consider an entire function

$$f(z) = \sum_{0}^{\infty} a_n z^n$$

satisfying the equation

$$(a - qz)f(q^2 z) - (1 + a)f(qz) + f(z) = 0\,, \quad 0 < |q| < 1.$$

Let z_n be the nth zero of $f(z)$ in order of nondecreasing moduli. Then

$$z_n = -q^{(1-2n)} \left\{ 1 + \sum_{\nu=1}^{k} b_\nu q^{n\nu} + O(|q|^{(k+1)n}) \right\},$$

where the b_ν are constants depending on a and q. This verifies a conjecture of Mourad Ismail [1], concerning the zeros of q–Bessel functions. The above result also contains as a special case an identity of Ramanujan [4].

The method builds on an earlier paper by Walter Bergweiler and the author [3] which applies to a wider class of functional equations but gives only the first term in the asymptotic series. In this case the zeros may approach a finite number of distinct geometric progressions. We compare the coefficients of $f(z)$ and so $f(z)$ itself with certain theta–functions.

References

[1] Letter to the author

[2] Ismail, Mourad E.H., The zeros of basic Bessel functions, the functions J(nu+ax)(x), and associated orthogonal polynomials, J. Math. Anal. Appl. 86, 1-19 (1982).

[3] Walter Bergweiler and Walter K. Hayman CMFT 3 (2003), 55–78.

[4] Third identity on p. 57 of Ramanujan's last notebook, see Bruce C. Berndt, Ramanujan's notebooks. Part I, Springer-Verlag, New York, 1985.

On the number of zeros of certain rational harmonic functions

Dmitry Khavinson
(joint work with Genevra Neumann)

A. Wilmshurst [Wil 98] showed that there is an upper bound on the number of zeros of a harmonic polynomial $f(z) = p(z) - \overline{q(z)}$, where p and q are analytic polynomials of different degree, answering the question of T. Sheil-Small [SS 92]. Let $n = \deg p > \deg q = m$. Wilmshurst showed that n^2 is a sharp upper bound when $m = n-1$ and conjectured that the upper bound is actually $m(m-1)+3n-2$. D. Khavinson and G. Świątek [KS 03] showed that Wilmshurst's conjecture holds for the case $n > 1, m = 1$ using methods from complex dynamics. When hearing of this result, P. Poggi-Corradini asked whether this approach can be extended to the case $f(z) = p(z)/q(z) - \overline{z}$, where p and q are analytic polynomials.

In this note, we apply the approach from [KS 03] to prove

Theorem *Let $r(z) = p(z)/q(z)$ be a rational function where p and q are relatively prime, analytic polynomials and such that $n = \deg\ r = \max(\deg\ p, \deg\ q) > 1$. Then*

$$\#\{z \in \mathbb{C} : \overline{r(z)} = z\} \leq 5n - 5$$

We note that the zeros of $\overline{r(z)} - z$ are isolated, because each zero is also a fixed point of $Q(z) = \overline{r(\overline{r(z)})}$, an analytic rational function of degree n^2. This also follows from a result of P. Davis [Da 74] (Chapter 14) concerning the Schwarz functions of analytic curves. (The Schwarz function is an analytic function $S(z)$ that gives the equation of a curve in the form $\overline{z} = S(z)$, *cf.* [Da 74].) A rational Schwarz function implies that the curve is a line or a circle, so the degree must be one.

We also note that $\overline{r(z)} - z$ will not have a zero at ∞.

L. Geyer [Ge 03] has recently shown that the $3n - 2$ bound on the number of zeros of $f(z) = p(z) - \overline{z}$ where $\deg p = n$ is sharp for all $n > 1$. D. Bshouty and A. Lyzzaik [BL 03] have recently given an elementary proof for $n = 4, 5, 6, 8$. Hence, a sharp bound on the number of zeros of $f(z) = \overline{r(z)} - z$ must be at least $3n - 2$.

Let us discuss applications of the result to gravitational microlensing. An n-point gravitational lens can be modeled as follows: Suppose that we have n point masses (such as stars). Construct a plane through the center of mass of these point masses, such that the line of sight from the observer to the center of mass is orthogonal to this plane. This plane is called the lens plane (or deflector plane). Suppose that the lens plane is between the observer and the light source. (We are assuming that the distance between the point masses is small compared to the distance between the observer and the lens plane, as well as the distance between the lens plane and the light source.) The plane containing our light source which is parallel to the lens plane is called the source plane. Due to deflection of light by masses multiple images of the light source are formed. This phenomenon is known as gravitational microlensing and is modeled by a lens equation. The lens equation defines a mapping from the lens plane to the source plane. Suppose that our light source is located at postion w in the source plane. In this model, if z satisfies the lens equation, then our gravitational lens will map z to w; hence z corresponds to the position of a lensed image. The number of lensed images is the number of solutions of the lens equation. See [Wa 98] for an introduction to gravitational lensing and [St 97] for an introduction to a complex formulation of lensing theory.

To set up a lens equation for our n-point gravitational lens, the point masses are projected onto positions in the lens plane. The projection of the j-th point mass has a scaled mass of m_j and is located at a scaled coordinate of z_j in the lens plane, where m_j is a positive constant and z_j is a complex constant. Suppose that we have a light source located at a scaled coordinate of w in the source plane.

Following [Wit 90], this lens equation will be given by

$$w = z + \gamma\overline{z} - sign(\sigma)\Sigma_{j=1}^{n} m_j/(\overline{z} - \overline{z_j}),$$

where the normalized shear γ and the optical depth (or normalized surface density) $\sigma \neq 0$ are real constants. See [Wit 90] and [Pa 86] for a derivation of the normalized lens equation for microlensing.

We can rewrite this lens equation in terms of the rational harmonic function $f(z) = \overline{r(z)} - z$ by letting $r(z) = \overline{w} - \gamma z + \mathrm{sign}(\sigma)\Sigma_{j=1}^{n} m_j/(z - z_j)$. We thus see that the zeros of $f(z)$ are solutions of the lens equation for a light source at position w. H. Witt [Wit 90] showed for $n > 1$ that the maximum number of observed images is at most $n^2 + 1$ when $\gamma = 0$ and $(n+1)^2$ when $\gamma \neq 0$. S. H. Rhie [Rh 01] conjectured that for $n > 1$ such a gravitational lens gives at most $5n - 5$ images for the case $\gamma = 0$ and $\sigma > 0$. In the $\gamma = 0$ case, deg $r = n$; hence, our theorem settles this conjecture. Further, for the case $\gamma \neq 0$, we see that deg $r = n + 1$, so our theorem gives an upper bound of $5(n+1) - 5 = 5n$ lensed images.

References

[BL 03] D. Bshouty and A. Lyzzaik, *On Crofoot-Sarason's conjecture for harmonic polynomials*, preprint (2003).

[Bu 81] W. L. Burke, *Multiple gravitational imaging by distributed masses*, Astrophys. J. **244** (1981), L1.

[CG 93] L. Carleson and T. Gamelin, *Complex Dynamics*, Springer-Verlag, New York-Berlin-Heidelberg (1993). MR 94h:30033.

[Da 74] P. J. Davis, *The Schwarz function and its applications*, The Carus Mathematical Monographs, No. 17, The Mathematical Association of America, Buffalo, N. Y., 1974. MR 53 #11031.

[Fo 81] O. Forster, *Lectures on Riemann Surfaces*, Translated from the German by Bruce Gilligan, Graduate Texts in Mathematics, 81, Springer-Verlag, New York-Berlin (1981). MR 83d:30046.

[Ge 03] L. Geyer, *Sharp bounds for the valence of certain harmonic polynomials*, preprint (2003).

[KS 03] D. Khavinson and G. Świątek, *On the number of zeros of certain harmonic polynomials*, Proc. Amer. Math. Soc. **131** (2003), 409-414. MR 2003j:30015.

[MPW 97] S. Mao, A. O. Petters, and H. J. Witt, *Properties of point mass lenses on a regular polygon and the problem of maximum number of images*, 'Proceeding of the Eighth Marcel Grossmann Meeting on General Relativity (Jerusalem, Israel, 1997)', edited by T. Piran, World Scientific, Singapore (1998), 1494-1496, arXiv:astro-ph/9708111.

[Pa 86] B. Paczyǹski, *Gravitational microlensing at large optical depth*, Astrophys. J. **301** (1986), 503-516.

[Rh 01] S. H. Rhie, *Can a gravitational quadruple lens produce 17 images?*, arXiv:astro-ph/0103463.

[Rh 03] S. H. Rhie, *n-point gravitational lenses with* $5(n-1)$ *images*, arXiv:astro-ph/0305166.

[SS 02] T. Sheil-Small, *Complex Polynomials*, Cambridge Studies in Advanced Mathematics 73, Cambridge University Press (2002). MR 2004b:30001.

[SS 92] T. Sheil-Small in *Tagesbericht*, Mathematisches Forsch. Inst. Oberwolfach, Funktionentheorie, 16-22.2.1992, 19.

[St 97] N. Straumann, *Complex formulation of lensing theory and applications*, Helvetica Physica Acta **70** (1997), 894-908, arXiv:astro-ph/9703103.

[ST 00] T. J. Suffridge and J. W. Thompson, *Local behavior of harmonic mappings*, Complex Variables Theory Appl. **41** (2000), 63-80. MR 2001a:30019.

[Wa 98] J. Wambsganss, *Gravitational lensing in astronomy*, Living Rev. Relativity **1** (1998) [Online Article]: cited on January 26, 2004, http://www.livingreviews.org/Articles/Volume1/1998-12wamb/.

[Wil 98] A. S. Wilmshurst, *The valence of harmonic polynomials*, Proc. Amer. Math. Soc. **126** (1998), 2077-2081. MR 98h:30029.

[Wit 90] H. J. Witt, *Investigation of high amplification events in light curves of gravitationally lensed quasars*, Astron. Astrophys. **236** (1990), 311-322.

An extension of the Schwarz–Carathéodory reflection principle

Oliver Roth

1. A reflection principle for conformal metrics

Let $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$ denote the open unit disk in the complex plane $\mathbb{C}$. An open subarc of the unit circle $\partial\mathbb{D} := \{z \in \mathbb{C} : |z| = 1\}$ is an open connected proper subset of $\partial\mathbb{D}$.

Theorem 1 *Let I be an open subarc of $\partial\mathbb{D}$ and let R be a Riemann surface, which carries a complete real analytic conformal Riemannian metric $\lambda(w)\,|dw|$. Then a non–constant analytic map $f : \mathbb{D} \to R$ can be continued analytically across I with $f(I) \subset R$ if and only if there exists a holomorphic function $h : I \to \mathbb{C}$ such that*

$$\lim_{z\to\xi} \frac{\lambda(f(z))\,|f'(z)|}{|h'(z)|} = 1, \qquad \xi \in I. \tag{1}$$

Remarks.

(a) Note that $\lambda(f(z))\,|f'(z)|$ in (1) is the pullback of the metric $\lambda(w)\,|dw|$ under the map f. Hence $\lambda(f(z))\,|f'(z)|$ is a well-defined function on $\mathbb{D}$.

(b) The phrase "$f : \mathbb{D} \to R$ can be continued analytically across I with $f(I) \subset R$" means there exists a domain $\Omega \supset \mathbb{D}$ with $I \subset \Omega$ and an analytic map $F : \Omega \to R$ such that $F = f$ in $\mathbb{D}$. This map F is the unique analytic continuation of f to Ω.

(c) A function $h : M \to \mathbb{C}$ is said to be holomorphic on a set $M \subseteq \mathbb{C}$, if it is defined and holomorphic in an open set $V \subseteq \mathbb{C}$ containing M.

(d) The special case $R = \mathbb{C}$ and $\lambda(w) = 1$ of Theorem 1 may be regarded as a version of the classical Schwarz–Carathéodory reflection principle [3, 7] for holomorphic functions $f : \mathbb{D} \to \mathbb{C}$. Just as with the Schwarz–Carathéodory reflection principle, Theorem 1 readily generalizes to non–constant analytic maps $f : D \to R$, where (i) D is a domain in $\mathbb{C}$ with an open free analytic boundary arc I or (ii) D is a bordered Riemann surface with border Γ and $I \subset \Gamma$.

(e) For the special case $R = \mathbb{D}$ and $\lambda(w) = 1/(1 - |w|^2)$ Theorem 1 reduces to the Fournier–Ruscheweyh reflection principle [4, 5].

(f) The restraint in Theorem 1 that $\lambda(w)\,|dw|$ is a *complete* and *real analytic* conformal Riemannian metric can slightly be relaxed. For the 'if' part it suffices to assume $\lambda(w)\,|dw|$ is a complete conformal Riemannian metric, which is real analytic in a neighborhood $U \subset R$ of $f(I)$. For the 'only if' part we need only $\lambda(w)\,|dw|$ is real analytic in a neighborhood of $f(I)$. These assumptions cannot further be weakened.

2. Analytic continuation of Beurling–Riemann maps

In 1953 Arne Beurling [2] proved the following extension of the Riemann mapping theorem[3].

Theorem A *Let $\Phi(w)$ be a positive, continuous and bounded function defined for $|w| < \infty$ and let w_0 be a given point in the w-plane. Then there exists an analytic and univalent function $f : \mathbb{D} \to \mathbb{C}$ normalized by*

$$f(0) = w_0, \qquad f'(0) > 0, \tag{2}$$

and satisfying the non–linear boundary condition

$$\lim_{|z|\to 1} \left(|f'(z)| - \Phi(f(z))\right) = 0. \tag{3}$$

Moreover, if $\log \Phi(w)$ *is superharmonic, then there is exactly one such function.*

We call any normalized, analytic and univalent function $f : \mathbb{D} \to \mathbb{C}$ satisfying (3) a *Beurling–Riemann mapping function* (for $\Phi(w)$). Note that every Beurling–Riemann mapping function $f(z)$ is a Lipschitz map from $(\mathbb{D}, |\cdot|)$ to $(\mathbb{C}, |\cdot|)$,

$$|f(z_1) - f(z_2)| \leq M \cdot |z_1 - z_2|, \qquad z_1, z_2 \in \mathbb{D},$$

with $M := \sup_{w\in\mathbb{C}} \Phi(w) < \infty$. Hence $f(z)$ has a continuous extension to $\overline{\mathbb{D}}$, and $\partial f(\mathbb{D})$ is a closed curve, which admits the conformal parametrization

$$\partial f(\mathbb{D}) : \quad f(e^{it}), \qquad 0 \leq t \leq 2\pi.$$

Moreover, $|f'(z)|$ has a continuous extension to $\overline{\mathbb{D}}$ with $|f'(z)| \neq 0$.

If a Beurling–Riemann mapping function can be continued analytically across an open subarc I of the unit circle, then the corresponding function $\Phi(f(z))$ will be real analytic on I since $\Phi(f(z)) = |f'(z)| > 0$ there. A partial converse is given by the following theorem, which is essentially another special case of Theorem 1.

Theorem 2 *Let $\Phi(w)$ be a positive, continuous and bounded function defined for $|w| < \infty$, let w_0 be a given point in the w-plane, and let $f(z)$ be a Beurling–Riemann mapping function for $\Phi(w)$ normalized by (2). If $\Phi(w)$ is real analytic in a neighborhood of $f(I)$ for some open subarc I of the unit circle, then $f(z)$ has an analytic continuation across I.*

In particular, if $\Phi(w)$ is real analytic in a neighborhood of $\partial f(\mathbb{D})$, then every Beurling–Riemann mapping function for $\Phi(w)$ has an analytic extension to some

[3] See [1, 4, 6] for recent generalizations of and variations on Beurling's theorem.

disk $|z| < \rho$, $\rho > 1$. Hence, at least in this special case, the analytic properties of the function $\Phi(w)$ are reflected by the analytic properties of the corresponding mapping functions.

References

[1] Farid G. Avkhadiev, Conformal mappings that satisfy the boundary condition of equality of metrics, *Doklady Akad. Nauk.* (1996), **347** no. 3, 295–297. English transl.: *Doklady Mathematics.* (1996), **53** no. 2, 194–196.

[2] Arne Beurling, An extension of the Riemann mapping theorem, *Acta Math.* (1953), **90**, 117–130.

[3] Constantin Carathéodory, Zum Schwarzschen Spiegelungsprinzip (Die Randwerte von meromorphen Funktionen), *Comment. Math. Helv.* (1946), **46**, 263–278.

[4] Richard Fournier and Stephan Ruscheweyh, Free boundary value problems for analytic functions in the closed unit disk, *Proc. Amer. Math. Soc.* (1999), **127** no. 11, 3287–3294.

[5] Richard Fournier and Stephan Ruscheweyh, A generalization of the Schwarz–Carathéodory reflection principle and spaces of pseudo–metrics, *Math. Proc. Cambridge Phil. Soc.* (2001), **130**, 353–364.

[6] Reiner Kühnau, Längentreue Randverzerrung bei analytischer Abbildung in hyperbolischer und sphärischer Geometrie, *Mitt. Math. Sem. Giessen* (1997), **229**, 45–53.

[7] Hermann Amandus Schwarz, Über einige Abbildungsaufgaben, *J. Reine Angew. Math.* (1869), **70**, 105–120.

Schwarzians of Hyperbolically Convex Functions

G. Brock Williams

(joint work with Roger W. Barnard, Leah Cole, and Kent Pearce)

The Schwarzian derivative S_f of an analytic function $f : \Omega \to \mathbb{C}$ is given by

$$S_f = \left(\frac{f''}{f'}\right)' - \frac{1}{2}\left(\frac{f''}{f'}\right)^2.$$

The Schwarzian itself contains a great deal of geometric information about the function f, but much more is encoded in the Schwarz norm

$$||S_f||_\Omega = \sup_{z\in\Omega}\{\eta_\Omega^{-2}(z)\,|S_f(z)|\},$$

where η_Ω is the hyperbolic density of Ω.

The Schwarz norm of f is completely Möbius invariant and is 0 if and only if f is a Möbius transformation. Thus the Schwarzian derivative provides an effective means of describing how much an analytic map differs from a Möbius transformation. For functions f defined on the unit disc $\mathbb{D}$, this also serves to describe how the range of f differs from a disc. Olli Lehto has made this notion precise, defining a pseudo-metric on the space of all simply connected proper subsets of $\mathbb{C}$ modulo Möbius transformations [1].

As a general principle, regions which are close to discs in Lehto's pseudo-metric share some of the properties of discs. Thus it is natural to ask "how far from a disc can a convex set be?" [6] For convex sets in euclidean geometry, this question was answered by Zeev Nehari who showed that if f is convex, then $||S_f||_\mathbb{D} \le 2$, with equality if and only if $f(\mathbb{D})$ is a euclidean strip [7]. Similarly, Diego Mejía and Christian Pommerenke proved that the extremal spherically convex domains are also strips [3].

In this talk, we complete the classification in all three classical geometries of the convex domains which are furthest from being a disc, by establishing the sharp upper bound on the Schwarz norm of functions from the disc onto hyperbolically convex regions. In particular, we show that the bound is attained by a map onto a domain bounded by two hyperbolic geodesics, a sort of "hyperbolic strip." This result had earlier been conjectured in several papers of Diego Mejía and Christian Pommerenke [2, 4, 5].

Our major tools are the Julia variation as extended by Roger Barnard and John Lewis, estimates on elliptic integrals, and a critical new Step Down Lemma. We formulate two new variations which preserve hyperbolic convexity. The first variation allows us to show there is an extremal domain with at most four sides. Our Step Down Lemma and the second variation then reduces the number of sides to at most two. We then directly compute the Schwarz norm for the remaining possibilities using special functions techniques.

This talk represents joint work with Roger W. Barnard, Leah Cole, and Kent Pearce.

References

[1] O. Lehto, Univalent Functions and Teichmüller Spaces, Springer-Verlag, 1987, Berlin - Heidelberg - New York

[2] Diego Mejía and Christian Pommerenke, Hyperbolically convex functions, dimension and capacity, Complex Var. Theory Appl., 47, 2002, 9, 803–814

[3] Diego Mejía and Christian Pommerenke, On spherically convex univalent functions, Michigan Math. J., 47, 2000, 1, 163–172

[4] Diego Mejía and Christian Pommerenke, On Hyperbolically Convex Functions, J. Geom. Anal., 2000, 10, 2, 365-378,

[5] Diego Mejía and Christian Pommerenke, Sobre la derivada Schawarziana de aplicaciones conformes hiperbólicamente, Revista Colombiana de Matemáticas, 2001, 35, 2, 51-60

[6] Ma, William and Minda, David, Hyperbolically convex functions. II, Ann. Polon. Math., 71, 1999, 3, 273–285

[7] Nehari, Zeev, A property of convex conformal maps, J. Analyse Math., 30, 1976, 390–393

Metric properties of Green's functions

Vilmos Totik

Extensions of the classical Markov inequality

$$\|P_n'\|_{[-1,1]} \le n^2 \|P_n\|_{[-1,1]}$$

(where P_n is a polynomial of degree at most n) to more general sets are closely related to smoothness of Green's functions. If E is a compact set on the plane, then the n-th Markoff constant M_n for E is defined as the smallest M_n for which

$$\|P_n'\|_E \le M_n \|P_n\|_E.$$

Let $g_{\mathbb{C}\setminus E}$ be the Green's function of the unbounded component of $\mathbb{C}\setminus E$ with pole at infinity (we assume that E is of positive logarithmic capacity). A standard way of estimating M_n is to use the Bernstein-Walsh lemma

$$|P_n(z)| \le e^{ng_{\mathbb{C}\setminus E}(z)}\|P_n\|_E, \qquad z\in\mathbb{C}$$

and then to use the Cauchy integral formula for the derivative of P_n. This approach gives e.g. that if $g_{\mathbb{C}\setminus E}$ is Hölder continuous: $g_{\mathbb{C}\setminus E}(z)\le C\ \mathrm{dist}(z,E)^\alpha$, then $M_n\le C'n^{1/\alpha}$. Thus, smoothness of Green's function implies a growth restriction on the Markov factors M_n. The converse is not clear, and in the talk first a situation is mentioned when the connection is completely known, and this is the case of Cantor type sets.

Let $\varepsilon_1,\varepsilon_2,\dots$ be a sequence from the interval $(0,1)$, and starting from $\mathcal{C}_0=[0,1]$ do the Cantor construction with the modification that at level n we remove the middle ε_n part of all remaining intervals. If $\mathcal{C}_n$ denotes the set after making n such steps, then $\mathcal{C}_n$ consists of 2^n intervals of total length $(1-\varepsilon_1)\cdots(1-\varepsilon_n)$. Consider the Cantor set $\mathcal{C}=\cap_n\mathcal{C}_n$. It is of measure zero if and only if $\sum_n\varepsilon_n=\infty$, and it is of positive capacity if and only if $\sum_k|\log(1-\varepsilon_k)|/2^k<\infty$ (see e.g. [5, Section V.6]). Now for Cantor sets we have (see [6], [7], [8])

(a): $M_n=e^{o(n)}\Longleftrightarrow g_{\mathbb{C}\setminus E}$ continuous $\Longleftrightarrow \sum_j 2^{-j}\log(1-\varepsilon_j)>-\infty$,
(b): $M_n=O(n^k)$ for some $k\Longleftrightarrow g_{\mathbb{C}\setminus E}\in$ Lip α for some $\alpha>0$
$\Longleftrightarrow \sum_{j=1}^n\log(1-\varepsilon_j)\ge -cn$,
(c): $M_n=O(n^2)\Longleftrightarrow g_{\mathbb{C}\setminus E}\in$ Lip$1/2\Longleftrightarrow\sum_j\varepsilon_j^2<\infty$.

Note that $M_n\ge cn^2$ and $g_{\mathbb{C}\setminus E}(-r)\ge cr^{1/2}$ for all $E\subseteq[0,1]$, i.e. the growth rates in (c) are optimal.

In the special case $\varepsilon_j=1/(j+1)$ we get a compact set $E\subset[0,1]$ of linear measure 0 such that $g_{\mathbb{C}\setminus E}\in$ Lip$1/2$ and $M_n=O(n^2)$.

As we can see, there is a big difference between the conditions on ε_j in (b) and (c). An explanation was given by V. Andrievskii [1] who proved that for $E\subset[0,1]$ the condition $g_{\mathbb{C}\setminus E}(z)\le C|z|^{1/2}$ implies that the set E is locally of full capacity at 0, i.e.

$$\lim_{t\to 0}\frac{\mathrm{cap}([0,t]\cap E)}{\mathrm{cap}([0,t])}=1.$$

Recently a characterization of optimal Hölder smoothness of Green's function was given by L. Carleson ([3]): for $E\subset[0,1]$ we have $g_{\mathbb{C}\setminus E}(z)\le C|z|^{1/2}$ if and only if $\sum_k\theta_k<\infty$, where with some $0<\varepsilon<1/3$

$$\theta_k=2^k\Big(\ \mathrm{cap}([0,2^{-k}])-\ \mathrm{cap}\big((E\cap[0,2^{-k}])\cup[0,\varepsilon 2^{-k}]\cup[(1-\varepsilon)2^{-k},2^{-k}]\big)\Big).$$

Returning to measuring density of sets with linear Lebesgue measure, T. Erdélyi, A. Kroó and J. Szabados [4] used for $E\subset[0,1]$ the function $\Theta_E(t)=|[0,t]\setminus E|$ to measure density, and they proved some local Markov inequalities in terms of this Θ_E. In [7] we used the same measure Θ_E (if E is not on $[0,1]$ then take its circular projection onto $\mathcal{R}_+$ and use the Θ function for the projected set), and proved that

$$g_{\mathbb{C}\setminus E}(z) \leq C\sqrt{|z|}\exp\left(C\int_{|z|}^{1}\frac{\Theta_E^2(u)}{u^3}du\right)\log\frac{2}{\mathrm{cap}(E)},$$

and this is sharp, for if $\Theta \nearrow$, $\Theta(t) \leq t$, then there is an $E \subset [0,1]$ such that $\Theta_E(t) \leq \Theta(t)$ and

$$g_{\mathbb{C}\setminus E}(-r) \geq c\sqrt{r}\exp\left(c\int_r \frac{\Theta^2(u)}{u^3}du\right).$$

This result was extended in [2] by V. Andrievskii.

Finally, we talk about characterization of Hölder continuity with some positive exponent in the spirit of Wiener's regularity test. Let E be a compact subset on the plane such that 0 is on the boundary of the unbounded component of $\mathbb{C}\setminus E$. With

$$E^n = \left\{z \in E \,\middle|\, 2^{-n} \leq |z| \leq 2^{-n+1}\right\}$$

the continuity of $g_{\mathbb{C}\setminus E}$ at 0 was characterized by Wiener (see e.g. [9, Theorem III.62]): $g_{\mathbb{C}\setminus E}$ is continuous at 0 if and only if

$$\sum_{n=1}^{\infty}\frac{n}{\log(1/\,\mathrm{cap}(E^n))} = \infty.$$

For $\varepsilon > 0$ set

$$\mathcal{N}_E(\varepsilon) = \{n \in \mathbb{N} \mid \ \mathrm{cap}(E^n) \geq \varepsilon 2^{-n}\},$$

and we say that a subsequence $\mathcal{N} = \{n_1 < n_2 < \ldots\}$ of the natural numbers is of positive lower density if

$$\liminf_{N\to\infty}\frac{|\mathcal{N}\cap\{0,1,\ldots,N\}|}{N+1} > 0,$$

which is clearly the same condition as $n_k = O(k)$. Now (see [3]) under the cone condition (i.e. there is a cone with vertex at 0 not intersecting E) Green's function $g_{\overline{\mathbb{C}}\setminus E}$ is Hölder continuous at 0 (i.e. $g_{\mathbb{C}\setminus E}(z) \leq C|z|^\alpha$ for some $\alpha > 0$) if and only if $\mathcal{N}_E(\varepsilon)$ is of positive lower density for some $\varepsilon > 0$. Here the cone condition cannot be omitted, but the rings $\{2^{-n} \leq |z| \leq 2^{-n+1}\}$ in the definition of E^n can be replace by the disks $\{|z| \leq 2^{-n}\}$.

References

[1] V. V. Andrievskii, The highest smoothness of the Green function implies the highest density of a set, *Arkiv för Math.* (to appear)

[2] V. V. Andrievskii, On the Green Function for a complement of a finite number of real intervals, *Constructive Approx.*, (to appear)

[3] L. Carleson and V. Totik, Hölder continuity of Green's functions, *Acta Sci. Math. (Szeged)* (to appear)

[4] T. Erdélyi, A. Kroó and J. Szabados, Markov–Bernstein type inequalities on compact subsets of **R**, *Analysis Math.*, **26**(2000), 17–34.

[5] R. Nevanlinna, *Analytic Functions*, Grundlehren der mathematischen Wissenschaften, **162**, Springer Verlag, Berlin, 1970

[6] W. Pleśnak, A Cantor regular set which does not have Markov's property, *Ann. Pol. Math.*, **LI**(1990), 269–274.
[7] V. Totik, *Metric Properties of Harmonic Measures*, (manuscript)
[8] V. Totik, Markoff constants for Cantor sets, *Acta Sci. Math. (Szeged)*, **60**(1995), 715–734.
[9] M. Tsuji, *Potential Theory in Modern Function Theory*, Maruzen, Tokyo, 1959

Random matrices in an external source and multiple orthogonal polynomials

Arno B.J. Kuijlaars
(joint work with Pavel Bleher)

We consider the random matrix ensemble

$$\frac{1}{Z_n} e^{-n\mathrm{Tr}(V(M)-AM)} dM \tag{1}$$

defined on $n \times n$ Hermitian matrices M, where A is a given Hermitian matrix, called the external source. The ensemble is unitary invariant if $A = 0$, and then the eigenvalue correlations can be described with orthogonal polynomials. The universal behavior of local eigenvalue statistics in the large n limit can then be obtained from precise asymptotic formulae for the orthogonal polynomials. This was done in [2, 8] with the steepest descent method for Riemann-Hilbert (RH) problems.

For a general external source A the ensemble (1) is not unitary invariant. Suppose A has p distinct eigenvalues $a_1, \ldots, a_p$ of multiplicity $n_1, \ldots, n_p$, respectively. Then the average characteristic polynomial $P_n(z) = \mathbb{E} \det[zI - M]$ satisfies

$$\int P_n(x) x^k e^{-n(V(x)-a_j x)} dx = 0, \qquad k = 0, \ldots, n_j - 1, \quad j = 1, \ldots, p,$$

and these relations characterize the polynomial P_n, see [3]. The polynomials are known as multiple orthogonal polynomials of type II and they are characterized by a $(p+1) \times (p+1)$-matrix RH problem [10]. For $p = 2$, the RH problem is to find an analytic $Y : \mathbb{C} \setminus \mathbb{R} \to \mathbb{C}^{3\times 3}$ such that

- for $x \in \mathbb{R}$, we have

$$Y_+(x) = Y_-(x) \begin{pmatrix} 1 & e^{-n(V(x)-a_1 x)} & e^{-n(V(x)-a_2 x)} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2}$$

- as $z \to \infty$, we have

$$Y(z) = \left(I + O\left(\frac{1}{z}\right)\right) \begin{pmatrix} z^n & 0 & 0 \\ 0 & z^{-n_1} & 0 \\ 0 & 0 & z^{-n_2} \end{pmatrix}. \tag{3}$$

This RH problem has a unique solution and $Y_{11}(z) = P_n(z)$.

The m-point correlation function for the eigenvalues of (1) has determinantal form [11]

$$R_m(\lambda_1, \ldots, \lambda_m) = \det\left(K_n(\lambda_i, \lambda_j)\right)_{1 \le i,j \le n}$$

with a kernel K_n built out of multiple orthogonal polynomials of type I and II, see [3]. For the case $p=2$ the kernel can be expressed in terms of the solution of the Riemann-Hilbert problem as follows

$$K_n(x,y) = \frac{e^{-\frac{1}{2}n(V(x)+V(y))}}{2\pi i(x-y)} \begin{pmatrix} 0 & e^{na_1 y} & e^{na_2 y} \end{pmatrix} Y^{-1}(y)Y(x) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{4}$$

The expression (4) is based on a Christoffel-Darboux formula for multiple orthogonal polynomials [3, 7].

The large n limit of the Gaussian case ($V(M) = \frac{1}{2}M^2$) with 2 eigenvalues $a_1 = a$, $a_2 = -a$ of equal multiplicity exhibits a phase transition for the value $a = 1$. For $a > 1$ the eigenvalues are asymptotically distributed on two disjoint intervals $[-z_1, -z_2] \cup [z_2, z_1]$, while for $a \le 1$ the eigenvalues accumulate on a single interval $[-z_1, z_1]$. The limiting mean eigenvalue density is given by $\rho(x) = \frac{1}{\pi}\Im|\xi(x)|$, where $\xi(x)$ satisfies the third order equation (Pastur's equation [9])

$$\xi^3 - x\xi^2 - (a^2-1)\xi + xa^2 = 0. \tag{5}$$

For $a = 1$, the density has a $|x|^{1/3}$ behavior near $x = 0$.

We establish universality of local eigenvalue correlations in the large n limit. In [4] we apply the steepest descent method to the RH problem (2), (3) with $V(x) = \frac{1}{2}x^2$, $a_1 = a$, $a_2 = -a$, $n_1 = n_2$, and we assume $a > 1$. A main tool is the Riemann surface for the equation (5) and the functions defined on it. The results are that for x_0 in the bulk,

$$\lim_{n\to\infty} \frac{1}{n\rho(x_0)} \hat{K}_n\left(x_0 + \frac{x}{n\rho(x_0)}, x_0 + \frac{y}{n\rho(x_0)}\right) = \frac{\sin\pi(x-y)}{\pi(x-y)}. \tag{6}$$

At the edge point z_1, we have for a certain $c > 0$,

$$\lim_{n\to\infty} \frac{1}{(cn)^{2/3}} \hat{K}_n\left(z_1 + \frac{x}{(cn)^{2/3}}, z_1 + \frac{y}{(cn)^{2/3}}\right) = \frac{\mathrm{Ai}(x)\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\mathrm{Ai}(y)}{x-y} \tag{7}$$

where Ai is the Airy function. Similar expressions are valid at $-z_1$ and at $\pm z_2$. The kernel $\hat{K}_n$ in (6) and (7) is a modification of K_n

$$\hat{K}_n(x,y) = e^{n(h(x)-h(y))}K_n(x,y)$$

for a certain function h, which does not affect the eigenvalue correlation functions.

For $0 < a < 1$, the steepest descent analysis of the RH problem proceeds in a different way [1], but we again find the sine kernel in the bulk and the Airy kernel at the edges. For $a = 1$, the local eigenvalue correlations near $x = 0$ are given in terms of Pearcey integrals [5, 6].

References

[1] A.I. Aptekarev, P.M. Bleher, and A.B.J. Kuijlaars, Large n limit of Gaussian random matrices with external source, part II, in preparation.

[2] P. Bleher and A. Its, Semiclassical asymptotics of orthogonal polynomials, Riemann-Hilbert problem, and the universality in the matrix model, Ann. Math. 150 (1999), 185–266.

[3] P.M. Bleher and A.B.J. Kuijlaars, Random matrices with external source and multiple orthogonal polynomials, Int. Math. Res. Not. 2004 (2004), 109–129.

[4] P.M Bleher and A.B.J. Kuijlaars, Large n limit of Gaussian random matrices with external source, part I, preprint arXiv:math-ph/0402042.

[5] E. Brézin and S. Hikami, Universal singularity at the closure of a gap in a random matrix theory, Phys. Rev. E 57 (1998) 4140–4149.

[6] E. Brézin and S. Hikami, Level spacing of random matrices in an external source, Phys. Rev. E 58 (1998), 7176–7185.

[7] E. Daems and A.B.J. Kuijlaars, A Christoffel-Darboux formula for multiple orthogonal polynomials, preprint arXiv:math.CA/0402031

[8] P. Deift, T. Kriecherbauer, K.T-R McLaughlin, S. Venakides, and X. Zhou, Uniform asymptotics of polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory, Commun. Pure Appl. Math. 52 (1999), 1335–1425.

[9] L.A. Pastur, The spectrum of random matrices (Russian), Teoret. Mat. Fiz. 10 (1972), 102–112.

[10] W. Van Assche, J.S. Geronimo, and A.B.J. Kuijlaars, Riemann-Hilbert problems for multiple orthogonal polynomials, in: Special Functions 2000 (J. Bustoz et al., eds.), Kluwer, Dordrecht, 2001, pp. 23–59.

[11] P. Zinn-Justin, Random Hermitian matrices in an external field, Nuclear Physics B 497 (1997), 725–732.

Behaviour of kernel functions under homotopies of planar domains

Eric Schippers

The main results are 1) a variational formula for Green's function of finitely connected planar domains, and 2) the demonstration of the monotonicity of various domain functions under set inclusion. The variational formula shows that up to first order, a general homotopy behaves like the normal variation of Hadamard [5]. The consideration of general homotopies is necessary in order to obtain monotonicity of the domain functions.

The variational formula is obtained by isolating the normal part of the variation. Let Γ_{t_0} and Γ_t be parametrize one of the boundary components of domains D_t and D_{t_0} (here t is the homotopy variable). For t is close to t_0, let $n_{t_0}(t,\tau)$ be the distance from $\Gamma_{t_0}(\tau)$ to the curve Γ_t along the normal to Γ_{t_0}. Let

$$\nu_{t_0}(\tau) = \left.\frac{d}{dt}\right|_{t_0} n_{t_0}(t,\tau);$$

we then have that

$$g_t(z,\zeta) - g_{t_0}(z,\zeta) = \frac{t-t_0}{2\pi}\int_{\partial D_{t_0}} \frac{\partial g_{t_0}}{\partial n_u}(u,z)\frac{\partial g_{t_0}}{\partial n_u}(u,\zeta)\nu_{t_0}(u)ds_u + O(|t-t_0|^2)$$

where ds is arc length and n is the outward unit normal. The remainder term is harmonic and bounded on compact sets. This idea was applied in special cases by Barnard and Lewis [1].

With the use of this formula, it is quite easy to prove the monotonicity of various expressions in the derivatives of Green's function simply by differentiating the expression in the homotopy variable. More precisely, one desires theorems of the form $D_1 \subset D_2 \implies \Phi(D_1) \geq \Phi(D_2)$, where Φ is some functional depending on the domain. If one can construct a homotopy D_t between D_1 and D_2, one can apply the variational formula above to show that $\Phi(D_t)$ is monotonic. For example, for Green's function g let

$$K(\zeta,\eta) = -\frac{2}{\pi}\frac{\partial^2 g}{\partial\zeta\partial\overline{\eta}} \quad \text{and} \quad L(\zeta,\eta) = -\frac{2}{\pi}\frac{\partial^2 g}{\partial\zeta\partial\eta}.$$

These are the familiar Bergman kernel and an analogue of the Garabedian kernel for the Bergman space. The following expression decreases as the domain increases:

$$\Re\Delta\left(\sum_{\mu,\nu}\alpha_\mu\alpha_\nu\frac{\partial^{2m}L}{\partial\zeta^m\partial\eta^n}(\zeta_\mu,\zeta_\nu)\right) - \Delta\left(\sum_{\mu,\nu}\alpha_\mu\overline{\alpha}_\nu\frac{\partial^{2m}K}{\partial\zeta^m\partial\overline{\eta}^m}(\zeta_\mu,\zeta_\nu)\right) \leq 0,$$

where ζ_μ are points in the domain and $\alpha_\mu \in \mathbb{C}$ for $\mu = 1,\ldots n$. For simply connected domains this result was obtained by the author in [4]. The case $m = 0$ is due to Nehari [3]. The theorem was obtained by Bergman and Schiffer [2] in the case that $m = 0$ and the outer domain is the plane.

The original motivation of the author for constructing monotonic quantities was in order to obtain distortion theorems for bounded univalent functions. In the simply connected case Green's function can be written in terms of the mapping function and vice versa. The monotonicity theorems for domain functions in some sense are intrinsic versions of inequalities for mapping functions; by choosing D_2 to be the unit disc, and D_1 to be the image of the unit disc under a mapping function, one recovers estimates for the mapping function.

Considering expressions in higher derivatives of Green's function is a natural way to generate inequalities for higher derivatives of the mapping function. Indeed the above inequality easily leads to sharp inequalities for odd derivatives of bounded univalent functions. Inequalities for even derivatives of the mapping function seem to be more difficult. The following quantity is monotonic for all λ, points ζ_μ and parameters α_μ, β_μ, and generates inequalities for even derivatives:

$$\sum_{\mu,\nu}\alpha_\mu\overline{\alpha}_\nu\frac{\partial^{2m}K}{\partial\zeta^m\partial\overline{\eta}^m}(\zeta_\mu,\zeta_\nu) + 2\lambda\Re\left(\sum_{\mu,\nu}\beta_\mu\alpha_\nu\frac{\partial^{2m+1}L}{\partial\zeta^{m+1}\partial\eta^m}(\zeta_\mu,\zeta_\nu)\right) + \lambda^2\sum_{\mu,\nu}\beta_\mu\overline{\beta}_\nu\frac{\partial^{2m+2}K}{\partial\zeta^{m+1}\partial\overline{\eta}^{m+1}}(\zeta_\mu,\zeta_\nu).$$

Although this expression appears complicated, it is the simplest monotonic quantity in which an odd derivative of L appears. Many more such monotonic quantities can be constructed.

Some questions arise naturally. 1) For this method, it is crucial that the boundaries of the domains must be homotopic, and hence two domains must be of the same topological type in order to compare them. For which expressions is this condition necessary for monotonicity to hold? 2) Can one detect the connectivity from these domain functions?

References

[1] Barnard, R and Lewis, J. *Subordination theorems for some classes of starlike functions.* Pacific J. Math. **56**, (1075), no 2. 55–69.
[2] Bergman, S. and Schiffer, M. *Kernel functions and conformal mapping.* Compositio Math. **8**, (1951). 205–249.
[3] Nehari, Z. *Some inequalities in the theory of functions.* Trans. Amer. Math. Soc. **75**, (1953). 256–286.
[4] Schippers, E. *Conformal invariants and higher-order Schwarz lemmas.* J. Anal. Math. **90**, (2003). 217–241.
[5] Hadamard, J. *Mémoire sur le problème conforme des plaques encastrées.* Acad. Sci. Paris, Mémoires des savants ètrangers. **33**, (1908) 1–128.

Zero distribution and asymptotics of Bergman orthogonal polynomials

Nikos Stylianopoulos
(joint work with Erwin Mina Diaz, Eli Levin and Ed Saff)

Let G be a bounded simply-connected domain in the complex plane $\mathbb{C}$, whose boundary $L := \partial G$ is a Jordan curve and let $\{P_n\}_{n=0}^{\infty}$ denote the sequence of Bergman polynomials of G. This is defined as the sequence

$$P_n(z) = \gamma_n z^n + \cdots, \quad \gamma_n > 0, \quad n = 0, 1, 2, \ldots,$$

of polynomials that are orthonormal with respect to the inner product

$$(f, g) := \int_G f(z)\overline{g(z)}dm(z),$$

where dm stands for the 2-dimensional Lebesgue measure.

One purpose of the talk is to report on results, obtained jointly with Eli Levin and Ed Saff in [2], concerning the asymptotic behaviour of the zeros of the Bergman polynomials $\{P_n\}$. In order to state these results we need to consider the two conformal maps associated with L. That is, with $\mathbb{D} := \{w : |w| < 1\}$, let $\Omega := \overline{\mathbb{C}} \setminus \overline{G}$ and $\Delta := \overline{\mathbb{C}} \setminus \overline{\mathbb{D}}$ denote, respectively, the exterior (in $\overline{\mathbb{C}}$) of $\overline{G}$ and $\overline{\mathbb{D}}$. Then, the *exterior* conformal map Φ associated with G is the conformal map $\Phi : \Omega \to \Delta$, normalised so that

$$\Phi(z) = cz + \mathcal{O}(1), \quad z \to \infty, \quad c > 0.$$

The constant

$$\operatorname{cap} L = 1/c,$$

is called the (logarithmic) capacity of L. With $\zeta \in G$, let φ_ζ be an *interior* conformal mapping of G onto the unit disk $\mathbb{D}$, such that $\varphi_\zeta(\zeta) = 0$. Our first result characterises the asymptotic behaviour of the zeros of P_n's in terms of the analytic properties of φ_ζ, by means of two measures. Namely, the *normalised counting measure of the zeros* of P_n, denoted by ν_{P_n}, and the *equilibrium measure* for L, denoted by μ_L. With the above notations, our result can be stated as follows (see [2, Thm 2.1]):

The following two statements are equivalent:

(i) φ_ζ *has a singularity on* L.
(ii) *There is a subsequence* $\mathcal{N} \subset \mathbb{N}$ *such that*

$$\nu_{P_n} \xrightarrow{*} \mu_L, \quad \text{as } n \to \infty,\ n \in \mathcal{N}.$$

Note that the fact φ_ζ has a singularity on L is independent of the choice of φ_ζ, since any two conformal mappings of G onto $\mathbb{D}$ are related by a Möbius transformation. The complimentary case where φ_ζ has no singularities on L is more complicated, and different situations may arise. Here, we consider the special case where the boundary L of G consists of two circular arcs, L_α and L_β, that meet each other at right angles at the points i and $-\mathrm{i}$. In this case we have the following result (see [2, Thm 3.3]):

There exists a Jordan arc Γ *joining the two vertices of* G*, and a certain measure* μ *supported on* Γ*, such that*

$$\nu_{P_n} \xrightarrow{*} \mu, \quad n \to \infty.$$

This "critical arc" Γ is characterised by the property that

$$\Gamma = \{z \in \overline{G} : |\Phi(z_\alpha)| = |\Phi(z_\beta)|\},$$

where for any point z on $\overline{G}$, z_α and z_β denote, respectively, the reflections of z with respect to L_α and L_β.

Another purpose of the talk is to report on, as yet unpublished, results obtained jointly with Erwin Mina Diaz and Ed Saff. These results concern the asymptotic behaviour of the zeros of the *weighted* Bergman polynomials $\{P_{n,w}\}_{n=0}^\infty$, of lens shaped-domains G of the type studied above. These are the polynomials orthonormal with respect to the weighted inner product

$$(f,g)_w := \int_G f(z)\overline{g(z)}|w(z)|^2 dm(z),$$

where w is an entire function with finitely many zeros in $\mathbb{C}$.

Finally, we present a conjecture concerning the asymptotic behaviour of the Bergman polynomials $\{P_n\}$. More precisely, consider the following two formulas:

$$\gamma_n = \sqrt{\frac{n+1}{\pi}} \frac{1}{\operatorname{cap} L^{n+1}} \{1+\alpha_n\},$$

$$P_n(z) = \sqrt{\frac{n+1}{\pi}} \Phi'(z)\Phi^n(z)\ \{1+\beta_n\},\ z \in \overline{\Omega}.$$

If the boundary L of G is an *analytic* Jordan curve, then a result due to T. Carleman gives,

$$\alpha_n = \mathcal{O}(\rho^{2n}) \text{ and } \beta = \mathcal{O}(\rho^n),\ n \to \infty,$$

for some $\rho < 1$; see e.g. [1, pp. 12–13]. In the case where L is *smooth*, typically $L \in C(p+1,s)$, where $p+1 \in \mathbb{N}$ and $p+s > \frac{1}{2}$, then a result of P.K. Suetin ([3, Thms 1.1 and 1.2]) gives,

$$\alpha_n = \mathcal{O}\left(\frac{1}{n^{2(p+s)}}\right) \text{ and } \beta_n = \mathcal{O}\left(\frac{\log n}{n^{p+s}}\right),\ n \to \infty.$$

Our *conjecture*, which is based on certain theoretical results and strong numerical evidence, is concerned with boundary curves that encountered very frequently in the applications and can be stated as follows:

If L is a piecewise analytic Jordan curve without cusps, then

$$\gamma_n = \sqrt{\frac{n+1}{\pi}}\frac{1}{\operatorname{cap} L^{n+1}}\,\{1+\mathcal{O}\left(\frac{1}{n^2}\right)\},\ n\to\infty,$$

$$P_n(z) = \sqrt{\frac{n+1}{\pi}}\Phi'(z)\Phi^n(z)\ \{1+\mathcal{O}\left(\frac{1}{n}\right)\},\ z\in\Omega,\ n\to\infty.$$

References

[1] D. Gaier, *Lectures on complex approximation*, Birkhäuser Boston Inc., Boston, MA, 1987. Translated from the German by Renate McLaughlin.

[2] A. L. Levin, E. B. Saff, and N. S. Stylianopoulos, *Zero distribution of Bergman orthogonal polynomials for certain planar domains*, Constr. Approx., 19 (2003), pp. 411–435.

[3] P. K. Suetin, *Polynomials orthogonal over a region and Bieberbach polynomials*, American Mathematical Society, Providence, R.I., 1974. Translated from the Russian by R. P. Boas.

Asymptotics of Hermite-Padé Polynomials to the Exponential Function

Herbert Stahl

1. Abstract of the Talk

Hermite-Padé polynomials and their associated approximants are in a very natural way generalizations of Taylor polynomials, Padé approximants, and continued fractions (cf. [2], [1]). Historically, they are, perhaps, most famous for their role in Hermite's proof of the transcendency of the number e (cf. [8], [11], [12]).

Within the last 15 years a considerable up-swing of interest and research in this topic could be observed in complex and constructive approximation theory, where the field is typically connected with questions like multiple orthogonality, higher order recurrence relations, and/or the approximation of functions with branch points (cf. surveys in [14], [3], [1], [7], [17]). Many of the basic questions about the convergence of the approximants and the asymptotics of the polynomials are still open.

The talk is based on recent research about quadratic Hermite-Padé polynomials associated with the exponential function. After a somewhat broader introduction to the subject, new results about the asymptotic behavior of the polynomials have been presented. The central element of the asymptotic relations is a concrete, compact Riemann surfaces with 3 sheets over $\overline{\mathbb{C}}$. Details of its definition can be found in [18], Subsection 2.2. Specific results will be summarized further below in the present abstract. First we repeat the definition of Hermite-Padé polynomials and the associated approximants.

2. Definition of Hermite-Padé Polynomials

Let $\mathfrak{f} = (f_0, \ldots, f_m)$, $m \geq 1$, be a system of $m+1$ functions; all functions are assumed to be analytic in a neighborhood of the origin.

Definition 1 Hermite-Padé Polynomials of Type I (Latin polynomials in K. Mahler's terminology in [13]): For any multi-index $n = (n_0, \ldots, n_m) \in \mathbb{N}^{m+1}$ there exists a vector of polynomials $(p_0, \ldots, p_m) \in \mathcal{P}^*_{n_0-1} \times \mathcal{P}_{n_1-1} \times \ldots \times \mathcal{P}_{n_m-1}$ such that

$$\sum_{j=0}^{m} p_j(z) f_j(z) \;=\; O(z^{|n|-1}) \quad \text{as} \quad z \to 0, \tag{1}$$

where $|n| := n_0 + \ldots + n_m$ and $\mathcal{P}^*_k := \{ p \in \mathcal{P}_k \,|\, p \text{ monic}, p \not\equiv 0 \}$. The vector $(p_0, \ldots, p_m)$ is called Hermite-Padé form of type I, and its elements are the Hermite-Padé polynomials of type I.

Definition 2 Hermite-Padé Polynomials of Type II (German polynomials in K. Mahler's terminology in [13]): For any multi-index $n = (n_0, \ldots, n_m) \in \mathbb{N}^{m+1}$ there exists a vector of polynomials $(\mathfrak{p}_0, \ldots, \mathfrak{p}_m) \in \mathcal{P}^*_{N_0} \times \mathcal{P}_{N_1} \times \ldots \times \mathcal{P}_{N_m}$ with $N_j := |n| - n_j$, $j = 0, \ldots, m$, such that

$$\mathfrak{p}_i(z) f_j(z) - \mathfrak{p}_j(z) f_i(z) \;=\; O(z^{|n|+1}) \quad \text{as} \quad z \to 0, \tag{2}$$

for $i, j = 0, \ldots, m$, $i \neq j$. The vector $(\mathfrak{p}_0, \ldots, \mathfrak{p}_m)$ is called Hermite-Padé form of type II, and its elements are the Hermite-Padé polynomials of type II.

The assumption $p_0 \in \mathcal{P}^*_{n_0-1}$ and $\mathfrak{p}_0 \in \mathcal{P}^*_{N_0}$ implies a normalization of the whole form $(p_0, \ldots, p_m)$ and $(\mathfrak{p}_0, \ldots, \mathfrak{p}_m)$, respectively. There may exist situations in which a normalization by the first component is not possible, however, one of the $m+1$ components always is appropriate for normalization.

3. Definition of Hermite-Padé Approximants

With each of the two types of Hermite-Padé polynomials a specific type of Hermite-Padé approximants is associated; these are the algebraic approximants in case of type I polynomials and the simultaneous rational approximants in case of type II polynomials. We start with the simultaneous rational approximants.

If $f_0(0) \neq 0$, then one can assume without loss of generality in Definition 2 that $f_0 \equiv 1$, and under this assumption the relations (2) reduce to

$$\mathfrak{p}_0(z) f_j(z) - \mathfrak{p}_j(z) \;=\; O(z^{|n|+1}) \quad \text{as} \quad z \to 0 \quad \text{for} \quad j = 1, \ldots, m. \tag{3}$$

Defintion 3 Hermite-Padé Simultaneous Rational Approximants: For a given multi-index $n \in \mathbb{N}^{m+1}$ let $\mathfrak{p}_0, \ldots, \mathfrak{p}_m$ be the Hermite-Padé polynomials of type II defined by (2) respectively (3). Then the vector of rational functions

$$\left(\frac{\mathfrak{p}_1}{\mathfrak{p}_0}(z), \ldots, \frac{\mathfrak{p}_m}{\mathfrak{p}_0}(z) \right) \tag{4}$$

with common denominator polynomial $\mathfrak{p}_0$ is called (Hermite-Padé) simultaneous rational approximant to the (reduced) system of functions $\mathfrak{f}_{red} = (f_1, \ldots, f_m)$.

One immediately sees that for $m = 1$ in Definition 3 we have the Padé approximant to f_1 with numerator and denominator degrees $(n_1,\, n_0)$.

As counterpart to the simultaneous rational approximants we have the algebraic Hermite-Padé approximants, which are defined with the help of polynomials of type I, but in their case the system of functions has to be an algebraic one.

Let f be a function analytic at the origin. We define the algebraic system of functions $\mathfrak{f}$ as

$$\mathfrak{f} = (f_0, \ldots, f_m) := (1, f, \ldots, f^m). \tag{5}$$

Defintion 4 Algebraic Hermite-Padé Approximants: For a given multi-index $n \in \mathbb{N}^{m+1}$ let $p_0, \ldots, p_m \in \mathcal{P}^*_{n_0-1} \times \ldots \times \mathcal{P}_{n_m-1}$ be the Hermite-Padé polynomials of type I defined by (1) with the special choice of (5). Let the algebraic function $y = y(z)$ be defined by the relation

$$\sum_{j=0}^{m} p_j(z) y(z)^j \;\equiv\; 0. \tag{6}$$

From the m branches of y we select the branch $y = y_n$ that has the highest contact to f at the origin; this branch y_n is the algebraic Hermite-Padé approximant to f associated with the multi-index n.

Again, it is immediate that for $m = 1$ Definition 4 leads to an Padé approximant, but this time with numerator and denominator degrees $(n_0 - 1, n_1 - 1)$.

4. The Special Case of the Exponential Function

In the talk we have reported about new research on asymptotics of Hermite-Padé polynomials of both types associated with systems of exponential functions. The order of the system is $m = 2$ and the multi-indices are all of the form $(n, \ldots, n) \in \mathbb{N}^{m+1}$ with $n \in \mathbb{N}$ and $n \to \infty$. Thus, we are dealing with quadratic diagonal Hermite-Padé polynomials to the system $\mathfrak{f} = (1, \exp, \exp^2)$.

After the investigations in the classical period, from where we here only mention [8], [11], [12], our specific line of research in Hermite-Padé approximants had been taken up P. B. Borwein in [4], and more or less the same problem has been studied from a point of view of special functions in [6] and [5]. In these later investigations several questions about the asymptotic distribution of the zeros of the polynomials, and especially about the asymptotic behavior of the larger zeros remained open, and these open questions have triggered our new research.

The leading idea in this new research is a rescaling of the independent variable in such a way that the zeros of the polynomials, which almost all normally diverge to infinity, now have finite asymptotic distributions.

The rescaling method was introduced by G. Szegö in [20] for the investigation of Taylor polynomials to the exponential function, and has later been taken up very successfully for the investigation of poles and zeros of Padé approximants by E.B. Saff and R.S. Varga; the results interesting here can be found in [15]

With the rescaling method it has become possible to prove asymptotic relations for quadratic Hermite-Padé polynomials. In these relations an algebraic function of third degree and the associated Riemann surface play a central role.

The new results for polynomials of type I have just been appeared in [18], very precise results about the asymptotic distributions of zeros will appear soon in [19], and results about the asymptotic behavior of polynomials of type II are in preparation.

An alternative approach to the asymptotic analysis based on a matrix Riemann-Hilbert problem has been developed by A.B.J. Kuijlaars, W. Van Assche, and F. Wielonsky in [9]. A survey of these results is contained in [10].

A generalisation of P. B. Borwein's investigations in [4] to general $m > 2$ has been done by F. Wielonsky in [21] and [22], and it has led to best results for the measure of irrationality of the number e. Investigations of quadratic Hermite-Padé approximants from a numerical point of view can be found in [16].

References

[1] Aptekarev, A.I., and Stahl, H., *Asymptotics of Hermite-Padé polynomials*, Progress in Approximation Theory (Gonchar, A.A. & Saff, E.B., eds.) Springer-Verlag 1992, 127–167.

[2] Baker, Jr. G.A.,Graves-Morris, P., *Padé Approximants*, Cambridge University Press (1996).

[3] Baker, Jr. G.A., Lubinsky, D.S., *Convergence theorems for rows of differential and algebraic Hermite–Padé approximants*, J. Comp. and Appl. Math. 18(1987), 29–52.

[4] Borwein, P.B., *Quadratic Hermite-Padé approximation to the exponential function*, Constr. Approx., 2(1986), 291–302.

[5] Driver, K., *Non-diagonal quadratic Hermite-Padé approximants to the exponential function*, J. Comp. Appl. Math. 65(1995), 125–134.

[6] Driver, K., and Temme, N.M., *On polynomials related with Hermite-Padé approximants to the exponential function*, J. Approx. Theory 95 (1998), 101–122.

[7] Geronimo, J.S., Kuijlaars, A.B.J., and Van Assche, W., *Riemann-Hilbert problems for multiple orthogonal polynomials*, in 'Special Functions 2000: Current Perspective and Future Directions' (J. Bustoz et al., eds.), NATO Science Series II. Mathematics, Physics and Chemestry 30, Kluwer, Dodrecht, 2001, 23-59.

[8] Hermite, C., *Sur la fonction exponentielle*, C.R.Acad. Sci. Paris 77(1873), 18–24, 74–79, 226–233.

[9] Kuijlaars, A.B.J., Van Assche, W., and Wielonsky, F., *Quadratic Hermite-Padé approximation to the exponential function: a Riemann-Hilbert approach*, preprint arxiv, math.CA/03032357.

[10] Kuijlaars, A.B.J., Stahl, H., Van Assche, W., and Wielonsky, F., *Asymptotique des approximants de Hermite-Padé quadratiques de la function exponentielle et problèmes de Riemann-Hilbert*, C. R. Acad. Sci. Paris, Ser. I 336 (2003), 893-896.

[11] Mahler, K., *Zur Approximation der Exponentialfunktion und des Logarithmus I, II*, J. Reine Angew. Math., 166 (1931), 118–37, 138–150.

[12] Mahler, K., *Application of some formulas by Hermite to the approximation of exponentials and logarithms*, Math. Ann., 168(1967), 200–227.

[13] Mahler, K., *Perfect systems*, Comp. Math., 19(1968), 95–166.

[14] Nuttall, J., *Asymptotics of diagonal Hermite-Padé approximants*, J. Approx. Theory 42(1984), 299–386.

[15] Saff, E.B. and Varga, R.S., *On the zeros and poles of Padé approximants to e^z III*, Numer. Math., 30 (1978), 241–266.
[16] Shafer, R.E., *On quadratic approximation*, SIAM . Numer. Anal., 11 (1974), 447–460.
[17] Stahl, H., *Asymptotics for Quadratic Hermite-Padé Polynomials Associated with the Exponential Function*, ETNA, Electr. Trans. Numer. Analysis, 14 (2002), 193-220.
[18] Stahl, H., *Quadratic Hermite-Padé Polynomials Associated with the Exponential Function*, J. Approx. Theory 125(2003), 238–296.
[19] Stahl, H., *Asymptotic Distributions of Zeros of Quadratic Hermite-Padé Polynomials Associated with the Exponential Function*, will appear in Constr. Approx., 2004.
[20] Szegö, G., *Über einige Eigenschaften der Exponentialreihe*, Sitzunggsberichte Berliner Math. Ges., 23(1924), 50–64.
[21] Wielonsky, F., *Asymptotics of diagonal Hermite-Padé approximants to e^z*, J. Approx. Theory 90(1997), 283–298.
[22] Wielonsky, F., *Hermite-Padé approximants to exponential functions and an inequality of Mahler*, J. Number. Theory 74(1999), 230–249.

Entire functions with no unbounded Fatou components

Aimo Hinkkanen

Let f be a transcendental entire function of order less than 1/2. We introduce a condition on the regularity of growth of f and show that it implies that every component of the Fatou set of f is bounded.

The Fatou set $\mathcal{F}(f)$ of f is defined to be the set of those points z in the complex plane $\mathbb{C}$ that have a neighbourhood U such that the family $\{f^n|U : n \geq 1\}$ of the restrictions of the iterates f^n of f to U is a normal family. The Julia set $\mathcal{J}(f)$ of f is $\mathcal{J}(f) = \mathbb{C} \setminus \mathcal{F}(f)$.

I.N. Baker asked in 1981 whether every component of $\mathcal{F}(f)$ is bounded if the growth of f is sufficiently small. This would then imply, in particular, that $\mathcal{F}(f)$ has no Baker domains and no completely invariant components. The best possible growth condition in terms of order would be of order 1/2, minimal type at most, as shown by the functions $f(z) = \cos\sqrt{\varepsilon z + (3\pi/2)^2}$, for $0 < \varepsilon < 3\pi$, for which $\mathcal{F}(f)$ has unbounded components. Baker proved that under this growth condition, a component D of $\mathcal{F}(f)$ is bounded except possibly if it is a wandering domain (that is, all $f^n(D)$ are contained in distinct components of $\mathcal{F}(f)$) or if D or one of its forward images is in a Baker domain cycle of length at least 2. Stallard extended Baker's result to cover Baker domain cycles of any length.

The problem remains if D is a wandering domain; one may then assume that D is simply connected for otherwise all components of $\mathcal{F}(f)$ are bounded for other reasons as shown by Baker.

A number of authors have shown that if f is a transcendental entire function of order less than 1/2 satisfying an extra condition on the regularity of growth of the maximum modulus $M(r, f)$ then all wandering domains and hence all components of $\mathcal{F}(f)$ are bounded. We prove that this conclusion holds if f has the following additional property where $m(r, f)$ denotes the minimum modulus of f: suppose that there exist positive numbers R_0, L, δ, and C with $R_0 > e$, $M(R_0, f) > e$,

$L > 1$, and $0 < \delta \leq 1$ such that for every $r > R_0$ there exists $t \in (r, r^L]$ with

$$\frac{\log m(t,f)}{\log M(r,f)} \geq L\left(1 - \frac{C}{(\log r)^\delta}\right). \tag{1}$$

One can ask whether every transcendental entire function of order less than 1/2 satisfies (1). This is still an open question. If we are not close to or inside an annulus containing very few zeros of f, it would seem plausible that the condition (1) should be easy to satisfy, with a wide margin, by taking t to be a value arising from the $\cos \pi\rho$–theorem. This is because then $\log m(t,f)/\log M(t,f)$ is greater than a fixed constant while $\log M(t,f)/\log M(r,f)$ should be quite large. So there should be a potential problem at most if we are in an annulus where f behaves like a polynomial. But in that case we should be able to take t close to r^L, and then the three numbers $\log m(t,f)$, $\log M(t,f)$, and $L \log M(r,f)$, should be close together. There may be some error term required to estimate $\log m(t,f)/(L\log M(r,f))$ from below, but (1) allows for such a term.

References

[1] J.M. Anderson and A. Hinkkanen, Unbounded domains of normality, Proc. Amer. Math. Soc. 126 (1998), 3243–3252.
[2] I.N. Baker, The iteration of polynomials and transcendental entire functions, J. Austral. Math. Soc. (Series A) 30 (1981), 483–495.
[3] P.D. Barry, On a theorem of Besicovitch, Quart. J. Math. Oxford (Series 2) 14 (1963), 292–302.
[4] A. Hinkkanen, Entire functions with no unbounded Fatou components, preprint (2003).
[5] Xinhou Hua and Chung-Chun Yang, Fatou components of entire functions of small growth, Ergodic Theory Dynam. Systems 19 (1999), 1281–1293.
[6] G. Stallard, The iteration of entire functions of small growth, Math. Proc. Camb. Phil. Soc. 114 (1993), 43–55.
[7] Jian-Hua Zheng, Unbounded domains of normality of entire functions of small growth, Math. Proc. Cambridge Phil. Soc. 128 (2000), 355–361.

Parameter Space of the Exponential Family and Infinite-Dimensional Thurston Theory

Markus Förster
(joint work with Lasse Rempe and Dierk Schleicher)

The talk deals with the investigation of the parameter space of the exponential family

$$\{E_\kappa : z \mapsto e^z + \kappa;\quad \kappa \in \mathbb{C}/2\pi i\mathbb{Z}\}\,.$$

For each parameter κ we consider the dynamical system generated by iteration of the function E_κ. The exponential family can be considered as a model family for transcendental dynamics in the spirit of quadratic polynomials, for every E_κ has only one singular value, the asymptotic value κ. We are interested in the set I of parameters for which the singular value is escaping, i.e. for which κ is contained in the set

$$I(E_\kappa) := \{z \in \mathbb{C} :\ |E_\kappa^{\circ n}(z)| \to \infty \ \text{ as } n \to \infty\}$$

of *escaping points*. We call such parameters *escaping parameters*. For the quadratic family $\{p_c : z \mapsto z^2 + c\}$, the set of escaping parameters can be viewed as a collection of external rays (*parameter rays*) which do or do not land on the bifurcation locus, the boundary of the Mandelbrot set. The parameter rays are the main tool of understanding the topological and bifurcation structure of the Mandelbrot set. In the case of the exponential family the bifurcation locus

$$\mathcal{B} := \{\kappa : \nexists \text{ neighborhood } U \ni \kappa \text{ s.t. } \forall \kappa' \in U \; E_\kappa \text{ and } E_{\kappa'} \text{ are conjugated}\}$$

also turns out to be the boundary of I, and I is still a disjoint union of *parameter rays*, see the theorem below.

The main idea to construct these parameter rays is to carry over structure from the dynamic plane into the parameter plane. Dierk Schleicher and Johannes Zimmer [SZ] have precisely described for any κ the set $I(E_\kappa)$ of escaping points, which consists of uncountably many *dynamic rays* $g^\kappa_{\underline{s}}(t)$ going off to $+\infty$ together with some (but not all) end points of them. This gives rise to a combinatorial description of $I(E_\kappa)$: each escaping point can be assigned a unique pair $(\underline{s}, t)$ of an integer sequence $\underline{s} \in \mathbb{Z}^{\mathbb{N}}$ (which codes the ray $g^\kappa_{\underline{s}}$ the point belongs to) and a real number $t \geq t_{\underline{s}}$ (which determines the position on the ray), where $t_{\underline{s}} \geq 0$ is independent of κ. The sequence $\underline{s} = (s_1, s_2, \dots)$ is derived from itineraries, i.e. symbolic dynamics, and the potential t indicates the speed of escape. Most importantly, the combinatorial data $(\underline{s}, t)$ gives a precise prediction of the orbit of $z = g^\kappa_{\underline{s}}(t)$: for large n we have

$$E^{\circ n}_\kappa(z) = F^{\circ n}(t) + 2\pi i s_{n+1} + O\left((F^{\circ(n+1)}(t))^{-1}\right) , \tag{1}$$

where $F(t) := e^t - 1$. Moreover, the set $X \subset \mathbb{Z}^{\mathbb{N}} \times \mathbb{R}^+_0$ of possible combinatorial pairs, endowed with the discrete topology in the first coordinate and the usual one in the second coordinate, is mapped for all κ bijectively onto $I(E_\kappa)$ by the continuous map

$$\phi_\kappa(\underline{s}, t) : X \to I(E_\kappa) \; ; \quad (\underline{s}, t) \mapsto g^\kappa_{\underline{s}}(t)$$

except if κ is an escaping parameter. We extended this result to the parameter space in the following sense.

Theorem (M. F., L. Rempe, D. Schleicher '03) *Let I be the set of* escaping parameters*:*

$$I := \{\kappa : \; \kappa \in I(E_\kappa)\} ,$$

the parameters for which the singular orbit escapes under E_κ. There is a continuous bijection $\phi : X \to I$ satisfying

$$\phi(\underline{s}, t) = \kappa \quad \Longleftrightarrow \quad g^\kappa_{\underline{s}}(t) = \kappa .$$

The maps $G_{\underline{s}}(t) := \phi(\underline{s}, t)$ are differentiable rays, which precisely form the path-connected components of I.

This result has been obtained by carefully estimating derivatives and winding numbers of dynamic rays ([FS1], [FRS]). Since the proof is technical and impossible to modify for any other setting, we reprove the existence and uniqueness of every combinatorial pair $(\underline{s}, t) \in X$ using *spider theory* [FS2]. The spider algorithm provides a constructive method of realizing given combinatorics and can be implemented as a computer program. It provides a much more conceptual proof which unlike the previous proof uses nothing but the asymptotics (1) of the escaping singular orbits in the dynamic plane. Spider theory is inspired by Thurston's topological characterization of rational maps [DH]. It establishes a correspondence between parameters assuming $(\underline{s}, t) \in X$ and fixed points of a certain self-mapping on Teichmüller space, which is easily described in terms of pull-backs of *spiders*. Spiders are a substantially simplified model of Teichmüller space. They have been invented by John H. Hubbard and have been used by several people in several contexts. However, this is the first time that spiders are applied to a case of infinite degree and an infinite-dimensional Teichmüller space.

The spiders constructed for this purpose are objects consisting of infinitely many feet, which model the escaping singular orbit and represent the projection into moduli space, as well as a leg attached to each foot modulo homotopy, which models the dynamic ray associated to the respective orbit point. By the asymptotic behavior (1) we have very good control of how the actual singular orbit and the dynamic rays eventually have to behave if κ assumes the prescribed combinatorics. This allows us to only consider legs and feet with rather special properties. The iterated map on the space of spiders (*spider map*) is defined by pulling back the spider along the inverse branches of E_κ as given by the entries of $\underline{s}$, where κ is the first foot.

Showing that the spider map possesses exactly one fixed point for a given pair $(\underline{s}, t) \in X$ consists of finding an invariant compact subset of spiders in order to apply the Banach fixed point theorem for the existence and a contraction argument for the uniqueness. The definition of the infinitesimal Teichmüller metric on the spider space involves the discussion of L^1-integrable meromorphic quadratic differentials, which describe the cotangent space and give rise to the dual norm on the tangent space. The push-forward of quadratic differentials turns out to be adjoint to the spider map acting on the tangent space, so that the contraction of the spider map can be understood in terms of mass loss of quadratic differentials. In order to find a compact invariant subset we carefully construct a configuration ofszi feet with definite estimates on absolute values and mutual distances as well as estimates on winding numbers of the feet.

References

[HS] John H. Hubbard and Dierk Schleicher: *The Spider Algorithm.* Proceedings of Symposia in Appl. Math. Vol. 49, 155-180 (1994).

[DH] Adrien Douady and John H. Hubbard: *A Proof of Thurston's Topological Characterization of Rational Functions.* Acta Math. 171, 263-297 (1993).

[FRS] Markus Förster, Lasse Rempe, Dierk Schleicher: *Classification of Escaping Exponential Maps.* Manuscript (2003), to be published; math.DS/0311427 at ArXiV.org.

[FS1] Markus Förster, Dierk Schleicher: *Parameter Rays for the Exponential Family.* Manuscript (2004), to be published.

[FS2] Markus Förster, Dierk Schleicher: *Spiders for Escaping Exponential Maps.* Manuscript (2004).

[SZ] Dierk Schleicher, Johannes Zimmer: *Escaping Points of Exponential Maps.* Journal of the London Math. Society Nr. 67 (2003), pp. 1-21.

Growth of harmonic functions in the unit disc and an application

Igor Chyzhykov

1. Analytic and harmonic functions in the unit disc. Let $D = \{z \in \mathbb{C} : |z| < 1\}$. We denote by $A(D)$ the class of analytic function in D. For $f \in A(D)$ let $M(r, f) = \max\{|f(z)| : |z| = r\}$, $0 < r < 1$, $T(r, f) = \frac{1}{2\pi}\int_0^{2\pi} \log^+ |f(re^{i\theta})|\, d\theta$.

Usually, the orders of the growth of analytic functions in D are defined as

$$\rho_M[f] = \limsup_{r\uparrow 1} \frac{\log^+ \log^+ M(r, f)}{-\log(1-r)}, \ \rho_T[f] = \limsup_{r\uparrow 1} \frac{\log^+ T(r, f)}{-\log(1-r)}.$$

It is well known that $\rho_T[f] \le \rho_M[f] \le \rho_T[f] + 1$, and all cases are possible.

In 1960th M. M. Djrbashian using the Riemann-Liouville fractional integral obtained a parametric representation of the class of analytic (meromorphic) functions f in D of finite order of the growth [Chap. IX, Dj].

Here we confine by the case when $f(z)$ has no zeros and of finite order of the growth, hence $\log|f(z)|$ is harmonic.

For $\psi\colon [0, 2\pi] \to \mathbb{R}$ we define the modulus of continuity $\omega(\delta; \psi) = \sup\{|\psi(x) - \psi(y)| : |x - y| \le \delta, x, y \in [0, 2\pi]\}$, $\delta > 0$.

Following [HL, Z] we say that $\psi \in \Lambda_\gamma$ if $\omega(\delta; f) = O(\delta^\gamma)$ $(\delta \downarrow 0)$.

The fractional integral of order $\alpha > 0$ for $h\colon (0, 1) \to \mathbb{R}$ is defined by the formulas [Dj, HL]

$$D^{-\alpha} h(r) = \frac{1}{\Gamma(\alpha)} \int_0^r (r - x)^{\alpha-1} h(x)\, dx, \quad D^0 h(r) \equiv h(r).$$

Let $H(D)$ be the class of harmonic functions in D.

We put $u_\alpha(re^{i\varphi}) = r^{-\alpha} D^{-\alpha} u(re^{i\varphi})$, where the fractional integral is taken on the variable r. Let $B(r, u) = \max\{u(z) : |z| \le r\}$.

Our starting point is the following theorem

Theorem B (M. Djrbashian). *Let $u \in H(D)$, $\alpha > -1$. Then*

$$u(re^{i\varphi}) = \frac{1}{2\pi} \int_0^{2\pi} P_\alpha(r, \varphi - \theta)\, d\psi(\theta), \tag{2}$$

where $\psi \in BV[0, 2\pi]$,

$$P_\alpha(r, t) = \Gamma(1 + \alpha)\Big(\Re \frac{2}{(1 - re^{it})^{\alpha+1}} - 1\Big),$$

if and only if

$$\sup_{0<r<1} \int_0^{2\pi} |u_\alpha(re^{i\varphi})|\, d\varphi < M_\alpha.$$

Remark 1. Actually, for $\alpha = 0$ it is the classical result of Nevanlinna on representation of $\log|F(z)|$ when $F \in N$.

Remark 2. Note that $P_0(r,t)$ is the Poisson kernel; $P_\alpha(r,t) = D^\alpha(r^\alpha P_0(r,t))$.

Applying methods from [Dj] and [HL] (see also [Chap.7, Z]), we prove the following theorem (cf. Theorem 40 [HL]).

Theorem 1. *Let $u(z) \in H(D)$, $\alpha \geq 0$, $0 < \gamma < 1$. Then $u(z)$ has form (2) where ψ is of bounded variation on $[0, 2\pi]$, and $\psi \in \Lambda_\gamma$, if and only if*

$$B(r,u) = O((1-r)^{\gamma-\alpha-1}), \quad r \uparrow 1$$

and

$$\sup_{0<r<1} \int_0^{2\pi} |u_\alpha(re^{i\varphi})|\, d\varphi < +\infty.$$

2. An application to growth of analytic functions. For $\psi \in BV[0, 2\pi]$ we denote

$$\tau[\psi] = \liminf_{\delta \downarrow 0} \frac{\log^+ \frac{1}{\omega(\delta;\psi)}}{-\log \delta} \geq 0.$$

The quantity $\tau[\psi]$ compares $\omega(\delta;\psi)$ with δ^γ as $\delta \to 0$.

Theorem 2. *Let $F \in A(D)$, and*

$$\log|F(re^{i\varphi})| = \frac{1}{2\pi} \int_0^{2\pi} P_\alpha(r, \varphi - t)\, d\psi(t),$$

where $\psi \in BV[0, 2\pi]$, $\tau[\psi] = \tau \in [0, 1)$. Then $\rho_M[F] = \alpha + 1 - \tau$, $\rho_T[F] \leq \alpha$. If, in addition, ψ is not absolutely continuous, then $\rho_T[F] = \alpha$.

Corollary. *Suppose that the conditions of Theorem 2 hold, and $\tau = 0$. Then $\rho_M[F] = \rho_T[F] + 1 = \alpha + 1$.*

References

[Dj] Djrbashian M.M. Integral transforms and representations of functions in a complex domain, Moscow, Nauka, 1966. (in Russian)

[HL] Hardy G.H., Litlewood J.E. Some properties of fractional integrals. II. Math. Zeit. **34** (1931/32), 403–439.

[Z] Zygmund A. Trigonometric series, V.1, Cambridge Univ. Press, 1959.

On conformal invariants in problems of constructive function theory

V. V. Andrievskii

This is a survey of some recent results by the author and his collaborators in the constructive theory of functions of a real variable. The results are achieved by the application of methods and techniques of modern geometric function theory and potential theory in the complex plane.

Let $E \subset \mathbb{C}$ be a compact set of positive logarithmic capacity $\mathrm{cap}(E)$ with connected complement $\Omega := \overline{\mathbb{C}} \setminus E$ with respect to $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, $g_\Omega(z) = g_\Omega(z, \infty)$ be the Green function of Ω with pole at infinity, and μ_E be the equilibrium measure for the set E. The properties of g_Ω and μ_E play an important role in many problems concerning polynomial approximation of continuous functions on E and the behavior of polynomials with a known uniform norm along E.

We discuss some of these problems for the case when E is a subset of the real line $\mathbb{R}$. The main idea of our approach is to use conformal invariants such as the extremal length and module of a family of curves. The basic conformal mapping can be described as follows.

Let $E \subset [0,1]$ be a regular set such that $0 \in E$, $1 \in E$. Then $[0,1] \setminus E = \sum_{j=1}^{N} (a_j, b_j)$, where N is finite or infinite.

Denote by $\mathbb{H} := \{z : \Im(z) > 0\}$ the upper half-plane and consider the function

$$F(z) = F_E(z) := \exp\left(\int_E \log(z-\zeta)\, d\mu_E(\zeta) - \log \mathrm{cap}(E) \right), \quad z \in \mathbb{H}.$$

Using the reflection principle we can extend F to a function analytic in $\overline{\mathbb{C}} \setminus [0,1]$ by the formula

$$F(z) := \overline{F(\overline{z})}, \quad z \in \mathbb{C} \setminus \overline{\mathbb{H}}.$$

F is univalent and maps $\overline{\mathbb{C}} \setminus [0,1]$ onto a (with respect to ∞) starlike domain $\overline{\mathbb{C}} \setminus K_E$ with the following properties: $\overline{\mathbb{C}} \setminus K_E$ is symmetric with respect to the real line $\mathbb{R}$ and coincides with the exterior of the unit disk with $2N$ slits.

Note that

$$g_\Omega(z) = \log |F(z)|, \quad z \in \mathbb{C} \setminus E.$$

There is a close connection between the capacities of the compact sets K_E and E, namely

$$4\mathrm{cap}(E)\mathrm{cap}(K_E) = 1.$$

The main idea of our results is the investigation of the local properties of the Green function g_Ω, i.e., local properties of conformal mapping F.

The lecture is organized as follows. In part 1 we describe the connection between uniformly perfect subsets in $\mathbb{R}$ and John domains. It allows us to extend well-known theorem about constructive description of functions with a given majorant of their best uniform polynomial approximations to the case of C-dense compact subset of $\mathbb{R}$.

In part 2 we give sharp uniform bounds for exponentials of logarithmic potentials if the logarithmic capacity of the subset, where they are at most 1, is known.

In part 3 we give a new interpretation (and a generalization) of recent remarkable result by Totik [7] concerning the smoothness properties of g_Ω and μ_E. We also demonstrate that if for $E \subset [0,1]$ the Green function satisfies the 1/2-Hölder condition locally at the origin, then the density of E at 0, in terms of logarithmic capacity, is the same as that of the whole interval $[0,1]$.

In part 4 the Nikol'skii-Timan-Dzjadyk theorem concerning polynomial approximation of functions on the interval $[-1,1]$ is generalized to the case of approximation of functions given on a compact set on the real line.

A new necessary condition and a new sufficient condition for the approximation of the reciprocal of an entire function by reciprocals of polynomials on $[0,\infty)$ with geometric speed of convergence are provided in part 5.

References

[1] Andrievskii V., On the Green function for a complement of a finite number of real intervals, Constr. Approx., to appear.
[2] Andrievskii V., The highest smoothness of the Green function implies the highest density of a set, Arkiv för Matematik, to appear.
[3] Andrievskii V., A note on a Remez-type inequality for trigonometric polynomials, J. Approx. Theory, 116 [2002], 416-424.
[4] Andrievskii V., A Remez-type inequality in terms of capacity, Complex Variables, 45 [2001], 35-46.
[5] Andrievskii V., The Nikol'skii-Timan-Dzjadyk theorem for functions on compact sets of the real line, Constr. Approx., 17 [2001], 431-454.
[6] Andrievskii V.V., Blatt H.-P., Kovacheva R.K., Approximation on an unbounded interval, Analysis, 17 [1997], 197-218.
[7] Totik V., Metric properties of harmonic measure. - Manuscript.

Dynamics on fractal spheres

Mario Bonk

The point of this talk was to argue the dynamics of Kleinian groups on the 2-sphere and the dynamics of a rational function under iteration lead to closely related problems in the theory of analysis on metric spaces.

We first introduce a setting for *Kleinian groups* that can be considered as the "standard picture" in this respect. Let M be a closed hyperbolic 3-orbifold, and $\Gamma = \pi_1(M)$ the fundamental group of M. The universal covering space of M is hyperbolic 3-space $\mathbb{H}^3$, the group Γ acts on $\mathbb{H}^3$ by deck transformations, and the orbifold is given by the quotient $M = \mathbb{H}^3/\Gamma$.

The action $\Gamma \curvearrowright \mathbb{H}^3$ is isometric, discrete, and cocompact. Let us call a group *standard* if it admits an action on $\mathbb{H}^3$ with these properties. The basic problem is to characterize this standard situation from the point of view of geometric group theory.

There is a well-developed theory due to Gromov of groups that resemble fundamental groups of negatively curved manifolds [Gr]. These groups are called *hyperbolic (in the sense of Gromov)* [GhHa]. If Γ is a group as above, then Γ is hyperbolic and its boundary at infinity $\partial_\infty\Gamma$ is homeomorphic to the standard

2-sphere $\mathbb{S}^2$ (abbreviated $\partial_\infty\Gamma \approx \mathbb{S}^2$). According to a conjecture by Cannon [Ca] this should characterize standard groups.

Cannon's conjecture. *Suppose G is a Gromov hyperbolic group with $\partial_\infty G \approx \mathbb{S}^2$. Then G is standard.*

In this situation it is enough to show that $\partial_\infty G$ is homeomorphic to $\mathbb{S}^2$ by a quasisymmetric homeomorphism (abbreviated $\partial_\infty G \overset{qs}{\approx} \mathbb{S}^2$). Indeed, G acts in a natural way on $\partial_\infty G$ by uniformly quasi-Möbius homeomorphisms. If $\partial_\infty G \overset{qs}{\approx} \mathbb{S}^2$, then this action $G \curvearrowright \partial_\infty G$ conjugates to an action $G \curvearrowright \mathbb{S}^2$ of G on the standard 2-sphere by uniformly quasiconformal homeomorphisms. A well-known theorem due to Sullivan [Su] and to Tukia [Tu] then implies that this action is conjugate to an action of G on $\mathbb{S}^2$ by Möbius transformations. From this it easily follows that G is standard.

We are lead to the general problem when a fractal 2-sphere such as $\partial_\infty G$ in the above situation is quasisymmetrically equivalent to the standard 2-sphere. This question was studied by B. Kleiner and myself [BK1]. As an application of our results we obtained the following partial result for Cannon's conjecture.

Theorem 1. *Suppose G is a Gromov hyperbolic group with $\partial_\infty G \approx \mathbb{S}^2$. If there exists an Ahlfors 2-regular 2-sphere Z such that $\partial_\infty G \overset{qs}{\approx} Z$, then G is standard.*

Recall that a (compact) metric space Z is called *Ahlfors Q-regular* for $Q > 0$ if the Hausdorff Q-measure of small balls $B(a, R)$ in Z behaves like R^Q up to multiplicative constants independent of the balls.

A stronger result can be obtained by using the concept of the *Ahlfors regular conformal dimension* $\dim_{AR} X$ of a metric space X. By definition this is the infimum of all numbers $Q > 0$ for which there exists an Ahlfors Q-regular space Y with $X \overset{qs}{\approx} Y$. Whenever X is the boundary of a Gromov hyperbolic group G, the set of these numbers Q is nonempty. In particular, $\dim_{AR} \partial_\infty G$ is well defined, and it is not hard to show that $\dim_{AR} \partial_\infty G$ is at least as large as the topolgical dimension of $\partial_\infty G$.

Theorem 2 [BK2]**.** *Suppose G is a Gromov hyperbolic group with $\partial_\infty G \approx \mathbb{S}^2$. If there exists an Ahlfors Q-regular 2-sphere Z such that $\partial_\infty G \overset{qs}{\approx} Z$ and $Q = \dim_{AR} \partial_\infty G$, then G is standard.*

In other words, if the infimum by which $\dim_{AR} \partial_\infty G$ is defined is attained as a minimum, then G is standard. Note that Theorem 2 contains Theorem 1, because we have $\dim_{AR} \partial_\infty G \geq 2$.

In view of these results it seems worthwhile to study the general question when the Ahlfors regular conformal dimension of a fractal 2-sphere is attained as a minimum. Interesting examples are provided by post-critically finite rational maps R on the Riemann sphere $\overline{\mathbb{C}}$. The analog of the standard picture in the Kleinian group case is given by the following setting.

Let $R\,\overline{\mathbb{C}} \to \overline{\mathbb{C}}$ a holomorphic map of $\overline{\mathbb{C}}$ into itself, i.e., a rational function. Let Ω_R denote the set of critical points of R, and $P_R = \bigcup_{n\in\mathbb{N}} R^n(\Omega_R)$ be the set of post-critical points of R (here R^n denotes the nth iterate of R). We make the following assumptions on R:

(i) R is post-critically finite, i.e., P_R is a finite set,
(ii) R has no periodic critical point; this implies that $J_R = \overline{\mathbb{C}}$ for the Julia set of R,
(iii) the orbifold $\mathcal{O}_R$ associated with R is hyperbolic (see [DoHu] for the definition of $\mathcal{O}_R$); this implies that the dynamics of R on $J_R = \overline{\mathbb{C}}$ is expanding.

A characterization of post-critically finite rational maps is due to Thurston. The right framework is the theory of topologically holomorphic self-maps $f\ \mathbb{S}^2 \to \mathbb{S}^2$ of the sphere. By definition these maps have the local form $z \mapsto z^n$ with $n \in \mathbb{N}$ in appropriate local coordinates, and one defines the critical set, the post-critical set, and the associated orbifold similarly as for rational maps. In our context, Thurston's theorem can be stated as follows [DoHu]:

Theorem (Thurston)**.** *Let $f\ \mathbb{S}^2 \to \mathbb{S}^2$ be a post-critically finite topologically holomorphic map with hyperbolic orbifold. Then f is equivalent to a rational map R if and only if f has no "Thurston obstructions".*

Equivalence has to be understood in an appropriate sense. If f and R are both expanding, this just means conjugacy of the maps.

The definition of a Thurston obstruction is as follows. A *multicurve* $\Gamma = \{\gamma_1, \dots, \gamma_n\}$ is a system of Jordan curves in $\mathbb{S}^2 \setminus P_f$ with the following properties: the curves have pairwise empty intersection, are pairwise non-homotopic in $\mathbb{S}^2 \setminus P_f$, and non-peripheral (this means that each of the complementary components of a curve contains at least two points in P_f). A multicurve Γ is called *f-stable* if for all j every component of $f^{-1}(\gamma_j)$ is either peripheral or homotopic in $\mathbb{S}^2 \setminus P_f$ to one of the curves γ_i.

If Γ is an f-stable multicurve, fix i and j and label by α the components $\gamma_{i,j,\alpha}$ of $f^{-1}(\gamma_j)$ homotopic to γ_i in $\mathbb{S}^2 \setminus P_f$. Then f restricted to $\gamma_{i,j,\alpha}$ has a mapping degree $d_{i,j,\alpha} \in \mathbb{N}$. Let $m_{i,j} = \displaystyle\sum_\alpha \frac{1}{d_{i,j,\alpha}}$ and define the *Thurston matrix* $A(\Gamma)$ of the f-stable multicurve Γ by $A(\Gamma) = (m_{ij})$. This is a matrix with nonnegative coefficients; therefore, it has a largest eigenvalue $\lambda(f,\Gamma) \geq 0$. Then Γ is a *Thurston obstruction* if $\lambda(f,\Gamma) \geq 1$.

Post-critically finite rational maps are related to *subdivision rules* [CFP]. For example, if R is a real rational map (i.e, $R(\overline{\mathbb{R}}) \subseteq \overline{\mathbb{R}}$) satisfying the above conditions (i)–(iii), then $R^{-1}(\overline{\mathbb{R}})$ is a graph providing a subdivision of the upper and lower half-planes whose combinatorics determines R (up to conjugacy by a real Möbius transformation). The combinatorics of the graphs $R^{-n}(\overline{\mathbb{R}})$ is determined by iterating the subdivisions of the upper and lower half-planes by the complementary components of $R^{-1}(\overline{\mathbb{R}})$ n-times. One can ask whether every rational map satisfying (i)–(iii) (or at least a sufficiently high iterate) is associated with a (two-tile) subdivision rule. This reduces to the following problem.

Problem. *Let R be a rational function satisfying (i)–(iii). Does there exist a quasicircle $C \subseteq \overline{\mathbb{C}}$ such that $P_R \subseteq C$ and $C \subseteq R^{-1}(C)$?*

Conversely, one can start with a (two-tile) subdivision rule of the sphere $\mathbb{S}^2$. One can associate a natural (class of) metric(s) on $\mathbb{S}^2$ associated with a subdivision rule (of appropriate type). If we denote by X the sphere $\mathbb{S}^2$ equipped with this metric, then the subdivision rule produces a topologically holomorphic expanding map $fX \to X$ which is post-critically finite. It turns out that f is conjugate to a rational function R if and only if $X \stackrel{qs}{\approx} \mathbb{S}^2$. So we have a situation that is very similar to the Kleinian group setting. In view of this it would be very interesting to find $\dim_{AR} X$ for these fractal spheres. In discussions with L. Geyer and K. Pilgrim we were lead to a conjecture on the Ahlfors regular conformal dimension of these spaces X. To state this conjecture let $Q \geq 2$ and Γ be an f-stable multicurve, define the modified Thurston matrix $A(\Gamma, Q)$ as $A(\Gamma, Q) = (m^Q_{i,j})$, where $m^Q_{ij} = \sum_\alpha \frac{1}{d^{Q-1}_{i,j,\alpha}}$, and let $\lambda(f, \Gamma, Q)$ be the largest nonnegative eigenvalue of $A(\Gamma, Q)$.

Conjecture. *If X comes from a subdivision rule with associated expanding map f, then $\dim_{AR} X$ is the infimum of all $Q \geq 2$ such that $\lambda(f, \Gamma, Q) < 1$ for all f-stable multicurves Γ. Moreover, $\dim_{AR} X$ is never attained unless $X \stackrel{qs}{\approx} \mathbb{S}^2$.*

References

[BK1] M. Bonk and B. Kleiner, *Quasisymmetric parametrizations of two-dimensional metric spheres*, Invent. Math. 150 (2002), 127–183.

[BK2] M. Bonk and B. Kleiner, *Conformal dimension and Gromov hyperbolic groups with 2-sphere boundary*, to appear in: Geom. Topol.

[Ca] J.W. Cannon, *The combinatorial Riemann mapping theorem*, Acta Math. 173 (1994), 155–234.

[CFP] J.W. Cannon, W.J. Floyd, and W.R. Parry, *Finite subdivision rules*, Conform. Geom. Dyn. 5 (2001), 153–196.

[DoHu] A. Douady and J.H. Hubbard, *A proof of Thurston's topological characterization of rational functions*, Acta Math. 171 (1993), 263–297.

[GhHa] E. Ghys and P. de la Harpe, *Sur les Groupes Hyperboliques d'après Mikhael Gromov*, Birkhäuser, Progress in Mathematics, Boston-Basel-Berlin, 1990.

[Gr] M. Gromov, *Hyperbolic Groups*, in: Essays in Group Theory, S. Gersten, Editor, MSRI Publication, Springer-Verlag, 1987, pp. 75–265.

[Su] D. Sullivan, *On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions*, in: Riemann surfaces and related topics. Proceedings of the 1978 Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978), Princeton Univ. Press, Princeton, N.J., 1981, pp. 465–496.

[Tu] P. Tukia, *On quasiconformal groups*, J. Analyse Math. 46 (1986), 318–346.

Inverse Source Problem in a 3-D Ball from Meromorphic approximation on 2-D Slices

L. Baratchart
(joint work with J. Leblond and E.B. Saff)

1. Notations and Preliminaries

Let $\mathbb{T}$ be the unit circle, $\mathbb{D}$ the unit disk, $\mathcal{P}_K$ the set of probability measures on a compact set K, $\mathcal{P}_n$ the space of algebraic polynomials of degree $\leq n$, H^∞ the Hardy space of the disk, $H_n^\infty = \{h/q_n;\ h \in H^\infty,\ q_n \in \mathcal{P}_n\}$ the set of meromorphic functions with n poles in $\mathbb{D}$ that are bounded near the boundary, Ω the unit ball of $\mathbb{R}^3$ and $\mathcal{S}_2$ the unit sphere.

The Green capacity of K is the nonnegative number $C_{\mathbb{T},K}$ given by

$$\frac{1}{C_{\mathbb{T},K}} = \inf_{\mu \in \mathcal{P}_K} \int \int \log \left| \frac{1-\bar{t}z}{z-t} \right| d\mu(t) d\mu(x).$$

If $C_{\mathbb{T},K} > 0$, there is a unique measure $\omega_K \in \mathcal{P}_K$ to meet the infimum, called the Green equilibrium measure on K. The measure ω_K is difficult to compute in general, but charges the endpoints if K is a system of arcs. We need the notion of extremal domain, which is specialized below to the case of a disk

Theorem [8] *Let f be holomorphic in $\overline{\mathcal{C}_\varepsilon} = \{z;\ 1-\varepsilon < |z| < 1\}$ and continuous in $\overline{\overline{\mathcal{C}_\varepsilon}}$. Set*

$\mathcal{V}_f = \{V;\ V$ connected open in $\overline{\mathbb{D}}$ with $\overline{\mathcal{C}_\varepsilon} \subset V,\ f$ extends holomorphically to $V\}$.

There is a unique $V_m \in \mathcal{V}_f$ such that $C_{\mathbb{T},\mathbb{D}\setminus V_m} = \inf_{V \in \mathcal{V}_f} C_{\mathbb{T},\mathbb{D}\setminus V}$ which contains every other member of $\mathcal{V}_f$ with this property.

We shall be concerned here with the class :

$\mathcal{BLP} \triangleq \{f$ continuous in $\overline{\mathcal{C}_\varepsilon}$, holomorphic in $\overset{\circ}{\overline{\mathcal{C}_\varepsilon}}$,can be continued analytically in $\mathbb{D}$ except for finitely many poles,branchpoints, and log singularities$\}$

For such functions, more is known on the structure of extremal domains.

Theorem [9] *If $f \in \mathcal{BLP}$, then $\overline{\mathbb{D}} \setminus V_f$ consists of its poles, its branchpoints, its* log *singularities, and finitely many analytic cuts. A cut ends up either at a branchpoint, a* log *singularity, or at an end of another cut. The diagram thus formed has no loop.*

For more than two points, $\overline{\mathbb{D}} \setminus V_f$ is a trajectory of a rational quadratic differential, but there is no easy computation. The situation is similar to that in problems of Tchebotarev-Lavrentiev type, where one must find the continuum of minimal capacity that connects prescribed groups of points [6, 7]. The difference is that, here, the connectivity is not known *a priori.*

2. Meromorphic Approximation

By a best meromorphic approximant with at most n poles of f, we mean some $g_n \in H^\infty_n$ such that :

$$\|f - g_n\|_{L^\infty(\mathbb{T})} = \inf_{g \in H^\infty_n} \|f - g\|_{L^\infty(\mathbb{T})}.$$

Clearly this notion is conformally invariant.

By the Adamjan-Arov-Krein theory [1], a best meromorphic approximant with at most n poles uniquely exists provided that $f \in C(\mathbb{T})$. Moreover, it can be computed from the singular value decomposition of the Hankel operator with symbol f.

If g_n is the sequence of best meromorphic approximants to f, whose poles are numbered as ξ_j, n for $1 \leq j \leq d_n \leq n$, we form the sequence of counting probability measures $\mu_n = \sum_j \delta_{\xi_{j,n}}/d_n$.

Theorem [2] *If $f \in \mathcal{BLP}$ is not single-valued, the counting measure μ_n of the poles of its best meromorphic approximants converges weak* to the Green equilibrium distribution of $\overline{\mathbb{D}} \setminus V_f$. Moreover, each neighborhood of a pole of f contains at least one pole of the approximant as $n \to \infty$, and only finitely many poles can remain in a compact subset of V_f.*

3. An Inverse Source Problem in 3-D.

If we are given m_1 monopolar sources $S_1, \ldots, S_{m_1}$ and m_2 dipolar sources $C_1, \ldots, C_{m_2}$ in Ω, the potential u satisfies :

$$\begin{cases} -\Delta u = F \quad \text{in } \Omega \\[2ex] \dfrac{\partial u}{\partial \nu}_{|\mathcal{S}_2} = \phi \quad \text{current flux} \\[2ex] u_{|\mathcal{S}_2} = g \quad \text{electric potential} \end{cases}$$

$$F = \sum_{j=1}^{m_1} \lambda_j \, \delta_{S_j} + \sum_{k=1}^{m_2} p_k \, . \, \nabla \, \delta_{C_k}$$

The inverse problem is to locate the monopolar sources S_j with their intensities λ_j and the dipolar sources C_k with their momentums p_k from the knowledge of Φ and u on $\mathcal{S}_2$. Such problems arise in Electro-Encephalography, see for instance [3, 5].

The fundamental solution is $(4\pi\|X\|)^{-1}$ so the potential assumes the form :

$$u(X) = h(X) - \sum_{j=1}^{m_1} \frac{\lambda_j}{4\pi \, \|X - S_j\|} + 3 \sum_{k=1}^{m_2} \frac{< p_k \, , \, X - C_k >}{4\pi \, \|X - C_k\|^3}$$

where h is harmonic. Using the Green formula and the expansion into spherical harmonics, one can then recover $h_{|\mathcal{S}_2}$, although we do not explain this in details here. This is just to say we can assume $h = 0$ by subtraction. We shall assume that all sources lie in general position, in the sense that none of them lies on the vertical axis $\{x = y = 0\}$.

Put : $\xi_j = x_{S_j} + iy_{S_j}$ where $S_j = (x_{S_j}, y_{S_j}, z_{S_j})^T$, $\xi_k = x_{C_k} + iy_{C_k}$ where $C_k = (x_{C_k}, y_{C_k}, z_{C_k})^T$, and let the dipolar moments be expressed in coordinates as : $p_k = (p_{k,x}, p_{k,y}, p_{k,z})$.

When we slice the ball Ω along the horizontal plane $\{z = z_p\}$, the intersection with $\mathcal{S}_s$ is a circle $\mathcal{C}_p$ of radius r_p with $r_p^2 = 1 - z_p^2$. If we let $\xi = x + iy$ be the complex variable in the plane $\{z = z_p\}$, the restriction $g_{|\mathcal{C}_p}$ is the trace on $\mathcal{C}_p$ of the function $f(\xi)$ given by

$$\frac{i}{4\pi} \times \left[-\sum_{j=1}^{m_1} \frac{\Lambda_{j,p}}{(\xi - \xi_{j,p}^-)^{1/2}} + 3 \sum_{k=1}^{m_2} \frac{R_{k,p}(\xi)}{(\xi - \xi_{k,p}^-)^{3/2}} \right]$$

where

$$Q_{l,p}(\xi) = |\xi - \xi_l|^2 + (z_p - z_l)^2 = -\frac{1}{\xi}\, \xi_l\, (\xi - \xi_{l,p}^-)\, (\xi - \xi_{l,p}^+),\ l = \{j, k\},$$

with

$$|\xi_{l,p}^-| < r_p,\ |\xi_{l,p}^+| > r_p,\ \xi_{l,p}^-/\xi_l \in \mathbb{R},\ \xi_{l,p}^+/\xi_l \in \mathbb{R}$$

and where

$$\Lambda_{j,p} = \frac{\lambda_j \sqrt{\xi}}{\sqrt{\xi_j (\xi - \xi_{j,p}^+)}},$$

$$R_{k,p}(\xi) = \frac{\sqrt{\xi}\, \left[\tilde{p}_k \xi^2 + 2(p_{k,z} h_{p,k} - \operatorname{Re}\{\tilde{p}_k \xi_k\})\xi + \overline{\tilde{p}_k}\, r_p^2 \right]}{2\sqrt{\xi_k} (\xi - \xi_{j,p}^+)^{3/2}}$$

with

$$\tilde{p}_k = p_{k,x} - ip_{k,y} \text{ and } h_{p,k} = z_p - z_k.$$

Although $f(\xi)$ may not lie in $\mathcal{BLP}$, its square does. We can in principle locate the branchpoints using the convergence of poles in meromorphic approximation from the previous section. To solve the inverse problem, it remains to connect $\xi_{l,p}^-$ with the original sources :

Proposition *For f as above, each branchpoint $\xi_{j,p}^-$ or $\xi_{k,p}^-$ has maximum modulus when $z_p = z_{S_j}$ in which case they coincide with the corresponding source.*

References

[1] V. M. Adamyan, D. Z. Arov, and M. G. Kreĭn, Analytic properties of Schmidt pairs, Hankel operators, and the generalized Schur-Takagi problem, Mat. Sb., 128, 1971, 34–75. English transl. Math. USSR Sb., 15, 1971.

[2] L. Baratchart and H. Stahl, Asymptotic Spectrum of Hankel Operator with Branched Symbol. In preparation.

[3] M. Chafik, A. El Badia and T. Ha-Duong, On some inverse EEG problems. Inverse Problems in engineering mechanics II, 2000, 537–544.

[4] A. El Badia and T. Ha-Duong, An inverse Source Problem in Potential Analysis. Inverse Problems, 16, 2000, 651–663.
[5] M. Hämäläinen, R. Hari, J. Ilmoniemi, J. Knuutila, O.V. Lounasmaa, Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. Reviews of Modern Physics, 65(2), 1993, 413–497.
[6] G.V. Kuzmina, Moduli of Curves and Quadratic Differentials. Trudy Mat. Int. Steklov, 139, 1980. Engl. transl. Proc. Steklov Inst. Math., 139, 1982.
[7] G.V. Kuzmina, Methods of Geometric Function Theory, I, II. Algebra and Analysis, 9, 1997. Engl. transl. St. Petersburg Math. J., 9, 1998, 455–507, 889–930. Trudy Mat. Int. Steklov, 139, 1980. Engl. transl. Proc. Steklov Inst. Math., 139, 1982.
[8] H. Stahl, Extremal Domains Associated with an Analytic Function, I, II. Complex variables, 4, 1985, 311–324, 325–338.
[9] H. Stahl, The Structure of Extremal Domains Associated with an Analytic Function. Complex variables, 4, 1985, 339–354.

Reporter: Oliver Roth

Participants

Prof. Dr. Vladimir Andrievskii
andriyev@mcs.kent.edu
Department of Mathematics
Kent State University
Kent OH 44242-0001 – USA

Prof. Dr. Alexander I. Aptekarev
aptekaa@spp.keldysh.ru
M.V. Keldysh Institute of Applied Mathematics
Russian Academy of Sciences
Miusskaya pl. 4
125047 Moscow – Russia

Dr. Laurent Baratchart
laurent.baratchart@sophia.inria.fr
INRIA Sophia Antipolis
B.P. 93
2004 Route des Lucioles
F-06902 Sophia Antipolis Cedex

Prof. Dr. Roger W. Barnard
barnard@math.ttu.edu
Department of Mathematics
Texas Tech. University
Lubbock, TX 79409-1042 – USA

Prof. Dr. Walter Bergweiler
bergweiler@math.uni-kiel.de
Mathematisches Seminar
Christian-Albrechts-Universität Kiel
D–24098 Kiel

Prof. Dr. Hans-Peter Blatt
mga009@ku-eichstaett.de
hans.blatt@ku-eichstaett.de
Mathematisch-Geographische Fakultät
Kath. Universität Eichstätt
Ostenstr. 26-28
D–85072 Eichstätt

Prof. Dr. Mario Bonk
mbonk@umich.edu
University of Michigan
Department of Mathematics
2074 East Hall
Ann Arbor MI 48109-1109 – USA

Prof. Dr. Rainer Brück
Rainer.Brueck@Math.uni-dortmund.de
Mathematisches Institut
Universität Dortmund
Campus Nord
Vogelpothsweg 87
D–44227 Dortmund

Dr. Ihor Chyzhykov
matstud@franko.lviv.ua
matstud@uli2.franko.lviv.ua
Faculty of Mechanics and Mathematics, Lviv Ivan Franko National University
1 Universytetska str.,
79000 Lviv – Ukraine

Dr. Michael Eiermann
eiermann@mathe.tu-freiberg.de
Fakultät für Mathematik und Informatik; Technische Universität Bergakademie Freiberg
Agricolastr. 1
D–09599 Freiberg

Prof. Dr. Alex E. Eremenko
eremenko@math.purdue.edu
Dept. of Mathematics
Purdue University
West Lafayette, IN 47907-1395 – USA

Markus Förster
m.foerster@iu-bremen.de
School of Engineering and Science
International University Bremen
D–28725 Bremen

Dr. Richard Fournier
fournier@dms.umontreal.ca
Departement de Mathematiques
Universite de Montreal
Pavillon Andre-Aisenstadt
2920 Chemin de la Tour
Montreal Quebec H3T 1J8 – Canada

Dr. Lukas Geyer
geyer@math.uni-dortmund.de
lgeyer@umich.edu
Dept. of Mathematics
The University of Michigan
525 East University Avenue
Ann Arbor, MI 48109-1109 – USA

Dr. Richard Greiner
greiner@mathematik.uni-wuerzburg.de
Mathematisches Institut
Universität Würzburg
Am Hubland
D–97074 Würzburg

Prof. Dr. Walter K. Hayman
w.hayman.ic.ac.uk
Department of Mathematics
Imperial College London
Huxley Building
GB-London SW7 2AZ

Martin Hemke
hemke@math.uni-kiel.de
Mathematisches Seminar
Christian-Albrechts-Universität Kiel
D–24098 Kiel

Prof. Dr. Aimo Hinkkanen
aimo@math.uiuc.edu
Dept. of Mathematics, University of
Illinois at Urbana-Champaign
273 Altgeld Hall MC-382
1409 West Green Street
Urbana, IL 61801-2975 – USA

Prof. Dr. Dmitry Khavinson
dmitry@uark.edu
Dept. of Mathematics,SE 301
University of Arkansas
Fayetteville, AR 72701 – USA

Daniela Kraus
dakraus@mathematik.uni-wuerzburg.de
Mathematisches Institut
Universität Würzburg
Am Hubland
D–97074 Würzburg

Dr. Arno Kuijlaars
arno.kuijlaars@wis.kuleuven.ac.be
Department of Mathematics
Katholieke Universiteit Leuven
Celestijnenlaan 200 B
B-3001 Leuven

Prof. Dr. Ilpo Laine
ilpo.laine@joensuu.fi
Department of Mathematics
University of Joensuu
P. O. Box 111
FIN-80101 Joensuu 10

Prof. Jim Langley
jkl@maths.nott.ac.uk
james.langley@nottingham.ac.uk
School of Mathematical Sciences
University of Nottingham
University Park
GB-Nottingham NG7 2RD

Prof. Dr. Norman Levenberg
levenber@math.auckland.ac.nz
nlevenbe@ucs.indiana.edu
8315 Parkview Avenue
Munster Indiana 46321 – USA

Prof. Dr. Eli A.L. Levin
elile@oumail.openu.ac.il
Department of Mathematics
The Open University of Israel
16, Klausner st.
P. O. Box 39 328
Tel Aviv 61392 – Israel

Prof. Dr. Yura Lyubarskii
yura@math.ntnu.no
Dept. of Mathematical Sciences
Norwegian University of Science
and Technology
A. Getz vei 1
N-7491 Trondheim

Prof. Dr. Raymond Mortini
mortini@poncelet.univ-metz.fr
mortini@poncelet.sciences.univ-metz.fr
Departement de Mathematiques
Universite de Metz
UFR M.I.M.
Ile du Saulcy
F-57045 Metz

Prof. Dr. Nicolas Papamichael
nickp@ucy.ac.cy
Department of Mathematics and
Statistics
University of Cyprus
P.O. Box 20537
1678 Nicosia – Cyprus

Prof. Dr. Franz Peherstorfer
franz.peherstorfer@jku.at
Institut für Mathematik
Universität Linz
Altenberger Str. 69
A-4040 Linz

Dr. Igor E. Pritsker
igor@math.okstate.edu
Dept. of Mathematics
Oklahoma State University
401 Math Science
Stillwater, OK 74078-1058 – USA

Prof. Dr. Mihai Putinar
mputinar@math.ucsb.edu
Department of Mathematics
University of California at Santa Barbara
Santa Barbara, CA 93106 – USA

Lasse Rempe
lasse@maths.warwick.ac.uk
lasse@math.uni-kiel.de
Mathematisches Seminar
Christian-Albrechts-Universität Kiel
Ludewig-Meyn-Str. 4
D–24118 Kiel

Dr. Oliver Roth
roth@mathematik.uni-wuerzburg.de
Mathematisches Institut
Universität Würzburg
Am Hubland
D–97074 Würzburg

Günter Rottenfußer
g.rottenfusser@iu-bremen.de
International University Bremen
School of Engineering and Science
Postfach 750561
D–28725 Bremen

Prof. Dr. Stephan Ruscheweyh
ruscheweyh@mathematik.uni-wuerzburg.de
Mathematisches Institut
Universität Würzburg
Am Hubland
D–97074 Würzburg

Prof. Dr. Edward B. Saff
esaff@math.usf.edu
esaff@math.vanderbilt.edu
Dept. of Mathematics
Vanderbilt University
1326 Stevenson Center
Nashville TN 37240-0001 – USA

Dr. Eric Schippers
schip@umich.edu
Department of Mathematics
University of Michigan
1863 East Hall
Ann Arbor MI 48109-1109 – USA

Prof. Dr. Dierk Schleicher
dierk@iu-bremen.de
School of Engineering and Science
International University Bremen
Postfach 750561
D–28725 Bremen

Prof. Dr. Gerhard Schmeisser
schmeisser@mi.uni-erlangen.de
Mathematisches Institut
Universität Erlangen-Nürnberg
Bismarckstr. 1 1/2
D–91054 Erlangen

Prof. Dr. Gerald Schmieder
schmieder@math.uni-oldenburg.de
schmieder@mathematik.uni-oldenburg.de
Fakultät V - Institut f. Mathematik
Carl-von-Ossietzky-Universität
Oldenburg
D–26111 Oldenburg

Gunter Semmler
semmler@math.tu-freiberg.de
Institut für Angewandte Mathematik I
TU Bergakademie Freiberg
Agricolastraße 1
D–09596 Freiberg

Prof. Dr. Herbert Stahl
stahl@tfh-berlin.de
Fachbereich 2 - Mathematik
Technische Fachhochschule Berlin
Luxemburger Str. 10
D–13353 Berlin

Prof. Dr. Norbert Steinmetz
stein@math.uni-dortmund.de
Fachbereich Mathematik
Universität Dortmund
D–44221 Dortmund

Prof. Dr. Kenneth Stephenson
kens@math.utk.edu
Department of Mathematics
University of Tennessee
121 Ayres Hall
Knoxville, TN 37996-1300 – USA

Prof. Dr. Marcus Stiemer
stiemer@math.uni-dortmund.de
Institut für Angewandte Mathematik
Universität Dortmund
Vogelpothsweg 87
D–44227 Dortmund

Dr. Nikos S. Stylianopoulos
nikos@ucy.ac.cy
Department of Mathematics & Statistics
University of Cyprus
P.O. Box 20537
1678 Nicosia – Cyprus

Prof. Dr. Vilmos Totik
totik@math.u-szeged.hu
Bolyai Institute
University of Szeged
Aradi Vertanuk Tere 1
H-6720 Szeged

Prof. Dr. Elias Wegert
wegert@math.tu-freiberg.de
Fakultät für Math. und Informatik
Institut für Angewandte Analysis
TU Bergakademie Freiberg
Agricolastr. 1
D–09596 Freiberg

Dr. G. Brock Williams
williams@math.ttu.edu
williams@koch.math.ttu.edu
Department of Mathematics and Statistics
Texas Tech University
Lubbock TX 79409-1042 – USA

Prof. Dr. Lawrence Zalcman
zalcman@macs.biu.ac.il
Dept. of Mathematics
Bar-Ilan University
52 900 Ramat-Gan – Israel

Mathematisches Forschungsinstitut Oberwolfach

Report No. 8/2004

Mini-Workshop: Nonlinear Spectral and Eigenvalue Theory with Applications to the p-Laplace Operator

Organised by
Jürgen Appell (Würzburg)
Pavel Drabek (Plzen)
Raffaele Chiappinelli (Siena)

February 15th – February 21st, 2004

Introduction by the Organisers

What is the state-of-the-art of abstract spectral and eigenvalue theory for nonlinear operators, and how may this theory be applied to nonlinear equations involving the p-Laplace operator? These two questions have provided the main focus of the Mini-Workshop. Accordingly, the main topics covered by the talks on this Mini-Workshop have been

- spectra for nonlinear operators,
- nonlinear eigenvalue problems, and
- equations involving the p-Laplace operator.

Of course, these three topics are not mutually independent, but there are various interconnections between them which are of particular interest. For example, sets of eigenvalues (point spectra) may be regarded, as in the linear case, as an important part of the spectrum; conversely, nonlinear eigenvalue theory is one of the historical roots of nonlinear spectral theory. Moreover, the p-Laplace operator is one of the most interesting homogeneous (though nonlinear) operators which may not only serve as a "model operator" in nonlinear eigenvalue problems, but also occurs quite frequently in various applications to physics, mechanics, and elasticity.

The aim and scope of the Mini-Workshop was to bring together experts on nonlinear spectral analysis and operator theory, on the one hand, and more application-oriented specialists in eigenvalue problems for nonlinear partial differential equations (like the p-Laplace equation), on the other. As a result, 15 leading experts in the field from 10 different countries discussed recent progress and open problems in the theory, methods, and applications of spectra and eigenvalues of nonlinear operators.

Mini-Workshop: Nonlinear Spectral and Eigenvalue Theory with Applications to the p-Laplace Operator

Table of Contents

Abstracts

Asymmetric Eigenvalue Problems with Weights for the p-laplacian with Neumann Boundary Conditions

M. Cuesta (Calais)
(joint work with M. Arias (Granada), J.-P. Gossez (Bruxelles))

The motivation of this work is the study of

$$-\Delta_p u = f(x,u) \text{ in } \Omega, \quad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega, \tag{1}$$

where $\Delta_p u := \operatorname{div}(|\nabla u|^{p-2}\nabla u)$, $1 < p < \infty$, and Ω is a bounded smooth domain of $\mathbb{R}^N$ and $|f(x,s)| \leq a(x)|s|^{p-1} + b(x)$ with a, b belonging to suitable Lebesgue spaces. Our ultimate goal is to find optimal conditions on the limits at $+\infty$ and $-\infty$ of the quotients $f(x,s)/|s|^{p-2}s$ and $pF(x,s)/|s|^p$ (where $F(x,s) := \int_0^s f(x,t)\,dt$) as $s \to +\infty$ and $s \to -\infty$ to assure solvability of (1). When considering $m(x) = \lim_{s\to+\infty} \frac{f(x,s)}{|s|^{p-2}s}$, $n(x) = \lim_{s\to-\infty} \frac{f(x,s)}{|s|^{p-2}s}$, we are lead to study *weighted asymmetric eigenvalue problems* of the form

$$-\Delta_p u = \lambda(m(x)(u^+)^{p-1} - n(x)\,(u^-)^{p-1}) \text{ in } \Omega, \quad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega \tag{2}$$

We will always assume that the weights $m(x)$ and $n(x)$ are possibly non constant, different, indefinite and belong to $L^r(\Omega)$ where $r > N/p$ if $p \leq N$ and $r = 1$ if $p > N$. We will also assume that m^+ and $n^+ \not\equiv 0$ and we are only interested on *positive eigenvalues.* Notice that 0 is always an eigenvalue of (2).

The case $m(x) \equiv n(x)$ have been studied [5]. When $m(x)$ are $n(x)$ are constant and different, (2) leads to the notion of Fučik spectrum and the so-called problems of Ambrosetti-Prodi type. Analogous problems (1) and (2) have been treated with *Dirichlet boundary conditions* by [1].

The study of (2) start with the following symmetric eigenvalue problem

$$-\Delta_p u = \lambda m(x)|u|^{p-2}u \text{ in } \Omega, \quad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega. \tag{3}$$

The following value introduced by [5] plays a crucial role:

$$\lambda^*(m) := \inf\{\int_\Omega |\nabla u|^p \, dx \; : \; \int_\Omega m\,|u|^p \, dx = 1\}.$$

which satisfies: (1) If $\int_\Omega m\,dx < 0$ then $\lambda^*(m) > 0$ is the unique non zero principal eigenvalue, it admits a non negative eigenfunction and there is no eigenvalue on $]0, \lambda^*(m)[$. (2) If $\int_\Omega m\,dx > 0$ then $\lambda^*(m) = 0$ is the unique non negative principal eigenvalue and (3) If $\int_\Omega m\,dx = 0$ then $\lambda^*(m) = 0$ is the unique principal eigenvalue. Besides a sequence of eigenvalues can be constructed using the Ljusternik-Schnirelmann critical point theory, cf. [3].

It follows straightforward that the principal eigenvalues of (2) are $\lambda = \lambda_1(m)$ and $\lambda = \lambda_1(n)$. We present in this work a construction of a *non principal eigenvalue of* (2) by considering the functionals $A(u) := \int_\Omega |\nabla u|^p$, $B_{m,n}(u) : \int_\Omega (m(u^+)^p +$

$n(u^-)^p)$ and $\tilde{A}$ the restriction of A to the C^1 manifold $M_{m,n} := \{u \in W_0^{1,p}(\Omega) : B_{m,n}(u) = 1\}$. We prove

Theorem 1. *Let* $\Gamma := \{\gamma \in C([0,1], M_{m,n}) \,:\, \gamma(0) \le 0$ *and* $\gamma(1) \ge 0\}$. *Then*
(1) $\Gamma \neq \emptyset$.
(2) *The value* $c(m,n) := \inf_{\gamma\in\Gamma} \max_{u\in\gamma[0,1]} \tilde{A}(u)$ *is an eigenvalue of* (2) *which satisfies*
$$c(m,n) > \max\{\lambda^*(m), \lambda^*(n)\}.$$
(3) *There is no eigenvalues of* (2) *between* $\max\{\lambda^*(m), \lambda^*(n)\}$ *and* $c(m,n)$.

The proof of this theorem relies on a critical point theorem of [2] for C^1 functionals restricted to C^1-manifolds that satisfy the *Palais-Smale condition of Cerami* (denoted (PSC) for short). This is one of main issues of the paper. Presicely we can prove that (1) $\tilde{A}$ satisfies $(PS)_c$ along bounded sequences $\forall c \ge 0$, (2) $\tilde{A}$ satisfies $(PSC)_c$ $\forall c > 0$, (3) if $\int_\Omega m\,dx \neq 0$ and $\int_\Omega n\,dx \neq 0$ then $\tilde{A}$ satisfies $(PS)_c$ for all $c \ge 0$, (4) if $\int_\Omega m\,dx = 0$ or $\int_\Omega n\,dx = 0$ then $\tilde{A}$ does not satisfy $(PSC)_0$ and (5) if $p = 2$ then $\tilde{A}$ satisfies $(PS)_c$ for all $c > 0$.

As an application of our main theorem we study the Fučik spectrum with weights. This spectrum is defined as the set $\Sigma(m,n)$ of those $(\alpha,\beta) \in \mathbb{R}^2$ such that

$$-\Delta_p u = \alpha m(x)(u^+)^{p-1} - \beta n(x)(u^-)^{p-1} \text{ in } \Omega, \quad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega \tag{4}$$

has a nontrivial solution. If we denote by $\Sigma^*(m,n)$ the set $\Sigma(m,n)$ amputed of the lines $\{\lambda^*(m)\} \times \mathbb{R}$ et $\mathbb{R} \times \{\lambda^*(n)\}$, we prove that for any $s > 0$, the line $\beta = s\alpha$ in the (α,β) plane intersects $\Sigma^*(m,n) \cap (\mathbb{R}^+ \times \mathbb{R}^+)$. Moreover the first point in this intersection is given by $\alpha(s) = c(m, sn)$, $\beta(s) = s\alpha(s)$.

We obtain in this way a first curve $\mathcal{C} := \{(\alpha(s), \beta(s)) \,:\, s > 0\}$ in $\Sigma^*(m,n) \cap (\mathbb{R}^+ \times \mathbb{R}^+)$.

References

[1] M. Arias, J. Campos, M. Cuesta and J.-P. Goessz, *Asymmetric elliptic problems with indefinite weights*, AIHP-AN **19,5** (2002), 581–616.

[2] M. Cuesta, *Mimimax Theorems on C^1-manifolds via Ekeland variational principle*, Abstract and Applied Analysis **13** (2003), 757–768.

[3] A. Dakkak, *Etude sur le spectre et la resonance pour des problemes elliptiques de Neumann*, These de 3eme cycle, Univ. Oujda, 1995.

[4] P. Drabek, *Solvability and bifurcations of nonlinear equations*, Pitman Research Notes in Mathematics, 264 (1992).

[5] T. Godoy, J.-P. Gossez and S. Paczka, *On the antimaximum principle for the p-laplacian with indefinite weight*, NonLinear Analysis **51** (2002), 449–467.

Antimaximum principle and Fučik spectrum
J.-P. Gossez (Bruxelles)

It is well-known that the antimaximum principle holds uniformly for the problem

$$\begin{cases} -u'' = \lambda u + h(x) & \text{on }]0,\pi[, \\ u'(0) = u'(\pi) = 0, \end{cases} \tag{5}$$

and that the interval of uniformity is $\lambda \in]0, 1/4[$. It is also well-known that the first curve in the Fučik spectrum for the problem

$$\begin{cases} -u'' = \alpha u^+ - \beta u^- & \text{on }]0,\pi[, \\ u'(0) = u'(\pi) = 0 \end{cases} \tag{6}$$

exhibits a gap at infinity with respect to the trivial horizontal and vertical lines, and that the value of this gap is equal to 1/4. When the Neumann conditions are replaced in (5) and (6) by the Dirichlet conditions, the antimaximum principle does not hold uniformly and there is no gap at infinity in the Fučik spectrum.

It is our purpose in this talk to survey some results which show that the above qualitative and quantitative correspondance between "uniformity of the antimaximum principle" and "gap at infinity in the Fučik spectrum" holds in various other situations (general elliptic operators, p-laplacian). However it does not hold anymore in general when weights are introduced.

References

[1] D. De Figueiredo and J.-P. Gossez, *On the first curve of the Fučik spectrum of an elliptic operator*, Diff. Int. Eq., **7** (1994), 1285–1302.

[2] F. De Thelin, J. Fleckinger, J.-P. Gossez and P. Takac, *Existence, nonexistence et principe de l'antimaximum pour le p-laplacien*, C. R. Ac. Sc. Paris, **321** (1995), 731–734.

[3] M. Arias, J. Campos and J.-P. Gossez, *Antimaximum principle and Fučik spectrum for the Neumann p-laplacian*, Diff. Int. Equat., **13** (2000), 217–226.

[4] M. Cuesta, D. De Figueiredo and J.-P. Gossez, *The beginning of the Fučik spectrum of the p-laplacian*, J. Diff. Equat., **159** (1999), 212–238.

[5] M. Arias, J. Campos, M. Cuesta and J.-P. Gossez, *Asymmetric elliptic problems with indefinite weights*, Annales Inst. H. Poincaré, Analyse Non Linéaire, **19** (2002), 581–616.

[6] T. Godoy, S. Paczka and J.-P. Gossez, *On the antimaximum principle for the p-laplacian with weight*, Nonlinear Analysis, TMA, **51** (2002), 449–467.

[7] T. Godoy, S. Paczka and J.-P. Gossez, *Minimax formula for the principal eigenvalue and application to the antimaximum principle*, Calculus Variations and P.D.E., to appear.

[8] J. Fleckinger, F. De Thelin and J.-P. Gossez, *Antimaximum principle in $\mathbb{R}^N$: local versus global*, J.Diff.Equat., **196** (2004), 119–133.

The Fredholm alternative for the p-Laplacian: bifurcation from infinity, existence and multiplicity

P. Drábek (Rostock), P. Girg (Plzeň), P. Takáč (Rostock)

We discuss the existence and multiplicity of solutions to the following boundary-value problem for the Dirichlet p-Laplacian in a bounded domain $\Omega \subset \mathbb{R}^N$:

$$\left\{ \begin{array}{rcll} -\Delta_p u - \lambda\, |u|^{p-2}\, u & = & f(x) & \text{in } \Omega\,; \\ u & = & 0 & \text{on } \partial\Omega\,. \end{array} \right. \tag{7}$$

Here, $\Delta_p u \stackrel{\text{def}}{=} \operatorname{div}(|\nabla u|^{p-2}\, \nabla u)$ where $p \in (1,\infty)$ is a fixed number, $f \in L^\infty(\Omega)$, and $\lambda \in \mathbb{R}$ is *spectral* parameter. Given $\lambda \in \mathbb{R}$, the solvability of (7) is closely related to the existence of a nontrivial solution of the corresponding eigenvalue problem

$$\left\{ \begin{array}{rcll} -\Delta_p u & = & \lambda\, |u|^{p-2}\, u & \text{in } \Omega\,; \\ u & = & 0 & \text{on } \partial\Omega\,, \end{array} \right. \tag{8}$$

which is nonlinear if $p \neq 2$ and linear for $p = 2$.

The first results applicable to the solvability of (7) go back to the works of Fučík et al. [10] and Pohozaev [11]: *If $\lambda \in \mathbb{R}$ is not an eigenvalue of* (8) *then* (7) *has at least one solution for any $f \in W^{-1,p'}(\Omega)$, $p = p/(p-1)$.*

Let $\lambda_1 > 0$ be the principal eigenvalue of $-\Delta_p$ subject to homogeneous Dirichlet boundary conditions. We concentrate on the behavior of the solutions under the assumption that λ is near λ_1 (and possibly $\lambda = \lambda_1$). Our main tool combines bifurcation theory and asymptotic estimates.

We first motivate our results by considering the linear boundary value problem

$$\left\{ \begin{array}{rl} -\Delta u - \lambda u = f & \text{in } \Omega\,; \\ u = 0 & \text{on } \partial\Omega\,, \end{array} \right. \tag{9}$$

which corresponds to $p = 2$ in (7). Let $f \in L^2(\Omega)$ be given, $f \not\equiv 0$. Then the set of all pairs $(\lambda, u) \in (-\infty, \lambda_2) \times W_0^{1,2}(\Omega)$ that satisfy (9) can be interpreted by means of a bifurcation diagram in $\mathbb{R} \times W_0^{1,2}(\Omega)$. Namely, let us write $u = c\varphi_1 + u^\top$ with $\int_\Omega u^\top \varphi_1 \,\mathrm{d}x = 0$. Here, φ_1 is the eigenfunction of the positive Dirichlet Laplacian $-\Delta$ associated with the (simple) principal eigenvalue λ_1 that is normalized by $\varphi_1 > 0$ in Ω and $\int_\Omega \varphi_1^2 \,\mathrm{d}x = 1$, and λ_2 stands for the second eigenvalue of $-\Delta$. Then problem (9) is equivalent to

$$\left\{ \begin{array}{rl} -\Delta u^\top - \lambda u^\top + (\lambda_1 - \lambda) c \varphi_1 = f^\top + a\varphi_1 & \text{in } \Omega\,; \\ u^\top = 0 & \text{on } \partial\Omega\,, \end{array} \right.$$

where $\int_\Omega f^\top \varphi_1 \,\mathrm{d}x = 0$ and $a = \int_\Omega f \varphi_1 \,\mathrm{d}x$. Clearly, $(\lambda_1 - \lambda)c = a$. The linear Fredholm alternative implies that the problem

$$\left\{ \begin{array}{rl} -\Delta u^\top - \lambda u^\top = f^\top & \text{in } \Omega\,; \\ u^\top = 0 & \text{on } \partial\Omega\,, \end{array} \right.$$

has a unique solution $u^\top \in W_0^{1,2}(\Omega)$ with $\int_\Omega u^\top \varphi_1 \,\mathrm{d}x = 0$. We have the following two different cases:

(i) If $\int_\Omega f\varphi_1 \,\mathrm{d}x = 0$ then
   (a) for any $\lambda \in (-\infty, \lambda_1) \cup (\lambda_1, \lambda_2)$, problem (9) has a unique solution $u_\lambda = u^\top$;
   (b) for $\lambda = \lambda_1$, all solutions of problem (9) can be written in the form $u_{\lambda_1} = c\varphi_1 + u^\top$ with $c \in \mathbb{R}$ arbitrary.

(ii) If $\int_\Omega f\varphi_1 \,\mathrm{d}x \neq 0$ then
   (a) there is no solution of (9) for $\lambda = \lambda_1$;
   (b) for any $\lambda \in (-\infty, \lambda_1)\cup(\lambda_1, \lambda_2)$ there is a unique solution of (9) expressed by $u_\lambda = c\varphi_1 + u^\top$ where
$$c = (\lambda_1 - \lambda)^{-1} \int_\Omega f\varphi_1 \,\mathrm{d}x.$$

The solution pairs $(\lambda, u) \in \mathbb{R} \times W_0^{1,2}(\Omega)$ of (9) can thus be sketched in the bifurcation diagrams indicated in Figure 1

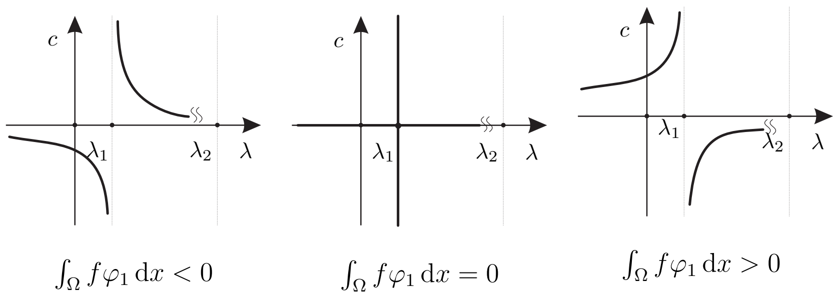


FIGURE 1. Bifurcations from infinity of solutions to (9), $c \stackrel{\text{def}}{=} \int_\Omega u\varphi_1 \,\mathrm{d}x$.

Motivated by this picture of the solution set of (9), we have decided to study the nonlinear problem (7) for $p \neq 2$ and to investigate the solution pairs $(\lambda, u) \in \mathbb{R} \times W_0^{1,p}(\Omega)$ for λ near λ_1. Again, φ_1 is the eigenfunction of the positive p-Laplacian associated with λ_1 and normalized by $\varphi_1 > 0$ and $\int_\Omega \varphi_1^p \,\mathrm{d}x = 1$. Notice that $a = \left(\int_\Omega \varphi_1^2 \,\mathrm{d}x\right)^{-1} \int_\Omega f\varphi_1 \,\mathrm{d}x$.

The existence of solutions $(\lambda, u) \in \mathbb{R} \times W_0^{1,p}(\Omega)$ to (7) with $\lambda \to \lambda_1$ and $\|u\|_{W_0^{1,p}(\Omega)} \to \infty$ is guaranteed by Dancer's type global bifurcation result for bifurcations from infinity at $\lambda = \lambda_1$. Roughly speaking, two continua $\mathcal{C}^\pm \subset R \times W_0^{1,p}(\Omega)$ of solutions to (7) emanate from (λ_1, ∞). Moreover, $\lambda \to \lambda_1$, $\|u\|_{W_0^{1,p}(\Omega)} \to \infty$ and $u \in \mathcal{C}^\pm$ imply $u/\|u\|_{W_0^{1,p}(\Omega)} \to \pm\varphi_1/\|\varphi_1\|_{W_0^{1,p}(\Omega)}$. If there is no sequence $\{(\lambda_1, u_n)\}_{n=1}^\infty$ of solutions to (7) such that $\|u_n\|_{W_0^{1,p}(\Omega)} \to \infty$, these two continua satisfy some very important global properties in addition; we refer to [4, 8] for a precise statement of this result.

We will establish an asymptotic estimate that plays the key role in the study of the structure of the solution set to (7). We assume $1 < p < \infty$, $p \neq 2$, if not explicitely mentioned otherwise. From now on, we denote by λ_2 ($\lambda_2 > \lambda_1$) the second eigenvalue of the positive Dirichlet p-Laplacian $-\Delta_p$. We use only the well-known fact from [2] that there is no eigenvalue of $-\Delta_p$ in the open interval (λ_1, λ_2), by a variational characterization of λ_2. All results presented here have been proved and reported in [8].

For the behavior of solutions u with large norm, the following a priori estimate plays the key role. We introduce some notation first. We introduce a new norm on $W_0^{1,p}(\Omega)$ by

$$\|v\|_{\mathcal{D}_{\varphi_1}} \stackrel{\text{def}}{=} \left(\int_\Omega |\nabla\varphi_1|^{p-2} |\nabla v|^2 \, \mathrm{d}x \right)^{1/2} \quad \text{for } v \in W_0^{1,p}(\Omega), \tag{10}$$

and denote by $\mathcal{D}_{\varphi_1}$ the completion of $W_0^{1,p}(\Omega)$ with respect to this norm. The Hilbert space $\mathcal{D}_{\varphi_1}$ is compactly imbedded in the Lebesgue space $L^2(\Omega)$; see [13, Lemma 4.2]. It is also shown there that the seminorm (10) is in fact a norm on $W_0^{1,p}(\Omega)$, if $2 < p < \infty$. For the case $1 < p < 2$ the space $\mathcal{D}_{\varphi_1}$ needs to be redefined. We do not need it for the formulation of any theorem here. Therefore its definition is omitted though it plays a key role in the proofs (see [8, 12, 13] for details).

For the sake of brevity, we also define

$$\mathcal{A}_{\varphi_1} \stackrel{\text{def}}{=} |\nabla\varphi_1|^{p-2} \left(\mathbf{I} + (p-2) \frac{\nabla\varphi_1 \otimes \nabla\varphi_1}{|\nabla\varphi_1|^2} \right) \quad \text{for } \nabla\varphi_1 \in \mathbb{R}^N, \ \nabla\varphi_1 \neq \mathbf{0} \in \mathbb{R}^N,$$

with $\mathbf{I}$ being the $n \times n$ identity matrix and $\otimes$ the tensor product.

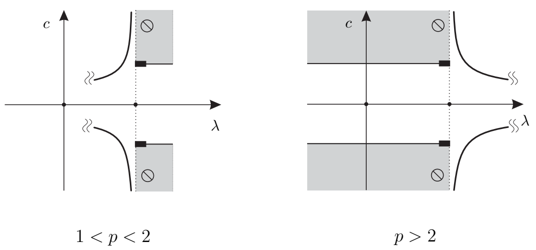


FIGURE 2. A priori bounds and bifurcations from infinity of solutions to (7) for $p > 1$, $p \neq 2$ and $a = 0$. There is no solution in the shaded regions (owing to a priori bounds).

Theorem 2. ([8, Thm. 4.1]) *Let* $\{\mu_n\}_{n=1}^\infty \subset \mathbb{R}$, $\{f_n\}_{n=1}^\infty \subset L^\infty(\Omega)$, $\{u_n\}_{n=1}^\infty \subset W_0^{1,p}(\Omega)$ *be sequences, and let* $\delta > 0$ *be such that*

(i) $\lambda_1 + \mu_n < \lambda_2 - \delta$ *for all* $n \in \mathbb{N}$*;*
(ii) $f_n \stackrel{*}{\rightharpoonup} f$ *weakly-star in* $L^\infty(\Omega)$*;*
(iii) $\|u_n\|_{W_0^{1,p}(\Omega)} \to \infty$ *as* $n \to \infty$*;*
(iv) *in addition, assume that for all* $n \in \mathbb{N}$ *and* $\phi \in W_0^{1,p}(\Omega)$,

$$\int_\Omega |\nabla u_n|^{p-2} \langle \nabla u_n, \nabla\phi\rangle \,\mathrm{d}x = (\lambda_1+\mu_n) \int_\Omega |u_n|^{p-2} u_n \,\phi \,\mathrm{d}x + \int_\Omega f_n \,\phi\,\mathrm{d}x\,. \tag{11}$$

Then $\mu_n \to 0$ and, writing $u_n = t_n^{-1}(\varphi_1 + v_n^\top)$ with $t_n \in \mathbb{R}$, $t_n \neq 0$, and $v_n^\top \in W_0^{1,p}(\Omega)^\top$, we have $t_n \to 0$, $|t_n|^{-p}\, t_n\, v_n^\top \to V^\top$ strongly in $\mathcal{D}_{\varphi_1}$ if $p>2$ and in $W_0^{1,2}(\Omega)$ if $1<p<2$, and

$$\begin{aligned} \mu_n = &-|t_n|^{p-2}\, t_n \int_\Omega f_n\,\varphi_1\,\mathrm{d}x + (p-2)\,|t_n|^{2(p-1)}\ \mathcal{Q}_0(V^\top, V^\top) \\ &+ (p-1)\,|t_n|^{2(p-1)} \left(\int_\Omega f\,\varphi_1\,\mathrm{d}x\right)\left(\int_\Omega \varphi_1^{p-1}\,V^\top\,\mathrm{d}x\right) + o\left(|t_n|^{2(p-1)}\right). \end{aligned} \tag{12}$$

In particular, if $\int_\Omega f_n\,\varphi_1\,\mathrm{d}x = 0$ for all $n\in\mathbb{N}$, then

$$\mu_n = (p-2)\,|t_n|^{2(p-1)}\ \mathcal{Q}_0(V^\top,V^\top) + o\left(|t_n|^{2(p-1)}\right).$$

Moreover, $V^\top \in \mathcal{D}_{\varphi_1} \cap \{\varphi_1\}^{\perp, L^2}$ is the (unique) solution to

$$2\cdot\mathcal{Q}_0(V^\top,\phi) = \int_\Omega f^\dagger\,\phi\,\mathrm{d}x \qquad \text{for all } \phi\in\mathcal{D}_{\varphi_1}, \tag{13}$$

where we have denoted

$$2\cdot\mathcal{Q}_0(V^\top,\phi) = \int_\Omega \left\langle \mathbf{A}_{\varphi_1}\nabla V^\top,\, \nabla\phi\right\rangle \,\mathrm{d}x - \lambda_1\,(p-1)\int_\Omega \varphi_1^{p-2}\,V^\top\,\phi\,\mathrm{d}x$$

and $f^\dagger = f - \left(\int_\Omega f\,\varphi_1\,\mathrm{d}x\right)\varphi_1^{p-1}$.

Remark 1. The linear equation (13) represents the weak form of the "limiting" Dirichlet boundary value problem for the limit function $|t_n|^{-p}\,t_n\,v_n^\top \to V^\top$ in the approximation scheme with $u_n = t_n^{-1}(\varphi_1 + v_n^\top)$. This is a resonant problem to which a standard version of the Fredholm alternative for a selfadjoint linear operator in a Hilbert space applies. More precisely, given a function $f\in L^2(\Omega)$, a weak solution $V\in\mathcal{D}_{\varphi_1}$ to the equation

$$2\cdot\mathcal{Q}_0(V,\phi) = \textstyle\int_\Omega f\,\phi\,\mathrm{d}x \qquad \text{for all } \phi\in\mathcal{D}_{\varphi_1}, \tag{14}$$

exists in $\mathcal{D}_{\varphi_1}$ if and only if $\int_\Omega f\,\varphi_1\,\mathrm{d}x = 0$. Such a solution is always unique under the orthogonality condition $\int_\Omega V\,\varphi_1\,\mathrm{d}x = 0$.

Note that (14) written in divergent form reads as follows (see e.g. [8,12,13])

$$\begin{aligned} \operatorname{div}(\mathbf{A}_{\varphi_1}\nabla V^\top) - \lambda_1\varphi_1^{p-2}V^\top &= f \quad \text{in } \Omega\,; \\ V^\top &= 0 \quad \text{on } \partial\Omega\,; \\ \int_\Omega V^\top\varphi_1 &= 0\,. \end{aligned}$$

Remark 2. In fact we also use a variant of Theorem 2 (see [8, Cor. 4.4] for details) in order to prove the following uniform result.

Let K be a closed bounded ball in $L^\infty(\Omega)$. Assume that $f_n \equiv f$ $(n = 1, 2, \dots)$ and $t_n \to 0$ as $n \to \infty$. Then there exists a sequence $\{\eta_n\}_{n=1}^\infty \subset (0, 1)$, $\eta_n \to 0$ as $n \to \infty$, such that for all $f \in K$ and for all $n = 1, 2, \dots$ we have

$$\begin{aligned} &\Bigg| |t_n|^{-2(p-1)} \left(\mu_n - |t_n|^{p-2}\, t_n \int_\Omega f\, \varphi_1 \,\mathrm{d}x \right) - (p-2) \cdot \mathcal{Q}_0(V^\top, V^\top) \\ &\quad - (p-1) \left(\int_\Omega f\, \varphi_1 \,\mathrm{d}x \right) \left(\int_\Omega \varphi_1^{p-1}\, V^\top \,\mathrm{d}x \right) \Bigg| \le \eta_n . \end{aligned} \tag{15}$$

The main results concerning the asymptotic behavior of the solution set to (7) as $\lambda \to \lambda_1$ are sketched in Figures 2 and 3. We assume that $f^\top \in L^\infty(\Omega)$ is a given function satisfying $\int_\Omega f^\top \varphi_1 \,\mathrm{d}x = 0$ and $f^\top \not\equiv 0$. In (7) we write $f = a\varphi_1 + f^\top$, $a \in \mathbb{R}$, and split the solution as $u = c\varphi_1 + u^\top$. Note, that there are no solutions in the shaded regions (we have a priori bounds) while there may be many other solutions in the nonshaded regions.

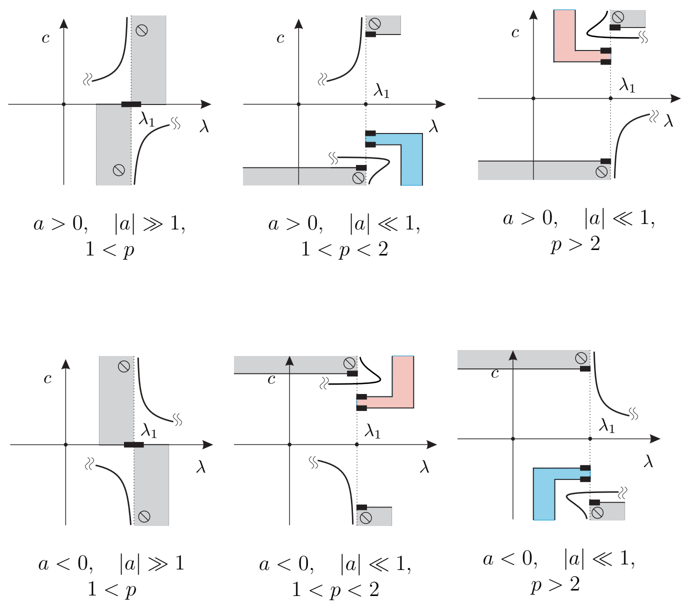


FIGURE 3. A priori bounds and bifurcations from infinity of solutions to (7) for $a \neq 0$, $1 < p < 2$ and/or $p > 2$.

We rewrite problem (7) as follows, with $f = f^\top + a\varphi_1$:

$$\begin{cases} -\Delta_p u - \lambda\, |u|^{p-2}\, u = f^\top + a\varphi_1 & \text{in } \Omega\,; \\ \qquad\qquad\qquad\qquad u = 0 & \text{on } \partial\Omega\,. \end{cases} \tag{16}$$

Here, $f^\top \in L^\infty(\Omega)$ is a given function, with $\int_\Omega f^\top \varphi_1 \,\mathrm{d}x = 0$ and $f^\top \not\equiv 0$, and $\lambda, a \in \mathbb{R}$ are real parameters. We split the solution as $u = c\varphi_1 + u^\top$. Basic multiplicity results are obtained from the shape of the continua emanating from (λ_1, ∞). Additional multiplicity results are deduced from the shape of the continua using the method of upper and lower solutions. For the convenience of the reader, we organize these results in following two tables. Dependence of the existence, multiplicity and *a priori* bounds of the solutions on the spectral parameter λ can easily be deduced from these tables.

Let us note that the theory developped in [8] can be used in the study of a more general boundary value problem

$$-\Delta_p u - \lambda\, |u|^{p-2}\, u = h(u,x) \text{ in } \Omega \quad u = 0 \text{ on } \partial\Omega\,. \tag{17}$$

Interested reader is refered to [7].

Finally, we would also like to note that the strongly nonlinear boundary value problems emphasize the importance of the interplay between numerical experiments and development of new theoretical methods, see e.g. [3].

References

[1] A. Anane, *Simplicité et isolation de la première valeur propre du p-laplacien avec poids*, Comptes Rendus Acad. Sc. Paris Série I, **305** (1987), 725–728.

[2] A. Anane, N. Tsouli, *On the second eigenvalue of the p-Laplacian*, In: “Nonlinear partial differential equations”, A. Benkirane - J-P. Gossez (Eds.), Pitman Research Notes in Mathematics Series **343**, Addison Wesley Longman, Essex, U.K., 1996, 1-9.

[3] J. Čepička, *Numerical Experiments for Nonlinear Problems*, Ph.D. Thesis, University of West Bohemia, Plzeň, 2001, p. 97 (in Czech).

[4] Drábek, P., *Solvability and Bifurcations of Nonlinear Equations*, Pitman Research Notes in Mathematics Series, Vol. 264, Longman Scientific & Technical, Essex, 1992.

[5] M. Del Pino, P. Drábek and R. F. Manásevich, *The Fredholm alternative at the first eigenvalue for the one-dimensional p-Laplacian*, J. Differential Equations, **151** (1999), 386–419.

[6] P. Drábek, P. Girg and R. F. Manásevich, *Generic Fredholm alternative for the one dimensional p-Laplacian*, Nonlin. Diff. Equations and Applications **8** (2001), 285–298.

[7] P. Drábek, P. Girg, P. Takáč, *Bounded Perturbations of Homogeneous Quasilinear Operators Using Bifurcations from Infinity,* J. Differential Equations, to appear.

[8] P. Drábek, P. Girg, P. Takáč and M. Ulm, *The Fredholm alternative for the p-Laplacian: bifurcation from infinity, existence and multiplicity,* Indiana Univ. Math. J., to appear.

[9] P. Drábek, G. Holubová, *Fredholm alternative for the p-Laplacian in higher dimensions*, J. Math. Anal. Appl. **263** (2001), 182–194.

[10] S. Fučík, J. Nečas, J. Souček and V. Souček, *Spectral Analysis of Nonlinear Operators*, Lecture Notes in Mathematics **346**, Springer-Verlag, New York-Berlin-Heidelberg, 1973.

[11] S. I. Pohozaev, *On the solvability of nonlinear equations involving odd operators,* Funkc. Anal. i Priloz. **1** (1967), pp.66–72 (in Russian).

[12] P. Takáč, *On the Fredholm alternative for the p-Laplacian at the first eigenvalue*, Indiana Univ. Math. J. **51** (2002), 187–237.

[13] P. Takáč, *On the number and structure of solutions for a Fredholm alternative with the p-Laplacian*, J. Differ. Equations **185** (2002), 306–347.

[14] P. Takáč, *A variational approach to the Fredholm alternative for the p-Laplacian near the first eigenvalue,* preprint.

$1 < p < 2,\ f \not\equiv 0$

<table>
<tr><td>$\lambda < 0$</td><td>$0 \le \lambda \le \lambda_1$</td><td>$\lambda_1$</td><td>$\lambda_1 < \lambda < \lambda_2 - \delta$</td></tr>
<tr><td colspan="4">$\|u\|_{C^{1,\beta(\overline{\Omega})}} \le M(f^\top, a, p, \lambda, \Omega) < \infty$</td></tr>
<tr><td colspan="2" rowspan="2">$M(f^\top, a, p, \lambda, \Omega) \to \infty$ as $\lambda \to \lambda_{1-}$</td><td colspan="2">$a = 0 \Rightarrow M(f^\top, a, p, \lambda, \Omega) < M < \infty \quad \forall\lambda$</td></tr>
<tr><td>$M(f^\top, a, p, \lambda_1, \Omega) < \infty \quad \forall a \in \mathbb{R}$</td><td>$a \ne 0 \Rightarrow M(f^\top, a, p, \lambda, \Omega) \to \infty$ as $\lambda \to \lambda_{1+}$</td></tr>
<tr><td rowspan="8">$\mathsf{S} = 1$</td><td>$\mathsf{S} \ge 1$</td><td></td><td>$\mathsf{S} \ge 1$</td></tr>
<tr><td>$f \prec 0$ $\quad \mathsf{S}, \mathsf{N} = 1, \mathsf{P} = 0$ (LMP)
$a < \underline{A}$ (LMP)
$\exists\eta > 0\,\forall\lambda \in (\lambda_1 - \eta, \lambda_1)$: $\quad \mathsf{N} \ge 1, \mathsf{S} - \mathsf{N} = 0$</td><td>$f \prec 0$ or
$a < \underline{A}$ $\quad \mathsf{S} = 0$</td><td>$f \prec 0$ or
$a < \underline{A}$ (LAMP)
$\exists\eta > 0\,\forall\lambda \in (\lambda_1, \lambda_1 + \eta)$: $\quad \mathsf{P} \ge 1, \mathsf{S} - \mathsf{P} = 0$</td></tr>
<tr><td>$\underline{A} < a < \underline{a}$ $\quad \mathsf{N} \ge 1$ PAB</td><td>$\underline{A} < a < \underline{a}$ $\quad \mathsf{S} \ge 2$ (UpLow)</td><td>$\underline{A} < a < \underline{a}$ $\quad \mathsf{P} \ge 1$, NAB</td></tr>
<tr><td>$\forall\varepsilon' \in (0, -\underline{a})\exists\eta = \eta(f^\top, \underline{a}, \varepsilon') \quad \forall\lambda$:
$\lambda_1 - \eta < \lambda < \lambda_1$
$\underline{a} < a < -\varepsilon'$ $\quad \mathsf{S} \ge 3, \mathsf{P}, \mathsf{N} \ge 1$</td><td>$\underline{a} < a < 0$ $\quad \mathsf{S} \ge 2, \mathsf{P} \ge 1$
$M(f^\top, a, p, \Omega) \to \infty$ as $a \to 0_-$</td><td>$\forall\varepsilon' \in (0, -\underline{a})\exists\eta = \eta(f^\top, \underline{a}, \varepsilon') \quad \forall\lambda$:
$\lambda_1 < \lambda < \lambda_1 + \eta$
$\underline{a} < a < -\varepsilon'$ $\quad \mathsf{S} \ge 2, \mathsf{P} \ge 2$</td></tr>
<tr><td>$\lambda_1 - \eta < \lambda < \lambda_1$
$a = 0$ $\quad \mathsf{S} \ge 3, \mathsf{P}, \mathsf{N} \ge 1$</td><td>$a = 0$ $\quad \mathsf{S} \ge 1$
$\|u\|_{C^{1,\beta(\overline{\Omega})}} \le M(f^\top, 0, p, \Omega) < \infty$</td><td>$\lambda_1 < \lambda < \lambda_1 + \eta$
$a = 0$ $\quad \mathsf{S} \ge 1$</td></tr>
<tr><td>$\forall\varepsilon' \in (0, \overline{a})\exists\eta = \eta(f^\top, \overline{a}, \varepsilon') \quad \forall\lambda$:
$\lambda_1 - \eta < \lambda < \lambda_1$
$\varepsilon' < a < \overline{a}$ $\quad \mathsf{S} \ge 3, \mathsf{P}, \mathsf{N} \ge 1$</td><td>$0 < a < \overline{a}$ $\quad \mathsf{S} \ge 2, \mathsf{N} \ge 1$
$M(f^\top, a, p, \Omega) \to \infty$ as $a \to 0_+$</td><td>$\forall\varepsilon' \in (0, \overline{a})\exists\eta = \eta(f^\top, \overline{a}, \varepsilon') \quad \forall\lambda$:
$\lambda_1 < \lambda < \lambda_1 + \eta$
$\varepsilon' < a < \overline{a}$ $\quad \mathsf{S} \ge 2, \mathsf{N} \ge 2$</td></tr>
<tr><td>$\overline{a} < a < \overline{A}$ $\quad \mathsf{P} \ge 1$, NAB</td><td>$\overline{a} < a < \overline{A}$ $\quad \mathsf{S} \ge 2$ (UpLow)</td><td>$\overline{a} < a < \overline{A}$ $\quad \mathsf{N} \ge 1$, PAB</td></tr>
<tr><td>$\overline{A} < a$ (LMP)
$\exists\eta > 0\,\forall\lambda \in (\lambda_1 - \eta, \lambda_1)$: $\quad \mathsf{P} \ge 1, \mathsf{S} - \mathsf{P} = 0$
$f \succ 0$ $\quad \mathsf{S}, \mathsf{P} = 1, \mathsf{N} = 0$ (LMP)</td><td>$\overline{A} < a$ or
$f \succ 0$ $\quad \mathsf{S} = 0$</td><td>$\overline{A} < a$ or
$f \prec 0$ (LAMP)
$\exists\eta > 0\,\forall\lambda \in (\lambda_1, \lambda_1 + \eta)$: $\quad \mathsf{N} \ge 1, \mathsf{S} - \mathsf{N} = 0$</td></tr>
</table>

Legend: S (N, P) number of (negative, positive) solution; $\mathsf{S} \ge n$ at least n solutions; $\mathsf{N} \ge 1$ at least one negative solution; $\mathsf{S} - \mathsf{N} = 0$ all solutions are negative; NAB (PAB) negative (positive) solutions are *a priori* bounded; *(LMP)* Local Maximum Principle, *(LAMP)* Local Anti-Maximum Principle, *(UpLow)* by upper and lower solutions argument.

$p > 2,\ f \not\equiv 0$

$\lambda < 0$	$0 \le \lambda \le \lambda_1$	λ_1	$\lambda_1 < \lambda < \lambda_2 - \delta$
	$\Vert u\Vert_{C^{1,\beta}(\overline{\Omega})} \le M(f^\top, a, p, \lambda, \Omega) < \infty$		
	$a = 0 \Rightarrow M(f^\top, a, p, \lambda, \Omega) < M < \infty \quad \forall\lambda$		$M(f^\top, a, p, \lambda, \Omega) \to \infty$ as $\lambda \to \lambda_{1+}$
	$a \ne 0 \Rightarrow M(f^\top, a, p, \lambda, \Omega) \to \infty$ as $\lambda \to \lambda_{1-}$	$M(f^\top, a, p, \lambda_1, \Omega) < \infty \quad \forall a \in \mathbb{R}$	
	$\mathsf{S} \ge 1$		$\mathsf{S} \ge 1$
	$f \prec 0$ $\mathsf{S}, \mathsf{N} = 1, \mathsf{P} = 0$ *(LMP)* $a < \underline{A}$ *(LMP)* $\exists\eta > 0\,\forall\lambda \in (\lambda_1 - \eta, \lambda_1)$: $\mathsf{N} \ge 1, \mathsf{S} - \mathsf{N} = 0$	$f \prec 0$ or $a < \underline{A}$ $\mathsf{S} = 0$	$f \prec 0$ or $a < \underline{A}$ *(LAMP)* $\exists\eta > 0\,\forall\lambda \in (\lambda_1, \lambda_1 + \eta)$: $\mathsf{P} \ge 1, \mathsf{S} - \mathsf{P} = 0$
	$\underline{A} < a < \underline{a}$ $\mathsf{N} \ge 1$ PAB	$\underline{A} < a < \underline{a}$ $\mathsf{S} \ge 2$ *(UpLow)*	$\underline{A} < a < \underline{a}$ $\mathsf{P} \ge 1$, NAB
	$\forall\varepsilon' \in (0, -\underline{a})\exists\eta = \eta(f^\top, \underline{a}, \varepsilon')\ \ \forall\lambda$: $\lambda_1 - \eta < \lambda < \lambda_1$ $\underline{a} < a < -\varepsilon'$ $\mathsf{S} \ge 2, \mathsf{N} \ge 2$	$\underline{a} < a < 0$ $\mathsf{S} \ge 2$, $\mathsf{P} \ge 1$ $M(f^\top, a, p, \Omega) \to \infty$ as $a \to 0_-$	$\forall\varepsilon' \in (0, -\underline{a})\exists\eta = \eta(f^\top, \underline{a}, \varepsilon')\ \ \forall\lambda$: $\lambda_1 < \lambda < \lambda_1 + \eta$ $\mathsf{S} \ge 3, \mathsf{P}, \mathsf{N} \ge 1$
$\mathsf{S} = 1$	$\lambda_1 - \eta < \lambda < \lambda_1$ $a = 0$ $\mathsf{S} \ge 1$	$a = 0$ $\mathsf{S} \ge 1$ $\Vert u\Vert_{C^{1,\beta}(\overline{\Omega})} \le M(f^\top, 0, p, \Omega) < \infty$	$\lambda_1 < \lambda < \lambda_1 + \eta$ $a = 0$ $\mathsf{S} \ge 3, \mathsf{P}, \mathsf{N} \ge 1$
	$\forall\varepsilon' \in (0, \overline{a})\exists\eta = \eta(f^\top, \overline{a}, \varepsilon')\ \ \forall\lambda$: $\lambda_1 - \eta < \lambda < \lambda_1$ $\varepsilon' < a < \overline{a}$ $\mathsf{S} \ge 2$, $\mathsf{P} \ge 2$	$0 < a < \overline{a}$ $\mathsf{S} \ge 2$, $\mathsf{N} \ge 1$ $M(f^\top, a, p, \Omega) \to \infty$ as $a \to 0_+$	$\forall\varepsilon' \in (0, \overline{a})\exists\eta = \eta(f^\top, \overline{a}, \varepsilon')\ \ \forall\lambda$: $\lambda_1 < \lambda < \lambda_1 + \eta$ $\varepsilon' < a < \overline{a}$ $\mathsf{S} \ge 3, \mathsf{P}, \mathsf{N} \ge 1$
	$\overline{a} < a < \overline{A}$ $\mathsf{P} \ge 1$, NAB	$\overline{a} < a < \overline{A}$ $\mathsf{S} \ge 2$ *(UpLow)*	$\overline{a} < a < \overline{A}$ $\mathsf{N} \ge 1$, PAB
	$\overline{A} < a$ *(LMP)* $\exists\eta > 0\,\forall\lambda \in (\lambda_1 - \eta, \lambda_1)$: $\mathsf{P} \ge 1, \mathsf{S} - \mathsf{P} = 0$ $f \succ 0$ $\mathsf{S}, \mathsf{P} = 1, \mathsf{N} = 0$ *(LMP)*	$\overline{A} < a$ or $f \succ 0$ $\mathsf{S} = 0$	$\overline{A} < a$ or $f \prec 0$ *(LAMP)* $\exists\eta > 0\,\forall\lambda \in (\lambda_1, \lambda_1 + \eta)$: $\mathsf{N} \ge 1, \mathsf{S} - \mathsf{N} = 0$

Legend: S (N, P) number of (negative, positive) solution; $\mathsf{S} \ge n$ at least n solutions; $\mathsf{N} \ge 1$ at least one negative solution; $\mathsf{S} - \mathsf{N} = 0$ all solutions are negative; NAB (PAB) negative (positive) solutions are *a priori* bounded; *(LMP)* Local Maximum Principle, *(LAMP)* Local Anti-Maximum Principle, *(UpLow)* by upper and lower solutions argument.

Perturbation of the simple eigenvalue by 1-homogeneous operators
Raffaele Chiappinelli (Siena, Italy)

Let T be a bounded linear operator acting in a real Banach space E and suppose that T has an isolated eigenvalue of finite multiplicity λ_0. If we add to T a perturbation term εB, with B (positively) homogeneous of degree 1, continuous and such that $B(0) = 0$, then we ask the following questions:

1) Do we find eigenvalues of $T + \varepsilon B$ near λ_0?

2) If this is the case, are these eigenvalues isolated themselves?

(An *eigenvalue* of an operator $F : E \to E$ such that $F(0) = 0$ is a $\lambda \in \mathbb{R}$ such that $F(u_0) = \lambda u_0$ for some *eigenvector* $u_0 \neq 0$; in this case we say that

$$N(F - \lambda I) \equiv \{u \in E : F(u) - \lambda u = 0\}$$

is the *eigenset* corresponding to λ. If F is (1-)homogeneous, this notion of eigenvalue coincides with that of *connected eigenvalue* proposed in [4]).

Simple examples in finite dimension show the answer to both questions may be "No". In particular, as for question 2) one may consider the equation

$$x + \varepsilon\phi\Big(\frac{x}{\|x\|}\Big)x = \lambda x, \qquad x \in \mathbb{R}^N \tag{18}$$

where $\phi : S \equiv \{x \in \mathbb{R}^N : \|x\| = 1\} \to \mathbb{R}$ is continuous. Then $Tx \equiv x$, $B(x) \equiv \phi(\frac{x}{\|x\|})x$ for $x \neq 0$, $B(0) = 0$ satisfy the above assumptions. However, each $x \in S$ is an eigenvector of (18) corresponding to the eigenvalue $\lambda = \lambda(x) = 1 + \varepsilon\phi(x)$; thus if $N > 1$, then - as S is connected in this case - $\{1 + \varepsilon\phi(x) : x \in S\}$ is an *interval* of eigenvalues (except when ϕ is constant on S) which for ε small is close as we wish to the "unperturbed" eigenvalue 1 of T.

On the other hand, it is possible to provide an affirmative answer when λ_0 is (algebraically) *simple* and B is Lipschitz continuous: indeed, in this case we essentially prove that λ_0 splits (for each $\varepsilon \neq 0$) into precisely two eigenvalues $\lambda_\pm(\varepsilon)$, while the eigenline $N(T-\lambda_0 I)$ correspondingly "breaks" into two eigenrays $N_\pm(\varepsilon)$. For the Hilbert space case, the precise statement is as follows:

Theorem 3. *Let T be a bounded linear operator in H (a real Hilbert space with scalar product denoted by $\langle , \rangle$), and let $B : H \to H$ be such that $B(0) = 0$. Suppose that:*

(i) *T is selfadjoint and λ_0 is an isolated and simple eigenvalue of T;*
(ii) *B is Lipschitz continuous of constant k;*
(iii) *B is homogeneous.*

Then there exist $\delta_0 > 0$, $\varepsilon_0 > 0$ (depending only on λ_0 and k) such that for every ε with $|\varepsilon| \leq \varepsilon_0$, $T + \varepsilon B$ has precisely two *(possibly coinciding) eigenvalues $\lambda_+(\varepsilon), \lambda_-(\varepsilon)$ in the interval $[\lambda_0 - \delta_0, \lambda_0 + \delta_0]$: that is, for $|\lambda - \lambda_0| \leq \delta_0$ nontrivial solutions of the equation*

$$Tu + \varepsilon B(u) = \lambda u \tag{19}$$

exist if and only if $\lambda = \lambda_\pm(\varepsilon)$. *Moreover, the eigensets* $N_\pm(\varepsilon) \equiv N(T + \varepsilon B - \lambda_\pm(\varepsilon)I)$ *corresponding to* $\lambda_\pm(\varepsilon)$ *are rays in* H*, that is, there exist nonzero vectors* $z_\pm(\varepsilon) \in H$ *such that*

$$N_\pm(\varepsilon) = \{tz_\pm(\varepsilon) : t \geq 0\}.$$

Finally $\lambda_\pm(\varepsilon)$ *and* $z_\pm(\varepsilon)$ *depend Lipschitz-continuously upon* ε *for* $|\varepsilon| \leq \varepsilon_0$*, and if* ϕ *is a normalized eigenvector of* T *corresponding to* λ_0*, then as* $\varepsilon \to 0$ $z_\pm(\varepsilon) \to \pm\phi$ *and*

$$\lambda_\pm(\varepsilon) = \lambda_0 + \varepsilon\langle B(\pm\phi), \pm\phi\rangle + o(\varepsilon).$$

Theorem 3 is proved in [1] by first using the Lyapounov-Schmidt reduction for (19), and then making full use of the homogeneity of B in the resulting bifurcation equation. In a sense, this generalizes a result of Ruf [3] concerning the existence and uniqueness of two eigenvalues $\mu_k^1, \mu_k^2 \in [\mu_k^0, \mu_{k+1}^0[$ for the problem (in a bounded open set $\Omega \subset \mathbb{R}^N$)

$$Lu = \gamma u^- + \mu u \quad \text{in} \quad \Omega, \qquad u = 0 \quad \text{on} \quad \partial\Omega \tag{20}$$

with L linear elliptic selfadjoint and $u^- = \max(-u, 0)$, near each simple eigenvalue μ_k^0 of L. In fact, similar results hold (see [1]) for the problem

$$Lu = \mu(u + \varepsilon(\alpha(x)u^+ - \beta(x)u^-)) \quad \text{in} \quad \Omega, \qquad u = 0 \quad \text{on} \quad \partial\Omega \tag{21}$$

with $u = u^+ - u^-$ and $\alpha, \beta \in L^\infty(\Omega)$. Moreover, in the special case $\alpha, \beta = const$ we obtain informations about the structure of the "Fučik spectrum" Σ of L near (μ_k^0, μ_k^0), which agree with classical results [2].

Open problem: Describe what happens when λ_0 is not simple. Also, single out a class of homogeneous mappings in $\mathbb{R}^N$ all of whose eigenvalues are isolated (as for linear mappings).

References

[1] R. Chiappinelli, *Isolated connected eigenvalues in nonlinear spectral theory*, Nonlinear Funct. Anal. Appl. **8** (2003), 557–579.

[2] T. Gallouet and O. Kavian, *Resultats d'existence et de non-existence pour certains problemes demi-lineaires a l'infini*, Ann. Fac. Sci. Toulouse, V Ser., Math., (1981), 201–246.

[3] B. Ruf, *On nonlinear elliptic problems with jumping nonlinearities*, Ann. Mat. Pura Appl. (IV) **128** (1981), 133–151.

[4] P. Santucci and M. Väth, *On the definition of eigenvalues for nonlinear operators*, Nonlinear Anal. TMA **40** (2000), 565–576.

Remarks on some inhomogeneus eigenvalue problems

Vesa Mustonen (Oulu)

We discuss the "principal" eigenvalues of the problem

$$\begin{cases} -\Delta_m(u) = \lambda m(u) & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{22}$$

where $m : [0,\infty) \mapsto [0,\infty)$ is nondecreasing continuous function with $m(0) = 0$, $m(t) > 0$ and $t > 0$, $\lim\limits_{t\to\infty} = \infty$, $m(-t) = -m(t)$ $\forall t \in \mathbb{R}^n$, $\Omega \subset \mathbb{R}^n$ bounded open subset and

$$\Delta_m(u) := \operatorname{div} \left(\frac{m(|\nabla u|)}{|\nabla u|} \nabla u \right) \qquad \text{(generalized Laplacian).}$$

It is known ([2], [1]) that for each $r > 0$ the solutions $u_r \in W_0^1 L_M(\Omega)$ of the minimization problem

$$\inf \left\{ \int_\Omega M(|\nabla u|) : u \in W_0^1 L_M(\Omega), \int_\Omega M(u) = r \right\}$$

are solutions for (22) with some $\lambda = \lambda_r > 0$. (Here $M(t) = \int_0^t m(s)ds$). Some examples for the ODE-case suggest that the set of "principal" eigenvalues $\lambda > 0$ obtained is not necessarily bounded from above or bounded from below away from zero ([3]) Therefore we suggest to modify the problem (22) to the form

$$\begin{cases} -\Delta_m(u) = \lambda m(\lambda u) & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{23}$$

For the ODE-case

$$\begin{cases} -(m(u'))' = \lambda m(\lambda u) & \text{in } (0,a) \\ u(0) = u(a) = 0 \end{cases} \tag{24}$$

one can use the time map which suggests that all principal eigenvalues for (24) are in the bounded interval $[2/a, 4/a]$. This is joint work with Matti Tienari, University of Oulu /Central Laboratory, Helsinki.

References

[1] J.- P. Gossez and R. Manásevich, *On nonlinear eigenvalueproblem in Orlicz-Sobolev spaces*, Proc. R. Soc. Edinb. **132A** (2002), 891–909.

[2] V. Mustonen and M. Tienari, *An eigenvalue problem for generalized Laplacian in Orlicz-Sobolev spaces*, Proc. R. Soc. Edinb. **129A** (1999), 153–163.

[3] V. Mustonen and M. Tienari, *Remarks on inhomogeneus elliptic eigenvalue problems*, Partial differential equations, 259–265, Lecture Notes in Pure and Appl. Math. **229**, Dekker, New York, 2002.

Applications of the degree for Fredholm maps to elliptic problems
C. A. Stuart (Lausanne)

A topological degree for C^1-Fredholm maps of index zero that are proper on closed bounded sets, has been defined by several approaches in a way that makes it possible to track the change in the degree under homotopy. See the work by Fitzpatrick, Pejsachowicz and Rabier [3,4,7] and then by Benevieri and Furi [1,2]. For the case of a map $F : X \to Y$ acting between two real Banach spaces X and Y, the following properties of the F are required.

(1) $F \in C^1(X, Y)$
(2) $F(u) \in B(X, Y)$ is a Fredholm operator of index zero for all $u \in X$
(3) $F : W \to Y$ is proper for all closed bounded subsets W of X.

In a series of papers written in collaboration with H. Jeanjean, M. Lucia, P. J. Rabier, S. Secchi and H. Gebran, we have used this degree to obtain results about the existence and bifurcation of solutions of systems of differential equations in several situations where the Leray-Schauder degree is not directly applicable. My lecture started with a summary of this work and then I presented in more detail the treatment of quasilinear systems that are elliptic in the sense of Petrovskii. I illustrated one of the keys steps in the case of a simple but typical example of a second order quasilinear elliptic equation.

References

[1] Benevieri, P. and Furi, M., *A simple notion of orientability for Fredholm maps of index zero between Banach manifolds and degree theory*, Ann. Sci. math. Québec, **22** (1998), 131–148.

[2] Benevieri, P. and Furi, M., *Bifurcation results for families of Fredholm maps of index zero between Banach spaces*, Nonlinear Analysis Forum, **6** (2001), 35–47.

[3] Fitzpatrick, P. M., Pejsachowicz, J. and Rabier, P. J., *The degree of proper C^2 Fredholm mappings, I*, J. reine angew. Math. **427** (1992) 1–33.

[4] Fitzpatrick, P. M., Pejsachowicz, J., Rabier, P. J., *Orientability of Fredholm families and topological degree for orientable nonlinear Fredholm mappings*, J. Funct. Anal. **124** (1994), 1–39.

[5] Gebran, H. and Stuart, C. A., *Fredholm and properness properties of quasilinear elliptic systems of second order*, submitted.

[6] Jeanjean, H., Lucia,M., Stuart, C. A., *Branches of solutions to semilinear elliptic equations*, Math. Z. **230** (1999), 79–105.

[7] Pejsachowicz, J., and Rabier, P. J., *Degree theory for C^1 Fredholm mappings of index* 0, J. Anal. Math. **76** (1997).

[8] Rabier, P. J., and Stuart, C. A., *Fredholm and properness properties of quasilinear elliptic operators on $\mathbb{R}^N$*, Math. Nachr. **231** (2001).

[9] Rabier, P. J., and Stuart, C. A., *Global bifurcation for quasilinear elliptic equations on $\mathbb{R}^N$*, Math. Z. **237** (2001).

[10] Rabier, P. J. and Stuart, C. A., *A Sobolev space approach to boundary value problems on the half-line*, Comm. Contemp. Math. (to appear).

[11] Rabier, P.J. and Stuart, C.A., *Boundary value problems for first order systems on the half-line*, submitted.

[12] Secchi, S. and Stuart, C. A., *Global bifurcation of homoclinic solutions of Hamiltonian systems*, Discrete Cont. Dynam. Syst., **9** (2003), 1493–1518.

On the sign-jump of one-parameter families of Fredholm operators and bifurcation

Massimo Furi (Florence, Italy)

In [1] (see also [2] for more details) a fairly simple notion of orientation for Fredholm linear operators of index zero between real vector spaces was introduced. Any such operator, invertible or noninvertible, admits exactly two orientations, and the choice of an orientation makes, by definition, the operator oriented. However, if an operator is invertible, one of the two orientations is more relevant than the other, and for this reason called *natural*. Thus it makes sense to assign to any oriented isomorphism a sign: 1 if the orientation is natural and -1 in the opposite case. For a singular Fredholm operator of index zero no one of the two orientations is more relevant than the other.

A crucial fact is that in the framework of Banach spaces the orientation has a sort of stability; in the sense that an orientation of an operator L induces, in a very natural way, an orientation to any operator which is sufficiently close to L. Using this fact, the notion of orientation was extended (in [1]) to the nonlinear case; namely, to the case of a C^1 Fredholm map of index zero between real Banach spaces (and Banach manifolds). Such an extension coincides (in the C^1 case) with the notion given by Dold in [4, exercise 6, p. 271] for maps between finite dimensional manifolds and, in the most important cases, with the notion due to Fitzpatrick, Pejsachowicz and Rabier for maps between Banach manifolds (see [5–9]).

In [1], by means of the concept of orientation, a degree theory for Fredholm maps between Banach manifolds was introduced. This theory is purely based on the Brouwer degree, and in the most important cases agrees with the theory developed by Fitzpatrick, Pejsachowicz and Rabier in a series of papers ranging from 1991 to 1998. The difference between the two theories is mainly in the construction method and in a different definition of orientation.

This talk is inspired by a recent joint work with Benevieri, Pera and Spadini (see [3]), and it concerns methods for computing the degree by counting the sign-jumps in a continuous curve of Fredholm operators of index zero.

Some consequences in global bifurcation theory are derived from the detection of a sign-jump.

References

[1] P. Benevieri, M. Furi, *A simple notion of orientability for Fredholm maps of index zero between Banach manifolds and degree theory*, Ann. Sci. Math. Québec **22** (1998), 131–148.

[2] P. Benevieri, M. Furi, *On the concept of orientability for Fredholm maps between real Banach manifolds*, Topol. Methods Nonlinear Anal. **16** (2000)2, 279–306.

[3] P. Benevieri, M. Furi, M. P. Pera, M. Spadini, *About the sign of oriented Fredholm operators between Banach spaces*, Z. Anal. Anwendungen **22** (2003)3, 619–645.

[4] A. Dold, *Lectures on algebraic topology*, Springer-Verlag, Berlin, 1973.

[5] P. M. Fitzpatrick, J. Pejsachowicz, *Parity and generalized multiplicity*, Trans. Amer. Math. Soc. **326** (1991)1, 281–305.

[6] P. M. Fitzpatrick, J. Pejsachowicz, *Orientation and the Leray-Schauder theory for fully nonlinear elliptic boundary value problems*, Mem. Amer. Math. Soc. **483** (1993).

[7] P. M. Fitzpatrick, J. Pejsachowicz, P. J. Rabier, *The degree of proper C^2 Fredholm mappings*, J. Reine Angew. Math. **427** (1992), 1–33.
[8] P. M. Fitzpatrick, J. Pejsachowicz, P.J. Rabier, *Orientability of Fredholm Families and Topological Degree for Orientable Nonlinear Fredholm Mappings*, J. of Funct. Anal. **124** (1994)1, 1–39.
[9] J. Pejsachowicz, P. Rabier, *Degree theory for C^1 Fredholm mappings of index* 0, J. Anal. Math. **76** (1998), 289–319.

Applications of nonlinear and semilinear spectral theory to boundary value problems

Wenying Feng (Peterborough, Canada)

We study the nonlinear spectrum $\sigma(f)$ and semilinear spectrum $\sigma(L, N)$, when L is Fredholm of index zero, f and N are asymptotically linear or positively homogeneous, thus close to a linear operator. The results generalize a previous result which required N to be a linear operator and L to be the identity map. To prove a theorem on the spectrum of asymptotically linear operator, we introduce the field of regularity for semilinear operators. When N is a positively homogeneous operator, we give a condition that ensures the existence of a positive eigenvalue for the semilinear pair (L, N).

The theorems can be applied to the study of some integral equations involving Urysohn and Hammerstein operators. Results on existence of solutions, bifurcation points, asymptotic bifurcation points are obtained. We also apply our theorems to the study of the second order differential equation:

$$u'' + f(t, u) = 0 \tag{25}$$

with one of the boundary conditions ($0 < \eta < 1$ fixed):

$$x(0) = 0,\ x(1) = \alpha x(\eta), \tag{26}$$

$$x'(0) = 0,\ x(1) = \alpha x(\eta). \tag{27}$$

By making use of a upper bound that involves the parameters α, η, we prove results on the existence of a solution, which in some cases are better than previous results (required a constant upper bound of f) of Gupta, Ntouyas and Tsamatos. Some examples show that there are equations that can be treated by our theorems but the previous results can not be applied. Moreover, with the assumption that f is positively homogeneous, we study the existence of an eigenvalue for the more general equation

$$u'' = f(t, u, u'),\ t \in (0, 1) \tag{28}$$

with one of the boundary condition (26) and (27). We give an alternative condition for existence of a positive eigenvalue and being a surjective map. Two examples are constructed to show that there are equations that satisfy our condition and so existence of an eigenvalue can be proved.

References

[1] J. Appell, E. De Pascale and A. Vignoli, *A comparison of different spectra for nonlinear operators*, Nonlinear Anal. TMA. **40** (2000), 73–90.
[2] M. Furi, M. Martelli, A. Vignoli, *Contributions to the spectral theory for nonlinear operators in Banach spaces*, Ann. Mat. Pura Appl. **118** (1978), 229–294.
[3] W. Feng, *Nonlinear spectral theory and operator equations*, Nonlinear Funct. Anal. & Appl., Vol. 8, No. 4 (2003), pp. 519-533.
[4] W. Feng, *A new spectral theory for nonlinear operators and its applications*, Abstr. Appl. Anal. **2** (1997), 163–183.
[5] W. Feng, *Nonlinear and semilinear spectrum for asymptotically linear or positively homogeneous operators*, to appear in Nonlinear Anal. TMA.
[6] W. Feng and J. R. L. Webb, *A spectral theory for semilinear operators and its applications*, Recent trends in nonlinear analysis, 149–163, Progr. Nonlinear Differential Equations Appl., **40**, Birkhäuser, Basel (2000).
[7] E. Giorgieri, J. Appell and M. Vath, *Nonlinear spectral theory for homogeneous operators*, Nonlinear Funct. Anal. Appl. **7** no. 4 (2002), 589–618.
[8] CH. P. Gupta, S. K. Ntouyas, and P. CH. Tsamatos, *Existence results for multi-point boundary value problems for second order ordinary differential equations*, Bull. Greek Math. Soc. **43** (2000), 105–123.
[9] G. Infante and J. R. L. Webb, *Three point boundary value problems with solutions that change sign*, J. Integral Equ. Appl. (2003).

Epi and Coepi Maps, and Further?

Martin Väth

This is a survey talk on the current state and possible developments of topological methods for coincidence points of function pairs which is one of the crucial ingredients of nonlinear spectral theory.

On the one hand, the concept of 0-epi maps (see e.g. [3]) can be considered as a definition of a homotopically stable coincidence point of two functions. On the other hand, there exist various degree theories for coincidence points which might be considered as a corresponding homologic approach: Degree theories for coincidence points of compact maps with linear Fredholm maps of zero or positive index [8, 9], with nonlinear Fredholm maps of index zero [2], or with monotone maps [11]. The link between these two kind of approaches (0-epi maps and degree theory) can be established by the famous Hopf theorems [5].

A third approach to coincidence points is given by various fixed point indices of multivalued maps (each of these indices is based on one of four essentiallz different ideas [1, 6, 7, 10] which are briefly sketched in the talk). This index approach is somewhat dual to the above coincidence degree theories and might be considered as an application of cohomology theory. It is possible to give a corresponding cohomotopic definition of a "coepi" concept which relates to these index theories by Hopf type theorems [13]. So, roughly, one has the following picture:

homotopic	homologic
Epi Maps Homotopy	*Degree* Homology
Coepi Maps Cohomotopy	*Index* Cohomology

All these concepts and Hopf theorems generalize also to noncompact but only condensing functions pairs. Moreover, it seems now that there is a natural notion of a degree for function triples which covers and extends all the above theories in a unified manner [4, 12].

As this is a survey talk, it would be too long to give a complete list of references in this abstract: For each referred subject only the historically first paper dealing with that topic is cited here.

References

[1] Bader, R. and Kryszewski, W., *Fixed-point index for compositions of set-valued maps with proximaly ∞-connected values on arbitrary ANR's*, Set-Valued Anal. **2** (1994), 459–480.

[2] Beneveri, P. and Furi, M., *Degree for locally compact perturbations of Fredholm maps in Banach spaces*, (submitted).

[3] Furi, M., Martelli, M., and Vignoli, A., *On the solvability of nonlinear operator equations in normed spaces*, Ann. Mat. Pura Appl. **124** (1980), 321–343.

[4] Gabor, D. and Kryszewski, W., *A coincidence theory involving Fredholm operators of nonnegative index*, Topol. Methods Nonlinear Anal. **15** (2000), no. 1, 43–59.

[5] Giorgieri, E. and Väth, M., *A characterization of* 0*-epi maps with a degree*, J. Funct. Anal. **187** (2001), 183–199.

[6] Kryszewski, W., *The fixed-point index for the class of compositions of acyclic set-valued maps on ANR's*, Bull. Soc. Math. France **120** (1996), 129–151.

[7] Kucharski, Z., *A coincidence index*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **24** (1976), 245–252.

[8] Mawhin, J. L., *Equivalence theorems for nonlinear operator equations and coincidence degree theory for some mappings in locally convex topological vector spaces*, J. Differential Equations **12** (1972), 610–636.

[9] Nirenberg, L., *Generalized degree and nonlinear problems*, 3ieme Coll. sur l'Analyse fonction., Liege 1970 (Louvain, Belgique), Centre Belge de Recherches Matheématiques. Vander éditeur, 1971, 1–9.

[10] Siegberg, H. W. and Skordev, G., *Fixed point index and chain approximations*, Pacific J. Math. **102** (1982), 455–486.

[11] Skrypnik, I. V., *Nonlinear elliptic boundary value problems*, Teubner, Leipzig, 1986.

[12] Väth, M., *Merging of degree and index theory*, (in preparation).

[13] Väth, M., *Coepi maps and generalizations of the Hopf extension theorem*, Topology Appl. **131** (2003), 79–99.

Spectral theory for homogeneous operators: part I

Elena Giorgieri (Rome)

The aim of this talk is to present a part of a joint work with J. Appell and M. Väeth, contained in the paper *Nonlinear spectral theory for homogeneous operators* [2], where, starting from the work [4], [3], [5], [6], and [1], we develop a parallel theory of spectra and phantoms which better describes properties of homogeneous operators of general degree. The above papers, with the exception of [5] and [6], deal with an operator F acting on a Banach space E and define the spectra by using some metric and topological characteristics and some notion of *solvability* of the equation

$$(\lambda I - F)(u) = G(u),$$

where $G : X \to Y$ varies in a suitable subset of the space of continuous operators. Moreover, the spectra introduced in [4] and in [1] depend essentially on the asymptotic properties of the operators involved and do not contain the eigenvalues (in the classical sense), while the spectrum in [3] is an example of "global" spectrum, because it is meant to contain all the eigenvalues. Regarding the papers [5] and [6], they deal with operators acting between two different Banach spaces and the spectra they define, called phantoms, describe the "local" behaviour of the operator. One of the main features of their work is the introduction of a new notion of eigenvalue for a pair of operators (F, J), where J replaces the identity (F acts between two different Banach spaces).

The basic idea of our work [2] is to modify the definitions given in the above papers in a way that takes into account the special behaviour of homogeneous operators. Indeed, we deal with *continuous* operators $F, J : X \to Y$ acting between two different Banach spaces X,Y (over the same field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) and satisfying $F(\theta) = J(\theta) = \theta$. Here J is some "well-behaved" operator that replaces the role of the identity, for example a homeomorphism, while F denotes the operator we want to analyse. The modified metric and topological characteristics we use are then the following.

Metric characteristics

$$M_\tau(F) = \sup_{u\neq\theta} \frac{\|F(u)\|}{\|u\|^\tau}, \qquad m_\tau(f) = \inf_{u\neq\theta} \frac{\|F(u)\|}{\|u\|^\tau},$$
$$|F|_\tau = \limsup_{\|u\|\to\infty} \frac{\|F(u)\|}{\|u\|^\tau}, \qquad d_\tau(F) = \liminf_{\|u\|\to\infty} \frac{\|F(u)\|}{\|u\|^\tau}, \qquad \tau > 0$$

Topological characteristics

$$\alpha_\tau(F) = \inf \left\{ \begin{array}{l} L \geq 0 : \alpha(F(M)) \leq L\alpha(M)^\tau \\ \text{for all bounded } M \subset X \end{array} \right\},$$
$$\beta_\tau(F) = \sup \left\{ \begin{array}{l} \ell \geq 0 : \alpha(F(M)) \geq \ell\alpha(M)^\tau \\ \text{for all bounded } M \subset X \end{array} \right\}, \qquad \tau > 0.$$

($\alpha(M)$ is the usual Kuratowski measure of noncompactness of the bounded subset M).

By adapting the definitions in [4], [3], [5], [6] and [1] to these new characteristics, we obtain spectra that maintain all the topological properties of the related ones (included compactness under some additional conditions), and this is precisely what I am going to present today. In the case when F and J are τ-homogeneous operators, these modified spectra say more on the properties of F then the previous spectra, as it will be shown in the second part of this talk by J. Appell.

References

[1] J. Appell, E. G., M. Väth, *On a class of maps related to the Furi-Martelli-Vignoli spectrum*, Annali Mat. Pura Appl. **179** (2001), 215–228.
[2] J. Appell, E. G., M. Väth, *Nonlinear spectral theory for homogeneous operators*, Nonlinear Funct. Anal. Appl., **4** (2002), 589–618.
[3] W. Feng, *A new spectral theory for nonlinear operators and its applications*, Abstr. Appl. Anal. **2** (1997), 163–183.
[4] M. Furi, M. Martelli, A. Vignoli, *Contributions to the spectral theory for nonlinear operators in Banach spaces*, Annali Mat. Pura Appl. **118** (1978), 229–294.
[5] P. Santucci, M. Väth, *Grasping the phantom: a new approach to nonlinear spectral theory*, Annali Mat. Pura Appl., **180** (2001), 3, 255–284.
[6] M. Väth, *The Furi-Martelli-Vignoli spectrum vs. the phantom*, Nonlin. Anal. TMA, **47, 9** (2001), 2237–2248.
[7] M. Väth, *Coincidence points of function pairs based on compactness properties*, Glasgow J. Math., **44** (2002), 2, 209–230.

Spectral theory for homogeneous operators: part II. Applications

Jürgen Appell (Würzburg)

This is a continuations of the previous talk by Elena Giorgieri on nonlinear spectral theory for homogenous operators. The following table gives a general comparison of the three spectra introduced in Elena's talk.

Author	*Spectrum*	*Point spectrum*	*Character*
Furi-Martelli-Vignoli [9]	FMV-spectrum $\sigma_{FMV}(F,J)$	asymptotic eigenvalues $\sigma_q(F,J)$	asymptotic $(\|\|u\|\| \to \infty)$
Feng [7]	Feng spectrum $\sigma_F(F,J)$	classical eigenvalues $\sigma_p(F,J)$	global $(u \in X)$
Väth [13]	phantom $\phi(F,J)$	connected eigenvalues $\phi_p(F,J)$	local $(u \in \overline{\Omega})$

As one could expect, there are some relations between all these spectra and point spectra. For example, the Väth phantom $\phi(F,J)$ is always contained in the Furi-Martelli-Vignoli spectrum $\sigma_{FMV}(F,J)$, which in turn is contained in the Feng spectrum $\sigma_F(F,J)$. Moreover, the point phantom $\phi_p(F,J)$ is contained im the asymptotic point spectrum $\sigma_q(F,J)$. So for general operators $F, J : X \to Y$ we get the following relations.

$$
\begin{array}{ccccc}
\phi(F,J) & \subseteq & \sigma_{FMV}(F,J) & \subseteq & \sigma_F(F,J) \\
\cup\!| & & \cup\!| & & \cup\!| \\
\phi_p(F,J) & \subseteq & \sigma_q(F,J) & & \sigma_p(F,J)
\end{array}
$$

In the linear case $L \in \mathfrak{L}(X)$ (and $J = I$) this table essentially simplifies. Here all the spectra in the first row coincide with the usual spectrum $\sigma(L)$, and both the point spectrum $\sigma_p(L, I)$ and point phantom $\phi_p(L, I)$ coincide with the usual point spectrum $\sigma_p(L)$.

Even if one restricts the class of nonlinear operators in consideration, the above table may simplify. We confine ourselves to the case of *τ-homogeneous operators* F and J, i.e.

$$F(tu) = t^\tau F(u), \quad J(tu) = t^\tau J(u) \qquad (t > 0,\, u \in X). \tag{29}$$

The following two theorems have been proved in [2].

Theorem 4 (Coincidence theorem). *Let X and Y be infinite dimensional Banach spaces, and suppose that $F, J : X \to Y$ satisfy* (29) *for some $\tau > 0$. Then*

$$\sigma_{FMV}(F,J) = \sigma_F(F,J) = \phi(F,J), \quad \sigma_q(F,J) \supseteq \sigma_p(F,J) = \phi_p(F,J).$$

Theorem 5 (Discreteness theorem). *Let X and Y be infinite dimensional Banach spaces, and suppose that $F, J : X \to Y$ are odd, $[F]_A = 0$ (i.e., F is compact), and $[J]_a > 0$. Then*

$$\sigma_{FMV}(F,J) \setminus \{0\} \subseteq \sigma_q(F,J), \quad \sigma_F(F,J) \setminus \{0\} \subseteq \sigma_p(F,J),$$

and

$$\phi(F,J) \setminus \{0\} \subseteq \phi_p(F,J).$$

Theorem 5 shows that, for F compact and odd, and J "sufficiently regular" and odd, each nonzero spectral value is actually an eigenvalue (in a sense to be made precise). For F compact and linear and $J = I$ this is a classical fact.

To illustrate how these theorems apply to nonlinear problems, we consider the *eigenvalue problem for the p-Laplacian* which consists in finding solutions $u \not\equiv 0$ of

$$\begin{cases} -\mathrm{div}\,(|\nabla u|^{p-2}\,\nabla u)(x) = \mu\,|u(x)|^{p-2}\,u(x) & \text{in } G \\ u(x) \equiv 0 & \text{on } \partial G \end{cases} \tag{30}$$

where $G \subset \mathbb{R}^n$ is a bounded domain. Although this problem makes sense for $1 < p < \infty$, we restrict ourselves to the case $2 \le p < \infty$. The problem (30) may be reformulated as equivalent operator equation in weak form

$$F_p(u) = \lambda J_p(u), \tag{31}$$

where $\lambda = 1/\mu$, and $F_p, J_p : W_0^{1,p}(G) \to W^{-1,p'}(G)$ $(p' = p/(p-1))$ are defined by $F_p(u) = |u|^{p-2}\, u$ and

$$\langle J_p(u), v\rangle = -\int_G (|\nabla u(x)|^{p-2}\, \nabla u(x), \nabla v(x))\, dx \quad (u, v \in W_0^{1,p}(G)),$$

respectively. Equation (31) has been studied by many authors, e.g. by Drábek et al. in [3–6]. Interestingly, the eigenvalue theory for the problem (30) has many features in common with the classical *linear* eigenvalue problem $-\Delta u(x) = \mu u(x)$, which is a special case of (30) for $p = 2$. For instance, the first eigenvalue μ_1 of (30) is always positive and simple and may be "calculated" as Rayleigh quotient

$$\mu_1 = \inf_{\substack{u \in W_0^{1,p}(G)\\ u \neq 0}} \frac{\displaystyle\int_G |\nabla u(x)|^p\ dx}{\displaystyle\int_G |u(x)|^p\ dx}.$$

Moreover, the corresponding eigenfunction $u_1 \in W_0^{1,p}(G)$ is positive on G and simple (in the sense that any other eigenfunction is a scalar multiple of u_1). This function has the same "variational characterization" as in the linear case $p = 2$: it minimizes the functional $\Psi_p : W_0^{1,p}(G) \to \mathbb{R}$ defined by $\Psi_p(u) = \frac{1}{p}\langle J_p(u), u\rangle$, subject to the constraint

$$\frac{1}{p}\int_G |u(x)|^{p-2}\, u(x)\, dx = 1.$$

Finally, we point out that there is a famous so-called *nonlinear Fredholm alternative* (see [8, 11, 12]) which implies that the operator $J_p - \mu F_p = \mu(\lambda J_p - F_p)$ is *onto* for $\mu < \mu_1$, while it is *not onto* for $\mu = \mu_1$. However, the coincidence and discreteness theorems given above allow us a more precise statement. The following is just a reformulation of Theorems 4 and 5.

Theorem 6 (Nonlinear Fredholm alternative). *Suppose that $J : X \to Y$ is an odd τ-homogeneous homeomorphism with $[J]_a > 0$, and $F : X \to Y$ is odd, τ-homogeneous and compact. Let $\lambda \neq 0$. Then the following four assertions are equivalent.*

(a) *The eigenvalue problem* (30) *has only the trivial solution $u = 0$.*
(b) *The operator $\lambda J - F$ is stably solvable, $[\lambda J - F]_a > 0$, and $[\lambda J - F]_q > 0$.*
(c) *The operator $\lambda J - F$ is epi on each $\Omega \in \mathfrak{D}(X)$, $[\lambda J - F]_a > 0$, and $[\lambda J - F]_b > 0$.*
(d) *The operator $\lambda J - F$ is strictly epi on some $\Omega \in \mathfrak{D}(X)$, and*

$$\inf\, \{\|\lambda J(u) - F(u)\| : u \in \partial\Omega\} > 0.$$

We claim that the operators F_p and J_p satisfy the hypotheses of Theorem 6 in the spaces $X = W_0^{1,p}(G)$ and $Y = X^* = W^{-1,p'}(G)$. In fact, since $J_p : X \to Y$ is continuous, strictly monotone, coercive (it is here that we use the restriction $p \geq 2$!), odd, and $(p-1)$-homogeneous, it is an *isomorphism*, by Minty's celebrated theorem [10]. Moreover, the coercivity also implies that $[J_p]_a > 0$. Finally, the operator $F_p : X \to Y$ is continuous, compact (by Krasnosel'skij's theorem and the compactness of the imbedding $X \hookrightarrow L_p(G)$), odd, and also $(p-1)$-homogeneous. So Theorem 6 implies that, whenever μ is not a classical eigenvalue of (2), then the operator $J_p - \mu F_p$ is not only onto, but even stably solvable and strictly epi. This makes it possible to obtain existence, uniqueness, and stability results for nonlinear perturbations of (31).

Several other applications of nonlinear spectra may be found in Chapter 12 of the recent monograph [1].

References

[1] J. Appell, E. De Pascale, A. Vignoli, *Nonlinear Spectral Theory*, Berlin: deGruyter-Verlag 2004.

[2] J. Appell, E. Giorgieri, M. Väth, *Nonlinear spectral theory for homogeneous operators*, Nonlin. Funct. Anal. Appl. **7** (2002), 589–614.

[3] P. A. Binding, P. Drábek, Y. X. Huang, *On the Fredholm alternative for the p-Laplacian*, Proc. Amer. Math. Soc **125, 12** (1997), 3555–3559.

[4] M. del Pino, P. Drábek, R. Manásievich, *The Fredholm alternative at the first eigenvalue for the one dimensional p-Laplacian*, J. Differ. Equ. **151, 2** (1999), 386–419.

[5] P. Drábek, *Fredholm alternative for the p-Laplacian: yes or no?*, in: Function Spaces, Diff. Oper. and Nonlin. Anal. [Ed.: V. Mustonen], Math. Inst. Acad. Sci., Prague 2000, 57–64.

[6] P. Drábek, G. Holubová, *Fredholm alternative for the p-Laplacian in higher dimensions*, J. Math. Anal. Appl. **263** (2001), 182–194.

[7] W. Feng, *A new spectral theory for nonlinear operators and its applications*, Abstr. Appl. Anal. **2** (1997), 163–183.

[8] S. Fučik, *Fredholm alternative for nonlinear operators in Banach spaces and its applications to differential and integral equations*, Comm. Math. Univ. Carol. **11, 2** (1970), 271–284.

[9] M. Furi, M. Martelli, A. Vignoli, *Contributions to the spectral theory for nonlinear operators in Banach spaces*, Annali Mat. Pura Appl. **118** (1978), 229–294.

[10] G. Minty, *Monotone nonlinear operators in Hilbert space*, Duke Math. J. **29** (1962), 341–346.

[11] J. Nečas, *Sur l'alternative de Fredholm pour les opérateurs non linéaires avec applications aux problèmes aux limites*, Annali Scuola Norm. Sup. Pisa **23** (1969), 331–345.

[12] S. I. Pokhozhaev, *Solvability of nonlinear equations with odd operators* [in Russian], Funkts. Anal. Prilozh. **1, 3** (1967), 66–73; Engl. transl.: Funct. Anal. Appl. **1, 3** (1967), 227–233.

[13] M. Väth, *The Furi-Martelli-Vignoli spectrum vs. the phantom*, Nonlin. Anal. TMA **47, 6** (2001), 2237–2248.

Numerical Ranges for Nonlinear Operators: A Survey

Jürgen Appell (Würzburg)

This talk was supposed to be given by E. De Pascale (Cosenza, Italy) who was unable to come to Oberwolfach.

The purpose of the talk is to give an overview of the definition and properties of numerical ranges for both linear and nonlinear operators in Hilbert or Banach spaces. For linear operators in Hilbert spaces this goes back to Toeplitz [12], for linear operators in Banach spaces to Bauer [1], and, independently, to Lumer [7]. In the nonlinear case, corresponding definitions have been given in the Hilbert space setting by Zarantonello [13–15], and in the Banach space setting by Rhodius [9–11], Martin [8], Dörfner [4], and Feng [5]. Somewhat different notions of numerical ranges are due to Furi, Martelli and Vignoli [6], Bonsall, Cain and Schneider [2], and Canavati [3].

Numerical ranges and radii have applications in matrix theory, numerical analysis, approximation theory, functional analysis, operator theory, system theory, and even in quantum mechanics. There are also useful for "localizing" the spectrum of an operator in the complex plane. This provides the connection with the topics dealt with in the Miniworkshop.

References

[1] F. L. Bauer, *On the field of values subordinate to a norm*, Numer. Math. 4 (1962), 103–111.

[2] F. F. Bonsall, B. E. Cain, H. Schneider, *The numerical range of a continuous mapping of a normed space*, Aequationes Math. 2 (1968), 86–93.

[3] J. Canavati, *A theory of numerical range for nonlinear operators*, J. Funct. Anal. (1979), 231–258.

[4] M. Dörfner, *A numerical range for nonlinear operators*, Z. Anal. Anw. 15 (1996), 445–456.

[5] W. Feng, *A new spectral theory for nonlinear operators and its applications*, Abstr. Appl. Anal. 2 (1997), 163–183.

[6] M. Furi, M. Martelli, A. Vignoli, *Contributions to the spectral theory for nonlinear operators in Banach spaces*, Annali Mat. Pura Appl. 118 (1978), 229–294.

[7] G. Lumer, *Semi-inner product spaces*, Trans. Amer. Math. Soc. 100 (1961), 29–43.

[8] R. H. Martin, *Nonlinear Operator and Differential Equations in Banach Spaces*, J. Wiley, New York 1976.

[9] A. Rhodius, *Der numerische Wertebereich für nicht notwendig lineare Abbildungen in nicht notwendig lokalkonvexen Räumen*, Math. Nachr. 72 (1976), 169–180.

[10] A. Rhodius, *Der numerische Wertebereich und die Lösbarkeit linearer und nichtlinearer Gleichungen*, Math. Nachr. 79 (1977), 343–360.

[11] A. Rhodius, *Über numerische Wertebereiche und Spektralwertabschätzungen*, Acta Sci. Math. 47 (1984), 465–470.

[12] O. Toeplitz, *Das algebraische Analogon zu einem Satz von Féjèr*, Math. Z. 2 (1918), 187–197.

[13] E. H. Zarantonello, *The closure of the numerical range contains the spectrum*, Bull. Amer. Math. Soc. 70 (1964), 781–787.

[14] E. H. Zarantonello, *The closure of the numerical range contains the spectrum*, Pacific J. Math. 22 (1967), 575–595.

[15] E. H. Zarantonello, *Proyecciones sobre conjuntos convexos en el espacio de Hilbert y teoría espectral*, Revista Unión Mat. Argentina 26 (1972), 187–201.

Reporter: Jürgen Appell

Participants

Prof. Dr. Jürgen Appell
appell@mathematik.uni-wuerzburg.de
Mathematisches Institut
Universität Würzburg
Am Hubland
97074 Würzburg

Prof. Dr. Raffaele Chiappinelli
chiappinelli@unisi.it
Dipartimento di Scienze Matematicae
ed Informatiche
Universita di Siena
Pian dei Mantellini, 44
I-53100 Siena

Prof. Dr. Mabel Cuesta
mcuesta@ulb.ac.be
Departement de Mathematiques
Universite du Littoral
50 Rue F. Buisson
B.P. 699
F-62228 Calais

Prof. Dr. Pavel Drabek
pdrabek@kma.zcu.cz
University of West Bohemia
Department of Mathematics
P.O. Box 314
30614 Plzen
CZECH REPUBLIC

Prof. Dr. Wenying Feng
wfeng@trentu.ca
Department of Mathematics and
Computer Science
Trent Univeristy
Peterborough ONT K9J 7B8
Canada

Prof. Dr. Massimo Furi
furi@dma.unifi.it
Dipart. di Matematica Applicata
Universita degli Studi di Firenze
Via S. Marta, 3
I-50139 Firenze

Prof. Dr. Elena Giorgieri
giorgier@mat.uniroma2.it
Dipartimento di Matematica
II. Universita di Roma
Via della Ricerca Scientifica
I-00133 Roma

Prof. Dr. Petr Girg
pgirg@kma.zcu.cz
c/o Prof. Dr. Peter Takac
Fachbereich Mathematik
Universität Rostock
Universitätsplatz 1
18055 Rostock

Prof. Dr. Jean-Pierre Gossez
gossez@ulb.ac.be
Dept. de Mathematiques
Unviersite Libre de Bruxelles
CP 214 Campus Plaine
Bd. du Triomphe
B-1050 Bruxelles

Prof. Dr. Vesa Mustonen
vesa.mustonen@oulu.fi
Department of Mathematical Sciences
University of Oulu
PL 3000
Linnanmaa
FIN-90014 Oulu

Stephen B. Robinson
sbr@wfu.edu
Department of Mathematics
Wake Forest University
P.O.Box 7388
Winston-Salem NC 27109
USA

Prof. Dr. Charles A. Stuart
charles.stuart@epfl.ch
Departement de Mathematiques
Ecole Polytechnique Federale
de Lausanne
CH-1015 Lausanne

Prof. Dr. Peter Takac
takac@hades.math.uni-rostock.de
peter.takac@mathematik.uni-rostock.de
Fachbereich Mathematik
Universität Rostock
18051 Rostock

Martin Väth
vaeth@mathematik.uni-wuerzburg.de
Mathematisches Institut
Universität Würzburg
Am Hubland
97074 Würzburg

Prof. Dr. Jeffrey Webb
jrlw@maths.gla.ac.uk
Department of Mathematics
University of Glasgow
University Gardens
GB-Glasgow, G12 8QW

Mathematisches Forschungsinstitut Oberwolfach

Report No. 9/2004

Mini-Workshop: Classification of Surfaces of General Type with Small Invariants

Organised by
Fabrizio Catanese (Bayreuth)
Ciro Ciliberto (Roma)

February 15th – February 21st, 2004

Introduction by the Organisers

This mini-workshop has been organized by Fabrizio Catanese and Ciro Ciliberto. Unfortunately Catanese was unable to participate.

The classification of algebraic surfaces is a long-standing research subject in algebraic geometry, started by Castelnuovo and Enriques more than one hundred years ago, and continued by the Italian school (Severi, de Franchis, etc.) until about 1950.

In more recent times, fundamental contributions have been given by Kodaira in the 1950's and later in the 1970's by Bombieri, whose works on pluricanonical maps gave a strong impulse in studying surfaces of general type, and Mumford.

Adding important information to classical results by Noether and Castelnuovo, sharp bounds on the invariants have been given by Miyaoka and Bogomolov-Yau, allowing many authors to develop a systematic study of the "geography" of surfaces of general type.

Interesting investigations about the moduli space of surfaces of general type have been worked out in the last twenty years by Catanese, Manetti, and others.

Despite the intensive effort made in the last decades in order to make more precise our knowledge about surfaces of general type, their fine classification is still an open problem, even for small invariants. It is actually rather embarassing that, after more than one century of research on the subject, a complete classification of surfaces with geometric genus zero or one is still lacking.

This mini-workshop carried together 14 mathematicians actively working on this subject, and related arguments, with the idea of updating the state-of-the-art, exchanging information, discussing interesting open problems and stimulating collaborations. In this respect, the workshop has been very successful.

The atmosphere has been lively and very collaborative. During every talk, several questions have been posed and interesting problems pointed out. It has been especially remarkable the active presence of young participants.

During the week, 16 formal lectures have been given by the participants. This report contains extended abstracts of all the talks and also a contribution by Catanese, in collaboration with Pignatelli, about the lecture he was supposed to give.

The topics include: pluri-canonical maps for surfaces of general type (M. Mendes Lopes), canonical rings, projective embeddings and birational techniques (C. Böhning, F. Catanese, S. Papadakis, U. Persson, R. Pignatelli), irregular surfaces with low invariants (F. Polizzi, F. Zucconi), surfaces with $p_g = 0$ (A. Calabri, C. Ciliberto, K. Keum, M. Mendes Lopes, C. Werner), general techniques (V. Brînzănescu, K. Konno). Ulf Persson chaired an "open problem and discussions" session, which especially concerned surfaces with $p_g = 0$.

The organizers thank the Institute staff for providing a comfortable environment to the participants.

Mini–Workshop on Classification of Surfaces of General Type with Small Invariants

Table of Contents

Abstracts

Canonical surfaces in $\mathbb{P}^4$ and Gorenstein algebras in codimension 2

Christian Böhning

Consider minimal surfaces of general type S with $p_g = 5$, $q = 0$ such that the 1-canonical map π is a birational morphism onto a surface $Y \subset \mathbb{P}^4$, the latter being referred to as a canonical surface in $\mathbb{P}^4$. The canonical ring $\mathcal{R} := \bigoplus_{n\geq 0} H^0(S, \mathcal{O}_S(nK))$ is then a Gorenstein algebra of codimension 2 with twist -6 over $\mathcal{A} := \mathbb{C}[x_0, \ldots, x_4]$, the homogeneous coordinate ring of $\mathbb{P}^4$. In general I make the

Definition. Let $S = S_0 \oplus S_1 \oplus S_2 \oplus \ldots$ be a positively graded ring with S_0 a field, S finitely generated over S_0 as an algebra; a finite graded perfect S-algebra B is called a *Gorenstein S-algebra of codimension c* (and *with twist $t \in \mathbb{Z}$*) if $B \cong \mathrm{Ext}^c_S(B, S(t))$ as B-modules where $c = \dim S - \dim_S B$.

By Castelnuovo's second inequality and Bogomolov-Miyaoka-Yau $8 \leq K^2 \leq 54$ for the above surfaces, the complete intersections of type $(2,4)$ resp. $(3,3)$ being the only solutions for $K^2 = 8$ resp. $= 9$. Moreover (cf. [Cil], [Cat4], [Böh1])

Theorem 1. *For a canonical surface in $\mathbb{P}^4$ with $q = 0$, $p_g = 5$, $K^2 \geq 10$ one has a resolution of the canonical ring $\mathcal{R}$*

$$(1)\ \ \mathbf{R}_\bullet :\ 0 \to \mathcal{A}(-6) \oplus \mathcal{A}(-4)^n \xrightarrow{\binom{-\beta^t}{\alpha^t}} \mathcal{A}(-3)^{2n+2} \xrightarrow{(\alpha\,\beta)} \mathcal{A} \oplus \mathcal{A}(-2)^n \to \mathcal{R} \to 0,$$

where $n := K^2 - 9$.

Resolution (1) displays the symmetry of a "generalized" Koszul complex (cf. [Gra]). The important point, however, is that knowledge of the resolution (1) easily allows us to reconstruct our entire geometric set-up; more precisely (cf. [Böh1], [Böh2])

Theorem 2. *Let $\mathcal{R}$ be some finite $\mathcal{A}$-module with minimal graded free resolution as in (1). Write $A := (\alpha\,\beta)$, $A' := A$ with first row erased, $I_n(A')$ = Fitting ideal of $n \times n$ minors of A', and assume* $\mathrm{depth}\, I_n(A') \geq 4$.

Then $\mathcal{R}$ is a Gorenstein algebra, and if one assumes that $\mathrm{Ann}_{\mathcal{A}}(\mathcal{R})$ *is a prime ideal, then $Y := \mathrm{Supp}(\mathcal{R}) \subseteq \mathbb{P}^4$ with its reduced induced subscheme structure (thus the ideal of polynomials vanishing on Y is $\mathcal{I}_Y = \mathrm{Ann}_{\mathcal{A}}\,\mathcal{R}$) is an irreducible surface, and if furthermore one assumes $X := \mathrm{Proj}(\mathcal{R})$ has only rational double points as singularities, then X is the canonical model of a surface S of general type with $q = 0$, $p_g = 5$, $K^2 = n + 9$. More precisely, writing $\mathcal{A}_Y$ for the homogeneous coordinate ring of Y, one has that the morphism $\psi : X \to Y \subset \mathbb{P}^4$ induced by the* *inclusion* $\mathcal{A}_Y \subset \mathcal{R}$ *is a finite birational morphism, and is part of a diagram*

$$\begin{tikzcd} S \arrow[rr, "\pi"] \arrow[dr, "\kappa"'] & & Y \subset \mathbb{P}^4 \\ & X \arrow[ur, "\psi"'] & \end{tikzcd}$$

where S *is the minimal desingularization of* X, κ *is the contraction morphism contracting exactly the (-2)-curves of* S *to rational double points on* X, *and the composite* $\pi := \psi \circ \kappa$ *is a birational morphism with* $\pi^*\mathcal{O}_{\mathbb{P}^4}(1) = \mathcal{O}_S(K_S)$ *(i.e. is 1-canonical for* S*).*

In some sense the most delicate part of the above theorem consists in recovering the ring structure of $\mathcal{R}$ from the resolution (1), cf. [Böh2], thm. 1.3 and 2.5.
To see how the above theorem may be applied, take $K^2 = 11$ as sample case: here the symmetry condition $\alpha\beta^t = \beta\alpha^t$ can be explicitly solved (cf. [Böh1], section 2) in order to re-prove by this method a result previously obtained by D. Roßberg (cf. [Roß]) with different techniques:

Theorem 3. *There is a unique irreducible component of the moduli space of regular surfaces of general type with* $p_g = 5$, $K^2 = 11$ *containing points corresponding to surfaces with canonical map a birational morphism onto a surface* $Y \subset \mathbb{P}^4$ *with only isolated singularities, which is unirational and of dimension 38.*

It may be hoped that this method will facilitate the study of canonical surfaces with higher K^2, the first unsolved case being $K^2 = 13$.

References

[Böh1] C. Böhning, *Canonical surfaces in* $\mathbb{P}^4$ *and Gorenstein algebras in codimension* 2, diploma thesis, Göttingen 2001, math.AG/0402369.

[Böh2] C. Böhning, *L. Szpiro's conjecture on Gorenstein algebras in codimension* 2, preprint, 2003, math.AC/0402370.

[Bom] E. Bombieri, *Canonical models of surfaces of general type*, I.H.E.S. Publ. Math. **42** (1973), 171–219.

[B-He] W. Bruns, J. Herzog, *Cohen-Macaulay rings*, CUP 1998.

[Cat1] F. Catanese, *Babbage's conjecture, contact of surfaces, symmetric determinantal varieties and applications*, Inv. Math. **63** (1981), 433–465.

[Cat1b] F. Catanese, *On the moduli spaces of surfaces of general type*, J. Differential Geom. **19** (1984), 483-515.

[Cat2] F. Catanese, *Commutative algebra methods and equations of regular surfaces*, Algebraic Geometry-Bucharest 1982, LNM 1056, Springer 1984, 68–111.

[Cat3] F. Catanese, *Equations of pluriregular varieties of general type*, Geometry today-Roma 1984, Progr. in Math. 60, Birkhäuser 1985, 47–67.

[Cat4] F. Catanese, *Homological Algebra and Algebraic surfaces*, Proc. Symp. in Pure Math. 62.1, 1997, 3–56.

[Cil] C. Ciliberto, *Canonical surfaces with* $p_g = p_a = 5$ *and* $K^2 = 10$, Ann. Sc. Norm. Sup. s.IV, **9**, 2 (1982), 287–336.

[D-E-S] W. Decker, L. Ein, F.-O. Schreyer, *Construction of surfaces in* $\mathbb{P}^4$, J. Alg. Geom. **2** (1993), 185–237.

[E-U] D. Eisenbud, B. Ulrich, *Modules that are Finite Birational Algebras*, Ill. Jour. Math. **41**, 1 (1997), 10–15.

[En] F. Enriques, *Le superficie algebriche*, Zanichelli, Bologna, 1949.

[Gra] M. Grassi, *Koszul modules and Gorenstein algebras*, J. Alg. **180** (1996), 918–953.

[dJ-vS] T. de Jong, D. van Straten, *Deformation of the normalization of hypersurfaces*, Math. Ann. **288** (1990), 527–547.

[M-P] D. Mond, R. Pellikaan, *Fitting ideals and multiple points of analytic mappings*, LNM 1414, Springer 1987, 107–161.

[P-S1] C. Peskine, L. Szpiro, *Dimension projective finie et cohomologie locale*, Publ. Math. IHES **42** (1972).

[P-S2] C. Peskine, L. Szpiro, *Liaison des variétés algébriques*, Inv. Math. **26** (1972), 271–302.

[Roß] D. Roßberg, *Kanonische Flächen mit $p_g = 5$, $q = 0$ und $K^2 = 11, 12$*, Doktorarbeit Bayreuth, November 1996.

On the classification of surfaces of general type with non birational bicanonical map and Du Val double planes

Giuseppe Borrelli

Let S be a minimal surface of general type and consider the bicanonical map φ_{2K} associated to the linear system $|2K_S|$. If there exists a rational map $S \to B$ onto a curve B with the general fiber a smooth irreducible curve of genus 2 then φ_{2K} is not birational, and in this situation one says that S presents the *standard case* (for the non birationality of φ_{2K}). By a theorem of I. Reider the standard case is the only possible exception to φ_{2K} being birational when $K_S^2 \geq 10$. In the 1950's P. Du Val [6] considered the problem for regular surfaces ($q = h^1(S, \mathcal{O}_S) = 0$), he obtained a list of possible surfaces with non birational bicanonical map and do not presenting the standard case. The examples of Du Val are as follows. Let X be a smooth surface and $G \subset X$ a reduced curve such that

- $\mathcal{B}$) *either $X = \mathbb{F}_2$ and $G = C_0 + G'$, where $G' \in |7C_0 + 14\Gamma|$ and G' has at most non essential singularities;*
- $\mathcal{D}$) *or $X = \mathbb{P}^2$ and G is a smooth curve of degree 8;*
- $\mathcal{D}_n$) *or $X = \mathbb{P}^2$ and $G = G' + L_1 + \cdots + L_n$, with $n \in \{0, 1, \ldots, 6\}$ ($G = G'$ if $n = 0$), where $L_1, \ldots, L_n$ are distinct lines meeting at a point γ and G' is a curve of degree $10 + n$. The singularities of G, besides the non essential ones, are a $(2n+2)$-tuple point at γ, a $[5,5]$-point lying on L_i, $i = 1, \ldots, n$, possibly some 4-tuple points or $[3,3]$-points;*

then S is the smooth minimal model of the double cover $X' \to X$ branched along G. Here $\mathbb{F}_2$ is the Hirzebruch surface $\mathbb{P}(\mathcal{O}_{\mathbb{P}^1} \oplus \mathcal{O}_{\mathbb{P}^1}(2))$ and Γ, C_0 its fibre and negative section with $C_0^2 = -2$. We will refer to X' as a *Du Val double plane* (of type $\mathcal{B}$, $\mathcal{D}$ or $\mathcal{D}_n$).

The exceptions to the standard case have been classified for surfaces with $p_g \geq 4$ by C. Ciliberto, P. Francia and M. Mendes Lopes [4]; F. Catanese, C. Ciliberto and M. Mendes Lopes classified those with $p_g = 3, q > 0$ [3], and C. Ciliberto and M. Mendes Lopes worked out the regular case with $p_g = 3$ [5]. Finally, I classified the regular case with $p_g = 2$ under the assumption that the canonical system has no fixed part [1]. It follows from [5, 4], that if $q = 0, p_g \geq 3$ and φ_{2K}

is non birational then S either presents the standard case or is one of the Du Val examples.

It is easy to see that if φ_{2K} is non birational, the conditions $p_g \geq 2, q = 0$ force φ_{2K} to be a map of degree 2 (generically) onto a rational ruled surface. Hence, it is natural to consider more in general a surface whose bicanonical map factors through a rational map of degree 2 onto a rational or ruled surface, that is if there exists the following commutative diagram

$$\begin{tikzcd} S \arrow[r, dashed, "\varphi_{2K}"] \arrow[d, dashed, "\phi"'] & S_2 \\ \Sigma \arrow[ur, dashed, "\phi_2"'] & \end{tikzcd}$$

where ϕ is a (generically finite) rational map of degree two and Σ is a rational or ruled surface. The result is the following,

Theorem 1. *Let S be a smooth minimal surface of general type which does not present the standard case. Then the following three conditions are equivalent:*

a) *the bicanonical map of S factors through a rational map of degree 2 onto a rational or ruled surface*
b) *the bicanonical map of S factors through a rational map of degree 2 onto a rational surface*
c) *S is the smooth minimal model of a Du Val double plane.*

Moreover, let S be as in (c) (resp. (a) or (b)) then:

d) $q(S) = 0$ *unless* $p_g(S) = q(S) = 1$;
e) *unless K_S is ample and $p_g(S) = 6, K_S^2 = 8$ or $p_g(S) = 3, K_S^2 = 2$, there is a rational pencil whose general member is a smooth hyperelliptic curve of genus 3 such that the bicanonical map of S induces the hyperelliptic involution on it.*

Sketch of the proof of Theorem 1, $(a) \Rightarrow (b), (c)$. (See [2] for the complete proof.) Consider the quotient Σ_σ of S by the involution σ induced by ϕ. Then Σ_σ is a rational or ruled surface birational equivalent to Σ whose only singularities are the k nodes, which corresponds to the isolated fixed points of σ. Let $\hat{\Sigma} \to \Sigma_\sigma$ be the minimal resolution, then we have the commutative diagram

$$\begin{tikzcd} \hat{S} \arrow[r] \arrow[d, "\rho"'] & S \arrow[d] \\ \hat{\Sigma} \arrow[r] & \Sigma_\sigma \end{tikzcd}$$

where $\hat{S}$ is the blow up of S at the isolated fixed points of σ and ρ is a finite double cover branched along a smooth curve B. Since $\hat{\Sigma}$ is smooth it is either $\mathbb{P}^2$ or ruled. When $\hat{\Sigma} \cong \mathbb{P}^2$ one has that $\hat{S} = S$ and B has degree 8 or 10. Otherwise we have that

i) $\hat{\Sigma}$ is rational,

ii) there exists a suitable birational morphism $\psi : \hat{\Sigma} \to X$ such that $G := \psi_*(B)$ and X are as in $\mathcal{B}$ or $\mathcal{D}_n$,

iii) $\hat{S}$ is the *canonical resolution* of the double cover $X' \to X$ branched along G.

For the proof of $i), ii)$ one uses a result of Xiao [9] who studied the possible images of the bicanonical map. □

As we remarked, the result for regular surfaces with $p_g \geq 3$ was already known and Theorem 1 extends the classification to regular surfaces with $p_g = 2$,

Theorem 2. *Let S be a regular surface of general type with $p_g \geq 2$ and non birational bicanonical map. Then either S presents the standard case or it is the smooth minimal model of a Du Val double plane.*

For $p_g = 0, 1$ we get some corollaries of Theorem 1.

Theorem 3 ([2, 9])**.** *Let S be a regular surface of general type with $p_g = 1$ and bicanonical map of degree 2. Then,*

i) *either S presents the standard case*

ii) *or S is the smooth minimal model of a Du Val double plane of type $\mathcal{D}_n$,*

iii) *or S_2 is a $K3$ surface.*

Theorem 4 ([2, 7, 9])**.** *Let S be a minimal surface of general type with $p_g = 0, K_S^2 \geq 2$ and bicanonical map of degree 2. Then,*

i) *either S presents the standard case*

ii) *or $K_S^2 = 3$ and $\varphi_{2K}(S)$ is an Enriques surface,*

iii) *or S is the smooth minimal model of a Du Val double plane of type $\mathcal{D}_n$ with K_S^2 and n as in the following table*

K_S^2	2	3	4	5	6	7	8
n	$0,1,2,3$	$1,2,3$	$2,3,4$	$3,4$	$4,5$	5	6

References

[1] G. Borrelli, *On regular surfaces of general type with $p_g = 2$ and non birational bicanonical map*, in Beltrametti et al. (eds.), *Algebraic Geometry, A volume in memory of Paolo Francia*, De Gruyter 2002, 65–78.

[2] G. Borrelli, *On the classification of surface of general type with non birational bicanonical map and Du Val double planes*, preprint AG/03123351.

[3] F. Catanese, C. Ciliberto, M. Mendes Lopes, *On the classification of irregular surfaces of general type with non birational bicanonical map*, Trans. Amer. Math. Soc. **350** (1998), 275–308.

[4] C. Ciliberto, P. Francia, M. Mendes Lopes, *Remarks on the bicanonical map for surfaces of general type*, Math. Z. **224** (1997), 137–166.

[5] C. Ciliberto, M. Mendes Lopes, *On regular surfaces of general type with $p_g(S) = 3$ and non birational bicanonical map*, J. Math. Kyoto Univ. **40** (2000), 79–117.

[6] P. Du Val, *On surfaces whose canonical system is hyperelliptic*, Canadian J. of Math. **4** (1952), 204–221.

[7] M. Mendes Lopes, R. Pardini, *Enriques surfaces with eight nodes*, Math. Z. **241** (2002), 673–683.
[8] I. Reider, *Vector bundles of rank 2 and linear systems on algebraic surfaces*, Ann. of Math. **127** (1988).
[9] G. Xiao, *Degree of the bicanonical map of a surface of general type*, Amer. J. of Math. **112** (5) (1990), 713–737.

Twisted Fourier-Mukai transforms on some elliptic surfaces

Vasile Brînzănescu
(joint work with Ruxandra Moraru)

Let X be a non-singular projective variety. The derived category $D(X)$ of X is a triangulated category whose objects are complexes of sheaves on X with bounded and coherent cohomology sheaves. In general, there exist pairs of non-singular projective varieties (X, Y) for which there are triangle-preserving equivalences $\Phi : D(Y) \to D(X)$. Such equivalences are called Fourier-Mukai transforms. In some cases, Φ takes sheaves to sheaves (not complexes) and this fact is used to study moduli spaces of some sheaves (for example, vector bundles). Sometimes, Fourier-Mukai transforms can be constructed on non-projective complex varieties.

Let $\pi : X \to B$ be a minimal non-Kähler elliptic surface (B a smooth compact connected curve). It is well-known that $X \to B$ is a quasi-bundle over B, i.e. all the smooth fibres are pairwise isomorphic and the singular fibres are multiples of elliptic curves (see [24], [8]). Let T denote the general fibre of π, which is an elliptic curve and let T^* denote the dual of T (i.e. $T^* := \mathrm{Pic}^0(T) \cong T$ non-canonically). It is known that the Jacobian surface associated to $\pi : X \to B$, in this case, is simply $J(X) = B \times T^* \to B$ and the surface $X \to B$ is obtained from its Jacobian surface $B \times T^*$ by a finite number of logarithmic transformations.

Now, we shall define a twisted Fourier-Mukai transform on non-Kähler elliptic surfaces. For simplicity, we shall consider that $\pi : X \to B$ has no multiple fibres, i.e. X is a principal elliptic bundle over B. Then, $X = \Theta^*/ < \tau >$, where Θ is a line bundle over B with positive Chern class l, Θ^* is the complement of the zero section in the total space of Θ, and $< \tau >$ is the multiplicative cyclic group generated by a fixed complex number τ with $| \tau |> 1$. The standard fibre of this bundle is $T \cong \mathbb{C}^*/ < \tau >$. Multiplication by τ defines a natural $\mathbb{Z}$-action on $X \times \mathbb{C}^*$ that is trivial on X, inducing the quotient $(X \times \mathbb{C}^*)/\mathbb{Z} = X \times T^* = X \times_B J(X)$.

Since X does not have multiple fibres, then the set of all holomorphic line bundles on X with trivial Chern class is given by the zero component of the Picard group $\mathrm{Pic}^0(X) \cong \mathrm{Pic}^0(B) \times \mathbb{C}^*$. In this case, any line bundle in $\mathrm{Pic}^0(X)$ is therefore of the form $H \otimes L_\alpha$, where H is the pullback to X of an element of $\mathrm{Pic}^0(B)$ and L_α is the line bundle corresponding to the constant automorphy factor $\alpha \in \mathbb{C}^*$; in particular, there exists a universal (Poincaré) line bundle $\mathcal{U}$ on $X \times \mathrm{Pic}^0(X)$ whose restriction to $X \times \mathbb{C}^* := X \times \{0\} \times \mathbb{C}^*$ is constructed in terms of constant automorphy factors (for details, see [10]).

Given a rank two vector bundle over X, its restriction to a generic fibre of π is semistable. More precisely, its restriction to a fibre $\pi^{-1}(b)$ is unstable on at most an isolated set of points $b \in B$; these isolated points are called the *jumps* of the bundle. Furthermore, there exists a divisor in the relative Jacobian $J(X) = B \times T^*$ of X, called the *spectral curve* or *cover* of the bundle, that encodes the isomorphism class of the bundle over each fibre of π. The spectral curve can be constructed as follows: we associate to the rank-2 vector bundle E the sheaf on $B \times \mathbb{C}^*$ defined by

$$\widetilde{\mathcal{L}} := R^1\pi_*(s^*E \otimes \mathcal{U}),$$

where $s : X \times \mathbb{C}^* \to X$ is the projection onto the first factor, id is the identity map, and π also denotes the projection $\pi := \pi \times id : X \times \mathbb{C}^* \to B \times \mathbb{C}^*$. This sheaf is supported on a divisor $\widetilde{S_E}$, defined with multiplicity, that descends to a divisor S_E in $J(X)$ of the form

$$S_E := \left(\sum_{i=1}^{k} \{x_i\} \times T^*\right) + \overline{C},$$

where $\overline{C}$ is a bisection of $J(X)$ and $x_1, \cdots, x_k$ are points in B that correspond to the jumps of E. The spectral curve of E is defined to be the divisor S_E. Note that there is also a natural $\mathbb{Z}$-action on $B \times \mathbb{C}^*$ defined as multiplication by τ on the second factor and $(B \times \mathbb{C}^*)/\mathbb{Z} \cong J(X)$. Moreover, this action extends to the torsion sheaf $\widetilde{\mathcal{L}} := R^1\pi_*(s^*E \otimes \mathcal{U})$, taking the stalk $\widetilde{\mathcal{L}}_{(x,\alpha)}$ to $\widetilde{\mathcal{L}}_{(x,\tau\alpha)} \otimes L_{\tau^{-1},x}$. Therefore, $\widetilde{\mathcal{L}}$ cannot descend to $J(X)$ because it is not invariant with respect to this action. To fix this problem, we construct a sheaf $\mathcal{N}$ on $B \times \mathbb{C}^*$ and a $\mathbb{Z}$-action that leaves the tensor product $\widetilde{\mathcal{L}} \otimes \mathcal{N}$ invariant (see [10], [11]). We denote the quotient sheaf

$$\mathcal{L} := (\widetilde{\mathcal{L}} \otimes \mathcal{N})/ \sim .$$

Note that the support of $\mathcal{L}$ is S_E; moreover, if we take the pull back of $\mathcal{L}$ to $B \times \mathbb{C}^*$ and tensor it by $\mathcal{N}^*$, then we recover $\widetilde{\mathcal{L}}$ (we also denote $\mathcal{N}^*$ the sheaf on $B \times \mathbb{C}^*$ obtained by extending the line bundle $\mathcal{N}^*$ on $\widetilde{S_E}$ by zero outside $\widetilde{S_E}$).

Given a locally free sheaf E on X, we define the twisted Fourier-Mukai transform to be the complex of sheaves $\Phi(E)$ on $J(X)$ given by

$$\Phi(E) := (R\pi_* (s^*E \otimes \mathcal{U}) \otimes \mathcal{N}) / \sim .$$

Conversely, if $\mathcal{L}$ is a sheaf on $J(X)$, we define the "inverse" twisted Fourier-Mukai transform as the complex of sheaves $\hat{\Phi}(\mathcal{L})$ on X given by

$$\hat{\Phi}(\mathcal{L}) := R\underline{s}_* \left((\pi^* \left((\rho^*\mathcal{L}) \otimes \mathcal{N}^* \right) \otimes \mathcal{U}^*) / \sim \right),$$

where $\underline{s} : X \times_B J(X) \to X$ is projection onto the first factor, $q : X \times \mathbb{C}^* \to X \times T^* = X \times_B J(X)$ and $\rho : B \times \mathbb{C}^* \to B \times T^* = J(X)$ are the natural quotient maps induced by the $\mathbb{Z}$-actions and π and s are the projections defined above.

We state some of their properties in:

Theorem 1. *(i) Suppose that E is a rank-2 vector bundle on X without jumps. Then, $\Phi^0(E) = 0$ and $\hat{\Phi}^0(\Phi^1(E)) = E$.*

(ii) If $\mathcal{L}$ is a torsion sheaf on $J(X)$, supported on a bisection $C \subset J(X)$, that has rank 1 on the smooth points of C and rank at most 2 on the singular ones, then $\hat{\Phi}^1(\mathcal{L}) = 0$ and $\Phi^1(\hat{\Phi}^0(\mathcal{L})) = \mathcal{L}$.

For the proof, see [11].

We use this result in the classification of rank two vector bundles over non-Kähler elliptic surfaces, including the study of moduli spaces of stable vector bundles (see [11], [12]).

References

[1] W. Barth, C. Peters, and A. Van de Ven, *Compact complex surfaces*, Springer-Verlag, Berlin, Heidelberg, New York, 1984.

[2] C. Bartocci, U. Bruzzo, D. Hernández Ruipérez, and J. Muños Porras, *Mirror symmetry on $K3$ surfaces via Fourier-Mukai transform*, Comm. Math. Phys. **195** (1) (1998), 79–93.

[3] K. Becker, M. Becker, K. Dasgupta, and P. S. Green, *Compactification of heterotic theory on non-Kähler complex manifolds: I*, preprint.

[4] M. Bershadsky, A. Johansen, T. Pantev, and V. Sadov, *On four-dimentional compactifications of F-theory*, Nuclear Phys. B **505** (1-2) (1997), 165–201.

[5] P. J. Braam and J. Hurtubise, *Instantons on Hopf surfaces and monopoles on solid tori*, J. reine Angew. Math. **400** (1989), 146–172.

[6] T. Bridgeland and A. Maciocia, *Fourier-Mukai transforms for $K3$ and elliptic fibrations*, J. Algebraic Geom. **11** (4) (2002), 629–657.

[7] T. Bridgeland, *Fourier-Mukai transforms for elliptic surfaces*, J. reine Angew. Math. **498** (1998) 115–133.

[8] V. Brînzănescu, *Holomorphic vector bundles over compact complex surfaces*, Lect. Notes in Math. 1624, Springer 1996.

[9] V. Brînzănescu and P. Flondor, *Holomorphic 2-vector bundles on non-algebraic 2-tori*, J. reine angew. Math. **363** (1985), 47–58.

[10] V. Brînzănescu and R. Moraru, *Holomorphic vector bundles on non-projective surfaces*, preprint.

[11] V. Brînzănescu and R. Moraru, *Twisted Fourier-Mukai transforms and bundles on non-Kähler elliptic surfaces*, preprint.

[12] V. Brînzănescu and R. Moraru, *Stable bundles on non-Kähler elliptic surfaces*, preprint.

[13] V. Brînzănescu and K. Ueno, *Néron-Severi group for torus quasi bundles over curves*, Moduli of vector bundles (Sanda, 1994; Kyoto, 1994), Lecture Notes in Pure and Appl. Math. 179, Dekker, New York, 1996, 11–32.

[14] A. Căldăraru, *Derived categories of twisted sheaves on Calabi-Yau manifolds*, PhD thesis, Cornell University, 2000.

[15] A. Căldăraru, *Derived categories of twisted sheaves on elliptic threefolds*, J. reine Angew. Math. **544** (2002), 161–179.

[16] G. L. Cardoso, G. Curio, G. Dall'Agata, D. Lüst, P. Manousselis and G. Zoupanos, *Non-Kähler string backgrounds and their five torsion classes*, preprint.

[17] R. Donagi, *Principal bundles on elliptic fibrations*, Asian J. Math. **1** (2) (1997), 214–223.

[18] R. Donagi, B. Ovrut, T. Pantev, and D. Waldram, *Standard-model bundles*, Adv. Theor. Math. Phys. **5** (3) (2001), 563–615.

[19] R. Donagi, B. Ovrut, T. Pantev, and D. Waldram, *Standard models from heterotic M-theory*, Adv. Theor. Math. Phys. **5** (1) (2001), 93–137.

[20] R. Donagi and T. Pantev, *Torus fibrations, gerbes, and duality*, preprint.

[21] R. Friedman, *Rank two vector bundles over regular elliptic surfaces*, Invent. Math. **96** (1989), 283–332.
[22] R. Friedman, J. Morgan, and E. Witten, *Vector bundles over elliptic fibrations*, J. Algebraic Geom. **2** (1999), 279–401.
[23] E. Goldstein anf S. Prokushkin, *Geometric model for complex non-Kähler manifolds with $SU(3)$ structure*, preprint.
[24] K. Kodaira, *On the structure of compact complex analytic surfaces I* , Amer. J. Math. **86** (1964), 751–798.
[25] J. Le Potier, *Fibrés vectoriels sur les surfaces K3*, Séminaire Lelong-Dolbeault-Skoda, LNM 1028, Springer 1983.
[26] R. Moraru, *Integrable systems associated to a Hopf surface*, Canad. J. Math. **55** (3) (2003), 609–635.
[27] S. Mukai, *Duality between $D(X)$ and $D(\hat{X})$ with its application to Picard sheaves*, Nagoya Math. J. **81** (1981), 153–175.
[28] A. Teleman, *Moduli spaces of stable bundles on non-Kähler elliptic fibre bundles over curves*, Expo. Math. **16** (1998), 193–248.

On the classification of numerical Godeaux surfaces with an involution

Alberto Calabri and Ciro Ciliberto
(joint work with Margarida Mendes Lopes)

In the one-century-and-a-half history of algebraic geometry in dimension two, projective surfaces with geometric genus $p_g = 0$ and irregularity $q = 0$ have been studied from the very beginning. They were supposed to be rational by Max Noether, until Enriques suggested the existence of the surfaces with $p_g = q = 0$ and bi-genus $P_2 = 1$ which now bear his name, and Castelnuovo proved in 1896 his celebrated *rationality criterion*, which states that a surface X is rational if and only if $P_2(X) = q(X) = 0$.

In 1931–32, Godeaux and Campedelli gave the first two examples of minimal surfaces of general type with $p_g = 0$ and $K^2 = 1, 2$, respectively. Godeaux considered a quotient of a quintic surface in $\mathbb{P}^3$ by a $\mathbb{Z}/5\mathbb{Z}$-action, whereas Campedelli constructed a double plane, i.e. a double cover of $\mathbb{P}^2$, branched along a degree 10 curve with six points of type $[3, 3]$, that is a triple point with another infinitely near triple point, not lying on a conic.

Campedelli also suggested the construction of a minimal surface of general type with $p_g = 0$ and $K^2 = 1$ as the smooth minimal model of a double plane branched along a curve C of degree 10 with a 4-tuple point and five points of type $[3, 3]$, not lying on a conic. The existence of a curve like C was proved only 50 years later by Kulikov, Oort and Peters. We will say that a double plane is *of Campedelli type* if the branch curve is of this type.

Minimal surfaces of general type with $p_g = 0$ and $K^2 = 1$, nowadays called *numerical Godeaux surfaces*, have been studied by several authors in the last 30 years: Miyaoka (1976), Dolgachev (1977), Reid (1978, 1988), Barlow (1984–85), which gave the first example of a simply connected one, Werner, Craighero-Gattazzo, Naie (1994), Stagnaro (1997), Dolgachev-Werner (1999), Catanese-Pignatelli, Keum-Lee (2000), and others (cf. e.g. [CP]).

Miyaoka proved that the subgroup Tors(S) of torsion elements of the Picard group of a numerical Godeaux surface S is a cyclic group of order strictly less than 6. He classified those with $\mathrm{Tors}(S) = \mathbb{Z}/5\mathbb{Z}$ by describing the canonical ring of the 5-tuple covering given by the torsion, and similarly Miles Reid classified those with $\mathrm{Tors}(S) = \mathbb{Z}/4\mathbb{Z}$ or $\mathbb{Z}/3\mathbb{Z}$.

Some examples of those with $\mathrm{Tors}(S) = \mathbb{Z}/2\mathbb{Z}$ or $\mathrm{Tors}(S) = 0$ have been found by Barlow, Werner and Craighero-Gattazzo (as shown by Dolgachev and Werner); nonetheless the classification problem is still open.

Note that these surfaces are interesting also because of Bloch's conjecture, which states that the Chow group of degree zero 0-cycles on a surface with $p_g = q = 0$ is trivial.

Here we report on a work in progress about the classification of numerical Godeaux surfaces S with an *involution*, i.e. with an automorphism $\sigma : S \to S$ of order 2. A first investigation of this subject has been done by J. Keum and Y. Lee in [KL]: under the assumption that the bicanonical system has no fixed components, they described all the possibilities for the fixed locus of the involution.

We make no assumption on fixed components of the bicanonical system $|2K_S|$ and we follow the ideas contained in joint works of the third author and Rita Pardini, namely we combine the topological and the holomorphic fixed point formulas for involutions on surfaces and the Kawamata-Viehweg vanishing theorem, in order to prove the following:

Theorem 1. *Let S be a minimal surface of general type with $p_g(S) = q(S) = 0$ and an involution $\sigma : S \to S$. The fixed locus of σ is composed of a smooth curve R and k isolated fixed points. Then:*

- $4 \le k \le K_S^2 + 4$;
- $k \equiv K_S^2 \pmod 2$;
- $K_S \cdot R \le K_S^2$ *and equality holds if and only if* $k = K_S^2 + 4$;
- *if* $k = K_S^2 + 4$, *then the bicanonical map* $\phi : S \dashrightarrow \mathbb{P}^{K_S^2}$ *is composed with* σ;
- *if* $|2K_S|$ *has no fixed component, then* ϕ *is composed with* σ *if and only if* $k = K_S^2 + 4$.

In particular, if S is a numerical Godeaux surface, i.e. $K_S^2 = 1$, then $k = 5$, ϕ is composed with σ, $K_S \cdot R = 1$ and $R = \Gamma + Z$, where Z are disjoint (-2)-curves, and $0 \le p_a(\Gamma) \le 2$ (cf. also [KL]*).*

Then we study the quotient surface S/σ, and, by a fine use of adjunction on S/σ and a deep analysis of some Del Pezzo surfaces, we prove the following:

Theorem 2. *A numerical Godeaux surface S with an involution σ is birationally equivalent to one of the following:*

(1) *a double plane of Campedelli type;*
(2) *a double plane branched along the union of two distinct lines r_1, r_2 and a curve B of degree 12 with the following singularities:*
    - *the point $p_0 = r_1 \cap r_2$ of multiplicity 4;*

- *a point* $p_i \in r_i$*,* $i = 1, 2$*, of type* $[4, 4]$*, where the tangent line is* r_i*;*
- *further three points* p_3, p_4, p_5 *of multiplicity 4 and a point* p_6 *of type* $[3, 3]$*, such that there is no conic through* $p_1, \ldots, p_6$*;*

(3) *a double cover of an Enriques surface.*

In case (3)*,* $\mathrm{Tors}(S) = \mathbb{Z}/4\mathbb{Z}$*, whilst in case* (2)*,* $\mathrm{Tors}(S)$ *is either* $\mathbb{Z}/2\mathbb{Z}$ *or* $\mathbb{Z}/4\mathbb{Z}$*.*

Moreover if the fixed locus R *of* σ *has an irreducible component* Γ *of genus* 2*, then* S *belongs to case* (3)*.*

All the previously known constructions of numerical Godeaux surfaces as double planes belong to case (1). Examples of case (3) have been given by Keum and Naie.

We show the existence of examples of case (2) by constructing degree 12 curves with the required singularities: we found out some examples with $\mathrm{Tors}(S) = \mathbb{Z}/2\mathbb{Z}$ and some with $\mathrm{Tors}(S) = \mathbb{Z}/4\mathbb{Z}$. Let us say that a double plane as in case (2) is *of Du Val type*, because it is the degeneration of a double plane, described by Du Val, whose smooth minimal model has $p_g = 4$ and $K^2 = 8$, with non-birational bicanonical map (see [Ci], [Bo]).

In both cases (1) and (2), it is possible to determine the possible configurations of components of the branch curve of the double planes.

In case (3) we prove that the double cover of the Enriques surface is branched along a curve which moves in a pencil whose general member is an irreducible curve of genus 2.

Theorem 1 suggests that it is possible to study in a similar way minimal surfaces of general type with an involution, $p_g = 0$ and $K^2 > 1$, in particular with $K^2 = 2$, i.e. *numerical Campedelli* surfaces.

References

[Bo] G. Borrelli, *On the classification of surfaces of general type with non-birational bicanonical map and Du Val double planes*, preprint, math.AG n. 0312351. See also his abstract in this report.

[CP] F. Catanese, R. Pignatelli, *On simply connected Godeaux surfaces*, in *Complex analysis and algebraic geometry. A volume in memory of M. Schneider*, de Gruyter 2000, 117–153.

[Ci] C. Ciliberto, *The bicanonical map for surfaces of general type*, in *Algebraic geometry—Santa Cruz 1995*, Proc. Sympos. Pure Math. 62, Part 1, Amer. Math. Soc. 1997, 57–84.

[KL] J. Keum, Y. Lee, *Fixed locus of an involution acting on a Godeaux surface*, Math. Proc. Cambridge Philos. Soc. **129** (2000), no. 2, 205–216. See also the abstract by J. Keum in this report.

[MP] M. Mendes Lopes, R. Pardini, *The bicanonical map of surfaces with* $p_g = 0$ *and* $K^2 \geq 7$, I, Bull. London Math. Soc. **33** (2001), 1–10, and II, Bull. London Math. Soc. **35** (2003), 337–343.

[We] C. Werner, *On numerical Godeaux surfaces constructed as double planes*, in this report.

On pencils of small genus

Fabrizio Catanese and Roberto Pignatelli

1. The relative canonical algebra

Throughout this abstract X will be a projective surface, $f : X \to B$ a morphism onto a smooth curve of genus b. Without loss of generality, we may assume that f has connected fibres F of genus g. These maps are studied (see, e.g., [Fuj1], [Fuj2], [Xia]) analyzing their relative canonical algebra.

Definition 1. Consider the relative dualizing sheaf

$$\omega_{X|B} := \omega_X(-f^* K_B).$$

Then the relative canonical algebra $\mathcal{R}(f)$ is the commutative graded algebra $\oplus_0^\infty V_n$, where V_n is the vector bundle on B given as the direct image sheaf $f_*(\omega_{X|B}^n)$

Definition 2. The multiplication maps $\mu_{n,m} : V_n \otimes V_m \to V_{n+m}$ yield natural sheaf homomorphisms

$$S^n(V_1) = S^n(f_*(\omega_{X|B})) \xrightarrow{\sigma_n} V_n = f_*(\omega_{X|B}^n),$$

and we define $\mathcal{T}_n = \text{ coker } \sigma_n$.

Remark 1. By Noether's theorem on canonical curves, $\mathcal{T}_n$ is a torsion sheaf if the general fibre of f is non-hyperelliptic.

Remark 1 shows that the hyperelliptic and the non hyperelliptic case should be treated separately; assume in fact for the time being that a general fibre is hyperelliptic. Then there is a birational involution σ on X, and σ acts linearly on the space of sections $\mathcal{O}_X(U, \omega_{X/B}^n)$, which splits as the direct sum of the $(+1)$-eigenspace and the (-1)-eigenspace. Accordingly, we get direct sums $V_n = V_n^+ \oplus V_n^-$: therefore, in the hyperelliptic case, where obviously $V_1 = V_1^-$, the cokernels $\mathcal{T}_n$ will be bigger than in the non hyperelliptic case.

2. The structure theorems

Let $f : X \to B$ be a genus 2 fibration. The rank 2 vector bundle $V_1 := f_*\omega_{X|B}$ induces a natural factorization of f as $\pi \circ \varphi$, where $\varphi : X \dashrightarrow \mathbb{P}(V_1)$ is a rational map of degree 2, and $\pi : \mathbb{P}(V_1) \to B$ is the natural projection.

The indeterminacy locus of φ is contained in the fibres of f which are not 2-connected, i.e., which split as $\mathcal{E}_1 + \mathcal{E}_2$ with $\mathcal{E}_1\mathcal{E}_2 = 1$. Then $\mathcal{E}_i^2 = -1$, $\mathcal{E}_i$ has arithmetic genus 1 and is called an elliptic cycle. These fibres are recognizable through $\mathcal{T}_2$ as follows.

Lemma 1. *Let $f : X \to B$ be a genus 2 fibration. Then $\mathcal{T}_2$ is the structure sheaf of an effective divisor $\tau \in Div_{\geq 0}(B)$, whose support is given by the points whose corresponding fibres of f are not 2-connected.*

The typical example is given by a fibre consisting of two smooth elliptic curves $\mathcal{E}_1, \mathcal{E}_2$ meeting transversally in a point P'. The blow-up of the point P' maps isomorphically to the fibre F'' of $\mathbb{P}$ over the point $P \in B$, while the elliptic curves $\mathcal{E}_1, \mathcal{E}_2$ are contracted to two distinct points of the fibre F''.

The resolution $\tilde{\varphi}$ of φ is the composition of the contraction of $\mathcal{E}_1, \mathcal{E}_2$ to two simple -2-elliptic singularities, with a finite double cover where the branch curve Δ in $\mathbb{P}$ contains the fibre and has two distinct 4-tuple points on it. More complicated fibres containing elliptic tails can produce different configurations of singularities of the branching divisor of φ: a complete list is the one given by Ogg and by Horikawa in [Ogg],[Hor]. This approach is widely used to construct genus 2 fibrations; the main difficulty is in the construction of Δ, often very singular.

Definition 3. We denote by $\mathcal{A}$ the graded subalgebra of $\mathcal{R}$ generated by V_1 and V_2; let $\mathcal{A}_n$ be its graded part of degree n, $\mathcal{A}_{even} = \oplus_k \mathcal{A}_{2k}$.

It is easy to see that the natural map $Sym(V_2) \to \mathcal{A}_{even}$ is surjective with kernel generated by the image of the map $i_2 : \det V_1^2 \hookrightarrow S^2(V_2)$ defined locally by $i_2(x_0 \wedge x_1)^2 = \sigma_2(x_0)^2 \sigma_2(x_1)^2 - \sigma_2(x_0 x_1)^2$.

Concretely, this gives explicit equations for $\mathbf{Proj}(\mathcal{A})$ as conic subbundle of the $\mathbb{P}^2$-bundle $\mathbb{P}(V_2)$. $\mathbf{Proj}(\mathcal{A})$ and $\mathbb{P}(V_1)$ are clearly birationally equivalent and biregularly equivalent outside the fibers over $\mathrm{supp}(\mathcal{T}_2)$. One can check that the fibres of $\mathrm{supp}(\mathcal{T}_2)$ are in fact the reducible fibres of the conic bundle.

If we consider the natural morphism $\varphi_{\mathcal{A}} : X \to \mathbf{Proj}(\mathcal{A})$ induced by the inclusion $\mathcal{A} \subset \mathcal{R}$ and the natural projection morphism $\pi_{\mathcal{A}} : \mathbf{Proj}(\mathcal{A}) \to B$ we get a new factorization of the fibration ('birational' to the previous one): $f = \pi_{\mathcal{A}} \circ \varphi_{\mathcal{A}}$. The advantage in considering $\varphi_{\mathcal{A}}$ instead of φ is that the branch curve $\Delta_{\mathcal{A}}$ has only simple singularities. In the typical example above described, the elliptic curves $\mathcal{E}_i$ will not be contracted by $\varphi_{\mathcal{A}}$ but they will be double covers of the two lines of the corresponding fibre of the conic bundle.

Lemma 2. *$\mathcal{A}_6$ is the cokernel of the map $\det V_1^2 \otimes V_2 \to S^3(V_2)$ naturally induced by the map i_2 above; note that $\mathcal{A}_6$ depends only on B, V_1 and σ_2. The branch curve $\Delta_{\mathcal{A}}$ is induced by a map $(\det(V_1) \otimes \mathcal{O}_B(\tau))^{\otimes 2} \to \mathcal{A}_6$.*

We can now introduce the building package of a genus 2 fibration:

Definition 4. Define the **associated 5-tuple** (B, V_1, τ, ξ, w) of a genus 2 fibration $f : X \to B$ as follows:

- B is the base curve;
- $V_1 = f_*(\omega_{X|B})$;
- τ is the effective divisor of B with $\mathcal{O}_\tau \cong \mathcal{T}_2$;
- $\xi \in Ext^1_{\mathcal{O}_B}(\mathcal{O}_\tau, S^2(V_1))/Aut_{\mathcal{O}_B}(\mathcal{O}_\tau)$ the class induced by σ_2;
- $w \in \mathbb{P}(H^0(B, \mathcal{A}_6 \otimes (\det(V_1) \otimes \mathcal{O}_B(\tau))^{\otimes -2}))$ inducing $\Delta_{\mathcal{A}}$ on $\mathbf{Proj}(\mathcal{A})$.

Definition 5. We will say that a a 5-tuple (B, V_1, τ, ξ, w) is **admissible** if

- B is a smooth curve;
- V_1 is a vector bundle on B of rank 2;

- $\tau \in Div^+(B)$;
- $\xi \in Ext^1_{\mathcal{O}_B}(\mathcal{O}_\tau, S^2(V_1))/Aut_{\mathcal{O}_B}(\mathcal{O}_\tau)$ yields a vector bundle V_2;
- $w \in \mathbb{P}(H^0(B, \mathcal{A}_6 \otimes (\det(V_1) \otimes \mathcal{O}_B(\tau))^{\otimes -2}))$ inducing $\Delta_{\mathcal{A}}$ on $\mathbf{Proj}(\mathcal{A})$, where $\mathcal{A}_6$ is the vector bundle induced by ξ;

and if moreover they satisfy some open conditions ensuring that the associated double cover has Rational Double Points as singularities.

We do not specify here the open conditions in detail for lack of space. The vector bundle $\mathcal{A}_6$ is 'induced' taking the map σ_2 induced by ξ and defining $\mathcal{A}_6$ as the cokernel of the map in lemma 2.

Theorem 1. *Let f be a relatively minimal genus 2 fibration. Then its associated 5-tuple is admissible. Viceversa, every admissible 5-tuple is the associated 5-tuple of a genus 2 fibration $f : X \to B$, and the surface X has invariants $\chi(\mathcal{O}_X) = \deg(V_1) + (b-1), K^2 = 2\deg V_1 + \deg \tau + 8(b-1)$. Two relatively minimal genus 2 fibration having the same associated 5-tuple are isomorphic.*

We can prove a very similar statement for a genus 3 fibrations f with non hyperelliptic general fibre, under the assumption that every fibre of f is 2-connected.

3. Applications

The first application of theorem 1 is a short proof of the following theorem (already proved by Bombieri ([Bom]) using Ogg's list of genus 2 fibres (cf. [Ogg])).

Theorem 2. *Let S be a Godeaux surface, and let $f : S \to \mathbb{P}^1$ be the fibration induced by the bicanonical pencil of S. Then the genus of the fibre can only be 3 or 4.*

We have an interesting application of theorem 1 to minimal surfaces of general type with $p_g = q = 1$. In this case $2 \le K_S^2 \le 9$ and the Albanese map is a morphism $f : S \to B$ where B is a smooth elliptic curve.

The case $K_S^2 = 2$ is completely described in [Cat1] where it is proved (among other things) that the moduli space is generically smooth, unirational of dimension 7.

The class of surfaces of general type with $K^2 = 3$, $p_g = q = 1$ is studied in [CC1], [CC2]. In [CC1] it is proved that for this class of surfaces the genus of the Albanese fibre is 2 or 3. The second case is completely classified in [CC2], where it is shown that the corresponding moduli space is generically smooth, unirational of dimension 5.

In [CC1] all surfaces with $p_g = q = 1$, $K^2 = 3$ and genus 2 of the Albanese fibre are described as double covers of $B^{(2)}$. It was conjectured there (see problem 5.5) that this family of surfaces should form an irreducible family of the moduli space. We can disprove this conjecture. More precisely (considering also the family in [CC2])

Theorem 3. *The family, in the moduli space of the minimal surfaces of general type, corresponding to the surfaces S with $p_g(S) = q(S) = 1$, $K_S^2 = 3$ has at least 4 connected components and at most 5 irreducible components, all of dimension 5.*

References

[Bom] E. Bombieri, unpublished manuscript.

[Cat1] F. Catanese, *On a class of surfaces of general type*, in Algebraic surfaces, 269–284, Fondazione C.I.M.E., Liguori Editore, Napoli 1981.

[CC1] F. Catanese, C. Ciliberto, *Surfaces with $p_g = q = 1$*, in Problems in the theory of surfaces and their classification (Cortona, 1988), 49–79, Sympos. Math., XXXII, Academic Press, London, 1991.

[CC2] F. Catanese, C. Ciliberto, *Symmetric products of elliptic curves and surfaces of general type with $p_g = q = 1$*, J. Algebraic Geom. **2** (1993), no. 3, 389–411.

[Fuj1] T. Fujita, *On Kähler fiber spaces over curves*, J. Math. Soc. Japan **30** (1978), no. 4, 779–794.

[Fuj2] T. Fujita, *The sheaf of relative canonical forms of a Kähler fiber space over a curve*, Proc. Japan Acad. Ser. A Math. Sci. **54** (1978), no. 7, 183–184.

[Hor] E. Horikawa, *On algebraic surfaces with pencils of curves of genus* 2, in Complex analysis and algebraic geometry, 79–90, Iwanami Shoten, Tokyo, 1977.

[Ogg] A.P. Ogg, *On pencils of curves of genus two*, Topology **5** (1966), 355–362.

[Xia] G. Xiao, *Surfaces fibrées en courbes de genre deux*, Lecture Notes in Mathematics 1137, Springer-Verlag, Berlin, 1985.

Numerical Godeaux surfaces with an involution

JongHae Keum

A minimal surface of general type with $p_g = 0$ and $K^2 = 1$ is called a numerical Godeaux surface, or simply a Godeaux surface. A joint work with Y. Lee [2] describes all possible fixed loci of an involution acting on a numerical Godeaux surface, under an assumption that the bicanonical system has no base components. Recently M. Mendes Lopes, R. Pardini [3] have proved the same result without the assumption.

Let X be a numerical Godeaux surface and σ be an involution acting on it. Its fixed locus consists of 5 isolated points, a curve l with $K_X l = 1$, and at most $g(l) + 2$ nodal curves. The genus $g(l)$ can take values 0, 1 and 2.

Let h denote the number of nodal curves. All known examples of Godeaux surfaces have an involution, and the corresponding $(g(l), h)$ is as follows:

a classical Godeaux surface from D_{10}-invariant quintic, Beauville's example, Barlow surface, and Craighero-Gattazzo-Dolgachev-Werner surface have $(0, 0)$*; Werner's example with Tors* $= \mathbb{Z}/2$, $(1, 1)$*; Stagnaro's example,* $(1, 2)$*; Oort-Peters' example,* $(1, 3)$.

In [2], two families of Godeaux surfaces with Tors $= \mathbb{Z}/4$ were constructed via canonical ring method due to M. Reid. These have involutions with $(g(l), h) = (1, 0)$, $(2, 0)$, respectively.

In this talk, I give an improvement as follows:

Theorem 1. *If $g(l) = 2$, then $h = 0$.*

Sketch of the proof of Theorem 1. If $g(l) = 2$, then the quotient surface X/σ is birational to an Enriques surface. This was one of the result presented by C. Ciliberto and A. Calabri [1] during this workshop. Let $W \to X/\sigma$ be a resolution of the five nodes. Then the branch $B \subset W$ is of the form $B = B_0 + N_1 + \cdots + N_5$, where N_i are nodal curves coming from the resolution. From the double covering formulas, we see that B_0 is a smooth curve of genus 2 with $B_0^2 = 2$. We also see that B_0 is disjoint from the exceptional curves on W which are to be blown down to an Enriques surface W'. On W', the branch consists of a genus 2 curve and 5 nodal curves. This means that no components other than l arise by the double covering process. □

I also suggest a way of constructing examples of Godeaux surfaces as double Enriques surfaces, whose covering involutions have $(g(l), h) = (0,1)$, $(0,2)$, the only missing cases.

References

[1] A. Calabri, C. Ciliberto, M. Mendes Lopes, *On the classification of numerical Godeaux surfaces with an involution*, in this report.

[2] J. Keum, Y. Lee, *Fixed locus of an involution acting on a Godeaux surface*, Math. Proc. Camb. Phil. Soc. **129** (2000), 205-216.

[3] M. Mendes Lopes, R. Pardini, *A new family of surfaces with $p_g = 0$ and $K^2 = 3$*, math. AG/0304181 (2003).

On fibred rational surfaces

Kazuhiro Konno

(joint work with Shinya Kitagawa)

Let X be a non-singular projective surface with $p_g = q = 0$ and $f : X \to \mathbb{P}^1$ a relatively minimal fibration of curves of genus $g \geq 2$. We denote by F a general fibre of f. Then $K_X + F$ is nef and the restriction map $H^0(X, K_X + F) \to H^0(F, \omega_F)$ is an isomorphism, because $p_g = q = 0$. In particular, $h^0(X, K_X+F) = g$. If $(K_X + F)^2 < 2g - 2$, then X is automatically a rational surface. Assume that the rational map defined by $|K_X + F|$ is generically finite onto the image W. Then,

Theorem 1. *$|K_X+F|$ is free from base points if $(K_X+F)^2 \leq 2g-4$. Furthermore, the ring $\oplus_{n\geq 0} H^0(X, n(K_X + F))$ is generated in degree 1 if $(K_X + F)^2 \leq 2g - 5$.*

Such an analysis is carried out by passing through the reduction (Y, G) obtained from (X, F) by blowing down all the (-1)-curves E satisfying $(K_X + F)E = 0$, where G is the image of F by the natural map $\mu : X \to Y$. The original fibration f is obtained from a pencil $\Lambda_f \subset |G|$ by blowing up the base points.

When X is a rational surface which is not $\mathbb{P}^2$, we can find a base point free pencil $|D|$ of rational curves on Y such that $c = (K_Y + G)D$ is minimal among such pencils. Then going down further to its #-minimal model $(Y^\#, G^\#)$, we get

$$(K_X + F)^2 = \frac{2c}{c+1}(g - c - 1) + \frac{1}{c+1}\sum_{i=1}^{N}(c + 1 - m_i)(m_i - 1),$$

where the m_i denotes the multiplicity of a singular point of $G^{\#}$, $m_i \leq c/2+1$. Furthermore, we can show the following by Serrano's theorem [5]:

Theorem 2. *Assume that $c \geq 2$ and $G^2 > (c+2)^2$. Then every morphism $\phi : G \to \mathbb{P}^1$ of degree at most $c+2$ can be extended to a morphism $\tilde{\phi} : Y \to \mathbb{P}^1$. Furthermore,*

(1) $\mathrm{gon}(F) = c+2$, *and*

(2) *the number of g^1_{c+2}'s on G is finite. In particular,* $\mathrm{Cliff}(F) = c$.

We use these results to study the Mordell-Weil lattice $\mathrm{MWL}(f)$ of f. Recall that the Mordell-Weil lattice is the group of sections of f endowed with a symmetric bilinear form coming from the intersection pairing on X. Put $r = \mathrm{rank}\ \mathrm{MWL}(f)$. Then Shioda [6] shows

$$r = \rho(X) - 2 - \sum_{P \in \mathbb{P}^1} (v_P - 1),$$

where $\rho(X)$ denotes the Picard number and v_P the number of irreducible components of the fibre $f^{-1}(P)$. In particular, we have $r = \rho(X) - 2$ if f has irreducible fibres only.

$\mathrm{MWL}(f)$ of maximal rank for fibred rational surfaces is determined so far by Saito-Sakakibara when f is hyperelliptic [3], by Saito-Nguyen Khac when f is of Clifford index one [4] and by Kitagawa when f is bi-elliptic [1]. As to the general fibrations of Clifford index two, we have the following:

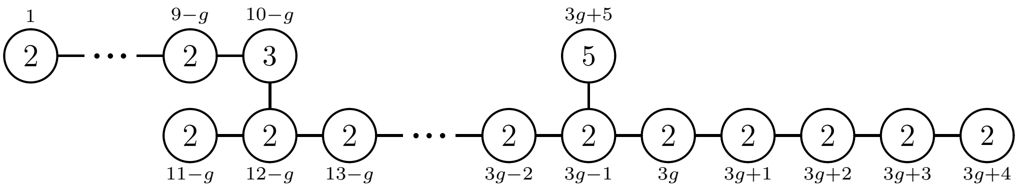
Theorem 3. *Let X be a non-singular rational surface, $f : X \to \mathbb{P}^1$ a relatively minimal fibration of genus g and of Clifford index 2. Let r be the Mordell-Weil rank of f.*

(1) *If $5 \leq g \leq 10$, then $r \leq 3g+5$.*

(2) *If $g \geq 11$, then $r \leq 3g + 8 - (g+\epsilon)/3$, where ϵ is the smallest non-negative integer with $g + \epsilon \equiv 0$ modulo 3.*

Assume that r attains the maximum. Then all the fibres of f are irreducible and the reduction Y is obtained as the image of $\Phi_{|K_X+F|}$. Furthermore, Y is a del Pezzo surface and $\Lambda_f \subset |-2K_Y|$ when $5 \leq g \leq 10$; it is a Hirzebruch surface blown up ϵ points and Λ_f comes from a linear system of quadruple sections when $g \geq 11$.

We can completely determine $\mathrm{MWL}(f)$ when the rank is maximum. For example, when $5 \leq g \leq 10$ and Y is obtained from $\mathbb{P}^2$ by blowing up $10-g$ points in general position, we get the following Dynkin diagram:



where the numbers in circles are self-pairing numbers of elements of a suitably fixed basis whose numbering is given near the circles. Therefore, it is an odd unimodular lattice of rank $3g+5$.

For $g \geq 11$, the maximal $\mathrm{MWL}(f)$ depends not only on g but also on ϵ and is much more complicated. We have four different types when $\epsilon = 0$, two types for each when $\epsilon = 1, 2$. The most interesting phenomena can be observed when $\epsilon = 0$, because the degree d of the Hirzebruch surface Y is an invariant of the fibration in this case. The parity of the lattice is the same as that of $g - d + 1$ and the structure of $\mathrm{MWL}(f)$ depends on the combination of g mod 4 and the parity of d. In particular, even and odd lattices both occur for a fixed g. See [2] for the detail.

References

[1] S. Kitagawa, *On Mordell-Weil lattices of bielliptic fibrations on rational surfaces*, to appear in J. Math. Soc. Japan (2004).

[2] S. Kitagawa and K. Konno, *Fibred rational surfaces with extremal Mordell-Weil lattices*, preprint, 2004.

[3] M.-H. Saito and K.-Sakakibara, *On Mordell-Weil lattices of higher genus fibrations on rational surfaces*, J. Math. Kyoto Univ. **34** (1994), 859–871.

[4] M.-H. Saito and V.-Nguyen Khac, *On Mordell-Weil lattices for nonhyperelliptic fibrations of surfaces with zero geometric genus and irregularity*, Izv. Ross. Akad. Nauk Ser. Mat. **66** (2002), 137–154.

[5] F. Serrano, *Extension of morphisms defined on a divisor*, Math. Ann. **277** (1987), 395–413.

[6] T. Shioda, *Mordell-Weil lattices for higher genus fibration over a curve*, in New trends in algebraic geometry (Warwick, 1996), London Math. Soc. Lecture Note Ser. 264, Cambridge Univ. Press 1999, 359–373.

The bicanonical map of surfaces of general type with $p_g = 0$ and $K^2 = 6$

Margarida Mendes Lopes
(joint work with Rita Pardini)

Many examples of complex surfaces of general type with $p_g = 0$ are known, but a detailed classification is still lacking, despite much progress in the theory of algebraic surfaces. Surfaces of general type are often studied using properties of their canonical curves. If a surface has $p_g = 0$, then there are of course no such curves, and it is natural to look instead at the bicanonical system, which is not empty.

Let S be a minimal surface of general type with $p_g = 0$. It is well known that $1 \leq K_S^2 \leq 9$. By a theorem of Xiao Gang [12], for $K_S^2 \geq 2$ the image of the bicanonical map of S is a surface Σ and, by Reider's theorem [11], the bicanonical map φ is a morphism if $K_S^2 \geq 5$.

Assume that $K_S^2 \geq 3$. Since $h^0(S, 2K_S) = K_S^2 + 1$, the bicanonical image of S is a surface of degree $m \geq K_S^2 - 1$ in $\mathbb{P}^{K_S^2}$. If, in addition, φ is a morphism (so, in particular, if $K_S^2 \geq 5$), one has $dm = (2K_S)^2 = 4K_S^2$, where d is the degree of φ. It is known that, if $K_S^2 \geq 3$ and φ is a morphism, then $d \leq 4$ [4]. Furthermore if

$K_S^2 = 9$, φ is birational [3], whilst if $K_S^2 = 7, 8$ the degree of φ is at most 2 and this bound is effective [6, 7, 9].

In the case $K_S^2 = 6$ one has the following numerical possibilities for the pair (d, m): $(1, 24), (2, 12), (3, 8), (6, 4)$.

The latter possibility occurs and in fact it can be completely characterized. Such surfaces turn out to be Burniat surfaces (see [2, 10]). More precisely one has the following:

Theorem 1. [5] *Let S be a minimal complex surface of general type such that $p_g(S) = 0$ and $K_S^2 = 6$ and let $\varphi\colon S \to \mathbb{P}^{K_S^2}$ the bicanonical map of S. Then $\deg \varphi = 4$ if and only if S is a Burniat surface.*

In particular, K_S is ample.

Theorem 2. [5] *Smooth minimal surfaces of general type S with $K_S^2 = 6$, $p_g(S) = 0$ and bicanonical map of degree 4 form an unirational 4-dimensional irreducible connected component of the moduli space of surfaces of general type.*

In this talk we discuss the other possible cases of non birationality of the bicanonical map, i.e., degrees 2 and 3. The results are the following:

Theorem 3. *Let S be a minimal surface of general type with $p_g(S) = 0$ and $K_S^2 = 6$ for which the bicanonical map φ is not birational. Then the degree of φ is either 2 or 4 and the image of φ is a rational surface.*

Theorem 4. *Let S be a minimal surface of general type with $p_g(S) = 0$ and $K_S^2 = 6$ for which the bicanonical map φ has degree 2. Then there is a fibration $f\colon S \to \mathbb{P}^1$ such that the general fibre F of f is hyperelliptic of genus 3 and f has 4 or 5 double fibres. Furthermore the bicanonical involution of S induces the hyperelliptic involution on F.*

Idea of the proof of Theorem 3. It is necessary to exclude the possibility that $d = 3$ occurs. For $d = 3$ the bicanonical image would be a rational surface of degree 8 in $\mathbb{P}^6$. By using repeated adjunction (an idea which dates back to Enriques), such surfaces are studied and their geometry is used to show that $d = 3$ does not occur. For details see [8]. $\square$

Idea of the proof of Theorem 4. Let σ be the bicanonical involution. The quotient surface $T := S/\sigma$ is a rational surface whose only singularities are nodes (corresponding to the isolated fixed points of σ). Since the bicanonical map factors through T it is possible to show that T has exactly 10 nodes. The statement of the theorem is obtained by a careful analysis of the binary linear code associated to the nodes. For details see again [8]. $\square$

Remark. Note that Theorem 4 is not a mere list of possibilities because there are examples of both situations (see again [8]). G. Borrelli (see [1]) has obtained recently with different methods the same list of possibilities and a description of them as double planes.

Remark. It would be very interesting to describe the moduli space of the surfaces appearing in Theorem 4 and in particular to find whether these surfaces deform to surfaces with birational bicanonical map (no such example is known for $K_S^2 = 6$).

References

[1] G. Borrelli, *On the classification of surfaces of general type with non-birational bicanonical map and Du Val Double planes*, preprint, math.AG/0312351.

[2] P. Burniat, *Sur les surfaces de genre $P_{12} > 0$*, Ann. Mat. Pura Appl., IV Ser., **71** (1966), 1–24.

[3] I. Dolgachev, M. Mendes Lopes, R. Pardini, *Rational surfaces with many nodes*, Compositio Math. **132** (2002), 349–363.

[4] M. Mendes Lopes, *The degree of the generators of the canonical ring of surfaces of general type with $p_g = 0$*, Arch. Math. **69** (1997), 435–440.

[5] M. Mendes Lopes, R. Pardini, *A connected component of the moduli space of surfaces of general type with $p_g = 0$*, Topology **40** (5) (2001), 977–991.

[6] M. Mendes Lopes, R. Pardini, *The bicanonical map of surfaces with $p_g = 0$ and $K^2 \geq 7$*, Bull. London Math. Soc. **33** (2001), 265–274.

[7] M. Mendes Lopes, R. Pardini, *The bicanonical map of surfaces with $p_g = 0$ and $K^2 \geq 7$, II*, Bull. London Math. Soc. **35** (2003), no. 3, 337–343.

[8] M. Mendes Lopes, R. Pardini, *The classification of surfaces with $p_g = 0$, $K^2 = 6$ and non birational bicanonical map*, to appear in Math. Ann., math.AG/0301138.

[9] R. Pardini, *The classification of double planes of general type with $K^2 = 8$ and $p_g = 0$*, J. Algebra **259** (2003), no. 1, 95–118.

[10] C. Peters, *On certain examples of surfaces with $p_g = 0$ due to Burniat*, Nagoya Math. J. **166** (1977), 109–119.

[11] I. Reider, *Vector bundles of rank 2 and linear systems on algebraic surfaces*, Ann. of Math. **127** (1988), 309–316.

[12] G. Xiao, *Finitude de l'application canonique des surfaces de type général*, Bull. Soc. Math. France **113** (1985), 23–51.

A new family of surfaces with $p_g = 0$ and $K^2 = 3$

Margarida Mendes Lopes
(joint work with Rita Pardini)

The starting point of the subject of this talk is the following Theorem:

Theorem 1 (Xiao Gang, [5])**.** *Let S be a minimal complex surface of general type with $p_g(S) = 0$ such that the bicanonical map φ is not birational and let T be the bicanonical image. If T is not a rational surface, then T is birational to an Enriques surface and φ is a degree 2 morphism.*

Furthermore $K_S^2 = 3$ or $K_S^2 = 4$.

This theorem lists possibilities and a natural question is whether it is sharp.

Both J. Keum, [1], and D. Naie, [4], constructed examples of surfaces S with $p_g(S) = 0$ and $K_S^2 = 3$ or $K_S^2 = 4$ as double covers of nodal Enriques surfaces. For these surfaces the bicanonical map, although it factorizes through the covering map, has degree 4 and the bicanonical image is a rational surface.

In [2], it is shown that, in fact, if the bicanonical image of a surface S with $p_g(S) = 0$ is birationally an Enriques surface then, necessarily, $K_S^2 = 3$. So the case with $K_S^2 = 4$ of Theorem 1 does not occur. Furthermore it is shown that the minimal surfaces S with $p_g(S) = 0$ and $K^2 = 4$ having an involution σ such that S/σ is birational to an Enriques surface and such that the bicanonical map is composed with σ are precisely the Keum-Naie examples.

No example of a surface S with $p_g(S) = 0$ and $K_S^2 = 3$, with bicanonical image birational to an Enriques surface appears in the literature, and so the question is whether it can occur at all. It turns out such surfaces exist.

The subject of this talk is not only showing the existence of surfaces S with $p_g(S) = 0$ and $K_S^2 = 3$, with bicanonical image birational to an Enriques surface, but also explaining an explicit construction of all such surfaces. This explicit construction enables us to show that the corresponding subset of the moduli space of surfaces of general type is irreducible and uniruled of dimension 6. Since the closure of this subset contains the Keum-Naie surfaces, whose fundamental group is isomorphic to $\mathbb{Z}_2^2 \times \mathbb{Z}_4$ (cf. [4]), also the fundamental group of all these surfaces is $\mathbb{Z}_2^2 \times \mathbb{Z}_4$.

The description of these surfaces is based on a very detailed study of the normalization of their bicanonical images. These are polarized Enriques surfaces of degree 6 with 7 nodes, satisfying some additional conditions.

For the proofs and details we refer to [3].

References

[1] J. H. Keum, *Some new surfaces of general type with $p_g = 0$*, (preprint 1988).

[2] M. Mendes Lopes, R. Pardini, *Enriques surfaces with eight nodes*, Math. Z. **241** (2002), 673–683.

[3] M. Mendes Lopes, R. Pardini, *A new family of surfaces with $p_g = 0$ and $K^2 = 3$* , to appear in Annales de L'Ecole Normale Supérieure, math.AG/0304181.

[4] D. Naie, *Surfaces d'Enriques et une construction de surfaces de type général avec $p_g = 0$*, Math. Z. **215** (2) (1994), 269–280.

[5] G. Xiao, *Degree of the bicanonical map of a surface of general type*, Amer. J. of Math. **112** (1990), 713–737.

Kustin–Miller unprojections

Stavros Papadakis

Kustin-Miller unprojection is a method that constructs more complicated Gorenstein rings from simpler data. Geometrically it corresponds to the inverse of the classical method of projection. The first talk was about the scheme–theoretic foundations of the simplest type of Kustin–Miller unprojection called Type I, which is joint work with M. Reid [3], and algebraically corresponds to the unprojection of a codimension one ideal I of a Gorenstein ring R with the quotient R/I being Gorenstein. In addition, I gave examples and mentioned a method, essentially due to A. Kustin and M. Miller [1], which calculates type I unprojection in the relative setting using projective resolutions and maps between complexes.

The second talk was about Tom and Jerry. They are two families of codimension four Gorenstein rings defined by M. Reid and studied by me at [2], which are constructed as Type I unprojection and appear in a variety of examples coming from Algebraic Geometry. Moreover, I talked about Type II unprojection, which is work in progress, and constructs a codimension $n+2$ conjecturally Gorenstein ring, starting from a codimension n complete intersection containing a certain codimension $n+1$ subscheme.

References

[1] A. Kustin and M. Miller, *Constructing big Gorenstein ideals from small ones*, J. Algebra **85** (1983), 303–322.

[2] S. Papadakis, *Kustin-Miller unprojection with complexes*, J. of Algebraic Geometry **13** (2004), 249-268.

[3] S. Papadakis and M. Reid, *Kustin-Miller unprojection without complexes*, to appear in J. of Algebraic Geometry.

Surfaces in your backyard

Ulf Persson

How do you give elementary examples of surfaces? Hypersurfaces in $\mathbb{P}^3$ are obvious candidates, but of course they are far too restrictive to present a wide variety of phenomena. It is e.g. impossible to give an example of a so called honestly elliptic surface (i.e. $\kappa = 1$ in the Kodaira classification). A natural thing is to consider imposing singularities. Just imposing ordinary double points (or more generally simple-singularities i.e. A-D-E singularities) does not give you anything new, although it certainly gives you interesting ecxamples with high Picard numbers. The next step is to consider ordinary triple points, i.e. points whose resolutions give you smooth elliptic curves with self-intersection -3. It is an elementary but instructive exercise to present the following list of quintics with ordinary triple points

Theorem 1. *If Q is a quintic with k ordinary triple points then $0 \le k \le 5$ and its resolution $\tilde{Q}$ satisfies*

$k=0$, $c_1^2 = 5$, $\chi = 5$ *(minimal of general type)*

$k=1$, $c_1^2 = 2$, $\chi = 4$ *(minimal of general type, a double octic)*

$k=2$ $c_1^2 = -1$, $\chi = 3$ *(an elliptic surface blown up once)*

$k=3$ $c_1^2 = -4$, $\chi = 2$ *(a K-3 surface blown up four times)*

$k=4$ $c_1^2 = -7$, $\chi = 1$ *(a rational surface)*

$k=5$ $c_1^2 = -10$, $\chi = 0$ *(a ruled surface over an elliptic curve blown up ten times)*

The proof is completely elementary. The interesting feature is the way those surfaces are geometrically realised. To take the example of $k = 2$. The line joining the two triple points becomes exceptional, and the elliptic fibration is given by the pencil of planes through it, intersecting the quintic residually in quartics with two double points. Those planes incidentally cut out the canonical divisors. In the case of $k = 3$ the canonical divisor consists of the plane through the three triple points, whose intersection is a triangle of lines and a circumscribed conic, all four easily seen to be exceptional. And finally the case of $k = 5$ the ruling consists of twisted cubics passing through the five triple points. By Bezout, any such twisted cubic having an additional intersection will be contained, and clearly through any six points, there is a twisted cubic. The degenerate fibers will be ten by choosing two points out of the five, defining a line and a residual conic through the remaining three. This distinction between the reducible components allow a canonical minimal model, which turns out to be a ruled surface over an elliptic curve defined by a stable rank-two bundle.

Now with my co-workers Endrass and Stevens I considered whether a similar classification can be effected for degree six, and the surprising answer is yes! However, the situation becomes more complicated. For one thing one can now no longer in general choose the locations of the triple points arbitrarily (there will be two many conditions). E.g. there will be no examples of eight generic triple points, but if the triple points happen to form the base points of a net of quadrics one can write down a simple example $C(Q_1, Q_2, Q_3)$ where Q_i span the net, and C is a plane cubic. This will actually be an honestly elliptic surface fibered over an elliptic curve (given by $C = 0$). Other special choices of eight points will also yield examples. In the case of nine triple points we get examples of non-minimal K-3 surfaces, as well as non-minimal fake K-3 surfces, namely honestly elliptic surfaces gotten from elliptic K-3 surfaces through logarithmic transforms. One may also find rational sextics with ten triple points, but ten is the upper limit.

For the complete classification I refer to the paper below. Let me only note that a typical construction is to consider a linear space made up by highly reducible, often not even reduced, hypersurfaces, such that the base points are of multiplicity three. (As a simple example consider a quintic Qu with five nodes on a conic $C = H \cap Q$, where H is a plane and Q a quadric. Then consider the generic member of the pencil spanned by HQu and Q^3).

One may wonder where to go from here? One may note that we prove that for degree seven or higher only minimal surfaces of general type occur in this way. Thus one should either consider other elementary constructions of low degree, like complete intersections in $\mathbb{P}^4, \mathbb{P}^5$ and maybe $\mathbb{P}^6$. The same thing for multi-projective spaces. In short, I suspect that there will be no more than perhaps a dozen different cases, similar to the ones I have refered to above. To be more specific, try to do a similar analysis for hypersurfaces of low degree in $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$. The case of tri-degree $(3, 3, 3)$ is analogous to the case of quintics, (but of course more involved). It turns out that its chern-invariants are given by $c_1^2 = 18, \chi = 9$.

So I would like to point this out by describing an analogy to the Godeaux quotient, which although elementary, has never been written down and published to my knowledge[1]. The key point is an action of $\mathbb{Z}_9$ on $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ inducing an action on the monomials $x_i y_j z_k$ involving an amalgation of the cyclic permutation of the co-ordinates and the action of a primitive 9-th root of unity. More precisely letting a generator of $\mathbb{Z}_9$ act accordingly

$$(x, y, z) \mapsto (\rho z, \rho x, \rho y)$$

It is easy to find the fixed points of the actions, and just like in the quintic case, avoid those by a judicious inclusion of certain extreme monomials. Once we have a fixed point free action the quotient will have $c_1^2 = 2, \chi = 1$. As the quotient is regular, we conclude that $p_g = 0$.

Finally instead of considering just triple points, one may take into account four-tuple points, or other more subtle singularities, one thinks of elliptic singularities with $E^2 = -2, -1$. Those two types are easily exhibited on double covers, by considering four-tuple points or so called infinitely close triple points.

All of those obviously are directed to the main question

Question. *Is it possible to classify all surfaces of small invariants?*

One first attempt would be to classify all such surfaces which can be deformed into double coverings, especially double planes.

References

[1] S. Endrass, U. Persson, J. Stevens, *Surfaces with Triple Points*, Journal of Algebraic Geometry **12** (2003), 367–404.

Extrasymmetric matrices and surfaces with $p_g = 4$ and $K^2 = 6$

Roberto Pignatelli
(joint work with Ingrid Claudia Bauer and Fabrizio Catanese)

Minimal surfaces with $p_g = 4$ have been studied by several mathematicians since the publication of the famous book of Enriques [Enr]. By the standard inequalities of Noether and Bogomolov-Miyaoka-Yau, for these surfaces it holds $4 \leq K^2 \leq 45$.

The case $K^2 = 4$ is completely described in [Hor2]. All these surfaces are double covers of an irreducible quadric in $\mathbb{P}^3$. Their moduli space is generically smooth, unirational, of dimension 42; its singular locus has codimension 1, and it is exactly the locus corresponding to the double covers of the quadric cone.

In [Hor1] (see also [Rei2], [Gri]) the case $K^2 = 5$ is completely described: the canonical map is either a birational morphism to a quintic in $\mathbb{P}^3$, or a rational map of degree 2 onto an irreducible quadric. Their moduli space has two irreducible unirational components of dimension 40 whose general point corresponds to surfaces with canonical image respectively a quintic or a smooth quadric. The

[1] I thought of it some twenty years ago, and may have circulated it around privately.

surfaces whose canonical image is a quadric cone form a 39-dimensional subvariety of this moduli space, the intersection of the two irreducible components.

The case $K^2 = 6$ is the first case not completely solved. In [Hor3] Horikawa listed all possibilities for the canonical map, dividing these surfaces in 11 classes (and therefore their moduli space in 11 strata). He proved that each of these cases occurs, and studying the local deformations of these surfaces (to understand how these strata can 'glue'), Horikawa proved that their moduli space has 4 irreducible components (one of dimension 39, the other three of dimension 38), and at most 3 connected components.

More precisely, Horikawa named the 11 classes as I_a, I_b, II, III_a, III_b, IV_{a_1}, IV_{a_2}, IV_{b_1}, IV_{b_2}, V_1, V_2 (see [Hor3] for precise definitions of each class). According to Horikawa's notation we define

Definition. Let A and B be two of the above introduced classes. If we write "$A \to B$", it means that there is a flat family with base a small disc $\Delta_\varepsilon \subset \mathbb{C}$ whose central fibre is of type B and whose general fibre is of type A.

With this notation Horikawa summarized its results in the following picture

$$\begin{tikzcd}
 & & & & III_a \arrow[dl] \arrow[d] & \\
 & IV_{a_1} \arrow[dl] \arrow[d] & I_a \arrow[dl] \arrow[dd] & V_1 \arrow[d] & III_b \arrow[dl] & II \\
IV_{a_2} \arrow[d] & IV_{b_1} & & V_2 & & \\
IV_{b_2} & & I_b & & &
\end{tikzcd}$$

He could disprove many other degenerations, but he could neither prove nor disprove the specializations $II \to III_b$, $II \to V$ and $I_a \to V$; we have shown that the degeneration $II \to III_b$ occurs.

Definition. A minimal surfaces of general type with $p_g = 4$ and $K^2 = 6$ is of type II if the canonical map has degree 3.

Horikawa proved that in this case the canonical image is a quadric cone.

Surfaces of type III_b are described by Horikawa as follows:

Theorem (5.2 in [Hor3]**).** *Let S be a surface of type III_b. Then S is birationally equivalent to a double covering of $\mathbb{F}_2$ whose branch locus B consists of the 0-section Δ_0 and $B_0 \in |7\Delta_0 + 14\Gamma|$ which has a quadruple point at $x \in \Gamma$ and a 2-fold triple point at $y \in \Gamma$ on a fibre Γ, with x and y being possibly infinitely near.*

The canonical ring of these surfaces is very complicated: it is a quotient of a polynomial ring of big (at least 6, maybe more) codimension. We do not know how to investigate the flat deformations of rings of high codimension. We look then for a 'bigger' and easier ring, a ring containing the canonical ring and of smaller codimension.

By standard computations one can show that the canonical system of S is $|2L| + Z$ where L is the genus 3 pencil pull-back of the ruling of $\mathbb{F}_2$, and Z is a fundamental cycle. Therefore, even if K_S is not 2-divisible in the Picard group, it can be divided by 2 when considered only as a Weil divisor on the canonical model.

Definition. Let S be a surface of type III_b, let Z be the fixed part of its canonical system, and let δ be a generator of $H^0(Z)$.

Let R be the graded ring whose homogeneous components are the spaces $R_d := H^0(dL + \lfloor \frac{d}{2} Z \rfloor)$, $d \in \mathbb{N}$, with product defined on the homogeneous elements as $ab = a \otimes b$ or $a \otimes b \otimes \delta$ according if the product of the degrees of a and b is even or odd.

Note that enlarging the ring 'restricts' the possible deformations. In fact, if the canonical rings induce, given a flat family of surfaces, a flat family of rings, the same does not hold for these 'half-canonical' rings, since the 2-divisibility of the canonical divisor (as a Weil divisor on the canonical model) is not necessarily preserved by a deformation.

As proved in [MP] (where these surfaces are studied in detail) the canonical system of a surface of type II, can be written again as $2L+Z$ with L genus 3 pencils and Z fundamental cycle. It is then natural to expect, if a family "$II \to III_b$" exists, that this family preserves the genus 3 pencils and the 'half-canonical' rings.

Theorem 1. *We have* $R \cong \mathbb{C}[x_0, x_1, y, z, w, v, u]/I$ *with* $\deg(x_0, x_1, y, z, w, v, u) = (1, 1, 2, 3, 4, 5, 6)$*, where* I *has codimension* 4*, generated by* 9 *equations yoked by* 16 *syzygies; the* 9 *generators of* I *are homogeneous polynomial of respective degrees* $(4, 5, 6, 7, 8, 9, 10, 11, 12)$.

Miles Reid and Duncan Dicks introduced in [Rei1] (see also [Rei2], [Rei3], [BCP]) the 'extrasymmetric format', for some Gorenstein rings of codimension 4 with 9 relations and 16 syzygies.

Roughly speaking, they noticed that the ideal generated by the pfaffians of order 4 of a 6×6 skewsymmetric matrix is, if the matrix has some further symmetry (it is 'extrasymmetric') of codimension 4 with 9 generators and 16 syzygies. This format is flexible, *i.e.* every deformation of the matrix preserving the symmetries induces a flat deformation of the ideal. This property allowed us to prove our main result.

Theorem 2. *Let* $(x_0, x_1, y, z, w, v, u)$ *variables of degrees* $(1, 1, 2, 3, 4, 5, 6)$*, Let* M *be the* 6×6 *skewsymmetric matrix*

$$M = \begin{pmatrix} 0 & t & z & v & y & x_1 \\ & 0 & w & u & P_3 & y \\ & & 0 & P_9 & u & v \\ & & & 0 & wP_4 & zP_4 \\ & & & & 0 & tP_4 \\ -sym & & & & & 0 \end{pmatrix}.$$

where the P_i*'s are homogeneous of degree* i *in the above introduced variables and* t *is the parameter on a small disc* $\Delta_\varepsilon \subset \mathbb{C}$*.*

For general choice of P_3, P_4 *and* P_9 *the* 4×4 *pfaffians of* M *define a variety* $X \subset \Delta_\varepsilon \times \mathbb{P}(1, 1, 2, 3, 4, 5, 6)$ *whose projection on* Δ_ε *is flat, with central fibre a surface of type* III_b *and with general fibre a surface of type* II*.*

Sketch of the proof of Theorem 2. The flatness of the above family (for general entries) follows directly from the flexibility of the format. One can check that for general choice of the polynomials P_i and for t small the above equations define a surface with only rational double points as singularities: the invariants can be easily computed.

Note that the pfaffians Pf_{1235} and Pf_{1236} are of the form $tu - \cdots$ and $tv - \cdots$, and that the pfaffian Pf_{1256} can, for general choice of P_4, be written as $t^2w - \ldots$. Therefore, for $t \neq 0$, we can 'eliminate' the variables u, v, w, and $R \cong \mathbb{C}[x_0, x_1, y, z]/J$ for some ideal J: a straightforward computation shows that J is a principal ideal generated by the equation obtained by Pf_{1234} after 'eliminating' u, v, w using Pf_{1235}, Pf_{1236} and Pf_{1256}.

We get then an hypersurface of degree 9 in $\mathbb{P}(1, 1, 2, 3)$, whose canonical system is induced by $\mathcal{O}(2)$: since for general entries of M the coefficient of the monomial z^3 in its equation does not vanish, we see that the canonical map has degree 3 (and image $\mathbb{P}(1, 1, 2)$, a quadric cone). This shows that the surface is of type II.

If $t = 0$, the canonical map is given again by the projection on $\mathbb{P}(1, 1, 2)$, but the surface meets the center of the projection in a point (if $P_4 = w + \cdots$, the point (0,0,0,0,1,0,1)), therefore the projection has only degree 2; one can easily check that the branch locus has the behavior described by Horikawa. □

As a corollary, we can improve Horikawa's bound on the deformation types

Corollary. *The number of deformation types of minimal surfaces of general type with* $p_g = 4$ *and* $K^2 = 6$ *is at most* 2*.*

References

[BCP] I. Bauer, F. Catanese, R. Pignatelli, *Canonical rings of surfaces whose canonical system has base points*, in Complex geometry (Göttingen, 2000), Springer 2002, 37–72.

[Enr] F. Enriques, *Le superficie algebriche*, Zanichelli, Bologna, 1949.

[Gri] E.E. Griffin, II, *Families of quintic surfaces and curves*, Compositio Math. **55** (1985), no. 1, 33–62.

[Hor1] E. Horikawa, *On deformations of quintic surfaces*, Proc. Japan Acad. **49** (1973), 377–379.

[Hor2] E. Horikawa, *Algebraic surfaces of general type with small c_1^2. I*, Ann. of Math. (2) **104** (1976), no. 2, 357–387.

[Hor3] E. Horikawa, *Algebraic surfaces of general type with small c_1^2. III*, Invent. Math. **47** (1978), no. 3, 209–248.

[MP] M. Mendes Lopes, R. Pardini, *Triple canonical surfaces of minimal degree*, Internat. J. Math. **11** (2000), no. 4, 553–578.

[Rei1] M. Reid, *Surfaces with $p_g = 3$, $K^2 = 4$ according to E. Horikawa and D. Dicks*, Proc. of Alg. Geometry mini Symposium, Tokyo Univ. Dec. 1989, 1–22.

[Rei2] M. Reid, *Infinitesimal view of extending a hyperplane section—deformation theory and computer algebra*, in Algebraic geometry (L'Aquila, 1988), Lecture Notes in Math. 1417, Springer 1990, 214–286.

[Rei3] M. Reid, *Graded rings and birational geometry*, Proc. of algebraic geometry symposium (Kinosaki, 2000), 1–72.

Surfaces of general type with $p_g = q = 1$, $K^2 = 8$ and bicanonical map of degree 2

Francesco Polizzi

In [Par03] R. Pardini classified the minimal surfaces S of general type with $p_g = q = 0$, $K_S^2 = 8$ and a rational involution, i.e. an involution $\sigma : S \longrightarrow S$ such that the quotient $T := S/\sigma$ is a rational surface. All the examples constructed by Pardini are *isogenous to a product*, i.e. there exist two smooth curves C, F and a finite group G acting faithfully on C, F and whose diagonal action is free on the product $C \times F$, in such a way that $S = (C \times F)/G$. Pardini's classification contains five families of such surfaces; in particular, four of them are irreducible components of the moduli space of surfaces with $p_g = q = 0$, $K_S^2 = 8$, and represent the surfaces with the above invariants and non- birational bicanonical map.

In this paper we deal with the irregular case, in fact we study the case $p_g = q = 1$, $K_S^2 = 8$. Surfaces with $p_g = q = 1$ are the minimal irregular surfaces of general type with the lowest geometric genus, therefore they are natural candidates to starting the investigation of irregular surfaces with $q = 1$ or, more generally, with an irrational pencil. However, such surfaces are still quite mysterious, and only a few families have been hitherto discovered. If S is a surface with $p_g = q = 1$, then $2 \leq K_S^2 \leq 9$; the case $K_S^2 = 2$ is studied in [Ca81], whereas [CaCi91] and [CaCi93] deal with the case $K_S^2 = 3$. For higher values of K_S^2 only some sporadic examples were so far known; see [Ca99], where a surface with $K_S^2 = 4$ and one with $K_S^2 = 5$ are constructed.

When $p_g = q = 1$, there are two basic tools that one can use in order to study the geometry of S: the *Albanese fibration* and the *paracanonical system*. First of all, $q = 1$ implies that the Albanese variety of S is an elliptic curve E, hence the Albanese map $\alpha : S \longrightarrow E$ is a connected fibration; we denote by F the general fibre of α and by $g = g(F)$ its genus. Let us fix a zero point $0 \in E$, and for any $t \in E$ let us write $K_S + t$ for the line bundle $K_S + F_t - F_0$. By Riemann-Roch and semicontinuity theorem we have $h^0(S,\ K_S + t) = 1$ for general $t \in E$, hence denoting by C_t the only element in the complete linear system $|K_S + t|$ we obtain a 1-dimensional algebraic family $\{K\} = \{C_t\}_{t \in E}$ parametrized by the elliptic curve E. We will call it the paracanonical system of S; according to [Be88], it is the irreducible component of the Hilbert scheme of curves on S algebraically equivalent to K_S which dominates E. The *index* $\iota = \iota(K)$ of the paracanonical system $\{K\}$ is the number of distinct curves of $\{K\}$ through a general point of S. The *paracanonical map* $\omega : S \longrightarrow E(\iota)$, where $E(\iota) := \mathrm{Sym}^{\iota} E$, is defined in the following way: if $x \in S$ is a general point, then $\omega(x) = t_1 + \cdots + t_\iota$, where $C_{t_1}, \ldots, C_{t_\iota}$ are the paracanonical curves containing x. The best result that one might obtain would be to classify the triples (K^2, g, ι) such that there exists a minimal surface of general type S with $p_g = q = 1$ and these invariants. Since by the results of Gieseker the moduli space $\mathcal{M}_{\chi,\, K^2}$ of surfaces of general type with fixed $\chi(\mathcal{O}_S),\ K_S^2$ is a quasiprojective variety, it turns out that there exist only finitely many such triples, but a complete classification is still missing.

By the results of [Re88], [Fr91] and [CaCi91] it follows that the bicanonical system $|2K_S|$ of a minimal surface of general type with $p_g = q = 1$ is base-point free, whence the bicanonical map $\phi := \phi_{|2K|} : S \longrightarrow \mathbb{P}^{K_S^2}$ of S is a morphism. Moreover such a morphism is generically finite by [Xi85], so $\phi(S)$ is a surface Σ. We will say that a surface S contains a *genus 2 pencil* if there is a morphism $f : S \longrightarrow B$, where B is a smooth curve and the general fibre Φ of f is a smooth curve of genus 2. Notice that in this case the bicanonical map ϕ of S is not birational, since $|2K_S|$ cuts out on Φ a subseries of the bicanonical series of Φ which is composed with the hyperelliptic involution. In this case we say that S presents the *standard case* for the non-birationality of the bicanonical map; otherwise, namely if ϕ is not birational but S does not contain any genus 2 pencils, we say that S presents the *non-standard* case. By the results of Bombieri (later improved by Reider, see [Bo73] and [Re88]) it follows that, if $K_S^2 \geq 10$ and the bicanonical map is not birational, then S contains a genus 2 pencil. Whence there exist only finitely many families of surfaces of general type presenting the non-standard case, and one would like to classify all of them; however, this problem is still open, although many examples are known. In the paper [Xi90] G. Xiao gave two list of possibilities for the bicanonical image of such a surface; later on several authors investigated about their real occurrence. For more details about this argument, we refer the reader to the paper [Ci97].

No examples of surfaces with $p_g = q = 1$ and presenting the non-standard case were hitherto known; if S is such a surface and $K_S^2 \geq 5$, then a result of Xiao ([see Xi90, Proposition 5]) implies that the degree of ϕ is either 2 or 4. In this work we describe the surfaces of general type with $p_g = q = 1,\ K_S^2 = 8$ and such that the degree of ϕ is 2. It will turn out that they belong to three distinct families, which provide as well the first known examples of surfaces which such invariants. None of these surfaces contains a genus 2 pencil, thus they are three substantially new pieces in the classification of surfaces presenting the non-standard case.

What we show is that, as in the case $p_g = q = 0$, the surfaces with $p_g = q = 1$, $K_S^2 = 8$ and bicanonical map of degree 2 are isogenous to a product. More precisely, our result is the following:

Theorem 1. *Let S be a minimal surface of general type with $p_g = q = 1$, $K_S^2 = 8$ and such that its bicanonical map has degree 2. Then S is a quotient of type $S = (C \times F)/G$, where C, F are smooth curves and G is a finite group acting faithfully on C, F and freely on $C \times F$. Moreover C is a curve of genus 3 which is both hyperelliptic and bielliptic, $E := C/G$ is an elliptic curve isomorphic to the Albanese variety of S and $F/G \cong \mathbb{P}^1$. The bicanonical map ϕ of S factors through the involution σ of S induced by the involution $\tau \times id$ on $C \times F$, where τ is the hyperelliptic involution of C. The occurrences for $g(F)$ and G are the following three:*

I. $g(F) = 3,\ G \cong \mathbb{Z}_2 \times \mathbb{Z}_2$;
II. $g(F) = 4,\ G \cong S_3$;
III. $g(F) = 5,\ G \cong D_4$.

The curve F is hyperelliptic in case I, whereas it is not hyperelliptic in cases II and III.

Surfaces of type I, II, III do exist and they form three generically smooth, irreducible component $\mathcal{S}_I$, $\mathcal{S}_{II}$, $\mathcal{S}_{III}$ of the moduli space $\mathcal{M}$ of surfaces with $p_g = q = 1$, $K_S^2 = 8$, whose respective dimensions are:

$$dim\ \mathcal{S}_I = 5, \quad dim\ \mathcal{S}_{II} = 4, \quad dim\ \mathcal{S}_{III} = 4.$$

The proof of Theorem 1 is somewhat involved as it requires the understanding of many different techniques.

Sketch of the proof of Theorem 1. **Step 1**. We analyze the *bicanonical involution* σ of S, following [Xi90] and [CM02]. It turns out that σ has 12 isolated fixed points and that the divisorial fixed locus of σ is contained in fibres of the Albanese pencil.

Step 2. Using the results obtained in Step 1 we prove that if S is a minimal surface of general type with $p_g = q = 1$, $K_S^2 = 8$ and bicanonical map of degree 2, then S contains a rational pencil of hyperelliptic curves of genus 3 with six double fibres. This in turn implies, by the results of Serrano contained in [Se90] and [Se96], that S is isogenous to a product, i.e. $S = (C \times F)/G$. We show moreover that there are at most three families of such surfaces, and we describe them.

Step 3. We show that the three families described in Step 2 actually exist, by constructing the two curves C, F and by exhibiting explicitly the actions of G on them.

Step 4. We study the moduli space of the surfaces S constructed in Step 3. This is not difficult because the group G acts separately on C and F, hence the Kuranishi family of S turns out to be smooth. □

References

[Be88] A. Beauville, *Annulation du H^1 et systèmes paracanoniques sur le surfaces*, J. reine angew. Math. **388** (1988), 149–157.

[Bo73] E. Bombieri, *Canonical models of surfaces of general type*, Publ. IHES **42** (1973), 171–219.

[BPV84] W. Barth, C. Peters, A. Van de Ven, *Compact Complex Surfaces*, Springer-Verlag 1984.

[Ca81] F. Catanese, *On a class of surfaces of general type*, in *Algebraic Surfaces*, CIME, Liguori 1981, 269–284.

[Ca99] F. Catanese, *Singular bidouble covers and the construction of interesting algebraic surfaces*, Contemporary Mathematics **241** (1999), 97–119.

[CaCi91] F. Catanese and C. Ciliberto, *Surfaces with $p_g = q = 1$*, Symposia Math. **32** (1991), 49–79.

[CaCi93] F. Catanese and C. Ciliberto, *Symmetric product of elliptic curves and surfaces of general type with $p_g = q = 1$*, J. of Algebraic Geometry **2** (1993), 389–411.

[Ci97] C. Ciliberto, *The bicanonical map for surfaces of general type*, Proc. of Symp. in Pure Mathematics 62.1, 1997, 57–84.

[CM02] C. Ciliberto, M. Mendes Lopes, *On surfaces with $p_g = q = 2$ and non-birational bicanonical map*, Adv. Geom. **2** (2002), no. 3, 281-300.

[Fr91] P. Francia, *On the base points of the bicanonical system*, Symposia Math. 32, 1991, 141–150.

[Par03] R. Pardini, *The classification of double planes of general type with $K^2 = 8$ and $p_g = 0$*, Journal of Algebra **259** (2003), no. 3, 95–118.

[Re88] I. Reider, *Vector bundles of rank 2 and linear systems on algebraic surfaces*, Ann. of Math. **127** (1988), 309–316.

[Se90] F. Serrano, *Fibrations on algebraic surfaces*, in A. Lanteri, M. Palleschi, D. C. Struppa (eds.), *Geometry of Complex Projective Varieties (Cetraro 1990)*, Mediterranean Press 1993, 291–300.

[Se96] F. Serrano, *Isotrivial fibred surfaces*, Annali di Matematica pura e applicata **171** (1996), 63–81.

[Xi85] G. Xiao, *Finitude de l' application bicanonique des surfaces de type générale*, Boll. Soc. Math. de France **113** (1985), 23–51.

[Xi90] G. Xiao, *Degree of the bicanonical map of a surface of general type*, Amer. J. of Math. **112** (5) (1990), 309–316.

On numerical Godeaux surfaces constructed as double planes

Caryn Werner

Let S be a minimal surface of general type with $p_g = q = 0, K_S^2 = 1$. The torsion of S, $\mathrm{Tors}(S)$, is cyclic of order at most five, and Reid has shown that in the cases of torsion $\mathbb{Z}_3$, $\mathbb{Z}_4$, and $\mathbb{Z}_5$ the moduli spaces are smooth and irreducible of dimension eight. In comparison, in the cases of $\mathrm{Tors}(S) = 0$ and $\mathrm{Tors}(S) = \mathbb{Z}_2$, little is known about the moduli space; while several examples of these surfaces have been found a more general classification is still unknown.

Surfaces with these invariants are called numerical Godeaux surfaces, after Godeaux who provided the first example, as the $\mathbb{Z}_5$-quotient of a quintic hypersurface in $\mathbb{P}^3$. Most known constructions of numerical Godeaux surfaces have an involution. One particular method for constructing these surfaces was proposed by Campedelli: as the minimal resolution of the double cover of the plane, branched along a degree ten curve with one quadruple point, five infinitely near triple points, such that these six singular points do not lie on a conic. In this talk we survey the known numerical Godeaux surfaces constructed as double planes; the cases of torsion equal to $0, \mathbb{Z}_2$, $\mathbb{Z}_4$, and $\mathbb{Z}_5$ all occur.

The first construction of a numerical Godeaux as a double plane is due to Oort and Peters, whose resulting surface has order four torsion. Reid proved that the classical Godeaux construction can also be realized as a Campedelli double plane; this construction has torsion of order five. As both the numerical Godeaux surfaces with torsion group $\mathbb{Z}_4$ and $\mathbb{Z}_5$ have irreducible moduli spaces, and constructions as double planes, one can ask if the same will be true for the other three cases.

For trivial torsion, a surface constructed as the resolution of a singular quintic in $\mathbb{P}^3$ by Craighero and Gattazzo has been shown to be a double plane. In the case of order two torsion there is a four dimensional family of double plane Godeaux surfaces.

After cataloguing these known double plane Godeaux surfaces, we classify the possible degree ten branch curves that are invariant under an involution of the plane. The idea of looking for branch curves with this additional symmetry was proposed by Stagnaro; following this idea one can prove

Theorem 1. *Let C be a degree ten plane curve with the singularities required for a numerical Godeaux double plane, and suppose C is invariant under involution. Then the resulting double cover branched along C has torsion group $\mathbb{Z}_4$.*

Moreover one can determine all possible decompositions of the branch curve; the example of Oort and Peters belongs to this class of constructions.

References

[1] M. Reid, *Campedelli vs. Godeaux*, in *Problems in the theory of surfaces and their classification (Cortona, 1988)*, Sympos. Math. XXXII, Academic Press, London 1991, 309–365.

[2] M. Reid, *Surfaces with $p_g = 0, K^2 = 1$*, J. Fac. Science Univ. Tokyo, Sect. IA, **25** (1978), 75–92.

[3] E. Stagnaro, *On Campedelli branch loci*, Ann. Univ. Ferrara Sez. VII (N.S.), **43** (1997), 1–26.

[4] C. Werner, *Branch curves for Campedelli double planes*, Rocky Mountain J. Math. (to appear).

A new proof for the adjoint theorem and a Castelnuovo's conjecture

Francesco Zucconi

Let $\xi \in H^1(X, \mathcal{T}_X)$ be the class of a first order deformation $\pi : \mathcal{X} \to \text{Spec } \mathbb{C}[\epsilon]/(\epsilon^2)$ being X a n-dimensional projective variety. Let $\langle \eta_1, \dots, \eta_{n+1} \rangle$ be an ordered set of $n+1$ linearly independent sections of $\text{Ker}(\delta_\xi : H^0(X, \Omega^1_X) \to H^1(X, \mathcal{O}_X))$ where δ_ξ is the coboundary map associated to the sequence: $0 \to \mathcal{O}_X \to \Omega^1_{\mathcal{X}|X} \to \Omega^1_X \to 0$. If $s_1, \dots, s_{n+1}$ are liftings in $H^0(X, \Omega^1_{\mathcal{X}|X})$ of respectively $\eta_1, \dots, \eta_{n+1}$ and $\Omega \in H^0(X, \wedge^{n+1}\Omega^1_{\mathcal{X}|X})$ is the form corresponding to $s_1 \wedge \ldots \wedge s_{n+1} \in \wedge^{n+1} H^0(X, \Omega^1_{\mathcal{X}|X})$ then via the isomorphism $L_\xi : H^0(X, \wedge^{n+1}\Omega^1_{\mathcal{X}|X}) \to H^0(X, \wedge^n \Omega^1_X)$ we obtain a top form $\omega_{\xi, \langle \eta_1, \dots, \eta_{n+1} \rangle} = L_\xi(\Omega)$. This form is called adjoint form of ξ and $\langle \eta_1, \dots, \eta_{n+1} \rangle$. If W is the subvector space generated by $\langle \eta_1, \dots, \eta_{n+1} \rangle$ and $\wedge^n W$

is the subvector space of $H^0(X, \wedge^n \Omega^1_X)$ given by $\langle \eta_1 \wedge \ldots \wedge \hat{\eta}_i \ldots \wedge \eta_{n+1} \rangle$ the adjoint theorem states that: if $\omega_{\xi, \langle \eta_1, \ldots, \eta_{n+1} \rangle} \in \wedge^n W$ then $\xi \in \mathrm{Ker}(H^1(X, \mathcal{T}_X) \to H^1(X, \mathcal{T}_X(D)))$ where D is the fixed component of the sublinear system given by $| \wedge^n W |$.

In this talk we present a new proof of this theorem based on the natural interpretation of the condition $s_1 \wedge \ldots \wedge s_{n+1} = 0$ as integrability condition for the system $s_1 \wedge \ldots \wedge \widehat{s_i *} \wedge \cdots \wedge s_{n+1} = 0$, $i = 1, \ldots, n+1$. We explain the relations between the solution of this system and the geometry of the natural map $\pi : \mathbb{P}(\Omega^n_{\mathcal{X}|X}) \to X$. In the second part of the talk we show the proof of the Castelnuovo conjecture stating that the number m of moduli of an irregular surfaces with $q \geq 4$ and Albanese map of degree 1 is less or equal to $p_g + 2q - 3$. In the final part we discuss some possible applications to surfaces with $q = 4$.

References

[1] G. Pirola, F. Zucconi, *Variations of the Albanese morphisms*, J. Alg. Geom. **12** (2003), 535–572.

Reporter: Alberto Calabri

Participants

Christian Böhning
boehning@btm8x5.mat.uni-bayreuth.de
Lehrstuhl für Mathematik VIII
Universität Bayreuth
NW - II
95440 Bayreuth

Prof. Dr. Giuseppe Borrelli
borrelli@mat.uniroma3.it
c/o Ciro Ciliberto
Dipartimento di Matematica
Universita di Roma Tor Vergata
via della Ricerca Scientifica
I-000133 Rome

Prof. Dr. Vasile Brinzanescu
Vasile.Brinzanescu@imar.ro
Institute of Mathematics of the
Romanian Academy
P.O. Box 1-764
70700 Bucharest
ROMANIA

Dr. Alberto Calabri
calabri@dm.unibo.it
Dipartimento di Matematica
Universita degli Studi di Bologna
Piazza di Porta S. Donato, 5
I-40126 Bologna

Prof. Ciro Ciliberto
cilibert@mat.uniroma2.it
Dipartimento di Matematica
Universita di Roma "Tor Vergata"
V.della Ricerca Scientifica, 1
I-00133 Roma

Prof. Dr. Jonghae Keum
jhkeum@kias.re.kr
School of Mathematics
Korea Inst. for Advanced Study
207-43 Cheongnyangri-dong.
Dongdaemun-gu
Seoul 130-722
Korea

Prof. Dr. Kazuhiro Konno
konno@math.wani.osaka-u.ac.jp
Department of Mathematics
Graduate School of Science
Osaka-University Machikaneyama 1-16
Toyonaka
Osaka 560-0043
JAPAN

Prof. Dr. Margarida Mendes Lopes
mmlopes@math.ist.utl.pt
Departamento de Matematica
Instituto Superior Tecnico
Avenida Rovisco Pais, 1
P-Lisboa 1049-001

Prof. Dr. Stavros Papadakis
spapad@maths.warwick.ac.uk
stavrospapadakis@hotmail.com
papadakis@math.uni-s
FB Mathematik und Informatik
Geb. 27
Universität des Saarlandes
66123 Saarbrücken

Prof. Dr. Ulf Persson
ulfp@math.chalmers.se
Department of Mathematics
Chalmers University of Technology
S-412 96 Göteborg

Roberto Pignatelli
Roberto.Pignatelli@uni-bayreuth.de
pignatel@science.unitn.it
Dipartimento di Matematica
Universita di Trento
Via Sommarive 14
I-38050 Povo (Trento)

Prof. Dr. Francesco Polizzi
polizzi@mat.uniroma2.it
Dipartimento di Matematica
Universita di Roma "Tor Vergata"
V.della Ricerca Scientifica, 1
I-00133 Roma

Prof. Dr. Caryn Werner
cwerner@allegheny.edu
Mathematics Department
Allegheny College
520 North Main St.
Meadville PA 16335
USA

Dr. Francesco Zucconi
zucconi@dimi.uniud.it
Dipartimento di Matematica e
Informatica
Universita di Udine
Via delle Scienze 206
I-33100 Udine

Mathematisches Forschungsinstitut Oberwolfach

Report No. 10/2004

Mini-Workshop: Wavelets and Frames

Organised by
H. Feichtinger (Vienna)
P. Jorgensen (Iowa City)
D. Larson (College Station)
G. Ólafsson (Baton Rouge)

February 15th – February 21st, 2004

Introduction by the Organisers

The workshop was centered around two important topics in modern harmonic analysis: "Wavelets and frames", as well as the related topics "time-frequency analysis" and "operator algebras". [1]

The theory of *frames*, or stable redundant non-orthogonal expansions in Hilbert spaces, introduced by Duffin and Schaeffer in 1952, plays an important role in *wavelet theory* as well as in *Gabor (time-frequency) analysis* for functions in $L^2(\mathbb{R}^d)$. Besides traditional and relevant applications of frames in signal processing, image processing, data compression, pattern matching, sampling theory, communication and data transmission, recently the use of frames also in numerical analysis for the solution of operator equation by adaptive schemes is investigated. These important applications motivated the study of frames as decompositions in classical Banach spaces, e.g. Lebesgue, Sobolev, Besov, and modulation spaces. Fundamental concepts on operator theory, as well as on the theory of representations of groups and algebras are also involved and they have inspired new directions within frame theory with applications in pseudodifferential operator and symbolic calculus and mathematical physics.

[1]Partial travel funding was provided by grants in Austria, Germany, and the USA. The three organizers from the US, are part of a Focused Research Group (FRG), funded by the US National Science Foundation (NSF), and two other participants are in this FRG group, Professors Chris Heil, GATECH, USA, and Akram Aldroubi, Vanderbilt University, USA. The organizers thank the US NSF for partial support.

Any element f of the Hilbert space $\mathcal{H}$ can be expanded as a series with respect to a frame $\mathcal{G} = \{g_n\}_{n\in\mathbb{Z}^{2d}}$ in $\mathcal{H}$, and the coefficients of such expansion can be computed as scalar products of f with respect to a *dual frame* $\tilde{\mathcal{G}} = \{\tilde{g}_n\}_{n\in\mathbb{Z}^{2d}}$:

$$f = \sum_{n\in\mathbb{Z}^{2d}} \langle f, \tilde{g}_n\rangle g_n, \text{ for all } f \in \mathcal{H}. \tag{1}$$

In particular, $\mathcal{G}$ is a frame if (and only if) the so called *frame operator*

$$Sf = \sum_{n\in\mathbb{Z}^{2d}} \langle f, g_n\rangle g_n,$$

is continuous and continuously invertible on its range. Then there exists a *canonical* choice of a possible dual frame (delivering the minimal norm coefficient) defined by the equation

$$S\tilde{\mathcal{G}} = \mathcal{G}.$$

The existence of a dual frame makes the expansion (1) work. On the other hand, it may be a hard problem to predict properties of the *canonical dual frame* since it is only implicitly defined by the previous equation, and not always is there an efficient way of computation approximations at hand. This motivated the so called *localization theory for frames*, making use of well-chosen Banach $*$-algebras of infinite matrices. They allow to deduce relevant properties of the canonical dual and to extend the Hilbert space concept of frames to *Banach frames* which characterize corresponding families of Banach spaces.

Another problem within frame theory concerns structured families of functions, depending perhaps on several parameters, and the question of whether such a family constitutes a frame for $L^2(\mathbb{R}^d)$. Classical examples are the following ones. Gabor frames are frames in $L^2(\mathbb{R}^d)$ constructed by *modulations* and *translations*: given a square-integrable function g our sequence is $g_{nm}(x) = e^{2\pi i(m,x)}g(x-n)$, $(n,m) \in \Lambda$, where Λ is a discrete subset of $\mathbb{R}^{2d}$. The wavelet frames are constructed using *dilations* and *translations*: given a set $\Delta \subset \mathrm{GL}(d,\mathbb{R})$ and $\Gamma \subset \mathbb{R}^d$, as well as a suitable square integrable function ψ, we set $\psi_{D,\gamma}(x) = |\det D|^{1/2}\psi(Dx+\gamma)$ for $D \in \Delta$ and $\gamma \in \Gamma$. They are canonically related to Besov spaces. The reader can find several interesting questions and problems related to those concepts in the following abstracts.

We would like to exemplify here two simple existence problems. If the density of the points in Λ is too small, then a Gabor frame cannot be constructed, and if the density is too large, then one can construct a frame, but not a basis. Suitable definitions of density and their relations with respect to the existence of frames in one of the current relevant topics in the frame theory.

In the wavelet case, an interesting problem has been the construction of *wavelet sets*. Given the set Δ and Λ, find the measurable subsets $\Omega \subset \mathbb{R}^d$ of positive, and finite measure, such that, with $\hat{\psi} = \chi_\Omega$, the sequence $\{\psi_{D,\gamma}\}_{D\in\Delta,\gamma\in\Gamma}$, is an orthogonal basis for $L^2(\mathbb{R}^d)$. Such a set is called a *wavelet set*. This line of work includes both geometry (tilings of $\mathbb{R}^d$) and analysis (the Fuglede conjecture). More general question is when $\{\psi_{D,\gamma}\}_{D\in\Delta,\gamma\in\Gamma}$ can be a frame.

Other more general frames, called *wave packets*, can be constructed as combinations of modulations, translations and dilations to *interpolate* the time-frequency properties of analysis of Gabor and wavelet frames. Interesting problems related to density and existence of such frames are an important direction of research and connections with new Banach spaces (for example α*-modulation spaces*), Lie groups (for example the *affine Weyl-Heisenberg group*), and representation theory (for example the *Stone-Von Neumann representation*) are currently fruitful fields of investigation. All these families of frames are generated by the common action of translations. *Shift invariant spaces* and *their generators* constituted the main building blocks from which to start the construction of more complicated systems. They showed relevant uses in engineering, signal and image processing, being one of the most prominent branch in the applications.

Rather than formal presentations of recent advances in the field, this workshop tried instead to aim at outlining the important problems and directions, as we see it, for future research, and to discuss the impact of the current main trends. In particular, the talks were often informal with weight on interaction between the speaker and the audience, both in form of discussion and general comments. A special problem session was organized by D. Larson one afternoon. Another afternoon session was devoted to talks and informal discussions of further open problems, new directions, and trends.

The topics that emerged in these discussions included the following general areas:

(1) Functional equations and approximation theory: wavelet approximation in numerical analysis, PDE, and mathematical physics. At the meeting, we discussed some operator theoretic methods that resonate with what numerical analysts want, and questions about localizing wavelets. We refer to the abstracts by M. Frank and K. Urban for more details. Two workshop lectures covered connections to numerical analysis and PDE.

(2) Gabor frames: We had many discussions, much activity, and several talks on aspects of this. H. Feichtinger explained some important results and discussed some open problems involving frames and Gelfand triples. K. Gröchenig gave a lecture on new formulations and results generalizing Wiener's inversion theorems, in particular for twisted convolution algebras and Gabor frames. The applications are striking in that they yield sharper frame bounds. And they involve non-commutative geometry and other operator algebraic tools. C. Heil discussed the basic properties of frames which are not bases, and in particular he discussed the current status of the still-open conjecture that every finite subset of a Gabor frame is linearly independent. A related problem is that there do not exist any explicit estimates of the frame bounds of finite sets of time-frequency shifts.

(3) Continuous vs. discrete wavelet transforms: We had several talks at the Oberwolfach workshop where the various operations, translation, scaling, phase modulation, and rotation, get incorporated into a single group. H.

Führ and G. Ólafsson gave talks, where links to Lie groups and their representations were discussed. This viewpoint seems to hold promise for new directions, and for unifying a number of current wavelet constructions, tomography, scale-angle representations, parabolic scaling, wavelet packets, curvelets, ridgelets, de-noising ... Wavelets are usually thought of as frames in function spaces constructed by translations and dilations. Much less is understood in the case of compact manifolds such as the n-dimensional sphere, where both "dilations" and "translations" are not obviously defined. The talk by Ilgewska–Nowak explained some of her joint work with M. Holschneider on the construction of discrete wavelet transforms on the sphere.

(4) Harmonic analysis of Iterated Function Systems (IFS): Several of the participants have worked on problems in the area, and P. Jorgensen spoke about past work, and directions for the future. The iterated function systems he discussed are closely related to the study of spectral pairs and the Fuglede problem. Recent work by Terence Tao makes the subject especially current.

(5) Multiplicity theory, spectral functions, grammians, generators for translation invariant subspaces, and approximation rates: We had joint activity at the workshop on problems in the general area, and we anticipate joint papers emerging from it. A. Aldroubi lectured on the engineering motivations. In particular he discussed translation invariant subspaces of $L^2(\mathbb{R})$ where two lattice-scales are involved, and issues about localizing the corresponding generating functions for such subspaces. O. Christensen presented an equivalent condition for two functions generating dual frame pairs via translation. The result lead to a way of finding a dual of a given frame, which belongs to a prescribed subspace. Several open questions related to this were discussed.

(6) Decompositions of operators and construction of frames: D. Larson discussed the problem of when is a positive operator a sum of finitely many orthogonal projections, and related it to frame theory. Problems and some recent results and techniques of D. Larson and K. Kornelson were discussed in this context, involving other related types of targeted decompositions of operators. In response, H. Feichtinger and K. Gröchenig pointed out that similar techniques just may lead to progress on a certain problem in modulation space theory. There are plans to follow up on this lead.

(7) Wave packets: We had two talks at the workshop about this broad research area. G. Kutyniok gave a talk about the role of the geometric structure of sets of parameters of wave packets for the functional properties of associated systems of functions. In this context some recent results of D. Speegle, G. Kutyniok, and W. Czaja were discussed. M. Fornasier presented the construction of a specific family of wave packet frames for $L^2(\mathbb{R})$ depending on a parameter $\alpha \in [0, 1)$, as a mixing tuner between

Gabor and wavelet frames. These more classical and well-known frames arise as special and extreme cases.

The organizers:

H. Feichtinger, P. Jorgensen, D. Larson, and G. Ólafsson

Mini–Workshop on Wavelets and Frames

Table of Contents

Abstracts

Almost Translation Invariant Spaces

Akram Aldroubi

Shift invariant spaces that are considered are of the form

$$V^2(\Phi) = \Big\{ \sum_{j\in\mathbb{Z}} D(j)^T \Phi(\cdot - j) : D \in (\ell^2)^{(r)} \Big\} \tag{1}$$

for some vector function $\Phi = (\phi_1, \dots, \phi_r)^T \in (L^2)^{(r)}$, where $D = (d_1, \dots, d_r)^T$ is a vector sequence such that $d_i := \{d_i(j)\}_{j\in\mathbb{Z}} \in \ell^2$, i.e., $D \in (\ell^2)^{(r)}$. Thus $\sum_{j\in\mathbb{Z}} D(j)^T \Phi(\cdot - j) = \sum_{i=1}^r \sum_{j\in\mathbb{Z}} d_i(j)\phi_i(\cdot - j)$.

We also assume that the Gramian satisfies

$$G_\Phi(\xi) := \sum_{k\in\mathbb{Z}} \widehat{\Phi}(\xi + k)\overline{\widehat{\Phi}(\xi + k)}^T = I, \quad a.e.\, \xi, \tag{2}$$

where I is the identity matrix.

An important and prototypical space is the space of band-limited functions where $r = 1$, $\phi = \frac{\sin(\pi x)}{\pi x}$. This space is translation invariant for all translates. This feature is important in applications since it allows the construction of signal/image processing algorithms that are invariant under time or space translations. However, band-limited functions are analytic and are not always well suited as signal models or for computational purposes. For this reason, we wish to investigate spaces that are almost translation invariant, thereby allowing for almost reproducibility and origin independence of the algorithms without the limitation of analyticity and the computational complexity of band-limited function space.

Let T_a be the translation operator by a factor a, i.e., $(T_a f)(x) = f(x - a)$, then obviously $T_1 V = V$. We would like to characterize the generators Φ such that $T_{1/n}V = V$ for some fixed integer n. This problem has been studied and Φ characterized for a particular case by Weber in [3] and for the general case by Chui and Sun in [1]. For the case $r = 1$ and $n = 2$ we have the following useful characterization:

Let $E_0 := \{\xi \in [0, 1) : \phi(\xi + 2j) \neq 0$ for some $j \in \mathbb{Z}\}$, $E_1 := \{\xi \in [0, 1) : \phi(\xi + 2j + 1) \neq 0$ for some $j \in \mathbb{Z}\}$, then $T_{1/n}V = V$ if and only if $E_0 \cup E_1 = [0, 1)$, and $E_0 \cap E_1 = \emptyset$. We conjecture that a similar characterization which is not an easy or direct consequence of [1] can be obtained for the general case.

Another direction that we will investigate is the problem of ϵ-$1/n$ translation invariant: Given $\epsilon > 0$ we wish to study the set A_ϵ of generators Φ such that

$$\sup\{\|f(\cdot - 1/n) - Pf(\cdot - 1/n)\|,\ f \in V, \|f\| = 1\} \leq \epsilon$$

where P is the orthogonal projection on V. This problem is related to the problem discussed in [2]. The problems under considerations are currently investigated in collaboration with C. Heil, P. Jorgensen, K. Kornelson, and G. Olafsson.

References

[1] C. K. Chui, and Q. Sun, Tight frame oversampling and its equivalence to shift-invariance of affine frame operators, *Proc.Amer.Math.Soc.*, **7**(2002), 1527–1538.

[2] J. A. Hogan and J.D. Lakey, Sampling and aliasing without translation-invariance, *Proceedings of the 2001 International Conference on Sampling Theory and Applications* (2001) p. 61–66.

[3] S. Schaffer Vestal, and E. Weber, Orthonormal wavelets and shift invariant generalized multi-resolution analysis. *Proc.Amer.Math.Soc.* , **10** (2003), 3089–3100.

Trends in Frame Theory

Ole Christensen

The increased flexibility (compared to orthonormal bases) is often an argument for the use of frames. However, in most cases we also want our frames to have some structure, and there are cases where this additional constraint limits (or removes) the freedom. For this reason we seek to extend classical frame theory by allowing duals belonging to a different space than the frame.

Given a frame for a subspace W of a Hilbert space H, we characterize the set of oblique dual frame sequences (i.e., dual frame sequences that are not constrained to lie in W). We then consider frame sequences in shift invariant spaces, and characterize the translation invariant oblique dual frame sequences. For a given translation invariant frame sequence an easily verifiable condition on another shift-invariant frame sequence implies that its closed linear span contains a generator for a translation invariant dual of the frame sequence we start with; in particular, this result shows that classical frame theory does not provide any freedom if we want the dual to be translation invariant. In the case of frame sequences generated by B-splines we can use our approach to obtain dual generators of arbitrary regularity.

Some open problems were presented during the lecture:

- It is well known that the canonical dual of a wavelet frame does not necessarily have the wavelet structure. Which conditions on the generator implies that the canonical dual has wavelet structure? The answer is known for quasi-affine systems, cf. [1].
- Frazier et. al have characterized all dual wavelet frame pairs for $L^2(R)$. How can this be extended to frames for subspaces?
- Is it possible to construct a tight Gabor frame for which the generator g as well as $\hat{g}$ decay exponentially and g is given explicitly in closed form as a linear combination of elementary functions?

References

[1] Bownik, M. and Weber, E.: *Affine frames, GMRA's, and the canonical dual.* Studia Math. **159** (2003), 453-479.

[2] Christensen, O. and Eldar, Y.: *Oblique dual frames and shift-invariant spaces.* 27 pages. To appear in Appl. Comp. Harm. Anal., 2004.

Banach frames, Banach Gelfand Triples, and Wiener Amalgam Spaces
Hans G. Feichtinger

The theory of *frames* is usually described in the context of Hilbert spaces. One may consider frames as those sequences in a Hilbert spaces which rich enough to allow the representations of all the elements in a given Hilbert space, using a series expansion with square summable coefficients. Equivalently to the standard definition one can say that the coefficient mapping $C : f \mapsto (\langle f, f_n \rangle)_{n \in N}$ establishes an isomorphism between the Hilbert space and a its closed range in ℓ^2. The natural inversion (the Moore-Penrose inverse to the coefficient mapping) - we will call it R - is defined on all of ℓ^2, projecting a given sequence onto the range of the coefficient mapping and then back to the uniquely determined function having the given coefficients. As a matter of fact R is realized by the usual (canonical) *dual frame*, called $(\tilde{f}_n)$, via $\mathbf{c} \mapsto \sum_n c_n \tilde{f}_n$. Obviously one has $R \circ C = Id_H$, which is just another form of describing the standard frame expansion for $f \in H$. Since R is bounded and any sequence in ℓ^2 is the norm limit of its finite sections, the convergence of the series is unconditional as well.

From a more abstract point one can say that the pair (C, R) establishes a **retract** between the Hilbert space H and the sequence space ℓ^2, making H isomorphic to a subspace of ℓ^2 (via C) and at the same time to a quotient of ℓ^2 (namely $\ell^2 / null(R)$).

The established notion of a *Banach frame* (as formalized by K. Gröchenig in [Grö91]) extends some aspects of this situation to the case where H is replaced be some Banach space and ℓ^2 by some Banach space of sequences (such as a weighted mixed-norm ℓ^p-space). We would like to suggest to add to these assumptions that the Banach space of sequences is also solid (i.e. $|x_n| \leq |y_n|$ for some sequence y, and all n should imply that $\|x\|_B \leq \|y\|_B$). This would imply unconditional convergence of the reconstruction process (which is not granted by the standard terminology). [2] We will see in a moment that this is not merely an abstract generalization of the frame concept but contributes very much to the actual usefulness of Gabor or wavelet frames.

It is however true that this is only part of the story. Wavelet and Gabor systems would not be so useful for applications if aside from the fact that their coefficients have a specific "meaning" in terms of time, frequency or scale they would not be useful to characterize various functions spaces (for example Besov-Triebel-Lizorkin spaces, with wavelet coefficients in suitable weighted mixed-norm spaces). So, in a way, the Banach frames for individual couples (one Banach space of functions and its corresponding Banach space of sequences) are just continuous extension of the corresponding mappings C and R defined on the smaller spaces. While "Banach frames for compatible families of Banach spaces" are an important mathematical concept they are not so easy to explain to non-experts, and therefore we discuss *Banach Gelfand Triples*: Given a Banach space $(B, \|\cdot\|_B)$ and some Hilbert space $\mathcal{H}$ are forming a *Banach Gelfand Triple* $(B, \mathcal{H}, B')$ if the following is true:

[2] We suggest the term *"unconditional Banach frame"*.

- $B \hookrightarrow \mathcal{H} \hookrightarrow B'$;
- B is norm dense in $\mathcal{H}$ and w^*-dense in B'.

The prototypical example consists of the sequence space $(\ell^1, \ell^2, \ell^\infty)$. For many applications in Gabor Analysis the (minimal TF-shift invariant) Segal algebra S_0 (cf. [Fei81]) plays an important role. Together with it's dual it establishes a Banach Gelfand triple (S_0, L^2, S_0'), the S_0-GT.

There is a natural concept of *"Gelfand triple morphism"*: bounded linear mappings at each level, mapping the corresponding "small spaces" into each other, also the corresponding Hilbert spaces, and finally the dual spaces with respect to two topologies, their standard norm topologies and their w^*-topologies respectively. If such a mapping is unitary at the level of Hilbert spaces we will call it a "unitary Gelfand triple isomorphism".

A really basic example of such a unitary GT-isomorphism is the Fourier transform, acting on (S_0, L^2, S_0'). While Plancherel's theorem takes care of the L^2 case, this statement includes the fact that the Fourier transform maps S_0 into itself, but also extends to the (not too large) dual space $S_0'(G)$. At the S_0 level one can use ordinary Riemannian integrals while at the $S_0'-$level one finds that "pure frequencies" are mapped into point-measures (i.e. Dirac Deltas). This is the correct analogue of the "linear algebra situation" (connected with the DFT or FFT), describing it simply as a (orthogonal) change of bases. Moreover, due to the w^*-density of the linear span of pure frequencies resp. discrete measures in S_0' the Fourier transform is uniquely determined by these properties as a unitary Gelfand triple isomorphism.

There are plenty of other Gelfand triple isomorphisms resp. Gelfand triple Banach frames (i.e. retracts between GTs of functions to sequence spaces GTs): Any Gabor frame of the form $(\pi(\lambda)g)_{\lambda\in\Lambda}$, with a Gabor atom in $S_0(R^d)$, and some lattice $\Lambda = \mathbf{A} * Z^{2d}$, for some non-singular $2d \times 2d$ matrix $\mathbf{A}$ has the property (as shown by Gröchenig and Leinert in their recent paper) that the canonical dual window $\tilde{g}$ also belongs to $S_0(R^d)$, and therefore the mappings establishing the standard frame diagram, $C(f) = V_g f(\lambda)$ and $R(\mathbf{c}) = \sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda)\tilde{g}$ extend to a retract between the Gelfand triples $(S_0, L^2, S_0')(R^d)$ and the GT $(\ell^1, \ell^2, \ell^\infty)(\Lambda)$. Wilson bases built from S_0 atoms are in fact establishing unitary GT isomorphisms between the same GTs (this is the perfect analogue to the statement of linear algebra: bases are in a one-to-one correspondence to isomorphisms between a finite dimensional vector space and its canonical version R^k).

In connection with operators (relevant for time-frequency analysis) one should point at various representations of operators. While we know from linear algebra that linear mappings from R^n to R^m can be uniquely determined by their matrices (with respect to given matrices) we have to look for a GT analogue in the case of non-finite groups. As already in the case of the finite groups (e.g Z_N, the cyclic group of order N) we have different choices by just making use of the standard basis (of "unit vectors" which then turn into Diracs, resp. pure frequencies, for example). We mention here only the the most important ones (one can find many applications in [FK98]).

Writing $\mathcal{L}$ for the space of bounded linear operators one finds that $\mathcal{L}(S_0', S_0)$ is identified with "smooth kernels", i.e. any such operator T has a nice (continuous and integrable) kernel $K = K(x,y)$ such that for functions f as input one has $Tf(x) = \int K(x,y)f(y)dy$. Just as one would identify the matrix of a linear mapping by realizing its columns as the images of the unit vectors, one expects that $K(x,y) = T(\delta_y)(x)$, which makes sense, because $\delta_y \in S_0'$ while $T(\delta_y)$ is a continuous function in S_0. Of course it is important to see that this functional connection can be extended to a unitary GT isomorphism. The Hilbert space (of operators) being now the space of Hilbert-Schmidt operators $\mathcal{HS}$. Since they are exactly the integral operators with kernel $K \in L^2(G \times G)$, acting on $L^2(G)$ they are also contained in $\mathcal{L}(S_0, S_0')$, which makes $(\mathcal{L}(S_0', S_0), \mathcal{HS}, \mathcal{L}(S_0, S_0'))$ a GT. The kernel theorem can be interpreted as a unitary GT-isomorphism between this triple and their kernels in $(S_0, L^2, S_0')(G \times G)$. The so-called *spreading symbol* of operators. It can be characterized as the uniquely determined unitary GT isomorphism between the GT of operator spaces given above to the S_0-GT over phase space (i.e. $G \times \hat{G}$), which identifies the pure time-frequency shifts $\pi(\lambda) = M_\omega T_t$ for $\lambda = (t, \omega)$ with δ_λ. An often used argument in Gabor analysis is the fact that Gabor frame operators commute with TF-shifts from a given TF-lattice Λ and therefore have a so-called Janssen representation: they can be written as an infinite series of the form $T = \sum_{\lambda^\circ \in \Lambda^\circ} c_{\lambda^\circ} \pi(\lambda^\circ)$ can be seen as a consequence of the following GT statement. Here Λ° is the "adjoint lattice" to Λ, which in the case of $aZ^d \times bZ^d$ equals $(1/b)Z^d \times (1/a)Z^d$. The operators in $\mathcal{L}(S_0, S_0')$ which commute with TF-shifts from Λ are exactly the ones having a Janssen representation. Moreover, the mapping between the operators in $\mathcal{HS}-$ GT of operator spaces with this extra property is isomorphic to the GT $(\ell^1, \ell^2, \ell^\infty)(\Lambda^\circ)$ through the mapping from T to it's Janssen coefficients (c_λ).

While the spreading function is an important tool in communication theory, because it is used to model slowly time-variant channels occurring in wireless communication, the Kohn-Nirenberg symbol of an operator is more popular in the context of pseudodifferential operators. However, it is not difficult to show that the symplectic Fourier transform, which is another unitary Gelfand triple isomorphism onto itself establishes in a natural link between spreading symbol and KN-symbol of a linear operator. Needless to say that, as a consequence of the statements above, the membership of the KN-symbol in the GT (S_0, L^2, S_0') is again equivalent to the membership of the operator in the corresponding member of the $\mathcal{HS}$-GT. It turns out to be also an appropriate tool to establish a connection between the theory of *Gabor multipliers* and the theory of spline type (resp. principal shift invariant) spaces. The most interesting case for Gabor multipliers, i.e. operators of the form $Tf = \sum_{\lambda \in \Lambda} m_\lambda P_\lambda f$, with $P_\lambda(f) = \langle f, \pi(\lambda)g \rangle \pi(\lambda)g$ arises when these operators form a Riesz basis for their closed linear span within $\mathcal{HS}$, which is the case if and only if the Λ- Fourier transform of the function $|V_g g(\lambda)|^2$ is free of zeros. One can show that in this case there is a canonical bi-orthogonal family (Q_λ) in their closed linear span $\mathcal{GM}_2$ (within $\mathcal{HS}$). Hence the orthogonal projection of $\mathcal{HS}$ onto $\mathcal{GM}_2$ takes the form $T \mapsto \sum_\lambda \langle T(\pi(\lambda)g), \pi(\lambda)g \rangle Q_\lambda$. If the atom g is in $S_0(R^d)$ then one can also show that $P_\lambda \in \mathcal{L}(S_0', S_0)$ and that that orthogonal projection extends to a bounded GT-mapping from the $\mathcal{HS}$-GT onto the Gelfand triple of Gabor multipliers $(\mathcal{GM}_1, \mathcal{GM}_2, \mathcal{GM}_\infty)$ with coefficients in the GT triple $(\ell^1, \ell^2, \ell^\infty)(\Lambda)$.

Finally we mention that *Wiener amalgam spaces* are at the technical level an important tool. It can be used to show the boundedness of coefficient operators (between suitable couples of Banach spaces), respectively the corresponding synthesis operators, but we cannot go into details here. A report on the use of Wiener amalgam spaces in the context of Gabor analysis is under preparation.

References

[Fei81] H. G. Feichtinger, *On a new Segal algebra*, Monatsh. f. Math. **92**, 1981, 269-289.

[Fei91] H. G. Feichtinger, *Wiener amalgams over Euclidean spaces and some of their applications*, In K. Jarosz, editor, Proc. Conf. Function spaces **136** of Lect. Notes in Math., Edwardsville, IL, April 1990, 1991. M. Dekker, 123-137.

[FGr88] H. G. Feichtinger, K. Gröchenig, *A unified approach to atomic decomposition via integrable group representations*, Springer Lect. Notes Math. **1302**, 1988.

[FGr89] H. G. Feichtinger, K. Gröchenig, *Banach spaces related to integrable group representations and their atomc decompositions I*, J. Funct. Anal. **86**, 1989, 307-340.

[FGr89I] H. G. Feichtinger, K. Gröchenig, *Banach spaces related to integrable group representations and their atomc decompositions II*, Monatsh. f. Math. **108**, 1989, 129-148.

[FGr92] H. G. Feichtinger, K. Gröchenig, *Iterative reconstruction of multivariate band-limited functions from irregular sampling values*, SIAM J. Math. Anal. **23**, No. 1, 1992, 244-261.

[FS98] H. G. Feichtinger, T. Strohmer (Eds.), *Gabor Analysis and Algorithms*, Birkhäuser, 1998.

[FS03] H. G. Feichtinger, T. Strohmer (Eds.), *Advances in Gabor Analysis*, Birkhäuser, 2003.

[FK98] H.G. Feichtinger and W. Kozek. Quantization of TF lattice-invariant operators on elementary LCA groups. In [FS98], 233-266.

[FZ98] H.G. Feichtinger and G. Zimmermann. A Banach space of test functions for Gabor analysis. In [FS98], 123-170.

[Grö02] K. Gröchenig, *Foundation of Time-Frequency Analysis*, Birkhäuser Verlag, 2002.

[Grö91] K. Gröchenig, *Describing functions: atomic decompositions versus frames*, Monatsh. Math. **112**, 1991, 1-41.

Building a Bridge between Gabor and Wavelet Worlds

Massimo Fornasier

The theory of *frames* or stable redundant non-orthogonal expansions in Hilbert spaces, introduced by Duffin and Schaeffer [DS52] plays an important role in wavelet theory [Dau92, Kai94] as well as in Gabor analysis [Grö02, FS98, FS03]. Many relevant contributions describe Gabor and wavelet analysis as two parallel theories with similar, but different structures and typically different applications. In [FGr88, FGr89, FGr89I, Grö91] Feichtinger and Gröchenig presented a unified approach to Gabor and wavelet analysis, which can*not* be used to describe any *intermediate* theory. Therefore, as a further [HN03, Tor91, Tor92] answer to the theoretical need of a common interpretation and framework between Gabor and wavelet frames, the author has recently proposed [FF04] the construction of frames, which allows to ensure that certain family of Schwartz functions on $\mathbb{R}$ obtained by a suitable combination of translation, modulation of dilation

$$\begin{aligned}
&T_x(f)(t) = f(t-x),\\
&M_\omega(f)(t) = e^{2\pi i\omega\cdot t} f(t),\\
&D_a(f)(t) = |a|^{-1/2} f(t/a), \quad x,\omega,t\in\mathbb{R}, a\in\mathbb{R}_+,
\end{aligned}$$

form Banach frames for the family of L^2-Sobolev spaces of any order. In the construction a parameter $\alpha\in[0,1)$ governs the dependence of the dilation factor on the frequency parameter. The well-known Gabor and wavelet frames arise as special case ($\alpha=0$) and limiting case ($\alpha\to 1$) respectively. One example of such intermediate families is given as follows. Consider the two functions

$$p_\alpha(j) := \operatorname{sgn}(j)\left((1+(1-\alpha)|j|)^{\frac{1}{1-\alpha}}-1\right), \quad s_\alpha(j) := (1+(1-\alpha)(|j|+1))^{\frac{\alpha}{1-\alpha}},$$

and g_0 is the Gaussian function. Then the family $\{g^\alpha_{j,k} := M_{p_\alpha(j)} D_{s_\alpha(j)^{-1}} T_{ak} g_0\}_{j,k\in\mathbb{Z}}$ is in fact a frame for $H^s(\mathbb{R})$ for $s>0$ and for $a>0$ small enough. The parameter α functions as a tuning tool of the mixture of the modulation and dilation operators, like "walking on a bridge" between the Gabor and wavelet worlds. Moreover, to frames endowed with intrinsic localization properties [FoGr04], i.e. the Gramian of the frame has nice off-diagonal decay, one can associate natural Banach spaces [Grö04] defined as the spaces of the frame series expansions with coefficients in suitable corresponding Banach sequence spaces. The associated spaces to Gabor and wavelet frames are the well-known families of modulation [Fei89I, Fei03] and Besov spaces [FJ85] respectively. A natural question arises: which are the associated spaces to the intermediate α-Gabor-wavelet frames? An answer to this question has been given in [For02, For04I], where it has been shown that the associated spaces are in fact the so called *α-modulation spaces*, introduced by Gröbner in 1992 [Grö92] in his Ph.D. thesis (see also [PS88]), as an intermediate family of spaces between modulation and Besov spaces This family is appearing also in other contributions and we refer to [For04I] for an extended literature. Let us just mention here that Borup [Bor04], Holschneider, and Nazaret [HN03] have recently described the mapping properties of pseudodifferential operators on α-modulation spaces as an extension of the earlier work of Cordoba and Fefferman [CF78]. From this, relevant open problems for applications arise, for example, on the behaviour of the spectrum of matrices $\left(\langle T g^\alpha_{j,k}, g^\alpha_{j',k'}\rangle\right)_{j,k,j',k'\in\mathbb{Z}}$, depending on $\alpha\in[0,1)$, associated to symmetric operators T acting on H^s. Anyway, even the more simple and related problem of discussing the behaviour of the *frame bounds* depending on $\alpha\in[0,1)$ might be indeed quite difficult. Also applications in best n-term approximation of functions with respect to the dictionary $\{g^\alpha_{j,k}\}_{j,k\in\mathbb{Z},\alpha\in[0,1)}$ might be investigated [DT01]. In particular the different approximation properties of such α-expansions can characterize different classes of functions, may be related by inclusions to α-modulation spaces.

References

[Bor04] L. Borup, *Pseudodifferential operators on α-modulation spaces*, to appear in J. Func. Spaces and Appl., **2**, no 2, May 2004.

[CF78] A. Cordoba, C. Fefferman, *Wave packets and Fourier integral operators*, Comm. Part. Diff. Eq. **3**, 1978, 979-1005.

[Dau92] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.

[DT01] L. Daudet, B. Torresani, *Hybrid representations for audiophonic signal encoding*, preprint, LATP 01-26, CNRS 6632, 2001.

[DS52] R. J. Duffin, A. C. Schaeffer, *A class of nonharmonic Fourier series*, Trans. Amer. Math. Soc. **72**, 1952, 341-366.

[Fei89I] H. G. Feichtinger, *Atomic characterization of modulation spaces through Gabor-type representations*, Proc . Conf. Constr. Function Theory, Rocky Mountain J. Math. **19**, 1989, 113-126.

[Fei03] H.G. Feichtinger, *Modulation spaces of locally compact Abelian groups*, In R.Radha, editor, *Proc. Internat. Conf. on Wavelets and Applications*, pages 1–56, Chennai, January 2002, 2003.

[FF04] H. G. Feichtinger, M. Fornasier, *Flexible Gabor-wavelet atomic decompositions for L^2-Sobolev spaces*, to appear in Annali di Matematica Pura e Applicata.

[FGr88] H. G. Feichtinger, K. Gröchenig, *A unified approach to atomic decomposition via integrable group representations*, Springer Lect. Notes Math. **1302**, 1988.

[FGr89] H. G. Feichtinger, K. Gröchenig, *Banach spaces related to integrable group representations and their atomc decompositions I*, J. Funct. Anal. **86**, 1989, 307-340.

[FGr89I] H. G. Feichtinger, K. Gröchenig, *Banach spaces related to integrable group representations and their atomc decompositions II*, Monatsh. f. Math. **108**, 1989, 129-148.

[FS98] H. G. Feichtinger, T. Strohmer (Eds.), *Gabor Analysis and Algorithms*, Birkhäuser, 1998.

[FS03] H. G. Feichtinger, T. Strohmer (Eds.), *Advances in Gabor Analysis*, Birkhäuser, 2003.

[For02] M. Fornasier, *Constructive Methods for Numerical Applications in Signal Processing and Homogenization Problems*, Ph.D. thesis, University of Padova, 2002.

[For04I] M. Fornasier, *Banach frames for α-modulation spaces*, preprint, 2004.

[FoGr04] M. Fornasier, K. Gröchenig, *Intrinsic localization of frames*, preprint, 2004.

[FJ85] M. Frazier, B. Jawerth, *Decomposition of Besov spaces*, *Indiana Univ. Math. J.* **34**, 1985, 777-799.

[Grö92] P. Gröbner, *Banachräume glatter Funktionen and Zerlegungmethoden*, Ph.D. thesis, University of Vienna, 1992.

[Grö02] K. Gröchenig, *Foundation of Time-Frequency Analysis*, Birkhäuser Verlag, 2002.

[Grö91] K. Gröchenig, *Describing functions: atomic decompositions versus frames*, Monatsh. Math. **112**, 1991, 1-41.

[Grö04] K. Gröchenig, *Localization of frames, Banach frames, and the invertibility of the frame operator*, J. Four. Anal. Appl. **10**. no. 2, 2004, 105–132.

[HL95] J.A. Hogan, J.D. Lakey, *Extensions of the Heisenberg group by dilations and frames*, Appl. Comp. Harm. Anal. **2**, 1995, 174-199.

[HN03] M. Holschneider, B. Nazaret, *An interpolation family between Gabor and wavelet transformations. Application to differential calculus and construction of anisotropic Banach spaces*, to appear in Advances In Partial Differential Equations, "Nonlinear Hyperbolic Equations, Spectral Theory, and Wavelet Transformations" (Albeverio, Demuth, Schrohe, Schulze Eds.), Wiley, 2003.

[Kai94] G. Kaiser, *A Friendly Guide to Wavelets*, Birkhäuser, 1994.

[PS88] L. Päivärinta, E. Somersalo, *A generalization of the Calderon-Vaillancourt theorem to L^p and h^p*, Math. Nachr. **138**, 1988, pag. 145-156.

[Tor91] B. TORRESANI, *Wavelets associated with representations of the affine Weyl-Heisenberg group*, J. Math. Phys. **32**, 1991, 1273-1279.

[Tor92] B. TORRESANI, *Time-frequency representation: wavelet packets and optimal decomposition*, Ann. Inst. H. Poincaré **56**, 1992, 215-234.

Frames for Hilbert C*-Modules
Michael Frank

There is growing evidence that Hilbert C*-module theory and the theory of wavelets and Gabor (i.e. Weyl-Heisenberg) frames are tightly related to each other in many aspects. Both the research fields can benefit from achievements of the other field. The goal of the talk given at the mini-workshop was to give an introduction to the theory of module frames and to Hilbert C*-modules showing key analogies, and how to overcome the existing obstacles of Hilbert C*-module theory in comparison to Hilbert space theory.

The theory of module frames of countably generated Hilbert C*-modules over unital C*-algebras was discovered and investigated studying an approach to Hilbert space frame theory by Deguang Han and David R. Larson [7]. Surprisingly, almost all of the concepts and results can be reobtained in the Hilbert C*-module setting. This has been worked out in joint work with D. R. Larson in [4, 5, 6]. Complementary results have been obtained by T. Kajiwara, C. Pinzari and Y. Watatani in [8] using other techniques and motivations. Frames have been also used by D. Bakić and B. Guljaš in [1] calling them quasi-bases. Meanwhile, the case of Hilbert C*-modules over non-unital C*-algebras has been investigated by I. Raeburn and S. J. Thompson [14], as well as by D. Bakić and B. Guljaš discovering standard frames even for this class of countably generated Hilbert C*-modules in a well-defined larger multiplier module. However, many problems still have to be solved.

How to link core C*-theory to wavelet theory was first observed by M. A. Rieffel in 1997, cf. [15]. His approach has been worked out by J. A. Packer and M. A. Rieffel [12, 13], and by P. J. Wood [16, 17] in great detail. As major results a framework in terms of Hilbert C*-modules has been obtained sharing most of the basic structures with generalized multi-resolution analysis for key classes of wavelet and Gabor frames. The Gabor case has been investigated by P. G. Casazza, M. A. Coco and M. C. Lammers [2, 3], and by F. Luef [11] obtaining an adapted to the Gabor situation variant of the Hilbert C*-module approach. In particular, the results by J. A. Packer and M. A. Rieffel in [13] indicate that the described operator algebraic approach to the wavelet theory in $L^2(\mathbb{R}^2)$ is capable to give new deep insights into classical wavelet theory.

To give an instructive example how to link a particular case of generalized multi-resolution analysis to Hilbert C*-module theory we explain one of the core ideas of M. A. Rieffel by example: Assume the situation of a wavelet sequence generated by a multi-resolution analysis in a Hilbert space $L_2(\mathbb{R}^n)$. Denote the mother wavelet by $\phi \in L_2(\mathbb{R}^n)$, $\|\phi\|_2 = 1$, and consider $\mathbb{R}^n$ as an additive group. The second group appearing in the picture is $\Gamma = \mathbb{Z}^n$ acting on $L_2(\mathbb{R}^n)$ by translations in the domains of functions, i.e. mapping $\phi(x)$ to $\phi(x-p)$ for $x \in \mathbb{R}^n$ and $p \in \mathbb{Z}^n$. The mother wavelet ϕ has to be supposed to admit pairwise orthogonal $\mathbb{Z}^n$-translates, i.e. $\int_{\mathbb{R}^n} \overline{\phi(x-q)}\phi(x-p)\,dx = \delta_{qp}$ for any $p, q \in \mathbb{Z}^n$. Introducing the group C*-algebras $A = C^*(\mathbb{Z}^n)$ of the additive discrete group $\mathbb{Z}^n$ into the picture and interpreting the set of all $\mathbb{Z}^n$-translates of ϕ as elements of the $*$-algebra $C_c(\mathbb{R}^n)$ we obtain a right action of A on $C_c(\mathbb{R}^n)$ by convolution and an A-valued inner product there defined by $\langle \phi, \psi \rangle_A(p) := \int_{\mathbb{R}^n} \overline{\phi(x)} \psi(x-p)\, dx$ for $\phi, \psi \in C_c(\mathbb{R}^n)$ and $p \in \mathbb{Z}^n$, (see below for details). The completion of $C_c(\mathbb{R}^n)$ with respect to the norm $\|\phi\| := \|\langle \phi, \phi \rangle_A\|_A^{1/2}$ is a (right) Hilbert C*-module $\mathcal{H} = \overline{C_c(\mathbb{R}^n)}$ over A.

Considering the dual Fourier transformed picture things become mathematically easier. The C*-algebra $A = C^*(\mathbb{Z}^n)$ is transformed to the C*-algebra $B = C(\mathbb{T}^n)$ of continuous functions on the n-torus. The right action of A on $\mathcal{H}$ by convolution becomes a right action of B on $\mathcal{H}$ by pointwise multiplication. Moreover, $\mathcal{H} = \overline{C_c(\mathbb{R}^n)}$ coincides with the set $B\phi$, i.e. it is a singly generated free B-module with B-valued inner product $\langle \phi, \psi \rangle_B(t) := \sum_{p \in \mathbb{Z}^n} (\overline{\phi}\psi)(t-p)$ for $t \in \mathbb{R}^n$. The set $\{\phi\}$ consisting of one element is a module frame, even a module Riesz basis. However, for $n \geq 2$ there exist non-free B-modules that are direct orthogonal summands of $\mathcal{H} = B$, cf. [12] for their construction. For them module Riesz bases might not exist, and module frames consist of more than one element. In a similar manner multi-wavelets give rise to Hilbert B-modules B^k of all k-tuples with entries from B and coordinate-wise operations. Since norm-convergence and weak convergence are in general different concepts in an infinite-dimensional C*-algebra B (whereas both they coincide in $\mathbb{C}$). Some more investigations have to be carried out to treat Gabor analysis, for example.

A *pre-Hilbert C*-module* $\mathcal{H}$ over a (unital) C*-algebra A is a (left) A-module equipped with an *A-valued inner product* $\langle .,. \rangle : \mathcal{H} \times \mathcal{H} \to A$ such that (i) $\langle x, x \rangle \geq 0$ for any $x \in \mathcal{H}$, (ii) $\langle x, x \rangle = 0$ if and only if $x = 0$, (iii) $\langle x, y \rangle = \langle y, x \rangle^*$ for any$x, y \in \mathcal{H}$, and (iv) $\langle .,. \rangle$ is A-linear in the first argument. The induced norm $\|.\| = \|\langle .,. \rangle\|^{1/2}$ opens up the opportunity to restrict attention to norm-closed A-modules of this kind, i.e. to *Hilbert* A-modules. The A-module $\mathcal{H}$ is algebraically finitely generated if there exists a finite set $\{x_i\}_{i=1}^N \subset \mathcal{H}$ such that $\mathcal{H} = \operatorname{span}\{a_i x_i : a_i \in A\}$. A Banach A-module is countably generated if there exists a finite or countable set $\{x_i\}_{i \in I} \subset \mathcal{H}$ such that $\operatorname{span}\{a_i x_i : a_i \in A\}$ is norm-dense in $\mathcal{H}$. For a comprehensive account to Hilbert C*-module theory we refer the reader to [10].

For unital C*-algebras A a finite or countable set $\{x_i\}_{i \in I} \subset \mathcal{H}$ is said to be a *frame* for the Hilbert C*-module $\mathcal{H}$ if there exist two real constants $C, D > 0$ such that the inequality $C \cdot \langle x, x \rangle \leq \sum_{i \in I} \langle x, x_i \rangle \langle x_i, x \rangle \leq D \cdot \langle x, x \rangle$ is valid for any $x \in \mathcal{H}$. The frame is called *standard* if the sum in the middle of the inequality converges in norm in A. A frame is *normalized tight* if $C = D = 1$. A sequence $\{x_i\}_{i \in I} \subset \mathcal{H}$ is a *standard Riesz basis of* $\mathcal{H}$ if it is a standard frame for $\mathcal{H}$ with the additional property that $\sum_{i \in S \subseteq I} a_i x_i = 0$ if and only if $a_i x_i = 0$ for any $i \in S$. Two frames $\{x_i\}_{i \in I}$ and $\{y_i\}_{i \in I}$ for a Hilbert A module $\mathcal{H}$ are *unitarily equivalent* (resp., *similar*) if there exists a unitary (resp., invertible adjointable) A-linear bounded operator T on $\mathcal{H}$ satisfying $T(x_i) = y_i$ for any $i \in I$. By Kasparov's stabilization theorem and by tensor product constructions one can easily see that standard (normalized tight) frames for Hilbert C*-modules over unital C*-algebras exist always and in abundance. For canonical examples of Hilbert C*-modules standard Riesz bases are found not to exist, and so orthogonal Hilbert bases often may not exist.

As the crucial result that makes the entire theory work one obtains two *reconstruction formulae* for standard (normalized tight) frames $\{x_i\}_{i\in I}$ of finitely or countably generated Hilbert C*-modules $\mathcal{H}$ over unital C*-algebras A. If $\{x_i\}_i$ is a standard normalized tight frame for $\mathcal{H}$ then the following reconstruction formula always holds for every $x \in \mathcal{H}$:

$$x = \sum_{i\in I} \langle x, x_i \rangle x_i \, .$$

The sum converges with respect to the norm of $\mathcal{H}$. If $\{x_i\}_{i\in I}$ is merely a standard frame for $\mathcal{H}$ then there exists a positive invertible A-linear bounded operator S on $\mathcal{H}$, the *frame operator*, such that the reconstruction formula

$$x = \sum_{i\in I} \langle x, S(x_i) \rangle x_i$$

is valid for any $x \in \mathcal{H}$. The sequence $\{S(x_i)\}_{i\in I}$ is a frame for $\mathcal{H}$ again, and it is said to be the *canonical dual frame* of for the frame $\{x_i\}_{i\in I}$. The key point of the proofs is the existence of the *frame transform* $\theta : \mathcal{H} \to l_2(A)$, $\theta(x) = \{\langle x, x_i \rangle\}_{i\in I}$, and its properties which can be found to be guaranteed in *any* situation - boundedness, A-linearity, and, most important, adjointability. The frame operator S can be expressed by $S = (\theta\theta^*)^{-1}$, and for every standard frame $\{x_i\}_{i\in I}$ the frame $\{S^{1/2}(x_i)\}_{i\in I}$ turns out to be a standard normalized tight one.

Starting from this point similarity of standard frames and the image of their frame transform can be investigated, leading to similar results about the canonical and alternate duals as in the Hilbert space situation. In the same manner as for Hilbert spaces results for complementary frames and inner sums of frames can be obtained giving rise to several types of disjointness of pairs of frames. Standard frames turn out to be precisely the inner direct summands of standard Riesz bases for Hilbert A-modules A^N, $N < \infty$, or $l_2(A)$. whereas standard normalized tight frames are the inner direct summands of orthonormal Hilbert bases of A^N or $l_2(A)$.

Establishing this key point of the theory of standard modular frames of countably Hilbert C*-modules over unital C*-algebras A one (re-)obtains an whole collection of frame theory results in this setting: Every standard frame of a countably generated Hilbert A-module is a set of generators. Every standard Riesz basis $\{x_i\}_{i\in I}$ with normalized tight frame bounds has the property $\langle x_j, x_k \rangle = \delta_{jk} \cdot \langle x_j, x_k \rangle^2$ for any $j, k \in I$, i.e. it is orthogonal and "normalized" in some sense. Every finite set of algebraic generators of a finitely generated Hilbert A-module is a frame for it. If the equality $x = \sum_{i\in I} \langle x, y_i \rangle x_i$ holds for any $x \in \mathcal{H}$ and for some standard frame $\{y_i\}_{i\in I}$ for $\mathcal{H}$ then this *alternate dual frame* fulfills the inequality

$$\sum_{i\in I} \langle x, S(x_i) \rangle \langle S(x_i), x \rangle < \sum_{i\in I} \langle x, y_i \rangle \langle y_i, x \rangle$$

for any $x \in \mathcal{H}$.

References

[1] D. Bakić and B. Guljaš, Wigner's theorem in Hilbert C*-modules over C*-algebras of compact operators, *Proc. Amer. Math. Soc.* **130**(2002), 2343-2349.

[2] P. G. Casazza, M. C. Lammers, Bracket products for Weyl-Heisenberg frames, in: *Advances in Gabor Analysis*, eds.: H. G. Feichtinger, T. Strohmer, Birkhäuser, Boston, Ma., 2003, 71-98.

[3] M. A. Coco, M. C. Lammers, A Hilbert C*-module for Gabor systems, preprint math.FA/0102165 at www.arxiv.org, 2001.

[4] M. Frank, D. R. Larson, A module frame concept for Hilbert C*-modules, in: *Functional and Harmonic Analysis of Wavelets (San Antonio, TX, Jan. 1999)*, *Contemp. Math.* **247**(2000), 207-233.

[5] M. Frank, D. R. Larson, Modular frames for Hilbert C*-modules and symmetric approximation of frames, in: *SPIE's 45th Annual Meeting, July 30 - August 4, 2000, San Diego, CA, Session 4119: Wavelet Applications in Signal and Image Processing VIII*, *Proc. of SPIE* **4119**(2000), 325-336.

[6] M. Frank, D. R. Larson, Frames in Hilbert C*-modules and C*-algebras, *J. Operator Theory* **48**(2002), 273-314.

[7] Deguang Han, D. R. Larson, Frames, bases and group representations, *Memoirs Amer. Math. Soc.* **147**(2000), no. 697, 94pp..

[8] T. Kajiwara, C. Pinzari, Y. Watatani, Jones index theory for Hilbert C*-modules and its equivalence with conjugation theory, preprint math.OA/0301259 at www.arxiv.org, 2003.

[9] A. Khosravi, A. A. Moslemipour, Modular standard frames in Hilbert A-modules, *Int. Math. J.* **3**(2003), 1139-1147.

[10] E. C. Lance, *Hilbert C*-modules - A toolkit for operator algebraists*, London Math. Soc. Lecture Notes Series v. 210, Cambridge University Press, Cambridge, UK, 1995.

[11] F. Luef, Gabor analysis, Rieffel induction, and Feichtinger's algebra as a link, in: *Workshop on Time-frequency Analysis and Applications, Sept. 22-26, 2003, National University of Singapore, Inst. of Math. Sciences, org.: H. G. Feichtinger, Say Song Goh, Zuowei Shen*, 2003
(www.ims.nus.edu.sg/Programmes/imgsci/files/luef1.pdf).

[12] J. A. Packer, M. A. Rieffel, Wavelet filter functions, the matrix completion problem, and projective modules over $C(\mathbb{T}^n)$, *J. Fourier Anal. Appl.* **9**(2003), 101-106.

[13] J. A. Packer, M. A. Rieffel, Projective multi-resolution analysis for $L^2(\mathbb{R}^2)$, preprint math.OA/0308132 at www.arxiv.org, 2003, submitted to *J. Fourier Anal. Appl.*.

[14] I. Raeburn, S. J. Thompson, Countably generated Hilbert modules, the Kasparov stabilization theorem, and frames in Hilbert modules, *Proc. Amer. Math. Soc.* **131**(2003), 1557-1564.

[15] M. A. Rieffel, Multiwavelets and operator algebras, *Talk given at AMS Special Session*, 1997.

[16] P. J. Wood, Wavelets and C*-algebras, Ph. D. Thesis, The Flinders University of South Australia, Adelaide, Australia, Sept. 2003, 116 pp..

[17] P. J. Wood, Wavelets and projective modules, preprint, The Flinders University of South Australia, Adelaide, Australia, 2001, submitted to *J. Fourier Anal. Appl.*.

Frame Generators and Traces on the Commuting Algebra

Hartmut Führ

Given a representation $(\pi, \mathcal{H}_\pi)$ of a unimodular, separable locally compact group G, we want to discuss the existence and characterization of vectors giving rise to coherent state expansions on $\mathcal{H}_\pi$.

For this purpose, a vector $\eta \in \mathcal{H}_\pi$ is called **bounded** if the coefficient operator

$$V_\eta : \mathcal{H}_\pi \to \mathrm{L}^2(G) \ , \ (V_\eta \varphi)(x) = \langle \varphi, \pi(x)\eta \rangle$$

is a bounded map. A pair of bounded vectors (η, ψ) is called **admissible** if $V_\psi^* V_\eta = \mathrm{Id}_{\mathcal{H}_\pi}$. This property gives rise to the weak-sense inversion formula

$$z = \int_G \langle z, \pi(x)\eta \rangle \ \pi(x)\psi \ d\mu_G(x) \ ,$$

which can be read as a continuous expansion of z in terms of the orbit $\pi(G)\psi \subset \mathcal{H}_\pi$. A single vector η is called admissible if (η, η) is an admissible pair. It is obvious from the definition that (η, ψ) is admissible iff (ψ, η) is. In such a case η is called the **dual vector** of ψ.

If G is a discrete group, the notions of bounded vectors and admissible pairs can be reformulated in terms of frames: Rewriting the inversion formula as

$$z = \sum_{x \in G} \langle z, \pi(x)\eta \rangle \ \pi(x)\psi \ ,$$

we see that the (η, ψ) are an admissible pair iff the systems $\pi(G)\eta$ and $\pi(G)\psi$ are a dual frame pair of $\mathcal{H}_\pi$. We want to discuss representation-theoretic criteria for frame generators. The study of discrete groups necessitates to go beyond the so-called *discrete series* or *square-integrable* representations [3], but also beyond the type I groups studied in [2].

It turns out that more general statements are possible by use of a particular trace on the right von Neumann algebra $VN_r(G)$, which is the commutant of the left regular representation λ_G on $\mathrm{L}^2(G)$. Indeed, the following observations can be made:

1. Up to unitary equivalence, any representation π having an admissible pair can be realized as a subrepresentation of λ_G, acting on some leftinvariant closed subspace $\mathcal{H} \subset \mathrm{L}^2(G)$. In particular, the projection onto $\mathcal{H}$ is in $VN_r(G)$.
2. Defining $f^*(x) = \overline{f(x^{-1})}$, the coefficient operators acting on $\mathrm{L}^2(G)$ (or subspaces) can be written as $V_f g = g * f^*$.
3. $VN_r(G)$ carries a natural faithful normal, semifinite trace defined for positive operators S

$$t_r(S) = \begin{cases} \|f\|_2 & : S = V_f^* V_f \text{ for a suitable bounded vector } f \in \mathrm{L}^2(G) \\ \infty & : \text{otherwise} \end{cases}$$

   Polarisation of the definition yields for bounded vectors

$$t_r(V_g^* V_f) = \langle g, f \rangle \ .$$

4. If G is discrete, any $T \in VN_r(G)$ is uniquely determined by its "impulse response" $T(\delta_e)$. In this case t_r is finite and given by

$$t_r(T) = T(\delta_e)(e) \ .$$

Given a particular trace tr on a von Neumann algebra $\mathcal{A}$, we call a pair of elements (η, ψ) of the underlying Hilbert space **tracial** if

$$\forall T \in \mathcal{A}^+ : tr(T) = \langle T\eta, \psi \rangle \ .$$

Then we have

Theorem 1. *Let $\mathcal{H} \subset \mathrm{L}^2(G)$ be a closed, leftinvariant subspace, with associated leftinvariant projection p, and let π denote the restriction of λ_G to $\mathcal{H}$.*

(a) There exists an admissible pair for $\mathcal{H}$ iff $t_r(p) < \infty$.

(b) For all pairs $(\eta, \psi) \in \mathcal{H} \times \mathcal{H}$ of bounded vectors: (η, ψ) is admissible iff (η, ψ) is tracial for $\pi(G)'$.

We shortly sketch two applications. The first concerns the central decomposition of λ_G. Let $\check{G}$ denote the space of quasi-equivalence classes of factor representations of G, and let

$$\lambda_G \simeq \int_{\check{G}}^{\oplus} \rho_\sigma d\nu_G(\sigma)$$

denote the central decomposition. This also provides the direct integral decompositions

$$\begin{aligned} VN_r(G) &\simeq \int_{\check{G}}^{\oplus} \mathcal{A}_\sigma d\nu_G(\sigma) \\ t_r(T) &= \int_{\check{G}} tr_\sigma(T_\sigma) d\nu_G(\sigma) \ , \end{aligned}$$

where A_σ is the commuting algebra of ρ_σ, $(T_\sigma)_{\sigma \in \check{G}}$ denotes the operator field corresponding to T under the central decomposition and tr_σ is a suitable faithful normal, semifinite trace on the factor $\mathcal{A}_\sigma$. Standard direct integral arguments then yield:

Proposition 2. *Let π denote the restriction of λ_G to a closed, leftinvariant subspace $\mathcal{H} \subset \mathrm{L}^2(G)$. Let P denote the projection onto $\mathcal{H}$, then P decomposes into a measurable field of projections $\widehat{P}_\sigma$, and $\pi(G)'$ decomposes under the central decomposition into the von Neumann algebras $\mathcal{C}_\sigma = \widehat{P}_\sigma \mathcal{A}_\sigma \widehat{P}_\sigma$.*

(a) For bounded $\eta, \psi \in \mathcal{H}$, we have

$$(\eta, \psi) \text{ is admissible for } \mathcal{H} \ \Leftrightarrow (\widehat{\eta}_\sigma, \widehat{\psi}_\sigma) \text{ is tracial for } \mathcal{C}_\sigma \ (\nu_G \, a.e.)$$

(b) $\mathcal{H}$ has an admissible pair of vectors iff $\int_{\check{G}} tr(\widehat{P}_\sigma) d\nu_G(\sigma) < \infty$. In particular, almost all C_σ are finite von Neumann algebras.

Generally the representations of interest are not realized as acting by left translations on subspaces of $\mathrm{L}^2(G)$. Therefore, applying Theorem 1 requires first to embed the representation into λ_G. The following corollary sketches an alternative approach. Roughly speaking, it derives a criterion for admissible pairs based on **one** explicitly known admissible pair.

Corollary 3. *Suppose we are given*

- *A family $(T_i)_{i \in I} \subset \pi(G)'$ spanning a weak-operator dense subspace of $\pi(G)'$.*
- *An admissible pair (η_0, ψ_0).*

Then for a pair of bounded vectors (η, ψ) *we have the following equivalence:*

(1) $$(\eta, \psi) \text{ is admissible } \iff \forall i \in I \ : \ \langle T_i \eta, \psi \rangle = \langle T_i \eta_0, \psi_0 \rangle \ .$$

The criterion is explicit as soon as the T_i and the admissible pair (η_0, ψ_0) are known explicitly. Using results from [1] it can be shown that the Wexler-Raz criteria for Gabor frames can be derived this way, thus yielding **explicit** criteria for a whole family of type-II representations.

References

[1] I. Daubechies, H.J. Landau and Z. Landau, *Gabor time-frequency lattices and the Wexler-Raz identity,* J. Fourier Analysis and Applications **1** (1995), 437-478.

[2] H. Führ, *Admissible vectors for the regular representation,* Proc. AMS **130** (2002), 2959-2970.

[3] A. Grossmann, J. Morlet and T. Paul, *Transforms associated to square integrable group representations I: General results*, J. Math. Phys. **26** (1985), 2473-2479.

Frames, Operators, and Banach Algebra Techniques

Karlheinz Gröchenig

Symbolic Calculus. A symbolic calculus is a mapping from a class of *symbols* to a class of *operators* acting on some Hilbert space (or subspace thereof):

$$\sigma \longrightarrow \mathrm{Op}(\sigma)$$

In many areas of mathematics one finds manifestations of the following principle.

Metatheorem. *If the symbol* σ *is nice and* $\mathrm{Op}(\sigma)$ *is invertible on Hilbert space, then* $\big(\mathrm{Op}(\sigma)\big)^{-1} = \mathrm{Op}(\tau)$ *for nice* τ*.*

An important consequence is the following extension principle.

Meta-Corollary. $\big(\mathrm{Op}(\sigma)\big)^{-1} = \mathrm{Op}(\tau)$ *is bounded on large class of Banach spaces.*

We give several examples of a symbolic calculus drawn from different fields of mathematics. Usually a symbolic calculus is proved by means of some "hard analysis", but we will emphasize the role of Banach algebra techniques in the analysis of symbolic calculi. A second aspect is the role of weights. Weighted versions of symbolic calculus can usually be derived from the corresponding unweighted versions and the growth properties of the weights.

1. Convolution Operators on Groups.
The prototype of a symbolic calculus is **Wiener's Lemma.** In its standard form asserts the following: If f has an non-vanishing absolutely convergent Fourier series, then so does $1/f$.

Wiener's Lemma can be recast as a statement about convolution operators defined by $T_{\mathbf{a}}\mathbf{c} = \mathbf{a} * \mathbf{c}$ for two sequences $\mathbf{a}, \mathbf{c}$ on $\mathbb{Z}^d$. In this case the symbol is the sequence $\mathbf{a}$ and the operator is $T_{\mathbf{a}}$. "Nice" symbols are sequences in the weighted ℓ^1 space $\ell^1_v(\mathbb{Z}^d)$ by the norm $\|\mathbf{a}\|_{\ell^1_v} = \sum_{k\in\mathbb{Z}^d} |c_k| v(k)$. The weight is always assumed to satisfy $v(0) = 1, v(k) = v(-k)$, and $v(k+l) \le v(k)v(l), k, l \in \mathbb{Z}^d$.

Theorem 1. *Assume that*
 (a) $\mathbf{a} \in \ell^1_v(\mathbb{Z}^d)$,
 (b) $T_{\mathbf{a}}$ *is invertible on* $\ell^2(\mathbb{Z}^d)$ *and*
 (c) $\lim_{n\to\infty} v(nx)^{1/n} = 1, \forall x \in \mathbb{Z}^d$ *(GRS-condition).*
Then $T_{\mathbf{a}}^{-1} = T_{\mathbf{b}}$ *for* $\mathbf{b} \in \ell^1_v(\mathbb{Z}^d)$ [2].

Let $\sigma_{\ell^p_m}(\mathbf{a})$ be the spectrum of the convolution operator $T_{\mathbf{a}}$ on the weighted ℓ^p-space $\ell^p_m(\mathbb{Z}^d)$. Then we have

Corollary 2. *If* $m(x+y) \le Cv(x)m(y)$, *then*

$$\sigma_{\ell^p_m}(\mathbf{a}) = \sigma_{\ell^2}(\mathbf{a})$$

The role of the GRS condition is illuminated by the following statement.

Theorem 3.

$$\sigma_{\ell^1_v}(\mathbf{a}) = \sigma_{\ell^2}(\mathbf{a})$$

if and only if v *satisfies the GRS-condition* $\lim_{n\to\infty} v(nx)^{1/n} = 1, \forall x \in \mathbb{Z}^d$.

Similar types of a symbolic calculus can be shown for "twisted convolution", for the rotation algebra [4], and for convolution operators on groups of polynomial growth [1, 6].

2. Matrix Algebras.
The second type of example concerns matrix algebras. In this case the "symbol" is an infinite matrix A, the associated operator is obtained simply by the action of A on a sequence c. "Nice" matrices are determined by their decay off the diagonal.

Theorem 4. [3, 5]. *Assume that* u *is a radial weight function on* $\mathbb{Z}^d$ *satisfying the GRS-condition and that* $v(x) = u(x)(1+|x|)^s$ *for some* $s > d$. *If the matrix* A *invertible on* $\ell^2(\mathbb{Z}^d)$ *and if*

$$|A_{kl}| \le Cv(k-l)^{-1}\,,$$

then

$$|(A^{-1})_{kl}| \le C' v(k-l)^{-1}$$

and

$$\sigma_{\mathcal{A}^1_v}(A) = \sigma(A) \qquad \forall A \in \mathcal{A}^1_v$$

where $\sigma(A)$ *is the spectrum of* A *as an operator on* ℓ^2.

As a consequence, A and A^{-1} are bounded on many weighted ℓ^p-spaces.

Theorems of this type are important in numerical analysis because they are used in error estimates, when infinite-dimensional matrix equations are approximated by finite-dimensional models (finite section method).

3. Self-Localized Frames.

In the final example the "symbols" are frames $\mathcal{E} = \{e_x : x \in \mathcal{X}\}$ and the associated operator is the frame operator $Sf = S_{\mathcal{E}}f = \sum_{x\in\mathcal{X}} \langle f, e_x\rangle e_x$. In the context of symbolic calculus, "nice" frames are frames with a localization property.

Definition: A frame $\{e_x : x \in \mathcal{X}\}$ is intrinsically s-self-localized, if

$$|\langle e_y, e_x\rangle| \leq C(1 + |x - y|)^{-s} \qquad \forall x, y \in \mathcal{X}\,.$$

Theorem 5 (Fornasier, Gröchenig, 2004)**.** *If $\{e_x : x \in \mathcal{X}\}$ is s-self-localized, then so is the canonical dual frame $\{\tilde{e}_x\}$, i.e.,*

$$|\langle \tilde{e}_y, \tilde{e}_x\rangle| \leq C(1 + |x - y|)^{-s} \qquad x, y \in \mathbb{R}^d$$

and

$$|\langle e_y, \tilde{e}_x\rangle| \leq C(1 + |x - y|)^{-s} \qquad x, y \in \mathbb{R}^d$$

This statement has wide applications in sampling theory, time-frequency analysis, and wavelet theory.

As further examples of a symbolic calculus we mention pseudodifferential operators and their spectral invariance on various function spaces, and new classes of matrix algebras that are dominated by a convolution operator.

All the above examples can be viewed as statements about the symmetry and inverse-closedness of the Banach algebra under discussion.

An involutive Banach algebra $\mathcal{A}$ is symmetric, if $\sigma(a^*a) \subseteq [0, \infty)$ for all $a \in \mathcal{A}$ (if and only if $\sigma(a) \subseteq \mathbb{R}$ for all $a = a^* \in \mathcal{A}$). Theorems 3 and 4 assert that $(\ell^1_v, *)$ and $\mathcal{A}_v$ are symmetric Banach algebras.

Another central concept is inverse-closedness. Let $\mathcal{A} \subseteq \mathcal{B}$ be two Banach algebras with a common identity. Then $\mathcal{A}$ is said to be inverse-closed in $\mathcal{B}$, if

$$a \in \mathcal{A} \text{ and } a^{-1} \in \mathcal{B} \quad \Longrightarrow \quad a^{-1} \in \mathcal{A}\quad .$$

Other terminology frequently used is that of a Wiener pair, a spectral subalgebra, or of spectral invariance. Theorems 1 and 4 state that $(\ell^1_v, *)$ is inverse-closed in ℓ^1 and $\mathcal{B}(\ell^2)$, and that $\mathcal{A}_v$ is inverse-closed in $\mathcal{B}(\ell^2)$.

References

[1] G. Fendler, K. Gröchenig, M. Leinert, J. Ludwig, and C. Molitor-Braun. Weighted group algebras on groups of polynomial growth. *Math. Z.*, 102(3):791 – 821, 2003.

[2] I. Gel'fand, D. Raikov, and G. Shilov. *Commutative normed rings.* Chelsea Publishing Co., New York, 1964.

[3] K. Gröchenig and M. Leinert. Symmetry of matrix algebras and symbolic calculus for infinite matrices. *Preprint*, 2003.

[4] K. Gröchenig and M. Leinert. Wiener's lemma for twisted convolution and Gabor frames. *J. Amer. Math. Soc.*, 17:1–18, 2004.

[5] S. Jaffard. Propriétés des matrices "bien localisées" près de leur diagonale et quelques applications. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 7(5):461–476, 1990.
[6] V. Losert. On the structure of groups with polynomial growth. II. *J. London Math. Soc. (2)*, 63(3):640–654, 2001.

The Zero Divisor Conjecture for the Heisenberg Group

Christopher Heil

The following conjecture was introduced in the paper [HRT96], and is still open today.

Conjecture 1. If $g \in L^2(\mathbf{R})$ is nonzero and $\{(\alpha_k, \beta_k)\}_{k=1}^N$ is any set of N distinct points in $\mathbf{R}^2$, then $\{e^{2\pi i \beta_k x} g(x - \alpha_k)\}_{k=1}^N$ is a linearly independent set of functions in $L^2(\mathbf{R})$.

The composition $M_b T_a g(x) = e^{2\pi i b x} g(x - a)$ of translation $T_a g(x) = g(x - a)$ and modulation $M_b g(x) = e^{2\pi i b x} g(x)$ is called a time-frequency shift of g, and the analysis and application of these operators is time-frequency analysis. A beautiful introduction to time-frequency analysis can be found in [Grö01]. Conjecture1 has many connections, to harmonic analysis, representation theory, functional analysis, the geometry of Banach spaces, and even more unexpected areas such as ergodic theory.

Today Conjecture 1 sometimes goes by the name of the *HRT Conjecture* or the *Zero Divisor Conjecture for the Heisenberg Group.* Despite attacks by a number of groups, the only published results specifically concerning the conjecture appear to be [HRT96], [Lin99], and [Kut02], which can be summarized as follows.

The paper [HRT96] introduced the conjecture and obtained some partial results, including the following.

(a) If a nonzero $g \in L^2(\mathbf{R})$ is compactly supported, or just supported on a half-line, then the independence conclusion holds for any value of N.

(b) The independence conclusion holds for any a nonzero $g \in L^2(\mathbf{R})$ if $N \leq 3$.

(c) If the independence conclusion holds for a particular $g \in L^2(\mathbf{R})$ and a particular choice of points $\{(\alpha_k, \beta_k)\}_{k=1}^N$, then there exists an $\varepsilon > 0$ such that it also holds for any h satisfying $\|g - h\|_2 < \varepsilon$, using the same set of points.

(d) If the independence conclusion holds for one particular $g \in L^2(\mathbf{R})$ and particular choice of points $\{(\alpha_k, \beta_k)\}_{k=1}^N$, then there exists an $\varepsilon > 0$ such that it also holds for that g and any set of points in $\mathbf{R}^2$ within ε of the original ones.

Another partial advance was made by Linnell in [Lin99]. He used C^*-algebra techniques to prove that if the points $\{(\alpha_k, \beta_k)\}_{k=1}^N$ are a subset of some translate of a lattice in $\mathbf{R}^2$, then the independence conclusion holds for any g (a lattice is a set of the form $A(\mathbf{Z}^2)$, the image of $\mathbf{Z}^2$ under an invertible matrix A). Note that any three points in the plane always lie on a translate of some lattice, so this recovers and extends the partial result (b) mentioned above. However, given four arbitrary points in the plane it is not always possible to find a translate of a lattice that contains those points. Indeed, the case $N = 4$ of the conjecture is still open. In fact, the following special case seems to be open.

Conjecture 2. If $g \in L^2(\mathbf{R})$ is nonzero then

$$\{g(x),\, g(x-1),\, g(x-\sqrt{2}),\, e^{2\pi i x} g(x)\}$$

is a linearly independent set of functions in $L^2(\mathbf{R})$.

Conjecture 2 remains open even if we impose the condition that g be continuous. The real-valued version obtained by replacing $e^{2\pi i x}$ by $\sin 2\pi x$ is likewise open.

One motivation for Conjecture 1 comes from looking at *frames*, which are possibly redundant or over-complete collections of vectors in a Hilbert space which nonetheless provide basis-like representations of vectors in the space. Thus a frame "spans" the space in some sense, even though it may be "dependent." However, in infinite dimensions there are many shades of gray to the meanings of "spanning" and "independence." Some of the most important frames are "dependent" taken as a whole even though have the property that every finite subset is linearly independent. One motivation for Conjecture 1 is the question of whether the the special class of *Gabor frames* have this property that every finite subset is independent.

Gabor frames are related to the Schrödinger representation of the Heisenberg group. If we instead use the affine group and the standard representation induced from dilations and translations, we obtain wavelets. However, the analogue of Conjecture 1 for wavelets fails in general. For example, a compactly supported refinable function φ satisfies an equation of the form

$$\varphi(x) = \sum_{k=0}^{N} c_k\, \varphi(2x-k).$$

This is an expression of linear dependence among the time-scale translates of φ. In particular, the box function $b = \chi_{[0,1)}$ satisfies the refinement equation

$$b(x) = b(2x) + b(2x-1).$$

The more general analogue of Conjecture 1 for the case of other groups is related to the Zero Divisor Conjecture in algebra; we refer to [Lin99] and the references therein for more on this connection.

References

[Grö01] K. Gröchenig, "Foundations of Time-Frequency Analysis," Birkhäuser, Boston, 2001.

[HRT96] C. Heil, J. Ramanathan, and P. Topiwala, Singular values of compact pseudodifferential operators, *J. Funct. Anal.*, **150** (1996), pp. 426–452.

[Kut02] G. Kutyniok, Linear independence of time-frequency shifts under a generalized Schrödinger representation, *Arch. Math. (Basel)*, **78** (2002), pp. 135–144.

[Lin99] P. A. Linnell, Von Neumann algebras and linear independence of translates, *Proc. Amer. Math. Soc.*, **127** (1999), pp. 3269–3277.

Poisson Wavelet Frames on the Sphere

Ilona Ilgewska–Nowak
(joint work with Matthias Holschneider)

People would like to create a mathematical repesentation of the Earth's magnetic field and how it is changing. One of the most broadly used possibilities is to represent the magnetic field in terms of spherical harmonics. This method has some disadvantages. One of them is a poor localization: coefficients obtained in Europe have influence on the representation of the field over Africa. On the other hand, it is difficult to distinguish the big–scale field component from the core and the small–scale field component from the crust. Moreover, changing the truncation level of spherical harmonics changes all the coefficients, according to spatial aliasing of the higher–order harmonics.

Another possibility is to use a wavelet representation of the magnetic field. This would solve some of the problems mentioned above. Here, we would like to introduce Poisson wavelets and give some ideas how frames of such wavelets could be constructed.

Note that there exists no natural dilation operator on the sphere, hence, we do not have a group structure of the wavelet coefficients. Here, the scales are defined in a more or less *ad hoc* way, but so that the wavelets behave like wavelets over the plane. The definition we use goes back to [2], in this talk we base on the simplified definition given in [3].

If Σ denotes the unit two–dimensional sphere, $\hat{e}$ the unit vector in direction of the north–pole, then Poisson wavelets are defined to be

$$g_a^n(x) = \sum_{l=0}^{n} (al)^n e^{-al} Q_l(x),$$

where $Q_l(x) = \frac{2l+1}{4\pi} P_l(x \cdot \hat{e})$, P_l – l-th Legendre polynomial. They are equal to the electromagnetic field caused by a sum of multipoles inside the unit ball:

$$g_a^n = a^n (2\Psi_{e^{-a}}^{n+1} + \Psi_{e^{-a}}^n), \quad \text{where } \Delta\Psi_\lambda^n = (\lambda\partial_\lambda)\delta_{\lambda\hat{e}}$$

(therefore the name Poisson wavelets.)

We obtain explicit expressions in terms of finite sums of Legendre polynomials if we develop g_a^n around the point $e^{-a}\hat{e}$:

$$g_a^n(x) = a^n \sum_{k=1}^{n+1} k!(2C_k^{n+1} + C_k^n) e^{-ka} P_k(\cos\chi) \frac{1}{|x - e^{-a}\hat{e}|^{k+1}},$$

where χ is the angle between $\hat{e}$ and $x - e^{-a}\hat{e}$, and C_k^n are constants defined through

$$(\lambda\partial_\lambda)^n = \sum C_k^n \lambda^k \partial_\lambda^k.$$

For small scales a the Euklidean limit holds:

$$\lim_{a\to 0} a^2 g_a^n(\Phi^{-1}(ax)) = g(x) \quad \text{for some } g \in \mathcal{L}^2(\mathbb{R}^2),$$

where Φ is the stereographic projection of the sphere onto the plane. This means that g_a^n are scaling like wavelets over $\mathbb{R}^2$ assymptitically for small a.

The wavelet transform of a function s is given by

$$\mathcal{W}_{g^n} s(x,a) = \int_\Sigma g_a^n(x\cdot y)s(y)d\sigma(y)$$

and the inverse wavelet transform is given by

$$\mathcal{M}_{g^n} r(x) = \int_{\mathbb{R}_+}\int_\Sigma r(y,a) g_a^n(y) d\sigma(y)\frac{da}{a}.$$

The following holds:

$$\mathcal{M}_{g^n}\mathcal{W}_{g^n} s = cs$$

for some constant $c = c(g^n)$, i.e., g^n build a continuous frame.

Remark: the wavelet transform with respect to this family can also be obtained as follows: take s as Dirichlet boundary data for the interior problem. Then apply a suitable radial derivative to the harmonic extension inside the unit ball.

The image of $\mathcal{W}$ is a Hilbert space with reproducing kernel. This reproducing kernel can be written in terms of the wavelets:

$$P_{g^n}(x,a;y,b) = \left(\frac{ab}{(a+b)^2}\right)^n g_{a+b}^{2n}(x\cdot y)$$

(if we identify $g(x)$ with $g(x\cdot\hat{e})$ for zonal functions g.)

In applications in geophysics this continuous family has to be discretized over some grid. We consider the following grid $\Lambda = \{(x,a)\}$ in $\Sigma\times\mathbb{R}_+$: for a fixed scale $a \in \{n\cdot 2^{-j},\, j\in\mathbb{N}_0\}$ (n – order of the wavelet) we take a cube centered with respect to the sphere , divide each of its six sides into 4^j similar squares and project the centers of the faces onto the sphere in order to define positions x. Question: is $A = \{g_{x,a},\,(x,a)\in\Lambda\}$ a frame for $\mathcal{L}^2(\Sigma)$ (for some set of weights $\mu(x,a)$)? Some approaches we have considered are:
(1) based on the atomic space decomposition of [1]: if

$$\begin{aligned}&|\sum_{(y,b)\in\Lambda} P_{g^n}(x,a;y,b)P_{g^n}(y,b;z,c)\mu(y,b)\\ &-\int_{\mathbb{R}_+}\int_\Sigma P_{g^n}(x,a;y,b)P_{g^n}(y,b;z,c)d\sigma(y)\frac{db}{b}| \le \frac{1}{c^2} f\left(\frac{x\cdot z}{c},\frac{a}{c}\right)\end{aligned}$$

for some function f which is $\mathcal{L}^2$-integrable with respect to $\theta d\theta da/a$, then A is a frame;
(2) transform the unit ball onto the upper half–plane (essentially by the Kelvin–transform) such that harmonic functions remain harmonic functions; consider the image of Poisson wavelets under this map and check if they build a frame of the weighted $\mathcal{L}^2(\mathbb{R}^2)$;

(3) based on quasi-frames: locally around each point of the sphere we obtain a quasi-frame; these however have to be patched together to a global frame.

In view of the remark above, having proven that $\{g_\lambda, \lambda \in \Lambda\}$ and alike grids build a frame for $\mathcal{L}^2(\Sigma)$, we automatically obtain some interesting results for harmonic functions (e.g. density of local maxima, sets of uniqueness, ...).

References

[1] H.G. Feichtinger, K.H. Gröchenig, Banach spaces related to integrable group representations and their atomic decompositions, *J. Funct. Anal.* (1989) **86**, 307-340

[2] M. Holschneider, Wavelet Analysis on the Sphere, *J. Math. Phys.* (1996) **37**, 4156-4165

[3] M. Holschneider, A. Chambodut, M. Mandea, From global to regional analysis of scalar and vector fields on the sphere using wavelet frames, *P.E.P.I* (2003) **135**, 107-124

Duality Principles in Analysis

Palle E. T. Jorgensen

Several versions of spectral duality are presented. On the two sides we present (1) a basis condition, with the basis functions indexed by a frequency variable, and giving an orthonormal basis; and (2) a geometric notion which takes the form of a tiling, or a Iterated Function System (IFS). Our initial motivation derives from the Fuglede conjecture, see [3, 6, 7]: For a subset D of $\mathbb{R}^n$ of finite positive measure, the Hilbert space $L^2(D)$ admits an orthonormal basis of complex exponentials, i.e., D admits a Fourier basis with some frequencies L from $\mathbb{R}^n$, if and only if D tiles $\mathbb{R}^n$ (in the measurable category) where the tiling uses only a set T of vectors in $\mathbb{R}^n$. If some D has a Fourier basis indexed by a set L, we say that (D, L) is a spectral pair. We recall from [9] that if D is an n-cube, then the sets L in (1) are precisely the sets T in (2). This begins with work of Jorgensen and Steen Pedersen [9] where the admissible sets $L = T$ are characterized. Later it was shown, [5] and [10] that the identity $T = L$ holds for all n. The proofs are based on general Fourier duality, but they do not reveal the nature of this common set $L = T$. A complete list is known only for $n = 1$, 2, and 3, see [9].

We then turn to the scaling IFS's built from the n-cube with a given expansive integral matrix A. Each A gives rise to a fractal in the small, and a dual discrete iteration in the large. In a different paper [8], Jorgensen and Pedersen characterize those IFS fractal limits which admit Fourier duality. The surprise is that there is a rich class of fractals that do have Fourier duality, but the middle third Cantor set does not. We say that an affine IFS, built on affine maps in $\mathbb{R}^n$ defined by a given expansive integral matrix A and a finite set of translation vectors, admits Fourier duality if the set of points L, arising from the iteration of the A-affine maps in the large, forms an orthonormal Fourier basis (ONB) for the corresponding fractal μ in the small, i.e., for the iteration limit built using the inverse contractive maps, i.e., iterations of the dual affine system on the inverse matrix A^{-1}. By "fractal in the small", we mean the Hutchinson measure μ and its compact support, see [4].

(The best known example of this is the middle-third Cantor set, and the measure μ whose distribution function is corresponding Devil's staircase.)

In other words, the condition is that the complex exponentials indexed by L form an ONB for $L^2(\mu)$. Such duality systems are indexed by complex Hadamard matrices H, see [9] and [8]; and the duality issue is connected to the spectral theory of an associated Ruelle transfer operator, see [1]. These matrices H are the same Hadamard matrices which index a certain family of quasiperiodic spectral pairs (D, L) studied in [6] and [7]. They also are used in a recent construction of Terence Tao [11] of a Euclidean spectral pair (D, L) in $\mathbb{R}^5$ for which D does not a tile $\mathbb{R}^5$ with any set of translation vectors T in $\mathbb{R}^5$.

We finally report on joint research with Dorin Dutkay where we show that all the affine IFS's admit wavelet orthonormal bases [2] now involving both the $\mathbb{Z}^n$ translations and the A-scalings.

References

[1] O. Bratteli, P. Jorgensen, Wavelets through a Looking Glass: The World of the Spectrum, Applied and Numerical Harmonic Analysis, Birkhäuser, Boston, 2002.
[2] D. Dutkay, P. Jorgensen, Wavelets on fractals, preprint June 2003, Univ. of Iowa, submitted to Rev. Mat. Iberoamericana.
[3] B. Fuglede, Commuting self-adjoint partial differential operators and a group-theoretic problem, J. Funct. Anal. 16 (1974), 101–121.
[4] J.E. Hutchinson, Fractals and self similarity, Indiana Univ. Math. J. 30 (1981), 713–747.
[5] A. Iosevich, S. Pedersen, Spectral and tiling properties of the unit cube, Internat. Math. Res. Notices 1998 (1998), no. 16, 819–828.
[6] P. Jorgensen, Spectral theory of finite-volume domains in $\mathbb{R}^n$, Adv. in Math. 44 (1982), 105–120.
[7] P. Jorgensen, S. Pedersen, Spectral theory for Borel sets in $\mathbb{R}^n$ of finite measure, J. Funct Anal. 107 (1992), 72–104.
[8] P. Jorgensen, S. Pedersen, Dense analytic subspaces in fractal L^2-spaces, J. Analyse Math. 75 (1998), 185–228.
[9] P. Jorgensen, S. Pedersen, Spectral pairs in Cartesian coordinates, J. Fourier Anal. Appl. 5 (1999), 285–302.
[10] J.C. Lagarias, J.A. Reeds, Y. Wang, Orthonormal bases of exponentials for the n-cube, Duke Math. J. 103 (2000), 25–37.
[11] T. Tao, Fuglede's conjecture is false in 5 and higher dimensions, preprint, June 2003, http://arxiv.org/abs/math.CO/0306134 .

Minimal Generator Sets for Finitely Generated Shift Invariant Subspaces of $L^2(\mathbb{R}^n)$

Norbert Kaiblinger
(joint work with Marcin Bownik)

Given a family of functions $\phi_1, \ldots, \phi_N \in L^2(\mathbb{R}^n)$, let $S = S(\phi_1, \ldots, \phi_N)$ denote the closed subspace of $L^2(\mathbb{R}^n)$ generated by their integer translates. That is, S is the closure of the set of all functions f of the form

$$f(t) = \sum_{j=1}^{N} \sum_{k\in\mathbb{Z}^n} c_{j,k}\ \phi_j(t-k), \qquad t \in \mathbb{R}^n, \tag{1}$$

where finitely many $c_{j,k} \in \mathbb{C}$ are nonzero. By construction, these spaces $S \subset L^2(\mathbb{R}^n)$ are invariant under shifts, i.e., integer translations and they are called finitely generated shift-invariant spaces. Shift-invariant spaces play an important role in analysis, most notably in the areas of spline approximation, wavelets, Gabor (Weyl-Heisenberg) systems, subdivision schemes and uniform sampling. The structure of this type of spaces is analyzed in [1], see also [2, 3, 4, 9]. Only implicitly we are concerned with the dependence properties of sets of generators, for details on this topic we refer to [7, 8]. [3]

The minimal number $L \leq N$ of generators for the space S is called the length of S. Although we include the case $L = N$, our results are motivated by the case $L < N$. In this latter case, there exists a smaller family of generators $\psi_1, \ldots, \psi_L \in L^2(\mathbb{R}^n)$ such that

$$S(\phi_1, \ldots, \phi_N) = S(\psi_1, \ldots, \psi_L), \qquad \text{with } L < N.$$

Since the new generators $\psi_1, \ldots, \psi_L$ belong to S, they can be approximated in the L^2-norm by functions of the form (1), i.e., by finite sums of shifts of the original generators. However, we prove that at least one reduced set of generators can be obtained from a linear combination of the original generators without translations. In particular, no limit or infinite summation is required. In fact, we show that almost every such linear combination yields a valid family of generators. On the other hand, we show that those combinations which fail to produce a generator set can be dense. That is, combining generators can be a sensitive procedure.

Let $M_{N,L}(\mathbb{C})$ denote the space of complex $N \times L$ matrices endowed with the product Lebesgue measure of $\mathbb{C}^{NL} \cong \mathbb{R}^{2NL}$.

Theorem. *Given $\phi_1, \ldots, \phi_N \in L^2(\mathbb{R}^n)$, let $S = S(\phi_1, \ldots, \phi_N)$ and let $L \leq N$ be the length of S. Let $\mathscr{R} \subset M_{N,L}(\mathbb{C})$ denote the set of those matrices $\Lambda = (\lambda_{j,k})_{1\leq j\leq N,\, 1\leq k\leq L}$ such that the linear combinations $\psi_k = \sum_{j=1}^N \lambda_{j,k}\, \phi_j$, for $k = 1, \ldots, L$, yield $S = S(\psi_1, \ldots, \psi_L)$.*
(i) *Then $\mathscr{R} = M_{N,L}(\mathbb{C}) \setminus \mathscr{N}$, where $\mathscr{N}$ is a null-set in $M_{N,L}(\mathbb{C})$.*
(ii) *The set $\mathscr{N}$ in* (i) *can be dense in $M_{N,L}(\mathbb{C})$.*

Remark. (i) The conclusions of the Theorem also hold when the complex matrices $M_{N,L}(\mathbb{C})$ are replaced by real matrices $M_{N,L}(\mathbb{R})$.
(ii) We note that our results are not restricted to the case of compactly supported generators.

We illustrate the Theorem by an example in the special case of $N = 2$ given generators for a principal shift-invariant space, i.e., $L = 1$. In this case, $M_{N,L}(\mathbb{C})$

[3] Our results presented here are available in more detail in the form of a preprint.

reduces to $\mathbb{C}^2$. We use the following normalization for the Fourier transform,

$$\widehat{f}(x) = \int_{\mathbb{R}} f(t)e^{-2\pi i t x}dt, \qquad x \in \mathbb{R}.$$

Example. For $x \in \mathbb{R}$, let $\lfloor x \rfloor$ denote the largest integer less or equal x. We define a discretized version of the Archimedean spiral by $\gamma\colon [0,1) \to \mathbb{Z}^2$,

$$\gamma(x) = (\lfloor u \cos 2\pi u \rfloor, \lfloor u \sin 2\pi u \rfloor), \qquad u = \tan \tfrac{\pi}{2} x, \quad x \in [0,1).$$

Next, let

$$\gamma^\circ(x) = \begin{cases} \gamma(x)/|\gamma(x)|, & \text{if } \gamma(x) \neq 0, \\ 0, & \text{otherwise,} \end{cases} \qquad x \in [0,1).$$

Now define $\phi_1, \phi_2 \in L^2(\mathbb{R})$ by their Fourier transforms, obtained from $\gamma^\circ = (\gamma_1^\circ, \gamma_2^\circ)$ by

$$\widehat{\phi}_j(x) = \begin{cases} \gamma_j^\circ(x), & x \in [0,1), \\ 0, & x \in \mathbb{R} \setminus [0,1), \end{cases} \qquad j = 1, 2.$$

Let $S = S(\phi_1, \phi_2)$. Then S is principal. In fact, the function $\psi = \lambda_1\phi_1 + \lambda_2\phi_2$ is a single generator, $S = S(\psi)$, if and only if λ_1 and λ_2 are rationally linearly independent. So here the set $\mathscr{N}$ of the Theorem is

$$\mathscr{N} = \{(\lambda_1, \lambda_2) \in \mathbb{C}^2 \colon \lambda_1 \text{ and } \lambda_2 \text{ rationally linear dependent}\}.$$

In particular, any rational linear combination of ϕ_1, ϕ_2 fails to generate S. This example illustrates the Theorem for the case of real coefficients, cf. Remark (i). Namely, $\mathscr{N} \cap \mathbb{R}^2$ is a null-set in $\mathbb{R}^2$ yet it contains $\mathscr{Q}^2$, so it is dense in $\mathbb{R}^2$.

Open Problem. It is interesting to ask whether the Theorem also holds for finitely generated shift-invariant subspaces of $L^p(\mathbb{R}^n)$, where $1 \leq p \leq \infty$ and $p \neq 2$. For a few properties of these spaces we refer to [5, 6]. Since the proof of the Theorem relies heavily on fiberization techniques for $p = 2$ and on the characterization of shift-invariant spaces in terms of range functions, this question remains open for $p \neq 2$.

References

[1] C. de Boor, R. A. DeVore, and A. Ron, *The structure of finitely generated shift-invariant spaces in $L_2(\mathbf{R}^d)$*, J. Funct. Anal. **119** (1994), no. 1, 37–78.
[2] ———, *Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^d)$*, Trans. Amer. Math. Soc. **341** (1994), no. 2, 787–806.
[3] M. Bownik, *The structure of shift-invariant subspaces of $L^2(\mathbf{R}^n)$*, J. Funct. Anal. **177** (2000), no. 2, 282–309.
[4] H. Helson, *Lectures on Invariant Subspaces*, Academic Press, New York, 1964.
[5] R.-Q. Jia, *Shift-invariant spaces on the real line*, Proc. Amer. Math. Soc. **125** (1997), no. 3, 785–793.
[6] ———, *Stability of the shifts of a finite number of functions*, J. Approx. Theory **95** (1998), no. 2, 194–202.
[7] R. Q. Jia and C. A. Micchelli, *On linear independence for integer translates of a finite number of functions*, Proc. Edinburgh Math. Soc. **36** (1993), no. 1, 69–85.
[8] A. Ron, *A necessary and sufficient condition for the linear independence of the integer translates of a compactly supported distribution*, Constr. Approx. **5** (1989), no. 3, 297–308.

[9] A. Ron and Z. Shen, *Frames and stable bases for shift-invariant subspaces of* $L_2(\mathbf{R}^d)$, Canad. J. Math. **47** (1995), no. 5, 1051–1094.

Geometry of Sets of Parameters of Wave Packets

Gitta Kutyniok
(joint work with Wojciech Czaja and Darrin Speegle)

The goal of our project is to describe completeness properties of wave packets via geometric properties of the sets of their parameters. Our research is motivated by the simple observation that for $L^2(\mathbb{R})$ the sets of parameters of Gabor and wavelet systems form discrete subsets of 2–dimensional linear subspaces in $\mathbb{R}^3$ and that there exists an abundance of sets of parameters which give rise to Gabor or wavelet frames. On the other hand, it is known that systems associated with either translations, dilations, or modulations of a single function do not form frames nor Riesz bases in $L^2(\mathbb{R})$, cf., [7] and [3] for systems consisting of translations (and equivalently modulations) of a single function, and see [4] for systems of dilations. Furthermore, it is known that systems associated with full lattices of translations, dilations, and modulations are infinitely over-complete. Therefore, we shall investigate the role of the geometric structure of sets of parameters of wave packets for the functional properties of associated systems of functions.

1. Wave packets. In [1], Córdoba and Fefferman introduced "wave packets" as those families of functions, which consist of a countable collection of dilations, translations, and modulations of the Gaussian function. Here we will generalize this definition to collections of dilations, translations, and modulations of an arbitrary function in $L^2(\mathbb{R})$.

Definition. Given a function $\psi \in L^2(\mathbb{R})$ and a discrete set $\mathcal{M} \subset \mathbb{R}^+ \times \mathbb{R}^2$, we define the discrete *wave packet* $\mathcal{WP}(\psi, \mathcal{M})$, associated with ψ and $\mathcal{M}$, to be:

$$\mathcal{WP}(\psi, \mathcal{M}) = \{D_x T_y M_z \psi : (x, y, z) \in \mathcal{M}\},$$

where D_x, T_y, and M_z are the $L^2(\mathbb{R})$ unitary operators of dilations, translations, and modulations, respectively:

$$D_x(f)(t) = \sqrt{x} f(xt), \quad T_y(f)(t) = f(t - y), \quad M_z(f)(t) = e^{2\pi i t z} f(t).$$

With this definition, Gabor systems ($\mathcal{M} = \{1\} \times \Lambda$, $\Lambda \subset \mathbb{R}^2$) as well as wavelet systems ($\mathcal{M} = \mathcal{B} \times \{0\}$, $\mathcal{B} \subset \mathbb{R}^+ \times \mathbb{R}$) are thus special examples of wave packets.

2. Density and Dimension. A successful approach to study Gabor frames utilizes the notion of *Beurling density* of the collection of parameters Λ. If $\Lambda = a\mathbb{Z} \times b\mathbb{Z}$, Rieffel proved in 1981 that an associated Gabor system is complete only if $ab \leq 1$. This result has been further extended and generalized, and Ramanathan and Steger in [8] proved that if a Gabor system associated with an arbitrary set Λ is a frame then the lower Beurling density of Λ satisfies $D^-(\Lambda) \geq 1$. Moreover, if this frame is a Riesz basis then $D^+(\Lambda) = D^-(\Lambda) = 1$. We refer to [3] for further results in this area and for additional references.

An analogous approach has been undertaken in [4] to study wavelet systems in terms of an appropriately redefined notion of density that is suitable for the structure of the affine group associated with the sets of parameters of wavelet systems. Using these notions, the authors were able to obtain necessary conditions for the existence of wavelet frames in $L^2(\mathbb{R})$.

Our approach shall be an analogue of the two above described methods of characterizations of special wave packets. We introduce a notion of density with respect to the geometry of the affine Weyl–Heisenberg group, which is the appropriate setting for sets of parameters of wave packets. Since the results for wavelet systems indicate that we cannot expect to have a critical density for general wave packets, cf., [4], we need to develop another tool to correlate the geometric properties of the sets of parameters of wave packets with their functional properties. Based on density considerations and motivated by the definition of Hausdorff dimension, we therefore introduce a notion of upper and lower dimension $\dim^{\pm}(\mathcal{M})$ for discrete subsets $\mathcal{M} \subset \mathbb{R}^+ \times \mathbb{R}^2$. The following result shows some basic properties of this notion.

Theorem. Let $\mathcal{M}$ be a subset of $\mathbb{R}^+ \times \mathbb{R}^2$. Then,

(i) $\dim^+(\mathcal{M}) \in [0, 3] \cup \{\infty\}$.
(ii) $\dim^-(\mathcal{M}) \in \{0\} \cup [3, \infty]$.

Moreover, by just employing the definition, we obtain the relation $\dim^-(\mathcal{M}) \leq \dim^+(\mathcal{M})$.

3. General Results. Although the situation we consider is more general than that considered in [4], an analogous necessary condition for the existence of an upper frame bound for wave packets $\mathcal{WP}(\psi, \mathcal{M})$ still holds.

Theorem. Let $\psi \in L^2(\mathbb{R})$ and let $\mathcal{M}$ be a discrete subset of $\mathbb{R}^+ \times \mathbb{R}^2$. If $\mathcal{WP}(\psi, \mathcal{M})$ possesses an upper frame bound, then

$$\mathcal{D}_A^+(\mathcal{M}) < \infty \quad \text{for all } A \geq 3. \tag{1}$$

This immediately leads to necessary conditions on the upper and lower dimension of sets of parameters of wave packets, since (1) implies that $\dim^-(\mathcal{M}) \in \{0, 3\}$ and $\dim^+(\mathcal{M}) \in [0, 3]$. Thus if $\mathcal{WP}(\psi, \mathcal{M})$ has an upper frame bound, then there are only two possible values for $\dim^-(\mathcal{M})$. Wavelet frames and Gabor frames are examples of wave packet frames that satisfy the condition $\dim^-(\mathcal{M}) = 0$. We conjecture that this is the only value, which can be attained by sets of parameters of frames in general.

Conjecture. Let $\psi \in L^2(\mathbb{R})$ and let $\mathcal{M}$ be a discrete subset of $\mathbb{R}^+ \times \mathbb{R}^2$. If $\mathcal{WP}(\psi, \mathcal{M})$ possesses an upper frame bound, then $\dim^-(\mathcal{M}) = 0$.

We can answer this question when the sets of parameters of wave packets have the special form $\mathcal{M} = \mathcal{B} \times \mathbb{Z}$. Wave packets with such sets of parameters have been recently studied by Guido Weiss and his collaborators, see, e.g., [5, 6].

4. Case of Integer Modulations. In this situation we can prove the conjecture and also obtain additional restrictions for the upper dimension.

Theorem. Let $\psi \in L^2(\mathbb{R})$ and let $\mathcal{B}$ be a discrete subset of $\mathbb{R}^+ \times \mathbb{R}$. If $\mathcal{WP}(\psi, \mathcal{B} \times \mathbb{Z})$ possesses an upper frame bound, then

$$\dim^-(\mathcal{B} \times \mathbb{Z}) = 0 \quad \text{and} \quad \dim^+(\mathcal{B} \times \mathbb{Z}) \in [1, 3].$$

It is natural to ask, whether each value in $[1, 3]$ is indeed attained. To study this question we need to construct wave packet frames with prescribed dimensions. Our investigation of this problem leads to multiple examples of non–standard wave packets.

We split our study into two cases. If $1 \leq d \leq 2$, we are even able to construct orthonormal wave packet bases, not only just wave packet frames. It turns out that the most difficult examples to construct are for large dimensions. In this situation by using a highly technical construction, we obtain wave packet frames but no orthonormal wave packet bases so far.

We obtain the following results:

Theorem.

(i) For every $1 \leq d \leq 2$, there exists a discrete subset $\mathcal{B} \subset \mathbb{R}^+ \times \mathbb{R}$ such that $\dim^+(\mathcal{B} \times \mathbb{Z}) = d$ and $\mathcal{WP}(\chi_{[0,1]}, \mathcal{B} \times \mathbb{Z})$ is an orthonormal basis for $L^2(\mathbb{R})$.

(ii) For every $2 < d \leq 3$, there exists a discrete subset $\mathcal{B} \subset \mathbb{R}^+ \times \mathbb{R}$ such that $\dim^+(\mathcal{B} \times \mathbb{Z}) = d$ and $\mathcal{WP}(\chi_{[-\frac{1}{2}, \frac{1}{2}]}, \mathcal{B} \times \mathbb{Z})$ is a frame for $L^2(\mathbb{R})$.

Thus we obtain a full description of which values the upper and lower dimension associated with a frame wave packet can attain in the case $\mathcal{M} = \mathcal{B} \times \mathbb{Z}$ under consideration.

For more detailed information on this project we refer to [2].

References

[1] A. Córdoba and C. Fefferman, *Wave packets and Fourier integral operators,* Comm. Partial Differential Equations 3 (1978), no. 11, 979–1005.

[2] W. Czaja, G. Kutyniok, and D. Speegle, *Geometry of sets of parameters of wave packets,* preprint.

[3] O. Christensen, B. Deng, and C. Heil, *Density of Gabor frames,* Appl. Comput. Harmon. Anal. 7 (1999), 292–304.

[4] C. Heil and G. Kutyniok, *Density of weighted wavelet frames,* J. Geom. Anal., 13 (2003), 479–493.

[5] E. Hernandez, D. Labate, G. Weiss, and E. Wilson, *Oversampling, quasi affine frames and wave packets,* Appl. Comput. Harmon. Anal. 16 (2004), 111–147

[6] D. Labate, G. Weiss, and E. Wilson, *An Approach to the Study of Wave Packet Systems,* to appear in Contemp. Mathematics (2004).

[7] T. E. Olson and R. A. Zalik, *Nonexistence of a Riesz basis of translates,* "Approximation Theory", pp. 401–408, Lecture Notes in Pure and Applied Math., Vol. 138, Dekker, New York, 1992.

[8] J. Ramanathan and T. Steger, *Incompleteness of sparse coherent states,* Appl. Comput. Harmon. Anal. 2 (1995), no. 2, 148–153.

Decomposition of Operators and Construction of Frames
David R. Larson

The material we present here is contained in two recent papers. The first was authored by a [VIGRE/REU] team consisting of K. Dykema, D. Freeman, K. Kornelson, D. Larson, M. Ordower, and E. Weber, with the title *Ellipsoidal Tight Frames*, and is to appear in Illinois J. Math. This article started as an undergraduate research project at Texas A&M in the summer of 2002, in which Dan Freeman was the student and the other five were faculty mentors. Freeman is now a graduate student at Texas A&M. The project began as a solution of a finite dimensional frame research problem, but developed into a rather technically deep theory concerning a class of frames on an infinite dimensional Hilbert space. The second paper, entitled *Rank-one decomposition of operators and construction of frames*, is a joint article by K. Kornelson and D. Larson, and is to appear in the volume of Contemporary Mathematics containing the proceedings of the January 2003 AMS special session and FRG workshop on *Wavelets, Frames and Operator Theory*, which took place in Baltimore and College Park.

We will use the term *spherical frame* for a frame sequence which is *uniform* in the sense that all its vectors have the same norm. Spherical frames which are tight have been the focus of several articles by different researchers. Since frame theory is essentially geometric in nature, from a purely mathematical point of view it is natural to ask: Which other surfaces in a finite or infinite dimensional Hilbert space contain tight frames? In the first article we considered ellipsoidal surfaces.

By an *ellipsoidal surface* we mean the image of the unit sphere S_1 in the underlying Hilbert space H under a bounded invertible operator A in $B(H)$, the set of all bounded linear operators on H. Let E_A denote the ellipsoidal surface $E_A := AS_1$. A frame contained in E_A is called an *ellipsoidal frame*, and if it is tight it is called an ellipsoidal tight frame (ETF) for that surface. We say that a frame bound K is *attainable* for E_A if there is an ETF for E_A with frame bound K.

Given an ellipsoidal surface $E := E_A$, we can assume $E = E_T$ where T is a positive invertible operator. Indeed, given an invertible operator A, let $A^* = U|A^*|$ be the polar decomposition, where $|A^*| = (AA^*)^{1/2}$. Then $A = |A^*|U^*$. By taking $T = |A^*|$, we see tht $TS_1 = AS_1$. Moreover, it is easily seen that the positive operator T for which $E = E_T$ is unique.

The starting point for the work in the first paper was the following Proposition. For his REU project Freeman found an elementary calculus proof of this for the real case. Others have also independently found this result, including V. Paulsen, and P. Casazza and M. Leon.

Proposition 1. *Let E_A be an ellipsoidal surface on a finite dimensional real or complex Hilbert space H of dimension n. Then for any integer $k \geq n$, E_A contains a tight frame of length k, and every ETF on E_A of length k has frame bound $K = k\left[trace(T^{-2})\right]^{-1}$.*

We use the following standard definition: For an operator $B \in H$, the *essential norm* of B is:

$$\|B\|_{ess} := \inf\{\|B - K\| \ : \ K \text{ is a compact operator in} B(H)\}$$

Our main frame theorem from the first paper is:

Theorem 2. *Let E_A be an ellipsoidal surface in an infinite dimensional real or complex Hilbert space. Then for any constant $K > \|T^{-2}\|_{ess}^{-1}$, E_T contains a tight frame with frame bound K.*

So, for fixed A, in finite dimensions the set of attainable ETF frame bounds is finite, whereas in infinite dimensions it is a continuum.

Problem. If the essential norm of A is replaced with the norm of A in the above theorem, or if the inequality is replaced with equality, then except for some special cases, and trivial cases, no theorems of any degree of generality are known concerning the set of attainable frame bounds for ETF's on E_A. It would be interesting to have a general analysis of the case where $A - I$ is compact. In this case, one would want to know necessary and sufficient conditions for existence of a tight frame on E_A with frame bound 1. In the special case $A = I$ then, of course, any orthonormal basis will do, and these are the only tight frames on E_A in this case. What happens in general when $\|A\|_{ess} = 1$ and A is a small perturbation of I?

We use elementary tensor notation for a rank-one operator on H. Given $u, v, x \in H$, the operator $u \otimes v$ is defined by $(u \otimes v)x = \langle x, v\rangle u$ for $x \in H$. The operator $u \otimes u$ is a projection if and only if $\|u\| = 1$.

Let $\{x_j\}_j$ be a frame for H. The standard frame operator is defined by: $Sw = \sum_j \langle w, x_j\rangle x_j = \sum_j (x_j \otimes x_j)\, w$. Thus $S = \sum_j x_j \otimes x_j$, where this series of positive rank-1 operators converges in the strong operator topology (i.e. the topology of pointwise convergence). In the special case where each x_j is a unit vector, S is the sum of the rank-1 projections $P_j = x_j \otimes x_j$.

For A a positive operator, we say that A has a *projection decomposition* if A can be expressed as the sum of a finite or infinite sequence of (not necessarily mutually orthogonal) self-adjoint projections, with convergence in the strong operator topology.

If x_j is a frame of unit vectors, then $S = \sum_j x_j \otimes x_j$ is a projection decomposition of the frame operator. This argument is trivially reversible, so a positive invertible operator S is the frame operator for a frame of unit vectors if and only if it admits a projection decomposition $S = \sum_j P_J$. If the projections in the decomposition are not of rank one, each projection can be further decomposed (orthogonally) into rank-1 projections, as needed, expressing $S = \sum_n x_n \otimes x_n$, and then the sequence $\{x_n\}$ is a frame of unit vectors with frame operator S.

In order to prove Theorem 2, we first proved Theorem 3 (below), using purely operator-theoretic techniques.

Theorem 3. *Let A be a positive operator in $B(H)$ for H a real or complex Hilbert space with infinite dimension, and suppose $\|A\|_{ess} > 1$. Then A has a projection decomposition.*

Suppose, then, that $\{x_n\}$ is a frame of unit vectors with frame operator S. If we let $y_j = S^{-\frac{1}{2}}x_j$, then $\{y_j\}_j$ is a *Parseval* frame (i.e. tight with frame bound 1). So $\{y_j\}_j$ is an ellipsoidal tight frame for the ellipsoidal surface $E_{S^{-\frac{1}{2}}} = S^{-\frac{1}{2}}S_1$. This argument is reversible: Given a positive invertible operator T, let $S = T^{-2}$. Scale T if necessary so that $\|S\|_{ess} > 1$. Let $S = \sum_j x_j \otimes x_j$ be a projection decomposition of S. Then $\{Tx_j\}$ is an ETF for the ellipsoidal surface TS_1. Consideration of frame bounds and scale factors then yields Theorem 2.

Most of our second paper concerned *weighted* projection decompositions of positive operators, and resultant theorems concerning frames. If T is a positive operator, and if $\{c_n\}$ is a sequence of positive scalars, then a weighted projection decomposition of T with weights $\{c_n\}$ is a decomposition $T = \sum_j P_j$ where the P_j are projections, and the series converges strongly. We have since adopted the term *targeted* to refer to such a decomposition, and generalizations thereof. By a *targeted decomposition* of T we mean any strongly convergent decomposition $T = \sum_n T_n$ where the T_n is a sequence of *simpler* positive operators with special prescribed properties. So a weighted decomposition is a targeted decomposition for which the scalar weights are the prescribed properties. And, of course, a projection decomposition is a special case of targeted decomposition.

After a sequence of Lemmas, building up from finite dimensions and employing spectral theory for operators, we arrived at the following theorem. We will not discuss the details here because of limited space. It is the *weighted* analogue of theorem 3.

Theorem 4. *Let B be a positive operator in $B(H)$ for H with $\|B\|_{ess} > 1$. Let $\{c_i\}_{i=1}^{\infty}$ be any sequence of numbers with $0 < c_i \leq 1$ such that $\sum_i c_i = \infty$. Then there exists a sequence of rank-one projections $\{P_i\}_{i=1}^{\infty}$ such that $B = \sum_{i=1}^{\infty} c_i P_i$.*

We refer the interested reader to the *Open Problems* section of this report for more on targeted decompositions. In the first problem, we raised the question of which positive operators admit *finite* projection decompositions. The second problem related to a completely different type of targeted decomposition than discussed in this abstract, or considered in the two papers we presented. It was motivated by talks and discussions in this Workshop, and just may be relevant to the theory of modulation spaces and Gelfand triples. We plan to pursue this further.

Groups, Wavelets, and Function Spaces
Gestur Ólafsson

In the talk we discussed several connections between the following topics:

(1) Representation theory of Lie groups;
(2) Linear action of Lie groups on $\mathbb{R}^d$;
(3) Wavelets and wavelet sets;
(4) Besov spaces associated to symmetric cones.

Let $H \subseteq \mathrm{GL}(d, \mathbb{R})$ be a closed subgroup, and hence a Lie group. Let G be the group of affine linear maps $(x, h)(t) := h(t) + x$, $h \in H$, $x, y \in \mathbb{R}^d$. Then G is the semi-direct product of $\mathbb{R}^d$ and H, $G = \mathbb{R}^d \times_s H$. Define a unitary representation of G on $L^2(\mathbb{R}^d)$ by

$$\pi(x,h)f(t) = |\det(h)|^{-1/2} f((x,h)^{-1}(t)) = |\det(h)|^{-1/2} f(h^{-1}(t-x))\,.$$

It is quite often useful to have an equivalent realization of π in frequency space. Define for $F \in L^2(\widehat{\mathbb{R}}^d)$

$$\widehat{\pi}(x,h)F(\omega) = |\det h|^{1/2} e^{-2\pi x\cdot\omega} F(h^T(\omega))\,.$$

Then the Fourier transform $\mathcal{F} : L^2(\mathbb{R}^d) \to L^2(\widehat{\mathbb{R}}^d)$, $\mathcal{F}(f)(\omega) = \int f(t) e^{-2\pi t\cdot\omega}\, dt$ is a unitary intertwining operator. Here, and elsewhere, we write $\widehat{\mathbb{R}}^d$ to underline, that we are looking at $\mathbb{R}^d$ as the *frequency domain*.

For $\psi \in L^2(\mathbb{R}^d)$ define $W_\psi : L^2(\mathbb{R}^d) \to C(G)$, by

$$W_\psi(f)(g) := (f, \pi(g)\psi) = |\det h|^{-1/2} \int f(t)\overline{\psi(h^{-1}(t-x))}\, dt \quad g = (x,h) \in G\,.$$

Note, that W_ψ depends on our choice of wavelet function ψ. In particular, if $\psi \in S(\mathbb{R}^d)$, then W_ψ extends to a linear map on $S'(\mathbb{R}^d)$, the space of tempered distributions. It is an important question in analysis to study spaces of functions or distribution using the wavelet transform. In particular, for a given weight function w on G, one can, if the representation π is integrable, define a Banach space of distribution by $\{f \in S'(\mathbb{R}^d) \mid W_\psi(f) \in L^p(G, w d\mu_G)$, with norm $\|f\| = \|f\|_{L^p(G, w d\mu_G)}$. Here $d\mu_G$ denotes a left invariant measure on G. Using the structure of G as a semi-direct product, one can even define mixed $L^{p,q}$-norm. This is related to the Feichtinger-Gröchenig co-orbit theory for group representations, which for the Heisenberg group has become quite important through the theory of Modulation spaces, [11, 12, 19].

The simples case is $p = 2$ and $w = 1$. A simple calculation shows, that

$$\|W_\psi(f)\|^2_{L^2(G)} = \int_{\widehat{R}^d} |\mathcal{F}(f)(\omega)|^2 \left(\int_H |\mathcal{F}(\psi)(h^T\omega)|^2\, d\mu_H(h) \right) d\omega\,.$$

It follows that $W_\psi(f) \in L^2(G)$ if and only if $\int_H |\mathcal{F}(\psi)(h^T(\omega))|^2\, d\mu_H(h) < \infty$ for almost all $\omega \in \widehat{R}^d$. Furthermore, $W_\psi : L^2(\mathbb{R}^d) \to \mathrm{Im}(W_\psi) \subset L^2(G)$ is an unitary isomorphism onto its image if and only if

$$\int_H |\mathcal{F}(\psi)(h^T(\omega))|^2 \, d\mu_H(h) = 1$$

for almost all $\omega \in \widehat{R}^d$. In this case we have $f = W_\psi^*(W_\psi(f))$, or

$$f = \int_G W_\psi(f)(g)\pi(g)\psi \, d\mu_G(g)$$

as an weak integral. We refer to [3, 8, 9, 10, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 26] for more detailed discussion.

Let $\Delta \subset H$ and $\Lambda \subset \mathbb{R}^d$ be countable subsets. Let $\Gamma := \Lambda \times \Delta \subset G$. Define a sequence of functions ψ_γ, $\gamma \in \Gamma$, by

$$\psi_{\lambda,\delta}(t) := \pi((\lambda, \delta)^{-1})\psi(t) = |\det \delta|^{1/2}\psi(\delta(t) + \lambda)\,.$$

Then ψ is a (subspace) *wavelet* if the sequence $\{\psi_\gamma\}_{\gamma\in\Gamma}$ is a orthonormal basis for its closed linear span. A measurable set $\Omega \subset \widehat{\mathbb{R}}^d$, $0 < |\Omega| < \infty$ is a (subspace) *wavelet set* if $\psi = \mathcal{F}^{-1}(\chi_\Omega)$ is a (subspace) wavelet. For discussion on wavelet sets see [1, 4, 5, 6, 7, 23, 24, 25]. A special class of groups H was studied in [10, 22, 23]. Here it was assumed, that H has finitely many open orbits $\mathcal{O}_1, \ldots, \mathcal{O}_r \subset \widehat{\mathbb{R}}^d$ of full measure, i.e., $(H, \widehat{\mathbb{R}}^d)$ is a *pre-homogeneous vector space*. We set

$$L_j^2 = L^2_{\mathcal{O}_j}(\mathbb{R}^d) := \{f \in L^2(\mathbb{R}^d) \mid \mathrm{Supp}(\mathcal{F}(f)) \subseteq \overline{\mathcal{O}_j}\}\,.$$

As an example, take $H = \mathbb{R}^+\mathrm{SO}(d)$. There is only one open orbit $\mathcal{O}_1 = \widehat{\mathbb{R}}^d \setminus \{0\}$. In particular, $L_1^2 = L^2(\mathbb{R}^d)$. Let $F \subset \mathrm{SO}(d)$ be a finite subgroup and $\lambda > 1$. Let $\Delta = \{\lambda^n R \mid n \in \mathbb{Z}, R \in F\}$, and let $\Lambda \subset \mathbb{R}^d$ a lattice. Then there exists a $\Gamma \times \Delta$-wavelet set for $L^2(\mathbb{R}^d)$. This follows from Theorem 1 [6] as was pointed out to me by my student M. Dobrescu. We get a more complicated example by taking $H = \mathbb{GL}(n, \mathbb{R})_o$ and $\mathbb{R}^d = \mathrm{Sym}(n, \mathbb{R})$, $d = n(n+1)/2$, the space of symmetric $n \times n$-matrices. The group H operates on $\mathbb{R}^d$ by $h \cdot X = gXg^T$. The open orbits are $\mathcal{O}_{p,q} = H \cdot I_{p,q}$. Here $\mathcal{O}_{p,q}$ stands for the open set of regular matrices of signature $(p, q = n - p)$, and $I_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$. The set $\mathcal{O}_{n,0}$ is an open symmetric cone. It is well known, that the group S of upper triangular matrices acts transitively on C. Let A be the group of diagonal matrices with positive diagonal elements, and let N be the group of upper triangular matrices (x_{ij}), with $x_{ii} = 1$, $i = 1, \ldots, n$. Then $S = AN = NA$. In [22, 23] a special choice for Δ was made. This set Δ is closely related to the structure of H. In our example this construction can be explained by taking $\Delta_N = \{(x_{ij} \in N \mid x_{ij} \in \mathbb{Z}\}$, and $\Delta_A = \{d(\lambda_1^{k_1}, \ldots, \lambda_n^{k_n}) \mid (k_1, \ldots, k_n) \in \mathbb{Z}^n\}$, where $\lambda_j > 1$. Then we set $\Delta = \Delta_A\Delta_N$. It follows by [23], Theorem 4.5, that, if Λ is a lattice in $\mathbb{R}^d$, then there exists a $L^2(\mathbb{R}^d)$ wavelet set for $\Lambda \times \Delta$. But it is an open problem, if there exists a $L^2_C(\mathbb{R}^d)$ wavelet set for $\Lambda \times \Delta$. One can even complicate this by adding a finite group of rotations that centralize A and normalize N.

It was also shown in [22, 23] that, for our special choice of Δ, there is always a set $\Omega \subset \mathcal{O}_j$ such that $\psi = \mathcal{F}^{-1}(\chi_\Omega)$ generates a tight frame for L_j^2. It is clear, that we can replace χ_Ω by a compactly supported function $\phi \geq 0$, $\phi|_\Omega = 1$, and get a frame generator that is rapidly decreasing. But it is an open problem if we can in fact get a rapidly decreasing function that generates a *tight frame*. A private note by D. Speegle indicates, that this might in fact be possible.

One of the reason I discuss the last example is, that this is just an example of H being the automorphism group of a symmetric cone $C \subset R^d$, i.e., $H = \mathrm{GL}(C) := \{h \in \mathrm{GL}(d,\mathbb{R}) \mid h(C) = C\}$. The general philosophy is, that wavelets are associated to Besov spaces. In fact, one sees easily, that the Besov spaces in [2] can also be defined by using the continuous wavelet transform. It is therefore a natural question, which we pose here as a third open problem, to study the Besov spaces, introduced in [2], using the theory of co-orbit spaces and the discrete wavelet transform using the results from [22, 23] applied to the group $\mathrm{GL}(C)$.

References

[1] P. Aniello, G. Cassinelli, E. De Vito, and A. Levrero: Wavelet transforms and discrete frames associated to semidirect products. *J. Math. Phys.* **39** (1998), 3965–3973

[2] D. Békollé, A. Bonami, G. Garrigós, F. Ricci: Littlewood-Paley decompositions and Besov spaces related to symmetric cones. Preprint, 2004

[3] D. Bernier, K. F. Taylor: Wavelets from square-integrable representations. *SIAM J. Math. Anal.* **27** (1996), 594–608

[4] C.K. Chui, X. Shi: Orthonormal wavelets and tight frames with arbitrary real dilations. *Appl. Comput. Harmon. Anal.* **9** (2000), no. 3, 243–264.

[5] X. Dai, Y. Diao, Q. Gu and D. Han: The existence of subspace wavelet sets. Preprint

[6] X. Dai, D.R. Larson and D. Speegle: Wavelet sets in $\mathbb{R}$, *J. Fourier Anal. Appl.* **3** (1997), no. 4, 451–456.

[7] X. Dai, D.R. Larson and D. Speegle: Wavelet sets in $\mathbb{R}$ II, *Wavelets, multiwavelets, and their applications (San Diego, CA, 1997),* 15–40, Contemp. Math., 216, *Amer. Math. Soc.,* Providence, RI, 1998.

[8] I. Daubechies, A. Grossmann, and Y. Meyer: Painless nonorthogonal expansions. *J. Math. Phys.* **27** (1986), 1271–1283

[9] M. Duflo, C.C. Moore: On the regular representation of a non-unimodular locally compact group. *J. Funct. Anal.* **21** (1976), 209–243

[10] R. Fabec, G. Ólafsson: The Continuous Wavelet Transform and Symmetric Spaces. *Acta Applicandae Math.* **77** (2003), 41–69

[11] H.G. Feichtinger, and K. Gröchenig: A unified approach to atomic decompositions via integrable group representations. In: *Proceedings, Conference on Functions, Spaces and Applications. Lund 1986* Lecture Notes in Mathematics **1302** (1986), 52–73, Springer-Verlag, New York/Berlin 1987

[12] ———: Banach spaces related to integrable group representations and their atomic decomposition. *J. Funct. Anal.* **86** (1989), 307–340

[13] H. Führ: Wavelet frames and admissibility in higher dimensions. *J. Math. Phys.* **37** (1996), 6353–6366

[14] ———: Continuous wavelet transforms with Abelian dilation groups. *J. Math. Phys.* **39** (1998), 3974–3986

[15] ———: Admissible vectors for the regular representation. To appear in *Proceedings of the AMS*

[16] H. Führ, M. Mayer: Continuous wavelet transforms from semidirect products: Cyclic representations and Plancherel measure. To appear in *J. Fourier Anal. Appl.*

[17] A. Grossmann, J. Morlet, and T. Paul: Transforms associated to square integrable group representations I. General results. *J. Math. Phys.* **26** (1985), 2473–2479

[18] ______: Transforms associated to square integrable group representations II. Examples. *Ann. Inst. Henri Poincaré: Phys. Theor.* **45** (1986), 293–309

[19] C. Heil: Integral operators, pseudodifferential operators, and Gabor frames. In "Advances in Gabor analysis", Eds. H.G. Feichtihger and T. Strohmer. Birkhäuser, 2002

[20] C.E. Heil, and D.F. Walnut: Continuous and discrete wavelet transform. *SIAMRev.* **31** (1989), 628–666

[21] R.S. Laugesen, N. Weaver, G.L. Weiss, E.N. Wilson: A characterization of the higher dimensional groups associated with continuous wavelets. *J. Geom. Anal.* **12** (2002), 89–102.

[22] G. Ólafsson: Continuous action of Lie groups on $\mathbb{R}^n$ and frames. To appear in *International Journal of Wavelets, Multi-resolution and Information Processing*

[23] G. Ólafsson and D. Speegle: Wavelets, wavelet sets, and linear actions on $\mathbb{R}^n$, To appear in *Contemporary Mathematics (AMS)*

[24] D. Speegle: On the existence of wavelets for non-expansive dilations, *Collect. Math.* **54**, 2, (2003), 163–179.

[25] Yang Wang: Wavelets, tiling, and spectral sets. *Duke Math. J.* **114** (2002), no. 1, 43–57.

[26] G. Weiss and E.N. Wilson: The Mathematical Theory of Wavelets. In: J.S. Byrnes (ed.) *Twenty Century Harmonic Analysis - A Celebration*, 329–366, Kluwer Academic Publisher, 2001

Adaptive Wavelet Methods for the Numerical Solutions of Operator Equations

Karsten Urban

We review recent results on the construction, analysis and realization of adaptive wavelet methods for the numerical solution of operator equations. The theoretic results are mainly based on on work by Cohen, Dahmen and DeVore, [4, 5, 6, 7].

Elliptic Operators. We consider (just for the sake of simplicity) the boundary value problem on a bounded, open domain $\Omega \subset \mathbb{R}^n$ determining $u : \Omega \to \mathbb{R}$ such that

$$-\Delta u(x) = f(x),\ x \in \Omega, \qquad u_{|\partial\Omega} = 0, \tag{1}$$

for a given function $f : \Omega \to \mathbb{R}$. The variational formulation reads: find $u \in H_0^1(\Omega)$ such that

$$a(u, v) := (\nabla u, \nabla v)_0 = (f, v)_0 \text{ for all } v \in H_0^1(\Omega) \tag{2}$$

for a given function $f \in H^{-1}(\Omega)$, where $(\cdot, \cdot)_0$ denotes the standard L_2-inner product on Ω. Introducing the differential operator

$$A : H_0^1(\Omega) \to H^{-1}(\Omega), \qquad \langle Au, v\rangle := a(u, v), \quad u, v \in H_0^1(\Omega), \tag{3}$$

we can rewrite (2) as an operator equation

$$Au = f \tag{4}$$

in the Sobolev space $H_0^1(\Omega)$. Note that (4) is an infinite-dimensional operator equation in a function space. We always assume in the sequel, that A is *boundedly invertible*, i.e.

$$\|Au\|_{-1} \sim \|u\|_1, \qquad u \in H_0^1(\Omega), \tag{5}$$

where we use the notation $A \sim B$ in order to abbreviate the existence of constants $0 < c \leq C < \infty$ such that $cA \leq B \leq CA$. At this point also general elliptic operators on Hilbert spaces are included.

Wavelet Characterization of Sobolev Spaces. The first step is to transform (4) into a (well-conditioned) problem in sequence spaces. This is done with the aid of a bi-orthogonal wavelet bases. Assume

$$\Psi := \{\psi_\lambda : \lambda \in \mathcal{J}\} \tag{6}$$

is a Riesz basis for $L_2(\Omega)$. Here $\mathcal{J}$ is an infinite set of indices and we always think of an index $\lambda \in \mathcal{J}$ as a pair (j,k), where $|\lambda| := j \in \mathbb{N}$ always denotes the *scale* or *level* and k (which possibly is a vector) contains information on the localization of ψ_λ (e.g. the center of its support). We assume that Ψ admits a characterization of a whole scale of Sobolev spaces in the sense, that the following estimates hold:

$$\left\| \sum_{\lambda\in\mathcal{J}} d_\lambda\, \psi_\lambda \right\|_s \sim \left(\sum_{\lambda\in\mathcal{J}} |d_\lambda|^{2s|\lambda|} \right)^{1/2}, \tag{7}$$

for $s \in (-\tilde{\gamma}, \gamma)$ and $\gamma, \tilde{\gamma} > 1$ depend on the properties of Ψ such as polynomial exactness and order of vanishing moments. Using the short hand notations

$$\boldsymbol{d}^T\Psi := \sum_{\lambda\in\mathcal{J}} d_\lambda\, \psi_\lambda, \quad \boldsymbol{d} = (d_\lambda)_{\lambda\in\mathcal{J}}, \qquad \boldsymbol{D} = \operatorname{diag}(2^{|\lambda|})_{\lambda\in\mathcal{J}} \tag{8}$$

we can rephrase (7) in the following way

$$\|\boldsymbol{d}^T\Psi\|_s \sim \|\boldsymbol{D}^s\boldsymbol{d}\|_{\ell_2(\mathcal{J})}. \tag{9}$$

Note that nowadays there are criteria known in order to ensure (9) and also constructions of wavelets also on complex domains are on the market, [2, 3, 10, 11].

Then, the Riesz Representation Theorem guarantees the existence of a *bi-orthogonal* wavelet basis $\tilde{\Psi} = \{\tilde{\psi} : \lambda \in \mathcal{J}\}$ such that

$$\|\boldsymbol{d}^T\tilde{\Psi}\|_{-s} \sim \|\boldsymbol{D}^{-s}\boldsymbol{d}\|_{\ell_2(\mathcal{J})}. \tag{10}$$

An equivalent well-conditioned problem in ℓ_2. This implies for any $u = \boldsymbol{u}^T\Psi$

$$\begin{aligned}
\|\boldsymbol{u}\|_{\ell_2(\mathcal{J})} &\overset{(9)}{\sim} \|(\boldsymbol{D}^{-1}\boldsymbol{u})^T\Psi\|_1 \overset{(5)}{\sim} \|A((\boldsymbol{D}^{-1}\boldsymbol{u})^T\Psi)\|_{-1} \\
&= \|(A((\boldsymbol{D}^{-1}\boldsymbol{u})^T\Psi),\Psi)_0\,\tilde{\Psi}\|_{-1} \overset{(10)}{\sim} \|\boldsymbol{D}^{-1}(A((\boldsymbol{D}^{-1}\boldsymbol{u})^T\Psi),\Psi)_0\|_{\ell_2(\mathcal{J})} \\
&= \|\boldsymbol{D}^{-1}(A\Psi,\Psi)_0\,\boldsymbol{D}^{-1}\boldsymbol{u}\|_{\ell_2(\mathcal{J})},
\end{aligned}$$

which shows that $\|\boldsymbol{u}\|_{\ell_2(\mathcal{J})} \sim \|\boldsymbol{A}\boldsymbol{u}\|_{\ell_2(\mathcal{J})}$ for $\boldsymbol{A} := \boldsymbol{D}^{-1}(A\Psi,\Psi)_0\,\boldsymbol{D}^{-1}$. In other words, $\boldsymbol{A} : \ell_2(\mathcal{J}) \to \ell_2(\mathcal{J})$ is a *boundedly invertible* operator on the sequence space $\ell_2(\mathcal{J})$. Defining $\boldsymbol{f} := (f,\Psi)_0$, we are led to the equivalent discrete problem

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{f}. \tag{11}$$

An infinte-dimensional convergent adaptive algorithm. Ignoring for a minute that an infinite ℓ_2-sequence can not be represented in a computer, we aim at constructing an iterative solution method for the discrete problem (11). This is done by a *Richardson*-type iteration: Given an initial guess $\boldsymbol{u}^{(0)} \in \ell_2(\mathcal{J})$ and some $\alpha \in \mathbb{R}^+$, we define

$$\boldsymbol{u}^{(i+1)} := \boldsymbol{u}^{(i)} + \alpha(\boldsymbol{f} - \boldsymbol{A}\boldsymbol{u}^{(i)}) = (\boldsymbol{I} - \alpha\boldsymbol{A})\boldsymbol{u}^{(i)} + \alpha\boldsymbol{f}. \tag{12}$$

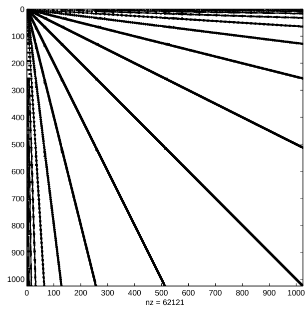
The convergence of this algorithm is easily seen:

$$\begin{aligned} \|\boldsymbol{u} - \boldsymbol{u}^{(i+1)}\|_{\ell_2(\mathcal{J})} &= \|\boldsymbol{u} + \alpha\underbrace{(\boldsymbol{f} - \boldsymbol{A}\boldsymbol{u})}_{=0} - \boldsymbol{u}^{(i)} - \alpha(\boldsymbol{f} - \boldsymbol{A}\boldsymbol{u}^{(i)})\|_{\ell_2(\mathcal{J})} \\ &= \|(\boldsymbol{I} - \alpha\boldsymbol{A})(\boldsymbol{u} - \boldsymbol{u}^{(i)})\|_{\ell_2(\mathcal{J})} \\ &\le \|\boldsymbol{I} - \alpha\boldsymbol{A}\|_{B(\ell_2(\mathcal{J}))}\|\boldsymbol{u} - \boldsymbol{u}^{(i)}\|_{\ell_2(\mathcal{J})}, \end{aligned}$$

i.e., this iteration converges if $\rho := \|\boldsymbol{I} - \alpha\boldsymbol{A}\|_{B(\ell_2(\mathcal{J}))} < 1$. This condition, in turns, can be guaranteed e.g. if $\boldsymbol{A}$ is s.p.d. which holds e.g. for wavelet representations of elliptic partial differential operators.

Approximate Operator Applications. Using the locality and the vanishing moment properties of wavelets, on can show that the wavelet representation of a large class of operators is *almost sparse*, i.e., one has

$$|a_{\lambda,\lambda'}| \le C\, 2^{-\left||\lambda|-|\lambda'|\right|\sigma}(1 + d(\lambda,\lambda'))^{-\beta}, \tag{13}$$



where $d(\lambda,\lambda') := 2^{\min(|\lambda|,|\lambda'|)}\mathrm{dist}(\mathrm{supp}\,\psi_\lambda, \mathrm{supp}\,\psi_{\lambda'})$ for some parameters σ and β. Roughly speaking this means that one has a decay in the level difference as well as in the spatial distance of wavelets. A typical structure is the well-known *finger structure*, see figure right.

For such kind of operators, an approximate application APPLY was constructed. Replacing any multiplication with $\boldsymbol{A}$ by the routine, yields a convergent adaptive method that moreover was proven to be asymptotically optimal in the sense that the rate of convergence stays proportional to the decay of the best N-term approximation at optimal cost.

Numerical results are shown for the Laplace [1] and the Stokes problem, [8, 12].

References

[1] Barinka, A., Barsch, T., Charton, P., Cohen, A., Dahlke, S., Dahmen, W., Urban, K., *Adaptive Wavelet Schemes for Elliptic Problems - Implementation and Numerical Experiments*, Siam J. Scient. Comput. **23**, No. 3 (2001), 910-939.

[2] Canuto, C., Tabacco, A., Urban, K., *The Wavelet Element Method, Part I: Construction and Analysis*, Appl. Comp. Harm. Anal. **6** (1999), 1–52.

[3] Canuto, C., Tabacco, A., Urban, K., *The Wavelet Element Method, Part II: Realization and additional features in 2D and 3D*, Appl. Comp. Harm. Anal. **8** (2000), 123–165.

[4] A. Cohen, W. Dahmen, and R.A. DeVore, *Adaptive Wavelet Schemes for Elliptic Operator Equations — Convergence Rates*, Math. Comput., **70**, No. 233 (2001), 27–75.

[5] A. Cohen, W. Dahmen, and R.A. DeVore, *Adaptive Wavelet Methods II — Beyond the Elliptic Case*, Found. Comput. Math. **2** (2002), 203–245.

[6] A. Cohen, W. Dahmen, and R.A. DeVore, *Adaptive Wavelet Schemes for Nonlinear Variational Problems*, SIAM J. Numer. Anal. **41**, No. 5 (2003), 1785–1823.

[7] A. Cohen, W. Dahmen, and R.A. DeVore, *Sparse Evaluation of Compositions of Functions Using Multiscale Expansions*, SIAM J. Math. Anal. **35**, No. 2 (2003), 279–303.

[8] Dahlke, S., Dahmen, W., Urban, K., *Adaptive wavelet methods for saddle point problems — Convergence Rates*, SIAM J. Numer. Anal. **40**, No. 4 (2002), 1230–1262.

[9] W. Dahmen, *Wavelet and Multiscale Methods for Operator Equations*, Acta Numerica **6** (1997), 55–228.

[10] W. Dahmen and R. Schneider, *Composite Wavelet Bases for Operator Equations*, Math. Comput., **68** (1999) 1533–1567.

[11] W. Dahmen and R. Schneider, *Wavelets on Manifolds I: Construction and Domain Decomposition*, SIAM J. Math. Anal., **31** (1999) 184–230

[12] Dahmen, W., Urban, K., Vorloeper, J., *Adaptive Wavelet Methods – Basic Concepts and Applications to the Stokes Problem*, in: Wavelet Analysis, D.-X,. Zhou (Hrsg.), World Scientific, New Jersey, 2002, 39-80.

Orthogonal Frames for Encryption

Eric Weber

(joint work with Ryan Harkins and Andrew Westmeyer)

There are several encryption algorithms in which randomness plays a role in the encryption process. The first example is the one time pad, which is an unconditionally secure cipher and is optimal in terms of key length. The process of the one time pad is the following: take a message m, expressed in some binary format, choose at random a binary sequence of the same length, and bitwise add the message to the random sequence. The recipient, knowing the random sequence, then adds the sequence to the cipher text again to recover the message. We remark here that this is actually the basis for quantum cryptography. The other example is the McEliece cipher. The encryption here is based on error correcting codes: choose a code which corrects N errors, encode the message, and introduce N randomly chosen errors. The cipher text then is the encoded message with the errors. The decryption then is to decode the ciphertext which corrects the errors. It is possible to actually alter this slightly to make it a public key encryption system.

Both ciphers have drawbacks: the one time pad is a private key system, and the key must change every time a message is encrypted. The McEliece cipher requires a prohibitively large key size compared to the size of the message.

We propose here a third encryption algorithm which utilizes randomness in the encryption process based on Hilbert space frames.

Let H be a separable Hilbert space over the field $\mathbb{F}$ with scalar product $\langle \cdot, \cdot \rangle$, where $\mathbb{F}$ denotes either $\mathbb{R}$ or $\mathbb{C}$. A frame for H is a sequence $\mathbb{X} := \{x_n\}_{n \in \mathbb{Z}}$ such that there exist constants $0 < A \leq B < \infty$ such that for all $v \in H$,

$$A\|v\|^2 \leq \sum_{n\in\mathbb{Z}} |\langle v, x_n\rangle|^2 \leq B\|v\|^2. \tag{1}$$

If $A = B = 1$, the frame is said to be *Parseval*, and then for all $v \in H$,

$$v = \sum_{n\in\mathbb{Z}} \langle v, x_n\rangle x_n.$$

For elementary frame theory, see [Han et al. 2000, Casazza 2000].

Let H be a finite dimensional Hilbert space. A finite frame is a frame $\mathbb{X} := \{x_i\}_{i=1}^M$ for H, where M is necessarily no smaller than the dimension of H. The analysis operator for $\mathbb{X}$ is given by

$$\Theta_{\mathbb{X}} : H \to \mathbb{F}^M : v \mapsto (\langle v, x_1\rangle, \langle v, x_2\rangle, \ldots, \langle v, x_M\rangle).$$

Definition 1. Let H and K be finite dimensional Hilbert spaces. Two frames $\mathbb{X} := \{x_n\}_{n=1}^M \subset H$ and $\mathbb{Y} := \{y_n\}_{n=1}^M \subset K$ are *orthogonal* if for all $v \in H$, $\sum_{n=1}^M \langle v, x_n\rangle y_n = 0$. Equivalently, $\mathbb{X}$ and $\mathbb{Y}$ are orthogonal if $\Theta_{\mathbb{Y}}^*\Theta_{\mathbb{X}} : H \to K$ is the 0 operator, where $\Theta_{\mathbb{Y}}^*$ denotes the Hilbert space adjoint.

Our encryption scheme, which is similar to a subband coding scheme, is an effort to approximate the One-Time Pad. The (private) key for this encryption scheme is two orthogonal Parseval frames $\{x_n\}_{n=1}^M \subset H$ and $\{y_n\}_{n=1}^M \subset K$. Let $\Theta_{\mathbb{X}}$ and $\Theta_{\mathbb{Y}}$ respectively denote their analysis operators. Suppose $m \in H$ is a message; let $g \in K$ be a non-zero vector chosen at random. The ciphertext $c \in \mathbb{F}^M$ is given as follows:

$$c := \Theta_{\mathbb{X}} m + \Theta_{\mathbb{Y}} g.$$

To recover the message, we apply $\Theta_{\mathbb{X}}^*$:

$$\begin{aligned}\Theta_{\mathbb{X}}^* c &= \Theta_{\mathbb{X}}^*\Theta_{\mathbb{X}} m + \Theta_{\mathbb{X}}^*\Theta_{\mathbb{Y}} g\\ &= \sum_{n=1}^M \langle m, x_n\rangle x_n + \sum_{n=1}^M \langle m, y_n\rangle x_n\\ &= m + 0 = m.\end{aligned}$$

Our experiments show that this encryption algorithm is robust against a brute force attack. However, the encryption algorithm is vulnerable to a chosen-plaintext attack.

A chosen-plaintext attack is an attack mounted by an adversary which chooses a plaintext and is then given the corresponding ciphertext. For convenience, assume that $H = K = \mathbb{R}^N$ and $M = 2N$. The attack on our scheme is as follows:

Step 1. Determine the range $\Theta_{\mathbb{Y}}(\mathbb{R}^N)$. Choose any plaintext m of size N. Encode the plaintext twice, with output, say, e_0 and e_1. Compute $e_1 - e_0 = \Theta_{\mathbb{X}} m + \Theta_{\mathbb{Y}} g_1 - (\Theta_{\mathbb{X}} m + \Theta_{\mathbb{Y}} g_0) = \Theta_{\mathbb{Y}}(g_1 - g_0)$. Notice that this yields a vector $f_1 = \Theta_{\mathbb{Y}}(g_1 - g_0)$ in the range of $\Theta_{\mathbb{Y}}$. Encode the plaintext a third time, with output e_2, and compute $f_2 = e_2 - e_0$. Compute $f_3, \ldots, f_m$ until the collection $\{f_1, \ldots, f_m\}$ contains a linearly independent subset of size N. This then determines the subspace $Z := \Theta_{\mathbb{Y}}(\mathbb{R}^N) \subset \mathbb{R}^M$.

Step 2. Determine the range $T := \Theta_{\mathbb{X}}(\mathbb{R}^N) \subset \mathbb{R}^M$. Choose any (non-zero) plaintext m_1 of size N; encode the plaintext, with output e_1; then project e_1 onto the orthogonal complement of Z. This yields a vector x_1 in T. Choose another plaintext m_2 and repeat, yielding vector $x_2 \in T$. Repeat until the collection $\{x_1, \ldots, x_q\}$ contains a linearly independent subset of size N. This set determines T.

Step 3. Determine the matrix $\Theta_{\mathbb{X}}$. Suppose in Step 2, $\{m_1, \ldots, m_N\}$ is such that $\{x_1, \ldots, x_N\}$ is linearly independent. We then have

$$\Theta_{\mathbb{X}} m_k = x_k \text{ for } k = 1, \ldots N.$$

Given this system of equations, now solve for $\Theta_{\mathbb{X}}$.

Step 4. Unencode cipher texts. Given any ciphertext c, the adversary computes the following:

$$\begin{aligned} \Theta_{\mathbb{X}}^* c &= \Theta_{\mathbb{X}}^*(\Theta_{\mathbb{X}} m + \Theta_{\mathbb{Y}} g) \\ &= \Theta_{\mathbb{X}}^* \Theta_{\mathbb{X}} m \\ &= m \end{aligned}$$

since $\mathbb{X}$ was a Parseval frame.

References

[Agaian 1985] Agaian, S.S.: *Hadamard Matrices and Their Applications.* New York: Springer-Verlag, 1985.

[Aldroubi et al. 2002] Aldroubi, A., D. Larson, W.S. Tang, and E. Weber, *Geometric Aspects of Frame Representations of Abelian Groups*, preprint (2002) (available on ArXiv.org; math.FA/0308250).

[Benedetto et al. 2003] Benedetto, J. and Fickus, M.: *Finite Normalized Tight Frames*, *Adv. Comput. Math.* **18**, (2003) no. 2-4, 357–385.

[Casazza 2000] Casazza, P.: *The Art of Frame Theory*, *Taiwanese Math. J.* **4** (2000) no. 2, 129–201.

[Casazza et al. 2001] Casazza, P., Kovacević, J.: *Uniform Tight Frames for Signal Processing and Communications*, SPIE Proc. vol. 4478 (2001), 129–135.

[Chabaud 1995] Chabaud, F., *On the security of some cryptosystems based on error-correcting codes Advances in Cryptology – EUROCRYPT '94*, Lecture Notes in Computer Science 950, Springer-Verlag, 1995, pp. 131–139.

[Delsarte et al. 1969] Delsarte, P. and Goethals, J. M.: *Tri-weight Codes and Generalized Hadamard Matrices*, Information and Control **15** (1969), p. 196-206.

[Dykema et al. 2003] Dykema, K., Freeman, D., Kornelson, K., Larson, D., Ordower, M., Weber, E.: *Ellipsoidal Tight Frames and Projection Decompositions of Operators*, preprint (2003).

[Han et al. 2000] Han, D. and Larson, D.: *Frames, Bases, and Group Representations* Mem. Amer. Math. Soc. **147**, (2000) no. 697.

[van Lint 1992] van Lint, J.H.: *A Course in Combinatorics*, Cambridge: Cambridge University Press, 1992.

[Menezes et al. 1997] Menezes, A., van Oorschot, P. and Vanstone, S.: *Handbook of Applied Cryptography* CRC Press, 1997.

[Wallis 1972] Wallis, W.D.: *Combinatorics: Room Squares, Sum-Free Sets, Hadamard Matrices*, Berlin: Springer-Verlag, 1972.

Open Problems

How close can an L^1-Function be to a Convolution Idempotent?

Hans G. Feichtinger

On one hand, it is well known that $(L^1(\mathbb{R}^d), *)$ is a convolution algebra which does not contain unit, in the sense that it does not exists any function $e \in L^1(\mathbb{R}^d)$ such that $g * e = g$, for all $g \in L^1(\mathbb{R}^d)$. On the other hand, it is always possible to construct a sequence $(e_n)_{n\in\mathbb{N}}$ of functions in $L^1(\mathbb{R}^d)$ such that $\|g * e_n - g\|_1 \to 0$, for $n \to +\infty$.

The open problem suggested by Feichtinger is the following (It has been first stated at an Oberwolfach conference in 1980):

What is the infimum for expression of the form $\|g * g - g\|_1$, where g is a symmetric function $g \in L^1(\mathbb{R})$ with $\|g\|_1 = 1$. Is there a function which minimizes $\|g * g - g\|$? (of course it cannot be uniquely determined, since the problem is invariant under L^1-normalized dilations, but maybe this is the only form of ambiguity).

Nowadays the problem appears again as very interesting because it not obvious how to attack it numerically, because it cannot be formulated in a non-trivial way over discrete groups.

Let us show that this problem well posed and, in particular, that

$$\inf_{g\in L^1(\mathbb{R}),\|g\|_1=1} \|g * g - g\|_1 \geq \frac{1}{4}. \tag{1}$$

Because of the assumptions on g, its Fourier transform $\mathcal{F}g$ has the following properties: $\mathcal{F}g$ is real valued, continuous, vanishing at infinity, and $\mathcal{F}g(0) = 1$. The norm $\|g * g - g\|_1$ can be estimated from below by

$$\|g * g - g\|_1 \geq \|(\mathcal{F}g)^2 - \mathcal{F}g\|_\infty.$$

In particular, one has $\|(\mathcal{F}g)^2 - \mathcal{F}g\|_\infty \geq \big\||\mathcal{F}g|^2 - |\mathcal{F}g|\big\|_\infty$. Since $\mathcal{F}g$ is continuous, vanishing at infinity, and $\mathcal{F}g(0) = 1$, there exists $\omega_0 \in \mathbb{R}^d$ such that $|\mathcal{F}g(\omega_0)| = \frac{1}{2}$. Therefore

$$\sup_{\omega\in\mathbb{R}^d} \big||\mathcal{F}g(\omega)|^2 - |\mathcal{F}g(\omega)|\big| \geq \frac{1}{4}.$$

This immediately implies (1). Numerical experiments indicate that (naturally) the Gaussian (up to dilation) is a “strong candidate” with a value around $0.31 < 1/3$, so one may expect that the “true value of the infimum” is in the interval $[1/4, 1/3]$.

Approximation of Frames by Normalized Tight Ones

Michael Frank

Let $\{x_i\}_i$ be a frame of a Hilbert subspace $K \subseteq H$ of a given (separable) Hilbert space H with upper and lower frame bounds B and A. The resulting frame transform is the map $\theta : K \to l_2\,,\ \theta(x) = \{\langle x, g_i \rangle\}_i$, and its adjoint operator is $\theta^* : l_2 \to K\,,\ \theta^*(e_i) = g_i$, $i \in \mathbb{N}$, for the standard orthonormal basis $\{e_i\}_i$ of l_2. Let $S = (\theta^*\theta)^{-1}$ be the frame operator defined on K. It is positive and invertible. There exists an orthogonal projection $P : l_2 \to \theta\theta^*(l_2) \subseteq l_2$ onto the range of the frame transform.

Problem:

Are there distance measures on the set of frames of all Hilbert subspaces L of H with respect to which a multiple of the normalized tight frame $\{S^{1/2}(x_i)\}_i$ is the closest normalized tight frame to the given frame $\{x_i\}_i$ of the Hilbert subspace $K \subseteq H$, or at least one of the closest normalized tight frames?

If there are other closest normalized tight frames with respect to the selected distance measures, do they span the same Hilbert subspaces of H? If not, how are the positions of the subspaces with respect to $K \subseteq H$?

To obtain at least partial results authors usually have applied some additional restrictions to the set of frames to be considered: (i) resort to similar frames, (ii) resort to the case $K = L = H$, (iii) resort to special classes of frames like Gabor (Weyl-Heisenberg) or wavelet frames, and others. So one goal might be to lessen the restrictions in the suppositions.

We would like to list some existing results from [1], [3] and [4] to give a flavor of the existing successful approaches and to outline the wide open field of research to be filled. From recent correspondences with R. Balan we know about new findings of him and Z. Landau to be published in the near future ([2]).

First recall the major results by R. Balan ([1]): The frame $\{x_i\}_i$ of the Hilbert space H is said to be *quadratically close* to the frame $\{y_i\}_i$ of H if there exists a non-negative number C such that the inequality

$$\left\| \sum\nolimits_i c_i(x_i - y_i) \right\| \leq C \cdot \left\| \sum\nolimits_i c_i y_i \right\|$$

is satisfied. The infimum of all such constants C is denoted by $c(y,x)$. In general, if $C \geq c(y,x)$ then $C(1-C)^{-1} \geq c(x,y)$, however this distance measure is not reflexive. Two frames $\{x_i\}_i$ and $\{y_i\}_i$ of a Hilbert space H are said to be *near* if $d(x,y) = \log(\max(c(x,y), c(y,x)) + 1) < \infty$. They are near if and only if they are similar, [1, Th. 2.4]. The distance measure $d(x,y)$ is an equivalence relation and fulfills the triangle inequality.

Theorem 1([1]) For a given frame $\{x_i\}_i$ of H the distance measures admit their infima at

$$\min c(y,x) = \min c(x,y) = \frac{\sqrt{B}-\sqrt{A}}{\sqrt{B}+\sqrt{A}}\,,\ \min d(x,y) = \frac{1}{4}(\log(B)-\log(A))\,.$$

These values are achieved by the tight frames

$$\left\{\frac{\sqrt{A}+\sqrt{B}}{2}S^{1/2}(x_i)\right\}_i\,,\ \left\{\frac{2\sqrt{AB}}{\sqrt{A}+\sqrt{B}}S^{1/2}(x_i)\right\}_i\,,\{\sqrt[4]{AB}S^{1/2}(x_i)\}_i\,,$$

in the same order as the three measures are listed above. The solution may not be unique, in general, however any tight frame $\{y_i\}_i$ of H that achieves the minimum of one of the three distance measures $c(y,x)$, $c(x,y)$ and $d(x,y)$ is unitarily equivalent to the corresponding solutions listed above. The difference of the connecting unitary operator and the product of minimal distance times either $S^{1/2}$ or $S^{-1/2}$ fulfills a measure-specific operator norm equality.

A second class of examples has been treated by T. R. Tiballi, V. I. Paulsen and the author in 1998 ([3]). The foundations were laid by T. R. Tiballi in his Master Thesis in 1991 ([6]). Therein he was dealing with the symmetric orthogonalization of orthonormal bases of Hilbert spaces in a way that did not make use of the linear independence of the elements. So his techniques have been extendable to the situation of frames giving rise to the symmetric approximation of frames by normalized tight ones.

Theorem 2([3]) The operator $(P-|\theta^*|)$ is Hilbert-Schmidt if and only if the sum $\sum_{j=1}^{\infty}\|\mu_j - x_j\|^2$ is finite for at least one normalized tight frame $\{\mu_i\}_i$ of a Hilbert subspace L of H that is similar to $\{x_i\}_i$. In this situation the estimate

$$\sum\nolimits_{j=1}^{\infty}\|\mu_j - x_j\|^2 \geq \sum\nolimits_{j=1}^{\infty}\|S^{1/2}(x_j)-x_j\|^2 = \|(P-|\theta^*|)\|_{c_2}^2$$

is valid for every normalized tight frame $\{\mu_i\}_i$ of any Hilbert subspace L of H that is similar to $\{x_i\}_i$. (The left sum can be infinite for some choices of subspaces L and normalized tight frames $\{\mu_i\}_i$ for them.)

Equality appears if and only if $\mu_i = S^{1/2}(x_i)$ for any $i \in \mathbb{N}$. Consequently, *the symmetric approximation of a frame* $\{x_i\}_i$ *in a Hilbert space* $K \subseteq H$ is the normalized tight frame $\{S^{1/2}(x_i)\}_i$ spanning the same Hilbert subspace $L \equiv K$ of H and being similar to $\{x_i\}_i$ via the invertible operator $S^{-1/2}$.

Remark: (see [5]) If $\{x_i\}_i$ is a Riesz basis, then $\{S^{1/2}(x_i)\}_i$ is *the symmetric orthogonalization* of this basis. This is why the denotation 'symmetric approximation' has been selected.

A third approach has been developed by Deguang Han investigating approximation of Gabor (Weyl-Heisenberg) and wavelet frames. His starting point are countable *unitary systems* $\mathcal{U}$ on separable Hilbert spaces that contain the identity operator. In particular, $\mathcal{U}$ is supposed to be *group-like*, i.e. *group*$(\mathcal{U}) \subseteq \mathbb{T}\mathcal{U} = \{\lambda U : \lambda \in \mathbb{T}, U \in \mathcal{U}\}$. A vector $\phi \in H$ is a *complete frame vector* (resp., a *normalized tight frame vector*) for $\mathcal{U}$ if the set $\mathcal{U}\phi := \{U(\phi) : U \in \mathcal{U}\}$ is a frame (resp., a

normalized tight frame) of H. Two *frame vectors* $\phi, \psi \in H$ are said to be *similar* if the two frames $\mathcal{U}\phi$ and $\mathcal{U}\psi$ are similar frames in H. Let $\mathcal{T}(\mathcal{U})$ denote the set of all normalized tight frame vectors of H with respect to the action of $\mathcal{U}$.
As a matter of fact the distance measure used in [3] gives $\sum_{U\in\mathcal{U}} \|U(\xi) - U(\eta)\|^2 = \infty$ if $\mathcal{U}$ is an infinite set and $\xi \neq \eta$. Also, $\mathcal{U}\xi$ and $\mathcal{U}\eta$ are not similar, in general, cf. [1]. So define a vector $\psi \in \mathcal{T}(\mathcal{U})$ to be a *best normalized tight frame (NTF) approximation* for a given complete frame vector $\phi \in H$ of $\mathcal{U}$ if

$$\|\psi - \phi\| := \mathrm{dist}(\phi, \mathcal{T}(\mathcal{U})) := \inf\{\|\eta - \phi\| \,:\, \eta \in \mathcal{T}(\mathcal{U})\}\,.$$

Theorem 3([4]) Let $\mathcal{U}$ be a group-like unitary system acting on a Hilbert space H. Let $\phi \in H$ be a complete frame vector for $\mathcal{U}$. Then the vector $S^{1/2}(\phi)$ is the unique best NTF approximation for ϕ, where $S = (\theta^*\theta)^{-1}$ is the frame operator for ϕ.

Theorem 4([4]) Let $\Lambda \subset \mathbb{R}^d \times \mathbb{R}^d$ be a full-rank lattice and g be a Gabor frame generator associated with Λ. Then the vector $S^{1/2}(g)$ is the unique best NTF approximation for g, where S is the frame operator for g. ($S^{1/2}(g)$ is a Gabor frame generator, again.)

Considering the wavelet situation where the generating unitary systems sometimes are not group-like some obstacles are encountered. For example, D. Han found that for an orthonormal wavelet $\mathcal{U}_{D,T}(g)$ the vector $\phi = 1/4 \cdot g$ possesses better NTF approximations than $S^{1/2}(\phi)$. In ongoing discussions of D. Han with I. Daubechies, J. Wexler and M. Bownik examples of wavelet frames have been found for which there does not exist any wavelet-type dual frame. It is unknown whether $\{S^{1/2}(x_i)\}_i$ has always wavelet structure for wavelet frames $\{x_i\}_i$, or not.

Theorem 5([4]) Suppose, ϕ is the generator of a semi-orthogonal wavelet frame, i.e. $\phi_{m,k} \perp \phi_{n,l}$ for $\phi_{m,k} := |\det(M)|^{m/2}\phi(M^m x - k)$ and for any $k, l \in \mathbb{Z}^d$, all $m, n \in \mathbb{Z}$ with $m \neq n$. Denote by $\mathcal{U}_{D,T}$ the unitary system generating the initial wavelet frame. Then there exists a unique normalized tight wavelet frame generated by ψ such that the equality

$$\|\phi - \psi\| = \min\{\|h - \phi\| \,:\, h \in \mathcal{T}(\mathcal{U}_{D.T}),\, h \sim \phi\}$$

holds. Moreover, $\psi = S^{1/2}(\phi)$.

References

[1] R. Balan, Equivalence relations and distances between Hilbert frames, *Proc. Amer. Math. Soc.* **127**(1999), 2353-2366.
[2] R. Balan, Z. Landau, preprint in preparation, 2004.
[3] M. Frank, V. I. Paulsen, T. R. Tiballi, Symmetric approximation of frames and bases in Hilbert spaces, *Trans. Amer. Math. Soc.* **354**(2002), 777-793.
[4] D. Han, Approximations for Gabor and wavelet frames, *Trans. Amer. Math. Soc.* **355**(2003), 3329-3342.
[5] P.-O. Löwdin, On the nonorthogonality problem, *Adv. Quantum Chem.* **5**(1970), 185-199.

[6] T. R. Tiballi, Symmetric orthogonalization of vectors in Hilbert spaces, *Ph.D. Thesis*, University of Houston, Houston, Texas, U.S.A., 1991.

A Reproducing Kernel without (?) Discretization

Hartmut Führ

For any measurable subset $B \subset \mathbb{R}$, let $\mathcal{H}_B$ denote the space of functions in $\mathrm{L}^2(\mathbb{R})$ whose Fourier transforms are supported in B. It is a translation-invariant closed subspace of $\mathrm{L}^2(\mathbb{R})$.

Now pick an open, dense subset $A \subset \mathbb{R}$ of finite measure, and let $\mathcal{H} = \mathcal{H}_A$. Then the fact that the characteristic function of A is in L^2 implies for all f that $f = f * g$, where g is the inverse Fourier transform of said characteristic function. Hence convolution with g acts as a reproducing kernel on $\mathcal{H}$. (Put differently: The function g is a coherent state.)

Question: Does there exist a subset $\Gamma \subset \mathbb{R}$ and the function $\eta \in \mathcal{H}$ such that the Γ-shifts of η are a frame? (A tight frame even?). The reason for choosing this particular set A is the following simple observation:

Proposition. *There exists $g \in \mathcal{H}_B$ such that the $\alpha\mathbb{Z}$-shifts of g are total in $\mathcal{H}_B$ iff $|\frac{k}{\alpha} + B \cap B| = 0$ for all nonzero $k \in \mathbb{Z}$, where $|\cdot|$ denotes Lebesgue measure.*

Hence, by choice of A, we have for all $\Gamma = \alpha\mathbb{Z}$, that the Γ-shifts of an arbitrary function are not total in $\mathcal{H}$. This fact suggests that there cannot exist a frame consisting of shifts of a single function, but I have not been able to prove it.

Understanding the problem could help to clarify the role of the integrability conditions which appear in the sampling theorems of Feichtinger and Gröchenig (e.g., [2] and related papers). The problem can also be phrased as follows: Does there exist a frame of exponentials for $\mathrm{L}^2(A)$? This formulation is reminiscent of spectral sets and the Fuglede conjecture [3]. Finally, the fact that regularly spaced sampling sets do not work, no matter how small the step-size, suggests using perturbation techniques (see e.g. [1]). Hence an understanding of this problem would also shed some light on the scope of these techniques.

References

[1] O. Christensen: *An introduction to frames and Riesz bases.* Birkhäuser, Boston, 2003.
[2] H.G. Feichtinger and K.H. Gröchenig, *Banach spaces related to integrable group representations and their atomic decompositions. I.* J. Funct. Anal. **86** (1989), 307-340.
[3] B. Fuglede, *Commuting self-adjoint partial differential operators and a group theoretic problem,* J. Funct. Anal. **16** (1974), 101-121.

Density for Gabor Schauder Bases

Christopher Heil

Let H be a Hilbert space. A sequence $\{f_i\}_{i\in\mathbf{N}}$ of vectors in H is a *Schauder basis* for H if for each $f \in H$ there exist unique scalars $c_i(f)$ such that $f = \sum_{i=1}^{\infty} c_i(f) f_i$.

In this case there exists a *dual basis* $\{\tilde{f}_i\}_{i\in\mathbf{N}}$ such that $f = \sum_{i=1}^{\infty}\langle f, \tilde{f}_i\rangle\, f_i$. However, in general this series may converge only conditionally, i.e., it might not converge if a different ordering of the series is used.

A sequence $\{f_i\}_{i\in\mathbf{N}}$ is a *frame* for H if there exist constants A, $B > 0$ such that $A\|f\|^2 \le \sum_{i=1}^{\infty} |\langle f, f_i\rangle|^2 \le B\|f\|^2$ for every $f \in H$. In this case there exists a *dual frame* $\{\tilde{f}_i\}_{i\in\mathbf{N}}$ such that $f = \sum_{i=1}^{\infty}\langle f, \tilde{f}_i\rangle\, f_i$. Moreover, this series converges unconditionally, i.e., every reordering converges. However, the coefficients $\langle f, \tilde{f}_i\rangle$ in the series need not be unique, i.e., there may exist some other coefficients c_i such that $f = \sum_{i=1}^{\infty} c_i f_i$.

A frame is a Schauder basis if and only if it is a Riesz basis, i.e., the image of an orthonormal basis under a continuous bijection of H onto itself. For more information on frames, Schauder bases, and Riesz bases, see [Hei97] or [Chr03].

Let $T_a f(x) = f(x-a)$ denote the operation of translation. In [Zal78], [Zal80], Zalik gave some necessary and some sufficient conditions on $g \in L^2(\mathbf{R})$ and countable subsets $\Gamma \subset \mathbf{R}$ such that $\{T_a g\}_{a\in\Gamma}$ is complete in $L^2(\mathbf{R})$. Olson and Zalik proved in [OZ92] that no such system of pure translations can be a Riesz basis for $L^2(\mathbf{R})$, and conjectured that no such system can be a Schauder basis. This conjecture is still open.

In [CDH99], it was observed that no such system of pure translations can form a frame for $L^2(\mathbf{R})$. This is a corollary of the following general result due to Ramanathan and Steger [RS95].

Theorem 1. Let $g \in L^2(\mathbf{R})$ and let $\Lambda \subset \mathbf{R}^2$ be given. Then the *Gabor system* $\mathcal{G}(g,\Lambda) = \{e^{2\pi i b x} g(x-a)\}_{(a,b)\in\Lambda}$ has the following properties.

(a) If $\mathcal{G}(g,\Lambda)$ is a frame for $L^2(\mathbf{R})$, then $1 \le D^-(\Lambda) \le D^+(\Lambda) < \infty$.

(b) If $\mathcal{G}(g,\Lambda)$ is a Riesz basis for $L^2(\mathbf{R})$, then $D^-(\Lambda) = D^+(\Lambda) = 1$.

(c) If $D^-(\Lambda) < 1$ then $\mathcal{G}(g,\Lambda)$ is not a frame for $L^2(\mathbf{R})$.

In this result, $D^{\pm}(\Lambda)$ denote the Beurling densities of Λ, which provide in some sense upper and lower limits to the average number of points of Λ inside unit squares. More precisely, to compute Beurling density we count the average number of points inside squares of larger and larger radii and take the limit, yielding the definitions

$$D^-(\Lambda) = \liminf_{r\to\infty} \inf_{z\in\mathbf{R}^2} \frac{|\Lambda \cap Q_r(z)|}{r^2},$$
$$D^+(\Lambda) = \limsup_{r\to\infty} \sup_{z\in\mathbf{R}^2} \frac{|\Lambda \cap Q_r(z)|}{r^2},$$

for the lower and upper Beurling densities of Λ. Here $Q_r(z)$ is the square in $\mathbf{R}^2$ centered at z with side lengths r and $|E|$ denotes the cardinality of a set E. In particular, the Beurling density of a rectangular lattice is $D^-(\alpha\mathbf{Z}\times\beta\mathbf{Z}) = D^+(\alpha\mathbf{Z}\times\beta\mathbf{Z}) = \frac{1}{\alpha\beta}$.

Some corrections and extensions to Ramanathan and Steger's result are given in [CDH99], and a suite of new results on redundancy of frames partly inspired by their proof are given in [BCHL03a], [BCHL03b], [BCHL04].

Since $\{T_a g\}_{a\in\Gamma} = \mathcal{G}(g, \Gamma \times \{0\})$ and $D^-(\Gamma \times \{0\}) = 0$, it follows from Theorem 1 that such a system can never be a frame for $L^2(\mathbf{R})$.

Little is known about Gabor systems that are Schauder bases but not Riesz bases for $L^2(\mathbf{R})$. One example of such a system is $\mathcal{G}(g, \mathbf{Z}^2)$ where

$$g(x) = |x|^\alpha \, \chi_{[-\frac{1}{2},\frac{1}{2}]}(x), \;\; 0 < \alpha < \frac{1}{2}.$$

It was conjectured in [DH00] that Gabor Schauder bases follow the same Nyquist-type rules as Gabor Riesz bases, i.e., if $\mathcal{G}(g, \Lambda)$ is a Gabor Schauder basis then $D^-(\Lambda) = D^+(\Lambda) = 1$. Some partial results were obtained in [DH00], but the conjecture remains open. If this conjecture is proved, then the Olson/Zalik conjecture follows as a corollary.

Another open question is whether there is an analogue of the Balian–Low Theorem for Gabor Schauder bases. Qualitatively, the Balian–Low Theorem states that any Gabor Riesz basis will be generated by a function which is either not smooth, or has very poor decay at infinity. For a survey of the Balian–Low Theorem, see [BHW95].

Finally, for recent wavelet versions of density theorems, see [HK03]. There are interesting differences between the density-type theorems for Gabor and wavelet frames, most notably that there is no Nyquist-like cutoff in the possible densities for wavelets. An open general problem is to derive more powerful necessary or sufficient conditions for the existence of Gabor or wavelet frames.

References

[BCHL03a] R. Balan, P. G. Casazza, C. Heil, and Z. Landau, Deficits and excesses of frames, *Adv. Comput. Math.*, **18** (2003), pp. 93–116.

[BCHL03b] R. Balan, P. G. Casazza, C. Heil, and Z. Landau, Excesses of Gabor frames, *Appl. Comput. Harmon. Anal.*, **14** (2003), pp. 87–106.

[BCHL04] R. Balan, P. G. Casazza, C. Heil, and Z. Landau, Density, over-completeness, and localization of frames, preprint (2004).

[BHW95] J. J. Benedetto, C. Heil, and D. F. Walnut, Differentiation and the Balian–Low Theorem, *J. Fourier Anal. Appl.*, **1** (1995), pp. 355–402.

[Chr03] O. Christensen, "An Introduction to Frames and Riesz Bases," Birkhäuser, Boston, 2003.

[CDH99] O. Christensen, B. Deng, and C. Heil, Density of Gabor frames, *Appl. Comput. Harmon. Anal.*, **7** (1999), pp. 292–304.

[DH00] B. Deng and C. Heil, Density of Gabor Schauder bases, in: "Wavelet Applications in Signal and Image Processing VIII" (San Diego, CA, 2000), A. Aldroubi at al., eds., Proc. SPIE Vol. 4119, SPIE, Bellingham, WA, 2000, pp. 153–164.

[Hei97] C. Heil, "A Basis Theory Primer," manuscript, 1997, `http://www.math.gatech.edu/~heil`.

[HK03] C. Heil and G. Kutyniok, Density of weighted wavelet frames, *J. Geometric Analysis*, **13** (2003), pp. 479–493.

[OZ92] T. E. Olson and R. A. Zalik, Nonexistence of a Riesz basis of translates, in: "Approximation Theory," Lecture Notes in Pure and Applied Math., Vol. 138, Dekker, New York, 1992, pp. 401–408.

[RS95] J. Ramanathan and T. Steger, Incompleteness of sparse coherent states, *Appl. Comput. Harmon. Anal.*, **2** (1995), pp. 148–153.

[Zal78] R. A. Zalik, On approximation by shifts and a theorem of Wiener, *Trans. Amer. Math. Soc.*, **243** (1978), pp. 299–308.

[Zal80] R. A. Zalik, On fundamental sequences of translates, *Proc. Amer. Math. Soc.*, **79** (1980), pp. 255–259.

Two Problems on Frames and Decomposition of Operators

David R. Larson

In the *Abstracts of Talks* section we showed a relation between frame theory and projection (and other) decompositions of positive operators on a Hilbert space H. If S is a positive invertible operator in $B(H)$, for H a real or complex separable Hilbert space with infinite dimension, and if $\|S\|_{ess} > 1$, then S can be written $S = \sum_n P_n$, where $\{P_n\}$ is a sequence of self-adjoint (i.e.*orthogonal*) projections. This is equivalent to the property that S is the frame operator for a frame (for all of H) consisting of unit vectors. More generally, it was shown that if T is a positive operator (not necessarily invertible) which has essential norm strictly greater than 1, then T admits such a projection decomposition. If T has closed range, then writing $T = \sum_n x_n \otimes x_n$,where the x_n are unit vectors, yields that $\{x_n\}$ is a frame of unit vectors for the range of T. If T does not have close range, then $\{x_n\}$ is a sequence of unit vectors which does not constitute a frame for its closed span (i.e. the closed range of T), but can be filled out in many ways with unit vectors to give a tight frame for its closed span. (Just choose a positive operator R of norm > 1 such that $T + R$ is a scalar multiple of P, where P is the orthogonal projection onto the closure of range(T).) Since projection decompositions of positive operators seem to be useful when they exist, this suggests some problems in single operator theory.

PROBLEM A: When does a positive operator $T \in B(H)$ have a *finite* projection decomposition? That is, when can it be written as a *finite* sum of orthogonal projections?

Suppose, in fact, we assume that T has an *infinite* projection decomposition. Is it a common occurrence for T to also admit a *finite* projection decomposition? Or does this rarely happen?

In the context of the above problem, it is clear that if T is an invertible operator which has a finite projection decomposition, then it is the frame operator for a frame which is the union of *finitely many orthonormal sets* of vectors.

Also in the context of Problem A, we mention that it is easy to show (it is a lemma in the second paper) that if a positive operator of norm exactly 1 has a projection decomposition, then in fact it *must be* a projection. So it has a finite projection decomposition consisting of one projection. On the other hand, if a positive operator has essential norm strictly greater than 1, then we know it has an infinite projection decomposition (by a theorem in the extended abstract), but does it also have a finite projection decomposition?

We now discuss a problem concerning more general *targeted* decompositions of positive operators. *Targeted* means that we are asking for a decomposition as a (strongly convergent) series of simpler positive operators (such as projections, or rank-one operators satisfying specified norm or other properties). In paper 2 that was discussed in the abstract, which was joint work with K. Kornelson, we found techniques to deal with problems of specified norm targeted decompositions. In this Workshop, in response to a short talk on targeted decompositions presented by D. Larson in a problem session, H. Feichtinger and K. Grochenig pointed out that similar techniques just may lead to progress on a certain problem in modulation space theory. Subsequently, Larson and C. Heil discussed this matter, and there are plans to follow up on this lead. The following problem *seems* to be pointing in the right direction. At the least, it seems to be a *representative problem* on the concept of *targeted decompositions*, which is mathematically interesting (at least to this investigator) as a problem in Hilbert space operator theory, and which was motivated by Workshop discussions. We present it in this spirit. It concerns targeted decompositions of trace-class operators, hence is a problem in a different direction from the results in both papers discussed in the *Abstract*.

PROBLEM B: Let H be an infinite dimensional separable Hilbert space. As usual, denote the Hilbert space norm on H by $\|\cdot\|$. If x and y are vectors in H, then $x \otimes y$ will denote the operator of rank one defined by $(x \otimes y)z = (z, y)x$. The operator norm of $x \otimes y$ is then just the product of $\|x\|$ and $\|y\|$.

Fix an orthonormal basis $\{e_n\}_n$ for H. For each vector v in H, define

$$|||v||| = \sum_n |(v, e_n)|$$

This may be $+\infty$.

Let L be the set of all vectors v in H for which $|||v|||$ is finite. Then L is a dense linear subspace of H, and is a Banach space in the *triple norm*. It is of course isomorphic to ℓ^1

Let T be any positive trace-class operator in $B(H)$. The usual eigenvector decomposition for T expresses T as a strongly convergent series of operators $h_n \otimes h_n$, where $\{h_n\}$ is an orthogonal sequence of eigenvectors of T. That is,

$$T = \sum_n h_n \otimes h_n$$

In this representation the eigenvalue corresponding to the eigenvector h_n is the square of the norm: $\|h_n\|^2$. The trace of T is then

$$\sum_n \|h_n\|^2$$

and since T is positive this is also the trace-class norm of T.

Let us say that T is of *typeA* with respect to the orthonormal basis $\{e_n\}$ if, for the eigenvectors $\{h_n\}$ as above, we have that $\sum_n |||h_n|||^2$ is finite. Note that this is just the (somewhat unusual) formula displayed above for the trace of T with the triple norm used in place of the usual Hilbert space norm of the vectors $\{h_n\}$. (So,in particular, such operators T must be of trace class.)

And let us say that T is of *typeB* with respect to the orthonormal basis $\{e_n\}$ if there is *some* sequence of vectors $\{v_n\}$ in H with $\sum_n |||v_n|||^2$ *finite* such that

$$T = \sum_n v_n \otimes v_n$$

where the convergence of this series is in the strong operator topology. (Of course, *typeA* wrt a basis implies *typeB* wrt that basis. It is the *converse* direction that we want to consider.)

The problem we wish to isolate is the following: Let $\{e_n\}$ be an orthonormal basis for H. Find a characterization of all positive trace class operators T that are of *typeB* with respect to $\{e_n\}$. In particular, is the class of *typeB* operators with respect to a fixed orthonormal basis for H much larger than the class of *typeA* operators (with respect to that basis)?

Quantitative Behaviour of Wavelet Bases

Karsten Urban

We demonstrate a typical wavelet discretization of an elliptic problem and give some examples of condition numbers indicating the corresponding research problems.

Wavelet Representation of Differential Operators. For simplicity, let us consider the periodic 1D problem finding a function $u : [0,1] \to \mathbb{R}$ such that

$$-u''(x) + u(x) = f(x), \quad x \in (0,1), \qquad u(0) = u(1). \tag{1}$$

Let $\psi \in H^1(\mathbb{R})$ be a sufficiently smooth wavelet, we consider the periodic wavelet basis

$$\psi_{j,k}(x) := 2^{j/2} \sum_{\ell \in \mathbb{Z}} \psi(2^j(x+\ell) - k)|_{[0,1]} \tag{2}$$

and the corresponding wavelet spaces $W_j := \text{span}\,(\Psi_j)$ for $\Psi_j := \{\psi_{j,k} : k = 0, \dots, 2^j - 1\}$, for $j \geq 0$ and $S_0 := \{c : c \in \mathbb{R}\} = \text{span}\,(\Phi_0)$, $\Phi_0 = \{\chi_{(0,1)}\}$, so that

$$H^1_{\text{per}}(\mathbb{R}) = S_0 \oplus \bigoplus_{j \in \mathbb{N}} W_j,$$

and $\Psi := \Phi_0 \cup \bigcup_{j \in \mathbb{N}} \Psi_j$ is a Riesz basis for $L_2(0,1)$. Let us describe the structure of the wavelet representation $a(\Psi, \Psi)$ of the differential operator in (1), where here we have $a(u,v) := (u',v')_0 + (u,v)_0$. In order to give an impression of the entries let us consider index pairs (j,k) and (ℓ,m) such that the corresponding shifted translates are completely located inside the interval $(0,1)$, i.e.,

$$\psi_{j,k}(x) = 2^{j/2}\,\psi(2^j x - k), \qquad \psi_{\ell,m}(x) = 2^{\ell/2}\,\psi(2^\ell x - m), \qquad x \in [0,1].$$

Then, we obtain by a simple change of variables $y = 2^j x - k$

$$\begin{aligned}(\psi_{j,k}, \psi_{\ell,m})_0 &= 2^{(j+\ell)/2} \int_{\mathbb{R}} \psi(2^j x - k)\, \psi(2^\ell x - m)\, dx \\ &= 2^{(\ell-j)/2} \int_{\mathbb{R}} \psi(y)\, \psi(2^\ell(2^{-j}(y+k) - m)\, dy \\ &= 2^{(\ell-j)/2} (\psi, \psi_{l-j, m-2^{l-j}k})_0,\end{aligned}$$

i.e., the inner product only depends on the *level difference* and the relative location in space. In particular, there is only a fixed number of non-zero values per level difference $\ell - j$.

For the second-order term in $a(\cdot,\cdot)$ we use the fact that for every (sufficiently smooth) wavelet ψ there exists a second wavelet ψ^* such that $\psi'(x) = 4\,\psi^*(x)$, [1]. Then, a similar calculation as above gives

$$(\psi'_{j,k}, \psi'_{\ell,m})_0 = 2^{j+\ell+4}\, 2^{(\ell-j)/2} (\psi^*, \psi^*_{l-j,m-2^{l-j}k})_0.$$

We observe the same behaviour as above, i.e., the values depend only on the level difference and there is only a fixed number of non-zero entries per level. This shows the finger structure of the matrix which is ordered level-wise. This block-structure is also shown in the figure.

Preconditioning. Recalling the norm equivalence

$$\left\| \sum_{j,k} d_{j,k}\, \psi_{j,k} \right\|_s^2 \sim \sum_{j,k} 2^{2js} |d_{j,k}|^2, \qquad s \in (-\tilde{\gamma}, \gamma),$$

where $\gamma, \tilde{\gamma} > 1$ depend on the wavelet ψ and its dual $\tilde{\psi}$, we obtain the following preconditioning for $u = \boldsymbol{d}^T \Psi = \sum_{j,k} d_{j,k} \psi_{j,k}$

$$a(u,u) = \|u'\|_0^2 + \|u\|_0^2 = \|u\|_1^2 \sim \sum_{j,k} 2^{2j} |d_{j,k}|^2,$$

i.e., we obtain the preconditioner $\boldsymbol{C}_1 = \mathrm{diag}(2^{|\lambda|})$. As an alternative, the norm equivalence also leads to the preconditioner $\boldsymbol{C}_2 = \mathrm{diag}(\sqrt{2^{|\lambda|+1}})$. Both preconditioners are asymptotically optimale, i.e.,

$$\mathrm{cond}(\boldsymbol{C}_i^{-1} a(\Psi,\Psi) \boldsymbol{C}_i^{-1}) < \infty, \quad i = 1, 2.$$

It is not clear a priori which one is better in a practical application.

From a practical point of view, one can also try to use the diagonal of $a(\Psi,\Psi)$ as a preconditioner which is easily accessible and already contains information of the matrix. In Table 1, we have listed condition numbers of slices of the matrices corresponding to the level j.

j	cond($a(\Psi, \Psi)$)	$\boldsymbol{C}_1$	$\Vert\boldsymbol{A}\Vert$	diag	$\boldsymbol{C}_2$
2	87.47	17.74	15.07	4.07	14.58
3	352.81	21.40	18.73	5.94	20.15
4	1.43e+3	25.19	20.43	6.67	24.52
5	5.75e+3	27.52	21.05	7.47	28.53
6	2.30e+4	29.12	21.26	8.45	30.50
7	9.20e+4	29.68	21.41	9.47	31.34
8	3.67e+5	29.93	21.46	9.62	32.06
9	1.47e+6	30.13	21.47	10.06	32.35
10	5.88e+6	30.20	21.48	10.76	32.57
11	2.35e+7	30.27	21.48	9.82	32.68

Table 1. Condition numbers.

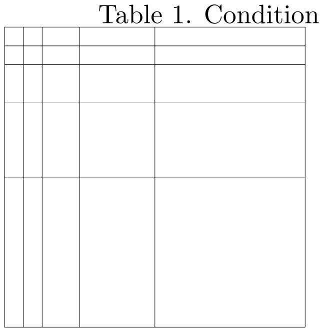


Figure 1. Block structure of $\boldsymbol{A}$.

Even though these numbers give a quite impressive rate of reduction, one should keep several facts in mind:

- The periodic case is the most simple one and of academic character only. When introducing Dirichlet boundary conditions even on an interval, the numbers increase significantly.
- When considering problems in 2D or 3D, even on the unit square or unit cube using tensor products, one has to square or cubic the condition numbers.
- When dealing with complex domains on needs several unit cubes and also certain combinations.

Adaptive Richardson Iteration. When using the adaptive wavelet method within the Richardson iteration directly, the condition number of $\boldsymbol{A}$ influences the error reduction factor ρ directly. In fact, it has been observed that $\rho \approx 1$ in many situations. Several attempts have been made in order to improve the condition numbers as listed in Table 1. However, none of them was really successful, which might also be explained by the low numbers for the diagonal preconditioners.

An alternative to the 'standard' adaptive Richardson iteration would be to combine the Richardson method using the adaptive approximate operator application APPLY from [2] as an outer iteration with an inner loop solving the Galerkin problem on a fixed set of unknowns. It has in fact been observed that such an iteration is quantitatively better in several situations:

- $[\boldsymbol{v}, \Lambda] = APPLY(\boldsymbol{u}^{(i)}, \varepsilon_i)$;
- solve $A_\Lambda u_\Lambda = f_\Lambda$ and call the numerical approximation $\boldsymbol{u}^{(i+1)}$.

So far there is no theoretical backup for the behaviour of this algorithm. It could be possible to explain the quantitative improvement with the aid of the *finite section method* known in frame theory.

References

[1] Lemarié-Rieusset, P.G. (1992): Analyses Multi-résolutions non Orthogonales, Commutation entre Projecteurs et Derivation et Ondelettes Vecteurs à Divergence Nulle. Rev. Mat. Iberoamericana **8**, 221–236

[2] A. Cohen, W. Dahmen, and R.A. DeVore, *Adaptive Wavelet Schemes for Elliptic Operator Equations — Convergence Rates*, Math. Comput., **70**, No. 233 (2001), 27–75.

Two Problems on the Generation of Wavelet and Random Frames

Eric Weber

Problem 1. If $g \in L^2(\mathbb{R})$ is the Gaussian, and $\Lambda := \{(a_z, b_z) : z \in \mathbb{Z}^2\} \subset \mathbb{R}^2$ is any set such that

$$1 < D^-(\Lambda) \leq D^+(\Lambda) < \infty,$$

then the Gabor system

$$G(g, \Lambda) = \{e^{-2\pi i a_z x} g(x - b_z) : z \in \mathbb{Z}^2\}$$

is a frame for $L^2(\mathbb{R})$, where $D^-(\Lambda)$, $D^+(\Lambda)$ are the lower and upper Beurling densities, respectively, of Λ.

Is there a corresponding statement regarding wavelet frames? More specifically, is there a function $\psi \in L^2(\mathbb{R})$ such that for any set $\Gamma := \{(a_z, b_z) : z \in \mathbb{Z}^2\} \subset \mathbb{R}^2$ with

$$1 < D^-(\Gamma) \leq D^+(\Gamma) < \infty,$$

then the wavelet system

$$W(\psi, \Gamma) = \{|a_z|^{1/2} g(a_z x - b_z) : z \in \mathbb{Z}^2\}$$

is a frame for $L^2(\mathbb{R})$, where $D^-(\Gamma)$, $D^+(\Gamma)$ are the lower and upper affine densities, respectively, of Γ (see [Heil et al. 2003]).

Problem 2. Let H be a finite dimensional Hilbert space.

i) What is a reasonable definition of "random frame"?
ii) How does one construct a "random frame"?

We make the following remarks:

a) The idea of a random orthonormal basis has a reasonably good definition. Fix any orthonormal basis of H; each orthonormal basis of H then corresponds to a unitary operator from the fixed basis to the new one. The group of unitary operators on H is a compact group, hence possesses a finite Haar measure, which can be normalized to give a probability measure. This probability measure would correspond to a uniform density, since the Haar measure is invariant under multiplication.

b) There are several ways of constructing random orthonormal bases for $\mathbb{R}^d$. Randomly choose $d(d+1)/2$ numbers and place in the upper triangle of a matrix B; fill in the remaining entries such that $B^T = -B$. The spectrum then is purely imaginary, whence e^B has spectrum on the unit circle and hence is unitary. (The matrix e^B can only be approximated). A second method is due to Stewart [Stewart 1980].

c) A frame with N elements for a Hilbert space H of dimension d can be obtained by choosing any basis of a Hilbert space K of dimension N and projecting the basis onto any subspace of dimension d. Thus, it is possible to construct a "random" frame from a random (orthonormal) basis.

References

[Heil et al. 2003] Heil, C. and Kutyniok, G.: *Density of weighted wavelet frames*, *J. Geom. Anal.* **13**, (2003), 479–493.

[Stewart 1980] Stewart, G. W.: *The efficient generation of random orthogonal matrices with an application to condition estimators*, *SIAM J. Numer. Anal.* **17**, (1980) no. 3, 403–409.

Reporter: Gitta Kutyniok and Gestur Ólafsson

Participants

Prof. Dr. Akram Aldroubi
aldroubi@math.vanderbilt.edu
Dept. of Mathematics
Vanderbilt University
Stevenson Center 1326
Nashville, TN 37240 – USA

Prof. Dr. Ole Christensen
Ole.Christensen@mat.dtu.dk
Department of Mathematics
Technical University of Denmark
Bldg. 303
DK-2800 Lyngby

Prof. Dr. Hans Georg Feichtinger
hans.feichtinger@univie.ac.at
Fakultät für Mathematik
Universität Wien
Nordbergstr. 15
A-1090 Wien

Dr. Massimo Fornasier
mfornasi@math.unipd.it
Institut für Mathematik
Universität Wien
Strudlhofgasse 4
A-1090 Wien

Prof. Dr. Michael Frank
mfrank@imn.htwk-leipzig.de
HTWK Leipzig
FB Informatik, Mathematik und
Naturwissenschaften
Gustav-Freytag-Str. 42A
D–04277 Leipzig

Dr. Hartmut Führ
fuehr@gsf.de
Institut für Biomathematik & Biometrie
GSF Forschungszentrum Neuherberg
Ingolstädter Landstr. 1
D–85764 Neuherberg

Prof. Dr. Karlheinz Gröchenig
karlheinz.groechenig@gsf.de
Institut für Biomathematik & Biometrie
GSF Forschungszentrum Neuherberg
Ingolstädter Landstr. 1
D–85764 Neuherberg

Prof. Dr. Christopher E. Heil
heil@math.gatech.edu
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332-0160 – USA

Prof. Dr. Matthias Holschneider
hols@ipgp.jussieu.fr
hols@math.uni-potsdam.de
Institut für Informatik
Universität Potsdam
Am Neuen Palais 10
D–14469 Potsdam

Ilona Iglewska-Nowak
iglewska@rz.uni-potsdam.de
Institut für Mathematik
Universität Potsdam
Am Neuen Palais 10
D–14469 Potsdam

Prof. Dr. Palle E.T. Jorgensen
jorgen@math.uiowa.edu
Dept. of Mathematics
University of Iowa
Iowa City, IA 52242-1466 – USA

Dr. Norbert Kaiblinger
norbert.kaiblinger@univie.ac.at
Fakultät für Mathematik
Universität Wien
Nordbergstr. 15
A-1090 Wien

Dr. Gitta Kutyniok
gittak@uni-paderborn.de
Institut für Mathematik
Fakultät EIM
Warburger Str. 100
D–33098 Paderborn

Prof. Dr. Dave Larson
larson@math.tamu.edu
Department of Mathematics
Texas A & M University
College Station, TX 77843-3368 – USA

Prof. Dr. Gestur Olafsson
olafsson@math.lsu.edu
Department. of Mathematics
Louisiana State University
Baton Rouge, LA 70803-4918 – USA

Prof. Dr. Karsten Urban
kurban@mathematik.uni-ulm.de
Abteilung Numerik
Universität Ulm
Helmholtzstr. 18
D–89069 Ulm

Prof. Dr. Eric Weber
esweber@iastate.edu
Department. of Mathematics
Iowa State University
400 Carver Hall
Ames, IA 50011 – USA

Mathematisches Forschungsinstitut Oberwolfach

Report No. 11/2004

Computational Electromagnetism

Organised by
Ralf Hiptmair (Zürich)
Ronald H.W. Hoppe (Augsburg)
Ulrich Langer (Linz)

February 22nd – February 28th, 2004

Introduction by the Organisers

The field of computational electromagnetism is dedicated to the design and analysis of numerical methods for the approximate solution of electromagnetic field problems. Since the exploitation of electromagnetic phenomena is one of the foundations of modern technology, computational electromagnetics is of tremendous industrial relevance: in a sense, it is peer to computational solid and fluid mechanics and huge research efforts are spent on developing and enhancing simulation methods and software for electromagnetic field computations.

For a long time, computational electromagnetism remained a realm of engineering research with applied mathematics shunning the area. This was in stark contrast to elasticity and fluid mechanics, where mathematicians have been involved in the development of numerical methods from the very beginning. Maybe, the blame has to be laid on the incorrect belief of mathematicians who thought that the laws governing the behavior of electromagnetic fields basically boil down to well understood second-order elliptic problems.

Fortunately, the past fifteen years have seen a real surge of mathematical research activities in the area of computational electromagnetism. This resulted in insights that have begun to have a big impact on the numerical methods used in engineering and industrial environments. A prominent example is the explanation of so-called spurious solutions that can arise when using continuous "nodal" finite elements for the discretization of certain electromagnetic boundary value or eigenvalue problems, respectively. Another example is the appreciation of so-called edge finite elements and the construction of multilevel iterative solvers for the low-frequency setting.

Meanwhile, computational electromagnetism can claim to be a major area of numerical mathematics and scientific computing in its own right. This prompted us to ask the Mathematisches Forschungsinstitut Oberwolfach to host a one week workshop on computational electromagnetism, the first of its kind. Reflecting the growing importance of the subject, this workshop has been one of a series of events dedicated to mathematical issues in the computation of electromagnetic fields. We would like to mention, the NSF-CBMS Regional Conference in the Mathematical Sciences about “Numerical Methods in Forward and Inverse Electromagnetic Scattering”, held in Golden, CO, June 3-7, 2002 (from which the book [2] arose), and the “LMS Durham Symposium on Computational methods for Wave Propagation in Direct Scattering”, Durham, England, July 15-25, 2002 (see [1]).

This Oberwolfach workshop brought together some 50 experts in computational electromagnetism. The majority of the participants were applied mathematicians, but a sizable number of people with a background in engineering also attended, as appropriate for a field with close ties to engineering and the applied sciences. Nevertheless, the workshop had a clear mathematical focus, emphasizing rigorous theory, principles and ideas. Throughout, the presentations matched these expectations. A total of 29 presentations were given, of which ten were survey lectures offering broader treatment of a particular subject.

As is typical of an event that targets a specific area of application, it arose that a broad range of mathematical issues and techniques was addressed. Although it will certainly not do justice to many presentations, we will try categorize the talks as follows:

- **Mathematical modelling.** This subject did not play a central role, because most presentations took the model equations for granted. Modelling for practical engineering calculations was described by O. Bíró in his survey talk about *Practical Aspects of FEM in Electromagnetics*, p. 559, and by M. Clemens when speaking on *Formulations and Efficient Numerical Solution Techniques for Transient 3D Magneto-and Electro-Quasistatic Field Problems*, p. 572. Homogenization was addressed in the presentation by A. Bendali about *Two Scale Asymptotic Expansion for the Scattering of a TM-Electromagnetic Wave by a Rough Surface and Applications*, p. 556.
- **Spatial discretizations.** This turned out to be one of the core subjects of the workshop. The survey lectures of D. Boffi about *Theoretical Aspects of Edge Finite Elements*, p. 564 and I. Perugia on *Discontinuous Galerkin Methods for Maxwell's Equations*, p. 608, addressed the topic. Particular issues were discussed by S. Christiansen in his talk about the *Div-Curl Lemma for Edge Elements*, p. 571, and by J. Pasciak about *The Approximation of the Maxwell Eigenvalue Problem using a Least-Squares Method*, p. 606. M. Kaltenbacher gave an account of observations concerning finite element schemes in his presentation on *Nodal and Edge Finite Element Discretization of Maxwell's Equations*, p. 590.

Several presentations were devoted to higher order spatial discretization: the survey lecture of M. Ainsworth gave an account of the *Dispersive Properties of High Order Nédélec/Edge Elements for Maxwell's Equations*, p. 553, L. Demkowicz spoke about $H^1, \boldsymbol{H}(\mathrm{curl})$ *and* $\boldsymbol{H}(\mathrm{div})$*-Conforming Projection-Based Interpolation in Three Dimensions*, p. 582, and P. Ledger about *Computation of Maxwell Eigenvalues with Exponential Rates of Convergence*.

- **Timestepping.** There was only one contribution dealing with temporal discretization, namely the talk by T. Driscoll on *High-Order Time Stepping Methods for Electromagnetics*, p. 585.
- **Regularity of solutions.** Here one of the pioneers in the field, M. Dauge, gave a survey talk about *Singularities of Electromagnetic Fields in the Eddy Current Limit*, p. 574.
- **Integral equation methods**. Boundary element methods in the frequency domain were treated by S. Kurz in his talk on *A New View on Collocation*, p. 599. Conversely, time-domain integral equation methods were examined in the survey lecture by E. Michielssen on *Fast Time Domain Integral Equation Solvers*, p. 603, and P. Davies in her contribution on *Convergence of Collocation Methods for Time Domain Boundary Integral Equations*, p. 579. S. Börm talked about $\mathcal{H}^2$*-Matrices with Adaptive Cluster Bases Applied to an Eddy Current Problem*, p. 562, and presented a fast summation method for discrete frequency-domain integral equations.
- **Electromagnetic Scattering.** This topic was treated by R. Kress in his survey lecture on *Inverse Obstacle Scattering for Time-Harmonic Electromagnetic Waves*, p. 596. Also the talk by A. Bendali on *Two Scale Asymptotic Expansion for the Scattering of a TM-Electromagnetic Wave by a Rough Surface and Applications*, p. 556, addressed a particular scattering problem.
- **Absorbing boundary conditions.** A special incarnation of these was examined in the survey talk by F. Teixeira on *Perfectly Matched Layers*, p. 621. Details of a PML approach were studied by Z. Chen in his talk about *An Adaptive Perfectly Matched Layer Technique for Time-harmonic Scattering Problems*, p. 568. Other techniques were outlined by M. Grote (*Nonreflecting Boundary Conditions for Computational Electromagnetics*, p. 588) and F. Schmidt (*Pole Condition: A New Approach to Solve Scattering Problems*, p. 615).
- **Topological issues**. These were discussed in the talks of R. Kotiuga (*The Hurewicz Map Distinguishes Intuitive vs. Computable Topological Aspects of Computational Electromagnetics*, p. 593) and F. Rapetti (*Smith Normal Form as an Adequate Tool to Detect Mesh Defects as well as to Build Basis Fields for Domains with Loops and Holes*, p. 612).
- **Fast solvers.** Several speakers discussed fast algorithms for the solution of linear systems of equations arising from discretized field equations: it was the subject of J. Schöberl's survey lecture on *Preconditioning for Maxwell*

Equations, p. 617, and O. Sterz' talk on *Adaptive Multigrid-Methods for the Solution of Time-Harmonic Eddy-Current Problems*, p. 618. The use of multigrid methods was discussed in the contributions by M. Clemens on *Formulations and Efficient Numerical Solution Techniques for Transient 3D Magneto-and Electro-Quasistatic Field Problems*, p.'572, and by M. Kaltenbacher on *Nodal and Edge Finite Element Discretization of Maxwell's Equations*, p. 590. An enhancement for algebraic multigrid was proposed by P. Arbenz (*Treatment of Nullspace in Maxwell Problem*, p. 553). J. Zou dealt with domain decomposition methods in his contribution on *Some New Inexact Uzawa Methods and Non-overlapping DD Preconditioners for Solving Maxwell's Equations in Non-homogeneous Media*, p. 624.
- **Adaptive techniques.** Only one presentation, that of Z. Chen on *An Adaptive Perfectly Matched Layer Technique for Time-harmonic Scattering Problems*, p. 568, dealt with a special adaptive scheme.
- **Optimization.** This important subject reaches beyond the core of computational electromagnetism. An aspect was discussed in the talk by D. Lukáš on *Computational Shape and Topology Optimization with Applications to 3–Dimensional Magnetostatics* p. 601.

We would like to add our personal impression that two families of methods have been received with particular interest during the workshop:
- Time-domain integral equation methods,
- High-order spatial discretization schemes.

We are sure that the workshop will have made a substantial contribution to the progress of research in these and all other areas of computational electromagnetism.

R. Hiptmair
R.H.W. Hoppe
U. Langer

References

[1] M. Ainsworth, P. Davis, D. Duncan, P. Martin, and B. Rynne, eds., *Topics in Computational Wave Propagation. Direct and inverse Problems*, vol. 31 of Lecture Notes in Computational Science and Engineering, Springer, Berlin, 2003.

[2] P. Monk, *Finite Element Methods for Maxwell's Equations*, Clarendon Press, Oxford, UK, 2003.

Workshop on Computational Electromagnetism

Table of Contents

Abstracts

Dispersive Properties of High Order Nédélec/Edge Elements for Maxwell's Equations.

Mark Ainsworth

The dispersive behaviour of high order Nédélec element approximation of the time harmonic Maxwell equations at a prescribed temporal frequency ω on tensor product meshes of size h is analysed. A simple argument is presented showing that the discrete dispersion relation may be expressed in terms of the discrete dispersion relation for the approximation of the scalar Helmholtz equation in one dimension. An explicit form for the one dimensional dispersion relation is given, valid for arbitrary order of approximation. Explicit expressions for the leading term in the error in the regimes where (a) ωh is small, showing that the dispersion relation is accurate to order $2p$ for a p-th order method; and (b) in the high wave number limit where $1 \ll \omega h$, showing that in this case the error reduces at a super-exponential rate once the order of approximation exceeds a certain threshold which is given explicitly. Details have been published in the following work [1–3]

References

[1] ———, *Discrete dispersion relation for hp-version finite element approximation at high wave number*, SIAM J. Numer. Anal., (accepted for publication).
[2] ———, *Dispersive and dissipative behaviour of high order discontinuous Galerkin finite element methods*, J. Comput. Phys., (accepted for publication).
[3] ———, *Dispersive properties of high order Nédélec/edge element approximation of the time-harmonic Maxwell equations*, Phil. Trans. Roy. Soc. Series A, (accepted for publication).

Treatment of Nullspace in Maxwell Problems

Peter Arbenz

1. Introduction

The weak form of the magnetostatic equation reads: Find $\mathbf{u} \in H_0(\mathbf{curl}\,,\Omega)$ such that

$$(1)\qquad \begin{array}{lll} \text{(a)} & (\mathbf{curl}\,\mathbf{u}, \mathbf{curl}\,\mathbf{\Psi}) = (\mathbf{r}, \mathbf{\Psi}), & \forall \mathbf{\Psi} \in H_0(\mathbf{curl}\,,\Omega), \\ \text{(b)} & (\mathbf{u}, \mathbf{grad}\,q) = 0, & \forall q \in H_0^1(\Omega), \end{array}$$

where $\Omega \in \mathbb{R}^3$ is a bounded domain with connected boundary $\partial\Omega$. We require that $(\mathbf{r}, \mathbf{grad}\,q) = 0$ for all $q \in H_0^1(\Omega)$ such that equation (1) is consistent.

The straightforward discretization of (1) by the finite element method yields the matrix equation

$$(2)\qquad \begin{array}{lll} \text{(a)} & A\mathbf{x} = M\mathbf{r}, & C^T\mathbf{r} = \mathbf{0} \\ \text{(b)} & C^T\mathbf{x} = \mathbf{0} & \end{array}$$

where $a_{ij} = (\mathbf{curl}\,\mathbf{\Psi}_i, \mathbf{curl}\,\mathbf{\Psi}_j)$, $m_{ij} = (\mathbf{\Psi}_i, \mathbf{\Psi}_j)$, $c_{i\ell} = (\mathbf{\Psi}_i, \mathbf{grad}\,\varphi_\ell)$. Here, the $\mathbf{\Psi}_i$, $i = 1, \ldots, n$, form a basis of the space N_h of lowest order Nédélec edge elements and the φ_ℓ, $\ell = 1, \ldots, m$, form a basis of the lowest order Lagrange elements L_h, see [4].

A has a m-dimensional nullspace $\mathcal{N}(A)$ that satisfies [4, §III.5.3]

$$\mathcal{N}(A) = \{\mathbf{v}_h \in N_h \mid \mathbf{curl}\,\mathbf{v}_h = \mathbf{0}\} = \mathbf{grad}\, L_h. \tag{3}$$

Thus, the gradient of each φ_ℓ can be written as a linear combination of the edge basis functions $\mathbf{\Psi}_j$,

$$\mathbf{grad}\,\varphi_\ell = \sum_{j=1}^{n} y_{j\ell}\mathbf{\Psi}_j.$$

Let $Y \in \mathbb{R}^{n\times m}$ be the matrix with elements $y_{j\ell}$. Then, $AY = 0$ and $C = MY$. The columns of Y form a *sparse* null space basis of A, see e.g. [2]. Notice that Y can be constructed from geometric properties of the finite element mesh.

In [1] we have investigated the numerical solutions of consistent semi-definite equations of the form (2). The key idea is to employ the sparse null space basis to extract a positive definite submatrix of A of order $n-m$, the rank of A.

Reitzinger and Schöberl [5] introduced an algebraic multigrid method to solve (2) regularized by a term that is positive on $\mathcal{N}(A)$. Here we present a way how to extend the ideas of [1] to all levels of the Reitzinger-Schöberl AMG algorithm. In this way we get an AMG algorithm that works entirely on the largest subspace of N_h on which $\|\mathbf{curl}\,(\cdot)\|$ is a norm. Its dimension $n-m$ is considerably smaller than n.

2. Elimination of the nullspace

Let's assume that the last m rows of Y are linearly independent. Then [1]

$$W := \begin{bmatrix} I_{n-m} & Y_1 \\ O & Y_2 \end{bmatrix}, \qquad Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}, \qquad Y_2 \in \mathbb{R}^{m\times m}, \tag{4}$$

is nonsingular. We split A, C, $\mathbf{x}$, and $\mathbf{r}$ according to Y. Then (2) becomes

$$W^T \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} W W^{-1}\mathbf{x} = \begin{bmatrix} A_{11} & O \\ O & O \end{bmatrix} W^{-1}\mathbf{x} = W^T M\mathbf{r} \iff \begin{cases} A_{11}\mathbf{x}_1 = \mathbf{r}_1, \\ \mathbf{x}_2 = \mathbf{0}. \end{cases} \tag{5}$$

A_{11} is symmetric positive definite. The general solution of (2) has the form

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{0} \end{bmatrix} + Y\mathbf{a}.$$

To satisfy the constraint $C^T\mathbf{x} = \mathbf{0}$ we determine $\mathbf{a}$ by solving

$$H\mathbf{a} = -C_1^T\mathbf{x}_1. \tag{6}$$

Here, H is the symmetric positive definite matrix with elements
$h_{ij} = (\mathbf{grad}\,\varphi_i, \mathbf{grad}\,\varphi_j)$.

3. Application to the Reitzinger-Schöberl AMG algorithm

Reitzinger and Schöberl [5] introduced an Algebraic Multigrid (AMG) method for solving (1)-(2) that properly treats the solenoidal and curl-free portions of the vector fields. The authors start from an AMG method for solving the Poisson (or a similar elliptic) problem in L_h. Coarse grids are constructed from fine grids by aggregating nodes into 'virtual nodes'. Two aggregates are defined connected (through 'virtual edges') if they contain nodes that are connected in the fine grid [6]. The system matrices on the various levels are denoted by H_k, where $H_0 = H$ corresponds to the finest level. Because of the Galerkin principle, among two consecutive levels the relation

$$H_{k+1} = Q_k^T H_k Q_k$$

holds. Q_k prolongates (interpolates) from level $k+1$ to the finer level k. Reitzinger and Schöberl then construct a sequence of levels for the curl-curl matrix. The matrices on the various levels are denoted by A_k with $A_0 = A$ and

$$A_{k+1} = P_k^T A_k P_k$$

where the P_k now prolongates from coarse to fine edge space. Q_k and P_k are related via the compatibility condition

$$P_k Y_{k+1} = Y_k Q_k \tag{7}$$

such that coarse grid gradients are prolongated to fine grid gradients. Here, Y_k is a sparse nullspace basis of A_k.

To eliminate the nullspace on all levels of the Reitzinger-Schöberl AMG we arrange the matrices A_k such that the nullspace bases can be written in the form

$$Y_k = \begin{bmatrix} Y_{k,1} \\ Y_{k,2} \end{bmatrix},$$

where $Y_{k,2}$ is a nonsingular submatrices of Y_k, cf. (4). With (5), we then get

$$\begin{aligned}
\begin{bmatrix} A_{k+1,11} & O \\ O & O \end{bmatrix} &= W_{k+1}^T \begin{bmatrix} A_{k+1,11} & A_{k+1,12} \\ A_{k+1,21} & A_{k+1,22} \end{bmatrix} W_{k+1} \\
&= W_{k+1}^T A_{k+1} W_{k+1} = W_{k+1}^T P_k^T A_k P_k W_{k+1} \\
&= W_{k+1}^T P_k^T \begin{bmatrix} A_{k,11} & A_{k,12} \\ A_{k,21} & A_{k,22} \end{bmatrix} P_k W_{k+1} \\
&= W_{k+1}^T P_k^T W_k^{-T} \begin{bmatrix} A_{k,11} & O \\ O & O \end{bmatrix} W_k^{-1} P_k W_{k+1}
\end{aligned}$$

and thus

$$W_k^{-1} P_k W_{k+1} = \begin{bmatrix} P_{k,11} - Y_{k,1} Y_{k,2}^{-1} P_{k,21} & O \\ Y_{k,2}^{-1} P_{k,21} & Q_k \end{bmatrix}.$$

So, the prolongator for the positive-definite portions of the systems is

$$\bar{P}_k := P_{k,11} - Y_{k,1} Y_{k,2}^{-1} P_{k,21}, \qquad A_{k+1,11} = \bar{P}_k^T A_{k,11} \bar{P}_k$$

These ideas can be incorporated in the Reitzinger-Schöberl AMG algorithm (or in a smoothed aggregation AMG algorithm like in [3]) as follows

(1) Build matrices the A_k and H_k. This implies that all the prolongators P_k and Q_k are available.
(2) Construct the nullspace bases Y_k on all levels.
(3) Reduce A_k to $A_{k,11}$
(4) Adapt the prolongators and smoothers.

A more memory-aware procedure works level by level starting with the finest.

References

[1] P. Arbenz and Z. Drmač. *On positive semidefinite matrices with known null space.* SIAM J. Matrix Anal. Appl., 24(1) (2002), pp132–149.

[2] P. Arbenz and R. Geus. *A comparison of solvers for large eigenvalue problems originating from Maxwell's equations.* Numer. Linear Algebra Appl., 6(1) (1999), pp3–16.

[3] P. B. Bochev, C. J. Garasi, J. J. Hu, A. C. Robinson, and R. S. Tuminaro. *An improved algebraic multigrid method for solving Maxwell's equations.* SIAM J. Sci. Comput., 25(2) (2003), pp623–642.

[4] V. Girault and P.-A. Raviart. *Finite Element Methods for the Navier-Stokes Equations.* Springer-Verlag, Berlin, 1986. (Springer Series in Computational Mathematics, 5).

[5] S. Reitzinger and J. Schöberl. *An algebraic multigrid method for finite element discretizations with edge elements.* Numer. Linear Algebra Appl., 9(3) (2002), pp223–238.

[6] P. Vaněk, J. Mandel, and M. Brezina. *Algebraic multigrid based on smoothed aggregation for second and fourth order problems.* Computing, 56(3) (1996), pp179–196.

Two Scale Asymptotic Expansion for the Scattering of a TM-Electromagnetic Wave by a Rough Surface and Applications

A. Bendali

(joint work with P. Borderies, J.-R. Poirier)

In this study, mainly of methodological interest, we show how the two-scale asymptotic expansion method [7] can be used as a powerful tool in the study of the scattering of an electromagnetic wave by a highly oscillating perfectly conducting surface both from the theoretical and the practical standpoint. More specifically, we consider the following simple 2D model related to the scattering of an E-polarized incident time-harmonic electromagnetic wave u^{inc}

$$(1)\qquad \begin{cases} \Delta u^\delta + k^2 u^\delta = 0 \ \text{ in } \ \Omega^\delta, \\ u^\delta = 0 \ \text{ on } \ \Gamma^\delta, \ x \to e^{-i\beta x} u^\delta(x,y) \text{ is periodic of period } L, \\ \text{Radiation Condition (RC) on } \ u^\delta - u^{\mathrm{inc}}. \end{cases}$$

The surface is considered as a periodic grating whose elementary cell is

$$\Omega^\delta := \left\{(x,y) \in \mathbb{R}^2 : 0 < x < L, \ \ y > \gamma_\delta(x)\right\}$$

in which $\Gamma^\delta := \left\{(x,y) \in \mathbb{R}^2 : 0 < x < L \mapsto \ \ y = \gamma_\delta(x)\right\}$ represents a sampling of the surface which is reproduced by periodicity. Data k and $\beta > 0$ are the wave number and the period respectively. The small parameter $\delta > 0$ characterizes the rapid oscillations of the surface and their small amplitude in the following way $\gamma_\delta(x) = s(x, x/\delta)$ where $(x,\sigma) \mapsto s(x,\sigma)$ is a given function, assumed to be smooth for simplicity, doubly periodic of period L in x and d in σ. The radiation condition is expressed by means of the Floquet expansion of u^δ (e.g., [5]). The existence and uniqueness of a solution to (1) are ensured by the stability estimates needed also to establish error bounds on the asymptotic expansion. Even much more involved, the general case can be treated along the same lines [4].

We briefly describe how to obtain a two-scale asymptotic expansion for u^δ, to derive an homogenized boundary condition on a flat surface from this expansion and finally to establish bounds on the error resulting from replacing the rough boundary condition by the homogenized one. The full details can be found in [3]. Actually, the results are known and can be obtained by the method of correctors [1, 2] or by the matching asymptotic expansions [6]. However, the correctors technique, a step by step process, does not give a clear overall idea of the full asymptotic expansion. In the matching asymptotic expansions method, slow and rapid variables are mixed in the boundary layer resulting in intricate analytical calculations to separate them.

For the two-scale asymptotic expansion that is considered here, its determination is first done by means of a formal process. Proven error bounds give it a complete justification a posteriori.

The unknown u^δ is decomposed in the following form

$$u^\delta(x,y) = U^\delta(x,y) + \Pi^\delta(x,\sigma,\tau)|_{\sigma=x/\delta,\tau=y/\delta}.$$

The variable x will play the role of a parameter in the part $\Pi^\delta(x,\sigma,\tau)$ containing the fast variables. It is assumed next that both U^δ and Π^δ have the following asymptotic expansions

$$U^\delta(x,y) = u^0(x,y) + \delta u^1(x,y) + \cdots + \delta^n u^n(x,y) + \cdots,$$
$$\Pi^\delta(x,\sigma,\tau) = \Pi^0(x,\sigma,\tau) + \delta\Pi^1(x,\sigma,\tau) + \cdots + \delta^n\Pi^n(x,\sigma,\tau) + \cdots.$$

Inserting this expansion in the Helmholtz equation and equating to zero the coefficients of powers of δ gives the following system

$$\left(\Delta_{\sigma,\tau}\Pi^n + 2\partial_x\partial_\sigma\Pi^{n-1} + \left(\partial_x^2 + k^2\right)\Pi^{n-2}\right)(x,\sigma,\tau) + \left(\Delta u^n + k^2 u^n\right)(x,y) = 0,$$
$$n = 0, 1, \ldots$$

with 0 for any term involving a negative index. Now, assuming that every derivative of Π^n satisfies

$$\lim_{\tau\to+\infty} \partial^\alpha_{x,\sigma,\tau}\Pi^n(x,\sigma,\tau) = 0$$

makes possible a separation of the functions depending on the slow and the rapid variables

$$\left(\Delta_{\sigma,\tau}\Pi^n + 2\partial_x\partial_\sigma\Pi^{n-1} + \left(\partial_x^2 + k^2\right)\Pi^{n-2}\right)(x,\sigma,\tau) = 0, \quad \left(\Delta u^n + k^2 u^n\right)(x,y) = 0.$$

Since Π^δ is living in a boundary layer of the surface, the radiation condition is on the slow variables functions only

$$\text{RC on } u^0 - u^{\text{inc}}, \quad \text{RC on } u^n \text{ for } n \geq 1.$$

However, the decisive advantage of the two scale asymptotic expansion is its suitability to deal with the boundary condition

$$\sum_{n\geq 0} \delta^n \left(u^n(x, \delta s(x,\sigma)) + \Pi^n(x,\sigma,s(x,\sigma))\right) = 0.$$

A simple Taylor expansion for $u^n(x, \delta s(x,\sigma))$ yields

$$\Pi^n(x,\sigma,s(x,\sigma)) + u^n(x,0) + \sum_{k=1}^{n} \frac{s(x,\sigma)^k}{k!}\partial_y^k u^{n-k}(x,0) = 0.$$

In this way, all the equations needed to determine the asymptotic expansion have been obtained at once. The following theorem is the main tool to do this determination.

Theorem 1. *Let F be a given periodic function of period d in σ in $C^\infty(\overline{D})$ verifying*

$$\Delta_{\sigma,\tau}{}^{m-1} F = 0 \text{ in } D \text{ and } |F(\sigma,\tau)| \leq ce^{-\gamma\tau},$$

and $G \in C^\infty(\mathbb{R})$, periodic of period d. Then, the boundary-value problem

$$\Delta\Pi = F \text{ in } D, \quad \Pi(\sigma, s(\sigma)) = G(\sigma),\ 0 < \sigma < d.$$

admits one and only one solution satisfying $|\Pi(\sigma,\tau) - \Pi^\infty| \leq ce^{-\gamma\tau}$.

Proof. The proof is based on a variational formulation in a weighted Sobolev space and elliptic interior estimates and Fourier series expansion. □

The different terms of the asymptotic expansion are then determined recursively by solving boundary-value problems in the slow and the rapid variables. For the zero order terms, we have $\Pi^0 = 0$ and u^0 is the solution to the problem with a flat boundary

$$\begin{cases} \Delta u^0 + k^2 u^0 = 0 \text{ for } y > 0 \\ u^0 = 0 \text{ for } y = 0,\ x \to e^{-i\beta x} u^0(x,y) \text{ is periodic of period } L, \\ \text{Radiation Condition (RC) on } u^0 - u^{\text{inc}}. \end{cases}$$

Note that, contrary to the corrector method, the flat plane problem has been obtained only by calculations without passing to any limit. Solving the auxiliary problem

$$\Delta_{\sigma,\tau} H = 0, H(x,\sigma,s(x,\sigma)) = s(x,\sigma).$$

yields $h(x) = \lim_{\tau\to\infty} H(x,\sigma,\tau)$. The term u^1 is then the solution of the following boundary-value problem

$$\begin{cases} \Delta u^1 + k^2 u^1 = 0 \text{ for } y > 0, \\ u^1(x,0) + h(x)\partial_y u^0(x,0) = 0, \text{ RC on } u^1. \end{cases}$$

Proceeding in the same way, one can determine the asymptotic expansion at any order. The rigorous justification of the method is then obtained through the error bound given in the following theorem

Theorem 2. *For any given $y_0 > 0$, there exists a constant independent of δ such that*

$$\left\|(u^\delta - (u^0 + \delta u^1)|_{y_0<y}\right\| \le c\delta^{3/2}.$$

Proof. Let be given a cut-off function $\chi \in \mathcal{D}(\mathbb{R})$ such that $\chi \equiv 1$ near 0 and $0 \le \chi \le 1$. The proof is obtained by means of an evaluation of the residuals of $u^\delta - (u^0 + \delta(u^1 + \chi(y)\Pi^1) + \delta^2\chi(y)\Pi^2)$ relatively to the equations of problem (1) and a suitable stability estimate for its solutions. □

The effective boundary condition at order 1 can then be written in terms of $h(x)$

$$u^{1,\delta}(x,0) + \delta h(x)\partial_y u^{1,\delta}(x,0) = 0. \tag{2}$$

and is used in place of the Dirichlet boundary condition in problem (1). The main result concerning the approximation by an effective boundary condition is stated in the following theorem.

Theorem 3. *As in the above theorem, the following bound holds*

$$\left\|(u^\delta - u^{1,\delta})|_{y_0<y}\right\| \le c\delta^{3/2}.$$

Proof. The main step is to obtain an asymptotic expansion for the problem related to the effective boundary condition. One can readily verify that the first two terms of the expansion are exactly u^0 and u^1. The bound is then obtained through a stability estimate for the approximate problem. □

References

[1] T. Abboud and H. Ammari, *Diffraction at a Curved Grating: TM and TE cases, Homogenization*, J. of Math. Analysis and Application, (1996), pp. 993–1026.

[2] Y. Achdou, *Effet d'un revetement métallisé mince sur la réflexion d'une onde électromagnétique*, C. R. Acad. Sciences, Série I **314** (1992) pp. 217-222. (1996), pp. 993–1026.

[3] A. Bendali, P. Borderies and J.-R. Poirier, *Effective boundary condition for the scattering of a TM time-harmonic wave by a rapidly oscillating surface*, in preparation.

[4] A. Bendali, P. Borderies and J.-R. Poirier, *Scattering of a time-harmonic wave by a rapidly oscillating surface: the three-dimensional case*, in preparation.

[5] A. F. Peterson and S. L. Ray and R. Mittra , *Computational Methods for Electromagnetics*, IEEE/OUP Series on Electromagnetic Wave Theory, IEEE Press and Oxford University Press, New-York, 1998.

[6] J. Sanchez-Hubert and E. Sanchez-Palencia , *Introduction aux méthodes asymptotiques et à l'homogénéisation* , Masson, Paris, 1993.

[7] A. B. Vasil'eva, V. F. Butuzov and L. V. Kalachev, *The Boundary Function Method for Singular Perturbation Problems*, SIAM Studies in Applied Mathematics, SIAM, Philadelphia, 1995.

Practical Aspects of FEM in Electromagnetics

Oszkár Bíró

The aim of this talk is to highlight two aspects of computational electromagnetism which concern practical low frequency applications. One of them is the question of taking account of the excitation through coils and the other is modelling magnetic nonlinearity.

In low frequency problems, the displacement current density can be neglected resulting in the so-called quasi-static approximation. If the electromagnetic field is generated by coils with known current density distribution, an *eddy current problem* is obtained. If, on the other hand, the current density of the coils is unknown, a *skin effect problem* is spoken of.

The boundary value problems are invariably formulated in terms of potentials. Scalar potentials are approximated on nodal elements and vector potentials on edge elements [6].

Magnetostatic problems in terms of a scalar potential. Magnetostatic fields are generated in nonconducting domains by coils with known current density. The magnetic field intensity can be described as the sum of the gradient of a scalar potential and of a vector function whose curl is the given current density (impressed vector potential). The scalar potential satisfies a Poisson equation with Dirichlet and Neumann boundary conditions. A great advantage of this approach is that the coils need not be modelled by the finite element mesh. It is shown, however, that if the smooth function representing the impressed vector potential is inserted into the finite element equations, wrong results are obtained. This is due to the fact that the impressed vector potential and the gradient of the scalar potential are in different function spaces. The remedy is representing the impressed vector potential in terms edge basis functions [2].

Magnetostatic problems in terms of a vector potential. An alternative to using the scalar potential is to describe the magnetic flux density as the curl of a magnetic vector potential. The current density of the coils appears then directly on the right hand side of the edge element equation system which is singular. Due to numerical integration errors the right hand side is not consistent and hence the equations cannot be solved by Krylov type iterative methods. Again, the remedy is to represent the current density by means of an impressed current vector potential and thus making the right hand side consistent. [4]

Using a reduced vector potential. A disadvantage of the vector potential approach is that the geometry of the coils has to be modelled by the finite element mesh. This can be avoided by writing the flux density as the sum of the curl of a known vector potential due to the coils in free space and of a reduced vector potential. [5].

Eddy current problems in terms of a current vector and a magnetic scalar potential or of a magnetic vector and an electric scalar potential. In case of eddy current problems, the excitation is represented by coils with given current density. Consequently, if their current is known, their treatment is similar to the approach followed in magnetostatic problems. In particular, if the eddy current field is represented by a current vector and a magnetic scalar potential, the coils are taken into account by means of an impressed current vector potential described by edge elements [6]. Similarly, if a magnetic vector and an electric scalar potential are used, the alternatives to represent the excitation are by means of an impressed current vector potential or a reduced vector potential [6]. If, on the other hand, the voltage of the coils is given, the current can be treated as an additional unknown and a circuit equation added to the system [8, 11].

Skin effect problems in terms of a current vector and a magnetic scalar potential. In case of skin effect problems, the excitation is either the current or the voltage of conductors acting as coils with their current density distribution unknown. Current excitation can be incorporated into the finite element formulation by means of prescribing appropriate boundary conditions if a current vector potential and a magnetic scalar potential act as system variables. Conversely, it is the voltage driven case that can be treated through boundary conditions within the frame of the formulation using a magnetic vector and an electric scalar potential [7]. The voltage excitation can be taken into account in the formulation in terms of the current vector potential and the magnetic scalar potential by means of treating the current as unknown and writing circuit equations [8–10].

Treatment of nonlinearity. Due to the nonlinear relationship between the magnetic flux density and field intensity, the finite element method leads to nonlinear algebraic equations in case of magnetostatic problems and to nonlinear ordinary differential equations for time dependent problems. Their solution can be carried out by means of standard techniques [3]. Frequently, it is more advantageous to write the eddy current equations in the frequency domain instead of the time domain. This leads to the harmonic balance method, see e.g. [1].

References

[1] F. Bachinger, M. Kaltenbacher, S. Reitzinger, *An efficient solution strategy for the HBFE method*, The 10th International IGTE Symposium on Numerical Field Calculation in Electrical Engineering, September 16-18, 2002, Graz, Austria, CD, ISBN 3-901351-65-5, pp. 385–389.

[2] O. Bíró, K. Preis, G. Vrisk, K.R. Richter and I. Tičar, *Computation of 3D magnetostatic fields using a reduced scalar potential*, IEEE Trans. Magnetics, 29 (1993), pp. 1329–1332.

[3] O. Bíró and K. Preis, *Finite element calculation of time-periodic 3d eddy currents in nonlinear media*, in vol. 9 of Studies in Applied Electromagnetics and Mechanics, Advanced Computational Electromagnetics, ed: T. Honma, Elsevier, IOS Press, 1995, pp. 62–74

[4] O. Bíró, K. Preis and K.R. Richter, *On the use of the magnetic vector potential in the nodal and edge finite element analysis of 3D magnetostatic fields*, IEEE Trans. Magnetics, 32 (1996), pp. 651–654.

[5] O. Bíró, C. Paul, K. Preis and S. Russenschuck, *A_r formulation using edge elements for the calculation of 3D fields in superconducting magnets*, IEEE Trans. Magnetics, 35 (1999), pp. 1391–1393.

[6] O. Bíró, *Edge element formulations of eddy current problems*, Comput. Methods Appl. Mech. Engrg., 169 (1999), pp. 391–405.

[7] O. Bíró, P. Böhm, K. Preis and G. Wachutka, *Edge finite element analysis of transient skin effect problems*, IEEE Trans. Magnetics, 36 (2000), pp. 835–839.

[8] O. Bíró, K. Preis, G. Buchgraber and I. Tičar, *Voltage-driven coils in finite-element formulations using a current vector and a magnetic scalar potential* , to be published in IEEE Trans. Magnetics, 40 (2004).

[9] A. Bossavit, *Most general "non-local" boundary conditions for the Maxwell equations in a bounded region*, COMPEL, 19 (2000), pp. 239–245.
[10] P. Dular, C. Geuzaine and W. Legros, *A natural method for coupling magnetodynamic H-formulations and circuit equations* , IEEE Trans. Magnetics, 35 (1999), pp. 1626–1629.
[11] P. J. Leonard and D. Rodger, *Modeling voltage forced coils using the reduced scalar potential method*, IEEE Trans. Magnetics, 28 (1992), pp. 1615–1618.

$\mathcal{H}^2$-Matrices with Adaptive Cluster Bases Applied to an Eddy Current Problem

Steffen Börm

$\mathcal{H}^2$-matrices [1, 5, 6, 16] can be used to find data-sparse representations of the densely populated matrices occurring, e.g., in boundary element methods.

The basic idea of hierarchical matrix techniques [3, 4, 10, 13–15] is to split the index set I into a hierarchy of subsets, the *cluster tree* $\mathcal{T}_I$, and to split the matrix into a hierarchy $\mathcal{T}_{I\times I}$ of subblocks $\tau \times \sigma$ corresponding to $\tau, \sigma \in \mathcal{T}_I$ that contains only small blocks and blocks that admit a separable approximation. The latter blocks are called *admissible*.

In a hierarchical matrix, an admissible block $\tau \times \sigma$ is approximated by a factorized rank-k-matrix $AB^\top$ ($A \in \mathbb{R}^{\tau\times k}$, $B \in \mathbb{R}^{\sigma\times k}$). The factorized form can be constructed by standard panel-clustering techniques [17], multipole expansion [11, 19] or interpolation [2].

In an $\mathcal{H}^2$-matrix, an admissible block $\tau \times \sigma$ is approximated by a special low-rank matrix of the form $V^\tau S^{\tau,\sigma} W^{\sigma\top}$ ($V^\tau \in \mathbb{R}^{\tau\times k}$, $W^\sigma \in \mathbb{R}^{\sigma\times k}$, $S^{\tau,\sigma} \in \mathbb{R}^{k\times k}$). By requiring the *row cluster bases* V^τ and the *column cluster bases* W^σ to be organized in a nested hierarchy (this is straightforward for polynomial approximation schemes [6, 9] and can also be achieved for multipole expansions [12]), we can reach algorithms with *linear* complexity in the number of degrees of freedom n.

While constructing an $\mathcal{H}^2$-matrix approximation of an integral operator by Lagrangian interpolation leads to a relatively general, simple and fast method, this approach also requires a large amount of storage, since polynomial bases are not adapted to the special characteristics of a given operator or a given geometry. This problem can be solved by combining the separable approximation scheme with an algebraic recompression algorithm that detects and eliminates redundant expansion functions by solving local symmetric eigenvalue problems [5, 7], which reduces the storage requirements significantly at the price of a moderate increase in computing time.

Since the recompression algorithm can be used without keeping the entire original $\mathcal{H}^2$-matrix approximation in memory, it is possible to treat boundary element problems with more than 100.000 degrees of freedom on standard PCs in less than ten minutes.

The combination of polynomial interpolation and algebraic recompression can not only be applied to standard Laplace problems, but also to more complicated vector-valued eddy-current models for Maxwell's equation [8, 18]. An example is the vector-valued double layer potential

$$b(\mathbf{E},\phi) = \int_\Gamma \int_\Gamma \langle \mathbf{curl}_\Gamma\, \phi(y), \mathbf{E}(x)\rangle \langle \mathbf{grad}_x\, \Phi(x,y), \mathbf{n}(x)\rangle\, dy\, dx \\ - \int_\Gamma \int_\Gamma \langle \mathbf{curl}_\Gamma\, \phi(y), \mathbf{n}(x)\rangle \langle \mathbf{grad}_x\, \Phi(x,y), \mathbf{E}(x)\rangle\, dy\, dx$$

for the fundamental solution $\Phi(x,y) := 1/(4\pi\|x-y\|)$. Even if Φ could be approximated by a single tensor product, the resulting matrix approximation would still have rank 3, since the variables x and y are coupled by a three-dimensional inner product. In practical approximation schemes, this implies that the rank required for the approximation of the vector-valued operator will be at least three times as high as in the case of scalar-valued operators.

Still, numerical experiments performed by applying the recompression algorithm to the vector-valued operator leads to storage requirements that are close to those of the scalar-valued operator. This result suggests that recompression is crucial for the efficient treatment of vector-valued problems, since the conventional fast approximation schemes like polynomial and multipole expansions seem to be incapable of taking advantage of their special structure.

References

[1] S. Börm, *$\mathcal{H}^2$-matrices — multilevel methods for the approximation of integral operators*, Tech. Rep. 7, Max Planck Institute for Mathematics in the Sciences, 2003. To appear in Computing and Visualization in the Sciences.

[2] S. Börm and L. Grasedyck, *Low-rank approximation of integral operators by interpolation*, Tech. Rep. 72, Max Planck Institute for Mathematics in the Sciences, 2002. To appear in Computing.

[3] S. Börm, L. Grasedyck, and W. Hackbusch, *Hierarchical Matrices*. Lecture Notes 21 of the Max Planck Institute for Mathematics in the Sciences, 2003.

[4] ———, *Introduction to hierarchical matrices with applications*, Engineering Analysis with Boundary Elements, 27 (2003), pp. 405–422.

[5] S. Börm and W. Hackbusch, *Data-sparse approximation by adaptive $\mathcal{H}^2$-matrices*, Computing, 69 (2002), pp. 1–35.

[6] ———, *$\mathcal{H}^2$-matrix approximation of integral operators by interpolation*, Applied Numerical Mathematics, 43 (2002), pp. 129–143.

[7] ———, *Approximation of boundary element operators by adaptive $\mathcal{H}^2$-matrices*, Tech. Rep. 5, Max Planck Institute for Mathematics in the Sciences, 2003. To appear in Foundations of Computational Mathematics.

[8] S. Börm and J. Ostrowski, *Fast evaluation of boundary integral operators arising from an eddy current problem*, Journal of Computational Physics, 193 (2003), pp. 67–85.

[9] K. Giebermann, *Multilevel approximation of boundary integral operators*, Computing, 67 (2001), pp. 183–207.

[10] L. Grasedyck and W. Hackbusch, *Construction and arithmetics of $\mathcal{H}$-matrices*, Computing, 70 (2003), pp. 295–334.

[11] L. Greengard and V. Rokhlin, *A fast algorithm for particle simulations*, Journal of Computational Physics, 73 (1987), pp. 325–348.

[12] L. Greengard and V. Rokhlin, *A new version of the fast multipole method for the Laplace in three dimensions*, in Acta Numerica 1997, Cambridge University Press, 1997, pp. 229–269.

[13] W. Hackbusch, *A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices*, Computing, 62 (1999), pp. 89–108.
[14] W. Hackbusch and B. Khoromskij, *A sparse $\mathcal{H}$-matrix arithmetic: General complexity estimates*, J. Comp. Appl. Math., 125 (2000), pp. 479–501.
[15] ———, *A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part II: Application to multi-dimensional problems*, Computing, 64 (2000), pp. 21–47.
[16] W. Hackbusch, B. Khoromskij, and S. Sauter, *On $\mathcal{H}^2$-matrices*, in Lectures on Applied Mathematics, H. Bungartz, R. Hoppe, and C. Zenger, eds., Springer-Verlag, Berlin, 2000, pp. 9–29.
[17] W. Hackbusch and Z. P. Nowak, *On the fast matrix multiplication in the boundary element method by panel clustering*, Numerische Mathematik, 54 (1989), pp. 463–491.
[18] R. Hiptmair, *Symmetric coupling for eddy current problems*, SIAM J. Numer. Anal., 40 (2002), pp. 41–65.
[19] V. Rokhlin, *Rapid solution of integral equations of classical potential theory*, Journal of Computational Physics, 60 (1985), pp. 187–207.

Theoretical Aspects of Edge Finite Elements

Daniele Boffi

Let us consider the time harmonic Maxwell system

$$\text{(TH)}\qquad \begin{cases} \operatorname{curl}(\mu^{-1}\operatorname{curl}\mathbf{u}) - \omega^2\varepsilon\mathbf{u} = \mathbf{f} & \text{in } \Omega \\ \operatorname{div}(\varepsilon\mathbf{u}) = 0 & \text{in } \Omega \\ \mathbf{u}\times\mathbf{n} = 0 & \text{on } \Omega \end{cases}$$

where ω is the fixed frequency, $\operatorname{div}\mathbf{f} = 0$, and Ω is a polyhedral (or polygonal) domain with outward normal $\mathbf{n}$.

It is well known that problem (TH) is well posed if and only if ω^2 is not an interior Maxwell eigenvalue. A variational formulation of the problem under consideration is obtained, for instance, by imposing the divergence free condition in a weak sense in the spirit of Kikuchi [21] as follows.

Find $(\mathbf{u}, p) \in H_0(\mathrm{curl};\Omega)\times H^1_0(\Omega) = V\times Q$ such that

$$\text{(TH-V)}\qquad \begin{cases} (\mu^{-1}\operatorname{curl}\mathbf{u}, \operatorname{curl}\mathbf{v}) - \omega^2(\varepsilon\mathbf{u},\mathbf{v}) + (\varepsilon\mathbf{v}, \operatorname{grad} p) = (\mathbf{f},\mathbf{v}) & \forall\mathbf{v}\in V \\ (\varepsilon\mathbf{u}, \operatorname{grad} q) = 0 & \forall q\in Q \end{cases}$$

A stability estimate of the solution of (TH-V) can be found, for instance, in [17].

$$\left(\|\mathbf{u}\|^2_{\mathrm{curl}} + \|p\|^2_1\right)^{1/2} \le \sup_{i=1,2,\dots}\left(1+\omega^2, \frac{1+\lambda_i}{|\lambda_i-\omega^2|}\right)\|\mathbf{f}\|_0$$

where λ_i $(i = 1, 2, \dots)$ are the interior Maxwell eigenvalues. Given $V_h \subset V$ and $Q_h \subset Q$ we consider the discretization of problem (TH-V).
(TH-Vh)
Find $(\mathbf{u}_h, p_h) \in V_h \times Q_h$ such that

$$\begin{cases} (\mu^{-1}\operatorname{curl}\mathbf{u}_h, \operatorname{curl}\mathbf{v}) - \omega^2(\varepsilon\mathbf{u}_h,\mathbf{v}) + (\varepsilon\mathbf{v}, \operatorname{grad} p_h) = (\mathbf{f},\mathbf{v}) & \forall\mathbf{v}\in V_h \\ (\varepsilon\mathbf{u}_h, \operatorname{grad} q) = 0 & \forall q\in Q_h \end{cases}$$

Assuming the compatibility condition

$$\text{(COMP)} \qquad \operatorname{grad} Q_h \subset V_h$$

which guarantees a discrete inf-sup condition for problem (TH-Vh), we have the error estimate

$$\|\mathbf{u}-\mathbf{u}_h\|_{\text{curl}}^2 + \|p-p_h\|_1^2 \le \gamma^2 \inf_{(\mathbf{v}_h,q_h)\in V_h\times Q_h} \left(\|\mathbf{u}-\mathbf{v}_h\|_{\text{curl}}^2 + \|p-q_h\|_1^2\right)$$

with

$$\gamma \le 1 + \max_{i=1,2,\dots} \left(1+\omega^2, \frac{1+\lambda_{i,h}}{|\lambda_{i,h}-\omega^2|}\right)$$

where $\lambda_{i,h}$ are the discrete Maxwell eigenvalues. We explicitly notice that $\operatorname{div}\mathbf{f}=0$ implies $p=p_h=0$.

Several numerical experiments and theoretical results (see [8, 10], for instance) show that standard nodal elements do not approximate Maxwell eigenvalues in a correct way, even on special two dimensional meshes where the compatibility condition (COMP) is satisfied [8, 28]. On the other hand, edge finite elements have been proven to satisfy the discrete compactness property which guarantees the good approximation of the eigensolutions [6, 7, 14, 22–24] (see [20, 26] for a review on this topic).

In this talk we review some of the most important theoretical properties of edge finite elements, including discrete compactness, commuting diagram (de Rham complex), interpolation estimates. The commuting diagram property (see, for instance, [7, 12, 13, 18, 19], [16, 29] for possible extensions and [3] for a review) on a simply connected domain reads

$$\begin{array}{ccccccccc}
0 \to & Q & \xrightarrow{\text{grad}} & V & \xrightarrow{\text{curl}} & U & \xrightarrow{\text{div}} & S/\mathbb{R} & \to 0 \\
 & \downarrow \Pi_h^Q & & \downarrow \Pi_h^V & & \downarrow \Pi_h^U & & \downarrow \Pi_h^S & \\
0 \to & Q_h & \xrightarrow{\text{grad}} & V_h & \xrightarrow{\text{curl}} & U_h & \xrightarrow{\text{div}} & S_h/\mathbb{R} & \to 0,
\end{array}$$

where $Q \subset H_0^1(\Omega)$, $V \subset H_0(\text{curl})$, $U \subset H_0(\text{div})$, and $S \subset L^2(\Omega)$ are suitable smooth function spaces, so that the corresponding interpolation operators can be defined and Q_h, V_h, U_h, and S_h are their discrete counterparts.

Standard interpolation estimates are (see, for instance, [1, 2, 15, 20, 25, 27])

$$\begin{aligned}
&\inf_{\mathbf{v}_h\in V_h} \|\mathbf{u}-\mathbf{v}_h\|_0 \le Ch^s(|\mathbf{u}|_s + \|\operatorname{curl}\mathbf{u}\|_s) \qquad 1/2 < s \le k+1\\
&\inf_{\mathbf{v}_h\in V_h} \|\operatorname{curl}\mathbf{u} - \operatorname{curl}\mathbf{v}_h\|_0 \le Ch^s|\operatorname{curl}\mathbf{u}|_s \qquad 0 < s \le k+1
\end{aligned}$$

When $\operatorname{curl}\mathbf{u}$ is discrete, the improved estimate

$$\inf_{\mathbf{v}_h\in V_h} \|\mathbf{u}-\mathbf{v}_h\|_0 \le Ch^s|\mathbf{u}|_s \quad 1/2 < s \le k+1$$

has been used in [6], see also [20]. Recent results [9] show the improved estimate

$$\begin{aligned}
&\|\mathbf{u} - \Pi_h^V \mathbf{u}\|_{L^2} \le Ch^s \left(|\mathbf{u}|_{H^s} + \| \operatorname{curl} \mathbf{u}\|_{L^p} \right) && 1/2 < s \le 1,\ p > 2 \\
&\|\mathbf{u} - \Pi_h^V \mathbf{u}\|_{L^2} \le Ch^s |\mathbf{u}|_{H^s} && 1 < s \le k+1 \\
&\| \operatorname{curl} \mathbf{u} - \operatorname{curl} \Pi_h^V \mathbf{u}\|_{L^2} \le Ch^s |\operatorname{curl} \mathbf{u}|_{H^s} && 0 < s \le k+1
\end{aligned}$$

These estimates, which do not require on $\operatorname{curl} \mathbf{u}$ more regularity than the one needed for the definition of the interpolant itself (see [2]), have been used in [9] for the analysis of the approximation of photonic crystals.

The last remark concerns the approximation properties achieved by edge finite elements on quadrilateral meshes. Recent results show that particular care has to be taken into account when dealing with general regular quadrilateral finite elements [4]. This issue is particularly significant for quadrilateral edge elements; the lowest order element does not achieve the convergence at all in the $H(\text{curl})$ norm, the higher order elements are substantially suboptimal [5]. Some modifications of standard edge element, which provide a solution to this phenomenon, have been recently proposed [5, 11].

References

[1] A. Alonso and A. Valli, *A domain decomposition approach for heterogeneous time-harmonic Maxwell equations*, Comput. Methods Appl. Mech. Engrg., 143 (1997), no. 1–2, pp. 97–112.

[2] C. Amrouche, C. Bernardi, M. Dauge, and V. Girault, *Vector potentials in three-dimensional non-smooth domains*, Math. Methods Appl. Sci., 21 (1998), no. 9, pp. 823–864.

[3] D.N. Arnold, *Differential complexes and numerical stability*, Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002), Higher Ed. Press, Beijing, (2002), pp. 137–157.

[4] D.N. Arnold, D. Boffi, and R.S. Falk, *Approximation by quadrilateral finite elements*, Math. Comp., 71 (2002), no. 239, pp. 909–922.

[5] D.N. Arnold, D. Boffi, and R.S. Falk, *Quadrilateral H(div) finite elements*, submitted.

[6] D. Boffi, *Fortin operator and discrete compactness for edge elements*, Numer. Math., 87 (2000), No. 2, pp. 229–246.

[7] D. Boffi, *A note on the de Rham complex and a discrete compactness property*, Appl. Math. Letters, 14 (2001), pp. 33–38.

[8] D. Boffi, F. Brezzi, and L. Gastaldi, *On the problem of spurious eigenvalues in the approximation of linear elliptic problems in mixed form*, Math. Comp., 69 (2000), no. 229, pp. 121–140.

[9] D. Boffi, M. Conforti, and L. Gastaldi, *Modified edge finite elements for photonic crystals*, submitted

[10] D. Boffi, P. Fernandes, L. Gastaldi, and I. Perugia, *Computational models of electromagnetic resonators: analysis of edge element approximation*, SIAM J. Numer. Anal., 36 (1999), pp. 1264–1290.

[11] D. Boffi, F. Kikuchi, and J. Schöberl, in preparation.

[12] A. Bossavit, *Mixed finite elements and the complex of Whitney forms*, in The mathematics of finite elements and applications, VI (Uxbridge, 1987), Academic Press, London, (1988), pp. 137–144.

[13] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer Series in Computational Mathematics, 15. Springer-Verlag, New York, 1991.

[14] S. Caorsi, P. Fernandes, and M. Raffetto, *On the convergence of Galerkin finite element approximations of electromagnetic eigenproblems*, SIAM J. Numer, Anal., 38 (2000), no. 2, pp. 580–607.

[15] P. Ciarlet Jr. and J. Zou, *Fully discrete finite element approaches for time-dependent Maxwell's equations*, Numer. Math., 82 (1999), no. 2, pp. 193–219.

[16] L. Demkowicz, P. Monk, L. Vardapetyan, and W. Rachowicz, *de Rham diagram for hp finite element spaces*, Comput. Math. Appl., 39 (2000), no. 7–8, pp. 29–38.

[17] L. Demkowicz and L. Vardapetyan, *Modeling of electromagnetic absorption/scattering problems using hp-adaptive finite elements*, Comput. Methods Appl. Mech. Engrg., 152 (1998), no. 1–2, pp. 103–124.

[18] J. Douglas Jr. and J.E. Roberts, *Mixed finite element methods for second order elliptic problems*, Mat. Apl. Comput., 1 (1982), no. 1, pp. 91–103.

[19] R. Hiptmair, *Canonical construction of finite elements*, Math. Comp., 68 (1999), no. 228, pp. 1325–1346.

[20] R. Hiptmair, *Finite elements in computational electromagnetism*, Acta Numer., 11 (2002), pp. 237–339.

[21] F. Kikuchi, *Mixed and penalty formulations for finite element analysis of an eigenvalue problem in electromagnetism*, Comput. Methods Appl. Mech. Engrg., 64 (1987), no. 1–3, pp. 509–521.

[22] F. Kikuchi, *On a discrete compactness property for the Nédélec finite elements*, J. Fac. Sci. Univ. Tokyo Sect. IA Math., 36 (1989), no. 3, pp. 479–490.

[23] F. Kikuchi, *Theoretical analysis of Nedelec's edge elements*, Recent topics in mathematics moving toward science and engineering. Japan J. Indust. Appl. Math., 18 (2001), no. 2, pp. 321–333.

[24] P. Monk and L. Demkowicz, *Discrete compactness and the approximation of Maxwell's equations in* $\mathbb{R}^3$, Math. Comp., 70 (2001), no. 234, pp. 507–523.

[25] P. Monk, *A finite element method for approximating the time-harmonic Maxwell equations*, Numer. Math., 63 (1992), no. 2, pp. 243–261.

[26] P. Monk, *Finite Element Methods for Maxwell's Equations*, Clarendon Press, Oxford, 2003

[27] J.C. Nédélec, *Mixed finite elements in* $\mathbb{R}^3$, Numer. Math., 35 (1980), no. 3, pp. 315–341.

[28] M.J.D. Powell, *Piecewise quadratic surface fitting for contour plotting*, in Software for numerical mathematics (Proc. Conf., Inst. Math. Appl., Loughborough Univ. Tech., Loughborough, 1973), Academic Press, London (1974), pp. 253–271.

[29] J. Schöberl, *Commuting quasi-interpolation operators for mixed finite elements*, Preprint ISC-01-10-MATH, Texas A& M University, 2001.

An Adaptive Perfectly Matched Layer Technique for Time-harmonic Scattering Problems

Zhiming Chen
(joint work with Xueze Liu)

We propose and study an adaptive perfectly matched layer (PML) technique for solving the Helmholtz-type scattering problems with perfectly conducting boundary:

$$\Delta u + k^2 u = 0 \quad \text{in } \mathbf{R}^2 \backslash \bar{D}, \tag{1a}$$

$$\frac{\partial u}{\partial \mathbf{n}} = -g \quad \text{on } \Gamma_D, \tag{1b}$$

$$\sqrt{r}\left(\frac{\partial u}{\partial r} - \mathbf{i}ku\right) \to 0 \quad \text{as } r = |x| \to \infty. \tag{1c}$$

Here $D \subset \mathbf{R}^2$ is a bounded domain with Lipschitz boundary Γ_D, $g \in H^{-1/2}(\Gamma_D)$ is determined by the incoming wave, and $\mathbf{n}$ is the unit outer normal to Γ_D. We assume the wave number $k \in \mathbb{R}$ is a constant. We remark that the results in this paper can be easily extended to solve the scattering problems with other boundary conditions such as Dirichlet or the impedance condition on Γ_D, or the acoustic wave propagation problems in inhomogeneous media which correspond to a variable wave number $k^2(x)$.

Since the work of Berenger [3] which proposed a PML layer for use with the time dependent Maxwell equations, various constructions of PML absorbing layers have been proposed and studied in the literature (cf. e.g. Turkel and Yefet [17] for a review). Under the assumption that the exterior solution is composed of outgoing waves only, the basic idea of the PML technique is to surround the computational domain by a finite thickness layer of the specially designed model medium that would either slow down or attenuate all the waves that propagate from inside the computational domain. The PML equation for the time-harmonic scattering problem (1a) is derived in Collino and Monk [8] by a complex extension of the solution u in the exterior domain. It is proved in Lassas and Somersalo [11], Hohage, Schmidt and Zschiedrich [10] that the resultant PML solution converges exponentially to the solution of the original scattering problem as the PML layer thickness tends to infinite. We remark that in practical applications involving PML method, one cannot afford to use a very thick PML layer because it requires excessive grid points and hence more computer time and more storage. On the other hand, a thin PML layer requires a rapid variation of the artificial material property which deteriorates the accuracy if two corse mesh is used in the PML layer.

A posteriori error estimates are computable quantities in terms of the discrete solution and data that measure the actual discrete errors without the knowledge of exact solutions. They are essential in designing algorithms for mesh modification which equi-distribute the computational effort and optimize the computation. Ever since the pioneering work of Babuška and Rheinboldt [2], the adaptive finite element methods based on a posteriori error estimates have become a central theme in scientific and engineering computations. The ability of error control and the asymptotically optimal approximation property (see e.g. Morin, Nochetto and Siebert [14], Chen and Dai [5]) make the adaptive finite element method attractive for complicated physical and industrial processes (cf. e.g. Chen and Dai [4], Chen, Nochetto and Schmidt [6]). For the efforts to solve scattering problems using adaptive methods based on a posterior error estimate, we refer to the recent work Monk [12], Monk and Süli [13].

It is proposed in Chen and Wu [7] for scattering problems by periodic structures, the grating problem, that one can use the a posteriori error estimate to determine the PML parameters. Moreover, the derived a posteriori error estimate in [7] has the nice feature of exponential decay in terms of the distance to the distance to the boundary of the fixed domain where the PML layer is placed. This property leads to coarse mesh size away from the fixed domain and thus makes the total computational cost insensitive to the thickness of the PML absorbing layer.

In this paper we extend the idea of using a posteriori error estimates to determine the PML parameters and propose an adaptive PML technique for solving the scattering problem (1a)-(1c). The first difficulty of the analysis is that in contrast to the grating problems in which there are only finite number of outgoing modes [7], now there are infinite number of outgoing modes expressed in terms of Hankel functions. We overcome this difficulty by using following uniform estimate for the Hankel functions H^1_ν, $\nu \in \mathbb{C}$,:

$$|H_\nu^{(1)}(z)| \le e^{-\mathrm{Im}\,(z)\left(1-\frac{\Theta^2}{|z|^2}\right)^{1/2}} |H_\nu^{(1)}(\Theta)|, \tag{2}$$

for any $z \in \mathbb{C}_{++}, \Theta \in \mathbf{R}$ such that $0 < \Theta \le |z|$, where $\mathbb{C}_{++} = \{z \in \mathbb{C} : \mathrm{Im}\,(z) \ge 0, \mathrm{Re}\,(z) \ge 0\}$. This sharp estimate, which seems first appeared in this paper, allows us to prove the exponentially decaying property of the PML solution without resorting to the integral equation technique in [11] or the representation formula in [10]. We remark that in [11], [10], it is required the fictitious absorbing coefficient must be linear after certain distance away from the bounary where the PML layer is placed.

The second difficulty is that the PML equation in the PML layer is not necessarily uniquely solvable for any wave number k^2. Let $\Omega^{\mathrm{PML}} = B_\rho \backslash \bar{B}_R$, where $0 < R < \rho$ and B_a denotes the circle of radius a for any $a > 0$. Let $\alpha = 1 + \mathbf{i}\sigma$ be the fictitious medium property. In practical applications, σ is usually taken as power functions:

$$\sigma = \sigma(r) = \sigma_0 \left(\frac{r-R}{\rho-R}\right)^m \quad \text{for some integer } m \ge 1, \tag{3}$$

where $\sigma_0 > 0$ is some constant. We prove that for any given R and ρ, the PML equation in the PML layer is uniquely solvable and its solution satisfies sharp stability estimates if σ_0 is chosen sufficiently large. This allows us to complete the proof of the following key estimate between the Dirichlet-to-Neumann mapping for the original scattering problem $T : H^{1/2}(\Gamma_R) \to H^{-1/2}(\Gamma_R)$ and the PML problem $\hat{T}$, where $\Gamma_R = \partial B_R$,

$$\| T - \hat{T} \|_{L(H^{1/2}(\Gamma_R), H^{-1/2}(\Gamma_R))} \leq C(1 + k^2 R^2)|\alpha_0|^2 e^{-k\mathrm{Im}\,(\tilde{\rho})\left(1-\frac{R^2}{|\tilde{\rho}|^2}\right)^{1/2}},$$

where $\alpha_0 = 1 + \mathbf{i}\sigma_0$, and $\tilde{\rho} = \int_0^\rho \alpha(t)dt$ is the complex radius corresponding to ρ.

References

[1] I. Babuška and A. Aziz. *Survey Lectures on Mathematical Foundations of the Finite Element Method.* in The Mathematical Foundations of the Finite Element Method with Application to Partial Differential Equations, ed. by A. Aziz, Academic Press, New York, 1973, 5-359.

[2] I. Babuška and C. Rheinboldt. *Error estimates for adaptive finite element computations.* SIAM J. Numer. Anal. 15 (1978), pp736-754.

[3] J.-P. Berenger. *A perfectly matched layer for the absorption of electromagnetic waves.* J. Comput. Physics 114 (1994), pp185-200.

[4] Z. Chen and S. Dai. *Adaptive Galerkin methods with error control for a dynamical Ginzburg-Landau model in superconductivity.* SIAM J. Numer. Anal. 38 (2001), pp1961-1985.

[5] Z. Chen and S. Dai. *On the efficiency of adaptive finite element methods for elliptic problems with discontinuous coefficients. SIAM J. Sci. Comput.* 24 (2002), pp443-462.

[6] Z. Chen, R.H. Nochetto and A. Schmidt. *A characteristic Galerkin method with adaptive error control for continuous casting problem. Comput. Methods Appl. Mech. Engrg.* 189 (2000), pp249-276.

[7] Z. Chen and H. Wu. *An adaptive finite element method with perfectly matched absorbing layers for the wave scattering by periodic structures.* SIAM J. Numer. Anal. 41, (2003), pp799-826.

[8] F. Collino and P.B. Monk. *The perfectly matched layer in curvilinear coordinates.* SIAM J. SCi. Comput. 19 (1998), pp2061-2090.

[9] D. Colton and R. Kress. *Integral Equation Methods in Scattering Theory.* John Wiley & Sons, New York, 1983.

[10] T. Hohage, F. Schmidt and L. Zschiedrich. *Solving time-harmonic scattering problems based on the pole condition. II: Convergence of the PML method. SIAM J. Math. Anal.*, to appear.

[11] M. Lassas and E. Somersalo. *On the existence and convergence of the solution of PML equations.* Computing 60 (1998), pp229-241.

[12] P. Monk. *A posteriori error indicators for Maxwell's equations.* J. Comput. Appl. Math. 100 (1998), 173-190.

[13] P. Monk and E. Süli. *The adaptive computation of far-field patterns by a posteriori error estimation of linear functionals.* SIAM J. Numer. Anal. 36 (1998), pp251-274.

[14] P. Morin, R.H. Nochetto and K.G. Siebert. *Data oscillation and convergence of adaptive FEM,* SIAM J. Numer. Anal. 38 (2000), pp466-488.

[15] A.H. Schatz. *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms.* Math. Comp. (1974), pp959-862.

[16] L.R. Scott and S. Zhang. *Finite element interpolation of nonsmooth functions satisfying boundary conditions.* Math. Comp. (1990), pp483-493.

[17] E. Turkel and A. Yefet. *Absorbing PML boundary layers for wave-like equations.* Appl. Numer. Math. 27 (1998), pp533-557.

[18] G.N. Watson. *A Treatise on The Theory of Bessel Functions,* Cambridge, 1922.

Div-Curl Lemma for Edge Elements
Snorre H. Christiansen

Given two sequences (u_h) and (u'_h) of vector fields converging weakly in L^2 on some open domain in $\mathbb{R}^3$ the div-curl lemma of Murat [5] and Tartar [7] gives sufficient conditions under which their scalar product converges in the weak-star sense of distributions to the right scalar field. Namely if the sequences $(\operatorname{div} u_h)$ and $(\operatorname{curl} u'_h)$ are relatively compact in H^{-1} then this convergence property holds. This lemma is useful in questions arizing in homogenization and certain non-linear PDEs and is an ingredient in the method of compensated compactness.

For the variational formulation of problems in electromagnetics on Nédélec's [6] edge element spaces X_h one can naturally obtain control over the L^2 norm of the involved fields. One can also expect to have sufficient control of the curl in H^{-1} (e.g. in the form of boundedness in L^2), due to energy considerations. However control of the divergence of a field $u_h \in X_h$ is obtained in the form of estimates on $\int u_h \cdot \operatorname{grad} p_h$ when p_h runs trough the maximal space Y_h of continuous piecewise polynomials which vanish on the boundary and such that the gradient operator maps Y_h into X_h. Since the space Y_h is smaller than H^1_0, the question arizes whether an L^2 bounded sequence of so-called discrete divergence free vector fields $u_h \in X_h$ has compact divergence in H^{-1}. This property is stronger than the discrete compactness property of Kikuchi which has come to play a central role in the numerical analysis of edge elements.

While we leave this question unanswered we prove in this talk the following div-curl lemma for edge elements on quasi-uniform meshes on bounded domains with smooth boundary[1]:

Lemma 1. *Suppose (u_h) and (u'_h) are sequences of vector fields $u_h, u'_h \in X_h$ converging weakly in L^2 to u and u'. Suppose furthermore that with the decomposition $u_h = v_h + \operatorname{grad} p_h$ with v_h in the L^2 orthogonal of $\operatorname{grad} Y_h$ in X_h, and $p_h \in Y_h$, (p_h) is relatively compact in H^1_0, and that $(\operatorname{curl} u'_h)$ is relatively compact in H^{-1}.*

Then $(u_h \cdot u'_h)$ converges to $u \cdot u'$ in the weak-star sense of distributions.

One of the main ingredients of the proof is a norm equivalence on a subspace of X_h which is uniform with respect to h and which strengthens the standard discrete compactness property (using a technique appearing in Lemma 4.1 in [4]). Another ingredient is a super-approximation property of the spaces Y_h. For the details of the proof I refer to the revised version of the preprint [2], which also contains bibliographical references in particular to the work by Boffi and Hiptmair on discrete compactness.

This work is related to a joint effort [3] to understand the variational formulation of constraints in the discretization of some non-linear PDEs, parts of which were presented in [1].

[1] The possibility of weaking these hypothesis was briefly discussed and is the object of current efforts.

References

[1] S. H. Christiansen, *On discretizations of Yang-Mills equations*; "Geometric and Structure Preserving Algorithms for PDEs", Workshop at the Centre for Advanced Study, Oslo, May 12 – 15, 2003,
(abstract available from http://www.cma.uio.no/reports/talks/index2003.html).

[2] S. H. Christiansen, *Div-curl lemma for edge elements*; University of Oslo, Dept. of Math., Preprint Pure Mathematics, ISBN 82-553-1400-8, No. 30, August 2003 (revised September 2003).

[3] S. H. Christiansen, G. Ellingsrud, R. Kozlov, R. Winther, *Numerical aspects of constraint preservation in Yang-Mills equations*; work in progress.

[4] P. Ciarlet Jr., J. Zou, *Fully discrete finite element approaches for time-dependant Maxwell's equations*; Numer. Math., Vol. 82, p. 193-219, 1999.

[5] F. Murat, *Compacité par compensation*; Ann. Scuola Norm. Sup. Pisa (4), Vol. 5, No. 3, p. 485-507, 1978.

[6] J.-C. Nédélec, *Mixed finite elements in* $\mathbb{R}^3$; Numer. Math., Vol. 35, p. 315-341, 1980.

[7] L. Tartar, *Compensated compactness and applications to partial differential equations*; in "Nonlinear analysis and mechanics: Heriot-Watt Symposium", Vol. IV, p. 136-212, Res. Notes in Math., Vol. 39, Pitman, Boston, Mass.-London, 1979.

Formulations and Efficient Numerical Solution Techniques for Transient 3D Magneto-and Electro-Quasistatic Field Problems

Markus Clemens

(joint work with Galina Benderskaya, Herbert De Gersem, Stefan Feigh, Markus Wilke, Jing Yuan and Thomas Weiland)

The simulation of 3D quasistatic electric high-voltage fields and magnetic eddy currents field problems typically involves nonlinear material properties such as field dependent electric conductivities of insulator materials or saturation effects within ferromagnetic materials which may be even of hysteretic nature. In these cases and, more generally, for any non-periodical field excitation, time domain formulations of these problems are preferred. Using spatial discretization schemes such as the Whitney Finite Element method [3], the Cell Method [18]or the Finite Integration Technique [9, 19], for electro-quasistatic problems this will result in large systems of stiff ordinary differential equations of the form

$$\mathbf{G}^T\mathbf{M}_\varepsilon\mathbf{G}\frac{d}{dt}\mathbf{\Phi}(t) + \mathbf{G}^T\mathbf{M}_\kappa\big(\mathbf{\Phi}(t)\big)\mathbf{G}\mathbf{\Phi}(t) = 0. \tag{1}$$

where $\mathbf{G}^T$ and $\mathbf{G}$ are the discrete divergence and gradient matrices with the vector of electric grid voltages as $\widehat{\mathbf{e}} = -\mathbf{G}\mathbf{\Phi}$ and $\mathbf{M}_\varepsilon$ and $\mathbf{M}_\kappa = \mathbf{M}_\kappa(\mathbf{\Phi})$ are material matrices combining the permittivities and field dependent electrical conductivities with the metric information of the grid [7]. Magneto-quasistatic fields can be described with systems of differential-algebraic equations of index 1

$$\mathbf{M}_\kappa\frac{d}{dt}\widehat{\mathbf{a}}(t) + \mathbf{C}^T\mathbf{M}_\nu\big(\widehat{\mathbf{a}}(t)\big)\mathbf{C}\widehat{\mathbf{a}}(t) = \widehat{\widehat{\mathbf{j}}}_s(t), \tag{2}$$

where $\widehat{\mathbf{a}}$ is the vector of path integrated magnetic vector potentials, $\mathbf{C}$ is the incidence matrix discretizing the curl operator to yield the vector of magnetic fluxes

$\widehat{\widehat{\mathbf{b}}} = \mathbf{C}\widehat{\mathbf{a}}$, $\mathbf{M}_\nu = \mathbf{M}_\nu(\widehat{\mathbf{a}})$ is the matrix of flux dependent reluctivities and the $\widehat{\widehat{\mathbf{j}}}_s$ is the vector of current excitations [8, 12]. Today, efficient numerical techniques for the solution of the large systems of equations (1) and (2) involve time step adaptive higher order embedded time integration schemes such as singly diagonal implicit Runge-Kutta methods (SDIRK) or linear-implicit Rosenbrock-Wanner (ROW) methods [13, 15]. In these schemes the repeated solution of the algebraic systems of equations involves a combination of advanced numerical methods. Such methods are geometric or algebraic multigrid preconditioners specifically designed to interact with the above mentioned geometric discretization methods [4–6, 17], multiple-righthand side Lanczos-projection techniques and a subspace projection extrapolation scheme for the generation of optimal start vectors of the iterative solution methods [11, 14]. Specialized projection methods are used for the inclusion of floating potential areas and other complicated boundary conditions [16] and non-standard time step-prediction schemes are developed for magnetodynamic field-circuit coupled formulations involving switching circuit elements [2]. Extensions of the magneto-quasistatic formulations also include models for motion-induced eddy currents as they occur e.g. in eddy current brakes using either Lagrangian or Eulerian coordinate descriptions [1, 10] and nonlinear iteration schemes adapted to hysteretic ferromagnetic material behavior described by Preisach or Jiles-Atherton hysteresis models [20–22].

References

[1] M. Bartsch, M. Clemens, T. Weiland, and M. Wilke, *Simulation of linear eddy current brakes using FI^{2}TD methods*, in Electromagnetic Fields in Electrical Engineering, A. Krawczyk and S. Wiak, eds., vol. 22, IOS Press, series Studies in Applied Electromagnetics and Mechanics, 2002, pp. 357–362.

[2] G. Benderskaya, H. De Gersem, M. Clemens, and T. Weiland, *Interpolating Technique for Effective Determination of Switching Time Instants for Field-Circuit Coupled Problems with Switching Elements.* To appear in Proc. XVIII Symposium on Electromagnetic Phenomena in Nonlinear Circuits (EPNC 2004), Poznan, Poland, June 2004.

[3] A. Bossavit, L. Kettunen, and T. Tarhassaari, *Some realizations of a discrete Hodge operator: A reinterpretation of the finite element technique*, IEEE Trans. Magn., 35 (1999), pp. 1494–1497.

[4] M. Clemens, S. Feigh, and T. Weiland, *Geometric multigrid algorithms using the Conformal Finite Integration Technique.* Conf. Rec. Compumag 2003, Saratoga Springs, USA, Vol. IV. To appear in IEEE Trans. Magn., June 2004.

[5] ———, *Divergence removing multigrid smoothers for curl-curl equations of the discrete electromagnetism.* Proc. PIERS 2004 (Progress in Electromagnetics Research), Pisa, March 2004.

[6] M. Clemens, S. Feigh, M. Wilke, and T. Weiland, *Non-nested geometric multigrid method using consistency error correction for discrete magnetic curl-curl formulations.* Proc. of the EMF 2003, Aachen, Germany, 06.-09. Oct. 2003. To appear in COMPEL, 2004.

[7] M. Clemens, H. De Gersem, W. Koch, T. Weiland, and M. Wilke, *Transient simulation of nonlinear electro-quasistatic problems using the Finite Integration Technique*, in Proc. IGTE 2002 Symposium, Graz, Austria, 2002, pp. 510–517.

[8] M. Clemens and T. Weiland, *Transient eddy current calculation with the FI-method*, IEEE Trans. Magn., 35 (1999), pp. 1163–1166.

[9] ———, *Discrete electromagnetism with the Finite Integration Technique*, in Geometric Methods for Computational Electromagnetics, F. L. Teixeira, ed., no. 32 in PIER, EMW Publishing, Cambridge, Massachusetts, USA, 2001, pp. 65–87.
[10] M. Clemens, T. Weiland, and M. Wilke, *Transient eddy current formulation including moving conductors using the Finite Integration method*, IEEE Trans. Magn., 36 (2000), pp. 804–808.
[11] M. Clemens, M. Wilke, R. Schuhmann, and T. Weiland, *Subspace projection extrapolation scheme for transient field simulations.* Conf. Rec. Compumag 2003, Saratoga Springs, USA, Vol. I. To appear in IEEE Trans. Magn., June 2004.
[12] M. Clemens, M. Wilke, and T. Weiland, *Advanced FI^2TD algorithms for transient 3d eddy current problems*, Compel, 20 (2001), pp. 365–379.
[13] ———, *3D transient eddy current simulations using FI^2TD with variable time step size selection schemes*, IEEE Trans. Magn., 38 (2002), pp. 605–608.
[14] ———, *Extrapolation strategies in transient magnetic field simulations*, IEEE Trans. Magn., 39 (2003), pp. 1171–1174.
[15] ———, *Linear–implicit time integration schemes for error-controlled transient nonlinear magnetic field simulations*, IEEE Trans. Magn., 39 (2003), pp. 1175–1178.
[16] H. De Gersem, M. Wilke, M. Clemens, and T. Weiland, *Efficient modelling techniques for complicated boundary conditions applied to structured grids.* Proc. EMF 2003, Aachen, Germany, 06.-09. Oct. 2003. Accepted for publication in COMPEL, 2004.
[17] S. Reitzinger and J. Schöberl, *An algebraic multigrid method for finite element discretizations with edge elements*, Num. Lin. Alg. Appl., 9 (2002), pp. 223–238.
[18] E. Tonti, *A direct formulation of field laws: The cell method*, CMES, 2 (2001), pp. 237–258.
[19] T. Weiland, *A discretization method for the solution of Maxwell's equations for six-component fields*, Electronics and Communications AEÜ, 31 (1977), pp. 116–120.
[20] J. Yuan, M. Clemens, and T. Weiland, *Simulation of hysteresis effects with the classical Preisach model in FI^2TD methods.* Proc. ISEM 2003, Versailles, France. Full paper to appear in special issue of the Int. J. Appl. Electromagn. Mech., IOS Press., 2004.
[21] ———, *The Jiles-Atherton model combined with the Newton-Raphson method for the simulation of transient hysteretic magnetic field problems.* To appear in Proc. XVIII Symposium on Electromagnetic Phenomena in Nonlinear Circuits (EPNC 2004), Poznan, Poland, June 2004.
[22] ———, *Solution of transient hysteretic magnetic field problems with hybrid Newton-Polarization methods.* To appear in Proc. 11th Biennial IEEE Conference on Electromagnetic Field Computation (CEFC 2004), Seoul, Korea, June 2004.

Singularities of Electromagnetic Fields in the Eddy Current Limit

Monique Dauge
(joint work with Martin Costabel and Serge Nicaise)

This talk discusses the notion of eddy current limit for a conductor surrounded by an exterior dielectric medium and presents results from [15–17] about the singularities of solutions when the conductor has corners and edges.

1. The eddy current limit

Let Ω_C be the conductor body. We assume that Ω_C is a three-dimensional polyhedron. To simplify the exposition we also assume that the boundary B of Ω_C has a single connected component. Let Ω be a ball, large enough to surround Ω_C. We consider the exterior domain $\Omega_E = \Omega \setminus \overline{\Omega}_C$. We denote by ε_C, μ_C and

σ_C the electric permittivity, the magnetic permeability and the conductivity of Ω_C, respectively, and by ε_E, μ_E and σ_E their values inside Ω_E. We assume that $\sigma_E = 0$. We consider the harmonic Maxwell equation at the given frequency ω:

$$(1) \qquad \begin{cases} (i) & \mathbf{curl}\ \mathsf{E} = -i\omega\,\mu\mathsf{H} & \text{in} \quad \Omega, \\ (ii) & \mathbf{curl}\ \mathsf{H} = (\sigma + i\omega\,\varepsilon)\mathsf{E} + \mathbf{j}_0 & \text{in} \quad \Omega, \\ (iii) & \mathsf{E} \times n = 0 \quad \text{and} \quad \mathsf{H}\cdot n = 0 & \text{on} \quad \partial\Omega. \end{cases}$$

Here $\mathbf{j}_0$ is a divergence free field (the source current density) with support inside Ω_C and σ denotes the piecewise constant equal to σ_C inside Ω_C and 0 inside Ω_E. Similar conventions hold for ε and μ. Taking the divergence of equation (1) *(ii)*, we obtain:

$$(2) \qquad \operatorname{div}(i\omega\varepsilon + \sigma)\mathsf{E} = 0 \quad \text{in} \quad \Omega.$$

The time-harmonic eddy current problem [2, 8, 11, 20] consists in neglecting $\omega\varepsilon$ in (1) in the case when $\sigma >> \omega\varepsilon$ and reads:

$$(3) \qquad \begin{cases} (i) & \mathbf{curl}\ \mathsf{E} = -i\omega\mu\mathsf{H} & \text{in} \ \Omega, \\ (ii) & \mathbf{curl}\ \mathsf{H} = \sigma\mathsf{E} + \mathbf{j}_0 & \text{in} \ \Omega, \\ (iii) & \mathsf{E} \times n = 0 \quad \text{and} \quad \mathsf{H}\cdot n = 0 & \text{on} \ \partial\Omega. \end{cases}$$

Let us write $\mathsf{E}_C = \mathsf{E}|_{\Omega_C}$ and $\mathsf{E}_E = \mathsf{E}|_{\Omega_E}$. Taking the divergence of (3) *(ii)*, we only obtain $\operatorname{div}\mathsf{E}_C = 0$ in Ω_C and $\mathsf{E}_C \cdot n = 0$ on B, which *has to be completed by the gauge conditions*:

$$(4) \qquad \operatorname{div}\mathsf{E}_E = 0 \ \text{in} \ \Omega_E \quad \text{and} \quad \int_B \mathsf{E}_E \cdot n \, \mathrm{d}S = 0.$$

Let us assume for simplicity that $\varepsilon_C \simeq \varepsilon_E$ and let us introduce our small parameter δ as

$$\delta = \frac{\varepsilon_C}{\sigma_C}\,.$$

Let us consider σ, μ and $\omega > 0$ as fixed and denote by $(\mathsf{E}^\delta, \mathsf{H}^\delta)$ the solution of (1) and by $(\mathsf{E}^0, \mathsf{H}^0)$ the solution of (3). We have proved in [16]

$$(5) \qquad \left\|\mathsf{E}^\delta - \mathsf{E}^0\right\|_{L^2(\Omega)} + \left\|\mathsf{H}^\delta - \mathsf{H}^0\right\|_{L^2(\Omega)} \le C\delta.$$

This notion of limit corresponds to that presented in [11, Ch.4], whereas it somewhat differs from the point of view adopted in [2] where a zero frequency limit is considered for both problems (1) and (3). However (5) does not answer completely the question of knowing whether the eddy current approximation is valid when we are given a set of parameters σ, μ, ε and ω. Let us set $\hat\varepsilon = \varepsilon/\delta$. The interior equations for the electric field E^δ take the form:

$$(6) \qquad \begin{cases} (i) \ \mathbf{curl}\ \mu^{-1}\,\mathbf{curl}\ \mathsf{E}^\delta + i\omega\sigma\mathsf{E}^\delta - \delta\omega^2\hat\varepsilon\,\mathsf{E}^\delta = -i\omega\mathbf{j}_0 & \text{in} \quad \Omega_C, \\ (ii) \qquad \mathbf{curl}\ \mu^{-1}\,\mathbf{curl}\ \mathsf{E}^\delta - \delta\omega^2\hat\varepsilon\,\mathsf{E}^\delta = 0 & \text{in} \quad \Omega_E. \end{cases}$$

We can see that *(i)* tends to its eddy current counterpart as soon as $\delta\omega$ is small, whereas for equation *(ii)* approaching the eddy current limit requires that $\omega^2\varepsilon\mu$ also is small at the scale of Ω_E. Another asymptotic effect may occur when

$\omega\mu\sigma >> 1$: The skin effect produces a strong concentration of the electromagnetic field inside the conductor in a very narrow layer around its surface [9].

2. Singularities

The equations (6) combined with the zero divergence constraint inside $\Omega_C \cup \Omega_E$ and transmission conditions on B, produce an elliptic boundary value problem on Ω. Like any elliptic boundary value problem in a domain with corners and edges [18, 19, 22, 24], the electric or magnetic Maxwell problems have singular solutions (the "singularities") [13, 15]. In the present situation of a polyhedral conductor surrounded by a dielectric medium, the issue is the investigation [16, 17] of the singularities of the eddy current problem (3) together with the way in which the singularities of the transmission problem (1) transform as $\delta \to 0$ in the eddy current limit. Let us define $\alpha = (\alpha_C, \alpha_E)$ by

$$i\omega\alpha = i\omega\delta\hat{\varepsilon} + \sigma.$$

The "electric" singularities of problems (1) and (3) are those of the operator

$$\left\{\begin{array}{lll} (i) \ \ \mathbf{curl}\, \mu_C^{-1}\, \mathbf{curl}\, \mathsf{E} - \nabla \operatorname{div} \mathsf{E} & \text{in} & \Omega_C, \\ (ii)\ \mathbf{curl}\, \mu_E^{-1}\, \mathbf{curl}\, \mathsf{E} - \nabla \operatorname{div} \mathsf{E} & \text{in} & \Omega_E, \end{array}\right. \tag{7}$$

with the essential transmission conditions:

$$[\mathsf{E} \times n] = 0 \quad \text{and} \quad [\alpha \mathsf{E} \cdot n] = 0 \qquad \text{on} \quad B, \tag{8}$$

which we complement by the Neumann type transmission conditions

$$[\mu^{-1}\, \mathbf{curl}\, \mathsf{E} \times n] = 0 \quad \text{and} \quad [\operatorname{div} \alpha \mathsf{E}] = 0 \qquad \text{on} \quad B. \tag{9}$$

Problem (7)-(9) is the principal part of one of the regularized operators associated with problem (1).

According to the classification of [13, 15], problem (7)-(9) has mainly two types of singularities, Type 1 and Type 2, at each corner and each edge of Ω_C. To each corner or edge we associate two cones Γ_C and Γ_E together with their interface I. For a corner point $\mathbf{c}$, Γ_C and Γ_E coincide with Ω_C and Ω_E, respectively, in a neighborhood of $\mathbf{c}$. For an edge we have similar definitions where Γ_C and Γ_E are plane sectors such that the diehedra $\Gamma_C \times \mathbb{R}$ and $\Gamma_E \times \mathbb{R}$ coincide with Ω_C and Ω_E in a neighborhood of the edge. The singularities are homogeneous functions on $\Gamma_C \cup \Gamma_E$.

The singularities of Type 1 are the gradients $\nabla\Phi$ in $\Gamma_C \cup \Gamma_E$ of a potential function $\Phi = (\Phi_C, \Phi_E)$ which is itself a singularity of the scalar transmission problem, *cf* [25, 26]:

$$\Delta\Phi_C = 0 \ \text{ in } \ \Gamma_C, \quad \Delta\Phi_E = 0 \ \text{ in } \ \Gamma_E, \quad [\Phi] = 0 \ \text{ and } \ [\alpha\partial_n\Phi] = 0 \ \text{ on } \ I, \tag{10}$$

the last transmission condition becoming $\partial_n\Phi_C = 0$ on I in the eddy current limit $\delta = 0$: In the latter case, either $\Phi_C = 0$ and Φ_E is a Dirichlet singularity of the Laplace problem on Γ_E, or Φ_C is a Neumann singularity of the Laplace problem on Γ_C and Φ_E has the same Dirichlet traces as Φ_C (and the same degree of homogeneity).

The singularities of Type 2 are electric fields associated with magnetic fields of the form $\nabla\Psi$ where the scalar potential $\Psi = (\Psi_C, \Psi_E)$ is a singularity of the transmission problem

$$(11) \quad \Delta\Psi_C = 0 \;\text{ in }\; \Gamma_C, \quad \Delta\Psi_E = 0 \text{ in } \Gamma_E, \quad [\Psi] = 0 \;\text{ and }\; [\mu\partial_n\Psi] = 0 \;\text{ on }\; I.$$

If the permeability μ has no jump, solutions of (11) still exist, but they are polynomials and do not decrease the regularity of Maxwell solutions.

3. Regularity

Let β_α and β_μ be the limiting regularity Sobolev exponents for the transmission Laplace operators $\operatorname{div}\alpha\nabla$, *cf* singularities (10), and $\operatorname{div}\mu\nabla$ respectively, *cf* singularities (11). Then if the data $\mathbf{j}_0$ is regular enough, the solution $\mathbf{E}$ of (1) satisfies

$$\mathbf{E}_C \in H^s(\Omega_C) \quad \text{and} \quad \mathbf{E}_E \in H^s(\Omega_E), \quad \forall s < \min\{\beta_\alpha - 1, \beta_\mu\}.$$

Moreover we have a decomposition of [4, 5]'s type: $\mathbf{E}$ can be split into $\nabla\Phi + \mathbf{E}^{\mathsf{reg}}$ with

$$\mathbf{E}_C^{\mathsf{reg}} \in H^s(\Omega_C) \quad \text{and} \quad \mathbf{E}_E^{\mathsf{reg}} \in H^s(\Omega_E), \quad \forall s < \min\{\beta_\alpha, \beta_\mu\}.$$

Concerning the eddy current problem (3), we introduce the limiting regularity Sobolev exponents β_E^{Dir} and β_C^{Neu} for the Dirichlet problem on Ω_E and the Neumann problem on Ω_C, respectively. Let us assume for simplicity that μ has no jump (which amounts to setting $\beta_\mu = \infty$). Then the solution $\mathbf{E}$ of (3) satisfies

$$\mathbf{E}_C \in H^s(\Omega_C),\ \forall s < \beta_C^{\mathsf{Neu}} - 1, \quad \text{and} \quad \mathbf{E}_E \in H^s(\Omega_E),\ \forall s < \min\{\beta_C^{\mathsf{Neu}}, \beta_E^{\mathsf{Dir}}\} - 1.$$

Moreover, we may split $\mathbf{E}$ into $\nabla\Phi + \mathbf{E}^{\mathsf{reg}}$ with

$$\mathbf{E}_C^{\mathsf{reg}} \in H^s(\Omega_C),\ \forall s < \beta_C^{\mathsf{Neu}} \quad \text{and} \quad \mathbf{E}_E^{\mathsf{reg}} \in H^s(\Omega_E),\ \forall s < \min\{\beta_C^{\mathsf{Neu}}, \beta_E^{\mathsf{Dir}}\}.$$

Thus, if the conductor Ω_C is convex, it may happen that, in the eddy current limit, the solution inside the conductor is more regular than outside. This effect does not occur for $\delta \neq 0$. In fact, the conductor part Φ_C of certain singularities of (1) is vanishing as $\delta \to 0$.

4. Short conclusion about the numerical approximation

The resolution of the eddy current problem is made by eliminating either the electric field ($\mathbf{H}$-formulation or magnetic approach [1, 8, 11]) or the magnetic field ($\mathbf{E}$-formulation or electric approach [2, 3, 8, 11, 20]) or combining both [12]. The magnetic approach can be preferred because the magnetic field in Ω_E is irrotational. Thus a coupled FEM-BEM method can be used to compute $\mathbf{H}$ [10, 23]. Concerning the use of edge elements, see [6, 7].

We would like to end by the "usual" warning: In the presence of reentrant corners (i.e. any situation where Ω_C is a polyhedron in $\mathbf{E}$-formulation, and the case when Ω_C is a non-convex polyhedron in $\mathbf{H}$-formulation) certain methods lead to wrong results. This is the case for the *plain regularization* by a divergence term, used with nodal elements, or, even, certain edge elements which do not satisfy the discrete compactness property, see the review papers [14, 21].

References

[1] A. Alonso Rodríguez, P. Fernandes, and A. Valli, *Weak and strong formulations for the time-harmonic eddy-current problem in general multi-connected domains*, European J. Appl. Math., 14 (2003), pp. 387–406.
[2] H. Ammari, A. Buffa, and J.-C. Nédélec, *A justification of eddy currents model for the Maxwell equations*, SIAM J. Appl. Math., 60 (2000), pp. 1805–1823 (electronic).
[3] A. Bermúdez, R. Rodríguez, and P. Salgado, *Numerical analysis of the electric field formulation of an eddy current problem*, C. R. Math. Acad. Sci. Paris, 337 (2003), pp. 359–364.
[4] M. Birman and M. Solomyak, *L^2-theory of the Maxwell operator in arbitrary domains*, Russ. Math. Surv., 42 (6) (1987), pp. 75–96.
[5] ———, *On the main singularities of the electric component of the electro-magnetic field in regions with screens*, St. Petersbg. Math. J., 5 (1) (1993), pp. 125–139.
[6] O. Bíró, *Edge element formulations of eddy current problems*, Comput. Methods Appl. Mech. Engrg., 169 (1999), pp. 391–405.
[7] O. Bíró and K. Preis, *Gauged current vector potential and reentrant corners in the FEM analysis of 3D eddy currents*, in Computational electromagnetics (Kiel, 2001), vol. 28 of Lect. Notes Comput. Sci. Eng., Springer, Berlin, 2003, pp. 1–10.
[8] A. Bossavit, *Two dual formulations of the* 3*-D eddy-currents problem*, COMPEL, 4 (1985), pp. 103–116.
[9] ———, *Stiff problems and boundary layers in electricity: a mathematical analysis of skin-effect*, in BAIL IV (Novosibirsk, 1986), vol. 8 of Boole Press Conf. Ser., Boole, Dún Laoghaire, 1986, pp. 233–240.
[10] ———, *The computation of eddy-currents, in dimension* 3*, by using mixed finite elements and boundary elements in association*, Math. Comput. Modelling, 15 (1991), pp. 33–42. Boundary integral equation methods (boundary element methods).
[11] ———, *Électromagnétisme, en vue de la modélisation*, vol. 14 of Mathématiques & Applications (Berlin) [Mathematics & Applications], Springer-Verlag, Paris, 1993.
[12] ———, *"Hybrid" electric-magnetic methods in eddy-current problems*, Comput. Methods Appl. Mech. Engrg., 178 (1999), pp. 383–391.
[13] M. Costabel and M. Dauge, *Singularities of electromagnetic fields in polyhedral domains*, Arch. Rational Mech. Anal., 151 (2000), pp. 221–276.
[14] ———, *Computation of resonance frequencies for Maxwell equations in non smooth domains*, Topics in Computational Wave Propagation, (M. Ainsworth, P. Davies, D. Duncan, P. Martin, B. Rynne, Eds.) Lecture Notes in Computational Science and Engineering., Vol. 31, Springer 2003, pp. 125–161.
[15] M. Costabel, M. Dauge, and S. Nicaise, *Singularities of Maxwell interface problems*, M2AN Math. Model. Numer. Anal., 33 (1999), pp. 627–649.
[16] ———, *Singularities of eddy current problems*, M2AN Math. Model. Numer. Anal., 37 (2003), pp. 807–831.
[17] ———, *Corner singularities of Maxwell interface and eddy current problems*, in Operator Theoretical Methods and Applications to Mathematical Physics The Erhard Meister Memorial Volume, I. Gohberg, A. Ferreira dos Santos, F.-O. Speck, F. Sepulveda Teixeira, and W. Wendland, eds., Operator Theory: Advances and Applications, Vol. 147, Springer-Birkhäuser, 2004, p. 473.
[18] M. Dauge, *Elliptic Boundary Value Problems in Corner Domains – Smoothness and Asymptotics of Solutions*, Lecture Notes in Mathematics, Vol. 1341, Springer-Verlag, Berlin, 1988.
[19] P. Grisvard, *Boundary Value Problems in Non-Smooth Domains*, Pitman, London, 1985.
[20] R. Hiptmair, *Symmetric coupling for eddy current problems*, SIAM J. Numer. Anal., 40 (2002), pp. 41–65 (electronic).

[21] ———, *Finite elements in computational electromagnetism*, Acta Numerica, (2002), pp. 237–339.
[22] V. A. Kondrat'ev, *Boundary-value problems for elliptic equations in domains with conical or angular points*, Trans. Moscow Math. Soc., 16 (1967), pp. 227–313.
[23] S. Meddahi and V. Selgas, *A mixed-FEM and BEM coupling for a three-dimensional eddy current problem*, M2AN Math. Model. Numer. Anal., 37 (2003), pp. 291–318.
[24] S. Nicaise, *Polygonal interface problems*, Methoden und Verfahren der Mathematischen Physik, 39, Verlag Peter D. Lang, Frankfurt-am-Main, 1993.
[25] S. Nicaise and A. M. Sändig, *General interface problems I/II*, Math. Methods Appl. Sci., 17 (1994), pp. 395–450.
[26] ———, *Transmission problems for the Laplace and elasticity operators: Regularity and boundary integral formulation*, Math. Methods Appl. Sci., 22 (1999), pp. 855–898.

Convergence of Collocation Methods for Time Domain Boundary Integral Equations

Penny J Davies
(joint work with Dugald B Duncan)

The problem of interest is to calculate the current induced on a perfectly conducting surface Γ when it is subjected to a transient electromagnetic field. Time–stepping solution schemes for this problem are often numerically unstable (see e.g. [2, 7, 9]), and our aim is to develop stable collocation approximations. Here we concentrate on the more straightforward case of acoustic scattering, where the same stability issues arise. This problem is to find the solution u of

$$\int_\Gamma \frac{u(\boldsymbol{x}', t-|\boldsymbol{x}'-\boldsymbol{x}|)}{|\boldsymbol{x}'-\boldsymbol{x}|}\, d\Gamma = a(\boldsymbol{x}, t) \tag{1}$$

given $a(\boldsymbol{x}, t)$ on $\Gamma\times(0,T)$, and assuming causality, namely that $u \equiv 0$ and $a \equiv 0$ for all $t \le 0$. Equation (1) is the single layer potential equation for acoustic scattering from the surface Γ, and we shall concentrate on the case in which Γ is a flat plate. Note that a can be calculated anywhere in space from (1) once $u|_\Gamma$ is known.

It follows from results of Ha-Duong [6, Thm. 3] and Lubich [8, §2.3] that for temporally smooth data $a(\cdot,\, t) \in H^{1/2}(\Gamma)$ which vanish near $t = 0$, equation (1) has a unique smooth solution $u(\cdot,\, t) \in H^{-1/2}(\Gamma)$.

Many authors have considered full Galerkin approximations (in time and space) for (1) and related boundary integral equations (see [7] for a description of the relevant theory and a survey of the literature). This approach is based on a sound theoretical framework, and stability is proved via an energy identity. However, the method is hard to implement (it involves evaluating integrals over complicated subregions of $\Gamma \times \Gamma \times (0, T)$), and collocation schemes are more frequently used in practice (the article [1] contains an overview of different solution methods for problems such as (1)).

In a collocation approximation we suppose that (1) holds at N_S points $\boldsymbol{x}_\beta \in \Gamma$ and at time $t^n = n\Delta t$ for $n = 1, 2, \ldots$

$$a(\boldsymbol{x}_\beta, t^n) = \int_\Gamma \frac{u(\boldsymbol{x}', t^n - |\boldsymbol{x}' - \boldsymbol{x}_\beta|)}{|\boldsymbol{x}' - \boldsymbol{x}_\beta|}\, d\Gamma\,. \tag{2}$$

The unknown u is then approximated (in time and space), and the integral is approximated or evaluated to give

$$\underline{a}^n = \sum_{m=0}^{n-1} Q_m\, \underline{U}^{n-m} \tag{3}$$

in terms of (very sparse) matrices $Q_m \in \mathbb{R}^{N_S \times N_S}$, where $\underline{U}^m = \left\{U_\beta^m\right\}_\beta$ and $U_\beta^m \approx u(\boldsymbol{x}_\beta, t^m)$. Rearranging gives the time–stepping algorithm

$$Q_0\, \underline{U}^n = \underline{a}^n - \sum_{m=1}^{n-1} Q_m\, \underline{U}^{n-m}\,.$$

The sparsity of Q_0 means that solving this equation for the unknown $\underline{U}^n$ is straightforward. However, numerical instability is often a problem for schemes of this type, with the computed solution typically exhibiting oscillating instabilities that grow exponentially in the time–step [2, 3, 9]. Insight can be obtained by comparing the continuous Fourier transform of (1) at spatial frequency $\boldsymbol{\omega}$ with the discrete Fourier transform of (3) at the same frequency, when Γ is assumed to be a flat infinite surface (i.e. $\Gamma = \mathbb{R}^2$), and the points $\boldsymbol{x}_\beta$ form a uniform square mesh. It can be shown [2, 3] that numerical stability in this case can be characterised by Fourier coefficients $p_n(\boldsymbol{\omega})$: if $|p_n(\boldsymbol{\omega})|$ remains bounded with n for all $\boldsymbol{\omega} \in S_h \equiv [-\pi/h, \pi/h]^2$, where h is the (spatial) grid size, then the scheme is stable. Unfortunately there appears to be no obvious way to verify this condition analytically, and we resort to testing it numerically for many individual frequencies $\boldsymbol{\omega} \in S_h$ to determine the stability of a collocation scheme for (1).

We have derived three new collocation schemes for (1), based on a piecewise linear approximation for u in space, and a piecewise linear or piecewise constant approximation for u in time. The resulting integrals are either evaluated exactly [5], or transformed to polar coordinates (R, θ) via the local change of variables $R = |\boldsymbol{x}' - \boldsymbol{x}_\beta|$ in (2), and then approximated using the trapezoidal rule in R and (nearly) exact integration in θ [4]. Numerical evaluation of the Fourier coefficients $p_n(\boldsymbol{\omega})$ for these three schemes indicate that they are all stable for any value of the mesh ratio $\Delta t/h$ [4, 5].

If a is assumed to be sufficiently smooth, then it can be shown that such stable schemes for (1) are second order convergent when Γ is an infinite flat plate. That is, there exists a constant C such that $\|u^n - U^n\|_h \le C\, h^2$ as $h \to 0$ for $t^n \le T$, where $\|\cdot\|_h$ denotes the discrete L^2-norm. The proof involves using estimates on (spatial) Fourier transforms [10], and (temporal) $Z-$transform techniques due to Lubich [8].

References

[1] M. Costabel, *Time-dependent problems with the boundary integral equation method*, in Encyclopedia of Computational Mechanics, E. Stein, R. de Borst and T. J. R. Hughes, eds., (in press).

[2] P. J. Davies, *Numerical stability and convergence of approximations of retarded potential integral equations*, SIAM J. Numer. Anal., 31 (1994), pp. 856–875.

[3] P. J. Davies and D. B. Duncan, *Averaging techniques for time marching schemes for retarded potential integral equations*, Applied Numerical Mathematics, 23 (1997), pp. 291–310.

[4] P. J. Davies and D. B. Duncan, *Numerical stability of collocation schemes for time domain boundary integral equations*, in Computational Electromagnetics: Proceedings of the GAMM Workshop, Kiel, 2001, C. Carstensen, S. A. Funken, W. Hackbusch, R. H. W. Hoppe, and P. Monk, eds., Springer-Verlag, 2003, pp. 51–66.

[5] P. J. Davies and D. B. Duncan, *Stability and convergence of collocation schemes for retarded potential integral equations*, SIAM J. Numer. Anal., (to appear).

[6] T. Ha-Duong, *On the transient acoustic scattering by a flat object*, Japan J. Appl. Math., 7 (1990), pp. 489–513.

[7] T. Ha-Duong, *On retarded potential boundary integral equations and their discretisation*, in Topics in Computational Wave Propagation: Direct and Inverse Problems, M. Ainsworth, P. J. Davies, D. B. Duncan, P. A. Martin, and B. P. Rynne, eds., Springer-Verlag, 2003, pp. 301–336.

[8] C. Lubich, *On the multistep time discretization of linear initial–boundary value problems and their boundary integral equations*, Numerische Mathematik, 67 (1994), pp. 365–389.

[9] B. P. Rynne and P. D. Smith, *Stability of time marching algorithms for the electric field equation*, J. Electromag. Waves & Appl., 4 (1990), pp. 1181–1205.

[10] V. Thomée, *Convergence estimates for semi-discrete Galerkin methods for initial-value problems*, in Numerische, inbesondere approximationstheoretische Behandlung von Funktionalgleichungen (Lecture Notes in Mathematics, 333), A. Dold and B. Eckmann, eds., Springer-Verlag, 1973, pp. 243–262.

$H^1, \boldsymbol{H}(\mathrm{curl})$ and $\boldsymbol{H}(\mathrm{div})$-Conforming Projection-Based Interpolation in Three Dimensions

L. Demkowicz
(joint work with A. Buffa)

The talk is concerned with optimal p- and hp-estimates for the *Projection Based Interpolation* operators [2, 3, 5].

Given a master tetrahedral element T, and a sequence of polynomial spaces reproducing the standard grad-curl-div exact sequence at the discrete level, we consider a family of projection-based interpolation operators [2, 3, 5], $\Pi, \Pi^{curl}, \Pi^{div}, P$ that make the de Rham diagram commute. The projection-based interpolation operators are defined through a sequence of projections done on edge, face, and element levels. A compact definition of the interpolation operators follows.

H^1-conforming::

$$\begin{cases} u_1(a) = u(a) \\ \|u - \Pi u\|_{\epsilon,e} \to \min \\ \|\boldsymbol{\nabla}_f(u - \Pi u)\|_{-\frac{1}{2}+\epsilon,f} \to \min \\ \|\boldsymbol{\nabla}(u - \Pi u)\|_{0,K} \to \min \end{cases}$$

$\boldsymbol{H}$(curl)-conforming::

$$\begin{cases} \displaystyle\int_e \boldsymbol{E}_t - \Pi^{curl}\boldsymbol{E}_t = 0 \\ \displaystyle\| \int (\boldsymbol{E}_t - \Pi^{curl}\boldsymbol{E}_t)\|_{0,\epsilon} \to \min \\ \begin{cases} \|\mathrm{curl}_f(\boldsymbol{E}_t - \Pi^{curl}\boldsymbol{E}_t)\|_{-\frac{1}{2}+\epsilon,f} \to \min \\ (\boldsymbol{E}_t - \Pi^{curl}\boldsymbol{E}_t, \boldsymbol{\nabla}_f\phi)_{-\frac{1}{2}+\epsilon,f} = 0, \quad \forall\phi \in P^{p_f+1}_{-1} \end{cases} \\ \begin{cases} \|\boldsymbol{\nabla} \times (\boldsymbol{E} - \Pi^{curl}\boldsymbol{E})\|_{0,T} \to \min \\ (\boldsymbol{E} - \Pi^{curl}\boldsymbol{E}, \boldsymbol{\nabla}\phi)_{0,T} = 0, \quad \forall\phi \in P^{p+1}_{p_f+1,p_e+1} \end{cases} \end{cases}$$

$\boldsymbol{H}$(div)-conforming::

$$\begin{cases} \|\boldsymbol{F}_n - \Pi^{div}\boldsymbol{F}_n\|_{-\frac{1}{2}+\epsilon,f} \to \min \\ \begin{cases} \|\boldsymbol{\nabla} \circ (\boldsymbol{F} - \Pi^{div}\boldsymbol{F})\|_{0,T} \to \min \\ (\boldsymbol{F} - \Pi^{div}\boldsymbol{F}, \boldsymbol{\nabla} \times \boldsymbol{\phi})_{0,T} = 0, \quad \forall\boldsymbol{\phi} \in P^{p+1}_{p_f+1} \end{cases} \end{cases}$$

The task is to develop optimal error estimates with respect to polynomial degree p. As the operators are polynomial preserving, this in turn, leads to optimal hp-estimates as well. A major difficulty in deriving such estimates in 3D is the "loss of traces" at vertices. The trace space of $H^{1+\epsilon}(\mathrm{T})$ for a face f is $H^{\frac{1}{2}+\epsilon}(f)$, and $H^{\epsilon}(e)$ for an edge e, but we have no trace at a vertex v. In other words, $H^{1+\epsilon}(T)$

is not continously embedded in the space of continuous functions. This lack of regularity prevents the use of the reasoning used in 2D [3].

The key idea in deriving the estimates, is to compare the interpolation errors with two families of commuting projections, defined on element T and face f levels, see the commuting diagrams below.

$$
\begin{array}{ccccccc}
\mathbb{R} \to H^{\frac{3}{2}+\epsilon}(T) & \xrightarrow{\boldsymbol{\nabla}} & \boldsymbol{H}^{\epsilon}(\mathrm{curl}, T) \cap \boldsymbol{H}^{\frac{1}{2}+\epsilon}(T) & \xrightarrow{\boldsymbol{\nabla}\times} & \boldsymbol{H}^{\epsilon}(\mathrm{div}, T) & \xrightarrow{\boldsymbol{\nabla}\circ} & L^2 \\
\big\downarrow id \quad P^1 \big\downarrow \Pi & & P^{curl} \big\downarrow \Pi^{\mathrm{curl}} & & P^{div} \big\downarrow \Pi^{\mathrm{div}} & & \big\downarrow P \\
\mathbb{R} \to P^{p+1}_{p_e+1, p_f+1} & \xrightarrow{\boldsymbol{\nabla}} & \mathbf{P}^{p}_{p_e, p_f} & \xrightarrow{\boldsymbol{\nabla}\times} & \mathbf{P}^{p-1}_{p_f-1, p_e} & \xrightarrow{\boldsymbol{\nabla}\circ} & P^{p-2}
\end{array}
$$

$$
\begin{array}{ccccccccc}
\mathbb{R} & \longrightarrow & H^{\frac{1}{2}+\epsilon}(f) & \xrightarrow{\boldsymbol{\nabla}} & \boldsymbol{H}^{-\frac{1}{2}+\epsilon}(\mathrm{curl}, f) & \xrightarrow{\boldsymbol{\nabla}\times} & H^{-\frac{1}{2}+\epsilon}(f) & \longrightarrow & \mathbf{0} \\
\big\downarrow id & & P^{\frac{1}{2}+\epsilon} \big\downarrow \Pi & & \big\downarrow \Pi^{\mathrm{curl}} & & \big\downarrow P & & \\
\mathbb{R} & \longrightarrow & P^{p_f+1}_{p_e+1} & \xrightarrow{\boldsymbol{\nabla}} & \mathbf{P}^{p_f}_{p_e} & \xrightarrow{\boldsymbol{\nabla}\times} & P^{p_f-2} & \longrightarrow & \mathbf{0}
\end{array}
$$

Besides the commuting projection operators, instrumental in deriving the estimates are

- the existence of polynomial preserving, extension operators defined for a tetrahedral face [1],

$$
\begin{array}{ccc}
H^{\frac{1}{2}+\epsilon}(f) & \xrightarrow{\boldsymbol{\nabla}} & \boldsymbol{H}^{-\frac{1}{2}+\epsilon}(\mathrm{curl}, f) \\
E^{grad} \big\uparrow \big\downarrow Tr^{grad} & & E^{curl} \big\uparrow \big\downarrow Tr^{curl} \\
H^{\epsilon}(\partial f) & \xrightarrow{\frac{\partial}{\partial s}} & H^{-1+\epsilon}(\partial f)
\end{array}
$$

$$
\begin{array}{ccc}
P^{p}_{p_e}(f) & \xrightarrow{\boldsymbol{\nabla}} & \boldsymbol{P}^{p-1}_{p_e-1}(f) \\
E^{grad} \big\uparrow \big\downarrow Tr^{grad} & & E^{curl} \big\uparrow \big\downarrow Tr^{curl} \\
P^{p_e}(\partial f) & \xrightarrow{\frac{\partial}{\partial s}} & P^{p_e-1}(\partial f)
\end{array}
$$

and on the element lecel (*conjectured*),

$$
\begin{tikzcd}
H^1(T) \arrow[r, "\boldsymbol{\nabla}"] \arrow[d, shift left, "Tr^{grad}"] & \boldsymbol{H}(\mathrm{curl}, T) \arrow[r, "\boldsymbol{\nabla}\times"] \arrow[d, shift left, "Tr^{curl}"] & \boldsymbol{H}(\mathrm{div}, T) \arrow[d, shift left, "Tr^{div}"] \\
H^{\frac{1}{2}}(\partial T) \arrow[r, "\boldsymbol{\nabla}_{\partial T}"] \arrow[u, shift left, "E^{grad}"] & \boldsymbol{H}^{-\frac{1}{2}}(\mathrm{curl}, \partial T) \arrow[r, "\boldsymbol{\nabla}_{\partial T}\times"] \arrow[u, shift left, "E^{curl}"] & H^{-\frac{1}{2}}(\partial T) \arrow[u, shift left, "E^{div}"]
\end{tikzcd}
$$

$$
\begin{tikzcd}
P^p_{p_f,p_e}(T) \arrow[r, "\boldsymbol{\nabla}"] \arrow[d, shift left, "Tr^{grad}"] & \boldsymbol{P}^{p-1}_{p_f-1,p_e-1}(T) \arrow[r, "\boldsymbol{\nabla}\times"] \arrow[d, shift left, "Tr^{curl}"] & \boldsymbol{P}^{p-2}_{p_f-2}(T) \arrow[d, shift left, "Tr^{div}"] \\
P^{p_f}_{p_e}(\partial T) \arrow[r, "\boldsymbol{\nabla}_{\partial T}"] \arrow[u, shift left, "E^{grad}"] & \boldsymbol{P}^{p_f-1}_{p_e-1}(\partial T) \arrow[r, "\boldsymbol{\nabla}_{\partial T}\times"] \arrow[u, shift left, "E^{curl}"] & P^{p_f-2}(\partial T) \arrow[u, shift left, "E^{div}"]
\end{tikzcd}
$$

and,

- the existence of polynomial preserving, right inverses G, K, D of grad, curl, and div operators,

$$
\begin{array}{ccccc}
H^{\frac{1}{2}+\epsilon}(f) & \xrightarrow{\boldsymbol{\nabla}} & \boldsymbol{H}^{-\frac{1}{2}+\epsilon}(\mathrm{curl}, f) & \xrightarrow{\boldsymbol{\nabla}\times} & H^{-\frac{1}{2}+\epsilon}(f) \\
P^{p_f+1}_{p_e+1} & \xleftarrow{G} & \mathbf{P}^{p_f}_{p_e} & \xleftarrow{K} & P^{p_f-1}
\end{array}
$$

that are instrumental in proving discrete versions of Friedrichs inequalities.

Under the conjecture on the existence of polynomial preserving, extension operators, we can prove the following interpolation error estimates.

$$
\begin{aligned}
\|u - \Pi u\|_{1,T} &\leq Cp^{-(r-\epsilon)}\|u\|_{1+r,T} \\
\|\boldsymbol{E} - \Pi^{curl}\boldsymbol{E}\|_{0,curl,T} &\leq Cp^{-(r-\epsilon)}\|\boldsymbol{E}\|_{r,curl,T} \\
\|\boldsymbol{F} - \Pi^{div}\boldsymbol{F}\|_{0,div,T} &\leq Cp^{-(r-\epsilon)}\|\boldsymbol{F}\|_{r,div,T}
\end{aligned}
$$

The interpolation theory is not crucial for the convergence analysis but it forms the backbone on fully automatic *hp*-adaptive startegies that deliver exponential convergence for both elliptic and Maxwell problems [4].

References

[1] M. Ainsworth, and L. Demkowicz, *Explicit Polynomial Preserving Trace Liftings on a Triangle*, ICES Report 03-47.

[2] L. Demkowicz, P. Monk, L. Vardapetyan, and W. Rachowicz, *De Rham Diagram for hp Finite Element Spaces*, Mathematics and Computers with Applications 39(7-8) (2000), pp29-38.

[3] L. Demkowicz, and I. Babuška, *Optimal p Interpolation Error Estimates for Edge Finite Elements of Variable Order in 2D*, SIAM J. Numer. Anal. 41(4) (2003), pp1195-1208.

[4] L. Demkowicz, *Fully Automatic hp-Adaptivity for Maxwell's Equations*, ICES Report 03-30.
[5] L. Demkowicz, A. Buffa, $H^1, \boldsymbol{H}(\text{curl})$ *and* $\boldsymbol{H}(\text{div})$*-Conforming Projection-Based Interpolation in Three Dimensions* ICES Report, in preparation.
[6] J. Gopalakrishnan and L. Demkowicz, *Quasioptimality of Some Spectral Mixed Methods*, TICAM Report 03-32, accepted to SIAM Journal on Numerical Analysis.

High-Order Time Stepping Methods for Electromagnetics

Tobin A. Driscoll

High-order and spectral methods for spatial discretization have significant advantages in accuracy and efficiency over first- and second-order schemes [6, 9]. Such discretizations are most naturally paired with high-order methods in time, which yield similar benefits.

There are two aspects of discretizing Maxwell's equations in particular that lead to consideration of special time-stepping methods: *staggering* and *linear stiffness*. Staggering in spacetime was suggested by Yee for his famous second-order scheme [12]; it improves equal-cost accuracy by a factor of four and stable time step size by a factor of two over the related collocated method. The benefits of staggering are significantly increased at higher orders of accuracy [5, 8]. Linear stiffness arises from perfectly matched absorbing layers [1] that decay signals rapidly. Such decay can severely constrain the allowable time step size of a standard method. However, since the stiff aspect of the problem is linear, there are several strategies for restoring large time step sizes in high-order methods.

1. Staggering

The nature of Maxwell's equations allows **E** and **H** field components to be interlaced in time, as Yee showed in [12]. (The staggering can be done in space as well; the choices of whether to stagger in space and time may be made independently.) Other pure propagation problems, such as elastic waves, follow the same pattern [8].

We represent the semidiscrete evolution of a system eligible for time staggering as

$$u_t = f(t, v), \qquad v_t = g(t, u). \tag{1}$$

For instance, the second-order leapfrog in time used by Yee can be expressed as

$$v_{n+1/2} - v_{n-1/2} = \Delta t\, g(t_n, u_n), \qquad u_{n+1} - u_n = \Delta t\, f(t_{n+1/2}, v_{n+1/2}).$$

This method has an error constant that is 1/4 of that for the same method on an integer-level-only grid, and a *stability ordinate* (extent of the stability region along the imaginary axis that represents propagation) twice as large. We can increase the order of accuracy of leapfrog by using either more past steps or more interior stages, as was shown in [8].

Multistep methods are created in two variants, corresponding to whether one uses past values of the solution or of its derivative. (As with nonstaggered classical methods, trying to use both simultaneously leads to unstable methods.) We

Table 1. Comparison of staggered to classical nonstaggered methods.

Error constants

order	AB	RK	ABS	BDS	RKS
2	0.417	0.667	0.042	0.042	0.042
3	0.375	1.125	0.042	0.042	0.646
4	0.349	2.133	0.039	0.037	0.133
7	0.304	–	0.031	unstable	?
8	0.295	–	0.029	unstable	?

Stability ordinates

order	AB	RK	ABS	BDS	RKS
2	0	0	2.00	2.00	2.00
3	0.72	0.58	1.71	1.67	1.04
4	0.43	0.71	1.33	1.00	1.43
7	0.06	–	0.37	unstable	?
8	0.03	–	0.21	unstable	?

call these variants *staggered backward differentiation* (BDS) and *staggered Adams–Bashforth* (ABS), respectively, by analogy with the classical methods.[2] For example, the fourth-order ABS formula is

$$v_{n+1/2} - v_{n-1/2} = \frac{\Delta t}{24}\big(26u_n - 5u_{n-1} + 4u_{n-2} - u_{n-3}\big)$$

The coefficients and stability regions of ABS and BDS methods are cataloged in [8]. Only BDS methods of orders 2–4 and ABS methods of orders 2 and 3,4,7,8,11,12,…, have nontrivial stability ordinates. Past second order, all these methods are dissipative.

Staggered multistage methods are constructed on a more ad-hoc basis. The best fourth-order method known is [8]

$$\begin{aligned}
d_1 &= \Delta t\, f(t_{n+1/2}, v_{n+1/2})\\
d_2 &= \Delta t\, g(t_n, u_n)\\
d_3 &= \Delta t\, f(t_{n+1/2} - \Delta t, v_{n+1/2} - d_2)\\
d_4 &= \Delta t\, g(t_n + \Delta t, u_n + d_1)\\
d_5 &= \Delta t\, f(t_{n+1/2} + \Delta t, v_{n+1/2} + d_4)\\
u_{n+1} &= u_n + \tfrac{11}{12} d_1 + \tfrac{1}{24} d_3 + \tfrac{1}{24} d_5,
\end{aligned}$$

with a symmetric formula for advancing v. This method requires four full function evaluations per full step, and it has a stability ordinate (normalized by the number of stages) of 1.43, compared to $1/\sqrt{2}$ for classical fourth-order Runge–Kutta.

Staggered methods are compared to their classical counterparts in Table 1. The table clearly demonstrates that the accuracy benefits of staggering increase with

[2] Note, however, that they are all explicit, unlike classical BD formulas.

the order of accuracy, and the stability ordinates controlling stable time step sizes are better by a factor of two or more.

2. Linear stiffness

Suppose a state $u(t)$ evolves according to

$$u_t = f(u) - Su, \tag{2}$$

where S is a linear operator and $\rho(S) \gg \rho(f')$, where ρ is spectral radius. Such systems arise quite frequently: the nonlinear Schrödinger, Korteweg–de Vries, Kuramoto–Sivashinsky, Gray–Scott, and Navier–Stokes equations are a few examples. In these cases the large spectrum of S is due to the presence of high-order spatial derivatives. In the Maxwellian case S represents the (perhaps nonphysical) conductive losses due to a perfectly matched layer [1].

The large spectrum of S creates an unacceptably strict stability condition on the time step size of explicit methods, but in most cases fully implicit methods are infeasible. A number of strategies have been devised to circumvent this difficulty at high orders of accuracy. They all work best when S is diagonal, as is the case in Maxwell's equations and in nonlinear PDEs under Fourier discretization.

One of the simplest ideas is the *integrating factor*, which transforms (2) to

$$\tfrac{d}{dt}(e^{St}u) = e^{St}f(u). \tag{3}$$

The evolution of $e^{St}u$ encounters no stiffness. However, the presence of the rapidly-varying exponential creates an accuracy penalty for a classical method. A better approach is to discretize (3) using a specialized method that explicitly incorporates the exponential. Such methods go by the name of *exact linear part* or *exponential time differencing*, are available in both multistep and multistage forms, and perform well in practice [2, 3, 10].

Another approach is to generalize the well known second-order Strang splitting, in which (2) over $[0, \Delta t]$ is replaced by

$$u_t = 2f(u) \text{ on } [0, \tfrac{1}{4}\Delta t]; \qquad u_t = 2Su \text{ on } [\tfrac{1}{4}\Delta t, \tfrac{3}{4}\Delta t];$$
$$u_t = 2f(u) \text{ on } [\tfrac{3}{4}\Delta t, \Delta t].$$

Fractional time step sequences can be found to give split-step methods of any even order [11, 13]. For fourth and sixth order, 7 and 15 substeps per step are needed, respectively, and in each case some steps must be negative, which makes these methods problematic for diffusion. However, they can be designed to conserve energy and symplecticness.

A third approach is to use *linearly implicit* methods, which marry explicit methods for the nonlinear term f and an implicit method for the stiff, linear—and hopefully diagonal—S. These have been shown to be quite effective when used in a heterogeneous discretization [4, 7], in which nonstiff components (e.g., free space propagation in Maxwell) are propagated by classical methods.

The best methods in each approach have mild or no stability restrictions and are orders of magnitude more efficient than their second-order counterparts. The composite method of [4], in particular, is easy to implement and appears to be at least as effective as any other of this type for a variety of test problems.

References

[1] J.-P. Berenger, *A perfectly matched layer for the absorption of electromagnetic waves*, J. Comput. Phys., 114 (1994), pp. 185–200.

[2] G. Beylkin, J. M. Keiser, and L. Vozovoi, *A new class of time discretization schemes for the solution of nonlinear PDEs*, J. Comput. Phys., 147 (1998), pp. 362–387.

[3] S. M. Cox and P. C. Matthews, *Exponential time differencing for stiff systems*, J. Comput. Phys., 176 (2002), pp. 430–455.

[4] T. A. Driscoll, *A composite Runge–Kutta method for the spectral solution of semilinear PDE*, J. Comput. Phys., 182 (2002), pp. 357–367.

[5] B. Fornberg, *High-order finite differences and the pseudospectral method on staggered grids*, SIAM J. Numer. Anal., 27 (1990), pp. 904–918.

[6] B. Fornberg, *A Practical Guide to Pseudospectral Methods*, Cambridge Universtiy Press, Cambridge, 1996.

[7] B. Fornberg and T. A. Driscoll, *A fast spectral algorithm for nonlinear wave equations with linear dispersion*, J. Comput. Phys., 155 (1999), pp. 456–467.

[8] M. Ghrist, B. Fornberg, and T. A. Driscoll, *Staggered time integrators for wave equations*, SIAM J. Numer. Anal., 38 (2000), pp. 718–741.

[9] J. S. Hesthaven, *Spectral penalty methods*, Appl. Numer. Math., 33 (2000), pp. 23–41.

[10] A.-K. Kassam and L. N. Trefethen, *Fourth-order time stepping for stiff pdes*, SIAM J. Sci. Comput., (to appear).

[11] R. I. McLachlan and G. R. W. Quispel, *Splitting methods*, Acta Numer., 11 (2002), pp. 341–434.

[12] K. S. Yee, *Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media*, IEEE Trans. Antennas and Prop., 14 (1966), pp. 302–307.

[13] H. Yoshida, *Construction of higher order symplectic integrators*, Phys. Lett. A, 150 (1990), pp. 262–268.

Nonreflecting Boundary Conditions for Computational Electromagnetics

Marcus J. Grote

(joint work with Wolfgang Bangerth, Joseph B. Keller and Christoph Kirsch)

For the numerical solution of wave scattering problems in unbounded media, a well-known approach is to enclose all obstacles, inhomogeneities and nonlinearities with an artificial boundary B. A boundary condition is then imposed on B, which leads to a numerically solvable boundary-value problem in a finite computational domain Ω. The boundary condition should be chosen such that the solution of the problem in Ω coincides with the restriction to Ω of the solution in the original unbounded region. Otherwise spurious reflections will appear at B, which will travel back into the interior computational region and spoil the numerical solution throughout Ω.

If the scatterer consists of several obstacles, which are well separated from each other, the use of a single artificial boundary to enclose the entire scattering region, becomes too expensive. Instead it is preferable to enclose every sub-scatterer by a separate artificial boundary B_j. Then we seek an exact boundary condition on $B = \bigcup B_j$, where each B_j surrounds a single computational sub-domain Ω_j. This boundary condition must not only let outgoing waves leave Ω_j without spurious reflection from B_j, but also propagate the outgoing wave from Ω_j to all other sub-domains Ω_ℓ, which it may reenter subsequently. To derive such an exact boundary condition, an analytic expression for the solution everywhere in the exterior region is needed. Neither absorbing boundary conditions [1, 2], nor perfectly matched layers [3] provide us with such a representation. Instead, we shall use exact Dirichlet-to-Neumann (DtN) conditions in the time-harmonic case, or nonreflecting boundary conditions (NBC) in the time dependent case, which are both based on a Fourier series representation of the solution in the exterior region.

In the time-harmonic case, Dirichlet-to-Neumann (DtN) maps yield exact nonreflecting conditions and thus avoid spurious reflections from B. They are explicitly known for various equations or geometries [4–8]. Once combined with a finite difference or finite element discretization inside Ω, they lead to a highly accurate and efficient numerical scheme. Here we extend the DtN approach to multiple scattering problems, where every scatterer is enclosed by a separate artificial boundary B_j [9]. Thus Ω consists of multiple disjoint components, Ω_j. We derive an exact DtN boundary condition on B, the disjoint union of all B_j, by combining multiple contributions from purely outgoing wave fields. We present theoretical results that show existence and uniqueness of the solution to the boundary value problem in Ω, as well as numerical results that demonstrate the accuracy and efficiency of our method.

In the time-dependent case, *exact nonreflecting boundary conditions* have been derived for the wave equation [10, 11] and Maxwell's equations [12]. These boundary conditions are *local in time* and involve only first derivatives of the solution. Therefore, they are easy to use with standard finite difference or finite element methods. As the accurate simulation of waves at high frequencies or the detailed representation of small scale geometric features requires the use of adaptive mesh strategies, explicit time integrators become prohibitively expensive because of the stringent CFL condition. Instead, implicit methods, such as Crank-Nicolson, are typically used, yet they require the solution of a large linear system of equations at every time step. Due to the nonreflecting boundary condition, this linear system is no longer symmetric, unlike the situation in bounded domains. However, it is possible to reformulate the discretized equations by decoupling the additional unknowns needed on the artificial boundary from the interior unknowns [13]. As a consequence the symmetry and positive definiteness of the linear system are restored, while the additional computational effort due to the nonreflecting boundary condition becomes negligible.

For time-dependent multiple scattering problems the use of a single artificial boundary surrounding all scatterers involved also becomes prohibitively expensive in memory requirement. Instead, it is judicious to enclose each scatterer within a single separate computational domain. Clearly waves that leave a certain domain,

Ω_1, will impinge upon a different domain, Ω_2, at later times; hence they are no longer purely outgoing waves. To transfer the time-retarded information from Ω_1 to Ω_2 an analytical representation of the solution in the unbounded medium becomes necessary. Again, such an analytical representation [14] is inherent to the exact nonreflecting boundary conditions described above.

References

[1] B. Engquist, A. Majda, *Absorbing boundary conditions for the numerical simulation of waves*, Math. Comp. 31 (1977), pp. 629–651.

[2] A. Bayliss, M. Gunzburger, E. Turkel, *Boundary conditions for the numerical solution of elliptic equations in exterior regions*, SIAM J. Appl. Math. 42 (1982), pp. 430–451.

[3] J. P. Bérenger, *A perfectly matched layer for the absorption of electromagnetic waves*, J. Comput. Phys. 114 (1994), pp. 185–200.

[4] J.B. Keller and D. Givoli, *Exact non-reflecting boundary conditions*, J. Comp. Phys. 82 (1989), pp. 172–192.

[5] D. Givoli, *Numerical Methods for Problems in Infinite Domains*, Elsevier, 1992.

[6] M.J. Grote and J.B. Keller, *On nonreflecting boundary conditions*, J. Comp. Phys. 122 (1995), pp. 231–243.

[7] D. Givoli, *Recent advances in the DtN-FE method*, Archives of Comput. Meth. Engin. 6 (1999), pp. 71–116.

[8] G.K. Gächter and M.J. Grote, *Dirichlet-to-Neumann map for three-dimensional elastic waves*, Wave Motion 37 (2003), pp. 293–311.

[9] M. J. Grote and C. Kirsch, *Dirichlet-to-Neumann boundary conditions for multiple scattering problems*, J. Comp. Phys, submitted. See preprint 2004-04 at `http://www.math.unibas.ch/preprints`.

[10] M. J. Grote and J. B. Keller, *Exact nonreflecting boundary conditions for the time dependent wave equation*, SIAM J. Appl. Math. 55 (1995), pp. 280–297.

[11] M. J. Grote and J. B. Keller, *Nonreflecting boundary conditions for time dependent scattering*, J. Comput. Phys. 127 (1996), pp. 52–65.

[12] M. J. Grote and J. B. Keller, *Nonreflecting boundary conditions for Maxwell's Equations*, J. Comput. Phys. 139 (1998), pp. 327–324.

[13] W. Bangerth, M. J. Grote, and C. Hohenegger, *Finite element method for time dependent scattering: nonreflecting boundary condition, adaptivity, and energy decay*, Comp. Meth. Appl. Mech. Eng., in press. See preprint 2003-06 at `http://www.math.unibas.ch/preprints`.

[14] M. J. Grote and C. Kirsch, *Far-field evaluation via nonreflecting boundary conditions.* In Proc. 9th Internat. Conf. on Hyperbolic Problems (Hyp2002), Eds. Th. Y. Hou, E. Tadmor, Springer, pp. 195–204, 2003.

Nodal and Edge Finite Element Discretization of Maxwell's Equations

Manfred Kaltenbacher, University of Erlangen, Germany
manfred@lse.eei.uni-erlangen.de
(joint work with Barbara Kaltenbacher and Stefan Reitzinger)

The numerical computation of electromagnetic fields is performed for more then 20 years. For the domain discretization nodal as well as edge finite elements have been used successfully. Nevertheless, in the last years inaccurate results at material parameter interfaces in the magnetostatic as well as in the eddy current case, and, spurious modes in Maxwell's eigenvalue problems have been reported. In this paper we will concentrate on the problems related with material parameter interfaces, where the magnetic reluctivity changes its value abruptly. We will describe a simple to implement method following the ideas reported in [3], which produces correct results. For the high frequency case we refer to [3].

The electromagnetic field is fully described by Maxwell's equations [9]. Restricting the problem class to the quasi-static (eddy current) case, we arrive at the following partial differential equation for the magnetic vector potential $\mathbf{A}$

$$\gamma \frac{\partial \mathbf{A}}{\partial t} + \nabla \times \nu \nabla \times \mathbf{A} = \mathbf{J}_{\mathrm{i}} \tag{1}$$

with boundary condition $\mathbf{n} \times \mathbf{A} = 0$ and $\mathbf{n}$ the unit outward normal vector. In (1) $\mathbf{J}_{\mathrm{i}}$ denotes the impressed current density, ν the magnetic reluctivity and γ the electric conductivity. Furtheron, the following interface conditions have to be fulfilled

$$[\mathbf{A} \times \mathbf{n}] = 0\ ; \quad [\nu\ \mathbf{n} \times \nabla \times \mathbf{A}] = 0\ ; \quad \left[\gamma \frac{\partial \mathbf{A}}{\partial t}\right] = 0 \tag{2}$$

with $[\mathbf{Z}] = \mathbf{Z}_{\mathrm{right}} - \mathbf{Z}_{\mathrm{left}}$. For further discussions let Ω be a bounded single connected convex domain with boundary $\partial\Omega = \Gamma$. Therewith, the variational formulation for (1) in the function space

$$\mathbf{H}_0^{\Sigma}(curl) = \{\mathbf{u} \in (L^2(\Omega))^3 \,|\nabla \times\ \mathbf{u} \in (L^2(\Omega))^3,\ \mathbf{u} \times \mathbf{n}|_{\Gamma} = 0,\ [\mathbf{n} \times \mathbf{u}]|_{\Sigma} = 0\} \tag{3}$$

reads as follows: Find $\mathbf{A} \in \mathbf{H}_0^{\Sigma}(curl)$ such that

$$\int_{\Omega} \gamma \mathbf{A}' \cdot \frac{\partial \mathbf{A}}{\partial t}\, d\Omega \quad + \quad \int_{\Omega} \nabla \times \mathbf{A}' \cdot \nu\, \nabla \times \mathbf{A}\ d\Omega = \int_{\Omega} \mathbf{A}' \cdot \mathbf{J}_{\mathrm{i}}\, d\Omega \tag{4}$$

for any $\mathbf{A}' \in \mathbf{H}_0^{\Sigma}(curl)$ is fulfilled.

It is well known, that an edge FE-discretization of (4) is $\mathbf{H}_0(curl)$-conform [6]. Nevertheless, the solution of the algebraic system requires special care in order to obtain an optimal multigrid solver (see e.g. [2], [5]). We suggest to add a fictive electric conductivity γ' to regions with zero electric conductivity to obtain a variational form, which is elliptic [8]. Of course, this fictive conductivity γ' has to be chosen small as compared to the reluctivity of the material. The proof of convergence even in the case of $\gamma' \to 0$ is given in [7].

For the application of nodal finite elements, we have to perform additional steps. According to [4] as well as [1] we decompose the magnetic vector potential $\mathbf{A}$ by

$$\mathbf{A} = \mathbf{w} + \nabla\phi\,, \quad \nabla \cdot \mathbf{w} = 0\,, \tag{5}$$

with $(\mathbf{w}, \phi) \in \big((H_{\mathrm{T}}^1(\Omega))^3, H_0^1(\Omega)\big)$ and Ω being a convex domain. The same decomposition is done for the test function $\mathbf{A}' = \mathbf{v} + \nabla\psi$. Since we have to guarantee $\nabla \cdot \mathbf{w} = 0$, we do so by adding the penalty term $\int_{\Omega} \nu\ (\nabla \cdot \mathbf{v}\, \nabla \cdot \mathbf{w})\ d\Omega$ to the variational formulation. Therewith, the variational formulation can be stated as follows: Find $(\mathbf{w}, \phi) \in ((H^1_{\mathrm{T}}(\Omega))^3,\, H^1_0(\Omega))$ such that

$$\begin{aligned}&\int\limits_{\Omega} \nu \nabla \times \mathbf{v} \cdot \nabla \times \mathbf{w}\; d\Omega + \int\limits_{\Omega} \nu\; \nabla \cdot \mathbf{v}\, \nabla \cdot \mathbf{w}\, d\Omega \\ &+ \int\limits_{\Omega} \gamma(\mathbf{v} + \nabla\psi) \cdot \frac{\partial}{\partial t}(\mathbf{w} + \nabla\phi)\, d\Omega = \int\limits_{\Omega} \mathbf{J}_{\mathrm{i}} \cdot \mathbf{v}\, d\Omega\,. \end{aligned} \tag{6}$$

for any $(\mathbf{v},\, \psi) \in ((H^1_{\mathrm{T}}(\Omega)^3,\, H^1_0(\Omega))$. Now, since for most practical eddy current problems the domain Ω is convex, the discretization of the above variational formulation with nodal finite elements will result in correct results. However, the question arises, if a domain Ω including subdomains of different material parameters (magnetic reluctivity or/and electric conductivity), is really convex? Let us consider the case of a ferromagnetic cube embedded in air (see Fig. 1). Assuming

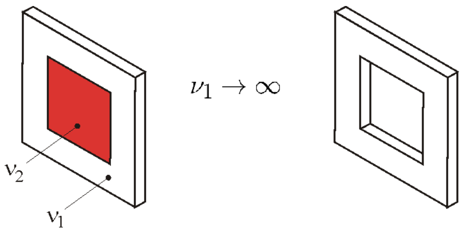


FIGURE 1. Ferromagnetic cube in air

the case $\nu_1 \to \infty$ (of course the limit of ν_1 is equal to $1/\mu_0$ with μ_0 being the permeability in vacuum), we arrive at a non-convex domain. Now according to [3], it is known, that for non-convex domains the discretization with nodal finite elements produces wrong solutions due to the non-density of smooth fields. In [3] the authors could proof, that by introducing a special weighting function inside the divergence integral, nodal finite elements can yet be used for the approximation. Therewith, the second term in the variational formulation (6) has to be changed to

$$\int\limits_{\Omega} \nu\; \nabla \cdot \mathbf{v}\, \nabla \cdot \mathbf{w}\, d\Omega \to \int\limits_{\Omega} \nu\; s\; \nabla \cdot \mathbf{v}\, \nabla \cdot \mathbf{w}\, d\Omega \tag{7}$$

with

$$s = \prod_{a \in \mathsf{Q}} r_a^{\alpha}\,. \tag{8}$$

In (8) Q denotes the set of all reentrant corners, r_a the distance to each reentrant corner, and α an exponent. We have implemented this idea in a simple way by setting the weighting function s to zero for finite elements near each interface of two subdomains with different material parameters.

The correctness of the weighted variational formulation have been demonstrated by numerical test cases (iron cube, thin iron plate) as well as industrial applications (electric power transformer, electromagnetic motor, magnetic resonance imaging scanner).

References

[1] C. Amrouche, C. Bernardi, M. Duage, and V. Girault, *Vector Potentials in Three-dimensional Non-smooth Domains*, Math. Meth. in Appl. Science, 21 (1998), pp. 823–864.

[2] D. Arnold, R. Falk, and R. Winther, *Multigrid in H(div) and H(curl)*, Numer. Math., 85 (2000), pp. 197–218.

[3] M. Costabel and M. Dauge, *Weighted regularization of Maxwell equations in polyhedral domains*, Numer. Math., 93 (2002), pp. 239–277.

[4] V. Girault and P.-A. Raviart, *Finite Element Approximation of the Navier-Stokes Equations*, Springer Verlag, Berlin Heidelberg New York, 1979.

[5] R. Hiptmair, *Multigrid method for Maxwell's equations*, SIAM J. Numer. Anal., 36 (1999), pp. 204–225.

[6] J.C. Nédélec, *A new family of mixed finite elements in R^3*, Numer. Math., 50 (1986), pp. 57–81.

[7] S. Reitzinger and J. Schöberl, *An Algebraic Multigrid Method for Finite Element Discretizations with Edge Elements*, Numer. Linear Algebra Appl., 31 (2002), pp. 223 – 238.

[8] M. Schinnerl, J. Schöberl, M. Kaltenbacher, and R. Lerch, *Multigrid Methods for the 3D Simulation of Nonlinear Magneto-Mechanical Systems*, IEEE Transactions on Magnetics, 38 (2002), pp. 1497–1511.

[9] J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, Inc., 1941.

The Hurewicz Map Distinguishes Intuitive vs. Computable Topological Aspects of Computational Electromagnetics

Robert Kotiuga

1. Abstract of talk

Answers to intuitive topological problems, such as checking if a space is contractible, are easily characterized in terms of homotopy groups. However, in four or more dimensions, such a characterization is provably computationally intractable. On the other hand, cohomology theory may not be intuitive, but it does provide a formal connection between Maxwell's equations and the lumped parameters occurring in Kirchhoff's laws. Furthermore, by exploiting sparse matrix algorithms and the Smith normal form, cohomological information is efficiently extracted from the data structures used in finite element analysis. A natural question then arises: Do engineers need to go beyond the linear algebra and sparse matrix techniques associated with homology calculations? It turns out that there are inverse problems involving near force-free magnetic fields where the conjectured characterization of the space of solutions, involves computationally intractable topological invariants such as the Thurston norm [4]. For this reason, it is imperative to investigate algebraic structures found in the data structures of finite element analysis, and

which yield topological insights not deducible from cohomological considerations alone.

The Hurewicz map takes representatives of generators of homotopy groups to their homology classes and is a well-defined map from homotopy groups to homology groups. In this sense, it provides a natural framework for comparing the intuitive but intractable aspects of homotopy theory with the computable but less intuitive aspects of homology theory. In particular, thorugh the use of the Hurewicz map, several important identifications can be made:

(1) The lower central series of the fundamental group is related to certain Massey products in the cohomology ring.
(2) The differential graded Lie algebras of rational homotopy theory are related to the minimal models of the cohomology ring.
(3) By Hopf's theorem, the cokernel of the second homology group under the Hurewicz map is characterized in terms of a presentation of the fundamental group.

The workshop talk concretely developed the relevance of these aspects of the Hurewicz map in the context of computational electromagnetics.

2. Putting my talk in the context of my previous work

Though originally developed as a natural outgrowth of multivariable calculus, algebraic topology and differential forms have become an essential tool used to formulate many basic laws of physics. Through my research this area of mathematics has found a natural application to many areas of electrical engineering and computational electromagnetics. A strong theme in my research is the identification of geometric and topological aspects, which shed light on dimensional dependence in the complexity of engineering problems and their algorithmic solution. This should be evident from the other publications I have selected to list below [7]-[15]. Much of my earlier work dealing with finite element analysis of electromagnetic fields and magnetic scalar potentials is summarized in the MSRI monograph coauthored with my Ph. D. student, Paul Gross [2].

If one were to seek a more mainstream characterization of my research interests, I could probably describe them in terms of the research interests listed on my resume:

- Electromagnetics;
- Numerical methods for 3-d vector fields;
- Whitney forms, the finite element method and the analysis of algorithms,
- Cuts for magnetic scalar potentials, formulation of eddy-current problems,
- Variational and symplectic techniques,
- Micromagnetics; nanoscale magnetics,
- Geometric inverse problems,
- Helicity functionals and near force-free magnetic fields; contact geometry,

My most recent research deals with how electromagnetic force constraints give rise to topological structures necessarily characterized by nonabelian algebraic structures [1], [4]. This reseach is interesting both in terms of applications, and in defining the data structures which are useful for the finite element analysis of electromagnetic fields. The abstract of my workshop presentation above, is an attempt to get a handle on the latter aspects.

References

[1] P.R. Kotiuga, *Topology-Based Inequalities and Inverse Problems for Near Force-Free Magnetic Field*, IEEE Trans. MAG., To appear March 2004.

[2] P.W. Gross and P.R. Kotiuga, *Electromagnetic Theory and Computation: A Topological Approach*, MSRI Monograph series # 48; Cambridge: University Press, 2004. ISBN # 0521801605

[3] S. Suuriniemi, L. Kettunen, P.R. Kotiuga, *Techniques for Systematic Treatment of Certain Coupled Problems*, IEEE Trans. MAG-39, (3), May 2003, pp. 1737–1740.

[4] J.C. Crager and P.R. Kotiuga, *Cuts for the Magnetic Scalar Potential in Knotted Geometries and Force-Free Magnetic Fields*, IEEE Trans. MAG-38, (2), Mar.2002, pp. 1309–1312.

[5] P.W. Gross and P.R. Kotiuga, *Data Structures for Geometric and Topological Aspects of Finite Element Algorithms*, Geometric Methods for Computational Electromagnetics, Teixeira, F.L. editor, Progress in Electromagnetics Res., 32, Kong, J.A. chief editor, EMW Publishing, Cambridge MA, 2001, pp. 151–169. (ISBN 09668143-6-3)

[6] P.W. Gross and P.R. Kotiuga, *Finite Element-based Algorithms to Make Cuts for Magnetic Scalar Potentials: Topological Constraints and Computational Complexity.* i.b.i.d., pp. 207–245.

[7] P.R. Kotiuga, *Metric Dependent Aspects of Inverse Problems and Functionals Based Helicity*, Journal of Applied Physics, 70(10), May 1993, pp. 5437–5439.

[8] P.R. Kotiuga, *Essential Arithmetic for Evaluating Three Dimensional Vector Finite Element Interpolation Schemes*, IEEE Transactions on Magnetics, MAG-27, (6), November, 1991, pp. 5208–5210.

[9] P.R. Kotiuga and R.C. Giles, *A Topological Invariant for the Accessibility Problem of Micromagnetics, Journal of Applied Physics*, 67(9), May 1990, pp. 5347–5349.

[10] P.R. Kotiuga, *Topological Duality in Three-Dimensional Eddy-Current Problems and its Role in Computer-Aided Problem Formulation*, Journal of Applied Physics, 67(9), May 1990, pp. 4717–4719.

[11] P.R. Kotiuga, *An Algorithm to Make Cuts for Magnetic Scalar Potentials in Tetrahedral Meshes Based on the Finite Element Method*, IEEE Transactions on Magnetics, MAG-25, (5), September 1989, pp. 4129–4131.

[12] P.R. Kotiuga, *Helicity Functionals and Metric Invariance in Three Dimensions*, IEEE Transactions on Magnetics, MAG-25, (4), July 1989, pp. 2813–2815.

[13] P.R. Kotiuga, *Topological Considerations in Coupling Magnetic Scalar Potentials to Stream Functions Describing Surface Currents*, IEEE Trans. MAG-25, (4), July 1989, pp. 2925–2927.

[14] P.R. Kotiuga, *Variational Principles for Three-Dimensional Magnetostatics Based on Helicity*, Journal of Applied Physics, 63(8), April 1988, pp.3360–3362

[15] . P.R. Kotiuga, *On Making Cuts for Magnetic Scalar Potentials in Multiply Connected Regions*, Journal of Applied Physics, 61(8), April 1987, pp. 3916–3918.

Inverse Obstacle Scattering for Time-Harmonic Electromagnetic Waves

Rainer Kress

This presentation provides a survey on some recent developments in the theory and numerical solution of time-harmonic inverse scattering problems. Roughly speaking, one can distinguish two groups of inverse problems in this field, namely the inverse medium problem and the inverse obstacle problem. For time reasons, only inverse obstacle scattering will be covered. However, most of the ideas that are presented for inverse obstacle scattering have counter parts in inverse medium scattering. After formulating the inverse problem, the issue of uniqueness, that is, identifyability will be addressed. The uniqueness question is of its own mathematical interest and also interrelates with some of the more recently developed reconstruction algorithms. By considering one or two of its representatives the basic ideas of three groups of methods will be outlined, namely decomposition methods, iterative methods and sampling and probe methods. For illustration a couple of numerical examples will be included.

Consider the scattering of a time-harmonic electromagnetic plane wave E^i, H^i from an impenetrable scatterer described by a bounded domain D in $\Re^3$ either with a perfect conductor or an impedance boundary condition. The inverse obstacle scattering problem consists of finding the shape and location of D from the knowledge of the electric far field pattern E_∞ of the scattered wave E^s, H^s for one or several incident plane waves. The corresponding uniqueness result due to Kirsch and Kress [17] (see also [6]) confirms that the domain D and the boundary condition are uniquely determined by the far field pattern for infinitely many incident plane waves. The main idea of the proof is to exploit the fact that for scattering of electric dipole fields the scattered wave develops singularities when the source and observation points approach the boundary. Uniqueness for one incident plane wave remains a challenging open problem. Partial results were recently obtained for scattering from balls [20] and polyhedral scatterers [1].

Decomposition methods, in principle, separate the inverse problem into an ill-posed linear problem to reconstruct the scattered wave E^s, H^s from its far field pattern E_∞ and a nonlinear problem for the subsequent determination of the boundary ∂D of the scatterer from the boundary condition. These methods do not require the solution of the forward problem and some of them perform well without a priori information on the geometry of the obstacle. A typical representative of this approach is the potential method of Kirsch and Kress (see [9, 16]).

Iteration methods interpret the inverse obstacle scattering problem as a nonlinear ill-posed operator equation $A(\partial D) = E_\infty$ and apply iterative schemes such as regularized Newton type, Landweber or conjugate gradient methods for its solution. Here, A denotes the operator that, for a fixed incident field, maps the boundary ∂D of the scatterer onto the far field pattern of the scattered wave. The theoretical foundation for this approach requires to establish the differentiability of the operator A with respect to the boundary and to explicitly characterize the derivative. For the perfect conductor boundary condition this was done by Potthast [22] via integral equation methods and by Kress [19] via a factorization formula. The latter method was recently extended to the impedance boundary condition by Haddar and Kress [10].

For details on the numerical implementation, among others, see [6, 8, 11, 14, 18]. The numerical examples provide amble evidence that iterative methods, in particular Newton iterations, yield very good reconstructions. However, they require the solution of the corresponding forward problem in each iteration step and a priori information on the geometry of the obstacle. Furthermore, although progress has been made through the work of Hohage [12] and Potthast [24], the convergence issue is not yet satisfactorily settled. A hybrid of Newton type iterations and decomposition methods was suggested in [21] and successfully tested for two-dimensional examples.

The main idea of the more recently developed so-called *sampling and probe methods* is to develop a criterium in terms of the behaviour of some ill-posed linear integral equation that decides on whether a point z lies inside or outside the scatterer D. Then the criterium is evaluated numerically for a grid of points to visualize the unknown scatterer. As opposed to the two previous types of methods that, in principle, only need the far field pattern for one incident direction, the sampling and probe methods need the far field pattern for all incident and observation directions and polarizations. However, as their main advantage they perform extremely well without any a priori information on the geometry. The linear sampling method as developed in acoustic scattering by Colton and Kirsch [5] has as its central piece the far field operator $F : L^2_t(\Omega) \to L^2_t(\Omega)$ on the space of tangential L^2 fields on the unit sphere Ω. This operator is defined as an integral operator with the kernel given by the far field pattern $E_\infty(\hat{x}, d)$ for all observation directions $\hat{x} \in \Omega$ and all incident directions d. With the explicitly available far field pattern $E^i_{\infty,dip}(\cdot\,, z)p$ of the field of an electric dipole with polarization p located at the point z the linear sampling method is based on the ill-posed linear integral equation $Fg(\cdot\,, z) = E^i_{\infty,dip}(\cdot\,, z)p$. Although, this integral equation, in general, is not solvable, it can be approximately solved in the sense that for every $p \in \Re^3$, $\varepsilon > 0$, and $z \in D$ there exists $g(\cdot\,, z) \in L^2_t(\Omega)$ such that $\|Fg(\cdot\,, z) - E^i_{\infty,dip}(\cdot\,, z)p\|_{L^2(\Omega)} \le \varepsilon$ and $\|g(\cdot\,, z)\|_{L^2(\Omega)} \to \infty$ as $z \to \partial D$. In the numerical implementation the far field integral equation is solved by Tikhonov regularization via Morozov's discrepancy principle and then ∂D is visualized through the points z where $\|g(\cdot\,, z)\|_{L^2(\Omega)}$ becomes large. For details on the theoretical foundation and numerical examples see [3, 4, 19].

A remaining gap in the theoretical foundation of the linear sampling method, namely, the question why the implementation via Tikhonov and Morozov actually picks the approximation g that is predicted by the above theoretical result was closed in acoustics through a recent contribution by Arens [2]. However, the gap remains open in electromagnetics, since Aren's analysis does not yet cover this case.

The factorization method may be considered as a variation of the linear sampling method in the sense that it replaces F in the far field equation by $(F^*F)^{1/4}$, that is, it is based on the equation $(F^*F)^{1/4}g(\cdot\,,z) = E^i_{\infty,dip}(\cdot\,,z)p$. This equation is more satisfying since it is to be expected that it is solvable if and only if $z \in D$. The corresponding result in acoustics is valid as shown in a pioneering paper by Kirsch [15]. However it is open for electromagnetics. The numerical implementation of the factorization is similar to that of the linear sampling method. The procedure is known as factorization method, since it relies on a factorization of the far field operator.

The linear sampling method and the factorization method may be viewed as dual to the uniqueness proof of Kirsch and Kress, since, in principle, their foundation is based on letting source points of electric dipole fields approach the boundary from inside of D whereas in the uniqueness proof the sourec points approch the boundary from outside of D. The latter idea is mimiced in the point source and singular source methods of Potthast [7, 23] and the probe method of Ikehata [13].

References

[1] G. Alessandrini and L. Rondi *Determing a sound-soft polyhedral scatterer by a single far-field measurement.* SIAM J. Appl. Math. (to appear).

[2] T. Arens. *Why linear sampling works. Inverse Problems* 20, (2004), pp163–173.

[3] F. Cakoni, D. Colton and P. Monk *The electromagnetic inverse scattering problem for partially coated Lipschitz domains,* Proc. Royal Soc. Edinburgh (to appear).

[4] D. Colton, H. Haddar and P. Monk *The linear sampling method for solving the electromagnetic inverse scattering problem,* SIAM J. Scientific Computation 24, (2002) pp719–731.

[5] D. Colton, and A. Kirsch *A simple method for solving inverse scattering problems in the resonance region*, Inverse Problems 12 (1996), pp383–393.

[6] D. Colton and R. Kress *Inverse Acoustic and Electromagnetic Scattering Theory.* 2nd. ed. Springer, Berlin 1998.

[7] K. Erhard, and R. Potthast A numerical study of the probe method. (to appear).

[8] C. Farhat, R. Tezaur and Djellouli *On the solution of three-dimensional inverse obstacle acoustic scattering problems by a regularized Newton method.* Inverse Problems 18, (2002) pp1229–1246.

[9] M. Haas, W. Rieger, W. Rucker and G. Lehner *Inverse 3D acoustic and electromagnetic obstacle scattering by iterative adaption.* In: Inverse Problems of Wave Propagation and Diffraction (Chavent and Sabatier, eds). Springer-Verlag, Heidelberg 1997.

[10] H. Haddar. and R. Kress. *On the Frechet derivative for obstacle scattering with an impedance boundary condition.* SIAM J. Appl. Math. (to appear).

[11] T. Hohage. *Logarithmic convergence rates of the iteratively regularized Gauss-Newton method for an inverse potential and an inverse scattering problem.* Inverse Problems 13, (1997), pp1279–1299.

[12] T. Hohage. *Iterative Methods in Inverse Obstacle Scattering: Regularization Theory of Linear and Nonlinear Exponentially Ill-Posed Problems.* Dissertation, Linz 1999.

[13] M. Ikehat. *Reconstruction of an obstacle from the scattering amplitude at a fixed frequency.* Inverse Problems 14, (1998), pp949–954.

[14] A. Kirsch. *The domain derivative and two applications in inverse scattering theory.* Inverse Problems 9, (1993) pp81–96.

[15] A. Kirsch *Characterization of the shape of a scattering obstacle using the spectral data of the far field operator.* Inverse Problems 14, (1998), pp1489–151.

[16] A. Kirsch and R. Kress *An optimization method in inverse acoustic scattering.* In: Boundary elements IX, Vol 3. Fluid Flow and Potential Applications (Brebbia et al, eds.) pp. 3–18, Springer-Verlag, Berlin Heidelberg New York, 1987.

[17] A. Kirsch and R. Kress. *Uniqueness in inverse obstacle scattering.* Inverse Problems 9, (1993) pp285–299.

[18] R. Kress. *Integral equation methods in inverse acoustic and electromagnetic scattering.* In: *Boundary Integral Formulations for Inverse Analysis,* (Ingham, Wrobel, eds.), pp 67–92, Computational Mechanics Publications, Southampton 1997.

[19] R. Kress *Electromagnetic waves scattering: Scattering by obstacles.* In: *Scattering* (Pike, Sabatier, eds.) pp. 191-210, Academic Press, London, 2001.

[20] R. Kress. *Uniqueness in inverse obstacle scattering for electromagnetic waves.* Proceedings of the URSI General Assembly, Maastricht 2002.

[21] R. Kress. *Newton's method for inverse obstacle scattering meets the method of least squares.* Inverse Problems 19, (2003) pp91-104.

[22] R. Potthast. *Domain derivatives in electromagnetic scattering.* Math. Meth. in the Appl. Sci. 19, (1996) pp1157–1175.

[23] R. Potthast.. *A point-source method for inverse acoustic and electromagnetic obstacle scattering problems.* IMA Journal of Appl. Math. 61, (1998), pp119–140.

[24] R. Potthast. *On the convergence of a new Newton-type method in inverse scattering.* Inverse Problems 17, (2001) pp1419–1434.

A New View on Collocation

S. Kurz

(joint work with O. Rain, V. Rischmüller, S. Rjasanow)

In recent years, a remarkable amount of papers has been published that treat continuous and discrete electromagnetics in terms of differential forms (DFs). For a good account on this topic, see, e.g., [2] and [7]. However, most of these papers focus on (generalised) finite difference and finite element methods. There are only rare papers that deal with the boundary element method [1, 3, 6, 11].

The aim of this talk is to show how the integral equations of electromagnetics can be expressed in the language of DFs. The integral kernels become double forms [5]. These are DFs in one space with coefficients that are DFs in another space, or DF-valued DFs [12]. We restrict ourselves to the static case. Similar schemes can be derived for time dependent problems. The formulation in terms of DFs enables a uniform treatment of electrostatics (Kirchhoff representation formula) and magnetostatics (Stratton-Chu representation formula).

Since DFs possess discrete counterparts, known as Whitney forms, such schemes lend themselves naturally to discretisation. As an example, a boundary integral equation for the double curl operator is considered. This equation has been investigated in a variational setting in [8]. A detailed discussion of the Sobolev spaces being involved can be found in
[4, 10].

In the present contribution we wish to highlight an alternative approach. The proposed discretisation scheme generalises the well-known collocation technique by using de Rham maps on dual grid systems [6, 11]. Depending on the integral operator to be discretised, the 1-form valued residual is forced to be zero either over the 1-chains of the primal or the dual grid. The viability of the method will be demonstrated by means of a numerical example, where a sphere is immersed in the field of a circular current loop.

For an extended version of this contribution see [9].

References

[1] A. Bossavit, *The 'scalar' Poincaré Steklov operator and the 'vector' one: Algebraic structures which underlie their duality*, in Domain Decomposition Methods for Partial Differential Equations, R. G. et al., ed., SIAM, 1991, ch. 2, pp. 19–26.

[2] ———, *Generalized finite differences in computational electromagnetics*, in Geometric Methods for Computational Electromagnetics, F. Teixeira, ed., vol. 32 of Progress in Electromagnetics Research, EMW Publishing, Cambridge, Massachusetts, 2001, pp. 45–64.

[3] ———, *On the representation of differential forms by potentials in dimension 3*, in Scientific Computing in Electrical Engineering, U. van Rienen, M. Günther, and D. Hecht, eds., vol. 18 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin, 2001, pp. 97–104.

[4] A. Buffa and P. Ciarlet, Jr., *On traces for functional spaces related to Maxwell's equations Part I: An integration by parts formula in Lipschitz polyhedra, Part II: Hodge decompositions on the boundary of Lipschitz polyhedra and applications*, Mathematical Methods in the Applied Sciences, 24 (2001), pp. 9–48.

[5] G. de Rham, *Differentiable Manifolds*, Springer-Verlag, Berlin, 1984.

[6] C. Geuzaine, T. Tarhasaari, L. Kettunen, and P. Dular, *Discretization schemes for hybrid methods*, IEEE Transactions on Magnetics, 37 (2001), pp. 3112–3115.

[7] R. Hiptmair, *Finite elements in computational electromagnetism*, Acta Numerica, (2002), pp. 237–339.

[8] ———, *Symmetric coupling for eddy current problems*, SIAM J. Numer. Anal., 40 (2002), pp. 41–65.

[9] S. Kurz, O. Rain, V. Rischmüller, and S. Rjasanow, *Discretization of boundary integral equations by differential forms on dual grids*, IEEE Transactions on Magnetics, 40 (2004).

[10] W. McLean, *Strongly Elliptic Systems and Boundary Integral Equations*, Cambridge University Press, Cambridge, UK, 2000.

[11] T. Tarhasaari, L. Kettunen, and C. Geuzaine, *Discretization of sources of integral operators*, IEEE Transactions on Magnetics, 36 (2000), pp. 659–662.

[12] K. Warnick and D. Arnold, *Electromagnetic Green functions using differential forms*, Journal of Electromagnetic Waves and Applications, 10 (1996), pp. 427–438.

Computation of Maxwell Eigenvalues with Exponential Rates of Convergence

Paul Ledger

Our interest in this work lies in the accurate calculation of Maxwell eigenvalues for closed cavities. The results are important for many applications such as the design of microwave devices and charged particle accelerators. The solution of such problems remains far from trivial due to the fact that realistic cavities often contain multi–materials, have small scale feature and contain many sharp corners, which all give rise to highly singular eigenfunctions.

Using a nodal finite element basis for each component of the electric field is known to be inappropriate, as the resulting solution is polluted by spurious modes. Instead, we choose to solve these problems using the $\vec{H}(\mathrm{curl})$ conforming finite elements that were first introduced by Nédélec [1]. Using such elements is known to overcome the problems of spurious modes and allow the easy incorporation of material interfaces and boundary conditions.

We follow a finite element approach which allows for arbitrary increases in polynomial order p. In particular we use the recent hierarchic basis of Ainsworth and Coyle [2, 3] with both p and h (mesh) refinements. Indeed, when the h and p refinements are correctly combined, we are able to observe the theoretically predicted exponential rates of convergence for the Maxwell eigenvalues. Numerical examples show that the exponential rates of convergence can be obtained in practice for a series of benchmark problems discretised with tetrahedral meshes in three–dimensions [4, 5].

Recent extensions include the application of *hp* finite elements to axisymmetric problems with rotational symmetry [6]. For such cases it is possible to reduce a three-dimensional problem to a sequence of two–dimensional problems. Again, exponential rates of convergence have been observed for the computed eigenvalues of closed cavities.

References

[1] J.C. Nédélec *Mixed elements in* $\mathbb{R}^3$, Numerische Mathematik, 35 (1980), pp315–341.

[2] M. Ainsworth and J. Coyle. *Hierarchic hp–edge element families for Maxwell's equations on hybrid quadrilateral/triangular meshes.* Computer Methods in Applied Mechanics in Engineering. 190 (2001), pp6709-6733.

[3] M. Ainsworth and J. Coyle. *Hierarchic finite element bases on unstructured tetrahedral meshes.* International Journal for Numerical Methods in Engineering. 58 (2003) pp2103-2130.

[4] M. Ainsworth, J. Coyle, P.D. Ledger and K. Morgan *Computation of Maxwell eigenvalues using higher order edge elements in three-dimensions.* IEEE Transactions on Magnetics 39 (2003), pp2149–2153.

[5] J. Coyle, P.D. Ledger. *Evidence of exponential convergence in the computation of Maxwell eigenvalues.* Accepted for publication in Computer Methods in Applied Mechanics and Engineering. (2004). Preprint available http://www.sam.math.ethz.ch/~ledger/publication.html

[6] R. Hiptmair, P.D. Ledger. *Computation of resonant modes for axisymmetric cavities using hp version finite elements.* Technical Report, Seminar for Applied Mathematics ETH Zürich (2003). Preprint available http://www.sam.math.ethz.ch/reports/2003

Computational Shape and Topology Optimization with Applications to 3–Dimensional Magnetostatics

D. Lukáš

(joint work with U. Langer, E. Lindner, R. Stainko, J. Pištora)

In the talk we mainly discussed computational aspects of shape and topology optimization governed with 3–dimensional linear and nonlinear magnetostatics, respectively. This is covered in the speaker's thesis [2] and in [3]. The acknowledgment is due to the Special Research Initiative SFB F013 "Numerical and symbolic scientific computing", subproject "Multilevel solvers for large scale discretized optimization problems" at the University of Linz, Austria. The speaker especially thanks to Dr. Joachim Schöberl for his kind software support during the week in Oberwolfach.

The presentation started with a motivation from physics. We described electromagnets that are used for measurements of magnetooptic effects on thin layers. We aim at designing their optimal topology and shape so that in the area where the measurements take place the magnetic field is as constant as possible and above a prescribed magnitude. Throughout the presentation we instantiate the ideas for this application.

Next, we recalled an abstract optimal shape design problem, its finite element approximation and we discussed the existence and convergence issues following the theory in [1], which is based on the compactness and continuity arguments. We optimize the interface between the air and ferromagnetics, rather than the boundary of the computational domain as usual in mechanics. We pointed out a drawback that on fine discretizations the non–design grid nodes cannot follow large perturbations of the design shape. The mapping from the shape to the grid nodes is carried over an artificial linear elasticity problem with the prescribed displacements along the design shape interface. Then, we presented the algebraic approach to the shape sensitivity analysis and its efficient software implementation, see [5]. The user is only supposed to dessignate the shape and to code the objective in terms of the state solution. The underlying finite element code provides the sensitivity of element contributions to the bilinear form with respect to the grid displacements. The optimization package is now to be included into the NgSolve, see [7].

Further, we presented numerical results for both 2– and 3–dimensional shape optimization problems. After the 2d optimized design the electromagnets were manufactured and the measurements of the magnetic field showed the 4.5–times improvements in terms of the objective functional, compared to the initial design.

We presented a multilevel optimization approach. Here, hierarchies of discretizations of both the state and design space are considered. We begin with the optimization on a coarse discretization for only two design parameters. The multilevel algorithm then proceeds such that the optimized shape is used on a finer level as the initial guess. Moreover, we prolonged the 2d coarse optimized shape to the third dimension and used that as the initial guess in the multilevel 3d optimization. In the 2d case for 7 design and 12.000 state unknowns we achieved the speedup 4.5. In the 3d case for 12 design and 30.000 state unknowns the speedup was more than 10–times.

Finally, we formulated a corresponding topology optimization problem governed by nonlinear magnetostatics. In the 2d case we solved for 3.920 design variables with 4.832 state ones and the computation typically proceeded within 8 steepest descent iterations and 8 nested nonlinear state Newton iterations. Just during the week in Oberwolfach we managed to run 3d topology optimization governed by linear magnetostatics and we were able to solve problems of up to 1 million design unknowns in hours. The optimal design is close to a sphere around the area where the constant magnetic field is required. The talk was ended with the outlook concerned on using nonlinear multigrid techniques, all–at–once optimization approach and preconditioning techniques for the arising KKT–systems and adaptivity with respect to the cost functional.

References

[1] J. Haslinger, P. Neitaanmäki, *Finite element approximation for optimal shape, material and topology design*, John Wiley & Sons, 1997.
[2] D. Lukáš, *Optimal shape design in magnetostatics*, PhD thesis, 163 pp., TU Ostrava, 2003, http://lukas.am.vsb.cz.
[3] D. Lukáš, *On solution to an optimal shape design problem in 3–dimensional linear magnetostatics*, Appl. Math., 30 pp, to appear in 2004.
[4] D. Lukáš, D. Ciprian, J. Pištora, K. Postava, and M. Foldyna, *Multilevel solvers for 3–dimensional optimal shape design with an application to magneto–optics*, in Proceedings of the 9th International Symposium on Microwave and Optical Technology, 5 pp., to appear in 2004.
[5] D. Lukáš, W. Mühlhuber, and M. Kuhn, *An object–oriented library for shape optimization problems governed by systems of linear elliptic partial differential equations*, in Trans. of VŠB–Technical University of Ostrava, Computer Science and Mathematics Series 1 (2001), pp. 115-128.
[6] D. Lukáš, *Shape optimization of homogeneous electromagnets*, Lect. Notes Comp. Sci. Eng. 18 (2001), pp. 145–152.
[7] J. Schöberl et al., *NgSolve – finite element multi–purpose package*, http://www.hpfem.jku.at.

Fast Time Domain Integral Equation Solvers

Eric Michielssen

(joint work with Mingyu Lu and Balasubramaniam Shanke)

Efficient schemes for analyzing transient electromagnetic wave scattering and short-pulse radiation phenomena are important in disciplines ranging from electromagnetics to acoustics, geophysics, and elastodynamics. The analysis of transient scattering from perfectly conducting as well as potentially inhomogeneous penetrable bodies often is effected using marching on in time (MOT) based time domain integral equation (TDIE) methods.

A typical TDIE solver for analyzing transient electromagnetic scattering from perfect electrically conducting (PEC) surfaces residing in unbounded 3D lossless environments operates as follows. The extinction theorem states that the electromagnetic field anywhere in space can be evaluated upon specification of the incident field and the total magnetic field, or, equivalently, the current, on the scatterer's surface. By enforcing the tangential component of the total electric field along the surface to vanish, the surface current can be related to the incident field through an electric field TDIE. To solve this TDIE by MOT methods, the surface current is represented in terms of N_s spatial basis functions with unknown amplitudes at N_t time steps. Then, the instantaneous total electric field is expressed as a superposition of the incident and scattered fields. The evaluation of the latter requires the computation of a retarded time boundary integral over the basis functions representing the field. This procedure leads to a system of equations that can be solved for the coefficients of the basis functions representing the surface field at a given time step. Depending on the choice of the time step size, the basis functions, and the testing procedure, the matrix to be inverted may be diagonal or sparse, yielding explicit or implicit time stepping schemes, respectively. It has been empirically shown that implicitness and accurate evaluation of retarded time boundary integrals contribute to the stability of a MOT scheme. Unfortunately, the overall computational cost of this procedure scales as $O(N_t N_s^2)$, which prevents the application of classical MOT-based TDIE solvers to the study of practical, real-world problems. It is noted that the above cost estimate is linear in only because the 3D lossless medium Green propagator is local in time. When the above procedure is applied to the study of scattering from 2D objects, or surfaces embedded in dissipative or structured (e.g., layered) environments, then the computational complexity would scale as $O(N_t^2 N_s^2)$, as Green propagators in such media all have a wake.

The recently introduced plane wave time-domain (PWTD) algorithm permits the efficient evaluation of transient wave fields generated by temporally bandlimited sources. The original PWTD scheme targeted sources residing in 3D homogeneous and lossless backgrounds [1]. This PWTD scheme constitutes the extension of the frequency domain (Helmholtz equation) fast multipole method [2, 3] to the time domain (wave equation) and, when coupled to the above described MOT-based TDIE solvers, reduces their computational complexity to $O(N_t N_s \log^2 N_s)$. To date, this PWTD scheme has been successfully used to construct (i) fast marching schemes for solving time domain integral equations [4] and (ii) fast boundary kernels for augmenting finite difference time domain simulators [5]. It even has been extended to 2D [6], layered [7], and dissipative environments [8] with only minor changes in the resulting computational complexity estimates. All PWTD schemes express wave fields as a superposition of plane waves. The evolution of these plane waves is either known analytically, or governed by one-dimensional wave equations. In 2D and in layered environments, a Hilbert transform acts on the plane wave superposition for it to yield the correct transient field. At present, spectral schemes have been developed that control the accuracy of each and every step in these various PWTD schemes; as a result, they can be hybridized with classical MOT-based TDIE solvers, thereby greatly improving their computational complexity and memory requirements, without affecting their accuracy. At present, PWTD-accelerated MOT-based TDIE solvers have been applied to the analysis of scattering and radiation from conducting [4, 9], resistive and impedance boundary condition surfaces [10], penetrable lossless [11], lossy [12], and dispersive volumes [13], and the analysis of hybrid lumped-distributed circuits [14, 15] involving up to hundreds of thousands of spatial unknowns, all this for thousands of time steps.

References

[1] A. A. Ergin, B. Shanker, and E. Michielssen, *Fast evaluation of three-dimensional transient wave fields using diagonal translation operators*, Journal of Computational Physics, 146 (1998), pp. 157-180.

[2] R. Coifman, V. Rokhlin, and S. Wandzura, *The fast multipole method for the wave equation: A pedestrian prescription*, IEEE Antennas and Propagation Magazine, 35 (1993), pp. 7-12.

[3] J. M. Song, C. C. Lu, and W. C. Chew, *MLFMA for electromagnetic scattering by large complex objects*," IEEE Transactions on Antennas and Propagation, 45 (1997), pp. 1488-1493.

[4] B. Shanker, A. Ergin, M. Lu, and E. Michielssen, *Fast analysis of transient scattering phenomena using the multilevel plane wave time domain algorithm*, IEEE Trans. Antennas Propagat, 51 (2003), pp. 628-641.

[5] B. Shanker, A. A. Ergin, and E. Michielssen, *Plane wave time domain acceleration of exact radiation boundary conditions in FDTD analysis of electromagnetic phenomena*, presented at URSI General Assembly, Toronto, Canada, 1999.

[6] M. Lu, J. Wang, A. A. Ergin, and E. Michielssen, *Fast evaluation of two-dimensional transient wave fields*, Journal of Computational Physics, 158 (2000), pp. 161-185.

[7] M. Lu and E. Michielssen, *A marching-on-in-time based transient electric field integral equation solver for microstrip structures*, presented at IEEE International Symposium on Antennas and Propagation, Boston, 2001.

[8] P. Jiang, K. Yegin, S. Li, B. Shanker, and E. Michielssen, *An improved plane wave time domain algorithm for dissipative media*, presented at IEEE Antennas and Propagation Symposium, Columbus, OH, 2003.

[9] N.-W. Chen, K. Aygun, and E. Michielssen, *Integral-equation-based analysis of transient scattering and radiation from conducting bodies at very low frequencies*, IEE Proceedings-Microwaves Antennas & Propagation, 148 (2001), pp. 381-387.

[10] Q. Chen, M. Lu, and E. Michielssen, *Integral equation based analysis of transient scattering from surfaces with impedance boundary condition*, to appear in Microwave and Optical Technology Letters, 2004.

[11] G. Kobidze and B. Shanker, *Integral equation based analysis of scattering from 3-d inhomogeneous anisotropic bodies*, presented at IEEE Antennas and Propagation Society International Symposium, Columbus, OH, 2003.

[12] A. E. Yilmaz, D. S. Weile, B. Shanker, J. M. Jin, and E. Michielssen, *Fast analysis of transient scattering in lossy media*, IEEE Antennas and Wireless Propagation Letters, 1 (2002), pp. 14-17.

[13] G. Kobidze, B. Shanker, and E. Michielssen, *A fast time domain integral equation based scheme for analyzing scattering from dispersive objects*, presented at IEEE Antennas and Propagation Society International Symposium, San Antonio, TX, 2002.

[14] K. Aygun, B. Shanker, and E. Michielssen, *Analysis of non-linear elements and circuitry using plane wave time domain enhanced MOT solvers*, presented at Proceedings of the ICEAA, Torino, Italy, 1999.

[15] K. Aygun, B. C. Fischer, J. Meng, B. Shanker, and E. Michielssen, *A fast hybrid field-circuit simulator for transient analysis of microwave circuits*, IEEE Transactions on Microwave Theory & Techniques, 52 (2004), pp. 573-583.

The Approximation of the Maxwell Eigenvalue Problem using a Least-Squares Method

Joseph E. Pasciak
(joint work with James H. Bramble and Tsanio V. Kolev)

In this talk, I consider the eigenvalue problem problem associated with Maxwell's equations. These equations can, for example, be used to determine the frequencies which will propagate through a medium such as a waveguide or photonic crystal [6, 11, 16].

Although two dimensional versions of Maxwell's eigenvalue problem often result in eigenvalue problems involving the Laplacian, three dimensional problems are significantly more complicated as they result in an eigenvalue problem involving curl-curl, an operator which is not elliptic. Accordingly, the inverse is no longer compact leading to a much more complicated analysis. However, as we shall see, a compact "pseudo" inverse can be constructed which has the same nonzero eigenvectors.

One of the more popular approaches for approximating Maxwell's eigenvalue is based on using curl-conforming spaces such as those developed by Nedelec (cf. [18, 19]). In such a method one looks for solutions to the problem in $\boldsymbol{H}(\mathbf{curl})$, the space of vector function which, along with their curls, are in $\boldsymbol{L}^2(\Omega)$. Analysis of the eigenvalue problem using these spaces either involves proving collective compactness [14, 17] or proving convergence in norm [1, 2].

Early engineering approximations to these equations were often attempted using conforming finite element spaces [3]. These were known to have problems due to low regularity solutions and multiple valued potentials [10, 12, 15]. Recently, new methods for dealing with these problems have been proposed [7, 8, 20]. The methods of [8] depend on weighted functional with weights depending on the strength of the singularities at corners and edges. In [20], discontinuous Galerkin methods are proposed.

The approach which we take in this talk is to first relate the problem to a block system involving the solution of div-curl systems. These div-curl systems are formulated as variational problems following [5] where the solution is posed in $\boldsymbol{L}^2(\Omega)$ and the (components of the) test functions are in various subspaces of the Sobolev space $H^1(\Omega)$. This results in a very weak formulation of the div-curl problem where the data can reside in a negative norm space, e.g., in the dual of the test spaces. That the test functions are in $H^1(\Omega)$ is a critical attribute of the method which we take advantage of in our subsequent analysis of the Maxwell eigenvalue problem. Indeed, this leads to solution operators for the div-curl problem which are bounded from $H^{-1}(\Omega)$ into $L^2(\Omega)$ in the continuous as well as the discrete case. Since the approximation is based in $\boldsymbol{L}^2(\Omega)$, our approximation subspaces can be very simple, for example, we can use discontinuous functions at the material interfaces where the solutions jump while using C^0 elements in the interior where the solution is smooth.

In this talk, I show how this variational form of the div-curl system can be used to develop a stable approximation to the Maxwell's eigenvalue problem. The eigenfunctions with non-zero eigenvalues are also eigenfunctions of a block compact skew-Hermitian problem where the blocks correspond to div-curl problems. We use the div-curl approximation to derive a sequence of approximation operators which converge in norm to the above mentioned compact operator.

Actual three dimensional applications necessarily contain large numbers of unknowns (on the order of millions). Such a large number of unknowns result from complicated device geometry and the mesh refinement necessary for resolving singular behavior in the solutions. Since the systems are too large for conventional direct eigensolvers, the eigenvalues must be computed iteratively. To obtain a system which is more amenable to iterative computation, we show that the original eigenpairs can be computed from those of a compact symmetric real operator. This system can be approximated in norm by the discrete operator for one div-curl system and its adjoint and results in a symmetric discrete eigenvalue problem. The development of effective iterative techniques for computing the eigenvalues of large symmetric problems has been the subject of intensive research in the past two decades, e.g., [4, 9, 13]. These methods are more efficient and robust than those developed for non-symmetric and/or indefinite systems. Thus, the reformulation of the problem as a symmetric real system represents a significant computational advantage.

Theorems on the rate of convergence of the discrete eigenvalues are given and supported by computational experiments.

References

[1] D. Boffi, F. Brezzi, and L. Gastaldi. *On the convergence of eigenvalues for mixed formulations.* Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 25(1-2):131–154 (1998), 1997.

[2] D. Boffi, P. Fernandes, L. Gastaldi, and I. Perugia. *Computational models of electromagnetic resonators: analysis of edge element approximation.* SIAM J. Numer. Anal., 36(4):1264–1290 (electronic), 1999.

[3] A. Bossavit. *Computational electromagnetism.* Academic Press Inc., San Diego, CA, 1998. Variational formulations, complementarity, edge elements.

[4] J. H. Bramble, D. Y. Kwak, and J. E. Pasciak. *Uniform convergence of multigrid V-cycle iterations for indefinite and nonsymmetric problems.* SIAM J. Numer. Anal., 31(6):1746–1763, 1994.

[5] J. H. Bramble and J. E. Pasciak. *A new approximation technique for div-curl systems.* Math. Comp., 2004. (to appear).

[6] e. C. M. Soukoulis. *Photonic Band Gap Materials.* Kluwer, Dordrecht, 1996.

[7] M. Costabel, M. Dauge, and D. Martin. *Numerical investigation of a boundary penalization method for Maxwell equations.* 1999. Preprint.

[8] M. Costabel, M. Dauge, and D. Martin. *Weighted regularization of Maxwell equations in polyhedral domains.* 2001. Preprint.

[9] E. D'yakonov and M. Orekhov. *Minimization of the computational labor in determining the first eigenvalues of differential operators.* Math. Notes, 27:382–391, 1980.

[10] C. Emson, J. Simkin, and C. Trowbridge. *Further developments in three dimensional eddy current analysis.* IEEE Trans. on Magnetics, MAG-21:2231–2234, 1985.

[11] J. D. Joannopoulos, R. D. Meade, and J. N. Winn. *Photonic Crystals.* Princeton University Press, Princeton NJ, 1995.

[12] A. Kameari. *Three dimensional eddy current calculation using finite element method with a-v in conductor and ω in vacuum.* IEEE Trans. on Magnetics, 24:118–121, 1988.
[13] A. Knyazev. *Convergence rate estimates for iterative methods for a mesh symmetric eigenvalue problem.* Sov. J. Num. Anal. Math. Modeling,, 2:371–396, 1987.
[14] F. Kukuchi. *On a discrete compactness property for the nedelec finite elements.* J. Fac. Sci. Univ. Tokyo, Sect. 1A, Math, 36:479–490, 1989.
[15] P. Leonard and D. Rodger. *Finite element scheme for transient 3d eddy currents.* IEEE Trans. on Magnetics, 24:58–66, 1988.
[16] P. Monk. *Finite element methods for Maxwell's equations.* Oxfor Science Pub., Oxford, 2003.
[17] P. Monk, L. Demkowicz, and L. Vardapetyan. *Discrete compactness and approximation of Maxwell's equations in $\mathbb{R}^3$.* Preprint, 2000.
[18] J. C. Nedelec. *Mixed finite elements in $\mathbf{R}^3$.* Numer. Math., 35:315–341, 1980.
[19] J. C. Nedelec. *A new family of mixed finite elements in $\mathbf{R}^3$.* Numer. Math., 50:57–81, 1986.
[20] I. Perugia, D. Schötzau, and P. Monk. *Stabilized interior penalty methods for the time-harmonic Maxwell equations.* 2001. Preprint.

Discontinuous Galerkin Methods for Maxwell's Equations

Ilaria Perugia

In recent years, there has been considerable interest in nonconforming finite element methods that are based on discontinuous piecewise polynomial approximation spaces and on local (element-by-element) variational formulations. Such approaches are referred to as discontinuous Galerkin (DG) methods. The main advantages of these methods lie in their ability to treat a wide range of problems within the same unified framework, and their great flexibility in the mesh-design. Indeed, DG methods can naturally handle non-matching grids and non-uniform, even anisotropic, polynomial approximation degrees; for this reason, DG methods are particularly suited within *hp*-adaptive procedures and for dealing with multi-material problems. In the following, a short survey on DG methods for the approximation of Maxwell's equation is presented.

The original DG method was introduced in [15] for the neutron transport equation. It is constructed by multiplying the equation by smooth test functions, integrating by parts element-by-element on a given mesh, replacing trial and test functions by discontinuous piecewise polynomial functions, and replacing interelement traces by *numerical fluxes.* Development of DG techniques in the context of conservation laws lead to the introduction of the Runge-Kutta (RK) DG method in [3], a high-order method based on a spatial approximation by means of discontinuous polynomials of order k with upwind *numerical fluxes*, and a special $(k+1)$-stage RK method for time-stepping, in combination with *slope limiters* in the case of nonlinear problems (see [4] for a review).

In the context of Maxwell's equations, RKDG-type methods have been applied to the problem in time-domain

$$\varepsilon_r \frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{H} - \mathbf{J}, \qquad \mu_r \frac{\partial \mathbf{H}}{\partial t} = -\nabla \times \mathbf{E},$$

written in conservation form:

$$Q(x)\frac{\partial \mathbf{q}}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{q}) = \mathbf{S}$$

with $\mathbf{q} = [\mathbf{E}, \mathbf{H}]^T$, $\mathbf{F}_i(\mathbf{q}) = [-\mathbf{e}_i \times \mathbf{H}, \mathbf{e}_i \times \mathbf{E}]^T$, and $Q = \mathrm{diag}(\varepsilon_r, \varepsilon_r, \varepsilon_r, \mu_r, \mu_r, \mu_r)$. The use of DG methods in this context is motivated by the possibility of using unstructured, even non-matching, meshes for dealing with complex geometries, by the simplicity of incorporating spatially varying coefficients, and by the possibility of constructing high order methods by simply choosing basis functions; moreover, the mass matrices are diagonal (or block diagonal), which is advantageous for time-stepping.

Schemes based on a DG discretization in space with upwind numerical fluxes and RK time-stepping have been presented several papers: in [14], in combination with a mortar method for treating nonmatching grids; in [6], together with stability analysis and *hp*-error bounds of the proposed scheme (a divergence-free variant of which can be found in [7]); and in [16], where a unified DG method is constructed within the computational and the PML regions. A method using centered numerical fluxes and leap-frog time-stepping in order to reduce dissipation has been introduced in [20].

Finally, a DG space-time approach has been adopted in [5] and in [17], in order to obtain schemes with only *local* CFL control of the time-step for stability, allowing for larger time-steps in larger space elements. These methods use space-time DG methods on meshes generated by advancing front techniques. In particular, in [5], for the case of smooth coefficients, an *explicit* mesh is constructed, allowing for an ordering of the elements with respect to domain of dependence, and therefore, for an explicit element-by-element advancing front solution. In [17], in order to deal with inhomogeneous media, the constraints on the meshes are weakened, allowing for meshes aligned with the discontinuities of the coefficients, and a *semi-implicit* method, based on an ordering of the mesh by macroelements, is constructed.

For the Maxwell's equations in frequency-domain, consider, to fix the ideas, the following electric field-based formulation:

$$\begin{aligned} \nabla \times (\mu^{-1}\nabla \times \mathbf{E}) - \omega^2 \varepsilon \mathbf{E} + i\omega\sigma\mathbf{E} &= -i\omega \mathbf{J}_s \quad &&\text{in } \Omega \\ \mathbf{n} \times \mathbf{E} &= \mathbf{0} &&\text{on } \partial\Omega. \end{aligned}$$

The term $\omega^2\varepsilon\mathbf{E}$ is neglected in the low-frequency case. The solutions of this problem are typically highly oscillatory or strongly singular. DG methods are particularly suited for capturing such solutions, since they allow for an easy implementation of high-order elements and *hp*-adaptive procedures. The main ingredient for the construction of DG schemes, in this context, is the DG approximation of the second order *curl-curl* operator. For a unified presentation of various DG methods for elliptic problems and their theoretical analysis, see [2].

For the low-frequency case, in the simple case of conductivity $\sigma \neq 0$, the problem is elliptic and optimal error estimates can be easily obtained (the case of irregular meshes and only piecewise smooth material coefficients is studied in [18], where *hp*-error bounds are derived). For the high-frequency case, optimal error estimated have been obtained in [9] in the case of smooth coefficients. Mixed methods for imposing the divergence-free constraint on the electric field in the regions where $\sigma = 0$, in the low-frequency case, and for providing control on the divergence of the electric field, in the high-frequency case, have been presented and analyzed in [11] and [19], respectively. An energy-norm *a posteriori* estimator for the mixed method in the low-frequency case has been studied in [10].

Finally, the Maxwell eigenvalue problem has been addressed in [8], where a nonstabilized local discontinuous Galerkin method is used. Numerical results have shown that, in the two-dimensional case, the method correctly captures the eigenmodes, and no spurious mode pollutes the spectrum. In the three-dimensional case, small spurious modes appear, which can be eliminated by adding a suitable stabilization to the scheme.

We conclude with some remarks. For the Maxwell equations in frequency-domain, the eddy-current and the stationary problems, extensive and comparative studies still need to be performed. The same for coupled field-based and potential-based formulations. Up to now, a rigorous analysis of the Maxwell eigenvalue problem has not been performed, as well as a theoretical analysis of the high-frequency problem in a framework which allows for treating discontinuous material coefficients. We finally mention that, in addition to some numerical studies (see [13], [21], [12]), a complete theoretical analysis of dispersion and dissipation errors for DG methods has been carried out in [1].

References

[1] M. Ainsworth. *Dispersive behaviour of high order discontinuous Galerkin finite element methods.* Technical Report 19, University of Strathclyde, Mathematics Department, 2003. To appear in J. Comput. Phys.

[2] D.N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini. *Unified analysis of discontinuous Galerkin methods for elliptic problems.* SIAM J. Numer. Anal., 39 (2001), pp1749–1779.

[3] B. Cockburn and C.-W. Shu. *Tvb Runge-Kutta local projection discontinuous Galerkin finite element method for scalar conservation laws II: General framework.* Math. Comp., 52 (1989), pp411–435.

[4] B. Cockburn and C.-W. Shu. *Runge–Kutta discontinuous Galerkin methods for convection–dominated problems.* J. Sci. Comp., 16 (2001), pp173–261.

[5] R.S. Falk and G.R. Richter. *Explicit finite element methods for symmetric hyperbolic equations.* SIAM J. Numer. Anal., 36 (1999) pp935–952.

[6] J.S. Hesthaven and T. Warburton. *Nodal high-order methods on unstructured grids. I. Time-domain solution of Maxwell's equations.* J. Comput. Phys., 181 (2002) pp186–221.

[7] J.S. Hesthaven and T. Warburton. *High-order accurate methods for time-domain electromagnetics.* Comput. Model. Engin. Sci., 2003. to appear.

[8] J.S. Hesthaven and T. Warburton. *High order nodal discontinuous Galerkin methods for the Maxwell eigenvalue problem.* Technical Report 2003-11, Brown Center for Fluid Mechanics, 2003.

[9] P. Houston, I. Perugia, A. Schneebeli, and D. Schötzau. *Interior penalty method for the indefinite time-harmonic Maxwell equations.* Technical Report 2003-15, Pacific Institute for the Mathematical Sciences, 2003.

[10] P. Houston, I. Perugia, and D. Schötzau. *Energy norm a posteriori error estimation for mixed discontinuous Galerkin approximations of the Maxwell operator.* Technical Report 2003/16, University of Leicester, Department of Mathematics, 2003. To appear in Comput. Methods Appl. Mech. Engrg.

[11] P. Houston, I. Perugia, and D. Schötzau. *Mixed discontinuous Galerkin approximation of the Maxwell operator: Non-stabilized formulation.* Technical Report 2003/17, University of Leicester, Department of Mathematics, 2003. To appear in J. Sci. Comp.

[12] F.Q. Hu and H.L. Atkins. *Eigensolution analysis of the discontinuous Galerkin method with non-uniform grids. Part 1: One space dimension.* J. Comput. Phys., 182 (2002) pp516–545.

[13] F.Q. Hu, M.Y. Hussaini, and P. Rasetarinera. *An analysis of the discontinuous Galerkin method for wave propagation problems.* J. Comput. Phys., 151 (1999) pp921–946.

[14] D.A. Kopriva, S.L. Woodruff, and M.Y. Hussaini. *Computation of electromagnetic scattering with a non-conforming discontinuous spectral element method.* Internat. J. Numer. Methods Engrg., 53 (2002) pp105–122.

[15] P. LeSaint and P.A. Raviart. *On a finite element method for solving the neutron transport equation.* In C. de Boor, editor, *Mathematical Aspects of Finite Elements in Partial Differential Equations*, pages 89–145. Academic Press, New York, 1974.

[16] T. Lu, P. Zhang, and W. Cai. *Discontinuous Galerkin methods for dispersive and lossy Maxwell's equations and PML boundary conditions.* Submitted to J. Comput. Phys. Available at `http://www.math.uncc.edu/~wcai/publication.html`.

[17] P. Monk and G.R. Richter. *A discontinuous Galerkin method for linear symmetric hyperbolic systems in inhomogeneous media finite volume scheme for the 3D heterogeneous.* To appear in J. Sci. Comp. Available at `http://www.math.udel.edu/~monk/`.

[18] I. Perugia and D. Schötzau. *The hp-local discontinuous Galerkin method for low-frequency time-harmonic Maxwell equations.* Math. Comp., 72 (2003), pp1179–1214.

[19] I. Perugia, D. Schötzau, and P. Monk. *Stabilized interior penalty methods for the time-harmonic Maxwell equations.* Comput. Methods Appl. Mech. Engrg., 191 (2002), pp4675–4697.

[20] S. Piperno and L. Fezoui. *A centered discontinuous Galerkin finite volume scheme for the 3D heterogeneous Maxwell equations on unstructured meshes.* Technical Report 4733, INRIA Tech. Rep., 2003.

[21] S. Sherwin. *Dispersion analysis of the continuous and discontinuous Galerkin formulations.* In B. Cockburn, G.E. Karniadakis, and C.-W. Shu, editors, *Discontinuous Galerkin Methods: Theory, Computation and Applications*, volume 11 of *Lect. Notes Comput. Sci. Engrg.*, pages 425–431. Springer–Verlag, 2000.

Smith Normal Form as an Adequate Tool to Detect Mesh Defects as well as to Build Basis Fields for Domains with Loops and Holes

Francesca Rapetti
(joint work with Alain Bossavit (L.G.E.P.) and François Dubois (C.N.A.M.))

A precise description of industrial geometries relies on the use of computer assisted design (C.A.D.) tools. Submeshes are generally created when complex domains with millions of element volumes are concerned. Accidental errors (due to human mistakes, to roundings, to bugs, ...) when gluing together separately created parts will result in spurious holes and/or loops. *How can we perform an automatic mesh defect detection ?*

The Hodge decomposition for a vector $\mathbf{u} \in L^2(\Omega)^3$ consists in its representation as the sum of three orthogonal components $\mathbf{u} = \operatorname{grad}\phi + \operatorname{curl} w + \theta$, the third component θ depending on the domain topology. *How can we build a basis for θ ?*

Algebraic topology and linear algebra help giving an answer to these or other questions.

Let $A : \mathcal{X} \to \mathcal{Y}$ be a linear operator between vector spaces of dimension m and n respectively. If bases are selected in both spaces, A is represented by a $(n \times m)$-matrix $\overline{A}$. One can choose bases in such a way that

$$\overline{A} = \begin{bmatrix} 0_{k,m-k} & \mathrm{Id}_{k,k} \\ 0_{n-k,m-k} & 0_{n-k,k} \end{bmatrix}.$$

This is the *Smith normal form* of A [6]. The normal form clearly exhibits the rank k, the null space (spanned by the first $m-k$ basis vectors in $\mathcal{X}$) and the range (spanned by the last $n-k$ basis vectors in $\mathcal{Y}$) of A (see Figure 1).

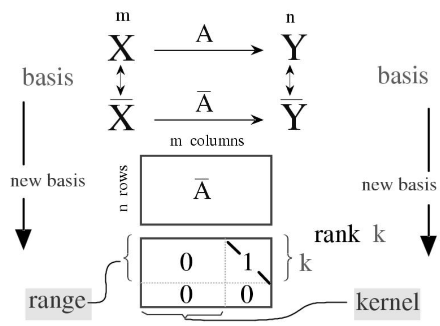


FIGURE 1. Smith normal form for the matrix $\overline{A}$.

Suppose now that one has a *complex* of linear maps $\partial_p : \mathcal{X}_p \to \mathcal{X}_{p-1}$, such that $\partial_p \partial_{p+1} = 0$. By a suitable choice of bases, one can put them all in Smith form,

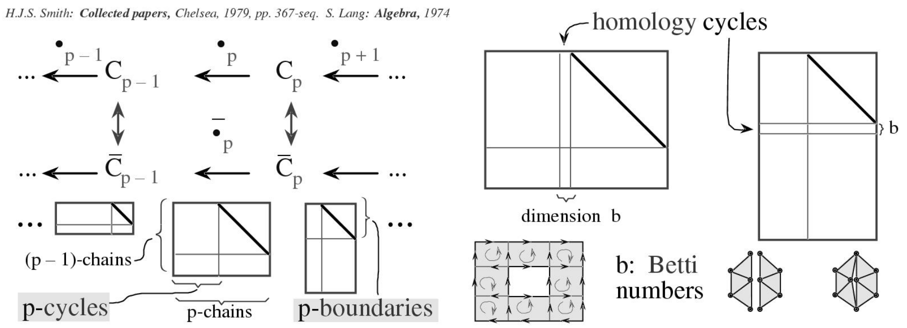


FIGURE 2. Computational configuration and analytical solution on the interface.

thus obtaining a complex of matrices on which one can spot the successive ranges and kernels, and most importantly the quotients $H_p = \ker(\partial_p)/\mathrm{ran}(\partial_{p+1})$.

A case in which this is valuable is when the ∂_ps are the boundary operators acting on chains based on p-cells of some discretization mesh (see Figure 2 (left) where the boundary operators are denoted by black dots carrying the dimension of the cells they act on). The original ∂_p are then the incidence matrices of this mesh, and take the above form when suitable bases are chosen in the chain spaces $\mathcal{X}_p$. One can then easily identify the *cycles* (chains with empty boundary), the *boundaries* (p-chains which bound a $(p+1)$-one), the *homology spaces* H_p and their dimensions b_p, the so-called *Betti numbers*, which are topological invariants (characteristics of the computational domain, not of the particular mesh), telling about the numbers of "holes" and "loops" in the meshed region (see Figure 2 (right)).

Such information is useful as a way to check whether the mesh has been consistently built. For istance, the mesh defects occurring when merging submeshes will result in spurious holes and/or loops, and thus can be detected this way [4]. Hence the interest for *algorithms* to reduce incidence matrices to normal form, with a competitive computational cost. They fall in two classes, depending on whether one works on the primal or the dual mesh.

In [4], we have proposed an algorithm working in $O(s^2)$ where $s = \max(n, m)$ for the considered $(n \times m)$-matrices. The results of the proposed algorithm applied to the incidence matrices of a simplicial discretization of a torus surface are shown in Figure 3. In this case, we are not looking for mesh defects but to an automatic way to compute the generators of H_p, $p = 1, 2$. We work with (incidence) matrices whose entries are *integers*, in particular 0, -1, +1. The Smith normal form of a (n, m)-matrix $\overline{A}$ is computed with *unimodular* transformations, represented by integer matrices with integer matrix inverses and determinants are ± 1. Elementary row operations

- exchange row i with row j
- multiply row i by -1

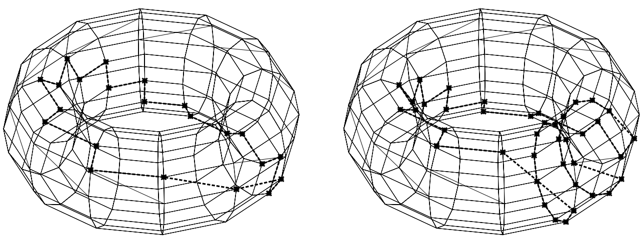


FIGURE 3. Wireframe representation of two loops, generators of the first homology group H_1 of the torus surface, see [5].

- replace row i by (row i) + α (row j), where α is an integer and $k \neq j$

Each of these operations corresponds to a change of basis in Y and similar column operations correspond to a change of basis in X.

These successive changes are stocked in two unimodular matrices, a (n, n)-matrix $\overline{Q}$ and a (m, m)-matrix $\overline{P}$. So, we look for $\overline{Q}$ and $\overline{P}$ such that $\overline{QAP}$ is in Smith form. Then, ker $(\overline{A})$ is spanned by the first $m - k$ column vectors of $\overline{P}$ and imag $(\overline{A})$ is spanned by the last $n - k$ row vectors of $\overline{Q}$ multiplied by the leading elements.

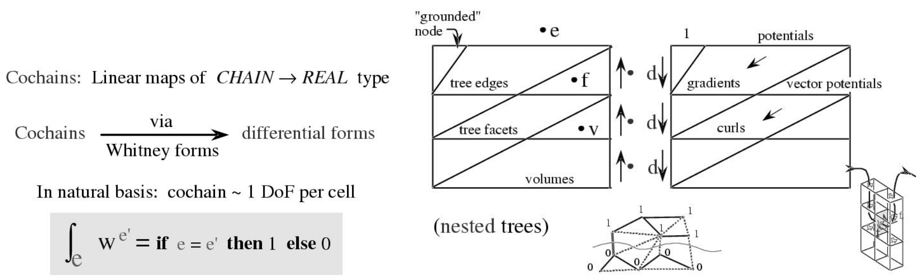


FIGURE 4. The dual side: cohomology.

There is more: by duality, a change of basis for chains induces one on cochains, which are the discrete representation of electromagnetic fields (see Figure 4). In particular, when loops or holes are present, there is a need [1] to construct "non-local" basis fields associated with them, which complete the basis of cell-related Whitney forms, as considered in [5]. Such fields can be read off from the Smith normal form, thanks to the geometric interpretation of the coefficients of the passage matrices. A classification of all possible ways to build representatives of the cohomology classes ("collars", "thick cuts", "tunnels", etc., as found in the work of Kotiuga [3], Kettunen [2], etc.) is thus obtained.

References

[1] P. Dular, W. Legros, A. Nicolet, *Coupling of local and global quantities in various finite element formulations and its application to electrostatics, magnetostatics and magnetodynamics*, IEEE Trans. on Magn., 34 (1998) pp. 3078-81.

[2] `http://natrium.em.tut.fi/kettunen/publications.html`

[3] P.R. Kotiuga, *Hodge decompositions and computational electromagnetics*, Ph.D., Dpt. of Electrical Engng., Mc Gill University, Montréal (1984).

[4] F. Rapetti, F. Dubois, A. Bossavit, *Integer matrix factorization for mesh defect detection*, C. R. Acad. Sc. Paris, Sèrie I, 334 (2002), pp. 717-20.

[5] F. Rapetti, F. Dubois, A. Bossavit, *Discrete vector potentials for non-simply connected three-dimensional domains*, SIAM J. on Numerical Analysis, 41(4) (2003), pp. 1505-1527.

[6] H.J.S. Smith, *On systems of linear indeterminate linear equations*, Phil. Trans., 151 (1861), pp. 293-326.

Pole Condition: A new Approach to Solve Scattering Problems

F. Schmidt

(joint work with T. Hohage and L. Zschiedrich)

The pole condition concept is an approach to investigate certain classes of wave propagation problems on unbounded domains, including the time-dependent Schrödinger equation, the Helmholtz equation and time-harmonic Maxwell equations. The basic idea has been developed originally to solve the 1D time-dependent Schrödinger equation with *non-constant* exterior potentials [8, 9]. The convenient handling of *heterogeneous* exterior domains in 1D situations obtained there was the motivation to extend this concept to higher space dimensions as well as to time-harmonic problems. It turned out that the desired generalization can be done a very natural way.

We discuss the pole condition concept for solving time-harmonic scattering problems modeled by Helmholtz and Maxwell's equations on unbounded domains. The essential aspects are the following. First, the entire space is decomposed into an interior domain containing the scatterer and an exterior domain. The exterior domain may have a heterogeneous structure. Among the admissible types of inhomogeneous exterior domains are waveguide-like inhomogeneities which play an important role in applications. For the special case of 1D problems it was shown [7] that even exterior domains with periodic permittivities can be treated.

The basic idea of the pole condition approach is to consider the Laplace transform of the field in the exterior domain in radial direction. Here, radial direction denotes the distance-like direction in the exterior when covered by a prismatoidal coordinate system [11]. If we fix the angular-like coordinate of the exterior system and let the distance-like coordinate tend to infinity, we move on a ray from the boundary of the interior domain towards infinity. We characterize the exterior fields by the poles of their Laplace transforms along all possible rays and say that a field satisfies the pole condition if its Laplace transform has no pole in the lower half of the complex plane. Fields which satisfy the pole condition are outgoing fields.

A formulation of a scattering problem based on the pole condition consists of three parts: the interior problem, the coupling to the exterior problem, and the exterior problem in its Laplace transformed version. Additionally we have to ensure that the solution of the Laplace part contains only functions that satisfy the pole condition. The latter can be achieved in a number of different ways. One way is to use an extra condition in form of an integral condition [7], another to restrict the possible space of Laplace transformed functions by construction. For the continuous form of the pole condition based formulation of the Helmholtz scattering problem we obtained a number of results. First, the pole condition is equivalent to Sommerfeld's radiation condition in case of homogeneous exterior domains [5], second, the pole condition yields a new representation formula for for the exterior solution, third, parts of theorems concerning the series expansions of exterior fields (theorems of Karp and Wilcox) could be extended. A further surprising result states that the pole condition and the famous PML method are very closely related to each other [6].

The different continuous formulations of the pole condition leads to different numerical algorithms. Until now we investigated mainly two realizations: the cut function approach [5, 7] which is also the basis if the theoretical analysis and the real axis method [3, 7]. Whereas the first one allows directly to compute the exterior fields from the obtained data, the second one yields only the interior solution but has a simpler structure and can easily be extended to solve, e. g., eigenproblems on unbounded domains. A first numerical comparison between the pole condition approach and the PML methods [3] shows that both cause roughly the same numerical costs with a slight favour for PML. However, PML is not able to reproduce the exterior solution.

There are a number of new theoretical results offering new application areas of the pole condition. In [1] it has been shown by Arens and Hohage that the pole condition and the *upward propagating radiation condition* are equivalent. This enables a new approach in solving scattering problems involving unbounded obstacles. In a recent paper [4] Hohage and Stratis proved the equivalence of the pole condition and the Silver-Müller condition for electromagnetic scattering problems. The discrete electromagnetic scattering problem in 2D has been considered in [2]. Another application area of the pole condition concept is the computation of eigensolutions and resonances of open systems. In [10] we develop a convergence theory for the 1D Schrödinger case which allows a safe determination of converged resonances. The complete algorithm and main parts of the theory apply to higher dimensions as well. In [11] Zschiedrich gives a review on the current state of results related to the pole condition concept, a number of new results for time-dependent equations and 2D and 3D applications of our code solving time-harmonic electromagnetic scattering problems.

References

[1] T. Arens and T. Hohage, *On Fourier-Laplace domain methods for rough surface scattering*, submitted, 2003.

[2] Mohammad-Hassan Farshbaf-Shaker, *Ein neues Verfahren zur Lösung des Streuproblems der Maxwell-Gleichungen*, Master's thesis, Zuse-Institute Berlin, 2003, URL: http://www.zib.de/bib/dipl/Shaker.ps.

[3] T. Hohage, F. Schmidt, and L. Zschiedrich, *A new method for the solution of scattering problems*, Proceedings of the JEE'02 Symposium (Toulouse) (B. Michielsen and F. Decavele, eds.), ONERA, 2002, pp. 251–256.

[4] T. Hohage and I. Stratis, *The pole condition for Maxwell's equations in chiral media*, in preparation, 2004.

[5] Thorsten Hohage, Frank Schmidt, and Lin Zschiedrich, *Solving Time-Harmonic Scattering Problems Based on the Pole Condition I:Theory*, SIAM J. Math. Anal. 35 (2003), no. 1, pp183–210.

[6] ———, *Solving Time-Harmonic Scattering Problems Based on the Pole ConditionII: Convergence of the PML Method*, SIAM J. Math. Anal. 35 (2003), no. 3, pp547–560.

[7] F. Schmidt, *Solution of Interior-Exterior Helmholtz-Type Problems Based on the Pole Condition Concept: Theory and Algorithms*, Ph.D. thesis, Free University Berlin, Fachbereich Mathematik und Informatik, 2002.

[8] F. Schmidt and P. Deuflhard, *Discrete transparent boundary conditions for the numerical solution of Fresnel's equation.*, Computers Math. Applic. 29 (1995), pp53–76.

[9] F. Schmidt and D. Yevick, *Discrete transparent boundary conditions for Schrödinger-type equations*, J. Comput. Phys. 134 (1997), pp96–107.

[10] F. Schmidt, L. Zschiedrich, S. Burger, and R. Klose, *Leaky Modes of the 1D Schrödinger Equation: Convergence of the Laplace Domain Method*, in preparation, 2004.

[11] L. Zschiedrich, *Transparent boundary conditions for electromagnetic scattering problems*, Ph.D. thesis, Free University Berlin, Facgbereich Mathematik und Informatik, 2004, in preparation.

Preconditioning for Maxwell Equations

Joachim Schöberl

In this talk, we discuss the construction and analysis of multigrid preconditioners for $H(curl)$ elliptic variational problems. We explain the smoothers of Hiptmair, and Arnold-Falk-Winther. These smoothers take care of components in the discrete kernel of the *curl*-operator, what is the gradient of the H^1 finite element space.

We sketch a new technique for the analysis of multi-level preconditioners in $H(curl)$. It is based on a multi-level decomposition by recently introduced commuting quasi-interpolation operators [1].

The second main topics in the talk is the discussion of algebraic multigrid methods in $H(curl)$. The idea is to define a coarsening algorithm for all finite element spaces in H^1, $H(curl)$, $H(div)$, and L_2, which maintains the complete sequence property on each multigrid level [2]. Thus, the the same smoothers work as in the geometric multigrid.

The last topics are new high order finite elements for all the spaces H^1, $H(curl)$, $H(div)$, and L_2. The high order H^1 elements have lowest order vertex functions, high order edge-, face-, and element-based shape functions. The $H(curl)$ elements have lowest order Nédélec (edge) shape functions, and high order edge-, face-, and element-based shape functions. Next, the $H(div)$ has lowest order Raviart-Thomas (face) shape functions, and high order face- and element-based ones. Finally, the

L_2 element has the constants, and high-order element functions. We stress the advantages of the new elements satisfying localized complete sequence properties for the lowerst order, edge-based, face-base, and element-based shape functions:

$$\begin{array}{llllllll}
W^V_{h,\,p+1=1} & \xrightarrow{\nabla} & V_h^{\mathcal{N}_0} & \xrightarrow{\text{curl}} & Q_h^{\mathcal{RT}_0} & \xrightarrow{\text{div}} & S_{h,\,0} \\
W^E_{p_E+1} & \xrightarrow{\nabla} & V^E_{p_E} & & & & \\
W^F_{p_F+1} & \xrightarrow{\nabla} & V^F_{p_F} & \xrightarrow{\text{curl}} & Q^F_{p_F-1} & & \\
W^I_{p_I+1} & \xrightarrow{\nabla} & V^I_{p_I} & \xrightarrow{\text{curl}} & Q^I_{p_I-1} & \xrightarrow{\text{div}} & S^I_{p_I-2}
\end{array}$$

For the linear system of equations obtained by these basis function, simple block-diagonal preconditioners (the blocks contain unknows associated with edges, faces, and elements) in connection with an good coarse grid solver is efficient for $H(curl)$-elliptic problems in the following sense: The condition number is independent of the relative scaling of the L_2-part and the *curl*-semi-norm in the quadratic form, as well as independent of the mesh size. The dependency of p depends on the choice of the basis functions, and is currently a major point in research.

An other advantage of these basis function is that the order of the gradient functions and rotational functions can be chosen independently. In the limit case of a magnetostatic problem, the gradient functions can be totally skipped, which improves computation time about by a factor of 4. These new high order basis functions are explained in the upcoming paper [3].

All results are available from `http://www.hpfem.jku.at`

References

[1] J. Schöberl, *Commuting quasi-interpolation operators for mixed finite elements*, Preprint ISC-01-10-MATH, Texas A&M University, 2001.

[2] S. Reitzinger and J. Schöberl, *Algebraic Multigrid for Edge Elements*, Numerical Linear Algebra with Applications, 9 (2002), pp. 223–238.

[3] J. Schöberl and J. Zaglmayr, *High order $H(curl)$ and $H(div)$ finite elements with localized complete sequence properties*, work in progress, will appear on `www.hpfem.jku.at`.

Adaptive Multigrid-Methods for the Solution of Time-Harmonic Eddy-Current Problems

O. Sterz

An important class of electromagnetic problems are low frequency applications where the magnetic energy dominates the electric energy. Examples are devices from power engineering like motors, generators, transformers and switch gears as well as medical hyperthermia applications in cancer therapy. Here, the eddy-current approximation of the full Maxwell equations can be employed to describe the electromagnetic fields.

An upper bound for the modeling error of the eddy-current approximation of the full Maxwell-equations at a fixed angular frequency ω, as well as an asymptotic analysis for $\omega \to 0$, is given in [10], further details will be presented in [8]. Concerning the justification of the eddy current model by an asymptotic analysis, we also want to mention the pioneering works [1] and [2].

Assuming perfect conductor conditions $\mathbf{n} \times \mathbf{E} = 0$ at the boundary of the domain Ω, a variational formulation based on the electric field reads: Find $\mathbf{E} \in \mathbf{H}_0(\mathbf{curl};\Omega)$, such that $\forall \mathbf{E}' \in \mathbf{H}_0(\mathbf{curl};\Omega)$

$$(\mu^{-1}\,\mathbf{curl}\,\mathbf{E}, \mathbf{curl}\,\mathbf{E}')_{\mathbf{L}^2(\Omega)} + i\omega(\sigma\mathbf{E}, \mathbf{E}')_{\mathbf{L}^2(\Omega_C)} = -i\omega(\mathbf{J_G}, \mathbf{E}')_{\mathbf{L}^2(\Omega)}\,.$$

¿From this formulation we do not get a unique electric field $\mathbf{E}$ in the insulating sub-domain, since we do not control the divergence of $\mathbf{E}$ and the total charges of the conductors. However, the magnetic field $\mathbf{H} = -(i\omega\mu)^{-1}\,\mathbf{curl}\,\mathbf{E}$, which is the interesting quantity in most cases of eddy current modeling, is unique. The discretization is done by edge elements on simplex grids (Whitney-1-forms) as the most natural choice.

To resolve local phenomena like singular behavior of the fields at edges and corners as well as small penetration depths (skin effect), we rely on an adaptive algorithm. With the help of an residual error estimator, see [5], the elements with the largest estimated error contribution are marked (maximum strategy) and refined (red/green-refinement). This results in a hierarchy of consistent grids.

The computation of real-world problems needs a large number of unknowns, up to several millions on a single processor machine are possible. Thus, for the solution of the linear systems of equations, as the most time consuming task, a fast method is essential. Therefore, multigrid methods are applied since they offer optimal complexity. For the smoothing in the multigrid cycles a standard algorithm like Gauß-Seidel is used in the insulating part of the domain ($\sigma = 0$), whereas the smoothing in the conductive part ($\sigma > 0$) needs a modification: Here, we may apply the idea proposed in [6], which is based on a Helmholtz decomposition and results in an additional smoothing step in the space of scalar potentials. Another possibility is the application of an overlapping block smoother, see [3].

In case of locally adapted grids the overall complexity may not be optimal unless the smoothing is restricted to the refined region. This leads to the implementation of *local* multigrid methods, which can be realized by grids that do not cover the whole computational domain at each level, see Fig. 5.

We finally mention, that the singularity of the arising linear system of equations is not a problem, as long as we take care of two things:

(1) The right hand side is in the range space of the matrix to guarantee solvability.
(2) During the iterative solution process the kernel components of the solution will not grow (or will grow slowly enough) to prevent cancellation errors.

The first condition can be satisfied by an adequate computation of the discrete excitation currents. To comply with the latter condition, we apply some approximate projections onto the orthogonal complement of the kernel of the **curl**-operator.

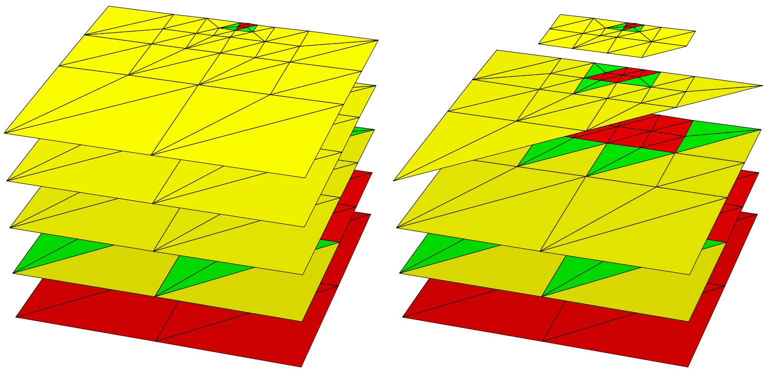


Figure 5. Example of a locally refined unit square: global grids (left) and local grids (right) of the multigrid hierarchy.

This can be realized with low costs by additional multigrid-sweeps on a Poisson-problem, see [10, 11].

All these concepts has been implemented in the adaptive finite element software EM$\mathcal{UG}$ (*electromagnetics on unstructured grids*), which is based on the simulation toolbox $\mathcal{UG}$, see [4, 10]. EM$\mathcal{UG}$ has been successfully applied to benchmark problems as well as realistic problems. A parallel prototype of the electromagnetic simulation tool is currently being developed.

References

[1] A. Alonso, *A mathematical justification of the low-frequency heterogeneous time-harmonic Maxwell equations*, Math. Models Methods Appl. Sci., 9 (1999), pp. 475–489.

[2] H. Ammari, A. Buffa, and J.-C. Nédélec, *A justification of eddy currents model for the Maxwell equations*, SIAM J. Appl. Math., 60 (2000), pp. 1805–1823.

[3] D. Arnold, R. Falk, and R. Winther, *Multigrid in* $H(\mathrm{div})$ *and* $H(\mathbf{curl})$, Numer. Math., 85 (2000), pp. 175–195.

[4] P. Bastian, K. Birken, K. Johannsen, S. Lang, N. Neuss, H. Rentz-Reichert, and C. Wieners, *UG—a flexible software toolbox for solving partial differential equations*, Computing and Visualization in Science, 1 (1997), pp. 27–40.

[5] R. Beck, R. Hiptmair, R. Hoppe, and B. Wohlmuth, *Residual based a-posteriori error estimators for eddy current computation*, $\mathrm{M}^2\mathrm{AN}$, 34 (2000), pp. 159–182.

[6] R. Hiptmair, *Multigrid method for Maxwell's equations*, SIAM J. Numer. Anal., 36 (1999), pp. 204–225.

[7] R. Hiptmair and O. Sterz, *Current and voltage excitations for the eddy currrent model*, Tech. Rep. 2003-07, SAM, ETH Zürich, July 2003. (submitted).

[8] ———, *The justification of the eddy current model*, (in preparation), (2004).

[9] O. Sterz, *Multigrid for time harmonic eddy currrents without gauge*, in Scientific Computing in Electrical Engineering. Proceedings of the 4th International Workshop Scientific Computing in Electrical Engineering, Eindhoven, The Netherlands, June 23–28, 2002, *LNCSE*, Berlin, 2002, Springer. (in print).

[10] ———, *Modellierung und Numerik zeitharmonischer Wirbelstromprobleme in 3D*, PhD thesis, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Universität Heidelberg, 2003. (in German) `http://www.ub.uni-heidelberg.de/archiv/4346`.

[11] ———, *Multigrid for time harmonic eddy currrents without gauge*, Preprint 2003-7, IWR, Universität Heidelberg, 2003.

Perfectly Matched Layers

Fernando L. Teixeira

The simulation of electromagnetic problems in unbounded regions with partial differential equation (PDE) based methods, such as finite element (FE) and finite difference (FD) methods, necessitates the use of an absorbing boundary condition (ABC) to emulate the radiation condition at infinity. Perfectly matched layers (PML) are absorption layers used toward his purpose. The PML achieves a reflectionless absorption of electromagnetic waves in the continuum limit as the mesh discretization size goes to zero. The absorption inside the PML operates through conductive losses, so that an exponential decay for the fields inside the PML is obtained. Therefore, when the computational domain is surrounded by a PML region, spurious reflections from the grid boundaries can be made exponentially smaller. Being a *local* ABC, the PML retains the nearest-neighbor interaction characteristic of PDE-based methods, and therefore it is particularly suited for PDE-bases simulations on parallel computers. Also because of this property, the PML retains the inherent sparsity and (low) computational complexity of PDE-based methods.

When first introduced in the literature [1], the PML relied upon the use of matched artificial electric and magnetic conductivities and the splitting of the electromagnetic field components into two subcomponents each (split-field formulation). Because of this, the resulting fields inside the PML layer were rendered nonphysical (non-Maxwellian). The PML was later shown to be equivalent to a complex coordinate stretching of the coordinate space [2] or a complex coordinate transformation (analytic continuation of the coordinate space) [3],[4],[5]. Via such transformation, the (real) spatial coordinates are mapped as

$$\zeta \to \tilde{\zeta} = \int_0^{\zeta} s_\zeta(\zeta')d\zeta'$$

where s_ζ, with $\zeta = x, y, z$, are the so-called *complex stretching variables*, given by

$$s_\zeta(\zeta, \omega) = a_\zeta(\zeta) + i\frac{\Omega_\zeta(\zeta)}{\omega}$$

with $a_\zeta \geq 1$ and $\Omega_\zeta \geq 0$ (profile functions). The first inequality ensures that evanescent waves will have a faster exponential decay in the PML region, and the second inequality ensures that propagating waves will also decay exponentially along the respective coordinates in the PML. The ordinary Maxwell's equations are recovered from the above when $s_\zeta = 1$. Therefore, the complex stretching variables can be seen as added degrees of freedom to Maxwell's equations.

The PML has also found an interesting dual formulation (Maxwellian PML) with a more clear physical interpretation whereby the PML is represented by frequency dependent material (constitutive) tensors $\overline{\overline{\epsilon}}$ and $\overline{\overline{\mu}}$ [6],[7]. These tensors also produce reflectionless absorption in the continuum limit. In addition to a more direct physical interpretation, the Maxwellian PML yields an easier interfacing with FE codes and a strongly well-posed formulation, as opposed to a weakly well-posed formulation in the original split-field PML [8].

The PML was first developed for planar grid terminations (Cartesian coordinates) [1],[2]. In order to be used with more general grid terminations, the PML later extended to curvilinear coordinates [3],[4],[5]. Although the first of such extensions have dealt with non-Maxwellian formulations only, it was later shown that Maxwellian PMLs could also be obtained in curvilinear geometries [9],[10]. In its most general form (for doubly curved surfaces), the curvilinear PML correspond to a *conformal* layer of anisotropic material tensors with inhomogeneous constitutive properties that depend on the local geometry (principal curvatures) of the mesh termination surface S [10],[11]. These PML constitutive parameters are given by $\overline{\overline{\mu}} = \mu\overline{\overline{\Lambda}}$ and $\overline{\overline{\epsilon}} = \epsilon\overline{\overline{\Lambda}}$, with [10]

$$\overline{\overline{\Lambda}} = \hat{t}_1\hat{t}_1\left(\frac{sh_1\tilde{h}_2}{\tilde{h}_1h_2}\right) + \hat{t}_2\hat{t}_2\left(\frac{s\tilde{h}_1h_2}{h_1\tilde{h}_2}\right) + \hat{n}\hat{n}\left(\frac{\tilde{h}_1\tilde{h}_2}{sh_1h_2}\right).$$

Here s is the complex tretching coordinate along the normal coordinate ξ_3 at a point P in the mesh termination surface S, and h_i and $\tilde{h}_i$, $i = 1, 2$ are the nonstretched and stretched, respectively, (local) metric coefficients [10]. The unit vectors $\hat{t}_i$, $i = 1, 2$ are tangent to S at P along the principal lines of curvature that define tangential orthogonal coordinates ξ_1 and ξ_2, and $\hat{n}$ is the unit normal vector at that point (outward). The metric coefficients are given by $h_i = r_i/r_{0i}$, where r_{0i} are the principal radii of curvature at P and $r_i = r_{0i} + \xi_3$, $i = 1, 2$. The conformal PML is hence constructed over *parallel* surfaces to S. A basic limitation that exist in this general case, however, is that both radii of curvature should be non-negative (i.e., the PML can only be defined over planar or concave termination surfaces as viewed from inside the computational domain). Otherwise, dynamical instabilities ensue [11]. We note that the Cartesian, cylindrical, and spherical PMLs are special cases of this general curvilinear case, followed (possibly) by a successive application of the analytic continuation in orthogonal directions, if needed to achieve absorption in corner regions.

It is also possible to generalize the PML to terminate problems in more complex media, such as linear interior media exhibiting frequency dispersion and/or (bi) anisotropy [12]. This is in contrast to other local ABC, where an exact extension is often not possible in such cases. This extension is particularly important, for example, in electromagnetic simulations involving subsurface problems or complex materials [13]. For example, given an arbitrary dispersive and/or (bi)anisotropic linear interior media in a Cartesian domain with constitutive tensors $\overline{\overline{\epsilon}}(\omega)$, $\overline{\overline{\xi}}(\omega)$, $\overline{\overline{\zeta}}(\omega)$, $\overline{\overline{\mu}}(\omega)$, the corresponding Maxwellian PML bianisotropic constitutive parameters are given as [12]

$$\overline{\overline{\lambda}}_{PML}(\omega) = (\det\overline{\overline{S}})^{-1}\left(\overline{\overline{S}}\cdot\overline{\overline{\lambda}}(\omega)\cdot\overline{\overline{S}}\right),$$

where the symbol $\overline{\overline{\lambda}}$ stands for any of the above four constitutive tensors, and

$$\overline{\overline{S}}(\omega) = diag\{s_x^{-1}, s_y^{-1}, s_z^{-1}\}$$

Finally, we note that PML concept also admits a *geometric* interpretation as a complexification of the *metric* of space in the Fourier domain [14]. By exploring this interpretation, it can be shown that the differential forms language [15],[16] provides an elegant framework to unify the various PML formulations and obtain further generalizations. This is because the metric invariance of Maxwell's equations (in the sense of [17],[18]) is explicitely manifest in such language. A modification in the metric (diffeomorphism) corresponds to a modification on the Hodge operator, which fully incorporates the constitutive relations. The existence of Maxwellian PMLs can be seen as a simple consequence of the metric invariance of Maxwell's equations. The various PML formulations in the vector calculus language arise from the different choices on how to map differential forms to vector fields. This map fixes an isomorphism between differential forms and vectors and it depends on a metric. If the real metric is chosen to define such map, then the Maxwellian PML is receovered. Alternatively, if the complex (stretched) metric is chosen, then the non-Maxwellian PML is recovered. This also reveals that if other consistent metrics are chosen to fix the form-vector isomorphism (e.g., hybridizations of the previous ones), other (indeed, infinitely many) PML formulations are possible, albeit more cumbersome for practical implementation in numerical algorithms [14]. In such context, the existing PML formulations are particular cases of these choices.

References

[1] J. P. Berenger, *A perfectly matched layer for the absorption of electromagnetic waves*, J. Comput. Phys., 114 (1994), pp. 185–200.

[2] W. C. Chew and W. Weedon, *A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates*, Microwave Opt. Techol. Lett., 7 (1994), pp. 599–604.

[3] W. C. Chew, J. M. Jin, , and E. Michielssen, *Complex coordinate stretching as a generalized absorbing boundary condition*, Microwave Opt. Technol. Lett., 15 (1997), pp. 363–369.

[4] F. L. Teixeira and W. C. Chew, *PML-FDTD in cylindrical and spherical grids*, IEEE Microwave Guided Wave Lett., 7 (1997), pp. 285–287.

[5] F. Collino and P. Monk, *The perfectly matched layer in curvilinear coordinates*, SIAM J. Sci. Comput., 19 (1998), pp. 2061–2090.

[6] Z. S. Sacks, D. M. Kingsland, R. Lee, and J.-F. Lee, *A perfectly matched anisotropic absorber for use as an absorbing boundary condition*, IEEE Trans. Antennas Propagat., 43 (1995), pp. 1460–1463.

[7] S. D. Gedney, *An anisotropic PML absorbing media for the FDTD simulation of fields in lossy and dispersive media*, Electromagn., 16 (1996), pp. 399–415.

[8] S. Abarbanel and D. Gottlieb, *A mathematical analysis of the PML method*, J. Comput. Phys., 134 (1997), pp. 357–363.

[9] F. L. Teixeira and W. C. Chew, *Systematic derivation of anisotropic PML absorbing media in cylindrical and spherical coordinates*, IEEE Microwave Guided Wave Lett., 7 (1997), pp. 371–373.

[10] F. L. Teixeira and W. C. Chew, *Analytical derivation of a conformal perfectly matched absorber for electromagnetic waves*, Microwave Opt. Technol. Lett., 17 (1998), pp. 231–236.

[11] F. L. Teixeira and W. C. Chew, *Conformal PML-FDTD schemes for electromagnetic field simulations: a dynamic stability study*, IEEE Trans. Antennas Propagat., 49 (2001), pp. 902–907.
[12] F. L. Teixeira and W. C. Chew, *General closed-form PML constitutive tensors to match arbitrary bianisotropic and dispersive linear media*, IEEE Microwave Guided Wave Lett., 8 (1998), pp. 223–225.
[13] F. L. Teixeira and W. C. Chew, *Finite-difference simulation of transient electromagnetic fields for cylindrical geometries in complex media*, IEEE Trans. Geosci. Remote Sensing, 38 (2000), pp. 1530–1543.
[14] F. L. Teixeira and W. C. Chew, *Differential forms, metrics, and the reflectionless absorption of electromagnetic waves*, J. Electromagn. Waves Applicat., 13 (1999), pp. 665-686.
[15] A. Bossavit, *Differential forms and the computation of fields and forces in electromagnetism*, Eur. J. Mech. B, 10 (1991), pp. 474–488.
[16] R. Hiptmair, *Canonical construction of finite elements*, Math. Comp., 68 (1999), pp. 1325–1346.
[17] D. Van Dantzig, *The fundamental equations of electromagnetism, independent of metrical geometry*, Proc. Cambridge Phil. Soc., 37 (1934), pp. 421–427.
[18] F. L. Teixeira and W. C. Chew, *Lattice electromagnetic theory from a topological viewpoint*, J. Math. Phys., 40 (1999), pp. 169–187.

Some New Inexact Uzawa Methods and Non-overlapping DD Preconditioners for Solving Maxwell's Equations in Non-homogeneous Media

Jun Zou
(joint work with Qiya Hu)

This talk will review some new preconditioned Uzawa iterative methods for solving saddle-point systems, and a non-overlapping domain decomposition preconditioner for solving three-dimensional Maxwell's equations in non-homogeneous media.

Iterative methods for saddle-point system. Consider the system

$$Ax + By = f\,, \quad B^t x = g \tag{1}$$

where A is a symmetric and positive definite $n \times n$ matrix, and B is an $n \times m$ matrix with $m \leq n$. The system (1) is assumed to be nonsingular, so the Schur complement matrix $C = B^t A^{-1} B$ is positive definite. Linear systems such as (1) arise often from finite element discretizations of Maxwell equations and Navier-Stokes equations. Solving the saddle-point system (1) is usually much more difficult than solving the SPD system $Ax = b$. Recently the following inexact preconditioned Uzawa-type algorithm:

$$x_{i+1} = x_i + \hat{A}^{-1}[f - (Ax_i + By_i)]\,, \qquad y_{i+1} = y_i + \hat{C}^{-1}(B^t x_{i+1} - g) \tag{2}$$

has been widely used and studied (cf. [1] [2] [3]) for solving (1). Here $\hat{A}$ and $\hat{C}$ are preconditioners for A and C. The existing convergence results indicate that these algorithms converges assuming some good knowledge of the spectrum of the preconditioned matrices $\hat{A}^{-1}A$ and $\hat{C}^{-1}C$ or under some proper scalings of the preconditioners $\hat{A}$ and $\hat{C}$. This "preprocessing" may not be easy and convenient to achieve in some applications.

Is it possible to introduce some relaxation parameters in (2) so that the resulting algorithm always converges for any SPD preconditioners $\hat{A}$ and $\hat{C}$, and converges with good rate when good preconditioners are available ? The following algorithm was proposed for this purpose (cf.[4]):

$$(3)\quad x_{i+1} = x_i + \omega_i \hat{A}^{-1}[f - (Ax_i + By_i)]\,, \qquad y_{i+1} = y_i + \tau_i \hat{C}^{-1}(B^t x_{i+1} - g)$$

where two parameters ω_i and τ_i can be updated using only the actions of $\hat{A}^{-1}$ and $\hat{C}^{-1}$.

The detailed convergence and convergence rate of algorithm (3) were given in terms of the condition numbers $\kappa(\hat{A}^{-1}A)$ and $\kappa(\hat{C}^{-1}C)$, without any conditions on preconditioner $\hat{C}$, see [4]. Unfortunately our proofs hold only with the condition that $\hat{A}$ is properly scaled so that the eigenvalues of $A^{-1}\hat{A}$ are bounded by one, although numerical experiments still demonstrated convergence when this condition is violated.

When a good preconditioner $\hat{A}$ is not available, one may replace the preconditioning part of $\hat{A}$ in (3) by some nonlinear iteration. This leads to the following algorithm (cf. [5]):

$$(4)\quad x_{i+1} = x_i + \Psi(f - (Ax_i + By_i))\,, \qquad y_{i+1} = y_i + \tau_i \hat{C}^{-1}(B^t x_{i+1} - g)$$

where Ψ is a nonlinear map in R^n such that for any $\phi \in R^n$, $\Psi(\phi)$ approximates the solution ξ of $A\xi = \phi$. And the parameter τ_i can be updated using only the actions of $\hat{C}^{-1}$ and Ψ.

The detailed convergence and convergence rate of the algorithm (4) can be given in terms of the condition number $\kappa(\hat{C}^{-1}C)$ and the tolerance parameter used for Ψ, and no any conditions on the preconditioner $\hat{C}$ are needed.

The algorithm (4) may not work well when the conditioning of the preconditioned Schur complement $\hat{C}^{-1}C$ is much worse than the conditioning of system $\hat{A}^{-1}A$. In this case, we may use a few PCG iterations with preconditioner $\hat{C}$ to improve the conditioning of $\hat{C}^{-1}C$, then apply the algorithm (4). This suggests the following algorithm (cf. [7]):

$$(5)\quad x_{i+1} = x_i + \Psi(f - (Ax_i + By_i))\,, \qquad y_{i+1} = y_i + \tau_i \Psi_H(B^t x_{i+1} - g),$$

where $\Psi_H(g_i)$ for any g_i is the iterate generated by the PCG method with preconditioner $\hat{C}$ for solving $H\psi = g_i$ with $H = B^T \hat{A}^{-1} B$ such that for some $\delta \in (0,1)$,

$$\|\Psi_H(g_i) - H^{-1}g_i\|_H \le \delta\, \|H^{-1}g_i\|_H\,.$$

The actual effect of $\Psi_H(g_i)$ amounts to generating a new preconditioner $\hat{Q}_i$ such that the conditioning of $\hat{Q}_i^{-1}C$ is much improved than the one of $\hat{C}^{-1}C$ and $\kappa(\hat{Q}_i^{-1}C)$ is about the same as $\kappa(\hat{A}^{-1}A)$ (cf. [7]). The convergence and convergence rate of this algorithm was given in [7] and also applied to solving nonlinear saddle-point system like

$$F(x) + By = f, \quad B^t x = g\,.$$

Non-overlapping domain decomposition methods. Consider the Maxwell system:

$$\left\{\begin{array}{ll} \nabla\times(\alpha\nabla\times\mathbf{u})+\gamma_0\beta\mathbf{u}=\mathbf{f} & \text{in}\quad\Omega \\ \nabla\cdot(\beta\mathbf{u})=g & \text{in}\quad\Omega \end{array}\right. \tag{6}$$

where Ω is a Lipschitz polyhedral domain in $\mathbf{R}^3$, not necessarily convex. $\alpha(x)$ and $\beta(x)$ are positive but may be discontinuous in Ω. The perfect boundary condition $\mathbf{u}\times\mathbf{n}=0$ is assumed on $\partial\Omega$. The constant γ_0 is non-negative, and it is allowed to be identically zero. It is this extreme case that causes the most troublesome technical difficulty in the analysis.

The variational saddle-point problem associated with system (6) is formulated as follows:

Find $(\mathbf{u},p)\in H_0(\mathbf{curl};\Omega)\times H_0^1(\Omega)$ such that

$$\left\{\begin{array}{l} (\alpha\nabla\times\mathbf{u},\nabla\times\mathbf{v})+\gamma_0(\beta\mathbf{u},\mathbf{v})+(\nabla p,\beta\mathbf{v})=(\mathbf{f},\mathbf{v}),\quad \forall\mathbf{v}\in H_0(\mathbf{curl};\Omega) \\ (\beta\mathbf{u},\nabla q)=(g,q),\qquad \forall q\in H_0^1(\Omega). \end{array}\right. \tag{7}$$

Domain decompositions and edge elements. Decompose Ω into N non-overlapping tetrahedral subdomains $\{\Omega_i\}_i^N$, with each Ω_i of size d. The common face of subdomains Ω_i and Ω_j is denoted by Γ_{ij}, and set $\Gamma=\cup\Gamma_{ij}$, and $\Gamma_i=\Gamma\cap\partial\Omega_i$. Then we divide each Ω_i into smaller tetrahedral elements of size h so that elements from two neighboring subdomains match with each other on the interface Γ. Let T_h be the resulting triangulation of the domain Ω. We shall approximate the field $\mathbf{u}$ and multiplier p by the Nédélec edge element space of lowest order and the piecewise linear nodal element space of $H_0^1(\Omega)$, denoted by $V_h(\Omega)$ and $Z_h(\Omega)$. Then the edge element approximation of system (7) is to find $(\mathbf{u}_h,p_h)\in V_h(\Omega)\times Z_h(\Omega)$ such that

$$\left\{\begin{array}{l} (\alpha\nabla\times\mathbf{u}_h,\nabla\times\mathbf{v}_h)+\gamma_0(\beta\mathbf{u}_h,\mathbf{v}_h)+(\nabla p_h,\beta\mathbf{v}_h)=(\mathbf{f},\mathbf{v}_h),\quad \forall\mathbf{v}_h\in V_h(\Omega) \\ (\beta\mathbf{u}_h,\nabla q_h)=(g,q_h),\qquad \forall q_h\in Z_h(\Omega). \end{array}\right. \tag{8}$$

For any face F of Ω_i, F_b denotes the union of all T_h-induced (closed) triangles on F, which have either one single vertex or one edge lying on $\partial\mathrm{F}$, and F_∂ denotes the open set $\mathrm{F}\backslash\mathrm{F}_b$. For any subdomain Ω_i, define $\Delta_i=\cup_{\mathrm{F}\subset\Gamma_i}\mathrm{F}_b$. With each Ω_i, we define a local operator A_i on $V_h(\Omega_i)$, a standard restriction space of $V_h(\Omega)$ on Ω_i, by

$$(A_i\mathbf{u},\mathbf{v})=(\alpha\nabla\times\mathbf{u},\nabla\times\mathbf{v})_{\Omega_i}+(\alpha\mathbf{u},\mathbf{v})_{\Omega_i},\quad \forall\mathbf{u},\ \mathbf{v}\in V_h(\Omega_i).$$

And $\tilde{A}$ is defined similarly to A_i but on the global space $V_h(\Omega)$. For any $\Phi\in V_h(\Gamma_i)$, we define its discrete A_i-extension $\mathbf{R}_h^i\Phi$ in $V_h(\Gamma_i)$: $\mathbf{R}_h^i\Phi\times\mathbf{n}=\Phi$ on Γ_i and solves

$$(A_i\mathbf{R}_h^i\Phi,\mathbf{v}_h)=0,\quad \forall\mathbf{v}_h\in V_h^0(\Omega_i).$$

We can write system (8) as the operator form:

$$(\bar{A}+\gamma_0\,\beta\,I)\mathbf{u}_h+Bp_h=\bar{\mathbf{f}}_h,\quad B^t\mathbf{u}_h=g_h. \tag{9}$$

Noting that the operator $\bar{A}$ is singular in $V_h(\Omega)$, we can not apply the existing Uzawa-type iterative solvers for solving the saddle-point system when $\gamma_0=0$. To avoid the difficulty, we rewrite (9) into the following equivalent system

$$A\,\mathbf{u}_h+Bp_h=\mathbf{f}_h,\quad B^t\mathbf{u}_h=g_h \tag{10}$$

where $A=\bar{A}+\gamma_0\,\beta\,I$ for $\gamma_0\neq 0$ and $A=\bar{A}+r_0B\hat{C}^{-1}B^t$ if $\gamma_0=0$. Now one can apply, for example, the inexact Uzawa algorithm (3) for (10). It is important to note that the action of $\hat{C}^{-1}$ needs only once in each iteration, and the convergence rate of this algorithm is determined by $\kappa(\hat{A}^{-1}A)$ and $\kappa(\hat{C}^{-1}B^t\hat{A}^{-1}B)$.

Construction of preconditioners for A and $B^tA^{-1}B$.

One can show (cf. [6]) that if $\hat{C}$ is a preconditioner for $B^tA^{-1}B$ such that $(\beta\nabla\phi, \nabla\phi) \lesssim (\hat{C}\phi, \phi) \lesssim G(d/h)(\beta\nabla\phi, \nabla\phi)$ for all $\phi \in Z_h(\Omega)$,
then we have $G(d/h)^{-1}(\tilde{A}\mathbf{v}_h, \mathbf{v}_h) \lesssim (A\mathbf{v}_h, \mathbf{v}_h) \lesssim (\tilde{A}\mathbf{v}_h, \mathbf{v}_h)$, for all $\mathbf{v}_h \in V_h(\Omega)$. So it suffices to construct a preconditioner for $\tilde{A}$, instead of A.

Let $\lambda_e(\mathbf{v})$ be the moment of $\mathbf{v}$ on any edge e, $V^H(\Omega) \subset V_h(\Omega)$, consisting of all discrete A_i-extensions in each Ω_i, and

$$\begin{aligned} V^p(\Omega) &= \prod_{k=1}^{N} V_h^0(\Omega_k), \quad V^{ij}(\Omega) = \Big\{\mathbf{v} \in V^H(\Omega);\ \mathrm{supp}(\mathbf{v}) \subset \Omega_i \cup \Omega_j \cup \Gamma_{ij}\Big\}, \\ V^0(\Omega) &= \Big\{\mathbf{v} \in V^H(\Omega);\ \lambda_e(\mathbf{v}) = 0 \text{ for each } e \in \mathrm{F}_\partial \text{ with } \mathrm{F} \subset \Gamma\Big\} \end{aligned}$$

while $\hat{A}_p$ and $\hat{A}_{ij}$ are operators on $V^p(\Omega)$ and $V^{ij}(\Omega)$, and $\hat{A}_0$ the coarse solver in $V^0(\Omega)$:

$$(\hat{A}_p\mathbf{v}, \mathbf{v}) \overline{\overline{\sim}} \sum_{k=1}^{N} (A_k\mathbf{v}, \mathbf{v})_{\Omega_k} \ \forall \mathbf{v} \in V^p(\Omega);$$

$$(\hat{A}_{ij}\mathbf{v}, \mathbf{v}) \overline{\overline{\sim}} (A_i\mathbf{v}_i, \mathbf{v}_i)_{\Omega_i} + (A_j\mathbf{v}_j, \mathbf{v}_j)_{\Omega_j} \ \forall \mathbf{v} \in V^{ij}(\Omega)$$

$$(\hat{A}_0\mathbf{v}, \mathbf{w}) = h[1 + \log(d/h)] \sum_{i=1}^{N} \alpha_i \Big\{ \langle \mathrm{div}_\tau(\mathbf{v} \times \mathbf{n})|_{\Gamma_i}, \mathrm{div}_\tau(\mathbf{w} \times \mathbf{n})|_{\Gamma_i} \rangle_{\Delta_i} + \langle \mathbf{v} \times \mathbf{n}, \mathbf{w} \times \mathbf{n} \rangle_{\Delta_i} \Big\}$$

Then the additive preconditioner $\hat{A}$ formed by $\hat{A}_0$, $\hat{A}_p$ and $\hat{A}_{ij}$ is nearly optimal, i.e. $\kappa(\hat{A}^{-1}A) \lesssim G(d/h)[1 + \log(d/h)]^2$, also independent of jumps of material coefficients (cf.[6]).

Acknowledgments. Jun Zou was supported by HKRGC Grants (No. CUHK4048/02P and 403403); Qiya Hu was supported by Special Funds for Major State Basic Research Projects of China G1999032804.

References

[1] R. Bank, B. Welfert and H. Yserentant, *A class of iterative methods for solving saddle point problems*, Numer. Math., 56 (1990), pp. 645–666.

[2] J. H. Bramble, J. E. Pasciak, and A. T. Vassilev, *Analysis of the inexact Uzawa algorithm for saddle point problems*, SIAM J. Numer. Anal., 34 (1997), pp. 1072–1092.

[3] H. C. Elman and G. H. Golub, *Inexact and preconditioned Uzawa algorithms for saddle point problems*, SIAM J. Numer. Anal., 31 (1994), pp. 1645–1661.

[4] Q. Hu and J. Zou, *An iterative method with variable parameters for saddle-point problems*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 317-338.

[5] Q. Hu and J. Zou, *Two new variants of nonlinear inexact Uzawa algorithms for saddle-point problems*, Numer. Math., 93 (2002), pp. 333-359.

[6] Q. Hu and J. Zou, *Substructuring preconditioners for saddle-point problems arising from Maxwell's equations in three dimensions*, Math. Comput., 73 (2004), pp. 35-61.

[7] Q. Hu and J. Zou, *Nonlinear inexact Uzawa algorithms for linear and nonlinear saddle-point problems*, Research Report 2003-03, June 2003, Institute of Computational Mathematics and Scientific/Engineering Computing, The Chinese Academy of Sciences, China.

Reporter: Ralf Hiptmair

Participants

Prof. Dr. Mark Ainsworth
m.ainsworth@strath.ac.uk
ma@maths.strath.ac.uk
Department of Mathematics
University of Strathclyde
Livingstone Tower
26, Richmond Street
GB-Glasgow, G1 1XH

Dr. Ana Alonso
alonso@science.unitn.it
Dipartimento di Matematica
Universita di Trento
Via Sommarive, 14
I-38050 Trento

Prof. Dr. Peter Arbenz
arbenz@inf.ethz.ch
Institut für Wissenschaftliches
Rechnen
ETH Zentrum
(HRS G 27)
CH-8092 Zürich

Prof. Dr. Abderrahmane Bendali
bendali@cerfacs.fr
Abderahman.Bendali@cerfacs.fr
CERFACS
42, Ave. Gustave Coriolis
F-31057 Toulouse Cedex

Prof. Dr. Oszkar Biro
biro@igte.tu-graz.ac.at
biro@TUGraz.at
IGTE
TU Graz
Kopernikusgasse 24
A-8010 Graz

Dr. Steffen Börm
sbo@mis.mpg.de
Steffen.Boerm@mis.mpg.de
Max-Planck-Institut für Mathematik
in den Naturwissenschaften
Inselstr. 22 - 26
D–04103 Leipzig

Prof. Dr. Daniele Boffi
boffi@dimat.unipv.it
Dipartimento di Matematica
Universita di Pavia
Via Ferrata 1
I-27100 Pavia

Prof. Dr. Alain Bossavit
bossavit@lgep.supelec.fr
Laboratoire de Genie Electrique
de Paris Supelec (LGEP)
Universite Paris VI et Paris XI
11, Rue Joliot-Curie
F-91192 Gif-sur-Yvette -CEDEX

Dr. Annalisa Buffa
annalisa@imati.cnr.it
IMATI - CNR
Via Ferrata, 1
I-27100 Pavia

Dr. Zhiming Chen
zmchen@lsec.cc.ac.cn
Inst. of Computational Mathematics
Chinese Academy of Sciences
P.O.Box 2719
Beijing 100080 – China

Dr. Snorre Harald Christiansen
snorrec@math.uio.no
CMA
c/o Department of Mathematics
University of Oslo
P.O.Box 1053 - Blindern
N-0316 Oslo

Dr. Patrick Ciarlet
ciarlet@ensta.fr
ENSTA/UMA
32, boulevard Victor
F-75739 Paris Cedex 15

Dr. Markus Clemens
clemens@temf.tu-darmstadt.de
Institut für Theorie
Elektromagnetischer Felder
Schloßgartenstraße 8
D–64289 Darmstadt

Prof. Dr. David L. Colton
colton@math.udel.edu
Department of Mathematical Sciences
University of Delaware
501 Ewing Hall
Newark, DE 19716-2553 – USA

Prof. Dr. Martin Costabel
costabel@univ-rennes1.fr
Departement de Mathematiques
Universite de Rennes I
Campus de Beaulieu
F-35042 Rennes Cedex

Prof. Dr. Monique Dauge
Monique.Dauge@univ-rennes1.fr
I.R.M.A.R.
Universite de Rennes I
Campus de Beaulieu
F-35042 Rennes Cedex

Dr. Penny J. Davies
penny@maths.strath.ac.uk
Department of Mathematics
University of Strathclyde
Livingstone Tower
26, Richmond Street
GB-Glasgow, G1 1XH

Prof. Dr. Leszek Demkowicz
leszek@ices.utexas.edu
Institute for Computational
Engineering and Sciences (ICES)
University of Texas at Austin
ACES 6.332
Austin, TX 78712-1085 – USA

Dr. Toby Driscoll
driscoll@math.udel.edu
Department of Mathematical Sciences
University of Delaware
501 Ewing Hall
Newark, DE 19716-2553 – USA

Prof. Dr. Marcus Grote
grote@math.unibas.ch
Mathematisches Institut
Universität Basel
Rheinsprung 21
CH-4051 Basel

Dr. Georg Hebermehl
hebermeh@wias.berlin.de
Weierstraß-Institut für
Angewandte Analysis und Stochastik
im Forschungsverbund Berlin e.V.
Mohrenstr. 39
D–10117 Berlin

Prof. Dr. Ralf Hiptmair
hiptmair@sam.math.ethz.ch
Seminar für Angewandte Mathematik
ETH-Zentrum
Rämistr. 101
CH-8092 Zürich

Prof. Dr. Ronald H.W. Hoppe
hoppe@math.uni-augsburg.de
rohop@math.uh.edu
Lehrstuhl für Angewandte
Mathematik I
Universität Augsburg
Universitätsstr. 14
D–86159 Augsburg

Prof. Dr. Manfred Kaltenbacher
manfred.kaltenabacher@
lse.e-technik.uni-erlangen.de
Lehrstuhl für Sensorik
Friedrich-Alexander-Universität
Erlangen
Paul-Gordan-Str. 3/5
D–91052 Erlangen

Prof. Dr. Robert Kotiuga
prk@bu.edu
Department of ECE
Boston University
8 Saint Mary's Street
Boston MA 02215 – SA

Prof. Dr. Rainer Kreß
kress@math.uni-goettingen.de
Institut für Numerische
und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16-18
37083 Göttingen

Prof. Dr. Stefan Kurz
Stefan.Kurz@gmx.de
Universität der Bundeswehr Hamburg
-Fachbereich Elektrotechnik-
Holstenhofweg 85
D–22043 Hamburg

Prof. Dr. Ulrich Langer
ulanger@numa.uni-linz.ac.at
Institut für Numerische Mathematik
Johannes Kepler Universität Linz
Altenbergstr. 69
A-4040 Linz

Dr. Paul David Ledger
P.D.Ledger@swansea.ac.uk
Ledger@sam.math.ethz.ch
Seminar for Applied Mathematics
HG E 13.1
ETH Zentrum
CH-8092 Zürich

Dr. Stephanie Lohrengel
lohrenge@math.unice.fr
Laboratoire J.A. Dieudonne
UMR CNRS 6621
Universite de Nice Sophia-Antipolis
Parc Valrose
F-06108 Nice Cedex 2

Prof. Dr. Dalibor Lukas
dalibor.lukas@vsb.cz
Spezialforschungsbereich F013
Institut für Numerische
Mathematik, Johannes-Kepler
Universität Linz
A-4040 Linz

Prof. Dr. Eric Michielssen
michiels@decwa.ece.uiuc.edu
Department of Electrical and
Computer Engineering
453 Everitt Lab.
1406 W. Green
Urbana IL 61801 – USA

Prof. Dr. Peter Monk
monk@math.udel.edu
Department of Mathematical Sciences
University of Delaware
501 Ewing Hall
Newark, DE 19716-2553 – USA

Prof. Dr. Jean-Claude Nedelec
nedelec@cmapx.polytechnique.fr
Centre de Mathematiques Appliquees
UMR 7641 - CNRS
Ecole Polytechnique
F-91128 Palaiseau Cedex

Prof. Dr. Serge Nicaise
serge.nicaise@univ-valenciennes.fr
ISTV, MACS
Universite de Valenciennes
Le Mont Houy
F-59313 Valenciennes Cedex 9

Prof. Dr. Joseph E.P. Pasciak
pasciak@math.tamu.edu
Department of Mathematics
Texas A & M University
College Station, TX 77843-3368 – USA

Prof. Dr. Ilaria Perugia
perugia@dimat.unipv.it
Dipartimento di Matematica
Universita di Pavia
Via Ferrata 1
I-27100 Pavia

Prof. Dr. Francesca Rapetti
frapetti@math.unice.fr
Laboratoire J.A. Dieudonne
UMR CNRS 6621
Universite de Nice Sophia-Antipolis
Parc Valrose
F-06108 Nice Cedex 2

Prof. Dr. Ursula van Rienen
ursula.van-rienen@etechnik.uni-rostock.de
Fakultät für Informatik und
Elektrotechnik
Universität Rostock
Albert-Einstein-Str. 2
D–18051 Rostock

Dipl.Phys. Werner E. Schabert
schabert@math.uni-augsburg.de
Institut für Mathematik der
Universität Augsburg
Universitätsstr. 14
D–86159 Augsburg

Dr. Frank Schmidt
frank.schmidt@zib.de
Konrad-Zuse-Zentrum für
Informationstechnik Berlin (ZIB)
Takustr. 7
D–14195 Berlin

Dr. Joachim Schöberl
js@jku.at
Institut für Mathematik
Universität Linz
Altenberger Str. 69
A-4040 Linz

Dr. Olaf Steinbach
steinbach@mathematik.uni-stuttgart.de
Institut für Angewandte Analysis
und Numerische Simulation
Universität Stuttgart
Pfaffenwaldring 57
D–70569 Stuttgart

Dr. Oliver Sterz
oliver.sterz@iwr.uni-heidelberg.de
IWR Technische Simulation
Universität Heidelberg
Im Neuenheimer Feld 368
D–69120 Heidelberg

Prof. Dr. Fernando Lisboa Teixeira
teixeira@ee.eng.ohio-state.edu
312 Dreese Laboratory
Ohio State University
2015 Neil Avenue
Columbus OH 43210-1272 – USA

Prof. Dr. Alberto Valli
valli@science.unitn.it
Dipartimento di Matematica
Universita di Trento
Via Sommarive 14
I-38050 Povo (Trento)

Prof. Dr. Jun Zou
zou@math.cuhk.edu.hk
Department of Mathematics
The Chinese University of Hong Kong
Shatin
Hong Kong – HONG KONG

Mathematisches Forschungsinstitut Oberwolfach

Report No. 12/2004

Algebraische Gruppen

Organised by
Michel Brion (Grenoble)
Jens Carsten Jantzen (Aarhus)

February 29th – March 6th, 2004

Introduction by the Organisers

The workshop was organized by Michel Brion (Grenoble) and Jens Carsten Jantzen (Aarhus). The schedule comprised twenty talks from a broad range of areas connected to algebraic groups, including but not limited to: arithmetic invariants of algebraic groups, compact Lie groups, embeddings of homogeneous spaces, Hilbert schemes, invariant theory, moduli spaces, nilpotent orbits, quantum groups, quotient singularities, rationality questions, reductivity and reducibility of algebraic groups, Schubert varieties, and symmetric varieties.

Workshop on Algebraische Gruppen

Table of Contents

Abstracts

LS-Galleries, the path model and MV-cycles
Peter Littelmann
(joint work with S. Gaussent)

The talk is a report on joint work [2] with Stéphane Gaussent (Nancy).

The aim of the work is to connect the combinatorics of the path model for representations of a complex semisimple algebraic group G [4] with the work of Mirković and Vilonen [6] on the intersection cohomology of Schubert varieties in the affine Grassmannian $\mathcal{G}$ of its Langlands dual group $\check{G}$.

Recall that $\mathcal{G}$ is the quotient $\mathcal{G} = \check{G}(\mathbb{C}((t)))/\check{G}(\mathbb{C}[[t]])$. As $\check{G}(\mathbb{C}[[t]])$–variety, $\mathcal{G}$ decomposes [3] into the disjoint union of orbits $\mathcal{G}_\lambda = \check{G}(\mathbb{C}[[t]]).\lambda$, where λ runs over all dominant characters of G (= co-characters of $\check{G}$).

The closure $X_\lambda = \overline{\mathcal{G}_\lambda}$ of such an orbit is a finite dimensional projective variety (in terms of Kac–Moody groups, it is a Schubert variety). The intersection cohomology of this variety is closely connected with the irreducible representation $V(\lambda)$ of G of highest weight λ. Lusztig [5] has shown that the Poincaré series of the stalks of the intersection cohomology sheaf in a point $x \in \mathcal{G}_\mu$, $\mu \preceq \lambda$, coincides with a q–version of the weight multiplicity of μ in $V(\lambda)$. Mirković and Vilonen construct in [6] a canonical basis of $I\!H^\bullet(X_\lambda)$, represented by certain cycles called MV–cycles in the following. This explicit basis has been used by Vasserot in [7] to construct an action of G on $I\!H^\bullet(X_\lambda)$ such that the latter is an irreducible representation of highest weight λ.

In our combinatorial setting, the language of paths is replaced by the language of galleries in an apartment, and LS-paths are replaced by LS–galleries. The translation between the two settings is rather straightforward.

Consider a Demazure–Hansen–Bott–Samelson desingularization $\hat{\Sigma}(\lambda)$ of X_λ. If λ is regular, fixing such a desingularization is equivalent to fixing a minimal gallery γ_λ joining the origin and λ. The homology of $\hat{\Sigma}(\lambda)$ has a basis given by Białynicki–Birula cells, which are indexed by the T–fixed points in $\hat{\Sigma}(\lambda)$. The connection with galleries is obtained as follows: by [1], the points of $\hat{\Sigma}(\lambda)$ can be identified with galleries of type γ_λ in the affine Tits building associated to $\check{G}$, and the T–fixed points correspond in this language to galleries of type γ_λ in the apartment fixed by the choice of T. We show that the retraction from $-\infty$ of the building onto the apartment induces on the level of galleries a map from $\hat{\Sigma}(\lambda)$ onto the set of galleries of type γ_λ, such that the fibres are precisely the Białynicki–Birula cells. We determine those galleries γ such that the associated cell has a non-empty intersection S_γ with $\mathcal{G}_\lambda$ (identified with an open subset of $\hat{\Sigma}(\lambda)$), and we show that the closure $\overline{S_\gamma} \subset X_\lambda$ is a MV-cycle if and only if γ is a LS-gallery. The galleries can also be used to derive more information about the cycles (dimension, affine open subsets of the form $\mathbb{C}^a \times (\mathbb{C}^*)^b, \ldots$).

References

[1] C. Contou-Carrère, *Géométrie des groupes semi-simples, résolutions équivariantes et lieu singulier de leurs variétés de Schubert*, Thèse d'état, (1983), Université Montpellier II.

[2] S. Gaussent and P. Littelmann, *LS-Galleries, the path model and MV-cycles*, Preprint arXiv:math.RT/0307122 v2.

[3] N. Iwahori and H. Matsumoto, *On some Bruhat decomposition and the structure of the Hecke rings of p−adic Chevalley groups*, Publ. I.H.E.S., **25**, (1965).

[4] P. Littelmann, *Paths and root operators in representation theory*, Annals of Mathematics **142**, (1995), 499–525.

[5] G. Lusztig, *Singularities, character formulas, and a q-analog of weight multiplicities*, Astérisque **101-102**, (1982), 208–229.

[6] I. Mirkovic and K. Vilonen, *Perverse sheaves on loop Grassmannians and Langlands duality*, preprint arXiv:math.RT/0401222.

[7] E. Vasserot, *On the action of the dual group on the cohomology of perverse sheaves on the affine Grassmannian*, preprint arXiv:math.AG/0005020v1 (2000).

Algebras with trace and Clebsch-Gordan coefficients for quantum groups

Corrado De Concini

(joint work with Procesi, Reshetikhin, Rosso)

We have reported on joint work with Procesi, Reshetikhin and Rosso. All rings will be algebras over a field of characteristic zero. An associative algebra with trace, over a commutative ring A is an associative algebra R with a 1-ary operation

$$t : R \to R$$

which is assumed to satisfy the following axioms:

(1) t is $A-$linear.
(2) $t(a)b = b\,t(a), \quad \forall a, b \in R.$
(3) $t(ab) = t(ba), \quad \forall a, b \in R.$
(4) $t(t(a)b) = t(a)t(b), \quad \forall a, b \in R.$

This operation is called a formal trace. We denote $t(R) := \{t(a),\ a \in R\}$ the image of t. We have:

1) $t(R)$ is an $A-$subalgebra and that t is also $t(R)-$linear.
2) $t(R)$ is in the center of R.
3) t is 0 on the space of commutators $[R, R]$.

The basic example is the algebra of $n \times n$ matrices over a commutative ring B. For the algebra of matrices one has the Cayley Hamilton theorem:

Every matrix M satisfies its characteristic polynomial $\chi_M(t) := \det(t - M)$.

The main remark that allows to pass to the formal theory is that, in characteristic 0, there are universal polynomials $P_i(t_1, \ldots, t_i)$ with rational coefficients, such that:

$$\chi_M(t) = t^n + \sum_{i=1}^{n} P_i(tr(M), \dots, tr(M^i))t^{n-i}.$$

Thus we can consider the Cayley Hamilton polynomial of an element in an arbitrary algebra with trace and we are led to make the following.

Definition 1. An algebra with trace R is said to be an $n-$Cayley Hamilton algebra, or to satisfy the n^{th} Cayley Hamilton identity if:
1) $t(1) = n$.
2) $\chi_a^n(a) = 0, \ \forall a \in R$.

A structure of Cayley Hamilton algebra can be given in the following situation. Let A be a domain and assume that $A \subset R$ and R is an $A-$module of finite type. Furthermore assume that A is integrally closed in its quotient field F.

Set $S := R \otimes_A F$. S is a finite dimensional division ring.

Let Z denote the center of S Set $\dim_Z S = h^2$ and $p := [Z : F] = \dim_F Z$.

Consider the $F-$linear operator $a^L : S \to S, \ a^L(b) := ab$ and put $t_{S/F}(a) = \frac{1}{hk} tr(a^L)$.

Theorem 1. *The reduced trace $t_{S/F}$ maps R into A, so we will denote by $t_{R/A}$ the induced trace.*

The algebras R, S with their reduced trace are $n-$Cayley Hamilton algebras of degree $n = hp = [S : F] = [R : A]$ (we set $[R : A] := [S : F]$).

Assume now that we have two domains $R_1 \subset R_2$ over two commutative rings $A_1 \subset A_2 \subset R_2$, that each R_i is finitely generated as A_i module and that the two rings A_i are integrally closed. We thus have the two reduced traces t_{R_i/A_i}. We say that the two algebras are compatible if denoting by Z_1 the center of R_1, $Z_1 \otimes_{A_1} A_2$ is a domain.

Theorem 2. *Given two compatible algebras $R_1 \subset R_2$ we have that for a positive integer r:*

$$r\, t_{R_1/A_1} = t_{R_2/A_2} \quad \text{on } R_1$$

One can give various applications of these ideas. One is the following. Let $\mathfrak{g}$ denote a semisimple Lie algebra and let U_ε be the quantized enveloping algebra with deformation parameter specialized at a primitive ℓ-th root of 1 (ℓ odd and prime with 3 if there are G_2 factors). U_ε is a Hopf algebra with comultiplication

$$\Delta : U_\varepsilon \to U_\varepsilon \otimes U_\varepsilon.$$

One knows that the center Z of U_ε contains a Hopf subalgebra Z_0 and that U_ε is a finite free Z_0 module. In particular by taking central characters, one can associate to every irreducible representation V of U_ε an element $\pi(V)$ in the algebraic group $H = \mathrm{Spec} Z_0$. Applying Theorem 2 with $R_1 = \Delta(U_\varepsilon)$, $A_1 = \Delta(Z_0)$, $R_2 = U_\varepsilon \otimes U_\varepsilon$, $A_2 = Z \otimes Z$, we deduce:

Theorem 3. *Given two generic irreducible representations V and W of U_ε with $h = \pi(V),\ k = \pi(W)$,*

$$V \otimes W \simeq \bigoplus_{U \in \pi^{-1}(hk)} U^{\oplus \ell^{(\dim \mathfrak{g} - \mathrm{rk}\, \mathfrak{g})/2}}.$$

Projective normality of complete symmetric varieties

Andrea Maffei

(joint work with R. Chirivì)

The results of this talk have been obtained together with Rocco Chirivì of the University of Pisa.

Let G be an adjoint semisimple algebraic group over $\mathbb{C}$ and $\sigma : G \to G$ an involution of algebraic groups. Denote by H the subgroup of points fixed by σ. A wonderful G-equivariant compactification X of the symmetric variety G/H has been constructed by De Concini and Procesi [5] in characteristic zero and by De Concini and Springer [6] in general. Our main result is the following.

Theorem A ([2])**.** *If $\mathcal{L}$ and $\mathcal{M}$ are line bundles on X, generated by global sections, then the multiplication $\Gamma(X,\mathcal{L}) \otimes \Gamma(X,\mathcal{M}) \to \Gamma(X,\mathcal{L}\otimes\mathcal{M})$ is surjective.*

This generalizes a result of Kannan [8] on the wonderful compactification of groups, and in characteristic zero it answers a question of Faltings [7]. In positive characteristic De Concini has given a counterexample to the same theorem. It is maybe worth observing here that these varieties are Frobenius split (and probably canonical Frobenius split).

We say that a line bundle $\mathcal{L}$ is bigger or equal to a line bundle $\mathcal{M}$ if $\mathcal{L}\otimes\mathcal{M}^{-1}$ is generated by global sections. We call this the dominant ordering. The proof of the theorem is essentially by induction on the dimension of X and on the dominant order on line bundles. Using the description of the boundary of G/H in X given in [5] it is possible to reduce the claim to a few cases controlled by some special triples of weights of a root system that we call “low triples”.

We can give the definition of root system for an abstract root system. Let Φ be a root system, Δ a basis of simple roots and Λ^+ the corresponding monoid of dominant weights and indicate with $\leq$ the dominant order. Given $\lambda,\mu,\nu \in \Lambda^+$, we say that (λ,μ,ν) is a *low triple* if the following conditions hold: (i) if $\lambda',\mu' \in \Lambda^+$ satisfy $\lambda' \leq \lambda$, $\mu' \leq \mu$ and $\nu \leq \lambda'+\mu'$, then $\lambda' = \lambda$, $\mu' = \mu$; (ii) $\nu + \sum_{\alpha\in\Delta} \alpha \leq \lambda+\mu$. We have the following classification which suffices to finish the proof of Theorem A.

Theorem B ([2])**.** *A triple (λ,μ,ν) of dominant weights is a low triple if and only if λ and μ are minuscule weights, $\mu = -w_0\lambda$ and $\nu = 0$, for the longest element w_0 of the Weyl group of Φ.*

Since X is smooth, Theorem A implies that for all line bundles $\mathcal{L}$ generated by global sections the cone over the image of X in $\mathbb{P}(\Gamma(X,\mathcal{L})^*)$ is normal.

Together with Corrado De Concini, we have applied Theorem A to the investigation of normality of cones of other immersions of X. Consider an irreducible representation V of $Lie\, G$ such that H has a fixed point h in $\mathbb{P}(V)$. Let C_V be the cone over the closure of the G-orbit X_V through h. The natural map $G/H \longrightarrow X_V$ induced by $g \longmapsto gh$ extends to X and defines a line bundle $\mathcal{L}_V$ generated by global sections. We have obtained the following description of the normalization of C_V.

Theorem C ([3]). *The integral closure of the coordinate ring of C_V is the ring $\bigoplus_{n\geq 0} \Gamma(X, \mathcal{L}_V^{\otimes n})$.*

In particular (by Theorem A above and the description of the sections of a line bundle given in [5]), we can classify the representations for which C_V is normal. generalizing the results obtained in [4] in the case of the compactification of a group.

A simple generalization of this result allows us to give a uniform proof of the normality of some classical varieties that appear in Lie theory.

References

[1] R. Chirivì and A. Maffei, *The ring of sections of a complete symmetric variety*, preprint arXiv:math.AG/0204354, in Journal of Algebra **261** (2003), 310–326.

[2] R. Chirivì and A. Maffei, *Projective normality of complete symmetric varieties*, preprint arXiv:math.AG/0206290, to appear in Duke Math. Journal.

[3] R. Chirivì , C. De Concini and A. Maffei, in preparation

[4] C. De Concini, *Normality and non normality of certain semigroups and orbit closures*, preprint.

[5] C. De Concini and C. Procesi, *Complete symmetric varieties*, In: Invariant Theory, Lect. Notes in Math., vol. 996, 1–44, Springer, 1983.

[6] C. De Concini and T. A. Springer, *Compactification of symmetric varieties*, Transform. Groups **4** (2000), no. 2-3, 273–300.

[7] G. Faltings, *Explicit resolution of local singularities of moduli spaces*, J. reine angew. Math. **483** (1997), 183–196.

[8] S. S. Kannan, *Projective normality of the wonderful compactification of semisimple adjoint groups*, Math. Zeit. **239** (2002), no. 4, 673–682.

[9] P. Littelmann, *Contracting modules and standard monomial theory for symmetrizable Kac-Moody algebras*, J. Amer. Math. Soc. **11** (1998), no. 3, 551–567.

Complete reducibility and strong reductivity

Gerhard Röhrle
(joint work with M. Bate and B. Martin)

Abstract. Let G be a connected reductive linear algebraic group and let H be a closed subgroup of G. Our main result shows that H is G-completely reducible if and only if H is strongly reductive in G. As a consequence we provide an affirmative answer to a question posed by J.-P. Serre, whether a normal subgroup of a G-completely reducible subgroup of G is again G-completely reducible. Apart from this we discuss other applications. In particular, we prove a converse to Serre's question, namely that H is G-completely reducible if and only if its normalizer $N_G(H)$ is.

This is a report on joint work with M. Bate and B. Martin.

Let G be a connected reductive linear algebraic group defined over an algebraically closed field k. Let H be a closed subgroup of G. Following R.W. Richardson, we say that H is *strongly reductive in G* provided H is not contained in any proper parabolic subgroup of $C_G(S)$, where S is a maximal torus of $C_G(H)$, [2, Def. 16.1]. Observe that this notion does not depend on the choice of the maximal torus S of $C_G(H)$. Richardson introduced this notion in order to characterize closed orbits for the diagonal action of G on the direct product of a finite number of copies of G or its Lie algebra $\operatorname{Lie} G$, [2, Thm. 16.4]. In [2, Lem. 16.2] Richardson showed that a closed subgroup H of $\mathrm{GL}(V)$ (where V is a finite dimensional k-space) is strongly reductive if and only if V is a semisimple H-module. Our aim is to extend this result to arbitrary reductive groups. For that purpose we require the notion of G-complete reducibility due to J.-P. Serre, [3]. Following Serre, a subgroup H of G is called *G-completely reducible* (G-cr) provided that whenever H is contained in a parabolic subgroup P of G, it is contained in a Levi subgroup of P. In case $G = \mathrm{GL}(V)$ a subgroup H is G-cr exactly when V is a semisimple H-module.

The principal result of this talk is

Theorem 1. *Let G be reductive and suppose H is a closed subgroup of G. Then H is G-completely reducible if and only if H is strongly reductive in G.*

The notion of G-complete reducibility is part of the philosophy developed by J.-P. Serre, J. Tits and others to extend standard results from representation theory to algebraic groups by replacing representations $H \to \mathrm{GL}(V)$ with morphisms $H \to G$, where the target group is an arbitrary reductive algebraic group. Theorem 1 is an example of such an extension.

Using Theorem 1 and existing results on strong reductivity, we immediately get new results on G-complete reducibility.

The following result which follows readily from Theorem 1 and [1, Thm. 2] gives an affirmative answer to a question posed by J.-P. Serre, [3, p. 24]. The special case when $G = \mathrm{GL}(V)$ is just a particular instance of Clifford theory.

Theorem 2. *Let G be reductive and let H be a closed subgroup of G with closed normal subgroup N. If H is G-completely reducible, then so is N.*

Serre proves a converse to Theorem 2 in [3, Property 5] under the assumption that the index of N in H is prime to $\operatorname{char} k$. Examples show that this restriction cannot be removed. For instance, let U be a finite unipotent subgroup of G contained in a Borel subgroup of G. Then, by a construction due to Borel and Tits there exists a parabolic subgroup P of G so that $U \subseteq R_u(P)$. In particular, U is not G-cr, but clearly $U^0 = \{1\}$ is. In Theorem 4 below we give a converse of Theorem 2 without characteristic restrictions but with the additional assumption that H contains the centralizer in G of N. In particular, we derive that a closed subgroup H of G is G-completely reducible if and only if its normalizer $N_G(H)$ is, cf. Corollary 5.

Much of the work in this paper is based on the following result:

Proposition 3. *Let $x_1, \dots, x_n \in G$ (for $n \in \mathbb{N}$) and let H be the subgroup of G (topologically) generated by $x_1, \dots, x_n$. Then H is G-completely reducible if and only if the orbit of $(x_1, \dots, x_n)$ under the diagonal action of G on G^n by simultaneous conjugation is closed.*

Proposition 3 allows us to use methods from geometric invariant theory to study G-completely reducible subgroups. E.g. it is crucial for our next

Theorem 4. *Let H be a closed G-completely reducible subgroup of G and suppose K is a closed subgroup of G satisfying $HC_G(H) \subseteq K \subseteq N_G(H)$. Then K is G-completely reducible.*

The following are immediate consequences of Theorems 2 and 4.

Corollary 5. *Let H be a closed subgroup of G. Then H is G-completely reducible if and only if $N_G(H)$ is.*

Corollary 6. *Let H be a closed subgroup of G. If H is G-completely reducible, then so is $C_G(H)$.*

Time permitting we shall discuss other applications of Theorem 1 and new results for G-completely reducible subgroups of G

Finally, we will indicate Serre's approach to G-complete reducibility by means of the homotopy type of the fixed point subcomplex of the building of G.

References

[1] B. M. Martin, *A normal subgroup of a strongly reductive subgroup is strongly reductive*, J. Algebra **265**, (2003), no. 2, 669–674.

[2] R. W. Richardson, *Conjugacy classes of n-tuples in Lie algebras and algebraic groups*, Duke Math. J. **57**, (1988), no. 1, 1–35.

[3] J.-P. Serre, *The notion of complete reducibility in group theory*, Moursund Lectures, University of Oregon, 1998.

On the canonical embeddings of certain homogeneous spaces

Dmitri A. Timashev

(joint work with I. V. Arzhantsev)

This is a joint work with I. V. Arzhantsev, see [3]. Let G be a connected reductive algebraic group over an algebraically closed field $\Bbbk$ of characteristic 0, and H its closed subgroup. The subgroup H is said to have the *Grosshans property* [1] if the homogeneous space G/H is quasiaffine and the coordinate algebra $\Bbbk[G/H]$ is finitely generated. In this situation among all equivariant open affine embeddings $X \hookleftarrow G/H$ one can distinguish a minimal one $X = \operatorname{Spec} \Bbbk[G/H]$, called the *canonical embedding*. The study of the canonical embedding is a geometric way to examine the properties of the coordinate algebra of G/H.

It is well known [2] that the unipotent radical P_{u} of a parabolic subgroup P of G is a Grosshans subgroup. We study the canonical embeddings of the spaces G/P_{u}. This interesting class of affine varieties includes the universal affine embedding of G/U, where U is a maximal unipotent subgroup of G, the space of linear maps to a symplectic vector space with isotropic image, etc. Our main results include: the description of the orbital decomposition for the canonical embedding $X \hookleftarrow G/P_{\mathrm{u}}$; computing the modality of the G-action; classification of the smooth canonical embeddings; construction of the minimal ambient G-module $V \supset X$ (in the algebraic language this is equivalent to the description of a minimal generating set for $\Bbbk[G/P_{\mathrm{u}}]$).

Our approach works for a wider class of affine embeddings of G/P_{u}. The idea is to consider G/P_{u} as a homogeneous space under $G \times L$, where the Levi subgroup $L \subseteq P$ acts by right translations. It is clear that this $(G \times L)$-action extends to the canonical embedding. More generally, we consider arbitrary $(G \times L)$-equivariant affine embeddings $X \hookleftarrow G/P_{\mathrm{u}}$. Several interesting varieties such as varieties of complexes belong to this class.

Such affine embeddings are classified by finitely generated semigroups S of G-dominant weights having the property that all highest weights of tensor products of simple L-modules with highest weights in S belong to S, too. Furthermore, every choice of the generators $\lambda_1, \dots, \lambda_m \in S$ gives rise to a natural G-equivariant embedding $X \hookrightarrow \mathrm{Hom}(V^{P_{\mathrm{u}}}, V)$, where V is the sum of simple G-modules of highest weights $\lambda_1, \dots, \lambda_m$. The convex cone Σ^+ spanned by S is nothing else but the dominant part of the cone Σ spanned by the weight polytope of $V^{P_{\mathrm{u}}}$. The variety X is normal iff S is the semigroup of all lattice points of Σ^+.

We prove that the $(G \times L)$-orbits in X are in bijection with the faces of Σ whose interiors contain dominant weights, orbit representatives being given by the projectors onto the subspaces of $V^{P_{\mathrm{u}}}$ spanned by eigenvectors of eigenweights in a given face. Also we compute the stabilizers of these points in $G \times L$ and in G, and the modality of the action $G : X$.

These results are applied to canonical embeddings as follows. The semigroup S here consists of all dominant weights, and Σ is the span of the dominant Weyl chamber by the Weyl group of L. The $(G \times L)$-orbits in X are in bijection with the subdiagrams in the Dynkin diagram of G such that no connected component of such a subdiagram is contained in the Dynkin diagram of L. In terms of these diagrams, we compute the stabilizers and the modality of $G : X$.

We prove that the only essential cases of smooth embeddings in the considered class are: $X_1 = G$, $X_2 = \mathrm{Mat}(n, n-1)$, $X_3 = \mathrm{Mat}(n, n)$, all other smooth cases being given by a product construction. The first two examples are canonical embeddings with $P = G$ for X_1; $G = SL(n)$, P the stabilizer of a hyperplane in $\Bbbk^n$ for X_2; $G = P = GL(n)$ for X_3.

The techniques used in the description of affine $(G \times L)$-embeddings of G/P_{u} are parallel to those in the study of equivariant embeddings of reductive groups

[4]. This analogy becomes more transparent in view of the bijection between these affine embeddings $G/P_{\mathrm{u}} \hookrightarrow X$ and algebraic monoids M with the group of invertibles L, given by $X = \operatorname{Spec} \Bbbk[G \times^P M]$.

Finally, returning to the case of the canonical embedding $X \hookleftarrow G/P_{\mathrm{u}}$, we describe the G-module structure on the tangent space of X at the G-fixed point, assuming that G is simply connected simple.
This space is obtained from $\bigoplus_i \operatorname{Hom}(V_i^{P_{\mathrm{u}}}, V_i)$, where V_i are the fundamental simple G-modules, by removing certain summands according to an explicit algorithm. The tangent space at the fixed point is at the same time the minimal ambient G-module for X.

References

[1] F. D. Grosshans, *Observable groups and Hilbert's fourteenth problem*, Amer. J. Math. **95:1** (1973), 229–253.

[2] F. D. Grosshans, *The Invariants of Unipotent Radicals of Parabolic Subgroups*, Invent. Math. **73** (1983), 1–9.

[3] I. V. Arzhantsev and D. A. Timashev, *On the canonical embeddings of certain homogeneous spaces*, arXiv:math.AG/0308201, to appear in AMS Translations.

[4] D. A. Timashev, *Equivariant compactifications of reductive groups*, Sbornik: Mathematics **194:4** (2003), 589–616.

Categorification of Weyl groups and Lie algebras

Raphaël Rouquier

(joint work with J. Chuang)

It is classical that various actions of Weyl groups or Lie algebras on vector spaces come from functors acting on abelian or triangulated categories of algebraic or geometric origin, whose Grothendieck group is that space. We want to explain that the natural transformations between these functors should satisfy certain algebraic relations, leading to a better control of the triangulated categories acted on. Namely, we believe there is a "canonical" categorification of a number of classical algebras or groups, in particular Kac-Moody algebras, Weyl groups, braid groups (a monoidal category given by generators and relations).

In a joint work with Joseph Chuang, we explain the setting for $\mathfrak{sl}_2$, which leads to a construction of equivalences of derived categories between blocks of Hecke algebras of type A.

An $\mathfrak{sl}_2$-*categorification* of an abelian category is the data of adjoint exact endofunctors E and F inducing an $\mathfrak{sl}_2$-action on the Grothendieck group and the data of endomorphisms X of E and T of E^2 satisfying the defining relations of (degenerate) affine Hecke algebras.

We prove a categorified version of the relation $[e, f] = h$. We construct divided powers of E and F and a categorification Θ of the simple reflection (following a construction of Rickard). Our main result is a proof that Θ is a self-equivalence at the level of derived categories.

We construct a minimal categorification of the simple $\mathfrak{sl}_2$-representations and show that the proof of the results above can be reduced to this case of a minimal categorification.

We apply these results to the sum of the module categories of all Hecke algebras of type A at an e-th root of unity in characteristic 0 (there are similar results in characteristic $p > 0$).

Recall that there is an action of $\hat{\mathfrak{sl}}_e$ on the sum of Grothendieck groups of categories of modules over Hecke algebras of type A at an e-th root of unity. The action of the generators e_i and f_i come from exact functors between modules ("i-restriction" and "i-induction"). The adjoint action of the simple reflections of the affine Weyl group can then be categorified as inversible endofunctors of the derived category, since every i leads to an $\mathfrak{sl}_2$-categorification. As a consequence, two blocks in the same affine Weyl group orbit have equivalent derived categories.

McKay equivalence for symplectic quotient singularities

Roman Bezrukavnikov
(joint work with D. Kaledin)

Let K be an algebraically closed field of characteristic 0, let V be a finite-dimensional K-vector space equipped with a non-degenerate skew-symmetric form $\omega \in \Lambda^2(V^*)$, and let $\Gamma \subset Sp(V)$ be a finite subgroup. Suppose that we are given a resolution of singularities of the quotient variety $\pi : X \to V/\Gamma$ such that the symplectic form on the smooth part of V/Γ extends to a non-degenerate closed 2-form $\Omega \in H^0(\Omega^2_X)$. In a joint work with D. Kaledin, see [BK], we prove the following

Theorem. *There exists an equivalence of $\mathcal{O}_V^\Gamma$-linear triangulated categories*

$$D^b(Coh(X)) \cong D^b(Coh^\Gamma(V)).$$

A conjecture of this type was first made by M. Reid [R]; a more general statement was conjectured by A. Bondal and D. Orlov, [BOr, §5].

When $\dim(V) = 2$ such an equivalence is well-known, [KV]; in fact, our argument relies on these results. Recently a similar statement was established by T. Bridgeland, A. King and M. Reid [BKR] for crepant resolutions of Gorenstein quotients of vector spaces of dimension 3. The result of [BKR] does not follow from our theorem, because a symplectic vector space can not be 3-dimensional. Notice though that our additional assumption on the resolution is not restrictive — every crepant resolution X of a symplectic quotient singularity in fact carries a non-degenerate symplectic form (see e.g. [Ka]).

Our proof uses reduction to positive characteristic, and quantization of the symplectic variety X_{k} over a field k of characteristic $p > 0$. Our method is suggested by the results of [BMR], where D-module technique is applied to study representations of simple Lie algebras in positive characteristic.

The key ingredient of the proof is a quantization of X_{k} whose global sections coincide with the standard quantization of $H^0(\mathcal{O}_X) = H^0(V, \mathcal{O}_V)^\Gamma$ (the role of this quantization in our picture is similar to the role played by the (crystalline) differential operators in [BMR]). By quantization we mean a deformation of the structure sheaf $\mathcal{O}_X$ to a sheaf of non-commutative $\mathsf{k}[h]$-algebras $\mathcal{O}_h(X)$ such that the algebra of global sections $H^0(X, \mathcal{O}_h)$ is identified to the subalgebra $\mathcal{W}^\Gamma \subset \mathcal{W}$ of Γ-invariant vectors in the (completed) Weyl algebra $\mathcal{W}$ of the vector space V.

It turns out that the reduction of $\mathcal{O}_h(X)$ at a non-zero value of the deformation parameter h (e.g. at $h = 1$) is an Azumaya algebra on $X_{\mathsf{k}}^{(1)}$ (a parallel statement for the ring of differential operators was discovered by Mirkovic and Rumynin, see [BMR]). The category of modules over the latter is the category of coherent sheaves on some gerb over $X^{(1)}$.

One then argues that the above Azumaya algebra on $X^{(1)}$ is *derived affine*, i.e. the derived functor of global sections provides an equivalence between the derived category of sheaves of modules, and the derived category of modules over its global sections; this algebra of global sections is identified with the algebra W^Γ, where W is the reduction of the Weyl algebra at $h = 1$.

Furthermore, for large p we have a Morita equivalence between W^Γ and $\mathsf{W}\#\Gamma$, the smash-product of W and Γ.
Thus we get an equivalence between $D^b(\mathcal{W}\#\Gamma\text{-mod}^{\mathsf{fg}})$ and the derived category of modules over the Azumaya algebra on $X^{(1)}$. The algebra W is an Azumaya algebra over $V^{(1)}$; thus, roughly speaking, the latter equivalence differs from the desired one by a twist with a certain gerb. We then use the norm map on Brauer groups, and Gabber's Theorem [G] to pass from sheaves over a gerb to coherent sheaves on the underlying variety.

Then the equivalence over k of large positive characteristic is constructed; by a standard procedure we derive the desired statement over a field of characteristic zero.

The above Theorem implies, more or less directly, that any crepant resolution X of the quotient V/Γ is the moduli space of Γ-equivariant Artinian sheaves on V satisfying some stability conditions (what is known nowadays as *G-constellations*).

In the case when $X = Hilb^n(\mathbb{A}^2)$ is the Hilbert scheme of n points on the affine plane our argument reproves some of the results by M. Haiman, which constitute a part of his proof of the $n!$ Conjecture.

Also, our methods were used by Finkelberg and Ginzburg to relate representation theory of (a graded version of) Cherednik double affine Hecke algebra (a.k.a. the symplectic reflection algebra) in characteristic p to geometry of the Hilbert scheme; this is explained in the talk by Victor Ginzburg at this conference.

References

[BK] R. Bezrukavnikov, D. Kaledin, *McKay equivalence for symplectic resolutions of singularities*, preprint math.AG/0401002, to appear in Proceedings of the Steklov Institute.

[BMR] R. Bezrukavnikov, I. Mirković, and D. Rumynin, *Localization of modules for a semisimple Lie algebra in prime characteristic*, math.RT/0205144.

[BOr] A. Bondal and D. Orlov, *Derived categories of coherent sheaves*, Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002), 47–56, Higher Ed. Press, Beijing, 2002.
[BKR] T. Bridgeland, A. King, and M. Reid, *The McKay correspondence as an equivalence of derived categories*, J. Amer. Math. Soc. **14** (2001), 535–554.
[G] O. Gabber, *Some theorems on Azumaya algebras*, The Brauer group (Sem., Les Plans-sur-Bex, 1980), Lecture Notes in Math., vol. 844, Springer, Berlin-New York, 1981.
[H] M. Haiman, *Hilbert schemes, polygraphs, and the Macdonald positivity conjecture*, JAMS **14** (2001), 941–1006.
[Ka] D. Kaledin, *Dynkin diagrams and crepant resolutions of singularities*, math.AG/9903157, to appear in Selecta Math.
[KV] M. Kapranov and E. Vasserot, *Kleinian singularities, derived categories and Hall algebras*, Math. Ann. **316** (2000), 565–576.
[R] M. Reid, *McKay correspondence*, alg-geom/9702016 v3, 1997.

Cherednik algebras and Hilbert schemes in characteristic p

Victor Ginzburg
(joint work with R. Bezrukavnikov, M. Finkelberg)

We prove a localisation theorem for the type A rational Cherednik algebra $H_c = H_{1,c}$ over an algebraic closure of the finite field F_p. In the most interesting special case where the parameter c takes values in F_p, we construct an Azumaya algebra A_c on $Hilb^n$, the Hilbert scheme of n points in the plane, such that the algebra of global sections of A_c is isomorphic to H_c. Our localisation theorem provides an equivalence between the bounded derived categories of H_c-modules and sheaves of coherent A_c-modules on the Hilbert scheme, respectively. Furthermore, we show that the Azumaya algebra splits on the formal completion of each fiber of the Hilbert-Chow morphism. This provides a link between our results and those of Bridgeland-King-Reid and Haiman.

Arithmetic birational invariants of linear algebraic groups over some geometric fields

Boris Kunyavskii
(joint work with M. Borovoi)

We discuss two birational invariants: the set of classes of R-equivalence $G(k)/R$, and the unramified Brauer group $\mathrm{Br}_{nr}\, G$. Our goal is to extend some results from the arithmetic case (where k is a number field) to the case where k is a field of cohomological dimension two. More precisely, for k of one of the following types:

(i) $k = k_0(X)$, $\dim X = 2$, $k_0 = \overline{k_0}$, $\operatorname{char} k_0 = 0$;
(ii) k = fraction field of a 2-dimensional, excellent, henselian local domain with residue field k_0;
(iii) $k = l((t))$, $\mathrm{c.\,d.}(l) = 1$, $\operatorname{char} l = 0$,

which were earlier studied by Colliot-Thélène, Gille, and Parimala, we prove that $G(k)/R$ and $\mathrm{Br}_{nr}\, G/\operatorname{Br} k$ can be expressed through the algebraic fundamental group $\pi_1(G)$. More precisely, in the above set-up, let

$$0 \to Q_G \to P \to \pi_1(G) \to 0$$

be a coflasque resolution of $\pi_1(G)$, that is, P is a permutation $\mathrm{Gal}(\overline{k}/k)$-module, and $\mathrm{H}^1(\Gamma', Q_G) = 0$ for all open $\Gamma' \subset \mathrm{Gal}(\overline{k}/k)$. Then $G(k)/R \cong \mathrm{H}^1(k, F_G)$, where F_G is the k-torus with cocharacter module Q_G, and $\mathrm{Br}_{nr}\, G/\operatorname{Br} k \cong \mathrm{H}^1(k, Q_G^\vee)$, where $Q_G^\vee = \mathrm{Hom}(Q_G, \mathbb{Z})$ is the dual module.

Furthermore, if $G \hookrightarrow V$ is a smooth compactification of G, $N_G = \mathrm{Pic}(V \times_k \overline{k})$ is the Picard module, S_G is the Néron-Severi torus (= the torus with character module N_G), then $G(k)/R \cong \mathrm{H}^1(k, S_G)$. This shows that the group $G(k)/R$ is a birational invariant of G.

To appear in J. of Algebra, 2004 (with an appendix by P. Gille).

A non rational group variety of type E_6
Philippe Gille

Let G/k be a semisimple algebraic group defined over a field k. The question whether the group variety G/k is k-rational (i.e. birational to an affine space) has been investigated for classical groups by several authors: Platonov, Yanchevskii, Merkurjev, Chernousov...

The talk deals with the rationality question for exceptional groups. For trialitarian groups of type D_4 and groups of type F_4, the rationality question is open. Using the Bruhat-Tits theory [T], we have found a simply connected group of type E_6 which is not a k-rational variety. The field k is then a 2-iterated power series field over some Merkurjev's suitable field. The proof of the non-rationality goes by a specialization argument involving Chow groups of 0-cycles on G/k.

References

[T] J. Tits, *Strongly inner anisotropic forms of simple algebraic groups*, J. Algebra **131** (1990), 648–677.

Rational points and curves on flag varieties

Emmanuel Peyre
(joint work with A. Chamber-Loir)

1. Heights

It is well known that there are many analogies between the rational points on a variety V defined over a number field K and the rational curves on a variety V over $\mathbf{C}$ and that one of the simplest way to make these links more precise is to consider rational points on a global field of finite characteristic.

In this talk we shall consider the three settings simultaneously:

(1) Over $\mathbf{Q}$ we may define several natural heights on the projective space, for example the height $H_N : \mathbf{P}^N(\mathbf{Q}) \to \mathbf{R}$ defined by

$$H_N((x_0 : \ldots : x_N)) = \sqrt{x_0^2 + \cdots + x_N^2},$$

if $x_0, \ldots, x_N$ are coprime integers. The corresponding logarithmic height is $h_N = \log H_N$.

More generally, if K is a number field, let M_K be the set of places of K. For any place v of K, we denote by K_v the completion of K for the topology defined by v and the absolute value $|\cdot|_v$ is normalized by $\mathrm{d}(ax)_v = |a|_v \mathrm{d}x_v$ for any Haar measure $\mathrm{d}x_v$. We then choose v-adic norms $\|\cdot\|_v : K_v^{N+1} \to \mathbf{R}$, for example we may define the norm $\|(x_0, \ldots, x_N)\|_v$ as $\sup_{0 \leq i \leq N} |x_i|_v$ if v is a finite place, as $\sqrt{\sum_{i=0}^N x_i^2}$ if K_v is isomorphic to $\mathbf{R}$, and as $\sum_{i=0}^N x_i \overline{x}_i$ if K_v is isomorphic to $\mathbf{C}$. Then $H_N : \mathbf{P}^N(K) \to \mathbf{R}$ is defined by

$$H_N(x_0 : \ldots : x_N) = \prod_v \|(x_0, \ldots, x_N)\|_v$$

and $h_N = \log H_N$.

(2) If $K = \mathbf{F}_q(\mathcal{C})$ where $\mathcal{C}$ is a smooth projective curve of genus g over $\mathbf{F}_q$, then there is a bijection from the set of points in the projective space $\mathbf{P}^N(K)$ to the set $\mathrm{Mor}(\mathcal{C}, \mathbf{P}^N_{\mathbf{F}_q})$. Let us denote by $\tilde{x}$ the image of a point x. Then

$$h_N(x) = \deg(\tilde{x}^*(\mathcal{O}(1)))$$

where $\tilde{x}^*(\mathcal{O}(1))$ belongs to the Picard group of the curve $\mathcal{C}$. We also put $H_N = q^{h_N}$.

(3) Similarly, if $K = k(\mathcal{C})$ where $\mathcal{C}$ is a smooth projective curve over a field k, we define

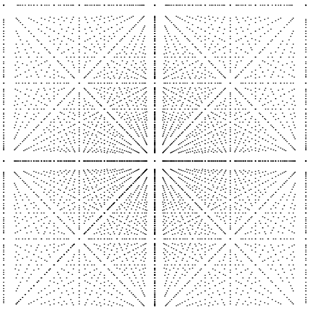
$$h_N(x) = \deg(\tilde{x}^*(\mathcal{O}(1)))$$

where $\tilde{x}^*(\mathcal{O}(1))$ belongs to the Picard group of the curve $\mathcal{C}$.

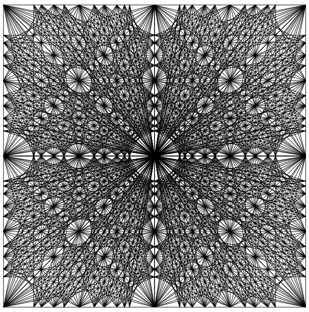
In all settings, if V is a variety over K, any morphism $\phi : V \to \mathbf{P}^N_K$ induces a map $h : V(K) \to \mathbf{R}$ defined by $h = h_N \circ \phi$. We want to study asymptotically the set

$$\{\, x \in V(K) \mid h(x) < \log(B) \,\}$$

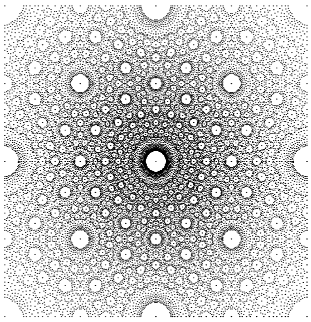
as B goes to $+\infty$. To illustrate this, I represented such sets as points on the projective plane, as lines on the plane and as points in $\mathbf{P}^1_{\mathbf{Q}(i)}$.



$\mathbf{P}^2_{\mathbf{Q}}$



Lines in $\mathbf{P}^2_{\mathbf{Q}}$



$\mathbf{P}^1_{\mathbf{Q}(i)}$

2. Height zeta functions

One of the main tool to study the asymptotic behavior of the number of points of bounded height is the height zeta function.

(1) Over a number field, it is defined for any open subset U of V by

$$\zeta_{U,H}(s) = \sum_{x\in U(K)} \frac{1}{H(x)^s}$$

where this series converges.

(2) Similarly, over $\mathbf{F}_q(T)$, for any open subset U of V

$$Z_{U,h}(T) = \sum_{x\in U(K)} T^{h(x)} \quad \text{and} \quad \zeta_{U,H}(s) = Z_{U,h}(q^{-s}).$$

(3) In the functional setting, we are in fact interested in moduli spaces of morphisms from the curve $\mathcal{C}$ to the variety V. Let $\mathcal{M}_k$ be the group generated by symbols $[V]$ for V variety over k with the relations $[V] = [V']$ if V and V' are isomorphic and

$$[V] = [F] + [V - F]$$

for any closed subset F of V.

If U is an open subset of V, for any integer n, there exists a variety U_n over k such that for any extension k' of k, there is a functorial bijection from $U_n(k')$ to the set of points of $U(k'(\mathcal{C}))$ of height n. The motivic height zeta function is the formal series in $\mathcal{M}_k[[T]]$ defined by

$$Z^{\mathrm{mot}}_{U,h}(T) = \sum_{n\in\mathbf{N}} [U_n] T^n.$$

If k is a finite field, one may go from the functional setting to the classical one by using the map

$$\begin{array}{ccc} \mathcal{M}_k & \to & \mathbf{Z} \\ [V] & \mapsto & \sharp V(\mathbf{F}_q). \end{array}$$

This map sends $Z_{U,h}^{\mathrm{mot}}$ to the classical zeta function $Z_{U,h}$.

3. The case of flag varieties

For flag varieties, one may use the fact, first discovered by Franke, Manin and Tschinkel [FMT] that, in that case, the height zeta function coincides with an Eisenstein series. One may then apply the difficult and deep results obtained for Eisenstein series by Langlands over number fields [Lan], by Harder [Harder] and Morris ([Mo1] and [Mo2]) over global fields of finite characteristic and by Kapranov [Ka] in the functional setting.

Notations 3.1. *Let G be a split semi-simple simply-connected algebraic group over K, let P be a smooth parabolic subgroup of G, let B be a Borel subgroup of G contained in P and let T be a split maximal torus of G contained in B. We denote by Φ the root system of T in G, by Φ^+ the positive roots corresponding to B and by Δ the corresponding basis of the root system. Let Φ_P be the roots of T in the Lie algebra $\mathrm{Lie}(R_u(P))$ of the unipotent radical of P. The set Φ_P is contained in the set of positive roots. We also put $\Delta_P = \Phi_P \cap \Delta$.*

Let $V = G/P$. There exists a canonical isomorphism from the character group $X^*(P)$ of P to $\mathrm{Pic}(V)$ sending the character χ to the line bundle $\mathcal{L}_\chi = G \times^P \mathbf{A}^1_K$ where P acts on the affine line via χ. There is also an injective restriction map $\mathrm{res} : X^*(P) \to X^*(T)$. Let ρ_P (resp. ρ_B) be the half-sum of the roots in Φ_P (resp. Φ^+) then $2\rho_P$ belongs to the image of $X^*(P)$ and we denote also by $2\rho_P$ its inverse image in $X^*(P)$. The line bundle $\mathcal{L}_{2\rho_P}$ is isomorphic to the anticanonical line bundle ω_V^{-1} and is very ample. From now on, all the heights used will be relative to ω_V^{-1}.

(1) In the number field case, it is possible to choose the height on V so that the height zeta function coincides with the value of an Eisenstein series:

$$\zeta_{V,H}(s) = \sum_{x \in G/P(K)} H(x)^{-s} = E_P^G((2s-1)\rho_P, e).$$

Franke, Manin and Tschinkel then applied the work of Langlands and have proven the following results:

- $\zeta_{V,H}(s)$ converges for $\mathrm{Re}(s) > 1$,
- It extends to a meromorphic function on the projective plane,
- It has a pole of order $t = \mathrm{rk}\,\mathrm{Pic}(V)$ at $s = 1$,
- There is a explicit formula for the leading term of the development of $\zeta_{V,H}(s)$ in Laurent series at $s = 1$:

$$\lim_{s\to 1}(s-1)^t \zeta_{V,H}(s) = \prod_{\alpha \in \Phi_P - \Delta_P} \frac{\xi_K(\langle \check\alpha, \rho_B\rangle)}{\xi_K(\langle \check\alpha, \rho_B\rangle + 1)} \prod_{\alpha \in \Delta_P} \frac{\mathrm{res}_{s=1}\,\xi}{\xi_K(2)\langle \check\alpha, 2\rho_P\rangle},$$

where

$$\xi_K(s) = d_K^{s/2}(\pi^{-s/2}\Gamma(s/2))^{r_1}((2\pi)^{-s}\Gamma(s))^{r_2}\zeta_K(s),$$

r_1 being the number of real places of K and r_2 the number of complex places. This limit may be reinterpreted as

$$\lim_{s\to 1}(s-1)^t \zeta_{V,H}(s) = \prod_{\alpha\in\Delta_P} \frac{1}{\langle\check{\alpha}, 2\rho_P\rangle} \boldsymbol{\omega}_H(V(\boldsymbol{A}_K))$$

where $\boldsymbol{\omega}_H$ is a Tamagawa measure on $V(\boldsymbol{A}_K)$.

(2) The connection with Eisenstein series is also valid for global fields of finite characteristic and we may apply the work of Harder and Morris to get the following results:

- $Z_{V,h}(z)$ converges for $|z| < q^{-1}$,
- $Z_{V,h}(T)$ is a rational function,
- $Z_{V,h}(z)$ has a pole of order t at $z = q^{-1}$
- the leading term at $z = q^{-1}$, that is $\lim_{z\to q^{-1}}(z-q^{-1})^t Z_{V,h}(z)$ is given by

$$q^{\dim(V)(1-g)} \prod_{\alpha\in\Phi_P-\Delta_P} \frac{Z_K(q^{-\langle\check{\alpha},\rho_B\rangle})}{Z_K(q^{-\langle\check{\alpha},\rho_B\rangle-1})} \prod_{\alpha\in\Delta_P} \frac{\operatorname{res}_{z=q^{-1}} Z_K}{Z_K(q^{-2})\langle\check{\alpha}, 2\rho_P\rangle}$$

which may be reinterpreted as in the number field case.

(3) In the functional setting, we need to use an extension of $\mathcal{M}_k$ constructed by Denef and Loeser to give an analog to the last assertion. Let $\mathcal{M}_k^{\mathrm{loc}}$ be the ring $\mathcal{M}_k[\mathbf{L}^{-1}]$ and, for any integer n, let $F^n\mathcal{M}_k^{\mathrm{loc}}$ be the subgroup of $\mathcal{M}_k^{\mathrm{loc}}$ generated by the elements of the form $\mathbf{L}^{-i}[V]$ where $i - dim(V) \geq m$. Then $\widehat{\mathcal{M}}_k$ is the completion of $\mathcal{M}_k^{\mathrm{loc}}$ for this filtration.

- The varieties V_n verify:

$$\overline{\lim_{n\to+\infty}} \frac{\dim(V_n)}{n} \leq 1,$$

- $Z_{V,h}^{\mathrm{mot}}(T)$ is a rational function,
- the formal series

$$\Big(\prod_{\alpha\in\Delta_P} \big(1-(\mathbf{L}T)^{\langle\check{\alpha},2\rho_P\rangle}\big)\Big) Z_{V,h}^{\mathrm{mot}}(T)$$

  converges in $\widehat{\mathcal{M}}_k$ at $T = \mathbf{L}^{-1}$,
- at this point it takes the value

$$\mathbf{L}^{\dim(V)(1-g)} \prod_{\alpha\in\Phi_P-\Delta_P} \frac{Z_K^{\mathrm{mot}}(\mathbf{L}^{-\langle\check{\alpha},\rho_B\rangle})}{Z_K^{\mathrm{mot}}(\mathbf{L}^{-\langle\check{\alpha},\rho_B\rangle-1})} \prod_{\alpha\in\Delta_P} \frac{Z_K^{\mathrm{mot}}(T)(1-\mathbf{L}T)(\mathbf{L}^{-1})}{Z_K^{\mathrm{mot}}(\mathbf{L}^{-2})}$$

  where Z_K^{mot} is the zeta function of the field defined by

$$Z_K^{\mathrm{mot}}(T) = \sum_{n\in\mathbf{N}} [\mathcal{C}^{(n)}]T^n,$$

  $\mathcal{C}^{(n)}$ being the n-th symmetric power of $\mathcal{C}$. Kapranov proved that Z_K^{mot} verifies

$$Z_K^{\mathrm{mot}}(T) = \frac{P(T)}{(1-T)(1-\mathbf{L}T)}$$

for a polynomial P in $\mathcal{M}_k[T]$ of degree $2g$ which satisfies a functional equation. Once again this may be interpreted in terms of a Tamagawa number in a motivic setting.

References

[FMT] J. Franke, Y. I. Manin, and Y. Tschinkel, *Rational points of bounded height on Fano varieties*, Invent. Math. **95** (1989), 421–435.

[Harder] G. Harder, *Chevalley groups over function fields and automorphic forms*, Ann. of Math. **100** (1974), 249–306.

[Ka] M. Kapranov, *The elliptic curve in the S-duality theory and Eisenstein series for Kac-Moody groups*, `http://front.math.ucdavis.edu/math.AG/0001005` (2001).

[Lan] R. P. Langlands, *On the functional equations satisfied by Eisenstein series*, Lecture Notes in Math., vol. 544, Springer-Verlag, Berlin, Heidelberg and New York, 1976.

[Mo1] L. E. Morris, *Eisenstein series for reductive groups over global function fields I. The cusp form case*, Can. J. Math. **34** (1982), 91–168.

[Mo2] ———, *Eisenstein series for reductive groups over global function fields II. The general case*, Can. J. Math. **34** (1982), 1112–1182.

Tangent cones to Schubert varieties

Jochen Kuttler

(joint work with J. B. Carrell)

Suppose G is a connected semisimple algebraic group over $\mathbb{C}$. Fix $T \subset B \subset G$, a maximal torus and a Borel subgroup respectively. For every $w \in W = N_G(T)/T$ (which is in one to one correspondence with $(G/B)^T$) denote by $X(w) \subset G/B$ the associated Schubert variety $\overline{Bw}$. Following ideas of Dale Peterson we obtain a generalization to Peterson's ADE-Theorem, which states that in types ADE every rationally smooth point of $X(w)$ is in fact smooth (see [1] for more details). If G has no component of type G_2, we generalize this to: $x \in X(w)^T$ is smooth if and only if for all $y \geq x$ (with respect to the Bruhat-Chevalley ordering on $(G/B)^T$) the reduced tangent cone $\mathcal{T}_y(X(w))$ is linear, ie. its linear span $\Theta_y(X(w))$ satisfies $\dim \Theta_y(X(w)) = \dim X$. Furthermore, we give a method to compute $\Theta_x(X(w))$ at a maximal singularity, provided G has no G_2 factor, which is assumed for the remainder of this text.

For this, let $TE(X(w), x)$ denote the span of tangent lines to the T-stable curves containing x, and let $\mathbb{T}_x(X(w))$ denote the B_x-submodule of $T_x(X(w))$ generated by $TE(X(w), x)$ with $B_x = \{b \in B \mid bx = x\}$. Then

$$TE(X(w), x) \subset \mathbb{T}_x(X(w)) \subset \Theta_x(X(w)).$$

Let μ, ϕ be two negative long orthogonal roots occurring as weights of $TE(X(w), x)$. Then $\{\mu, \phi\}$ is called an *orthogonal B_2-pair* (for $X(w)$ at x) if $\{\mu, \phi\}$ is contained in a copy of $B_2 \subset \Phi$, the roots of G, and if the following holds: suppose α, β are the (unique) positive generators of this B_2 with α short, then we require $r_\alpha x < x$ and $r_\alpha r_\beta x \leq w$. Here $r_\alpha \in W$ is the reflection associated to a root α.

We show that every T-weight of $\Theta_x(X(w))/\mathbb{T}_x(X(w))$ arises as $\frac{1}{2}(\mu + \phi)$ where $\{\mu, \phi\}$ is an orthogonal B_2-pair, whenever x is a maximal singularity.

A second method to compute $\Theta_x(X(w))$ is given by using so called *Peterson translates*: following an idea of Dale Peterson, to a T-stable curve C containing x we associate

$$\tau_C(X(w),x) = \lim_{\substack{z \to x \\ z \neq x}} T_z(X(w)),$$

a T-stable linear subspace of $T_x(X(w))$, which can be computed explicitly at maximal singularities x, whenever $C^T = \{x, y\}$ with $y > x$. We then have

$$\Theta_x(X(w)) = \sum_{\substack{C \\ C^T = \{x,y\} \\ \text{for some } y > x}} \tau_C(X(w),x),$$

where, again, x is a maximal singularity. A more thorough investigation of the schematic tangent cone itself yields a description of the T-fixed points in the fiber of the Nash blowing up over a maximal singular point: each such fixed point is a Peterson translate, provided G is simply laced or the number of T-stable curves containing x equals $\dim X(w)$. This is proved by showing that the blowing up $B_x(S)$ with center x of the (T-stable affine) slice S of $X(w)$ at x is nonsingular, admitting an equivariant surjective map to $N(S)$, the Nash blowing up of S ([3]): Since every T-fixed point in $B_x(S)$ lies on $B_x(C)$ for some T-stable curve $C \subset S$ ([2]), it follows that each T-fixed point of $N(S)$ lies on a T-stable curve which lifts a curve in S, and therefore is a Peterson translate.

References

[1] J. B. Carrell and J. Kuttler, *Smooth points of T-stable varieties in G/B and the Peterson map*, Invent. Math. **151** (2003), no. 2, 353–379.

[2] J. B. Carrell and J. Kuttler, *Singularities of Schubert varieties, tangent cones and Bruhat graphs*, preprint arXiv:math.AG/0404393 (2004).

[3] J. Kuttler, *The Singular Loci of T-Stable Varieties in G/P*, thesis, U. of Basel, 2004

Torsion in intersection cohomology of Schubert varieties

Tom Braden

To a reduced word $a = (s_{i_1}, s_{i_2}, \dots, s_{i_k})$ for an element w in the Weyl group W of a semisimple complex algebraic group G, we can associate a Bott-Samelson variety

$$BS(a) = P_{i_1} \times_B P_{i_2} \times_B \cdots \times_B P_{i_k}/B.$$

The multiplication map π to the flag variety G/B makes $BS(a)$ a resolution of singularities of the Schubert variety $X_w = \overline{BwB/B}$.

We study the question: does the statement of the Decomposition Theorem, proved by Beilinson, Bernstein, Deligne, and Gabber for sheaves with coefficients in $k = \mathbb{Q}$, hold for other coefficient rings k? In other words, is there an isomorphism

$$R\pi_* k_{BS(a)} \cong \bigoplus_\alpha \mathbf{IC}^\bullet(X_{y_\alpha}; k)[n_\alpha], \tag{1}$$

where $y_\alpha \in W$ and $n_\alpha \in \mathbb{Z}$? Soergel has shown that an affirmative answer for k an algebraically closed field of characteristic p would prove Lusztig's conjectured character formula for modular representations of a simply connected split algebraic group of the same type as G, for weights around the Steinberg weight.

We have the following criterion to decide if a splitting (1) is possible. For an element $x \in W$ with $x \leq w$, we let $i_x \colon C_x \to X_w$ be the inclusion.

Theorem 1. *Let a, k be as above. A decomposition of the form* (1) *is possible if and only if every prime p for which p-torsion appears in the cokernel of the natural map*

$$\mathbb{H}^\bullet(i_x^! R\pi_* \mathbb{Z}_{BS(a)}) \to \mathbb{H}^\bullet(i_x^* R\pi_* \mathbb{Z}_{BS(a)}) \tag{2}$$

for some $x \leq w$ is a unit in k.

Furthermore, if $y \in W$, then there is a decomposition (1) *for every $w \leq y$ and every reduced word a for w if and only if for all $w \leq y$, the stalks and costalks of $\mathbf{IC}^\bullet(X_w; k)$ are free k-modules which vanish in odd degrees and whose ranks are given in the usual way by Kazhdan-Lusztig polynomials.*

The theorem that rationally smooth Schubert varieties are smooth in types A, D, E proved by Deodhar [D] (type A), Peterson, and Carrell-Kuttler [CK] might suggest that for these groups the decomposition should hold for all rings and all characteristics. However, we have the following examples where the decomposition theorem fails for $\mathbb{Z}$ or $\mathbb{Z}/2$ coefficients.

Let G be the group $SL_8(\mathbb{C})$; then $W = S_8$ is the symmetric group on 8 letters. Let s_i denote the transposition of i and $i+1$. Then we consider the *hexagon permutations*:

$$w = s_4^c s_3 s_2 s_1 s_5 s_4 s_3 s_2 s_6 s_5 s_4 s_3 s_8 s_7 s_6 s_5 s_4^d,\ c, d \in \{0, 1\}.$$

Billey and Warrington [BW] have shown that avoiding these four permutations along with the longest element $s_1 s_2 s_1$ in S_3 characterizes the permutations for which X_w has a small Bott-Samelson resolution. For this choice of w, we let $x = s_4^c s_2 s_3 s_4 s_5 s_6 s_5 s_4^d$.

For our other example, let G be a group of type D_4. Order the simple reflections $\{s_1, s_2, s_3, s_4\}$ so that s_1, s_2, and s_3 all commute. Then we consider the pair of elements in W: $w = s_1 s_2 s_3 s_4 s_3 s_2 s_1$, and $x = s_1 s_2 s_3$.

Theorem 2. *Let w, x be any of the five pairs described above. Then the cokernel of the map* (2) *has 2-torsion.*

References

[BW] S. Billey and G. Warrington, *Kazhdan–Lusztig Polynomials for 321-hexagon-avoiding permutations*, J. Alg. Comb. **13** (2000), 111–136.

[B] T. Braden, *Bott-Samelson varieties and torsion in intersection cohomology*, in preparation.

[CK] J. Carrell and J. Kuttler, *Smooth points of T-stable varieties in G/B and the Peterson map*, Invent. Math. **151** (2003), no. 2, 353–379.

[D] V. Deodhar, *Local Poincaré duality and nonsingularity of Schubert varieties*, Comm. Algebra **13** (1985), no. 6, 1379–1388.

Normality of nilpotent varieties

Eric Sommers

For the classical groups, Kraft and Procesi have resolved the question of when the closure of a nilpotent orbit is normal, except in the case of some of the orbits for a special orthogonal group which are not invariant under the full orthogonal group. For example, the normality of the closure of an orbit with Jordan block sizes $(4,4,2,2)$ can not be decided using previously known methods.

In this talk we show that these remaining orbits do have normal closure by showing that the regular functions on these orbits are naturally a quotient of the regular functions on an orbit whose closure is known to be normal. Along the way we prove and use a new result concerning the vanishing of the higher cohomology of vector bundles on flag varieties.

References

[1] E. Sommers, *Normality of very even nilpotent varieties in D_{2l}*, submitted to London Math Society, available at http://www.math.umass.edu/~esommers/papers.html

Standard Monomial basis for nilpotent orbit closures

V. Lakshmibai

(joint work with V. Kreiman, P. Magyar, and J. Weyman)

We shall work over the base field $k := \mathbb{C}$. Let $G = GL_n(k), \mathcal{N} =$ the variety of nilpotent $n \times n$ matrices. For the action of G on $\mathcal{N}$ given by conjugation, the G-orbits are indexed by partitions of n. For a partition λ of n, let $\mathcal{N}_\lambda$ denote the corresponding orbit closure. Note that $\mathcal{N} = \mathcal{N}_\lambda$, where $\lambda = (n, 0, \cdots, 0)$. We construct a basis for $k[\mathcal{N}_\lambda]$, the ring of regular functions on $\mathcal{N}_\lambda$. For this construction, we use the following result of Lusztig:

Theorem ([5])**.** *A nilpotent orbit closure $\mathcal{N}_\lambda$ gets identified with an open subset of a certain affine Schubert variety in the affine Grassmannian.*

We first recall the following classical result of Hodge on the Grassmannian.

1. The Grassmannian & its Schubert varieties

Let us fix the integers $1 \le d < n$ and let $V = k^n$. The *Grassmannian $G_{d,n}$* is the set of d-dimensional subspaces $U \subset V$; with respect to a basis $a_1, \ldots, a_d$ of U, where

$$a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \cdots \\ a_{nj} \end{pmatrix}, \text{ with } a_{ij} \in k, \text{for } 1 \le i \le n,\ 1 \le j \le d$$

(here, each vector a_j is written as a column vector with respect to the standard basis of k^n), U may be represented by the $n \times d$ matrix $A = (a_{ij})$ (of rank d), whose columns are the vectors $a_1, \ldots, a_d$.

1.1. **Plücker embedding and Plücker co-ordinates.** Let

$$p : G_{d,n} \to \mathbb{P}(\wedge^d V)$$

be the *Plücker embedding.* Define the set

$$I_{d,n} := \{\underline{i} = (i_1, \ldots, i_d) \,|\, 1 \leq i_1 < \cdots < i_d \leq n\}$$

For $\underline{i} \in I_{d,n}$, the $\underline{i}$-th component of p is denoted by $p_{\underline{i}}$; the $p_{\underline{i}}$'s, with $\underline{i} \in I_{d,n}$, are called the *Plücker coordinates*. If a point U in $G_{d,n}$ is represented by the $n \times d$ matrix A (as above), then $p_{i_1,\ldots,i_d}(U) = det(A_{i_1,\ldots,i_d})$, where $A_{i_1,\ldots,i_d}$ denotes the $d \times d$ sub matrix of A consisting of the rows with indices $i_1, \ldots, i_d$.

1.2. **Schubert Varieties of $G_{d,n}$.** For $1 \leq t \leq n$, let V_t be the subspace of V spanned by $\{e_1, \ldots, e_t\}$. For each $\underline{i} \in I_{d,n}$, the *Schubert variety associated to* $\underline{i}$ is defined to be

$$X_{\underline{i}} = \{U \in G_{d,n} \mid \dim(U \cup V_{i_t}) \geq t, \; 1 \leq t \leq d\}.$$

Remark 1.2.1. We have that under the set-theoretic bijection between the set of Schubert varieties and the set $I_{d,n}$, the partial order on the set of Schubert varieties given by inclusion induces the partial order $\geq$ on $I_{d,n} : \underline{i} \geq \underline{j} \Leftrightarrow i_t \geq j_t, \forall t.$

1.3. **Standard Monomial Basis for Schubert varieties in $G_{d,n}$.** Let R be the homogeneous co-ordinate ring of $G_{d,n}$ for the Plücker embedding, and for $\tau \in I_{d,n}$, let $R(\tau)$ be the homogeneous co-ordinate ring of the Schubert variety $X(\tau)$.

Definition 1.3.1. A monomial $f = p_{\tau_1} \cdots p_{\tau_m}$ is said to be *standard* if

$$(*) \qquad \tau_1 \geq \cdots \geq \tau_m.$$

Such a monomial is said to be *standard on* $X(\tau)$, if in addition to condition (*), we have $\tau \geq \tau_1$.

Theorem 1.3.2 ([2, 3])**.** *Standard monomials on $X(\tau)$ of degree m give a basis for $R(\tau)_m$.*

As a corollary, we obtain that if L denotes the tautological line bundle on $\mathbb{P}(\wedge^d V)$ (as well as its restriction to $X(\tau)$), then the standard monomials on $X(\tau)$ of degree m form a basis for $H^0(X(\tau), L^m)$.

2. The Affine Grassmannian & its Schubert varieties

Let $F = k((t))$, the field of Laurent series, $A = k[[t]]$, the ring of formal power series. Let $G = SL_n(k), B$ the Borel subgroup consisting of upper triangular matrices, and T the maximal torus consisting of diagonal matrices. Let $\mathcal{G} = SL_n(F), \mathcal{P} = SL_n(A), \mathcal{B} = ev^{-1}(B)$, where ev is the evaluation map $ev : SL_n(A) \to SL_n(k), t \mapsto 0$. Let $\hat{W}$ be the affine Weyl group. Then $\mathcal{G}/\mathcal{B}$ is an ind-variety; further, $\mathcal{G}/\mathcal{B} = \cup_{w \in \hat{W}} \mathcal{B}w\mathcal{B}\,(mod\,\mathcal{B})$. Set $X(w) = \cup_{\tau \leq w} \mathcal{B}\tau\mathcal{B}\,(mod\,\mathcal{B})$. Then, $X(w)$ is the *affine Schubert variety* associated to w. Even though, $\mathcal{G}/\mathcal{B}$ is infinite dimensional, $X(w)$ is a finite dimensional projective variety. Similarly, one defines Schubert varieties inside the *affine Grassmannian* $\mathcal{G}/\mathcal{P}$.

2.1. **Sketch of the construction of the basis:** The first step in our construction is to give a matrix presentation for the elements of $\mathcal{G}/\mathcal{P}$, the affine Grassmannian (as in the case of the classical Grassmannian). Once we have a matrix presentation, then we could talk about Plücker co-ordinates on the affine Grassmannian; we could then define standard monomials in the Plücker co-ordinates on the affine Grassmannian (as well as on its Schubert varieties) similar to the classical situation.

Let $k^\infty = \overline{span}\{e_i, i \in \mathbb{Z}\}$, where by $\overline{span}$, we mean that we allow for rightward-infinite linear combinations. For $i \in \mathbb{Z}$, let $E_i = \overline{span}\{e_j, j \geq i\}$. For a subset S such that $\mathbb{N}_p \supset S \supset \mathbb{N}_q$, for some $p, q \in \mathbb{Z}$, let $E_S = \overline{span}\{e_j, j \in S\}$ (here, $\mathbb{N}_p = \{p, p+1, \cdots\}$); note that $E_p \supset E_S \supset E_q$. Define Gr_∞, the *infinite Grassmannian* $= \{$subspaces $E \subset k^\infty | \, E_p \supset E \supset E_q,$ for some $p, q \in \mathbb{Z}\}$. Let $GL_\infty = \{A = (a_{ij})_{\mathbb{Z}\times\mathbb{Z}} \,|$ all but a finite number of $a_{ij} - \delta_{ij}$ are 0, and $detA \neq 0\}$. Then GL_∞ acts transitively on Gr_∞, the isotropy at $E_1 (= \overline{span}\{e_j, j \geq 1\})$ being a certain parabolic subgroup P_∞. Let $\sigma : k^\infty \to k^\infty$ be the map, $e_j \mapsto e_{j+n}$. Define $\hat{Gr}_n$, the *affine Grassmannian* $= \{\sigma-$stable subspaces of $k^\infty\}$.

Identify $k^\infty \cong F^n, e_j \mapsto t^c e_i$, where c and i are given by $j = i + nc, 0 \leq i \leq n-1$ (here, we denote the standard basis for F^n by $\{e_0, \cdots, e_{n-1}\}$). Let $\tau : F^n \to F^n$ be the map $v \mapsto tv$. Via this identification, $\hat{Gr}_n$ may be identified with $\{A-$lattices in $F^n\}$.

Connected components of $\hat{Gr}_n$:

For $i \in \mathbb{Z}$, let $\hat{Gr}^i_n := \{A-$lattices $L \,|\, dim_k(L/L \cap L_0) = dim_k(L_0/L \cap L_0)\}$, where L_0 is the A-lattice $Ae_0 \oplus \cdots \oplus Ae_{n-1}$. The $\hat{Gr}^i_n$'s, $i \in \mathbb{Z}$ give the connected components of $\hat{Gr}_n$. We have an identification $\mathcal{G}/\mathcal{P} \cong \hat{Gr}^0_n, g\mathcal{P} \mapsto gL_0 (= Age_0 \oplus \cdots \oplus Age_{n-1})$. Thus via the embedding $\mathcal{G}/\mathcal{P} \hookrightarrow Gr_\infty$, we obtain a $\mathbb{Z} \times \mathbb{Z}$ matrix presentation for elements of $\mathcal{G}/\mathcal{P}$. Let

$$Y_0 = \{S \,|\, \mathbb{N}_p \supset S \supset \mathbb{N}_q, |S \setminus \mathbb{N}| = |\mathbb{N} \setminus S|\}.$$

For $S \in Y_0$, let p_S denote the Plücker co-ordinate on $\hat{Gr}^0_n$ defined in the obvious way.

Definition 2.1.1. Let $S = (s_1 < s_2 < \cdots) \in Y_0$. We say, S is *admissible* if $s_{i+1} - s_i \leq n$.

Theorem 2.1.2 (cf.[1])**.** $\{p_S, S$ admissible$\}$ *is a basis for* $H^0(\mathcal{G}/\mathcal{P}, L), L$ *being the basic line bundle on* $\mathcal{G}/\mathcal{P}$.

We define operators $\{e_\alpha, f_\alpha, \alpha$ simple$\}$ similar to Kashiwara's crystal operators (cf. [4]), using which we associate to each $S \in Y_0$ a canonical pair $(\lceil S \rceil > \lfloor S \rfloor)$ of elements of $\hat{W}$, and call these respectively the *ceiling* and the *floor* of S.

Definition 2.1.3. A monomial $p_{S_1} \cdots p_{S_m}, S_i \in Y_0$ is standard on $X(\tau)$ if $\tau \geq \lceil S_1 \rceil \geq \lfloor S_1 \rfloor \geq \lceil S_2 \rceil \geq \cdots \geq \lfloor S_m \rfloor$.

Theorem 2.1.4. *Monomials in* p_S*'s standard on* $X(\tau)$ *of degree* m *form a basis of* $H^0(X(\tau), L^m)$.

As a corollary, we obtain a basis for $k[\mathcal{N}_\lambda]$.

References

[1] Date-Jimbo-Kuniba-Miwa-Okado, *Paths, Maya diagrams and representations of $\hat{sl}(r,\mathbb{C})$*, Advanced studies in Pure Math. **19** (1989), Integrable systems in quantum field theory and statistical mechanics, 149–191.

[2] W. V. D. Hodge, *Some enumerative results in the theory of forms*, Proc. Camb. Phil. Soc. **39** (1943), 22–30.

[3] W. V. D. Hodge and D. Pedoe, *Methods of Algebraic Geometry, Vol.II*, Cambridge University Press, (1952).

[4] M. Kashiwara, *Crystalizing the q-anaiogue of universal enveloping algebras*, Comm. Math. Phys. **133** (1990), 249–260.

[5] G. Lusztig, *Canonical bases arising from quantized enveloping algebras*, Journal AMS **3** (1990), 447–498.

Singularities of moduli spaces of vector bundles in char. 0 and char. p

Vikram Mehta

(joint work with V. Balaji)

We study the singularities of the moduli spaces of vector bundles on curves. These moduli spaces over arbitrary base have been constructed by Seshadri. If $U_W \to W$ is the relative moduli space over W, then $U_W \otimes_W \overline{K}$ gives the "correct" moduli space over $\overline{K}$, as this is a flat base change. We prove that $U_W/(p)$ also gives the correct moduli space in characteristic p.

This is achieved by studying the action of $\mathrm{Aut}(V)$ on the local moduli space. The key point is that these local moduli spaces have a good filtration relative to $\mathrm{Aut}(V)$. This enables us to conclude that these invariants in char. 0 specialize to the invariants in char. p. Also we prove that the moduli spaces in char. p are strongly F-regular and consequently, the moduli spaces in char. 0 have canonical singularities.

We conclude with some remarks on the moduli spaces of principal G-bundles in arbitrary characteristic.

Good quotients for reductive group actions

J. Hausen

Good quotients. In the sequel, G is a reductive complex algebraic group, and X is a normal complex algebraic G-variety. A *good quotient* for X is a morphism $p\colon X \to X/\!\!/G$ to an algebraic space $X/\!\!/G$ such that for every affine étale neighbourhood $V \to X/\!\!/G$ the inverse image $p^{-1}(V)$ is affine, G-invariant, and the structure sheaf $\mathcal{O}_V$ equals the sheaf of invariants $p_*(\mathcal{O}_{p^{-1}(V)})^G$.

We are interested in the family $\mathfrak{F}^G$ of all G-invariant open subsets admitting a good quotient $p\colon U \to U/\!\!/G$. Inside $\mathfrak{F}^G$, we will consider subfamilies with prescribed properties on the quotient space, e.g., the family $\mathfrak{F}^G_{\mathrm{sep}} \subset \mathfrak{F}^G$ of subsets with a separated quotient space, or the family $\mathfrak{F}^G_{\mathrm{qp}} \subset \mathfrak{F}^G$ of subsets with a quasiprojective quotient space.

Let "$*$" stand for a property imposed on the quotient space, e.g., separatedness or quasiprojectivity. Then a set $U_1 \in \mathfrak{F}_*^G$ is called *maximal* if there exists no $U_2 \in \mathfrak{F}_*^G$ containing U_1 as a proper subset such that U_1 is saturated with respect to the quotient map $p_2 \colon U_2 \to U_2 /\!\!/ G$, i.e., satisfies $U_1 = p_2^{-1}(p_2(U_1))$. By a basic result of A. Białynicki-Birula, there are only finitely many maximal $U \in \mathfrak{F}_{\mathrm{sep}}^G$.

The central task of the theory of good quotients is to describe or even construct all maximal sets $U \in \mathfrak{F}_*^G$. In the sequel, we report on some results obtained since the appearance of Białynicki-Birula's survey article [2].

Around Mumford's GIT. Here we consider the family $\mathfrak{F}_{\mathrm{qp}}^G \subset \mathfrak{F}^G$ of open subsets with quasiprojective quotient spaces. In his fundamental book, D. Mumford introduces the notion of a G-linearized line bundle L on X, and to any such L he associates a set of semistable points $X^{ss}(L)$. The main features of this construction are well known:

- $X^{ss}(L) \in \mathfrak{F}_{\mathrm{qp}}^G$ holds, that means that there is a good quotient $X^{ss}(L) \to X^{ss}(L)/\!\!/G$, and the quotient space is quasiprojective;
- if X is smooth and $U \in \mathfrak{F}_{\mathrm{qp}}^G$ is maximal, then $U = X^{ss}(L)$ for some G-linearized line bundle L on X;
- for projective X and ample L, the Hilbert–Mumford Criterion characterizes semistability in terms of one parameter subgroups $\mathbb{C}^* \to G$;
- for projective X, the sets $X^{ss}(L)$ with L ample correspond order reversingly to the cones of a fan subdivision of the G-ample cone, see [4] and [9].

In the case of a projective X and an ample L, the quotient space $X^{ss}/\!\!/L$ is projective, but even for smooth projective X, there may exist $U \in \mathfrak{F}_{\mathrm{qp}}^G$ with $U/\!\!/G$ projective that do not arise from ample bundles. Moreover, if X is not smooth, then there may exist maximal $U \in \mathfrak{F}_{\mathrm{qp}}^G$ that do not arise from any G-linearized line bundle.

To overcome the latter problem, we propose in [8] the following approach: consider a Weil divisor D on X, the subsemigroup $\Lambda \subset \mathrm{WDiv}(X)$ generated by D, and the data

$$\mathcal{A} := \bigoplus_{E \in \Lambda} \mathcal{O}_X(E), \qquad \widehat{X} := \mathrm{Spec}(\mathcal{A}).$$

Then a G-linearization of D is a certain lifting of the G-action to $\widehat{X}$, and $x \in X$ is semistable, written $x \in X^{ss}(D)$, if there is a G-invariant $f \in \Gamma(X, \mathcal{O}_X(nD))$, where $n > 0$, such that X_f is affine, $x \in X_f$ holds, and D is Cartier on X_f. The basic features are the following:

- $X^{ss}(D) \in \mathfrak{F}_{\mathrm{qp}}^G$ holds, and, conversely, if $U \in \mathfrak{F}_{\mathrm{qp}}^G$ is maximal, then $U = X^{ss}(D)$ for some G-linearized Weil divisor D on X;
- given a maximal torus $T \subset G$, we have a generalized Hilbert-Mumford Criterion

$$X^{ss}(D,G) = \bigcap_{g \in G} g \cdot X^{ss}(D,T).$$

Divisorial quotient spaces. Borelli calls a prevariety Y *divisorial* if any $y \in Y$ admits an affine neighbourhood $Y \setminus D$ with an effective Cartier divisor D on Y. This concept comprises quasiprojective as well as smooth varieties. It turns out that a prevariety is divisorial if and only if it is the quotient of a quasiaffine variety by a free torus action.

Let $\mathfrak{F}^G_{\mathrm{div}} \subset \mathfrak{F}^G$ denote the family of subsets with a divisorial quotient space. We discuss the construction of such sets presented in [5] and [8]. Consider a finitely generated subgroup $\Lambda \subset \mathrm{WDiv}(X)$ of the group of Weil divisors of the G-variety X. Then we have the data

$$\mathcal{A} := \bigoplus_{D \in \Lambda} \mathcal{O}_X(D), \qquad \widehat{X} := \mathrm{Spec}(\mathcal{A}).$$

Similarly as before, a G-linearization of Λ is a certain lifting of the G-action to $\widehat{X}$. A point $x \in X$ is semistable, written $x \in X^{ss}(\Lambda)$, if there is a G-invariant homogeneous $f \in \Gamma(X, \mathcal{O}_X(\Lambda))$ such that X_f is affine with $x \in X_f$, all $D \in \Lambda$ are Cartier on X_f, and almost all $D \in \Lambda$ admit an invertible $h \in \Gamma(X_f, \mathcal{A}_D)^G$. The basic features are

- $X^{ss}(\Lambda) \in \mathfrak{F}^G_{\mathrm{div}}$ holds, and, conversely, if $U \in \mathfrak{F}^G_{\mathrm{div}}$ is maximal, then $U = X^{ss}(\Lambda)$ for some G-linearized group Λ of Weil divisors on X;
- given a maximal torus $T \subset G$, we have a generalized Hilbert-Mumford Criterion:

$$X^{ss}(\Lambda, G) = \bigcap_{g \in G} g \cdot X^{ss}(\Lambda, T).$$

An application of this construction is the following algebraicity criterion for orbit spaces, see [5]: Suppose that X is $\mathbf{Q}$-factorial and that G acts properly. Then the algebraic space X/G is an algebraic variety if and only if the induced action of the Weyl group $W(T)$ on the algebraic variety X/T has an algebraic variety as orbit space.

CombinaTorics. Here, X is a toric variety, and T is a subtorus of the big torus $T_X \subset X$. In this setting, J. Święcicka observed that any maximal $U \in \mathfrak{F}^T_{\mathrm{sep}}$ is already T_X-invariant. Thus, the description of the maximal sets $U \in \mathfrak{F}^T_{\mathrm{sep}}$ becomes a purely toric problem.

A first approach is the language of fans. Let X arise from a fan Σ in the lattice of one parameter subgroups of T_X. Then the maximal $U \in \mathfrak{F}^T_{\mathrm{sep}}$ correspond to the subfans $\Sigma' \subset \Sigma$ that are maximal with the property that any two maximal cones admit a separating linear form, which is invariant under the lattice of one parameter subgroups of the small torus $T \subset T_X$, see e.g. [6].

Another approach is the language of bunches of cones in the character lattice of the small torus, compare [3] and [1]. Suppose that $X = \mathbb{C}^n$ holds. Then T acts via

$$t \cdot z = (\chi_1(t) z_1, \ldots, \chi_n(t) z_n).$$

A weight cone is a cone generated by some of the weights $\chi_i \in \mathrm{Char}(T)$. A *bunch* is a collection Φ of weight cones such that for any arbitrary weight cone σ we have:

$$\sigma \in \Phi \iff \emptyset \neq \mathrm{relint}(\sigma) \cap \mathrm{relint}(\tau) \neq \mathrm{relint}(\tau) \text{ for all } \sigma \neq \tau \in \Phi.$$

The possible bunches are in one to one correspondence with the maximal $U \in \mathfrak{F}^T_{\mathrm{sep}}$. The translation from the language of bunches to the language of fans is based on a linear Gale transformation. Moreover, the language of bunches has an analogue for $X = \mathbb{P}_n$, obtained by replacing the weight cones with weight polytopes, see [2].

Finally, consider an arbitrary **Q**-factorial T-variety X that has the A_2-property, i.e., any two $x, x' \in X$ have a common affine neighborhood in X. Then there are only finitely many maximal $U_1, \dots, U_r \in \mathfrak{F}^T_{A_2}$, and X admits a T-equivariant closed embedding into a toric variety Z such that $U_i = X \cap W_i$ for some maximal open $W_i \subset Z$ having a good quotient $W_i \to W_i /\!\!/ T$ with $W_i /\!\!/ T$ separated, see [6].

Reduction Theorems. Now G is a connected reductive group, and $T \subset G$ is a maximal torus. Consider a given family $\mathfrak{F}^T_*$, let $U \in \mathfrak{F}^T_*$ be maximal, and set

$$W(U) := \bigcap_{g \in G} g \cdot U.$$

Then Białynicki-Birula asks when we have (a) $W(U) \in \mathfrak{F}^G$, or (b) $W(U) \in \mathfrak{F}^G_*$, or even when (c) $W(U)$ is maximal in $\mathfrak{F}^G_*$.

A first couple of results can be derived using Mumford's GIT and the generalizations presented before: Suppose that G is semisimple and that U is invariant under the normalizer $N(T) \subset G$. Then

- for $U \in \mathfrak{F}^T_{\mathrm{proj}}$, one has (c), see [2] for smooth X, and [8] for normal X;
- for $U \in \mathfrak{F}^T_{\mathrm{qp}}$, one has (b), see [8];
- for a maximal $U \in \mathfrak{F}^T_{\mathrm{div}}$, one has (b), see [8].

An interesting question is that of complete quotients — here only for $G = \mathrm{SL}_2$ is something known, see [2]. Further reduction theorems are the following:

- for $U = X \in \mathfrak{F}^T_{\mathrm{sep}}$, one has (a), see [2];
- for $X = \mathbb{P}_n, \mathbb{C}^n$ and $U \in \mathfrak{F}^T_{\mathrm{sep}}$, one has (a), see [2];
- for **Q**-factorial X and $U \in \mathfrak{F}^T_{A_2}$, one has (a), see [7].

References

[1] F. Berchtold, J. Hausen, *Bunches of cones in the divisor class group — A new combinatorial language for toric varieties* Int. Math. Res. Not. Vol. 2004, No. 6, 261–302 (2004).

[2] A. Białynicki-Birula, *Algebraic Quotients* In: R.V. Gamkrelidze, V.L. Popov (Eds.), Encyclopedia of Mathematical Sciences, Vol. 131, 1–82 (2002).

[3] A. Białynicki-Birula, J. Święcicka, *A recipe for finding open subsets of vector spaces with a good quotient*, Colloq. Math. **77**, No. 1, 97–114 (1998).

[4] I.V. Dolgachev, Y. Hu, *Variation of geometric invariant theory quotients*, Publ. Math., Inst. Hautes Etud. Sci. **87**, 5–56 (1998).

[5] J. Hausen, *A Generalization of Mumford's Geometric Invariant Theory*, Documenta Math. **6**, 571–592 (2001).

[6] J. Hausen, *Producing good quotients by embedding into toric varieties*, Semin. Congr. **6**, 193–212 (2002).
[7] J. Hausen, *A general Hilbert-Mumford criterion*, Ann. Inst. Fourier **53**, No. 3, 701–712 (2003).
[8] J. Hausen, *Geometric Invariant Theory based on Weil divisors*, To appear in Compositio Math.
[9] N. Ressayre, *The GIT-equivalence for G-line bundles*, Geom. Dedicata *81*, No. 1–3, 295–324 (2000).

Universal denominators of invariant rings
Harm Derksen

Suppose that R is a finitely generated graded ring and M is a finitely graded module. It is not always true that the denominator of the Hilbert series of M divides the denominator of the Hilbert series of R. (By Hilbert's Syzygy Theorem this is true if R is a polynomial ring.) This observation leads to my notion of "the universal denominator of a module". Usually the universal denominator of a module is the denominator of the Hilbert series but there are exceptions. The universal denominator behaves much nicer. I could present various formulas for universal denominators of Hilbert series. There are various applications:

(a) I can prove a statement that is very close to one of the conjectures of Dixmier about the denominator of the Hilbert series for invariants of binary forms.
(b) In case the coefficients of the Hilbert series have combinatorial interpretation, one can prove properties about these combinatorial numbers. For example, Jerzy Weyman and I proved the polynomiallity of Littlewood-Richardson numbers. The notion of the universal denominator sheds new light on this result.
(c) Universal denominators can be used to bound the denominator of invariant rings. For example, Nolan Wallach's computation of the Hilbert series for the invariants of 4 qubits can be simplified by using better bounds for the denominator.

Alternating signs of quiver coefficients
Anders S. Buch

Let $E_0 \to E_1 \to \cdots \to E_n$ be a sequence of vector bundles and bundle maps over a non-singular variety X. A set of *rank conditions* for this sequence is a collection $r = \{r_{ij}\}$ of non-negative integers, for $0 \leq i \leq j \leq n$, such that r_{ii} is the rank of E_i for every i. This data defines the *quiver variety*

$$\Omega_r = \{x \in X \mid \operatorname{rank}(E_i(x) \to E_j(x)) \leq r_{ij} \ \forall i < j\}\,.$$

I demand that the rank conditions can *occur*, i.e. they describe an orbit in a quiver representation, and that the bundle maps are sufficiently general, so that the quiver variety obtains its expected codimension $d(r) = \sum_{i<j}(r_{i,j-1} - r_{ij})(r_{i+1,j} - r_{ij})$.

In earlier work with Fulton [3], we proved a formula for the cohomology class of the quiver variety Ω_r in the cohomology ring of X. I later generalized this to the following formula for the *Grothendieck class* of Ω_r in the Grothendieck ring of algebraic vector bundles on X [1]:

$$[\mathcal{O}_{\Omega_r}] = \sum_{\mu} c_\mu(r)\, G_{\mu_1}(E_1; E_0) \cdot G_{\mu_2}(E_2; E_1) \cdots G_{\mu_n}(E_n; E_{n-1}) \in K(X)\,.$$

The sum is over sequences μ of partitions, and the elements G_{μ_i} are K-theoretic generalizations of Schur determinants called *stable Grothendieck polynomials.* The *quiver coefficients* $c_\mu(r)$ appearing in this formula are uniquely determined by the fact that the formula is true for all varieties X, as well as the condition that these coefficients do not change when the same number is added to all the rank conditions.

The quiver coefficients are indexed by sequences of partitions μ for which the sum of the weights is greater than or equal to the codimension $d(r)$. The coefficients $c_\mu(r)$ for which $\sum |\mu_i| = d(r)$ also appear in the cohomology formula and are called *cohomological quiver coefficients.* It was conjectured that cohomological quiver coefficients are non-negative and that the general quiver coefficients have signs that alternate with codimension, that is

$$(-1)^{\sum |\mu_i| - d(r)} c_\mu(r) \geq 0\,.$$

The conjecture for cohomological quiver coefficients has been proved by Knutson, Miller, and Shimozono [4]. In my talk (based on [2]) I present a proof of the general conjecture, which furthermore results in a combinatorial formula for K-theoretic quiver coefficients. My main result is a K-theoretic generalization of the *component formula* of [4]. It writes the Grothendieck class of a quiver variety as an alternating sum of products of stable Grothendieck polynomials given by permutations:

$$[\mathcal{O}_{\Omega_r}] = \sum (-1)^{\sum \ell(u_i) - d(r)}\, G_{u_1}(E_1; E_0) \cdot G_{u_2}(E_2; E_1) \cdots G_{u_n}(E_n; E_{n-1})\,.$$

This sum is over certain sequences of permutations $(u_1, \ldots, u_n)$ which I call *KMS-factorizations* for the rank conditions r. Since Lascoux has proved that any stable Grothendieck polynomial given by a permutation is an alternating linear combination of stable Grothendieck polynomials for partitions [5], this formula implies that K-theoretic quiver coefficients have alternating signs.

I also introduce and prove some new variants of the *factor sequences conjecture* from [3], and I prove Knutson, Miller, and Shimozono's conjecture that their *double ratio formula* agrees with the original quiver formulas.

References

[1] A. S. Buch, *Grothendieck classes of quiver varieties*, Duke Math. J. **115** (2002), no. 1, 75–103.

[2] A. S. Buch, *Alternating signs of quiver coefficients*, preprint, 2003.

[3] A. S. Buch and W. Fulton, *Chern class formulas for quiver varieties*, Inv. Math. **135** (1999), 665–687.

[4] A. Knutson, E. Miller, and M. Shimozono, *Four positive formulas for type A quiver polynomials*, preprint, 2003.

[5] A. Lascoux, *Transition on Grothendieck polynomials*, Physics and combinatorics, 2000 (Nagoya), World Sci. Publishing, River Edge, NJ, 2001, 164–179.

On the values of the characters of compact Lie groups

Jean-Pierre Serre

The lecture discussed three loosely related theorems on the characters of a compact Lie group G.

1. A generalization of a theorem of Burnside

Theorem 1. *Let χ be the character of an irreducible complex representation of G. Assume $\chi(1) > 1$. Then there exists an element x of G, of finite order, with $\chi(x) = 0$.*

When G is finite, this is a well-known result of Burnside.

2. Positive characters with mean value equal to 1

Let f be a virtual character of G having the following two properties:

(a) $f(x)$ is real ≥ 0 for every $x \in G$.
(b) The mean value $\langle f, 1\rangle$ of f is equal to 1.

There are many examples of such characters when G is finite (e.g. permutation characters relative to a transitive action). Not so when G is connected. More precisely:

Question. If G is connected and simply connected, is it true that every character f having properties (a) and (b) is equal to $\chi.\bar{\chi}$, where χ is an irreducible (complex) character of G?

Theorem 2. *The answer to the question above is* "yes" *when G is of rank 1, i.e. when $G = \mathbf{SU}_2(\mathbf{C})$.*

3. The character of the adjoint representation

Consider the adjoint representation $\mathrm{Ad} : G \to \mathrm{Aut}\ (\mathrm{Lie}\ G)$.

Theorem 3. *One has* $\mathrm{Tr}\ \mathrm{Ad}(x) \geq -\ \mathrm{rank}(G)$ *for every* $x \in G$.

The minimal value of $\mathrm{Tr}\ \mathrm{Ad}(x)$ can be determined explicitly:

Choose a maximal torus T of G ; let N be its normalizer and let W be the quotient N/T (so that W is the Weyl group if G is connected). For any $w \in W$, let $\mathrm{Tr}_T(w)$ be the trace of w acting on Lie T. Theorem 3 can be refined as:

Theorem 3$'$. *One has* $\inf_{x\in G} \mathrm{Tr}\ \mathrm{Ad}(x) = \inf_{w\in W} \mathrm{Tr}_T(w)$.

This shows in particular that inf Tr Ad(x) is *an integer*, a fact which was not *a priori* obvious. It also shows that inf Tr Ad(x) is equal to $-$ rank(G) if and only if W contains an element which acts on T by $t \mapsto t^{-1}$.

When G is connected and simple, Theorem 3$'$ gives:

$$\inf\ \mathrm{Tr}\ \mathrm{Ad}(x) = -\ \mathrm{rank}(G) \ \text{ if } G \text{ is of type } A_1, B_n, C_n, D_n\ (n \text{ even}), G_2, F_4, E_7, E_8\ ,$$

$$\inf\ \mathrm{Tr}\ \mathrm{Ad}(x) = \begin{cases} -1 & \text{if } G \text{ is of type } A_n\ (n \geq 1) \\ 2-n & \text{if } G \text{ is of type } D_n\ (n \text{ odd} \geq 3) \\ -3 & \text{if } G \text{ is of type } E_6\ . \end{cases}$$

4. Proofs

They are not published yet. Here are some hints for the interested reader :

Theorem 1: Use the properties of the “principal A_1 subgroup” of G.

Theorem 2: An exercise on positive-valued trigonometric polynomials.

Theorem 3$'$: If $w \in W$ is such that $\mathrm{Tr}_T(w)$ is minimum, any representative x of w in N is such that Tr Ad$(x) = \mathrm{Tr}_T(w)$; this proves the inequality inf Tr Ad$(x) \leq$ inf $\mathrm{Tr}_T(w)$. The opposite inequality can be checked by a case-by-case explicit computation. The classical groups are easy enough, but F_4, E_6, E_7 and E_8 are not (especially E_6, which I owe to Alain Connes). I hope there is a better proof.

References

[1] W. Burnside *On an arithmetical theorem connected with roots of unity and its application to group characteristics*, Proc. London Math. Soc. **1** (1903), 112–116.

Reporter: Jochen Kuttler

Participants

Dr. Annette A'Campo-Neuen
annette@math-lab.unibas.ch
Mathematisches Institut
Universität Basel
Rheinsprung 21
CH-4051 Basel

Prof. Dr. Henning Haahr Andersen
mathha@imf.au.dk
Matematisk Institut
Aarhus Universitet
Ny Munkegade
Universitetsparken
DK-8000 Aarhus C

Prof. Dr. Eva Bayer-Fluckiger
eva.bayer@epfl.ch
Institut de Mathematiques
Ecole Polytechnique Federale
de Lausanne
MA-Ecublens
CH-1015 Lausanne

Prof. Dr. Roman Bezrukavnikov
roman@math.tau.ac.il
bezrukav@math.northwestern.edu
Dept. of Mathematics
Lunt Hall
Northwestern University
2033 Sheridan Road
Evanston, IL 60208-2730 – USA

Prof. Dr. Thomas C. Braden
braden@math.umass.edu
Dept. of Mathematics & Statistics
University of Massachusetts
Amherst, MA 01003-9305 – USA

Dr. Michel Brion
Michel.Brion@ujf-grenoble.fr
Laboratoire de Mathematiques
Universite de Grenoble I
Institut Fourier
B.P. 74
F-38402 Saint-Martin-d'Heres Cedex

Prof. Dr. Anders Buch
abuch@imf.au.dk
Matematisk Institut
Aarhus Universitet
Ny Munkegade
Universitetsparken
DK-8000 Aarhus C

Dr. Rocco Chirivi
chirivi@dm.unipi.it
Dipartimento di Matematica
Universita di Pisa
Via Buonarroti, 2
I-56127 Pisa

Prof. Dr. Corrado De Concini
deconcin@mat.uniroma1.it
Dipartimento di Matematica
Istituto "Guido Castelnuovo"
Universita di Roma "La Sapienza"
Piazzale Aldo Moro 2
I-00185 Roma

Prof. Dr. Harm Derksen
hderksen@umich.edu
Dept. of Mathematics
The University of Michigan
525 East University Avenue
Ann Arbor, MI 48109-1109 – USA

Prof. Dr. Stephen Donkin
S.Donkin@qmul.ac.uk
School of Mathematical Sciences
Queen Mary College
University of London
Mile End Road
GB-London, E1 4NS

Dr. Philippe Gille
gille@math.u-psud.fr
Philippe.Gille@math.u-psud.fr
Mathematique
Universite Paris Sud (Paris XI)
Centre d'Orsay, Batiment 425
F-91405 Orsay Cedex

Prof. Dr. Victor Ginzburg
ginzburg@math.uchicago.edu
Department of Mathematics
The University of Chicago
5734 South University Avenue
Chicago, IL 60637-1514 – USA

Prof. Dr. William Graham
wag@math.uga.edu
Dept. of Mathematics
University of Georgia
Athens, GA 30602 – USA

Prof. Dr. Günter Harder
harder@mpim-bonn.mpg.de
harder@math.uni-bonn.de
Mathematisches Institut
Universität Bonn
Beringstr. 1
D–53115 Bonn

Prof. Dr. Jürgen Hausen
hausen@mfo.de
Mathematisches Forschungsinstitut
Oberwolfach
Lorenzenhof
D–77709 Oberwolfach

Dr. Lutz Hille
hille@math.uni-hamburg.de
Mathematisches Seminar
Universität Hamburg
Bundesstr. 55
D–20146 Hamburg

Prof. Dr. Jens Carsten Jantzen
jantzen@imf.au.dk
Matematisk Institut
Aarhus Universitet
Ny Munkegade
Universitetsparken
DK-8000 Aarhus C

Prof. Dr. Friedrich Knop
knop@math.rutgers.edu
Department of Mathematics
Rutgers University
Hill Center, Busch Campus
110 Frelinghuysen Road
Piscataway, NJ 08854-8019 – USA

Prof. Dr. Hanspeter Kraft
hanspeter.kraft@unibas.ch
Mathematisches Institut
Universität Basel
Rheinsprung 21
CH-4051 Basel

Prof. Dr. Shrawan Kumar
kumar@math.unc.edu
shrawan@email.unc.edu
Dept. of Mathematics
University of North Carolina
Phillips Hall
Chapel Hill, NC 27599-3250 – USA

Prof. Dr. Boris Kunyavskii
kunyav@macs.biu.ac.il
Department of Mathematics
Bar-Ilan University
52900 Ramat Gan – Israel

Dr. Jochen Kuttler
kuttler@math.unibas.ch
Mathematisches Institut
Universität Basel
Rheinsprung 21
CH-4051 Basel

Prof. Dr. Venkatramani Lakshmibai
laksmibai@neu.edu
lakshmibai@research.neu.edu
Dept. of Mathematics
Northeastern University
567 Lake Hall
Boston, MA 02115-5000 – USA

Prof. Dr. Peter Littelmann
littelmann@math.uni-wuppertal.de
Fachbereich Mathematik
Universität Wuppertal
Gauss-Str. 20
D–42097 Wuppertal

Prof. Dr. Dominique Luna
dluna@ujf-grenoble.fr
Dominique.Luna@ujf-grenoble.fr
Laboratoire de Mathematiques
Universite de Grenoble I
Institut Fourier
B.P. 74
F-38402 Saint-Martin-d'Heres Cedex

Dr. Andrea Maffei
amaffei@mat.uniroma1.it
Dipt. di Matematica
G. Castelnuovo
Universita di Roma "La Sapienza"
P.le A. Moro 5
I-00185 Roma

Prof. Dr. Vikram Mehta
vikram@math.tifr.res.in
Tata Institute of Fundamental
Research
School of Mathematics
Homi Bhabha Road, Colaba
400 005 Mumbai – India

Prof. Dr. Dmitri I. Panyushev
dmitri@panyushev.mccme.ru
panyush@mccme.ru
Independent University of Moscow
Bolshoi Vlasjevskii Pereulok 11
Moscow 119002 – Russia

Prof. Dr. Raman Parimala
parimala@math.tifr.res.in
School of Mathematics
Tata Institute of Fundamental
Research
Homi Bhabha Road
Mumbai 400 005 – India

Dr. Emmanuel Peyre
Emmanuel.Peyre@ujf-grenoble.fr
Laboratoire de Mathematiques
Universite de Grenoble I
Institut Fourier
B.P. 74
F-38402 Saint-Martin-d'Heres Cedex

Dr. Guido Pezzini
pezzini@mat.uniroma1.it
Dipartimento di Matematica
Universita di Roma "La Sapienza"
Istituto "Guido Castelnuovo"
Piazzale Aldo Moro, 2
I-00185 Roma

Prof. Dr. Alexander Premet
sashap@ma.man.ac.uk
Dept. of Mathematics
The University of Manchester
Oxford Road
GB-Manchester M13 9PL

Prof. Dr. Gerhard Röhrle
ger@for.mat.bham.ac.uk
School of Maths and Statistics
The University of Birmingham
Edgbaston
GB-Birmingham, B15 2TT

Prof. Dr. Raphael Rouquier
rouquier@math.jussieu.fr
Inst. de Mathematiques de Jussieu
Universite Paris VII
175 rue du Chevaleret
F-75013 Paris

Prof. Dr. Joachim Schwermer
joachim.schwermer@univie.ac.at
Institut für Mathematik
Universität Wien
Strudlhofgasse 4
A-1090 Wien

Prof. Dr. Jean-Pierre Serre
serre@dma.ens.fr
6, Avenue Montespan
F-75116 Paris

Prof. Dr. Wolfgang Soergel
wolfgang.soergel@math.uni-freiburg.de
Mathematisches Institut
Universität Freiburg
Eckerstr.1
D–79104 Freiburg

Dr. Eric Sommers
esommers@math.umass.edu
Dept. of Mathematics & Statistics
University of Massachusetts
Amherst, MA 01003-9305 – USA

Prof. Dr. Tonny A. Springer
springer@math.uu.nl
Mathematisch Instituut
Universiteit Utrecht
Budapestlaan 6
P. O. Box 80.010
NL-3508 TA Utrecht

Dr. Donna M. Testerman
donna.testerman@epfl.ch
Institut de Mathematiques
Ecole Polytechnique Federale
de Lausanne
MA-Ecublens
CH-1015 Lausanne

Prof. Dr. Dimitri Timashev
timashev@mech.math.msu.su
timashev@mccme.ru
Department of Mathematics
Moscow State University
Vorobevy Gory
119992 Moscow – Russia

Prof. Dr. Michela Varagnolo
Michela.Varagnolo@math.u-cergy.fr
Departement de Mathematiques
Universite de Cergy-Pontoise
Site Saint-Martin
2, avenue Adolphe Chauvin
F-95302 Cergy-Pontoise Cedex

Prof. Dr. Eric Vasserot
eric.vasserot@math.u-cergy.fr
vasserot@math.u-cergy.fr
Departement de Mathematiques
Universite de Cergy-Pontoise
Site Saint-Martin
2, avenue Adolphe Chauvin
F-95302 Cergy-Pontoise Cedex

Prof. Dr. Ernest Boris Vinberg
vinberg@ztel.ru
Department of Mechanics and
Mathematics
Moscow Lomonosov State University
Vorobiovy Gory, GSP-2
Moscow 119992 GSP-2 – Russia

Prof. Dr. Andrei V. Zelevinsky
andrei@neu.edu
Dept. of Mathematics
Northeastern University
567 Lake Hall
Boston, MA 02115-5000 – USA

Mathematisches Forschungsinstitut Oberwolfach

Report No. 13/2004

Discrepancy Theory and Its Applications

Organised by
Bernard Chazelle (Princeton)
William Chen (Sydney)
Anand Srivastav (Kiel)

March 7th – March 13th, 2004

Introduction by the Organisers

The meeting was organized by Bernard Chazelle (Princeton), William Chen (Sydney) and Anand Srivastav (Kiel), and was attended by some twenty participants from over ten countries and three continents.

The purpose of the meeting was to encourage and enhance dialogue and collaboration between the theoretical and practical aspects of discrepancy theory. The topics covered included:

(1) Classical discrepancy theory, including low discrepancy sequences, geometric discrepancy and number theoretical aspects.
(2) Combinatorial discrepancy theory, including coloring of hypergraphs and arithmetic structures.
(3) Algorithms and complexity, including relations of discrepancy theory to derandomization of probabilistic algorithms and pseudorandomness, complexity classes, data structures in computational geometry and applications in combinatorial optimization.
(4) Numerical integration in high dimension and its complexity.

Nineteen talks were presented, including a few of a survey nature as well as others that concentrated on specific recent results. These talks demonstrated the diversity on all four areas and their inter-relationships, as well as the vitality of these areas of research.

The organizers and participants would like to take this opportunity to thank again the "Mathematisches Forschungsinstitut Oberwolfach" for having provided a comfortable and inspiring environment for the meeting and the scientific work.

The pleasant atmosphere and superb facilities contributed to the overall success of the meeting.

We include the abstracts of all the talks in alphabetical order of the speakers.

Workshop on Discrepancy Theory and Its Applications

Table of Contents

Abstracts

Balanced partitions of vector sequences

Imre Bárány

Let $d, N \in \mathbb{N}$. Let $\|\cdot\|$ be any norm on $\mathbb{R}^d$ and $B = \{v \in \mathbb{R}^d : \|v\| \leq 1\}$ its unit ball. Some time ago I proved the following result [2]: Let $v_1, v_2, \ldots, v_N \subseteq B$ be a finite sequence of vectors. Then there are signs $\varepsilon_i \in \{-1, 1\}$ such that

$$\left\| \sum_{i \in [n]} \varepsilon_i v_i \right\| \leq 2d$$

for all $n \in [N] = \{1, 2, \ldots, N\}$. In other words, there is a partition $[N] = I_1 \cup I_2$ such that

$$\left\| \sum_{i \in I_j \cap [n]} v_i - \frac{1}{2} \sum_{i \in [n]} v_i \right\| \leq d$$

for all $n \in [N]$ and $j \in [2]$.

This partitioning version of the theorem was extended to partitions into $r > 2$ classes with error bound $(r-1)d$ in [3]. In my talk, I explained how the factor $(r-1)$ can be replaced by a constant. To state this result precisely, we introduce some convenient notation: Let V be the given sequence of vectors $v_1, v_2, \ldots, v_N$. We use the (non-standard) notation

$$\sum_k V = \sum_{i=1}^{k} v_i.$$

Further, for a subsequence X of V, we define

$$\sum_k X = \sum_{i \leq k,\ v_i \in X} v_i.$$

Theorem 1. *For every sequence $V \subset B$, and for every integer $r \geq 2$, there is a partition of V into r subsequences $X_1, \ldots, X_r$ such that for all k and j,*

$$\sum_k X_j \in \frac{1}{r} \sum_k V + C(r) dB,$$

where $C(r)$ is a constant depending only on r. In particular, $C(2) \leq 1$, $C(3) \leq 1.5$, and $C(r) \leq 2.005$ always.

It is worth mentioning here that the result holds for all norms in $\mathbb{R}^d$. This is due to the fact that proofs use linear dependence among some vectors, with the norm playing very little role. But most likely, much better bounds are valid for particular norms. For instance, it is conjectured that for the $r = 2$ and Euclidean norm case the best bound is of order $\sqrt{d}$.

Moreover, for some norms the bound given above is tight, apart from the precise value of $C(r)$. An example showing this is the ℓ_1 norm, when the sequence is just $e_1, e_2, \dots, e_d$ (the standard basis) and r is much smaller than d.

The proof of the theorem, with further results of this type, is to appear in [1].

References

[1] I. Bárány, B. Doerr. Balanced partitions of vector sequences (to appear).

[2] I. Bárány, V.S. Grinberg. On some combinatorial questions in finite-dimensional spaces. *Linear Algebra Appl.*, 41 (1981), 1–9.

[3] B. Doerr, A. Srivastav. Multicolour discrepancies. *Combinatorics, Probability and Computing*, 12 (2003), 365–399.

Limitations to regularity

József Beck

In 1981 I [1] proved the following "irregularity" result. For every n, there is an n-element point set in the unit square which does not have a balanced 2-coloring in the following quantitative sense: Whatever way one 2-colors the n points red and blue, there is always an axis-parallel rectangle in which the number of red points differs from the number of blue points by at least $(\log n)/100$.

My argument was nonconstructive. I could not provide an explicit example of such an n-set. Note that the usual grid is not good. A chessboard type alternating 2-coloring is so balanced that the two color classes differ by at most one.

Roth [4] later gave the following explicit example. Consider the tilted $\sqrt{n} \times \sqrt{n}$ grid, where the slope is a quadratic irrational like $\sqrt{2}$. An equivalent reformulation of Roth's theorem goes as follows. Given any 2-coloring of the $n \times n$ square lattice ("grid"), there is always a tilted rectangle of slope $\sqrt{2}$, say, such that the number of red points differs from the number of blue points by at least $c \log n$, where $c > 0$ is an absolute constant. Note, however, that the size of the "unbalanced" rectangle cannot be specified in advance.

The following questions arise naturally:

(1) What happens in the case of circles?
(2) Can one specify the radius of the circle in advance?
(3) How about one-sided discrepancy for circles?
(4) Is there any other "natural geometric shape" for which translated copies alone give "unbounded irregularity"?

The first question was basically solved by Schmidt [5] many years earlier. His integral equation method, developed in the late 1960's, can be easily adapted to show that the 2-coloring discrepancy of circles is as large as a power of n, rather than $\log n$. Unfortunately Schmidt's method does not work for circles of fixed radius, and cannot handle one-sided discrepancy.

In the early 1990's I could answer Question 4. My natural shape was a "hyperbola segment". Consider the region between the two curves $y = 1/x$ and $y = -1/x$ where $1 \le x \le n$; I call it the hyperbolic needle of length n. It has area $2 \log n$, and it has the following remarkable extra large irregularity property: Given any 2-coloring of the $n \times n$ square lattice, there is always a translated copy of the hyperbolic needle of length n with slope $\sqrt{2}$, say, in which the number of red points differs from the number of blue points by at least $c \log n$, where $c > 0$ is an absolute constant. Since the area of the hyperbolic needle is $2 \log n$, it means that the irregularity is proportional to the area. This explains the name "extra large irregularity".

I could even prove a one-sided version [3] as follows. Assume that n is even, and the $n \times n$ square lattice has a globally balanced 2-coloring, in the sense that there are $n^2/2$ red points and $n^2/2$ blue points. Assume also that the $n \times n$ square lattice is a torus, so that we can "wrap around" the hyperbolic needles. Then there is always a translated copy of the hyperbolic needle of length n with slope $\sqrt{2}$, say, on the $n \times n$ torus in which the number of red points is more than the number of blue points by at least $c \log n$, where $c > 0$ is an absolute constant.

Note that the theorem holds for hyperbolic needle of any length $\ell < n$. Then the corresponding irregularity is constant times $\log \ell$ instead of $\log n$.

Recently I could give an affirmative answer to Questions 2 and 3. The main result, which answers both questions at the same time, goes as follows: Again assume that n is even, and that the $n \times n$ square lattice has a globally balanced 2-coloring in the sense that there are $n^2/2$ red points and $n^2/2$ blue points. Let R be an arbitrary real number with $2 < R < n/2$. Also assume that the $n \times n$ square lattice is a torus, so that we can "wrap around" the circles of radius R. Then there is always a circle of radius R on the $n \times n$ torus in which the number of red points is more than the number of blue points by at least $c\sqrt{\log R}$, where $c > 0$ is an absolute constant. In the case of varying radius, a weaker result, see [2].

The order $\sqrt{\log R}$ is almost certainly very far from optimal. I conjecture that the truth is a power of R rather than a power of $\log R$, but I do not have the slightest idea how to prove it. But I am not complaining – I was more than happy to prove anything "tending to infinity".

References

[1] J. Beck. Balanced two-colorings of finite sets in the square I. *Combinatorica*, 1 (1981), 327–335.

[2] J. Beck. On a problem of W.M. Schmidt concerning one-sided irregularities of point distributions. *Math. Ann.*, 285 (1989), 29–55.

[3] J. Beck. Randomness in lattice point problems. *Discrete Math.*, 229 (2001), 29–55.

[4] K.F. Roth. On a theorem of Beck. *Glasgow Math. J.*, 27 (1985), 195–201.

[5] W.M. Schmidt. Irregularities of distribution IV. *Invent. Math.*, 7 (1969), 55–82.

Classical discrepancy

William Chen

(joint work with Maxim Skriganov)

Let $\mathcal{P}$ be a distribution of N points in the unit square $[0,1]^2$. For every $\mathbf{x} = (x_1, x_2)$ in $[0,1]^2$, let $Z[\mathcal{P}; B(\mathbf{x})] = |\mathcal{P} \cap B(\mathbf{x})|$ denote the number of points of the distribution $\mathcal{P}$ that fall into the rectangle $B(\mathbf{x}) = [0, x_1) \times [0, x_2)$, and consider the corresponding discrepancy function $D[\mathcal{P}; B(\mathbf{x})] = Z[\mathcal{P}; B(\mathbf{x})] - N x_1 x_2$.

Theorem 1.

(i) *There exists a positive absolute constant c such that for every positive integer N and every distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} > c \log N.$$

(ii) *There exists a positive absolute constant C such that for every integer $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$ such that*

$$\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} < C \log N.$$

The lower bound was obtained by Roth [7] in 1954, while the upper bound was obtained by Davenport [5] in 1956. Indeed, the lower bound can be extended to point distributions in the k-dimensional unit cube for arbitrary $k \geq 2$ without any extra difficulty, as shown in Roth [7] with lower bound $c(k)(\log N)^{k-1}$. However, ideas different from those of Davenport are necessary to extend the upper bound to the k-dimensional unit cube for arbitrary $k \geq 2$.

Much work in connection with the upper bound involves the van der Corput point sets and their generalizations. The van der Corput point set of 2^h points in $[0,1]^2$ is given by

$$\mathcal{P}(2^h) = \{(0.a_1 \ldots a_h, 0.a_h \ldots a_1) : a_1, \ldots, a_h \in \{0,1\}\}, \tag{1}$$

where we have used digit expansion base 2 on the right hand side. However,

$$\int_{[0,1]^2} |D[\mathcal{P}(2^h); B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} = 2^{-6} h^2 + O(h), \tag{2}$$

as shown by Halton and Zaremba [6], and so this does not give a proof of the upper bound.

This difficulty was studied in detail by Chen and Skriganov [2], using classical Fourier analysis, since the van der Corput point sets have nice periodicity properties. Recall that $\mathbf{x} = (x_1, x_2)$ denotes the top right vertex of the rectangle $B(\mathbf{x})$. Suppose that $x_1 \neq 1$. Then it can be shown that there exists a finite set $\mathcal{I}(x_1) \subseteq \{1, \ldots, h\}$ such that

$$D[\mathcal{P}(2^h); B(\mathbf{x})] = \sum_{s \in \mathcal{I}(x_1)} \left(c_s - \psi\left(\frac{x_2 + z_s}{2^{s-h}} \right) \right) + O(1).$$

One therefore needs to study sums of the form

$$\sum_{s'\in\mathcal{I}(x_1)}\sum_{s''\in\mathcal{I}(x_1)}\left(c_{s'}-\psi\left(\frac{x_2+z_{s'}}{2^{s'-h}}\right)\right)\left(c_{s''}-\psi\left(\frac{x_2+z_{s''}}{2^{s''-h}}\right)\right).$$

Using Fourier analysis and integrating with respect to the variable x_2 over the interval $[0,1]$, one can show that each of the summands above gives rise to an integral

$$\int_0^1\left(c_{s'}-\psi\left(\frac{x_2+z_{s'}}{2^{s'-h}}\right)\right)\left(c_{s''}-\psi\left(\frac{x_2+z_{s''}}{2^{s''-h}}\right)\right)\mathrm{d}x_2=c_{s'}c_{s''}+O\left(\frac{2^{2\min\{s',s''\}}}{2^{s'+s''}}\right).$$

Unfortunately, the sum

$$\sum_{s'\in\mathcal{I}(x_1)}\sum_{s''\in\mathcal{I}(x_1)}c_{s'}c_{s''}$$

leads to the term $2^{-6}h^2$ in (2).

There are various ways of overcoming this difficulty. In Roth [8], one uses a translation variable t and translates the point set $\mathcal{P}(2^h)$ vertically modulo 1 to obtain the point set $\mathcal{P}(2^h;t)$ and a corresponding discrepancy function

$$D[\mathcal{P}(2^h;t);B(\mathbf{x})]=\sum_{s\in\mathcal{I}(x_1)}\left(\psi\left(\frac{z_s+t}{2^{s-h}}\right)-\psi\left(\frac{w_s+t}{2^{s-h}}\right)\right)+O(1),$$

where z_2 and w_2 are constants that depend on x_2. Squaring and integrating with respect to the variable t over the interval $[0,1]$, we now handle integrals of the form

$$\int_0^1\psi\left(\frac{z_{s'}+t}{2^{s'-h}}\right)\psi\left(\frac{z_{s''}+t}{2^{s''-h}}\right)\mathrm{d}t=O\left(\frac{2^{2\min\{s',s''\}}}{2^{s'+s''}}\right).$$

In Chen [1], one uses digit translations to modify the point set $\mathcal{P}(2^h)$ horizontally to obtain the point set $\mathcal{P}(2^h;\chi)$ and a corresponding discrepancy function

$$D[\mathcal{P}(2^h;\chi);B(\mathbf{x})]=\sum_{s\in\mathcal{I}(x_1)}\left(c_s(\chi)+\psi\left(\frac{x_2+z_s(\chi)}{2^{s-h}}\right)\right)+O(1).$$

Squaring and integrating with respect to the variable x_2 over the interval $[0,1]$ and being economical with the truth, we now essentially handle integrals of the form

$$\begin{aligned}&\int_0^1\left(c_{s'}(\chi)+\psi\left(\frac{x_2+z_{s'}(\chi)}{2^{s'-h}}\right)\right)\left(c_{s''}(\chi)+\psi\left(\frac{x_2+z_{s''}(\chi)}{2^{s''-h}}\right)\right)\mathrm{d}x_2\\&\quad=c_{s'}(\chi)c_{s''}(\chi)+O\left(\frac{2^{2\min\{s',s''\}}}{2^{s'+s''}}\right).\end{aligned}$$

Furthermore, over a large collection of digit translations χ, the sum

$$\sum_{s'\in\mathcal{I}(x_1)}\sum_{s''\in\mathcal{I}(x_1)}c_{s'}(\chi)c_{s''}(\chi)$$

has a small average. However, both of these involve probabilistic variables, and so no explicit point sets $\mathcal{P}$ satisfying the conclusion of Theorem 1(ii) are obtained.

The van der Corput point sets (1) also possess nice group structure. Clearly $\mathcal{P}(2^h)$ forms a group under coordinatewise and digitwise addition modulo 2, and is isomorphic to the direct product $\mathbb{Z}_2^h$. This observation immediately invites the use of Fourier-Walsh functions and series. The discussion can be conducted in general in base p, where p is a fixed prime. In other words, we consider the generalization of the classical van der Corput point sets $\mathcal{P}(2^h)$ to sets of the form

$$\mathcal{P}(p^h) = \{(0.a_1 \ldots a_h, 0.a_h \ldots a_1) : a_1, \ldots, a_h \in \{0, 1, \ldots, p-1\}\},$$

where we now use digit expansion base p on the right hand side. Clearly $\mathcal{P}(p^h)$ forms a group of p^h elements under coordinatewise and digitwise addition modulo p, and is isomorphic to the direct product $\mathbb{Z}_p^h$. This suggests the use of Fourier-Walsh functions and series base p. Using the abbreviation $\mathcal{P}$ for the point set $\mathcal{P}(p^h)$, one can show that an approximation $D_h[\mathcal{P}; B(\mathbf{x})]$ of the discrepancy function $D[\mathcal{P}; B(\mathbf{x})]$ satisfies

$$D_h[\mathcal{P}; B(\mathbf{x})] = \sum_{\substack{\ell_1=0 \\ (\ell_1,\ell_2)\neq(0,0)}}^{p^h-1} \sum_{\ell_2=0}^{p^h-1} \left(\sum_{\mathbf{p}\in\mathcal{P}} \overline{w_{\ell_1}(p_1) w_{\ell_2}(p_2)} \right) \widetilde{\chi}_{\ell_1}(x_1) \widetilde{\chi}_{\ell_2}(x_2).$$

Here w_ℓ, where $\ell \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, denotes the ℓ-th base p Walsh function, and $\widetilde{\chi}_\ell(x)$ denotes the ℓ-th coefficient of Fourier-Walsh series of the characteristic function of the interval $[0, x)$. Since the Walsh functions are characters of the group $\mathcal{P}$, the orthogonality relationship

$$\sum_{\mathbf{p}\in\mathcal{P}} w_{\ell_1}(p_1) w_{\ell_2}(p_2) = \begin{cases} p^h & \text{if } (\ell_1, \ell_2) \in \mathcal{P}^\perp, \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathcal{P}^\perp \subseteq \mathbb{N}_0^2$ is the orthogonal dual to the group $\mathcal{P}$, gives

$$D_h[\mathcal{P}; B(\mathbf{x})] = p^h \sum_{(\ell_1,\ell_2)\in\mathcal{P}^\perp\setminus\{(0,0)\}} \widetilde{\chi}_{\ell_1}(x_1) \widetilde{\chi}_{\ell_2}(x_2).$$

One would like to square this expression and then integrate with respect to $\mathbf{x} = (x_1, x_2)$ over the unit square $[0, 1]^2$. Unfortunately, the Fourier-Walsh coefficients

$$\widetilde{\chi}_{\ell_1}(x_1) \widetilde{\chi}_{\ell_2}(x_2), \quad (\ell_1, \ell_2) \in \mathcal{P}^\perp \setminus \{(0,0)\}, \tag{3}$$

are not orthogonal in $L^2([0, 1]^2)$ in general. In Chen and Skriganov [3], it is shown that as long as the prime p is chosen large enough, there exist groups $\mathcal{P}$ of p^h elements in the square $[0, 1]^2$, in the spirit of van der Corput, such that the Fourier-Walsh coefficients (3) are quasi-orthonormal in $L^2([0, 1]^2)$. Indeed, they are able to establish Theorem 1(ii) for arbitrary dimensions with explicitly constructed point sets. More recently, Chen and Skriganov [4] have shown that in fact, as long as the prime p is chosen large enough, there exist groups $\mathcal{P}$ of p^h elements in the square $[0, 1]^2$, in the spirit of van der Corput, such that the Fourier-Walsh coefficients (3) are orthogonal in $L^2([0, 1]^2)$, so that

$$\int_{[0,1]^2} |D_h[\mathcal{P}; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} = p^{2h} \sum_{(\ell_1,\ell_2)\in\mathcal{P}^\perp\setminus\{(0,0)\}} \int_{[0,1]^2} |\widetilde{\chi}_{\ell_1}(x_1) \widetilde{\chi}_{\ell_2}(x_2)|^2 \, \mathrm{d}\mathbf{x}.$$

Furthermore, they have shown that, corresponding to the group $\mathcal{P}$ of p^h elements in the square $[0,1]^2$, there is a group $\mathcal{G}$ of order p^{2h} of digit shifts such that

$$\frac{1}{|\mathcal{G}|}\sum_{\mathbf{t}\in\mathcal{G}}\int_{[0,1]^2}|D_h[\mathcal{P}\oplus\mathbf{t};B(\mathbf{x})]|^2\,\mathrm{d}\mathbf{x}=p^{2h}\sum_{(\ell_1,\ell_2)\in\mathcal{P}^\perp\setminus\{(0,0)\}}\int_{[0,1]^2}|\widetilde{\chi}_{\ell_1}(x_1)\widetilde{\chi}_{\ell_2}(x_2)|^2\,\mathrm{d}\mathbf{x}.$$

This is a consequence of the orthogonality relationship

$$\sum_{\mathbf{t}\in\mathcal{G}}\overline{w_{\ell_1'}(t_1)w_{\ell_2'}(t_2)}w_{\ell_1''}(t_1)w_{\ell_2''}(t_2)=\begin{cases}p^{2h} & \text{if } (\ell_1',\ell_2')=(\ell_1'',\ell_2''),\\ 0 & \text{otherwise.}\end{cases}$$

We therefore now have a better understanding of the probabilistic argument of Chen [1].

References

[1] W.W.L. Chen. On irregularities of distribution II. *Quart. J. Math. Oxford*, 34 (1983), 257–279.

[2] W.W.L. Chen, M.M. Skriganov. Davenport's theorem in the theory of irregularities of point distribution. *Zapiski Nauch. Sem. POMI*, 269 (2000), 339–353.

[3] W.W.L. Chen, M.M. Skriganov. Explicit constructions in the classical mean squares problem in irregularities of point distribution. *J. Reine Angew. Math.*, 545 (2002), 67–95.

[4] W.W.L. Chen, M.M. Skriganov. Explicit constructions in the classical mean squares problem in irregularities of point distribution II (preprint).

[5] H. Davenport. Note on irregularities of distribution. *Mathematika*, 3 (1956), 131–135.

[6] J.H. Halton, S.K. Zaremba. The extreme and L_2 discrepancies of some plane sets. *Monatsh. Math.*, 73 (1969), 316–328.

[7] K.F. Roth. On irregularities of distribution. *Mathematika*, 1 (1954), 73–79.

[8] K.F. Roth. On irregularities of distribution IV. *Acta Arith.*, 37 (1980), 67–75.

Multi-color discrepancies

Benjamin Doerr

(joint work with Anand Srivastav)

We extend the notion of combinatorial discrepancy of hypergraphs to arbitrary numbers of colors. Unless otherwise stated, the following results appeared in [5]. Let $\mathcal{H}=(X,\mathcal{E})$ denote a finite hypergraph, *i.e.*, X is a finite set and $\mathcal{E}$ is a family of subsets of X. Put $n=|X|$ and $m=|\mathcal{E}|$. A c-coloring of $\mathcal{H}$ is a mapping $\chi : X\to M$, where M is any set of cardinality c. Usually, we take $M=[c]:=\{1,\ldots,c\}$. The basic idea of measuring the deviation from perfect balance motivates these definitions of the discrepancy of $\mathcal{H}$ with respect to χ and the discrepancy of $\mathcal{H}$ in c colors:

$$\begin{aligned}\operatorname{disc}(\mathcal{H},\chi,c) &:= \max_{i\in M,\, E\in\mathcal{E}}\left||\chi^{-1}(i)\cap E|-\frac{|E|}{c}\right|,\\ \operatorname{disc}(\mathcal{H},c) &:= \min_{\chi:X\to[c]}\operatorname{disc}(\mathcal{H},\chi,c).\end{aligned}$$

Let us start with an example which shows that a hypergraph may have very different discrepancies in different numbers of colors. Let $k \in \mathbb{N}$ and $n = 4k$. Set

$$\mathcal{H}_n = ([n], \{X \subseteq [n] : |X \cap [n/2]| = |X \setminus [n/2]|\}).$$

Obviously, $\mathcal{H}_n$ has 2-color discrepancy zero, but $\operatorname{disc}(\mathcal{H}_n, 4) = \frac{1}{8}n$.

In fact, such examples exist for nearly any two numbers of colors. Unless c_1 divides c_2, there are hypergraphs $\mathcal{H}_n$ on n vertices having discrepancy $\Theta(n)$ in c_1 colors and zero discrepancy in c_2 colors. This has been investigated in [2].

For some 2-color discrepancy results, the proofs seem to rely heavily on the fact that only two colors are used. This applies in particular to those where the partial coloring method introduced by Beck [1] is used. A key step there is to construct a low discrepancy partial coloring $\chi := \frac{1}{2}(\chi_1 - \chi_2)$ from two colorings χ_1, χ_2 with $\chi_1(E) \approx \chi_2(E)$ for all $E \in \mathcal{E}$. It is not clear to us how this idea can be extended to c colors.

The idea of recursive coloring is to successively enlarge the number of partition classes. We start with a suitable 2-coloring of X with color classes X_1, X_2 and then iterate this process on the subhypergraphs induced by X_1 and X_2. If the weighted 2-color discrepancies of the induced subhypergraphs are bounded, such a recursive approach can be analyzed, even if c is not a power of 2. For $p \in [0,1]$, denote the discrepancy of $\mathcal{H}$ with respect to the weight $(p, 1-p)$ by

$$\operatorname{disc}(\mathcal{H}, (p, 1-p)) = \min_{\chi: X \to [2]} \max_{E \in \mathcal{E}} \left| |E \cap \chi^{-1}(1)| - p|E| \right|.$$

Theorem 1. *Let* $\operatorname{disc}(\mathcal{H}_0, (p, 1-p)) \le K$ *for all induced subgraphs* $\mathcal{H}_0$ *of* $\mathcal{H}$ *and all* $p \in [0,1]$. *Then the inequality* $\operatorname{disc}(\mathcal{H}, c) \le 2.0005K$ *holds for all numbers* c *of colors.*

For many classical results, a refinement of the above ideas yields even stronger bounds that decrease for larger numbers of colors. For reasons of space we are not able to state the general result precisely. Roughly speaking, we have that if induced subhypergraphs on n_0 vertices have 2-color discrepancy at most $O(n_0^\alpha)$ for some $\alpha \in]0,1[$, then $\operatorname{disc}(\mathcal{H}, c) = O((n/c)^\alpha)$. This gives, among many others, the following bounds, where in all cases, the implicit constants do not depend on c:

- General bound: $\operatorname{disc}(\mathcal{H}, c) \le 45\sqrt{(n/c)\log(4m)} + 1$.
- Spencer's six standard deviations [7]: For all hypergraphs $\mathcal{H}$ having $n = m$ vertices and hyperedges, $\operatorname{disc}(\mathcal{H}, c) = O(\sqrt{(n/c)\log c})$.
- Arithmetic progressions: The hypergraph $\mathcal{A}_n$ of arithmetic progressions in $[n]$ satisfies $\operatorname{disc}(\mathcal{A}_n, c) = O(c^{-0.16} n^{0.25})$ for $c \le n^{0.25}$. This extends the bound of Matoušek and Spencer [6].

A second general approach is to mimic the proofs of two-color results. Since the choice of the colors ± 1 for two colors allows several powerful arguments, the key problem is to choose a suitable set of colors for the general case. The colors we use are vectors in $\mathbb{R}^c$. We obtain a multi-color analogue of the Beck-Fiala theorem showing that $\operatorname{disc}(\mathcal{H}, c) \le 2\Delta(\mathcal{H})$ and one of the Bárány-Grunberg theorems. The latter was improved by Bárány in his talk by reducing the multiplicative dependence on the number of colors to a constant.

An analogue of an eigenvalue bound attributed to Lovász and Sós shows that

$$\mathrm{disc}(\mathcal{H}, c) \geq \sqrt{\frac{n(c-1)}{mc^2}\lambda_{\min}(A^\top A)},$$

where $A^\top$ is an incidence matrix of A. This can be used to show a lower bound of $0.04c^{-1/2}n^{1/4}$ for the c-color discrepancy of the arithmetic progressions in $[n]$.

For hypergraph having $n = m$ vertices and edges, using a random construction we recently showed that our upper bound in Spencer's six standard deviations is sharp apart from constant factors [4].

Theorem 2. *For all $c \in \mathbb{N}_{\geq 2}$ and $n \geq c \log c$, there is a hypergraph having n vertices, n hyperedges and c-color discrepancy at least $\frac{1}{40}\sqrt{(n/c)\log c}$.*

In contrast to the (ordinary) c-color discrepancy, there is a strong correlation between the hereditary discrepancies of a hypergraph in different numbers of colors.

Theorem 3. *For any two numbers of colors $c_1, c_2 \in \mathbb{N}_{\geq 2}$ and all hypergraphs $\mathcal{H}$, we have*

$$\mathrm{herdisc}(\mathcal{H}, c_2) \leq 3c_1^2 \,\mathrm{herdisc}(\mathcal{H}, c_1).$$

Hence $\mathrm{herdisc}(\cdot, c_2) = \Theta_{c_1,c_2}(\mathrm{herdisc}(\cdot, c_1))$. The proof given in [3] actually solves a more general problem, namely it reduces the color rounding problem in c_2 colors to the hereditary discrepancy problem in c_1 colors. We currently have no purely combinatorial proof.

References

[1] J. Beck. Roth's estimate of the discrepancy of integer sequences is nearly sharp. *Combinatorica*, 1 (1981), 319–325.

[2] B. Doerr. Discrepancy in different numbers of colors. *Discrete Math.*, 250 (2002), 63–70.

[3] B. Doerr. The hereditary discrepancy is nearly independent of the number of colors. *Proc. Amer. Math. Soc.* (to appear in 2004).

[4] B. Doerr. Lower bounds for multi-color discrepancies (submitted).

[5] B. Doerr, A. Srivastav. Multicolour discrepancies. *Combinatorics, Probability and Computing*, 12 (2003), 365–399.

[6] J. Matoušek and J. Spencer. Discrepancy in arithmetic progressions. *J. Amer. Math. Soc.*, 9 (1996), 195–204.

[7] J. Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289 (1985), 679–706.

Digital expansions and uniformly distributed sequences modulo 1

Michael Drmota

Let $s_q(n)$ denote the sum-of-digits function of the q-ary digital expansion of the non-negative integer n. Then it is well known that the sequence $(s_q(n)\alpha)$ is uniformly distributed modulo 1 if and only if α is irrational. The purpose of this talk is to present a survey of recent results of this kind and also to present the methods that are used. We will deal with the following topics:

(1) Uniform distribution of $(f(n)\alpha)$ for additive functions related to various digital expansions.
(2) Discrepancy bounds for $(s_q(n)\alpha)$ in terms of the continued fraction expansion of α.
(3) Uniform distribution of $(s_{q_1}(n)\alpha_1, .., s_{q_d}(n)\alpha_d)$ for coprime bases $q_1, \dots, q_d$ and irrational $\alpha_1, \dots, \alpha_d$.
(4) Uniform distribution of $(s_q(n)\alpha)_{n\in S}$ for certain subsequences S; for example, squares.

Let $f(n)$ be a q-additive function, *e.g.*, the q-ary sum-of-digits function $s_q(n)$. Then it is worth considering the generating function

$$F_N(x) = \sum_{n<N} x^{f(n)}.$$

Owing to the recursive structure of q-ary digital expansion, one directly gets recurrences for $F_N(x)$ that (usually) lead to (more or less) explicit (or asymptotic) expressions for $F_N(x)$. For example, for the binary sum-of-digits function one has

$$\sum_{n<2^k} x^{s_2(n)} = (1+x)^2.$$

The advantage of these kinds of representation is that they directly imply results on

- the distribution $\#\{n < N : f(n) \le x\}$ as $N \to \infty$ (Gaussian limiting distributions), and
- uniform distribution and discrepancy estimates of the sequence $(f(n)\alpha)$ for irrational numbers α.

For example, in order to treat uniform distribution of $(f(n)\alpha)$, one has to evaluate $F_N(x)$ for $x = \mathrm{e}^{2\pi \mathrm{i} h\alpha}$.

Of course, with the help of this method one obtains upper bounds for exponential sums and for the discrepancy, however, usually not optimal ones. Nevertheless, it is possilbe to get more precise discrepancy estimates by using the continued fraction expansion of α (see [3]). For example, if α has bounded continued fraction expansion, then one gets

$$\frac{1}{\sqrt{\log N}} \ll D_N(s_q(n)\alpha) \ll \frac{\log\log\log N}{\sqrt{\log N}}.$$

It is also an interesting problem to consider d-dimensional sequences

$$(s_{q_1}(n)\alpha_1, \dots, s_{q_d}(n)\alpha_d)$$

for coprime bases $q_1, \dots, q_d$ and irrational $\alpha_1, \dots, \alpha_d$. With the help of exponential sum estimates (see [3]), it follows that these kinds of sequences are uniformly distributed modulo 1 for all irrational numbers $\alpha_1, \dots, \alpha_d$.

In other words, this is a mathematical formulation of the well accepted fact that q-ary digital expansions are independent if the bases $q_1, \dots, q_d$ are coprime. Interestingly one can be even more precise. With the help of methods of Bassily and Katai [1] and by proper use of Baker's theorem on linear forms in logarithms, it follows that the joint distribution of $(s_{q_1}(n), s_{q_2}(n))$ is asympotically Gaussian and independent if q_1 and q_2 are coprime (see [2]). It is even possible to derive asymptotic expansions for the numbers $\#\{n < N : s_{q_1}(n) = k_1,\ s_{q_2}(n) = k_2\}$.

Finally we consider the (binary) sum-of-digits function $s(n^2)$ of squares. There are no precise results on the distribution of squares. For example, it is an unsolved problem of Gelfond [5] whether the asymptotic frequency of $s(n^2)$ being even is 1/2 or not. Equivalently we can ask whether

$$\sum_{n<N} (-1)^{s(n^2)} = o(N)?$$

We could not answer this question. However, in joint work with Rivat [4], the sum of binary digits $s(n^2)$ is split into two parts $s_{[<k]}(n^2)+s_{[\geq k]}(n^2)$, where $s_{[<k]}(n^2) = s(n^2 \bmod 2^k)$ collects the first k digits and $s_{[\geq k]}(n^2) = s(\lfloor n^2/2^k \rfloor)$ collects the remaining digits. With the help of the generating function approach mentioned above, we derive very precise results on the distribution on $s_{[<k]}(n^2)$ and $s_{[\geq k]}(n^2)$. We provide asymptotic formulas for the numbers $\#\{n < 2^k : s_{[<k]}(n^2) = m\}$ and $\#\{n < 2^k : s_{[\geq k]}(n^2) = m\}$ and show that the sequences $(s_{[<k]}(n^2)\alpha)_{n<2^k}$ and $(s_{[\geq k]}(n^2)\alpha)_{n<2^k}$ are very well distributed modulo 1.

References

[1] N.L. Bassily, I. Kátai. Distribution of the values of q-additive functions on polynomial sequences. *Acta Math. Hung.*, 68 (1995), 353–361.
[2] M. Drmota. The joint distribution of q-additive functions. *Acta Arith.*, 100 (2001), 17–39.
[3] M. Drmota, G. Larcher. The sum-of-digits-function and uniform distribution modulo 1. *J. Number Theory*, 89 (2001), 65–96.
[4] M. Drmota, J. Rivat. The sum-of-digits function of squares (manuscript, 2004).
[5] A.O. Gelfond. Sur les nombres qui ont des propriétés additives et multiplicatives données. *Acta Arith.*, 13 (1968), 259–265.

Geometric discrepancies and δ-covers

Michael Gnewuch
(joint work with Benjamin Doerr)

It is of interest to derive bounds for geometric discrepancies, *e.g.*, the $*$- or the unanchored discrepancy, with good behaviour in the parameter of dimension d. An upper bound for the $*$-discrepancy with a nearly optimal behaviour in d and explicitly known constants was proved; see Theorem 1 in [1]. Here we introduce the notion of δ-covers on the d-dimensional unit cube $[0,1]^d$ and give bounds for their minimal cardinality. From these estimates we obtain upper bounds for the $*$-discrepancy and its inverse, which improve the results of [1]. We achieve similar results for the unanchored discrepancy.

For $x, y \in [0,1]^d$, we write $[x, y[= \prod_{i\in[d]} [x_i, y_i[$. Let $\delta > 0$. We say that some finite subset Γ is a δ-cover of $[0,1]^d$ if for all $y \in [0,1]^d$, there are $x, z \in \Gamma \cup \{0\}$

such that $x_i \le y_i \le z_i$ for all i and $\mathrm{vol}([0,z[) - \mathrm{vol}([0,x[) \le \delta$. We denote the minimal cardinality of all δ-covers by $N(d,\delta)$.

We get a first bound on $N(d,\delta)$ by considering an equidistant grid Γ_m with mesh size $1/m$, where $m = \lceil d/\delta \rceil$. Obviously Γ_m is a δ-cover of $[0,1]^d$ with cardinality $(m+1)^d$. We then derive a better bound by calculating the coordinates of a non-equidistant grid $\Gamma = \{x_0, \ldots, x_{\kappa(\delta,d)}\}^d$ with the following recursive procedure:

$$\begin{aligned}
&x_0 := 1\\
&x_1 := (1-\delta)^{1/d}\\
&\qquad \text{for } i \ge 1 \text{ do}\\
&\qquad x_{i+1} := (x_i - \delta)x_1^{1-d}\\
&\qquad \text{if } x_{i+1} \le \delta, \text{ then } \kappa(\delta,d) := i+1 \text{ and stop}
\end{aligned} \tag{1}$$

The sequence $x_0, x_1, \ldots$ is finite and strictly decreasing. Γ is a δ-cover of $[0,1]^d$, which establishes the following bound on $N(d,\delta)$.

Theorem 1. *Let $d \ge 2$ and $0 \le \delta < 1$. Then $N(d,\delta) \le (\kappa(\delta,d)+1)^d$, where*

$$\kappa(\delta,d) = \left\lceil \frac{d}{d-1} \frac{\log(1-(1-\delta)^{1/d}) - \log(\delta)}{\log(1-\delta)} \right\rceil . \tag{2}$$

The estimate

$$\kappa(\delta,d) \le \left\lceil \frac{d}{d-1} \frac{\log d}{\delta} \right\rceil$$

holds, and the quotient of the left and the right hand sides of the inequality converges to 1 as $\delta \to 0$.

Another recursive construction gives us a bound with better asymptotic behaviour in d. The construction in dimension d uses the $(d-1)$-dimensional construction and a scaling property and leads to the next theorem. Note that all O-notation refer to the variable δ^{-1} only.

Theorem 2. *Let $d \ge 2$ and $0 < \delta < 1$. Then, with a constant $C \le 2\mathrm{e}$,*

$$N(d,\delta) \le 2^d \frac{d^d}{d!} \left(\delta^{-1} + \frac{d+1}{4} - \frac{1}{2d} \right)^d \le C^d \delta^{-d} + O(\delta^{-d+1}). \tag{3}$$

A lower bound for the cardinality of each δ-cover is stated in the next theorem.

Theorem 3. *Let $\delta \in]0,1]$. Then, with a constant $c \ge \mathrm{e}^{-1}$,*

$$N(d,\delta) \ge \frac{2}{5} \frac{d!}{d^d} \delta^{-d} - \frac{2}{5} d! \sum_{k=0}^{d-1} \frac{d^k (\log(d\delta)^{-1})^k}{k!} \ge c^d \delta^{-d} + O((\log \delta^{-1})^{d-1}).$$

We now discuss applications to $*$-discrepancy. The L^∞-$*$-discrepancy is given by

$$d^*_\infty(n,d) = \inf_{t_1,\ldots,t_n \in [0,1]^d} d^*_\infty(t_1,\ldots,t_n),$$

where

$$d^*_\infty(t_1, ..., t_n) = \sup_{x\in[0,1]^d} \left| \mathrm{vol}([0,x[) - \frac{1}{n}\sum_{k=1}^{n} 1_{[0,x[}(t_k) \right|.$$

The inverse of the $*$-discrepancy is defined by

$$n^*_\infty(\varepsilon, d) = \min\{n \in \mathbb{N} : d^*_\infty(n, d) \le \varepsilon\}$$

for given $\varepsilon > 0$. For any δ-cover Γ of $[0,1]^d$, the following approximation property holds: For every $t_1, \ldots, t_n \in [0,1]^d$, we have

$$d^*_\infty(t_1, ..., t_n) \le \max_{x\in\Gamma} \left| \mathrm{vol}([0,x[) - \frac{1}{n}\sum_{i=1}^{n} 1_{[0,x[}(t_i) \right| + \delta.$$

Using this and our results on δ-covers, we obtain the following result.

Theorem 4. *Let $d \ge 2$ and $\varepsilon > 0$. If $\varepsilon \le 8/(d+1)$, then there exists a constant $C \le 8\mathrm{e}$, independent of ε and d, with*

$$n^*_\infty(\varepsilon, d) \le \left\lceil 2\varepsilon^{-2}\left(d\log\left(\frac{C}{\varepsilon}\right) + \log 2\right)\right\rceil.$$

For all $0 < \varepsilon \le 1$, we have

$$n^*_\infty(\varepsilon, d) \le \left\lceil 2\varepsilon^{-2}\left(d\log\left(\kappa\left(\frac{\varepsilon}{2}, d\right) + 1\right) + \log 2\right)\right\rceil, \tag{4}$$

where $\kappa(\varepsilon/2, d)$ is defined as in (2). If

$$n \ge 2\left(d\log\left(\left\lceil \frac{2d}{d-1}\log d\right\rceil + 1\right) + \log 2\right),$$

then, with $\rho = 2\sqrt{2\log 2/5} < 1.0532$,

$$d^*_\infty(n, d) \le \sqrt{2}n^{-1/2}(d\log(\lceil \rho n^{1/2}\rceil + 1) + \log 2)^{1/2}.$$

We verified this theorem by adapting the proof idea from Theorem 1 of [1]. The proof considers n equally distributed independent random variables representing the possible point configurations, and in this situation our approximation property above allows us to make use of Hoeffding's inequality. Note that the same proof technique was also employed in [2] and [3].

Using (4), we can give explicit bounds for the inverse of the $*$-discrepancy. Corresponding to the same values of d and ε as in Section 2 of [1], we have the following bounds:

$n^*_\infty(0.45, 5) \le 116$	$n^*_\infty(0.1, 5) \le 3828$
$n^*_\infty(0.45, 10) \le 244$	$n^*_\infty(0.1, 10) \le 8003$
$n^*_\infty(0.45, 20) \le 514$	$n^*_\infty(0.1, 20) \le 16648$
$n^*_\infty(0.45, 40) \le 1103$	$n^*_\infty(0.1, 40) \le 34679$
$n^*_\infty(0.45, 60) \le 1686$	$n^*_\infty(0.1, 60) \le 53020$
$n^*_\infty(0.45, 80) \le 2291$	$n^*_\infty(0.1, 80) \le 71777$

The bounds in [1] were achieved by using the same technique that we adapted in the proof of Theorem 4 and by analysing the average behaviour of the L^p-$*$-discrepancy for even integers p. Our bounds are smaller by factors between 5 and 8.1 than the bounds in [1] that make use of Hoeffding's inequality, and they are still smaller than the bounds resulting from the average L^p-$*$-discrepancy analysis – roughly by a factor 3 for $\varepsilon = 0.45$ and 1.6 for $\varepsilon = 0.1$.

We conclude by making some remarks on unanchored discrepancy. Instead of δ-covers of $[0,1]^d$ we can define δ-covers for characteristic functions of axis-parallel boxes in $[0,1]^d$. This definition is a special case of the more general notion of one-sided (μ, δ)-covers in [3]. We use our results about $N(d, \delta)$ to get bounds for the minimal cardinality of these modified δ-covers, which lead to upper bounds for the unanchored discrepancy. Those bounds are similar to the ones for the $*$-discrepancy in Theorem 4 – more or less, we just have to substitute d by $2d$ in each estimate.

References

[1] S. Heinrich, E. Novak, G.W. Wasilkowski, H. Woźniakowski. The inverse of the star-discrepancy depends linearly on the dimension. *Acta Arith.*, 96 (2001), 279–302.

[2] F.J. Hickernell, I.H. Sloan, G.W. Wasilkowski. On tractability of weighted integration over bounded and unbounded regions in $\mathbb{R}^s$. *Math. Comp.* (to appear).

[3] H.N. Mhaskar. On the tractability of multivariate integration and approximation by neural networks. *J. Complexity* (to appear).

Discrepancy and declustering

Nils Hebbinghaus

(joint work with Benjamin Doerr and Sören Werth)

The declustering problem is to assign data blocks from a multi-dimensional grid system to one of M storage devices in a balanced manner. More precisely, we consider a grid $V = [n_1] \times \ldots \times [n_d]$ for some positive integers $n_1, \ldots, n_d$. Here we use the notations $[n] := \{1, 2, \ldots, n\}$ and $[n..m] := \{n, n+1, \ldots, m\}$ for $n, m \in \mathbb{N}$ with $n \leq m$.

A query Q requests the data assigned to a sub-grid $[x_1..y_1] \times \ldots \times [x_d..y_d]$ for some integers $1 \leq x_i \leq y_i \leq n_i$. We assume that the time to process such a query is proportional to the maximum number of requested data blocks that are stored in a single device. If we represent the assignment of the data blocks to the devices by a mapping $\chi : V \to [M]$, then the query time of the query above is

$$\max_{i \in [M]} |\chi^{-1}(i) \cap Q|,$$

where we identify the query Q with its associated sub-grid. Clearly, no declustering scheme can do better than $|Q|/M$. Hence a natural performance measure is the additive deviation from this lower bound.

Thus the problem turns out to be a combinatorial discrepancy problem in M colors. Denote by $\mathcal{E}$ the set of all sub-grids in V. Then $\mathcal{H} = (V, \mathcal{E})$ is a hypergraph.

For a coloring $\chi : V \to [M]$, the positive discrepancy of $\mathcal{H}$ with respect to χ and the positive discrepancy of $\mathcal{H}$ in M colors are respectively

$$\begin{aligned}\operatorname{disc}^+(\mathcal{H},\chi) &:= \max_{i\in[M],\ E\in\mathcal{E}} \left(|\chi^{-1}(i) \cap E| - \frac{1}{M}|E| \right),\\ \operatorname{disc}^+(\mathcal{H},M) &:= \min_{\chi:V\to[M]} \operatorname{disc}^+(\mathcal{H},\chi).\end{aligned}$$

A similar definition was introduced by Srivastav and the first author in [4]. The only difference is that we regard positive instead of absolute deviations. Independently, Anstee, Demetrovics, Katona and Sali [1] and Sinha, Bhatia and Chen [8] proved a lower bound of $\Omega(\log M)$ for the additive error of any declustering scheme in dimension 2. Sinha *et al.* [8] also gave the bound $\Omega(\log^{(d-1)/2} M)$ for arbitrary $d \geq 3$, but their proof contains a crucial error.

The current best upper bounds in arbitrary dimension for the declustering schemes are proposed by C.-M. Chen and C. Cheng [3]. They present two schemes for d-dimensional problems with an additive error $O(\log^{d-1} M)$. The first one works if $M = p^k$ for some $k \in \mathbb{N}$ and p a prime such that $p \geq d$, whereas the second works for arbitrary M, but the error increases with N.

For the upper bounds, we present an improved scheme that yields an additive error of $O(\log^{d-1} M)$ for a broader range of values of M, which is independent of the data size. Our requirement on M is that if $M = q_1 \dots q_k$, where $q_1 < \dots < q_k$, is the canonical factorization of M into prime powers, we require $d \leq q_1 + 1$. Thus, in particular, our schemes work for M being a power of 2 (such that $M \geq d-1$) and without any restriction on M in dimensions 2 and 3, which is very useful from the viewpoint of application. We also show that the latin hypercube construction used by Chen *et al.* [3] is much better than claimed. Where they show that the latin hypercube coloring extended to the whole grid has an error of at most 2^d times the one of the latin hypercube, we show that both errors are the same.

For the lower bound, we present the first correct proof of the $\Omega(\log^{(d-1)/2} M)$ bound. Again, a more careful analysis shows that the positive discrepancy is at least $1/2d$ times the normal discrepancy instead of 3^{-d} as claimed in [8]. Note that in typical applications with M between 16 and 1024, these 2^d and 3^d factors are at least as important as finding the right exponent of the $\log M$ term.

Since a central result of our investigation is on discrepancy bounds that are independent of the size of the grid, we usually work with the hypergraph $\mathcal{H}_N^d = ([N]^d, \mathcal{E}_N^d)$, where

$$\mathcal{E}_N^d = \left\{ \prod_{i=1}^{d} [x_i..y_i] : 1 \leq x_i \leq y_i \leq N \right\}$$

for some sufficiently large integer N. We have the following result.

Theorem 1. *Let M, $d \geq 2$ be positive integers and q_1 the smallest prime power in the canonical factorization of M into prime powers. We have*

(i) $\operatorname{disc}^+(\mathcal{H}_N^d, M) = O(\log^{d-1} M)$ *for $d \leq q_1 + 1$, independently of $N \in \mathbb{N}$; and*

(ii) $\operatorname{disc}^+(\mathcal{H}_N^d, M) = \Omega(\log^{(d-1)/2} M)$ *for* $N \geq M$.

The combinatorial discrepancy results are shown via strong results from geometric discrepancy theory. The problem of geometric discrepancy in the unit cube $[0,1[^d$ is to distribute $n \in \mathbb{N}$ points evenly with respect to axis-parallel boxes: In every box R there should be approximately $n \operatorname{vol}(R)$ points, where $\operatorname{vol}(R)$ denotes the volume of R. Again, discrepancy quantifies the distance to a perfect distribution. The discrepancy of a point set $\mathcal{P}$ with respect to a box $R \subseteq [0,1[^d$ and the set of all axis-parallel boxes $\mathcal{R}_d$ are defined by

$$\begin{aligned} D(\mathcal{P}, R) &= ||\mathcal{P} \cap R| - n \operatorname{vol}(R)|, \\ D(\mathcal{P}, \mathcal{R}_d) &= \sup_{R \in \mathcal{R}_d} |D(\mathcal{P}, R)|. \end{aligned}$$

The general idea in the proofs of the lower bound in Sinha *et al.* [8] and Anstee *et al.* [1] is the same, described here in two dimensions.

Starting with an arbitrary M-coloring of $[M]^2$, there is a monochromatic set $\hat{P}$ with M vertices. Based on this set, an M-point set $\mathcal{P}$ in $[0,1[^2$ is constructed. By discrepancy theory [7], there is a rectangle R such that $D(\mathcal{P}, R) = \Omega(\log M)$. Rounding R to the $[M]^2$ grid, they construct a hyperedge $\hat{R}$ that has almost the volume as R. Additionally $\hat{R}$ contains as many vertices of $\hat{P}$ as R points of $\mathcal{P}$. With the help of $\hat{R}$ and a short calculation the lower bound of the additive error $\Omega(\log M)$ is shown.

The small, but crucial, mistake in the proof of Sinha *et al.* [8] lies in the transfer from the geometric discrepancy setting back to the combinatorial one. Recall that the authors started with a color class of exactly M^{d-1} points. They scaled it down by a factor of M to a set in the unit cube (note that this is a subset of $\{0, 1/M, 2/M, \ldots, (M-1)/M\}^d$). Then their geometric discrepancy argument yields a rectangle of polylogarithmic discrepancy. However, the rectangle $[0, (M-1)/M]^d$ has a much larger discrepancy: It contains all M^{d-1} points, but has a volume of $((M-1)/M)^d$ only. This immediately yields a discrepancy of $M^{d-1}(1 - ((M-1)/M)^d) = \Omega(M^{d-2})$. For dimension $d \geq 3$, this is larger than the upper bound, also indicating an error in the proof of Sinha *et al.* [8]. The last argument also shows that rounding an arbitrary box to a box in the grid can cause a roundoff error, which is of magnitude larger than the discrepancy. For this reason, a direct generalization using the lower bound of Roth [6] is not possible. A more careful analysis is needed. In particular, we have to ensure the existence of a small box having large discrepancy. Using ideas of Beck [2], we show the following.

Theorem 2. *For any n-point set $\mathcal{P}$ in the unit cube $[0,1[^d$, there is an axis-parallel cube Q with side at most $n^{-(2d-3)d/(d-1)^2(2d+1)}$ fully contained in $[0,1[^d$ with*

$$D(\mathcal{P}, Q) = \Omega(\log^{(d-1)/2} n).$$

Now Theorem 1(ii) follows from Theorem 2 using the roundoff reduction of Anstee *et al.* [1] and Sinha *et al.* [8].

For the proof of our upper bound, we use geometric discrepancies to construct a declustering scheme. The notation of Niederreiter [5] is used in the following.

For an integer $b \geq 2$, an elementary interval in base b is an interval of the form

$$E = \prod_{i=1}^{d} [a_i b^{-d_i}, (a_i + 1) b^{-d_i}[,$$

with integers $d_i \geq 0$ and $0 \leq a_i < b^{d_i}$ for $1 \leq i \leq d$. For integers t, m such that $0 \leq t \leq m$, a (t, m, d)-net in base b is a point set of b^m points in $[0, 1[^d$ such that all elementary intervals with volume b^{t-m} contain exactly b^t points. Note that any elementary interval with volume b^{t-m} has discrepancy zero in a (t, m, d)-net. Since any subset of an elementary interval of volume b^{t-m} has discrepancy at most b^t and any box can be packed with elementary intervals in a way that the uncovered part can be covered by $O(\log^{d-1} n)$ elementary intervals of volume b^{t-m}, the following is immediate.

Theorem 3. *A (t, m, d)-net $\mathcal{P}$ has discrepancy $D(\mathcal{P}, \mathcal{R}_d) = O(\log^{d-1} n)$.*

The central argument in our proof of the upper bound is the following result of Niederreiter [5] on the existence of $(0, m, d)$-nets. From the viewpoint of application it is important that his proof is constructive.

Theorem 4. *Let $b \geq 2$ be an arbitrary base and $b = q_1 q_2 \ldots q_u$ be the canonical factorization of b into prime powers such that $q_1 < \ldots < q_u$. Then for any $m \geq 0$ and $d \leq q_1 + 1$, there exists a $(0, m, d)$-net in base b.*

We construct colorings of $\mathcal{H}_N^d$ from $(0, m, d)$-nets with small discrepancy. We start with colorings for $\mathcal{H}_M^d$.

Theorem 5. *Let $\mathcal{P}_{\text{net}}$ be a $(0, d-1, d)$-net in base M in $[0, 1[^d$. Then there is an M-coloring χ_M of $\mathcal{H}_M^d = ([M]^d, \mathcal{E}_M^d)$ such that all rows of $[M]^d$ contain every color exactly once and $\operatorname{disc}(\mathcal{H}_M^d, \chi_M) \leq D(\mathcal{P}_{\text{net}}, \mathcal{R}_d)$.*

In Theorem 6 below, we show that it is sufficient to consider the discrepancy of $\mathcal{H}_M^d$ with respect to these colorings for determining the upper bound of the discrepancy of $\mathcal{H}_N^d$. Theorem 6 is a reasonable improvement of Theorem 4.2 in [3], where

$$\operatorname{disc}(\mathcal{H}_N^d, \chi) \leq 2^d \operatorname{disc}(\mathcal{H}_M^d, \chi_M)$$

is shown. Note that this reduces the implicit constant in the upper bound by factor of 2^d.

Theorem 6. *Let χ_M be an M-coloring of $\mathcal{H}_M^d$ such that all rows of $[M]^d$ contain every color exactly once and χ a coloring of $\mathcal{H}_N^d$ defined by $\chi(x_1, \ldots, x_d) = \chi_M(y_1, \ldots, y_d)$ such that $x_i \equiv y_i \pmod{M}$ for $i \in [d]$, $x_i \in [N]$ and $y_i \in [M]$. Then*

$$\operatorname{disc}(\mathcal{H}_N^d, \chi) = \operatorname{disc}(\mathcal{H}_M^d, \chi_M).$$

The upper bound in Theorem 1 follows from the above.

References

[1] R. Anstee, J. Demetrovics, G.O.H. Katona, A. Sali. Low discrepancy allocation of two-dimensional data. *Foundations of Information and Knowledge Systems*, 1762 (2000), 1–12.

[2] J. Beck, W.W.L. Chen. *Irregularities of Distribution* (Cambridge Tracts in Mathematics 89, Cambridge University Press, Cambridge, 1987).

[3] C.-M. Chen, C. Cheng. From discrepancy to declustering: near optimal multidimensional declustering strategies for range queries. *ACM Symp. on Database Principles*, pp. 29–38 (Madison, WI, 2002).

[4] B. Doerr, A. Srivastav. Multicolour discrepancies. *Combinatorics, Probability and Computing*, 12 (2003), 365–399.

[5] H. Niederreiter. Point sets and sequences with small discrepancy. *Monatsh. Math.*, 104 (1987), 273–337.

[6] K.F. Roth. Remark concerning integer sequences. *Acta Arith.*, 9 (1964), 257–260.

[7] W.M. Schmidt. On irregularities of distribution VII. *Acta Arith.*, 21 (1972), 45–50.

[8] R.K. Sinha, R. Bhatia, C.-M. Chen. Asymptotically optimal declustering schemes for 2-dim range. *Theoret. Comput. Sci.*, 296 (2003), 511–534.

Quantum algorithms for numerical integration

Stefan Heinrich

One of the most challenging questions of today, in the overlap of computer science, mathematics, and physics, is the exploration of potential capabilities of quantum computers. Milestones which intensified and enlarged research considerably were the algorithm of Shor [6], who showed that quantum computers could factor large integers efficiently (which is widely believed to be infeasible on classical computers) and the quantum search algorithm of Grover [1], which provides a quadratic speedup over deterministic and randomized classical algorithms of searching a database.

So far research was mainly concentrated on discrete problems like the above and many others one encounters in computer science. Much less is known about computational problems of analysis, including such typical field of application as high dimensional integration. We seek to understand how well these problems can be solved in the quantum model of computation (that is, on a – hypothetical – quantum computer) and how the outcome compares to the efficiency of deterministic or Monte Carlo algorithms on a classical (*i.e.* non-quantum) computer.

First steps were made by Novak [5], who considered integration of functions from Hölder spaces. This line of research was continued by the author [2], where quantum algorithms for the integration of L_p-functions and, as a key prerequisite, for the computation of the mean of p-summable sequences were constructed. In [2], a rigorous model of quantum computation for numerical problems was developed, as well. The case of integration of functions from Sobolev spaces is considered in [3], and more on the computation of the mean was presented in [4]. These papers also established matching lower bounds.

Combining these results with previous ones of information-based complexity theory about the best possible ways of solving the respective problems deterministically or by Monte Carlo on classical computers, we are now in a position to fairly well answer the question where quantum computation can provide a speedup in high dimensional integration and where not. We know cases among the above where quantum algorithms yield an exponential speedup over deterministic algorithms and a quadratic speedup over randomized ones (on classical computers). The talk gives an overview about the state of the art in this field.

References

[1] L. Grover. A fast quantum mechanical algorithm for database search. *Proc. 28th Annual ACM Symp. on the Theory of Computing*, pp. 212–219 (ACM Press, New York, 1996); *Physical Review Letters*, 79 (1996), 325–328. See also http://arXiv.org/abs/quant-ph/9706033.

[2] S. Heinrich. Quantum Summation with an Application to Integration. *J. Complexity*, 18 (2002), 1–50. See also http://arXiv.org/abs/quant-ph/0105116.

[3] S. Heinrich. Quantum integration in Sobolev classes. *J. Complexity*, 19 (2003), 19–42. See also http://arXiv.org/abs/quant-ph/0112153.

[4] S. Heinrich, E. Novak. On a problem in quantum summation. *J. Complexity*, 19 (2003), 1–18. See also http://arXiv.org/abs/quant-ph/0109038.

[5] E. Novak. Quantum complexity of integration. *J. Complexity*, 17 (2001), 2–16. See also http://arXiv.org/abs/quant-ph/0008124.

[6] P.W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pp. 121–131 (IEEE Computer Society Press, Los Alamitos, CA, 1994). See also http://arXiv.org/abs/quant-ph/9508027.

Geometric transversal problems

Jiří Matoušek

A fairly general formulation of the basic problem in discrepancy theory is this: We are given a ground set X (often $\mathbb{R}^d$), a system $\mathcal{F}$ of subsets of X (such as all axis-parallel boxes), a probability measure μ on X such that all sets of $\mathcal{F}$ are measurable, and a parameter $\varepsilon > 0$, and we want to find a probability measure ν supported on n points of X, with $n = n(\mathcal{F}, \mu, \varepsilon)$ as small as possible, such that $|\mu(F) - \nu(F)| \leq \varepsilon$ for all $F \in \mathcal{F}$. A related problem discussed in this talk is that of finding a small transversal for all large sets in $\mathcal{F}$, that is, a set $N \subseteq X$ with $N \cap F \neq \emptyset$ for all $F \in \mathcal{F}$ with $\mu(F) \geq \varepsilon$. Such an N is called a weak ε-net for $\mathcal{F}$ with respect to μ.

Well known results of Vapnik and Chervonenkis and of Haussler and Welzl show that if the VC-dimension of $\mathcal{F}$ is finite, then there is ν as above supported on $O(\varepsilon^{-2} \log \varepsilon^{-1})$ points and N of size $O(\varepsilon^{-1} \log \varepsilon^{-1})$. Moreover, finite VC-dimension is necessary if we want ν or N of size bounded in terms of ε and $\mathcal{F}$ hereditarily (also for $\mathcal{F}$ restricted to any subset $Y \subseteq X$).

Interestingly, if we do not consider restrictions of $\mathcal{F}$ to subsets of X, then weak ε-nets of bounded size exist for some set systems of infinite VC-dimension too. A prime example is the system of all convex sets in $\mathbb{R}^d$. Let us denote by

$f(d,\varepsilon)$ the smallest number such that every probability measure μ in $\mathbb{R}^d$ admits a weak ε-net for convex sets with respect to μ. It is nontrivial to prove that $f(d,\varepsilon) < \infty$ for all d and $\varepsilon > 0$ (this was first done by Alon, Bárány, Füredi, and Kleitman). The best known upper bound is $f(d,\varepsilon) = O(\varepsilon^{-d}(\log \varepsilon^{-1})^{c(d)})$ for every fixed d with a suitable constant $c(d)$. The only known nontrivial lower bound is $f(d, 0.01) = \mathrm{e}^{\Omega(\sqrt{d})}$ as $d \to \infty$. It would be very interesting to improve the quadratic upper bound on $f(2,\varepsilon)$, say, or to provide a superlinear lower bound. A nice (and perhaps hard) problem in high-dimensional convex geometry is to improve bounds on the minimum size of a weak ε-net for convex sets in $\mathbb{R}^d$ with respect to μ, the uniform measure on S^{d-1} (this is the example used for the $\mathrm{e}^{\Omega(\sqrt{d})}$ lower bound mentioned above).

For general set systems $\mathcal{F}$, the existence of weak ε-nets of bounded size seems closely related to the fractional Helly property, which is weaker than finite VC-dimension, but no satisfactory characterization is known.

Most of the material of this talk is covered in detail, for example, in the book [2], where detailed references are also provided. Some more recent results from [3], [1], and [4] are also reported.

References

[1] N. Alon, G. Kalai, J. Matoušek, R. Meshulam. Transversal numbers for hypergraphs arising in geometry. *Adv. Appl. Math.*, 130 (2003), 2509–2514.

[2] J. Matoušek. *Lectures on Discrete Geometry* (Springer-Verlag, New York, 2002).

[3] J. Matoušek. A lower bound for weak ε-nets in high dimension. *Discrete Comput. Geom.*, 28 (2002), 45–48.

[4] J. Matoušek, U. Wagner. New constructions of weak epsilon-nets. *Proc. 19th Annual ACM Symposium on Computational Geometry*, pp. 129–135 (2003).

New bounds for the star discrepancy

Erich Novak

(joint work with Aicke Hinrichs)

Can we compute, up to some error $\varepsilon > 0$, the integral

$$I_d(f) = \int_{[0,1]^d} f(x)\,\mathrm{d}x$$

for $f : [0,1]^d \to \mathbb{R}$ from F_d in polynomial time, *i.e.*,

$$\mathrm{cost}(\varepsilon, F_d) \leq C\varepsilon^{-\gamma} d^{\beta}?$$

In some applications the dimension d is (very) large. The answer depends on the classes F_d, see the survey [4]. For certain F_d, we have to study the star-discrepancy.

Let $M_n = \{t_1, \dots, t_n\} \subset [0,1]^d$. The star-discrepancy $\mathrm{disc}_\infty(M_n)$ is defined by

$$\mathrm{disc}_\infty(M_n) = \sup_{x \in [0,1]^d} \left| x_1 \dots x_d - \frac{1}{n} \sum_{i=1}^{n} 1_{[0,x)}(t_i) \right|.$$

Low discrepancy sequences are quite often used in numerical analysis for the so called quasi-Monte Carlo methods. One obtains

$$\mathrm{disc}_\infty(M_n) \le C_d\, n^{-1}(\log n)^{d-1},$$

or similar upper bounds. It is not known whether these known M_n have a small discrepancy if d is large (say $d > 10$) and n is moderate (say $n \approx 10d^2$). In this direction we present and comment on the following main results, established respectively in [1] and [2].

Theorem 1. *There exists $c > 0$ such that for any $n, d \in \mathbb{N}$, there exists M_n with*

$$\mathrm{disc}_\infty(M_n) \le c\sqrt{\frac{d}{n}}. \tag{1}$$

Theorem 2. *There exists $k > 0$ such that*

$$\mathrm{disc}_\infty(M_n) \ge k \min\left(\frac{d}{n}, 1\right)$$

for all M_n and all $n, d \in \mathbb{N}$.

Both results are proved using the Vapnik-Červonenkis dimension. The proof of the upper bound is probabilistic. It is not known how we can construct points $M_n \subset [0,1]^d$ in polynomial (in n and d) time such that (1), or a slightly weaker estimate, holds.

Can we prove results with "less randomness"? Can we find a "small" subset of $[0,1]^{dn}$ containing a low discrepancy set M_n? We now discuss how the p-discrepancy might be of some help for these. The p-discrepancy of M_n is defined by

$$\mathrm{disc}_p(M_n) = \left(\int_{[0,1]^d} \left| x_1 \dots x_d - \frac{1}{n}\sum_{i=1}^n 1_{[0,x)}(t_i)\right|^p \mathrm{d}x\right)^{1/p}.$$

The discrepancy function is not "too peaked", one can obtain upper bounds for $\mathrm{disc}_\infty(M_n)$ from upper bounds of $\mathrm{disc}_p(M_n)$. The idea is to compute the expectation $\mathbb{E}(\mathrm{disc}_p^p(M_n))$ for even p with different distributions on $[0,1]^{nd}$. We consider the Lebesgue measure λ and another measure. For even p, we obtain

$$\mathrm{disc}_p^p(M_n) = \sum_{j=0}^p \binom{p}{j}(-n)^{-j} \sum_{(u_1,\dots,u_j)\in\{1,\dots,n\}^j} (p-j+1)^{-d} \prod_{m=1}^d \min_{k=1,\dots,j}(1-t_{u_k,m}^{p-j+1}),$$

and so

$$\begin{aligned}&\mathbb{E}(\mathrm{disc}_p^p(M_n))\\ = \;& \sum_{j=0}^p \binom{p}{j}(-n)^{-j} \sum_{(u_1,\dots,u_j)\in\{1,\dots,n\}^j} (p-j+1)^{-d}\, \mathbb{E}\left(\prod_{m=1}^d \min_{k=1,\dots,j}(1-t_{u_k,m}^{p-j+1})\right).\end{aligned}$$

We consider first the case of Lebesgue measure. We obtain

$$\mathbb{E}_\lambda\left(\prod_{m=1}^d \min_{k=1,\dots,j}(1-t_{k,m}^{\alpha})\right) = \left(\frac{\alpha}{\alpha+j}\right)^d.$$

Let $\#(j,k,n)$ be the number of tuples $(u_1,\ldots,u_j)\in\{1,\ldots,n\}^j$ such that k different elements occur. Then

$$\mathbb{E}_\lambda(\mathrm{disc}_p^p(M_n))=\sum_{j=0}^{p}\binom{p}{j}(-n)^{-j}\sum_{k=0}^{j}(k+p-j+1)^{-d}\#(j,k,n).$$

The numbers $\#(j,k,n)$ can be written with the Stirling numbers of first and second type. Using the fact that

$$\sum_{k=0}^{p-r+j}\binom{p}{r+k-j}(-1)^k s(k,k-j)S(k-j+r,k)=0$$

for $p=2m$ even, $r=0,\ldots,m-1$ and $j=0,\ldots,r$, we obtain

$$(\mathbb{E}_\lambda(\mathrm{disc}_p^p(M_n)))^{1/p}\le 4p(p+2)^{1/p}2^{-d/p}n^{-1/2}.$$

Compared with Theorem 1, one obtains a slightly weaker upper bound, see [1].

An improvement is possible using symmetrization. Let $X_i(M_n)(x)=1_{[0,x)}(t_i)$. Then $X_1,\ldots,X_n$ are i.i.d. random variables with values in L_p and one gets, see [3],

$$\mathbb{E}_\lambda\left(\left\|\frac{1}{n}\sum_{i=1}^{n}(X_i-\mathbb{E}X_i)\right\|^p\right)\le\mathbb{E}_{\lambda,\varepsilon}\left(2\left\|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_iX_i\right\|\right)^p.$$

Similar computation as above yields

$$\mathbb{E}_{\lambda,\varepsilon}\left(2\left\|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_iX_i\right\|\right)^p=2^pn^{-p}\sum_{k=0}^{p/2}(k+1)^{-d}\#(p/2,k,n).$$

Observe that now there is no cancellation of positive and negative terms, and one gets

$$(\mathbb{E}_\lambda(\mathrm{disc}_p^p(M_n)))^{1/p}\le 2\sqrt{p}(p+2)^{1/p}2^{-d/p}n^{-1/2}.$$

The upper bound (n proportional to d) follows.

Next, we consider generalized lattices with shift. Now

$$M_n^{z,\Delta}=\{t_j=jz+\Delta \pmod 1 : j=0,\ldots,n-1\}$$

with $z,\Delta\in[0,1]^d$. Consider

$$\mathbb{E}_{z,\Delta}(\mathrm{disc}_p^p(M_n^{z,\Delta}))=\int_{[0,1]^{2d}}\mathrm{disc}_p^p(M_n^{z,\Delta})\,\mathrm{d}z\mathrm{d}\Delta.$$

Is it true that

$$\mathbb{E}_{z,\Delta}(\mathrm{disc}_p^p(M_n^{z,\Delta}))\le\mathbb{E}_\lambda(\mathrm{disc}_p^p(M_n))?$$

One would need the numbers

$$\mathbb{E}_{z,\Delta}\left(\prod_{m=1}^{d}\min_{k=1,\ldots,j}(1-t_{u_k,m}^\alpha)\right),$$

i.e., the two-dimensional integrals

$$\int_0^1\int_0^1\max((j_1z+\Delta)_1^\alpha,(j_2z+\Delta)_1^\alpha,\ldots,(j_kz+\Delta)_1^\alpha)\,\mathrm{d}z\mathrm{d}\Delta,$$

with j_i different natural numbers and $(j_i z+\Delta)_1$ are modulo 1, *i.e.*, $x = \lfloor x\rfloor + (x)_1$, $\alpha \in \{1,\dots,p+1\}$ and $k \in \{0,1,\dots,p\}$.

An open problem is to prove an upper bound, such as (1), for lattices M_n.

References

[1] S. Heinrich, E. Novak, G. W. Wasilkowski, H. Woźniakowski. The inverse of the star-discrepancy depends linearly on the dimension. *Acta Arith.*, 96 (2001), 279–302.

[2] A. Hinrichs. Covering numbers, Vapnik-Červonenkis classes and bounds for the star-discrepancy. *J. Complexity* (to appear).

[3] M. Ledoux, M. Talagrand. *Probability in Banach spaces: Isoperimetry and Processes* (Springer, Berlin, 1991).

[4] E. Novak, H. Woźniakowski. When are integration and discrepancy tractable? *Foundation of Computational Mathematics, Oxford, 1999*, pp. 211–266, R.A. DeVore, A. Iserles, E. Süli, eds. (Cambridge University Press, Cambridge, 2001).

Discrepancy of $(0,1)$-sequences

Friedrich Pillichshammer

(joint work with Gerhard Larcher)

For a sequence $x_0, x_1, \dots$ of points in the 1-dimensional unit interval $[0,1)$, the discrepancy function Δ_N, where $N \geq 1$, is defined as $\Delta_N(\alpha) := A_N([0,\alpha))/N - \alpha$, for $0 \leq \alpha \leq 1$, where $A_N([0,\alpha))$ denotes the number of indices i satisfying $0 \leq i \leq N-1$ and $x_i \in [0,\alpha)$. Now the L_p-discrepancy $L_{p,N}$, for $p \geq 1$, of the sequence is defined as the L_p-norm of the discrepancy function Δ_N, *i.e.*, for $1 \leq p < \infty$, we set

$$L_{p,N} = L_{p,N}(x_0, x_1, \dots) := \left(\int_0^1 |\Delta_N(\alpha)|^p \, \mathrm{d}\alpha \right)^{1/p}.$$

For $p = \infty$, we get the usual star discrepancy

$$D_N^* = D_N^*(x_0, x_1, \dots) := \sup_{0 \leq \alpha \leq 1} |\Delta_N(\alpha)|$$

of the sequence.

We consider the discrepancy of a special class of sequences in $[0,1)$, namely the class of the so-called digital $(0,1)$-sequences. Digital $(0,1)$-sequences or, more generally, digital (t,s)-sequences were introduced by Niederreiter [3, 4], and they provide at the moment the most efficient method for generating sequences with small discrepancy.

We consider the discrepancy of digital $(0,1)$-sequences over $\mathbb{Z}_2$. Choose an $\mathbb{N}\times\mathbb{N}$ matrix C over $\mathbb{Z}_2$ such that every left upper $m \times m$ matrix $C(m)$ has full rank over $\mathbb{Z}_2$. For $n \geq 0$, let $n = n_0 + n_1 2 + n_2 2^2 + \dots$ be the base 2 representation of n. Then multiply the vector $\vec{n} = (n_0, n_1, \dots)^T$ with the matrix C to obtain

$$C\vec{n} =: (y_1(n), y_2(n), \dots)^T \in \mathbb{Z}_2^\infty,$$

and set

$$x_n := \frac{y_1(n)}{2} + \frac{y_2(n)}{2^2} + \dots .$$

Every sequence constructed in this way is called digital $(0,1)$-sequence over $\mathbb{Z}_2$.

The most famous digital $(0,1)$-sequence over $\mathbb{Z}_2$ is the well known van der Corput sequence which is generated by the $\mathbb{N}\times\mathbb{N}$ identity matrix.

Niederreiter [3, 4] proved that for any digital $(0,1)$-sequence over $\mathbb{Z}_2$, we have

$$ND_N^* \leq \frac{\log N}{2\log 2} + O(1)$$

for any $N \in \mathbb{N}$. There is also a well known lower bound due to Schmidt [8] which tells us that for any sequence in $[0,1)$, for the star discrepancy D_N^*, we have

$$ND_N^* \geq \frac{\log N}{66\log 4}$$

for infinitely many values of $N \in \mathbb{N}$. Hence the star discrepancy of any digital $(0,1)$-sequence over $\mathbb{Z}_2$ is of best possible order in N.

Our first result [5] is the following improvement of Niederreiter's result.

Theorem 1. *Let $\widetilde{D}_N^*$ denote the star discrepancy of any digital $(0,1)$-sequence over $\mathbb{Z}_2$. For every $N \geq 1$, we have*

$$N\widetilde{D}_N^* \leq ND_N^* \leq \frac{\log N}{3\log 2} + 1,$$

where D_N^* denotes the star discrepancy of the van der Corput sequence.

Hence the van der Corput sequence is the worst distributed digital $(0,1)$-sequence over $\mathbb{Z}_2$ with respect to star discrepancy. For the star discrepancy of the van der Corput sequence we can say even more.

Theorem 2. *Let D_n^* denote the star discrepancy of the first n elements of the van der Corput sequence. For every $\varepsilon > 0$, we have*

$$\lim_{N\to\infty} \frac{1}{N}\left|\left\{n \leq N : \frac{\log 2}{4} - \varepsilon \leq \frac{nD_n^*}{\log n} \leq \frac{\log 2}{4} + \varepsilon\right\}\right| = 1.$$

Finally we consider the L_2-discrepancy of digital $(0,1)$-sequences. We can prove the following.

Theorem 3. *For the L_2-discrepancy of any digital $(0,1)$-sequence over $\mathbb{Z}_2$ generated by a non-singular upper triangular (NUT) matrix, we have*

$$(NL_{2,N})^2 \leq \left(\frac{\log N}{6\log 2}\right)^2 + O(\log N).$$

This is a generalization of a result of Faure [2] who proved this bound for the L_2-discrepancy of the van der Corput sequence. Further, we know from [1, 5, 6] that

$$\limsup_{N\to\infty} \frac{NL_{2,N}}{\log N} = \frac{1}{6\log 2}$$

for the L_2-discrepancy of the van der Corput sequence. Hence we have the following consequence.

Theorem 4. *We have*

$$\limsup_{N\to\infty} \sup \frac{N L_{2,N}}{\log N} = \frac{1}{6\log 2},$$

where the sup *is extended over all digital* $(0,1)$*-sequences generated by an NUT matrix. In other words, the van der Corput sequence is essentially the worst distributed digital* $(0,1)$*-sequence over* $\mathbb{Z}_2$ *which is generated by an NUT matrix.*

We compare this result with the lower bound of Roth [7] which tells us that there exists a constant $c>0$ such that for the L_2-discrepancy, for any sequence in $[0,1)$, we have

$$L_{2,N} \geq c\frac{\sqrt{\log N}}{N}$$

for infinitely many values of $N \in \mathbb{N}$. So our upper bound is not best possible in the sense of Roth's lower bound. The following question arises: Is there a digital $(0,1)$-sequence over $\mathbb{Z}_2$ generated by an NUT matrix C such that for the L_2-discrepancy of this sequence, we have

$$L_{2,N} \leq c_1\frac{\sqrt{\log N}}{N}$$

for any $N \geq 2$, where $c_1 > 0$?

Until now no such sequence is known. Motivated by results from [5], we consider the digital $(0,1)$-sequence generated by the matrix

$$C = \begin{pmatrix} 1 & 1 & 1 & \dots \\ 0 & 1 & 1 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{1}$$

Theorem 5. *For the* L_2*-discrepancy of the digital* $(0,1)$*-sequence generated by the matrix* C *from (1), we have, for any* $\varepsilon > 0$,

$$\lim_{N\to\infty} \frac{1}{N}\left|\left\{n \leq N : L_{2,n} \leq c\frac{(\log n)^{1/2+\varepsilon}}{n}\right\}\right| = 1. \tag{2}$$

Theorem 6. *For the digital* $(0,1)$*-sequence from Theorem 5, we have*

$$L_{2,N} > c\frac{\log N}{N}$$

for infinitely many $N \in \mathbb{N}$*, where* $c > 0$.

To summarize, it is well known that the star discrepancy of any digital $(0,1)$-sequence over $\mathbb{Z}_2$ is best possible in the order of magnitude in N. On the other hand, the question on whether there exist digital $(0,1)$-sequences with best possible order of L_2-discrepancy (in the sense of Roth) or not seems to be a very difficult open problem.

References

[1] H. Chaix, H. Faure. Discrépance et diaphonie en dimension un. *Acta Arith.*, 63 (1993), 103–141.

[2] H. Faure. Discrépance quadratique de la suite de van der Corput et de sa symétrique. *Acta Arith.*, 55 (1990), 333–350.

[3] H. Niederreiter. Point sets and sequences with small discrepancy. *Monatsh. Math.*, 104 (1987), 273–337.

[4] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods* (CBMS-NSF Series in Applied Mathematics, 63, SIAM, Philadelphia, 1992).

[5] F. Pillichshammer. On the discrepancy of (0, 1)-sequences. *J. Number Theory*, 104 (2004), 301–314.

[6] P.D. Proinov, E.Y. Atanassov. On the distribution of the van der Corput generalized sequences. *C. R. Acad. Sci. Paris Sér. I Math.*, 30 (1988), 895–900.

[7] K.F. Roth. On irregularities of distribution. *Mathematika*, 1 (1954), 73–79.

[8] W.M. Schmidt. Irregularities of distribution VII. *Acta Arith.*, 21 (1972), 45–50.

Combinatorial complexity of convex sequences and some other hard Erdős problems

Mischa Rudnev

I am not a 100% aware whether there is a precise definition of what constitutes a "hard Erdős problem". However, there is a sort of general agreement about some of those great many questions posed by Erdős. Take for example the "distance conjecture": Let $P_N \subset \mathbb{R}^d$ be a point set of N elements, where N is large. Let

$$\Delta(P_N) = \{t = \|x - y\| : x, y \in P_N\}$$

be the Euclidean distance set of P_N. Prove that its cardinality

$$|\Delta(P_N)| = \Omega_\varepsilon(N^{2/d}). \tag{1}$$

Above and below, the symbols Ω, $\gtrsim$ or Ω_ε, $\gtrapprox$ (O, $\lesssim$ or O_ε, $\lessapprox$) are used to indicate lower (upper) bounds in the usual way. The symbol $\approx$ stands for equality up to a constant (depending on d).

The distance conjecture has been mostly approached by methods of combinatorial geometry. See for example the book of Matoušek [10] for the state-of-the-art. The best result so far, specifically in $d = 2$, is ε slightly below $1/7$, due to Solymosi and Toth [12], improved a bit by Tardos [14].

The case when P_N is well-distributed (*i.e.*, when there exists a cube Q containing P_N and a pair of constants (c, C) such that a ball of radius c centered at any $p \in P_N$ contains no other points of P_N, while any ball of radius C, centered anywhere in Q does contain some $p \in P_N$) is of special interest. For example, if $P_N = \mathbb{Z}^2 \cap [0, \sqrt{N}]^2$, then we have $|\Delta(P_N)| \approx N/\sqrt{\log N}$, so for $d = 2$, the bound (1) is best possible.

It is expected [6] that in the well-distributed case, the distance conjecture should be true with the bound $|\Delta(P_N)| = \Omega(N^{2/d}/\log^2 N)$, using the methods of Fourier analysis. Turn each $p \in P_N$ into a small ball, endow the resulting set with a natural probability measure μ, and then study the distance measure ν_μ, generated by μ.

The theory was neatly set up by Mattila [11], in connection with the question of Falconer [3], whether any Borel set of Hausdorff dimension $s > d/2$ has a distance set of positive Lebesque measure. An example with integer lattice points shows that $d/2$ would be the best possible, see [3]. The present status of Falconer's conjecture is $s = d(d+1)/2(d+1)$ due to Wolff [15] and Erdogan [2].

Application of Fourier analysis methods to the well-distributed set case prompts one to try to appeal to the methods developed for mean square discrepancy of the lattice point distribution, see, *e.g.*, [8]. However, in the latter case, there is the Poisson summation formula, which results in a curious fact that the corresponding distance measure ν_μ are commensurable point-wise.

The results obtained via Fourier analysis are easily extendible to non-isotropic distances, determined by a symmetric strictly convex body K, with a smooth boundary (K is a Euclidean ball for the Euclidean distance $\|\cdot\|$), as long as there is a lower bound for Gaussian curvature on ∂K. The effect of curvature is crucial and displays itself in a variety questions, one of which is discussed below. The motivation for it comes from estimating the L^p-norms of some trigonometric polynomials and a theorem of Konyagin [9]. See below.

The rest exposes the results of our recent work [7]. Let $B = \{1, 2, \ldots, N\}$ be a "base" set. Let $S = \{s_j\}_{j=1}^N$ be a strictly convex sequence, *i.e.*, the differences $s_{j+1} - s_j$ are strictly monotone in j. One can assume that $s_j = f(j)$, $j \in B$, for some strictly convex function f. There is no bound on $D^2 f$ from below, except for $D^2 f > 0$.

Consider the equation

$$s_{j_1} + \ldots + s_{j_d} = s_{j_{d+1}} + \ldots + s_{j_{2d}}. \tag{2}$$

Let $\mathcal{C}_d$ be the number of solutions of (2), with all j's in B. It appears reasonable to conjecture that without any algebraic assumptions on f, one has

$$\mathcal{C}_d = O_\varepsilon(N^{2d-2}). \tag{3}$$

Traditionally, problems like this one have been studied by algebraic methods, see the survey [5]. For example, (3) follows easily if $f(x) = x^m$, where $m = 2, 3, \ldots$. But Konyagin used combinatorics, namely the Szemerédi-Trotter (henceforth ST theorem – see [13] for a proof - bounding the number of incidences I for an arrangement $(\mathcal{L}, \mathcal{P})$ of lines (curves) and points in $\mathbb{R}^2$ as $I \lesssim |\mathcal{L}| + |\mathcal{P}| + (|\mathcal{L}||\mathcal{P}|)^{2/3}$) to get a robust bound

$$\mathcal{C}_2 = O(N^{5/2}),$$

no matter what S, as long as it is strictly convex. A paper by Elekes *et al.* [1] falls short of proving this result, instead giving the lower bound $N^{3/2}$ for the number of the elements of the sumset $2S = S + S$. Konyagin's result was repeated by Garaev [4], who removed ST as the (only) prerequisite for the proof.

The following theorem of Iosevich, Ten and the author generalizes Konyagin's theorem for $d \geq 2$.

Theorem 1. *For $d \geq 2$, let $\alpha = 2(1 - 2^{-d})$. Then $|\mathcal{C}_d| = O(N^{2d-\alpha})$.*

At the moment, the author is confident that conjecture (3) can be vindicated for a wide class of f's by using the Fourier transform. The approach was roughly outlined in the final section of [7]. Our recent proof [6] of the Erdős distance conjecture for well-distributed sets uses the same ideas. However the constant in estimate (3) may end up being dependent on the lower bound for $D^2 f$.

If the sequence $\{s_i\}_{i\in B}$ is integer-valued, we deduce an estimate for the L^{2d}-norm of the Dirichlet kernel associated with S.

Theorem 2. *If $S \subset \mathbb{Z}$, for $\theta \in \mathbb{T}^1$, let*

$$F_N(\theta) = \sum_{j=1}^{N} \mathrm{e}^{2\pi \mathrm{i} b_j \theta}.$$

Then

$$\|f_N\|_{2d} = O\left(N^{1-\frac{1-2^{-d}}{d}}\right).$$

Theorem 1 was proved by induction in dimension, starting off $d = 2$. However, the higher-dimensional set-up is not amenable to the standard ST, unless one adds weights to it. Reduction to weighted ST is not obvious, as one is tempted to turn towards higher-dimensional versions of ST, which are nothing as good as the case $d = 2$. Though Fourier analysis is much more robust, as far as the dimension is concerned, see [6].

In fact, [1] proves a very similar estimate $\Omega(N^{2-2^{-d+1}})$ for cardinality of the sumset dS. However, for the latter estimate, no weighted ST turns out to be necessary; it also arises as a by-product in [7]. From the harmonic analysis point of view the two estimates end up being equivalent, see [6].

Weighted (in some sense) versions of ST have been around for a while, see [13]. But it is in connection with the weights where the main difficulty arises. On the inductive step $d \to d+1$ of our proof, the lines involved in the incidences have weights, which they have inherited from the previous step. These weights are equal to the multiplicity of $c \in dS$, available from the previous step via a certain majorant, bounding the distribution function $\nu(t)$ of multiplicities (weights) over the elements of the sumset dS. Complexity $\mathcal{C}_d$ is just the square of the L^2-norm of ν.

If one writes down the incidence bound in the weighted set-up, it incorporates the L^∞-norm of ν, which is too large. So the most non-trivial part of the proof of Theorem 1 is a lemma, which states that one can partition the weighted set of curves $\mathcal{L}$ into some $\log\log N$ pieces, so that eventually one can use the L^1-norm of ν in the estimate for the total number of incidences. This enables one to get an exponentially small error 2^{-d+1}, with respect to the conjectured bound.

Both works [7] and [6] are in essence based on the same simple principle: the L^2-norm of the function ν in the former case and the distance measure ν_μ in the latter case should not be too large in comparison with the L^1-norm, at most $O_\varepsilon(N)$ times greater. In both cases, this prevents the quantity in question from being supported on a thin set, yielding the desired result.

References

[1] G. Elekes, M. Nathanson, I.Z. Ruzsa. Convexity and sumsets. *J. Number Theory*, 83 (2000), 194–201.
[2] M.B. Erdogan. A note on the Fourier transform of fractal measures (preprint, 2003).
[3] K.J. Falconer. On the Hausdorff dimensions of distance sets. *Mathematika*, 32 (1986), 206–212.
[4] M. Garaev. On lower bounds for the L^1-norm of exponential sums. *Math. Notes*, 68 (2000), 713–720.
[5] D.R. Heath-Brown. Counting rational points on algebraic varieties (preprint, 2002).
[6] A. Iosevich, M. Rudnev. (in preparation, 2004).
[7] A. Iosevich, M. Rudnev, V. Ten. Combinatorial complexity of convex sequences (submitted in 2003).
[8] A. Iosevich, E. Sawyer, A. Seeger. Mean square discrepancy bounds for the number of lattice points in large convex bodies. *J. D'Anal. Math.*, 87 (2002), 209–230.
[9] S. Konyagin. An estimate of the L_1-norm of an exponential sum (in Russian). *The Theory of Approximations of Functions and Operators – Abstracts of Papers of the International Conference Dedicated to Stechkin's 80th Anniversary, Ekaterinbourg, 2000*, pp. 88–89.
[10] J. Matoušek. *Lectures on Discrete Geometry* (Springer, 2002).
[11] P. Mattila. Spherical averages of Fourier transforms of measures with finite energy; dimension of intersection and distance sets. *Mathematica*, 34 (1987), 207–228.
[12] J. Solymosi, Cs.D. Tóth. Distinct distances in the plane. *Discrete Comput. Geom.*, 25 (2001), 629–634.
[13] L. Székely. Crossing numbers and hard Erdős problems in discrete geometry. *Combinatorics, Probability, and Computing*, 6 (1997), 353–358.
[14] G. Tardos. On distinct sums and distinct distances. *Adv. Math.*, 180 (2003), 275–289.
[15] T. Wolff. Decay of circular means of Fourier transforms of measures. *Internat. Math. Res. Notices*, 10 (1999), 547–567.

Extremal additive intersective sets

Tomasz Schoen

For a set $S = \{s_1, s_2, \dots\} \subseteq \mathbb{N}$, denote its counting function by $S(n) = |S \cap [n]|$, where $[n] = \{1, 2, \dots, n\}$. As usual, let $A+B$ be the set of all numbers represented in the form $a+b$, where $a \in A$ and $b \in B$. Let

$$\underline{\mathrm{d}}(A) = \liminf_{n\to\infty} \frac{A(n)}{n} \quad \text{and} \quad \overline{\mathrm{d}}(A) = \limsup_{n\to\infty} \frac{A(n)}{n}.$$

Define

$$\underline{\mathrm{int}}(S) = \sup_{(A+A)\cap S=\emptyset} \underline{\mathrm{d}}(A).$$

We say that a set S has no intersective property if there is a set A such that

$$(A+A)\cap S = \emptyset \quad \text{and} \quad \underline{\mathrm{d}}(A) = \tfrac{1}{2}.$$

Consider the following question of Erdős. Put $S(d,r,n) = |\{s \in S \cap [n] : s \equiv r \pmod{d}\}|$ and suppose that S satisfies the two conditions

$$\frac{S(d,r,n)}{S(n)} \to \frac{1}{d} \quad \text{as } n \to \infty \text{ for all } d, r \in \mathbb{N}, \tag{1}$$

and

$$\frac{s_n}{s_{n+1}} \to 1 \quad \text{as } n \to \infty. \tag{2}$$

Is it true that for every set $A \subseteq \mathbb{N}$ with $(A+A) \cap S = \emptyset$, we have $\underline{\mathrm{d}}(A) = \frac{1}{2}$?

We prove the following result which solves the problem of Erdős in the negative.

Theorem 1. *Let $\omega(n)$ be any increasing function tending to infinity as $n \to \infty$. Then there is a set $S \subseteq \mathbb{N}$ satisfying (1) and (2), having no intersective property and such that $S(n) \geq n/\omega(n)$ for every $n \in \mathbb{N}$.*

We also show that every sufficiently sparse set has no intersective property.

Theorem 2. *For every $\varepsilon > 0$, there is a set $S \subseteq \mathbb{N}$ such that $S(n) \leq \varepsilon \log n$ for every $n \in \mathbb{N}$, and*

$$\underline{\mathrm{int}}(S) \leq \frac{1}{2} - \frac{1}{4 \cdot 2^{300/\varepsilon}}.$$

Theorem 3. *Let $S \subseteq \mathbb{N}$ be any set with $S(n) = o(\log n)$. Then S has no intersective property, so that there exists a set $A \subseteq \mathbb{N}$ with $\underline{\mathrm{d}}(A) = \frac{1}{2}$ such that $(A+A) \cap S = \emptyset$.*

Theorem 4. *Let $\frac{1}{10} > \varepsilon > 0$, and let $S \subseteq \mathbb{N}$ be an arbitrary set with $S(n) \leq \varepsilon \log n$ for all sufficiently large n. Then*

$$\underline{\mathrm{int}}(S) \geq \frac{1}{2} - \frac{4}{2^{1/\varepsilon}}.$$

A set S is called sum-intersective if for every set A with $\overline{\mathrm{d}}(A) > 0$, we have $(A+A) \cap S \neq \emptyset$ (or $\overline{\mathrm{int}}(S) = 0$). We know from Erdős and Sárközy [1] that if S is sum-intersective, then $S(n) = o(\log^2 n)$ is impossible. We also know from Ruzsa [2] that if $\omega(n) \to \infty$, then there is a sum-intersective set S with $S(n) = O(\omega(n) \log^2 n)$.

Our next result shows that Ruzsa's theorem is sharp.

Theorem 5. *If there is a constant C such that the inequality $S(n) \leq C \log^2 n$ has infinitely many solutions, then*

$$\overline{\mathrm{int}}(S) \geq \frac{1}{2^{20C}}.$$

References

[1] P. Erdős, A. Sárközy. On differences and sums of integers II. *Bull. Soc. Math. Gréce (N.S.)*, 18 (1977), 204–223.

[2] I.Z. Ruzsa. Probabilistic constructions in additive number theory. *Astérisque*, 147–148 (1987), 173–182.

Variation of the number of lattice points in large balls

Maxim Skriganov
(joint work with Alexander Sobolev)

Let $\Gamma \subset \mathbb{R}^d$, $d \geq 2$, be a lattice in the d-dimensional Euclidean space. For any bounded set $\mathcal{C} \subset \mathbb{R}^d$, we denote by $\mathcal{N}[\mathcal{C}]$ the number of lattice points in $\mathcal{C}$, that is

$$\mathcal{N}[\mathcal{C}] = \#\{\boldsymbol{\gamma} \in \Gamma : \boldsymbol{\gamma} \in \mathcal{C}\}.$$

Denote by

$$B(r; \mathbf{k}) = \{\boldsymbol{\xi} : |\boldsymbol{\xi} - \mathbf{k}| < r\}$$

the open ball of radius $r > 0$ centered at the point $\mathbf{k} \in \mathbb{R}^d$. The function $\mathcal{N}[B(r; \mathbf{k})]$ is a periodic function of the variable $\mathbf{k}$ with the period lattice Γ, and hence it is bounded. We are interested in the variation of the quantity $\mathcal{N}[B(r; \mathbf{k})]$ as a function of $\mathbf{k}$. Define for all $r > 0$

$$\mathcal{N}^+(r) = \max_{\mathbf{k}} \mathcal{N}[B(r; \mathbf{k})], \qquad \mathcal{N}^-(r) = \min_{\mathbf{k}} \mathcal{N}[B(r; \mathbf{k})],$$

and introduce the δ-variation of the counting function by writing

$$V(\lambda, \delta) = \mathcal{N}^+(\sqrt{\lambda - \delta}\,) - \mathcal{N}^-(\sqrt{\lambda + \delta}\,)$$

for $\lambda \geq 0$ and $\delta \in [0, \lambda]$. Our objective is to find out when the δ-variation is non-negative and to obtain lower bounds for $V(\lambda, \delta)$ for small δ and large λ under the assumption that the lattice Γ is rational.

A lattice $\Gamma \subset \mathbb{R}^d$ is said to be rational if for any two vectors $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \in \Gamma$, the inner product satisfies the relation

$$\langle \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \rangle = \beta_\Gamma r_{12},$$

where $\beta_\Gamma \neq 0$ is a real-valued constant independent of $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$, and where $r_{12} = r_{21}$ is an integer. Otherwise the lattice is called irrational.

For the cases $d = 2, 3$ quite precise lower bounds for $V(\lambda, \delta)$ are known to hold without any assumptions on the arithmetic properties of Γ. However, in higher dimensions these become important. Our main results are contained in Theorems 1, 2 and 3.

Theorem 1. *Let $\Gamma \subset \mathbb{R}^d$ be a rational lattice and let $d \geq 5$. Then there are three positive constants $\delta_0 = \delta_0(\Gamma)$, $\lambda_0 = \lambda_0(\Gamma)$ and c_Γ such that for all $\delta \in [0, \delta_0]$ and all $\lambda \geq \lambda_0$, we have*

$$V(\lambda, \delta) \geq c_\Gamma \lambda^{(d-2)/2}. \tag{1}$$

The bound (1) is sharp.

Theorem 2. *Let $\Gamma \subset \mathbb{R}^4$ be a rational lattice. Then there are three positive constants $\delta_0 = \delta_0(\Gamma)$, $\lambda_0 = \lambda_0(\Gamma)$ and c_Γ such that for all $\delta \in [0, \delta_0]$ and all $\lambda \geq \lambda_0$, we have*

$$V(\lambda, \delta) \geq c_\Gamma \lambda (\log \log \lambda)^{-1}. \tag{2}$$

It is not yet clear whether one can get rid of the log log-factor in (2) for general rational lattices. However, for the case of a cubic lattice Γ, this can be done.

Theorem 3. *Let $\Gamma = \mathbb{Z}^4$. Then for each $\delta \in [0, 10^{-4}]$, all sufficiently large $\lambda \geq \lambda_0 > 0$ and some $c > 0$, one has the bound*

$$V(\lambda, \delta) > c\lambda.$$

The proofs of Theorems 1, 2 and 3 are based on the classical results on representation of integers by the integer quadratic forms and some arguments from the geometry of numbers.

One-sided discrepancy of hyperplanes in $\mathbb{F}_q^r$

Anand Srivastav
(joint work with Nils Hebbinghaus and Tomasz Schoen)

We study the one-sided discrepancy and discrepancy of the hypergraph $\mathcal{H}_{q,r} = (\mathbb{F}_q^r, \mathcal{E}_{q,r,1})$ of linear hyperplanes in $\mathbb{F}_q^r$, where $\mathbb{F}_q^r$ is the r-dimensional vector space over $\mathbb{F}_q$ and $\mathcal{E}_{q,r,1}$ is the set of all its linear hyperplanes, *i.e.*, the subspaces of codimension one. Let $n := q^r$.

The bounds on the discrepancy can be derived with standard methods (lower bound with the eigenvalue technique and the upper bound via the VC-dimension) and are given by

$$\frac{\sqrt{z(1-z)}}{q}\sqrt{n} - \frac{c-1}{c} \geq \operatorname{disc}(\mathcal{H}_{q,r,1}, c) \geq \alpha\sqrt{\frac{n}{qc}}\, c^{1/2(r-1)}.$$

Since the one-sided discrepancy satisfies $\operatorname{disc}^+(.) \leq \operatorname{disc}(.)$, we have the same upper bound. Our main result is the proof of the lower bound for the one-sided discrepancy $\operatorname{disc}^+(\mathcal{H}_{q,r,1}, c)$, given by

$$\operatorname{disc}^+(\mathcal{H}_{q,r,1}, c) \geq \frac{\sqrt{z(1-z)}}{4q(q-1)\sqrt{c}}\sqrt{n} - \frac{q-1}{q}.$$

This is accomplished by Fourier analysis on the additive group $\mathbb{F}_q^r$. Note that for $q = O(1)$ and $c = O(1)$, the bounds are tight and give a new example for Spencer's six-standard-deviation theorem [6].

Finally, we generalise our main result for the one-sided discrepancy to the hypergraph $\mathcal{H}_{q,r,m} = (\mathbb{F}_q^r, \mathcal{E}_{q,r,m})$, where $\mathcal{E}_{q,r,m}$ is the set of all subspaces of $\mathbb{F}_q^r$ of codimension m, where $m \leq r-3$.

Let V be a finite set and $\mathcal{E}$ a subset of 2^V. Then $\mathcal{H} := (V, \mathcal{E})$ is called a hypergraph. A c-coloring of $\mathcal{H}$ is a function $\chi : V \to M_c$, where M_c is any set of cardinality c. For convenience we take $M_c = \{1, 2, \ldots, c\} =: [c]$, but in applications a different choice of M_c can be helpful (see [1]).

Let $A_i := \chi^{-1}(i)$ be the color-class of color $i \in [c]$ in V. The c-color discrepancy of $\mathcal{H}$ with respect to χ is defined by

$$\operatorname{disc}(\mathcal{H}, \chi, c) = \max_{i\in[c]} \max_{E\in\mathcal{E}} \left| |A_i \cap E| - \frac{|E|}{c} \right|,$$

and the c-color discrepancy of $\mathcal{H}$ is

$$\operatorname{disc}(\mathcal{H}, c) = \min_{\chi:V\to[c]} \operatorname{disc}(\mathcal{H}, \chi, c).$$

For $c = 2$, the c-color discrepancy is exactly half of the common two-color discrepancy where the two colors are represented by 1 and -1. For further information on discrepancies, we refer to Beck and Sós [2] and Matoušek [4]. For our purposes, a related discrepancy notion will be relevant. The one-sided c-color discrepancy of $\mathcal{H}$ with respect to χ is

$$\operatorname{disc}^+(\mathcal{H}, \chi, c) = \max_{i\in[c]} \max_{E\in\mathcal{E}} \left(|A_i \cap E| - \frac{|E|}{c} \right),$$

and the one-sided c-color discrepancy of $\mathcal{H}$ is

$$\operatorname{disc}^+(\mathcal{H}, c) = \min_{\chi:V\to[c]} \operatorname{disc}^+(\mathcal{H}, \chi, c).$$

Trivially we have $\operatorname{disc}^+(\mathcal{H}, c) \le \operatorname{disc}(\mathcal{H}, c)$, where equality holds for $c = 2$.

Let $\mathbb{F}_q$ be the field of q elements, where $q = p^k$ is a power of a prime p, $V := \mathbb{F}_q^r$ the r-dimensional vector space over $\mathbb{F}_q$, and let $\mathcal{E}_{q,r,m}$ be the set of all subspaces of V of codimension m. Put $n := |V| = q^r$. For a set $S \subseteq \mathbb{F}_q^r$ define $S^\sharp := S \setminus \{0\}$. We investigate the discrepancy of the hypergraph $\mathcal{H}_{q,r,m} = (V, \mathcal{E}_{q,r,m})$. Note that $\mathcal{H}_{q,r} = (V, \mathcal{E}_{q,r,1})$ is an (n/q)-uniform hypergraph on n vertices with $|\mathcal{E}_{q,r,1}| = (n-1)/(q-1)$ hyperedges.

We define a new hypergraph $\mathcal{H}' := (V', \mathcal{E}')$ with $V' := V \setminus \{0\}$ and

$$\mathcal{E}' := \{E \cap V' : E \in \mathcal{E}_{q,r,1}\}.$$

For $q = 2$, this hypergraph has constant pair-degree, *i.e.*, there exists a $\lambda \in \mathbb{N}$ with

$$|\{E \in \mathcal{E} : i, j \in E\}| = \lambda,$$

for all $i, j \in V$, $i \neq j$. For such hypergraphs $\mathcal{H} = (V, \mathcal{E})$, we can extend the "trace"-lower bound of Beck and Sós [2] to c-colors and obtain

$$\operatorname{disc}(\mathcal{H}, c) \ge \left(\frac{1}{c^2|\mathcal{E}|} \sum_{v\in V} (d_v - \lambda) \right)^{1/2},$$

with d_v as degree of v. In fact, this lower bound can be extended to cover also the incidence matrix of $\mathcal{H}_{q,r,1}$ for $q > 2$, where we do not have constant pair-degree, but the pair-degree cannot vary too much. This yields the bound of the following theorem. The upper bound is obtained by a c-color generalisation of a theorem of Matoušek [4] for hypergraphs with bounded VC-dimension by Doerr and Srivastav [3].

Theorem 1. *Let* $z := \frac{(q-1) \bmod c}{c}$. *Then there is a constant* $\alpha > 0$ *with*

$$\frac{\sqrt{z(1-z)}}{q}\sqrt{n} - \frac{c-1}{c} \leq \operatorname{disc}(\mathcal{H}_{q,r,1}, c) \leq \alpha\sqrt{\frac{n}{qc}}\, c^{1/2(r-1)}.$$

For one-sided discrepancy, we invoke the Fourier transform on $\mathbb{F}_q^r$ in the following way. For simplicity, we take here $q = 2$. A subspace $E \subseteq \mathbb{F}_q^r$ of codimension 1 is uniquely determined by a vector $z \in \mathbb{F}_q^r$, where $E^\perp = \langle z \rangle$. Thus, for $A \subseteq \mathbb{F}_q^r$, the function

$$E \to |A \cap E| - \frac{|E|}{c}$$

is a function of z, denoted by $f(z)$, and we may build $\widehat{f}(z)$. A sophisticated interplay between the growth of Fourier coefficents and the size of color classes leads to the following main result.

Theorem 2. *Let* $z := \frac{(q-1) \bmod c}{c}$ *and* $q^{r-1} \geq q^{r/2} + 6q^2$. *There exists a constant* $\alpha > 0$ *such that for every* $c \geq 2$, *we have*

$$\frac{\sqrt{z(1-z)}}{4q(q-1)\sqrt{c}}\sqrt{n} - \frac{q-1}{q} \leq \operatorname{disc}(\mathcal{H}_{q,r}, c) \leq \alpha\sqrt{\frac{n}{qc}}\, c^{1/2(r-1)}.$$

Note that the lower bounds for discrepancy and one-sided discrepancy differ by a factor of about $4(q-1)\sqrt{c}$.

Using our theorems, we can extend the result from linear hyperplanes to subspaces of codimensions $m \leq r - 3$.

Theorem 3. *Let* $z := \frac{(q-1) \bmod c}{c}$. *If* $q^{r-m} \geq q^{(r-m+1)/2} + 6q^2$, *there is a constant* $\alpha > 0$ *such that for* $m \leq r-3$, *we have*

$$\frac{\sqrt{z(1-z)}}{4(q-1)\sqrt{c}}\sqrt{\frac{n}{q^{m+1}}} - \frac{q-1}{q} \leq \operatorname{disc}(\mathcal{H}_{q,r,m}, c) \leq \alpha\sqrt{\frac{n}{qc}}\, c^{1/2(r-m)}.$$

References

[1] L. Babai, T.P. Hayes, P.G. Kimmel. The cost of the missing bit: Communication complexity with help. *Combinatorica*, 21 (2001), 455–488.

[2] J. Beck, V.T. Sós. Discrepancy theory. *Handbook of Combinatorics*, pp. 1405–1446, R. Graham, M. Gröschel, L. Lovász, eds. (Elsevier, 1995).

[3] B. Doerr, A. Srivastav. Multi-color discrepancies. *Combinatorics, Probability and Computing*, 12 (2003), 365–399.

[4] J. Matoušek. *Geometric Discrepancy* (Springer, 1999).

[5] R. Lidl, H. Niederreiter. *Introduction to finite fields and their applications* (Revised edition, Cambridge University Press, 1994).

[6] J. Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289 (1985), 679–706.

Metric discrepancy theory

Robert Tichy

In the first part of the lecture a survey on normal numbers and metric theory of uniform distribution is given. In the second part metric theorems for distribution measures of pseudorandom sequences are discussed – joint work with W. Philipp.

Let $\chi(x) = 2\mathbf{1}_{[0,\frac{1}{2})}(\{x\}) - 1$, where $\{x\}$ denotes the fractional part of x and $\mathbf{1}_A$ the indicator function of the set A. Throughout this abstract (n_k) denotes an increasing sequence of positive integers and $\omega \in [0,1)$. For $k \geq 1$, we define

$$e_k := \chi(n_k\omega). \tag{1}$$

The well-distribution measure of stage N of the sequence (1) is defined as

$$W_N := \max_{a,b,t} \left| \sum_{j \leq t} e_{a+bj} \right|, \quad N \geq 1, \tag{2}$$

where the maximum is extended over all $a \in \mathbb{Z}$ and $b, t \in \mathbb{N}$ such that $1 \leq a + b \leq a + bt \leq N$. This measure of pseudorandomness was first introduced by Mauduit and Sárkőzy [1]. As was already noted by them, there is nothing special about the interval $[0, \frac{1}{2})$ since W_N can be bounded by the discrepancy D_t of the defining sequence $(n_k\omega,\ k \geq 1)$ in the form

$$W_N \leq \max_{a,b,t} tD_t(\{n_{a+bj}\omega\}). \tag{3}$$

Here, for a fixed sequence (x_j) with $0 \leq x_j < 1$,

$$D_t(x_j) := \sup \left(\left| \frac{1}{t} \sum_{j \leq t} \Big(\mathbf{1}_{[\alpha,\beta)}(x_j) - (\beta - \alpha) \Big) \right| : 0 \leq \alpha < \beta \leq 1 \right)$$

denotes the discrepancy in the sense of uniform distribution mudulo 1. In view of relation (3) we will formulate our results in terms of discrepancies.

Among other things, Mauduit and Sárkőzy [2, 3] prove metric results for sequences $n_k = k^d$, where $d \in \mathbb{N}$. Our first result can handle arbitrary increasing sequences (n_k) and for $d \geq 3$ it yields a sharper error term.

Theorem 1. *Let $(n_k,\ k \geq 1)$ be an increasing sequence of positive integers. Then for almost all ω and arbitrary $\varepsilon > 0$,*

(4)

$$\max \Big(tD_t(\{n_{a+bj}\omega\}) : a \in \mathbb{Z},\ b, t \in \mathbb{N},\ 1 \leq a+b \leq a+bt \leq N \Big) \ll N^{2/3}(\log N)^{1+\varepsilon}.$$

The third part is devoted to the analysis of pair correlations as studied by Rudnick, Sarnak and Zaharescu. Here some joint results of I. Berkes, W. Philipp and R. Tichy are presented.

We prove a Glivenko-Cantelli type strong law of large numbers for the pair correlation of independent random variables. Except for a few powers of logarithms the results obtained are sharp. Similar estimates hold for the pair correlation of lacunary sequences $\{n_k\omega\}$ modulo 1.

References

[1] C. Mauduit, A. Sárközy. On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol. *Acta Arith.*, 82 (1997), 365–377.
[2] C. Mauduit, A. Sárközy. On finite pseudorandom binary sequences V: On $(n\alpha)$ and $(n^2\alpha)$ sequences. *Monatsh. Math.*, 129 (2000), 197–216.
[3] C. Mauduit, A. Sárközy. On finite pseudorandom binary sequences VI: On $(n^k\alpha)$ sequences. *Monatsh. Math.*, 130 (2000), 281–298.

Average decay of Fourier transforms, geometry of planar convex bodies, and discrepancy theory

Giancarlo Travaglini

A number of facts in discrepancy theory depends on estimates for the decay of the Fourier transform (see [2, 8, 10, 14]). Our first example is given by the following result, which extends a theorem of D. Kendall (see [9, 7]): Let $B \subset \mathbb{R}^d$ be a convex body. For $\rho \geq 2$, $\sigma \in SO(d)$ and $t \in \mathbb{T}^d$, consider the discrepancy

$$D_{\sigma(B)-t}(\rho) = \mathrm{card}((\rho\sigma(B) - t) \cap \mathbb{Z}^d) - \rho^d |B|.$$

Then

$$\int_{\mathbb{T}^d} \int_{SO(d)} |D_{\sigma(B)-t}(\rho)|^2 \, \mathrm{d}\sigma \mathrm{d}t \leq c\rho^{d-1}. \tag{1}$$

The inequality (1) depends on the inequality

$$\int_{\Sigma_{d-1}} |\widehat{\chi}_B(\rho\gamma)|^2 \, \mathrm{d}\gamma \leq c\rho^{-d-1}$$

(see [11, 5]). L^2 results such as (1) do not depend on the shape of B, which instead plays a role when we replace L^2 with L^p, $p < 2$. We consider L^1 and we state our second example, which depends on upper and lower estimates for

$$\int_{\Sigma_1} |\widehat{\chi}_B(\rho\gamma)| \, \mathrm{d}\gamma$$

(see [12, 13, 3, 4, 7]).

Theorem 1. *Let P be a convex polygon and let K be a planar convex body with piecewise smooth boundary, different from a polygon. Then*

$$c_1 \log \rho \leq \int_{\mathbb{T}^2} \int_{SO(2)} |D_{\sigma^{-1}(P)-t}(\rho)| \, \mathrm{d}\sigma \mathrm{d}t \leq c_2 \log^2 \rho, \tag{2}$$

$$c_1 \rho^{1/2} \leq \int_{\mathbb{T}^2} \int_{SO(2)} |D_{\sigma^{-1}(K)-t}(\rho)| \, \mathrm{d}\sigma \mathrm{d}t \leq c_2 \rho^{1/2}. \tag{3}$$

Here we wish to show nearly best possible results (see [6]) for intermediate cases between (2) and (3). In order to do this, we scale between discs and polygons in two different, although related, ways. The first one consists of approximating the convex body B with certain polygons, especially tailored for the Fourier transform, and then counting the number of sides of these polygons. The second one consists of a fractal measure of the image of the Gauss map on ∂B. In both cases we need estimates of the Fourier transform of the characteristic function of a polygon with many sides. These estimates depend partially on a development of the following remark.

Remark. Let T be a triangle and let $\widehat{\chi}_T(\rho\Theta)$ be the Fourier transform of its characteristic function, written in polar coordinates $\rho \geq 2$ and $\Theta = (\cos\theta, \sin\theta)$. Then we have $|\widehat{\chi}_T(\rho\Theta)| \leq c_\theta \rho^{-2}$ when Θ is not orthogonal to a side of T, while we only have $|\widehat{\chi}_T(\rho\Theta)| \leq c\rho^{-1}$ in the three remaining directions. Then (see [4, 7]) one can prove that

$$\int_0^{2\pi} |\widehat{\chi}_T(\rho\Theta)| \,\mathrm{d}\theta \leq c\rho^{-2} \log \rho.$$

Now let $P = P_N$ be a polygon with N sides, of lengths not greater than 1. By splitting P into triangles, we obviously get

$$\int_0^{2\pi} |\widehat{\chi}_P(\rho\Theta)| \,\mathrm{d}\theta \leq cN\rho^{-2} \log \rho,$$

with c independent of N. It turns out that this last “trivial” inequality is nearly sharp, since for any $\varepsilon > 0$, we cannot replace N in the right hand side above by $N^{1-\varepsilon}$ (see [14]).

Estimates for the decay of the Fourier transform can be also used to prove lower bounds for irregularities of distribution. As an example we consider the following theorem, which is a basic result in the theory and it has been independently proved in [1] and [10].

Theorem 2. *Let B be a convex body in $\mathbb{T}^2$. For every finite set $\{u(j)\}_{j=1}^N \subset \mathbb{T}^2$, we have*

$$\int_0^1 \int_{SO(2)} \int_{\mathbb{T}^2} \left| -Ns^2|B| + \sum_{j=1}^N \chi_{s\sigma^{-1}(B)-t}(u(j)) \right|^2 \mathrm{d}t\mathrm{d}\sigma\mathrm{d}s \geq cN^{1/2}. \tag{4}$$

It is possible to prove that for certain choices of B the inequality (4) holds and it is best possible even without averaging over dilations.

Theorem 3. *Let T be a triangle in $\mathbb{T}^2$. For every finite set $\{u(j)\}_{j=1}^N \subset \mathbb{T}^2$, we have*

$$\int_{SO(2)} \int_{\mathbb{T}^2} \left| -N|T| + \sum_{j=1}^N \chi_{\sigma^{-1}(T)-t}(u(j)) \right|^2 \mathrm{d}t\mathrm{d}\sigma \geq cN^{1/2}.$$

The lower bound depends on the argument in [10] and on estimates in [4], while the upper bound runs as in (1).

References

[1] J. Beck. Irregularities of distribution I. *Acta Math.*, 159 (1987), 1–49.
[2] J. Beck, W.W.L. Chen. *Irregularities of Distribution* (Cambridge University Press, 1987).
[3] J. Beck, W.W.L. Chen. Note on irregularities of distribution II. *Proc. London Math. Soc.*, 61 (1990), 251–272.

[4] L. Brandolini, L. Colzani, G. Travaglini. Average decay of Fourier transforms and integer points in polyhedra. *Ark. Mat.*, 35 (1997), 253–275.
[5] L. Brandolini, S. Hofmann, A. Iosevich. Sharp rate of average decay of the Fourier transform of a bounded set. *Geom. Funct. Anal.*, 13 (2003), 671–680.
[6] L. Brandolini, A. Iosevich, G. Travaglini. Planar convex bodies, Fourier transform, lattice points, and irregularities of distribution. *Trans Amer. Math. Soc.*, 355 (2003), 3513–3535.
[7] L. Brandolini, M. Rigoli, G. Travaglini. Average decay of Fourier transforms and geometry of convex sets. *Rev. Mat. Iberoam.*, 14 (1998), 519–560.
[8] W.W.L. Chen. Fourier techniques in the theory of irregularities of point distribution. *Fourier Analysis and Convexity*, L. Brandolini, L. Colzani, A. Iosevich, G. Travaglini, eds. (Birkhauser).
[9] D.G. Kendall. On the number of lattice points in a random oval. *Quart. J. Math. (Oxford)*, 19 (1948), 1–26.
[10] H.L. Montgomery. *Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis* (CBMS Regional Conf. Ser. in Math. 84, Amer. Math. Soc., Providence, 1994).
[11] A.N. Podkorytov. The asymptotic of a Fourier transform on a convex curve. *Vestn. Leningrad Univ. Mat.*, 24 (1991), 57–65.
[12] B. Randol. On the Fourier transform of the indicator function of a planar set. *Trans. Amer. Math. Soc.*, 139 (1969), 271–278.
[13] M. Tarnopolska-Weiss. On the number of lattice points in a compact n-dimensional polyhedron. *Proc. Amer. Math. Soc.*, 74 (1979), 124–127.
[14] G. Travaglini. Average decay of the Fourier transform. *Fourier Analysis and Convexity*, L. Brandolini, L. Colzani, A. Iosevich, G. Travaglini, eds. (Birkhauser).

Polynomial-time algorithms for multivariate linear problems with finite-order weights

Grzegorz Wasilkowski
(joint work with Henryk Woźniakowski)

There is a host of practical problems that deal with functions of very many variables. In many cases, the required error tolerance for such problems is not too small. Then the classical estimates are asymptotic for n going to ∞ and for fixed the number d of variables, and they are usually of no practical value if n is fixed and d is very large. For instance, the classical discrepancy bounds are of the form $n^{-1}(\log n)^{d-1}$ and become meaningful only when the number n of function evaluations significantly exceeds e^d. This is why, since its introduction in 1994, see [12], there has been an increasing interest in the study of tractability of multivariate problems. Recall that a problem is tractable if it is possible to reduce the initial error ε-times by using a polynomial number of evaluations in ε^{-1} and d; and it is strongly tractable if this number is independent of d. We stress that the upper bound on the number of evaluations should hold for all $\varepsilon \in (0,1)$ and all $d = 1, 2, \ldots$, including the case of huge d and relatively large ε, say $\varepsilon = 10^{-1}$. Algorithms that compute an ε-approximation and use a polynomial number of evaluations in ε^{-1} and d are called polynomial-time algorithms; and if this number does not depend on d they are called strongly polynomial-time algorithms.

There are many results on the tractability of multivariate problems. However, quite a few of them are not constructive, see the survey paper [4] and many papers cited there. Most of the results are obtained for problems defined over general tensor product spaces, including Banach spaces, see, *e.g.*, [3]. As observed in a number of papers, see, *e.g.*, [1, 5, 7, 8], there are important problems, including problems in mathematical finance and physics, that deal with functions which only depend on groups of few variables. That is, the functions depend on all d variables; however, they are sums of terms each of which depends only on few, say q^*, variables. For some applications, the number q^* is fairly small, *e.g.*, $q^* = 1$ or 2. An example of such functions with $q^* = 2m$ is provided by the Coulomb potential function where

$$f(\mathbf{x}) = \sum_{i \neq j,\ i,j=1}^{d} (\|x_i - x_j\|_2 + \alpha)^{-1}$$

for vectors $x_j \in \mathbb{R}^m$ and a positive α. That is, f only depends on groups of two variables each being an m-dimensional vector.

Functions of d variables can be written as the sum of functions of groups x_u of variables with u varying through all subsets of the index set $\{1, 2, \ldots, d\}$. That is, for $\mathbf{x} = [x_1, x_2, \ldots, x_d]$, we have

$$f(\mathbf{x}) = \sum_{u \subset \{1,2,\ldots,d\}} \gamma_{d,u} f_u(\mathbf{x})$$

for some functions f_u depending only on x_j for $j \in u$, and non-negative weights $\gamma_{d,u}$. The essence of the example with the Coulomb potential function is that $\gamma_{d,u} = 0$ for all u with cardinality greater than $2m$. If such a special structure of functions is present in a specific problem, it is said that the problem has finite-ordered weights; see [2, 5] where the concept of finite-order weights has been introduced.

When such a structure is properly used we might be able to obtain efficient algorithms that are polynomial-time or even strongly polynomial time algorithms. Indeed, it has recently been shown in [2, 5] that this is the case for approximating integrals

$$\int_{[0,1]^d} f(\mathbf{x})\, \mathrm{d}\mathbf{x}$$

for Sobolev and Korobov spaces of functions equipped with finite-order weights. In this case, the quasi-Monte Carlo algorithms based on such classical low discrepancy points as Niederreiter, Halton, Sobol, lattice rules and shifted lattice rules are polynomial or even strongly polynomial-time algorithms.

More general problems, including the weighted L_2-approximation problem, have been studied in a recent paper [10]. It was shown there that, under a special assumption (1), these problems are tractable or even strongly tractable for reproducing kernel Hilbert spaces equipped with finite-order weights. More specifically, an upper bound on the number of evaluations needed to compute an ε-approximation was shown to be independent on d and of order ε^{-2} or ε^{-4}; the former dependence for algorithms that use properly chosen linear functional evaluations, and the latter for algorithms that use only function evaluations at properly chosen points. For some problems these bounds are not sharp; however, in full generality, the bound of order ε^{-2} cannot be improved. The bound ε^{-4} is probably not sharp, and the proof of it is non-constructive.

The research presented here may be viewed as a continuation of [10]. Indeed, under slightly different assumptions and using different proof techniques, we provide constructions of polynomial-time algorithms that use only function evaluations for linear problems over reproducing kernel Hilbert spaces equipped with finite-order weights. These algorithms are derived for arbitrary $d \geq 2$ in terms of tensor products of algorithms for $d = 1$ in a way similar to weighted Smolyak algorithms studied in [9], see also [6]. Upper bounds on the number of evaluations needed to compute an ε-approximation for general d are practically the same as for $d = 1$ as far the dependence on ε^{-1} is concerned. Hence, these upper bounds are sharp in ε^{-1} if we use optimal algorithms for $d = 1$. The dependence on d is polynomial and the degree of this polynomial depends on the order of the weights, *i.e.*, on the largest cardinality of u for which $\gamma_{d,u}$ is still non-zero.

We explain our results in more technical terms for the following simplified version of weighted approximation problem, where one wants to recover f with the error measured in a weighted L_2-norm

$$\sqrt{\int_{D_d} |f(\mathbf{x}) - (Af)(\mathbf{x})|^2 \rho_d(\mathbf{x})\, \mathrm{d}\mathbf{x}}.$$

Here $D_d = D \times \ldots \times D$ with $D \subset \mathbb{R}$, $\rho_d = \prod_{k=1}^{d} \rho(x_k)$ is a probability density function on D_d, and Af is an approximation given by an algorithm A. We assume that functions f belong to a reproducing kernel Hilbert space F_d whose formal definition was presented during the talk, see also [11]. The condition (1) from [10] relates the kernel K defining the space and the probability density ρ by assuming that

$$\int_D K(x,x)\rho(x)\, \mathrm{d}x < \infty. \tag{1}$$

Let $A_{d,\varepsilon}$ be one of the proposed algorithms that computes an ε-approximation for the d-dimensional case. Letting $\mathrm{card}(A_{d,\varepsilon})$ denote the corresponding number of function evaluations used by the algorithm $A_{d,\varepsilon}$, we show that for any positive δ, there exists a positive number a_δ such that

$$\mathrm{card}(A_{d,\varepsilon}) \leq a_\delta \varepsilon^{-p(1+\delta)} d^{q^*} \quad \forall\, d, \varepsilon. \tag{2}$$

Here p can be chosen as the smallest exponent for the case $d = 1$, and q^* is the order of the weights, *i.e.*, $\gamma_{d,u} = 0$ for all $u \in \{1, 2, \ldots, d\}$ with the cardinality $|u| > q^*$. In particular, this means that, modulo $(1 + \delta)$, the exponent of ε^{-1} is as small as possible. Since we do not assume that the condition (1) is satisfied, the exponent p can be arbitrarily large. As in [10], $p < 4$ if (1) holds. For smooth problems, however, p is much smaller than 4.

We now comment on the results concerning algorithms that may use arbitrary functional evaluations. As already mentioned, general results with constructive proofs have been obtained in [10] with the exponent $p = 2$. Under an additional assumption and using different proof techniques, we construct optimal algorithms with $\mathrm{card}(A_{d,\varepsilon})$ bounded as in (2). Hence, we may have the exponent p much smaller than 2. We also show that this bound is sharp in both ε^{-1} and d.

Under yet an additional assumption that the weights $\gamma_{d,u}$ depend on u only via $|u|$, we show a necessary and sufficient condition for the approximation problem to be strongly tractable and we present strongly polynomial-time algorithms. We also show that sometimes there is a tradeoff between the minimal exponents of ε^{-1} and d. Indeed, for strongly tractable problems we have a sharp bound of the form

$$\mathrm{card}(A_{d,\varepsilon}) \le c\varepsilon^{-p'} \quad \forall\, d, \varepsilon.$$

Furthermore, (2) also holds; however, the exponent p' is in general larger than p in (2). This means that by increasing the exponent of ε^{-1} we can obtain the bound independent of d.

The results discussed here will be published in [11].

References

[1] R.E. Caflisch, W. Morokoff, A. Owen. Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension. *J. Comp. Finance*, 1 (1997), 27–46.

[2] J. Dick, I.H. Sloan, X. Wang, H. Woźniakowski. Good lattice rules in weighted Korobov spaces with general weights (submitted in 2003).

[3] F.J. Hickernell, I.H. Sloan, G.W. Wasilkowski. On tractability of weighted integration for certain Banach spaces of functions. *Monte Carlo and Quasi-Monte Carlo Methods 2002*, pp. 51–71, H. Niederreiter, ed. (Springer, 2004).

[4] E. Novak, H. Woźniakowski. When are integration and discrepancy tractable? *Foundation of Computational Mathematics, Oxford, 1999*, pp. 211–266, R.A. DeVore, A. Iserles, E. Süli, eds. (Cambridge University Press, Cambridge, 2001).

[5] I.H. Sloan, X. Wang, H. Woźniakowski. Finite-order weights imply tractability of multivariate integration. *J. Complexity*, 20 (2004), 46–74.

[6] S.A. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl. Akad. Nauk SSSR*, 4 (1963), 240–243.

[7] X. Wang, K.T. Fang. Effective dimensions and quasi-Monte Carlo integration. *J. Complexity*, 19 (2003), 101–124.

[8] X. Wang, I.H. Sloan. Why are high-dimensional finance problems often of low effective dimension? (submitted in 2003).

[9] G.W. Wasilkowski, H. Woźniakowski. Weighted tensor-product algorithms for linear multivariate problems. *J. Complexity*, 15 (1999), 402–447.

[10] G.W. Wasilkowski, H. Woźniakowski. Finite-order weights imply tractability of linear of linear multivariate problems (submitted in 2003).

[11] G.W. Wasilkowski, H. Woźniakowski. Polynomial-time algorithms for multivariate linear problems with finite-order weights (in progress).

[12] H. Woźniakowski. Tractability and strong tractability of linear multivariate problems. *J. Complexity*, 10 (1994), 96–128.

Integration, tractability, discrepancy
Henryk Woźniakowski

In this talk we discuss recent progress on solving multivariate integration when the number d of integrand variables is in hundreds or thousands. Such high dimensional integrals occur in many applications including financial mathematics and computational physics. We want to approximate

$$I_d(f) = \int_{D_d} \rho_d(t) f(t)\,\mathrm{d}t,$$

where $D_d \subset \mathbb{R}^d$, the function ρ_d is non-negative and its integral over D_d is one, and real f belongs to a normed class F_d of integrable functions.

We restrict our attention to the worst case setting although different settings such as average, randomized and quantum are also studied. We approximate $I_d(f)$ by a quadrature rule

$$Q_{n,d} = \sum_{j=1}^{n} a_j f(t_j).$$

Here, t_j are sample points from the domain of f, and a_j are real numbers. For QMC (quasi-Monte Carlo) rules we have $a_j = 1/n$. The number n denotes the total number of function values used by $Q_{n,d}$.

The worst case error of $Q_{n,d}$ is defined as its worst performance for approximating integrals for the unit ball of F_d,

$$e(Q_{n,d}) = \sup_{f \in F_d,\ \|f\| \le 1} |I_d(f) - Q_{n,d}(f)|.$$

Clearly, the cost of using $Q_{n,d}$ is proportional to n, and therefore we would like to use n as small as possible with the worst case error below a given threshold. For $n = 0$, we formally set $Q_{0,d} = 0$, and then the worst case error is called the initial error which is the norm $\|I_d\|$ of the integration in the space F_d.

We consider two error criteria. The first one is the absolute error criterion in which we want to guarantee that the worst case error is at most ε, *i.e.*, $e(Q_{n,d}) \le \varepsilon$. The second one is the normalized error criterion in which we want to guarantee that the worst case reduces the initial error by a factor of ε, *i.e.*, $e(Q_{n,d}) \le \varepsilon \|I_d\|$. Here, the error parameter $\varepsilon \in (0,1)$.

Define $n(\varepsilon, F_d)$ as the minimal number of function values needed to satisfy the absolute or normalized error criterion. If $n(\varepsilon, F_d)$ can be bounded by a polynomial in ε^{-1} and d, then multivariate integration in F_d is called tractable, *i.e.*, there exist non-negative numbers C, p and q such that

$$n(\varepsilon, F_d) \le C \varepsilon^{-p} d^{\,q} \qquad \forall\, \varepsilon \in (0,1),\ d = 1, 2, \ldots.$$

If $q = 0$ in the bound above, then $n(\varepsilon, F_d)$ is bounded by a polynomial in ε^{-1} independently of d, and then multivariate integration in F_d is called strongly tractable.

The study of tractability, not only for multivariate integration, has recently become a popular research subject. The main point is to identify classes F_d for which strong tractability or tractability hold. A survey of current results and approaches may be found in [4].

For some spaces F_d, the worst case error is the same as the L_2 or star discrepancy. In this case, tractability is equivalent to finding discrepancy bounds of n sample points with polynomial dependence on d and converging to zero as a positive power of n^{-1}. For instance, consider the Sobolev space of functions defined over $[0,1]^d$ which are one time differentiable with respect to all variables and satisfying the boundary condition $f(t) = 0$ if at least one of the components of t is zero. The norm of f in this space is defined by the L_2 norm of $\partial^d f/\partial t_1 \ldots \partial t_d$. Then the worst case of $Q_{n,d}$ is exactly the L_2 discrepancy of t_j. It turns out that for the absolute error criterion, multivariate integration is strongly tractable. On the other hand, for the normalized error criterion, multivariate integration is intractable.

If we remove the boundary condition and redefine the norm in the L_p sense by taking projections of f as in the Zaremba and Koksma-Hlawka (in)equalities, the situation changes. For the L_2 case, multivariate integration is intractable for the two error criteria. Surprisingly enough, if we switch to the L_1 case, then the worst case error is the same as the star discrepancy. In this case, the two error criteria are the same since the initial error is one. It turns out that we now have tractability, but not strong tractability, as proven in [2].

It was observed in many papers that integrands of practical importance have additional properties which are not properly modeled by classical spaces. Namely, in many cases, integrands are sums of functions that depend only on groups of a few variables, or that they depend on the successive variables in the diminishing sense. This additional structure of integrands may be modeled by weighted spaces of functions in which each group of variables has a weight moderating its importance. Tractability for weighted spaces has been initiated in [5]. For some spaces we know necessary and sufficient conditions on the weights to obtain strong tractability or tractability. For instance, take the Sobolev space without the boundary condition with the L_2 norm as above, and equip the space with the weight γ_j for each variable. This means that the norm of the space is redefined and $\|f\| \leq 1$ with small γ_j means that f weakly depends on the jth variable. Then, in particular, strong tractability of multivariate integration holds iff $\sum_{j=1}^{\infty} \gamma_j < \infty$ as proven in [5, 3].

What seems especially promising is the idea of finite-order weights as introduced in [1] and [6]. The weights are finite-order if they are zero for all groups of variables of cardinality greater than, say, k. Here k is independent of d and usually relatively small. For instance, for some financial problems, $k = 1$ or $k = 2$, and for some problems in computational physics, $k = 6$. It turns out that for finite-order weights multivariate integration is tractable and often strongly tractable. However, the error bounds are exponential in k. This, in turn, is not dangerous as long as k is not large. Furthermore, classical sample points such as Halton, Niederreiter or Sobol lead to tractable error bounds.

References

[1] J. Dick, I.H. Sloan, X. Wang, H. Woźniakowski. Good lattice rules in weighted Korobov spaces with general weights (submitted in 2003).

[2] S. Heinrich, E. Novak, G.W. Wasilkowski, H. Woźniakowski. The inverse of the star-discrepancy depends linearly on the dimension. *Acta Arith.*, 96 (2001), 279–302.

[3] E. Novak, H. Woźniakowski. Intractability results for integration and discrepancy. *J. Complexity*, 17 (2001), 388–441.

[4] E. Novak, H. Woźniakowski. When are integration and discrepancy tractable? *Foundation of Computational Mathematics, Oxford, 1999*, pp. 211–266, R.A. DeVore, A. Iserles, E. Süli, eds. (Cambridge University Press, Cambridge, 2001).

[5] I.H. Sloan, H. Woźniakowski. When are quasi-Monte Carlo algorithms efficient for high dimensional integrals? *J. Complexity*, 14 (1998), 1–33.

[6] I.H. Sloan, X. Wang, H. Woźniakowski. Finite-order weights imply tractability of multivariate integration. *J. Complexity*, 20 (2004), 46–74.

Reporter: William W.L. Chen

Participants

Prof. Dr. Imre Barany
barany@renyi.hu
barany@math.ucl.ac.uk
Alfred Renyi Institute of Mathematics
Hungarian Academy of Sciences
P.O.Box 127
H-1364 Budapest

Prof. Dr. Jozsef Beck
jbeck@math.rutgers.edu
Dept. of Mathematics
Rutgers University
Busch Campus, Hill Center
New Brunswick, NJ 08903 – USA

Prof. Dr. William W.L. Chen
wchen@maths.mq.edu.au
Department of Mathematics
Macquarie University
Sydney NSW 2109 – Australia

Dr. Benjamin Doerr
bed@numerik.uni-kiel.de
Mathematisches Seminar
Bereich 2
Christian-Albrechts-Universität
Christian-Albrechts-Platz 4
D–24098 Kiel

Dr. Michael Drmota
michael.drmota@tuwien.ac.at
Michael.drmota@dmg.tuwien.ac.at
Institut für Diskrete Mathematik
und Geometrie
Technische Universität Wien
Wiedener Hauptstr. 8 - 10
A-1040 Wien

Dr. Michael Gnewuch
mig@numerik.uni-kiel.de
Mathematisches Seminar
Christian-Albrechts-Universität
Kiel
D–24098 Kiel

Nils Hebbinghaus
nhe@numerkik.uni-kiel.de
Mathematisches Seminar
Bereich 2
Christian-Albrechts-Universität
Christian-Albrechts-Platz 4
D–24098 Kiel

Prof. Dr. Stefan Heinrich
heinrich@informatik.uni-kl.de
Fachbereich Informatik
Universität Kaiserslautern
D–67653 Kaiserslautern

Prof. Dr. Jiri Matousek
matousek@kam.mff.cuni.cz
Department of Applied Mathematics
Charles University
Malostranske nam. 25
118 00 Praha 1 – Czech Republic

Prof. Dr. Erich Novak
novak@minet.uni-jena.de
novak@mathematik.uni-jena.de
Fakultät für Mathematik
und Informatik
Friedrich-Schiller-Universität
Ernst-Abbe-Platz 2
D–07743 Jena

Prof. Dr. Friedrich Pillichshammer
friedrich.pillichshammer@jku.at
Institut für Analysis
Universität Linz
Altenbergstr. 69
A-4040 Linz

Prof. Dr. Michael Rudnev
M.Rudnev@bris.ac.uk
School of Mathematics
University of Bristol
University Walk
GB-Bristol BS8 1TW

Dr. Tomasz Schoen
schoen@amu.edu.pl
Zakad Matematyki Dyskretnej
Adam Mickiewicz University
pokoj nr B3-31
ul. Umultowska 87
61-614 Poznan – Poland

Prof. Dr. Maxim Skriganov
skrig@pdmi.ras.ru
St. Petersburg Branch of
Mathematical Institute of
Russian Academy of Science
Fontanka 27
191023 St. Petersburg – Russia

Prof. Dr. Anand Srivastav
asr@numerik.uni-kiel.de
Mathematisches Seminar
Christian-Albrechts-Universität
Kiel
Ludewig-Meyn-Str. 4
D–24118 Kiel

Prof. Dr. Robert F. Tichy
tichy@weyl.math.tu-graz.ac.at
tichy@tugraz.at
Institut für Mathematik
Technische Universität Graz
Steyrergasse 30
A-8010 Graz

Prof. Dr. Giancarlo Travaglini
travaglini@matapp.unimib.it
Dipt. Matematica e Applicazioni
Universita' Milano Bicocca
Via Bicocca degli Arcimboldi 8
I-20126 Milano

Prof. Dr. Grzegorz W. Wasilkowski
greg@cs.engr.uky.edu
Computer Science Department
University of Kentucky
773 Anderson Hall
Lexington, KY 40506-0046 – USA

Prof. Dr. Emo Welzl
emo@inf.ethz.ch
Theoretische Informatik
ETH-Zürich
ETH-Zentrum
CH-8092 Zürich

Prof. Dr. Henryk Wozniakowski
henryk@cs.columbia.edu
Department of Computer Science
Columbia University
450 Computer Science Building
1214 Amsterdam Avenue
New York, NY 10027-7003 – USA

Mathematisches Forschungsinstitut Oberwolfach

Report No. 14/2004

Analysis and Design of Electoral Systems

Organised by
Michel L. Balinski (Paris)
Steven J. Brams (New York)
Friedrich Pukelsheim (Augsburg)

March 7th – March 13th, 2004

Introduction by the Organisers

Electoral systems transform numbers into numbers. In order not to be blinded by the intricacies of any particular case, as interesting as particular cases may be, it is imperative to view theses transformations as functions, or as relations, and to study their properties and structure in their full generality. Promoting the mathematical foundations of the topic cannot be a goal in and for itself, but must reflect the practical needs defined by the problems themselves that are traditionally treated by such non-mathematical fields as political science, economics, social choice theory, constitutional law, etc. This diversity was well reflected in the expert fields represented by the conference participants, as well as in the wide range of topics covered. They may broadly be summarized under three headings: (1) voting schemes that aggregate many individual preference rankings of a given set of alternatives into a single global ranking; (2) proportional representation schemes that map vote counts (or population counts) into parliamentary representation; and (3) the determination of electoral districts that reasonably reflect geographical, political, and social structures.While mathematics at large provides the appropriate language to investigate these problems, the tools that are used draw on specific mathematical fields, including:

- discrete mathematics, in that preference rankings are (usually) partial orderings or permutations;
- combinatorial optimization, when finitely many units are assigned in a accordance with some optimality criterion;

- stochastics, where weights that are virtually continuous (proportions of votes) are to be approximated by weights that are intrinsically discrete (numbers of seats); and
- geometry, identifying the intricacies of voting with the symmetry structure of permutation groups, and with the geometry of the probability simplex.

The mathematical analysis of electoral systems has direct repercussions on current political issues. To mention but three, there is the question whether elections of the US president would be less likely to be disputed, and their outcomes considered more legitimate, were he to be elected directly. Another is the analysis of qualified majority rules in the EU Council of Ministers as stipulated in the draft constitution proposed by the European Convention 2003. A third relates to developing biproportional apportionment methods that simultaneously accommodate partisan and regional representation of the entire electorate.

Somewhat unusual for an Oberwolfach conference was the rather broad mix of participants, representing the fields of mathematics, physics, economics, political science, and psychology. It was this breadth that virtually all participants experienced as challenging and fruitful. As always, the atmosphere was congenial for scientific discussions, and the Institute in its Schwarzwald setting made people feel creative by just being there . . . particularly so, since the sky was blue and the sun was shining (most of the time).

M.L. Balinski
S.J. Brams
F. Pukelsheim

Analysis and Design of Electoral Systems

Table of Contents

Abstracts

Measuring Power in At-Large Representation

Paul Edelman

This talk presents a formal analysis of a voting game inspired by a common type of local legislature: a legislative body in which some of the seats are allocated by majority vote in equipopulous districts, and some of the seats are elected by an at-large majority vote. Such legislatures are common in city councils of large metropolitan areas and county boards. For instance, the Metropolitan Council of Nashville and Davidson County in the state of Tennessee consists of 40 members, 5 of whom are at-large and the remaining 35 are elected from separate districts.

The motivation for this study is the question of how to decide how many of each type of representative is optimal, given a fixed total number of representatives. The analysis will follow in the tradition of Banzhaf [1]. The legislature will be modelled by a weighted voting game and I will compute the power of an individual voter by using the composition of this weighted voting game with the majority game in each component. It will be evident by symmetry that the actual Banzhaf power of each voter is identical, but the sum of the Banzhaf power over all of the voters, what Felsenthal and Machover call the *sensitivity* (see [4, Def. 3.3.1]) of this composite game, is maximized when the number of at-large representatives is the square-root of the total number of representatives.

Two novel mathematical issues will arise in the analysis. In the composite game that I study, the underlying sets of players are not disjoint and the behaviour of a player who appears multiple times can be different in different coordinates of the game. The usual definition of composition requires disjoint sets of players [5, XI.2.2] or at least that if the player sets are not disjoint then a player must behave the same way in all coordinates [4, Definition 2.3.12]. To my knowledge there has not been a theoretical account of games without these requirements and so I will provide one.

What is even more interesting is that even though there has not been a theoretical account, results concerning the Banzhaf measure of such games have been used. In *New York Board of Estimate v. Morris*, the U. S. Supreme Court considers (and ultimate rejects) a Banzhaf analysis of a composite game involving three at-large representatives and separate representatives for each borough [2, p. 697]. Banzhaf has considered a districted presidential election game with similar features. Even though in these two different situations the Banzhaf power was computed in similar ways, we will see that they are different types of games and perhaps require different measures.

In this talk I will concentrate solely on the game theoretic aspects of this result. In particular I will not enter the fray as to whether the Banzhaf measure is a realistic measure of power in a voting game. What this paper does show is that the Banzhaf measure provides a way to give a theoretical justification for a certain number of at-large representatives in a legislative body. I know of no other model that does so. In order to do this, a new theoretical account of composite games is required.

The results discussed in this talk have recently appeared [3].

References

[1] Banzhaf, J. F. III: *Multi-member electoral districts — do they violate the "one man, one vote" principle.* Yale L. J. **75** (1966), 1309–1338.

[2] *Board of Estimate of City of New York et. al. v. Morris et. al.* (1989), 489 US 688.

[3] Edelman, Paul H.: *Voting power and at-large representation.* Math. Soc. Sci. **47** (2004), 219–232.

[4] Felsenthal, D. S. and Machover, M.: *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes.* Edward Elgar, Cheltenham 1998.

[5] Owen, G.: *Game Theory.* 3^{rd} ed., Academic Press, San Diego 1995.

Voter Souvereignty

Remzi Sanver

(joint work with Steven J. Brams)

The thesis of this paper is that *several* outcomes of single-winner elections may be acceptable. Perhaps the most dramatic recent example illustrating this proposition is the 2000 U.S. presidential election, in which George W. Bush won the electoral vote — disputed though it was in Florida — and Al Gore won the popular vote. To be sure, the extreme closeness of this election was unusual. But many elections, especially those with three or more candidates, may have more than one acceptable outcome. For example, even when there is a Condorcet winner, who can defeat every other candidate in pairwise contests, there may be a different Borda-count winner, who on the average is ranked higher than a Condorcet winner. If there is no Condorcet winner because of cyclical majorities, the Condorcet cycle may be broken at its weakest link to select the strongest candidate in the cycle, who need not be the Borda winner.

That different voting systems can give different outcomes is, of course, an old story. The observation that different outcomes may satisfy different social-choice criteria is also old hat (Nurmi, 1999, 2002, give many examples). What is new here is our claim that in an election with three or more candidates, other outcomes — not just the Condorcet winner, the Borda-count winner, or the strongest candidate in a cycle — may be *more* acceptable to the electorate. In fact, even a Condorcet loser, who would lose in pairwise contests to every other candidate, may turn out to be the most acceptable candidate. To justify this last statement, we need to define some measure of "acceptability." If voters rank candidates from best to worst, where they draw the line in their rankings between acceptable and unacceptable candidates offers one such measure. It is precisely this information that is elicited under approval voting (AV), whereby voters can approve of as many candidates as they like or consider acceptable. This gives them the opportunity to be *sovereign* by expressing their approval for any *set* of candidates, which no other voting system permits[1]. Call a candidate a *Pareto candidate* if there is no other candidate that all voters rank higher. We demonstrate that candidates selected under AV always include at least one Pareto candidate. In fact, AV *dominates* so-called scoring systems, including plurality voting (PV) and the Borda count (BC), with respect to the election of Pareto candidates: a Pareto candidate elected by a scoring system is always elected by AV for some sincere and admissible strategies, but not vice versa. This is also true for ranking systems that do not rely on scoring, including the Hare system of single transferable vote (STV) and the majoritarian compromise (MC). But if AV does a better job of finding Pareto candidates, doesnt it open the door to a plethora of possibilities? Isnt this a vice rather than a virtue, as some have argued (e.g., Saari and Van Newenhizen, 1988a; Saari, 1994, 2001)?[2] This argument might have merit if the plethora of possibilities were haphazard choices that could easily be upset when voters are manipulative. But we show that AV often leads to Nash-equilibrium outcomes, from which voters with the same preferences will have no incentive to depart. Moreover, if voters with different preferences are able to coordinate their choices and none has an incentive to depart, AV *guarantees* the election of a unique Condorcet winner (if one exists).

The latter notion of stability is that of a *strong* Nash equilibrium, which yields outcomes that are invulnerable to departures by *any* set of voters. None of the other voting systems we assay guarantees that a unique Condorcet winner, and only a Condorcet winner, will be a strong Nash equilibrium outcome when voters are sincere. While AV offers this guarantee, however, it also allows for other Nash-equilibrium outcomes, including even a Condorcet loser, who may be the most acceptable candidate, even in equilibrium. In section 2, we define preferences and strategies under AV and give an example that illustrates the choice of sincere, admissible strategies. In section 3 we characterize AV outcomes, describing the "critical strategy profile" that produces them, and compare these outcomes with those given by other voting systems. Among other things, we show that no "fixed rule," in which voters vote for a predetermined number of candidates, always elects a unique Condorcet winner, suggesting the need for a more flexible system. The stability of outcomes under the different voting systems is analyzed in section 4, where we show that Nash equilibria and strong Nash equilibria may vary from system to system. Also, Condorcet voting systems, which guarantee the election of Condorcet winners when voters are sincere, may not elect Condorcet candidates in

[1]Voter sovereignty should be distinguished from Arrow's (1963) condition of "citizen sovereignty," whereby for any two alternatives a and b, if all voters prefer a to b, a cannot be prohibited as the social choice. If voters are "sincere," AV satisfies citizen sovereignty, because all voters who approve of b will also approve of a. Note that voter sovereignty describes the behaviour of individual voters whereas citizen sovereignty is a property of a voting system.

[2]The critique of AV by Saari and Van Newenhizen (1988a) provoked an exchange between Brams, Fishburn, and Merrill (1988a, 1988b) and Saari and Van Newenhizen (1988b) over whether the plethora of AV outcomes more reflected AVs "indeterminacy" (Saari and Van Newenhizen) or its "responsiveness" (Brams, Merrill, and Fishburn); other critiques of AV are referenced in Brams and Fishburn (2003). Here we argue that which outcome is chosen should depend on voters judgments about the acceptability of candidates rather than standard social-choice criteria, whichas we will show may clash with these judgments.

equilibrium. In section 5 we show that rational departures by voters from unstable outcomes under other voting systems may not induce AV outcomes, but rational departures under AV always do. Hence, outcomes under AV form a closed set.

Nonstrong Nash equilibria might be thought of as possessing a kind of local stability, whereas strong Nash equilibria possess a global stability. These different kinds of equilibria may coexist, which is to say that *which* stable outcome is chosen will depend on which candidates voters consider acceptable and whether they coordinate their choices. In large-scale public elections, coordination is typically done when voters draw inferences from polls, not by face-to-face communication, which is commonplace in smaller settings like committees. That a Condorcet candidate is a globally stable choice under AV should not be surprising. What is more surprising is that such a candidate can be upset if (i) coordination is difficult and (ii) many voters consider another candidate more acceptable.

Speaking normatively, we believe that voters *should* be sovereign, able to express their approval of any set of candidates. Likewise, a voting system *should* allow for the possibility of multiple acceptable outcomes, especially in close elections. That AV more than other voting systems is responsive in this way we regard as a virtue. That it singles out as strong Nash equilibrium outcomes unique Condorcet winners may or may not be desirable. We discuss these and other questions related to the nature of acceptable outcomes in section 6, where we suggest that "acceptability" replace the usual social-choice criteria for assessing the satisfactoriness of election outcomes chosen by sovereign voters.

A Minimax Procedure for Negotiating Multilateral Treaties

Marc Kilgour

(joint work with Steven J. Brams and Remzi Sanver)

In this paper we propose a procedure for reaching agreement on multilateral treaties by finding a compromise that is as close as possible to the preferences of all negotiating states. By "close" we mean that the maximum (Hamming) distance from the compromise to the position of any negotiator is a minimum. This compromise, which we call a *minimax* outcome, is most likely to be an acceptable resolution because it leaves no state too aggrieved. We argue also that it reduces any incentive states might have to misrepresent their preferences to induce a better outcome and is, therefore, relatively invulnerable to strategizing.

What we propose differs from the usual method of reaching an agreement in multilateral treaty negotiations. Normally, states vote separately on each provision of a treaty, often starting from a "single negotiating text." If a simple or qualified majority supports a provision, it is included in the treaty; otherwise, it is excluded. We call this the *Majority Voting* (MV) procedure. MV chooses compromises that may differ from minimax outcomes; we show that it produces all outcomes that minimize the total (or average) distance to the negotiators' positions, or *minisum* outcomes.

Our analysis assumes that negotiations have reached a stage whereby the provisions of a treaty in dispute

- can be specified;
- are of approximately equal significance to all states;
- are relatively independent of each other.

Moreover, we assume that bargaining on all provisions is binary: each provision must be included in the final treaty (coded as 1) or rejected (coded as 0). We recognize that it may be difficult to achieve these conditions in practice, but we suggest how they might be approximated. Moreover, it is certainly feasible to extend our definitions to non-binary issues, but we do not pursue such an extension here.

Assume that a treaty to be negotiated by n states (players, or countries) has k possible provisions. The possible treaties can therefore be represented as binary k-vectors, $(p_1, p_2, \ldots, p_k)$, where each p_i equals 0 or 1. Such binary vectors will be called *combinations*. We simplify notation by writing combinations without punctuation, so that $(1, 1, 0)$ becomes 110. Note that the total number of combinations is 2^k.

Because the procedure we propose for forging consensus in multilateral treaty negotiations requires bargainers' preferences as inputs, we begin with a model of preferences over combinations. We assume that each state has a most preferred treaty; we call this combination its ideal point or *top preference*. We further assume that the state ranks other combinations according to a *spatial model* — that is, the ranking of a combination depends only on its distance from the top preference. (As distance between two binary k-vectors, p and q, we use the Hamming distance, $d(p, q)$, which equals the number of components on which p and q differ.) So a state prefers a combination that lies closer to its top preference, and equally prefers combinations that are at equal distances from its top preference. For example, if $k = 2$ and a state approves of the first provision but disapproves of the second, its top preference is 10, and its complete preference ranking is $10 \succ 11 \sim 00 \succ 01$, where "$\succ$" indicates preference and "$\sim$" indicates indifference.

The procedure we propose for forging consensus in multilateral treaty negotiations is based on "fallback bargaining" [1]. Let r be an integer satisfying $0 < r \leq n$. The FB_r outcomes are those attained by applying fallback bargaining with parameter r, which is the following procedure (starting with $d = 0$):

> Assume bargainers approve only combinations at distance d or less from their top preferences. If one or more combinations is approved by at least r bargainers, then those combinations with the most approvals are the FB_r outcomes. If no combination is approved by at least r bargainers, increase d by 1 and repeat.

Thus, in fallback bargaining each state begins by considering only its most preferred outcome acceptable. If there is not sufficient consensus (as measured by r) in the resulting acceptability sets, the condition is relaxed incrementally; now each state finds acceptable not only its top preference, but also any other combination at distance $d = 1$ from it. Acceptability sets expand in this way until there is some outcome that is acceptable to at least r states.

Although we discuss all fallback bargaining outcomes, we focus on the case of fallback bargaining with unanimity, or $r = n$, which we show produces exactly the minimax outcomes, and on the comparison of these outcomes with the minisum outcomes produced by MV. Whereas FB_n outcomes are in the Rawlsian tradition of minimizing the largest deviations from a compromise, MV outcomes are in the utilitarian tradition of minimizing average departures.

We examine some of the social-choice properties of these procedures. The first step in applying any procedure is for the bargainers to report their top preferences; a procedure is *manipulable* if it may be in a side's interest to report its top preference falsely. We show that the MV procedure is non-manipulable, whereas all FB_r procedures are vulnerable to manipulation. But in any realistic situation with incomplete information about the preferences of two or more other parties, FB_n would be extremely difficult to manipulate. Moreover, and unlike MV, FB_n is completely unaffected by clones, or duplicates of an existing bargainer.

Maximin outcomes seem superior compromises in many bargaining situations, such as OilPol 54, the 1954 negotiation of $n = 32$ states over oil-pollution controls on the high seas. This negotiation concerned $k = 4$ binary issues. Because states could abstain, as well as vote yes or no, on any provision, it was necessary to extend our preference model to account for issues on which a state expresses no preference. Using public statements before and during the conference and other information, we were then able to estimate the top preferences of each participant, as well as any indifferences. We then applied the FB_n procedure, and found six minimax outcomes, all of which differed from the MV outcome, which was the historical outcome of OilPol 54. We argue that one of these outcomes might have been a better choice than the actual outcome, and we suggest approval voting as a way of choosing among the six.

Complex multilateral negotiations frequently involve individual states or overlapping blocs of states, scores of provisions, and considerable maneuvering by the bargainers to try to achieve a strategic advantage. We believe that our proposed procedure would encourage them to be honest, render their negotiations more open, and make the compromises they achieve as acceptable as possible to all bargainers.

References

[1] Brams, Steven J. and D. Marc Kilgour: *Fallback Bargaining.* Group Decision and Negotiation **10** 4 (2001), 287 – 316.

From Principles of Representation to Electoral Methods

Victoriano Ramírez

Politicians habitually establish certain objectives or principles when designing an electoral system. Some of their decisions are influenced by national tradition and history. The system may be of the majority rule type, proportional, or some mixture thereof, and the design of the electoral circumscriptions may respond to previously established geographical boundaries or administive distinctions.

One objective that should not be forgotten is that of governability. A Parliament comprising a large number of political parties, with the ensuing potential for a number of winning coalitions, may give way to continual changes of government, and therefore to the instability of the regime, an undesirable consequence in any case.

Another objective of politicians aiming for systems of proportional representation is to achieve a high proportionality among the global votes for the parties and the total seats obtained by each. Or, taken one step further, the aim might be to obtain a double proportionality between votes and seats, insofar as the parties and the circumscriptions are concerned (as in Mexico from 1989 onward). In other countries, the system of mixed election represents the search for global proportionality. Germany is the most distinguished example of this. Despite the fact that half their Upper House is chosen in uninominal districts and the rest according to the party lists with conditioned proportionality, the German electoral system presently stands as one guaranteeing high proportionality.

Similarly, in the context of social election, certain prior objectives are usually set forth [6, 7]. When it is necessary to choose a representative, be it the Presidency of France or the position of Rector in a given Spanish University, it would be logical to establish as a basic principle that the Condorcet winner will be declared the winner of the election. Many other objectives and principles [1] enter into play as well in the different electoral processes that take place in modern day democracies.

On occasion however, the various principles established are not entirely compatible, while at other times there may be different ways of attaining them, some more satisfactory than others. This is why rigorous analysis is a prerequisite for any electoral stipulations. Politicians should agree on which objectives and principles are the most adequate and appropriate, and in fact they are often the very persons who establish the rules and define the electoral system. Ignoring or underestimating the variables involved in electoral processes, together with their different properties and interrelations, can lead to a failure in achieving the stated objectives, as well as other unpleasant surprises, such as the discovery that certain laws are impossible to apply. A good number of contradictions and inconsistencies can be seen with a careful look at the Mexican Constitutions of 1989 and 1994 and the corresponding Electoral Legislation [2]. But these are not the only cases. In many countries there are confusing electoral systems with scarcely acceptable implications.

I would like to present here several examples of electoral systems under which the ultimate consequences are not necessarily a reflection of the original principals that inspired the electoral design, or which have given rise to unforeseen negative consequences. All the examples refer to different electoral processes that have taken place within Spain in recent years. In Spain we have two Houses of Parliament: the Congress (*Congreso de los Diputados*) and the Senate (*Senado*).

The first example is relative to the size of constituencies for the Congressional elections. The actual Electoral Law of 1985 has led to tremendous differences between the value of a vote from a large constituency or a small one, the latter reaping the relative benefits. The vote of five citizens from Madrid holds as much electoral weight as that of one inhabitant of Soria — far from the democratic principal of "one person, one vote! We must bear in mind the requirement that each constitutncy should receive at least two seats, and that there is no maximum restriction, then using the Webster method we obtain an alternative for which the previous ratio is reduced to 2.6.

The second example is relative to the allotment of Congressional seats to the parties in Spain. Thus, the spanish electoral system produces some important imbalances between certain regional parties and those national parties that obtain similar results insofar as the number of votes. It does not guarantee a significant bonus to the most voted party, although absolute majorities have resulted from half of the elections held. For example in the last election (March, 2004): the national party IU has obtained $1,269,532$ votes and 5 seats, the regional party PNV has obtained $417,154$ votes and 7 seats. Also the winner party, the PSOE has obtained $10,909,687$ votes and 164 seats. In average, each seat of PSOE costs $66,500$ votes, each seat of PNV costs $59,600$ votes and each seat of IU costs $253,900$ votes. It is possible to define an electoral system that leads to a greater proportionality while avoiding unfairness such as that cited for the above cases, even offering a bonus to the most voted party in order to contribute to governability [4]. One mean of ensuring high proportionality is to carry out a second allotment reflecting the total votes of the parties, with minimal conditioning factors based on the results of the initial allotment, as done in Germany.

The third example is the Senatorial election in Spain. In Spain the elections to the Congress and the Senate are celebrated simultaneously. Most of the senators are elected in the same constituencies as the members of Congress. But in this case four are chosen, regardless of the circumscriptions population. The voters usually behave along the same lines as for Congress, despite the fact that the voting system is very different. In the Senate, there are candidate names from all the parties on a single list, with groupings by party. A maximum of three names for each party is established, and the voter can mark only up to three candidate names. This method is similar to Approval Voting [3], but the number of candidates who can be approved, for every elector, is *limited to* 3. Therefore, most voters choose those three names proposed by the party they voted for in the congressional contest. If they opt to mark just one or two names, it would tend to be in favor of the first candidate(s) presented by the preferred party for Congress. The consequence of that behavior is that the most voted party in the lists for the Congress obtains, systematically, three senators and the second party more voted a senator obtains, independently of the relation of votes between both and with respect to the other parties.

The strategy of the Spanish in the elections to the Senate meets moreover reinforced in the university elections, where there is used a method of election very similar to Approval Voting (with different limitations). If greater proportionality in the Senate is desired (or in the universities elections), it would be preferable, for a behavior of the voters like the one mentioned previously, to use as method of social election a Borda-type method with the following weights: $1, \ldots, 1/3, \ldots$.

This paper sheds some critical light on several electoral systems and practices that can be seen in Spain (the constituencies size, the advantage of the main regional parties over the similar national parties, the election of the Senators, the higly manipulable electoral system to determinig university representatives or Juntas). Notwithstanding all these drawbacks, the electoral processes of the Congress, the Senate, and the municipal, regional or European elections do function in a positive sense in that they are applicable in all cases.

On the other hand, I introduce a new property for the proportionality: *Limited loss of seats in coalitions*. We put forth that a method has a limited loss of seats in the case of coalitions: the fusion of $2r$ or $2r+1$ parties does not entail a loss of more than r seats. Then, a necessary condition, for a divisor method, to imply a limited loss of a seat is that

$$d(s) \in \left[s + \frac{1}{2}; s + 1\right].$$

(If $d(s) = s+t$ or the fusion is of $2r$ parties, the previous condition is also sufficient).

In accordance with this property and the properties of the parametric methods [5], I think that the most reasonable option is to use divisor methods of the parametric family from Webster to Jefferson in approaching problems of proportional allotment.

References

[1] Balinski, M. L., Young, H. P.: *Fair representation: Meeting the ideal of One Man, One Vote*. New Haven 1982.

[2] Balinski, M. L., Ramírez, V.: *A case study of electoral manipulation: The Mexican laws of 1989 and 1994*. Electoral Studies **15** (1996), 203–217.

[3] Brams, S. J, Fishburn, P. C.: *Approval Voting*. Birkhäuser Boston 1983.

[4] Márquez, M. L., Ramírez, V.: *The Spanish electoral system: Proportionality and governability*. Annals of Operations Research (1998), 45–59.

[5] Ramírez, V., Márquez, M.L., Pérez, R.: *Parametric subfamilies of apportionment methods*. Advances in Computational Mathematics, Marcel Dekker (1999) 471–479.

[6] Saari, D.: *Geometry of voting*. Springer 1994.

[7] Taylor, A.: *Mathematics and politics*. Springer 1985.

BAZI — A Java Program for Proportional Representation

Friedrich Pukelsheim

BAZI is a freely available JAVA-Program, permitting the user to experiment with various apportionment methods, and to assess their relative merits on the basis of real data rather than abstract theory.

The pertinent theory is available in the seminal monograph [4] by Balinski and Young. Among all possible apportionment methods, the authors single out two important subclasses. The first class consists of divisor methods, the second of quota methods. BAZI features just two quota methods, the method of greatest remainders (Hamilton, Hondt, Hagenbach-Bischoff), and the Droop method.

However, a central message of the Balinski/Young monograph is that divisor methods are generally more appropriate for the apportionment problem. Of these, BAZI offers two parametric families, the divisor methods with stationary roundings, and the divisor methods with powermean roundings; for details see [5, p.357].

The powermean methods are more important from a historical point of view, comprising the five traditional methods of Adams, Dean, Hill, Webster, and Jefferson. In contrast, the stationary methods are more amenable to a mathematical analysis. BAZI relies on an algorithm [5, p. 378] whose computational complexity is minimum [6, p. 154].

On the computer screen, BAZI comes up with the graphical user interface split into three panels, the input field to the left, the methods field in the middle, and the output field on the right.

The input field invites the user to key in data of his or her own, or to read in a data file that the user has created, or to load data from the extensive data base.

In the methods field the user can select a house size (district magnitude) and, in particular, one or more apportionment methods.

Whenever the user chooses a divisor method, BAZI outputs the resulting apportionment along with a pertinent divisor. This way the user may easily confirm the results with paper and pencil (or a pocket calculator), rather than being forced to believe what the machine says.

A particular feature of BAZI is that it offers three options for multiple electoral districts. The user may choose between (1) separate evaluations for each district, (2) biproportional apportionments using divisor methods, and (3) a variant of the latter that is specifically tailored to the needs of the new Zurich electoral law of 2003.

For these matrix apportionments BAZI uses an algorithm akin to the one reported by Balinski and coauthors in [1],[2] and [3]. More precisely, BAZI implements a discrete variant of the iterative proportional fitting procedure, also known as alternating scaling. A paper to report on the specific properties of the BAZI algorithm is under preparation.

The BAZI homepage and download site is

www.uni-augsburg.de/bazi.

The site also includes the pseudocode of the program, a detailed description of the district options (1)–(3) mentioned above, and an extensive list on the Proportional Representation literature.

References

[1] Balinski, M.L. / Demange, G.: *Algorithms for proportional matrices in reals and integers.* Mathematical Programming **45** (1989a), 193–210.

[2] Balinski, M.L. / Demange, G.: *An axiomatic approach to proportionality between matrices.* Mathematics of Operations Research **14** (1989b), 700–719.

[3] Balinski, M.L. / Rachev, S.T.: *Rounding proportions: Methods of rounding.* Mathematical Scientist **22** (1997), 1–26.
[4] Balinski, M.L. / Young, H.P.: *Fair Representation — Meeting the Ideal of One Man, One Vote.* Brookings Institution Press, Washington, D.C. Second Edition 2001.
[5] Happacher, M. / Pukelsheim, F.: *Rounding Probabilities: Unbiased Multipliers.* Statistics & Decisions **14** (1996), 373–382.
[6] Happacher, M. / Pukelsheim, F.: *Rounding probabilities: maximum probability and minimum complexity multipliers.* Journal of Statistical Planning and Inference **85** (2000), 145–158.

Seat Biases of Apportionment Methods for Proportional Representation

Mathias Drton
(joint work with K. Schuster, F. Pukelsheim and N. R. Draper)

In proportional representation systems, apportionment methods are used to translate the electoral votes into specific seat allocations. The seat allocations are of course integer numbers, and the votes are almost continuous quantities, by comparison. One of the pertinent problems is to measure the effect of the use of a given apportionment method. Whereas previous studies have made inferences about the proportionality of apportionment methods from empirical data, this paper (Schuster *et al.* [6]) derives the information deductively.

We concentrate on the three most popular apportionment methods (cf. Balinski/Young [1], Kopfermann [4]):

(H) the quota method of greatest remainders (Hamilton, Hare),
(W) the divisor method with standard rounding (Webster, Sainte-Laguë),
(J) and the divisor method with rounding down (Jefferson, Hondt).

Assuming repeated applications of each method, we evaluate the seat biases of the various parties. These seat biases are averages, over all possible electoral outcomes, of the differences between the (integer) seats actually apportioned, and the (fractional) ideal share of seats that would have been awarded, had fractional seats been possible.

More formally, we consider ℓ parties, numbered $1, \ldots, \ell$, with respective vote counts $v_1, \ldots, v_\ell$. In proportional representation, the number of seats allocated to a party ought to be proportional to the relative weight of their vote counts. Hence, if $V = \sum_{k=1}^{\ell} v_k$ is the total number of votes cast, there is no loss of generality to convert the vote counts into vote ratios, or *weights*, $w_k = v_k/V$, $1 \leq k \leq \ell$. Assuming that the weights $w_1, \ldots, w_\ell$ follow a uniform distribution over the set of any ℓ non-negative numbers summing to one, we calculate the average behavior of the seat allocations. This distributional assumption can be traced back to Pólya [5].

The *district magnitude*, that is, the total number of seats to be apportioned is denoted by M. The numbers $w_1M, \ldots, w_\ell M$ are the *ideal shares of seats of parties* $1, \ldots, \ell$. These would be the "fractional numbers of seats" to which, ideally, each party would be entitled if that were possible. In real life, the parties are apportioned an integral number of seats $m_1, \ldots, m_\ell$, using the apportionment method in the applicable electoral law.

A common approach for evaluating the goodness of an apportionment method is to compare, for each party k, their actual seat allocation m_k with their ideal share of seats $w_k M$. This results in the *seat excess* $m_k - w_k M$ of party k. We are interested in whether an apportionment method systematically favors larger over smaller parties. Hence, we condition the averaging process on the event that party 1 is largest, party 2 is second-largest, etc., where "largeness" refers to party weights. Under this condition $w_1 \geq \ldots \geq w_\ell$, we

study the expected value of the seat excess $m_k - w_k M$ as a function of the district magnitude M. The resulting quantity

$$B_k(M) = E\big[m_k - w_k M \big| w_1 \geq \ldots \geq w_\ell\big],$$

is called the *seat bias of the k-th largest party*. The standard statistical term "bias" indicates an expected difference between all possible observable values of a quantity and its ideal value. The main results of our paper are formulas for the seat biases, for each party k, under a given apportionment method.

For the quota method of greatest remainders (Hamilton, Hare), the seat biases $B_k^{\mathrm{H}}(M)$ turn out to be identical and slightly positive, for parties $k = 1, \ldots, \ell - 1$ from the largest down to the second-smallest:

$$B_k^{\mathrm{H}}(M) = \frac{\ell + 1}{24M} + O\left(\frac{1}{M^2}\right), \tag{1}$$

$$B_\ell^{\mathrm{H}}(M) = -(\ell - 1)\frac{\ell + 1}{24M} + O\left(\frac{1}{M^2}\right). \tag{2}$$

The ℓ-th, smallest party carries the deficit that balances the positive accumulation. Even though the special role of the smallest party may appear disconcerting, its seat bias remains so small numerically as to be invisible in practice. Thus the quota method of greatest remainders is practically unbiased.

For the divisor method with standard rounding (Webster, Sainte-Laguë), the seat biases of the largest $\ell - 1$ parties $k = 1, \ldots, \ell - 1$ are given in (3), while the seat bias of the ℓ-th, smallest party is given in (4):

$$B_k^{\mathrm{W}}(M) = \frac{\ell + \frac{2}{\ell}}{24M} + \frac{\ell + 2}{24M}\left\{\left(\sum_{j=k}^{\ell-1}\frac{1}{j}\right) - 1\right\} + O\left(\frac{1}{M^2}\right), \tag{3}$$

$$B_\ell^{\mathrm{W}}(M) = -(\ell - 1)\frac{\ell + \frac{2}{\ell}}{24M} + O\left(\frac{1}{M^2}\right). \tag{4}$$

Here a certain amount of balancing goes on between the $\ell - 1$ largest parties alone. The accumulated contribution of the terms $(\ell + \frac{2}{\ell})/(24M)$ is evened out by the negative seat bias of the smallest party. However, all these theoretical seat biases are so small numerically that we do not consider them practically relevant. That is, the Webster seat allocations are practically unbiased.

For the divisor method with rounding down (Jefferson, Hondt) the situation changes dramatically. The leading term in the seat-bias is independent of the district magnitude M:

$$B_k^{\mathrm{J}}(M) = \frac{1}{2}\left\{\left(\sum_{j=k}^{\ell}\frac{1}{j}\right) - 1\right\} + O\left(\frac{1}{M}\right). \tag{5}$$

The remainder term, bounded of order $1/M$, appears to be practically irrelevant.

Now, the largest party clearly enjoys a positive seat bias and can expect

seats in excess of their ideal share. The seat biases become successively smaller, as we pass from the largest party ($k = 1$) to the smallest party ($k = \ell$). The biasedness of Jefferson's method has been observed over many years on the basis of empirical data, but our formulas permit specific calculations about the numerical sizes of the seat biases. For example, the largest party in a three-party system can expect five extra seats per twelve elections in excess to their ideal share, under the Jefferson method.

Our seat bias results depend on the assumption of uniformly distributed weights. However, Schuster *et al.* [6] confirm the theoretical findings via empirical data from the German State of Bavaria, the Swiss Canton Solothurn, and the U.S. House of Representatives. Furthermore, Schuster *et al.* [6] give illustrations of the seat biases and provide details on their interpretation. Mathematical details are provided in Drton and Schwingenschlögl [2, 3].

References

[1] Balinski, M.L. / Young, H.P.: *Fair Representation — Meeting the Ideal of One Man, One Vote.* Second Edition. Washington D.C. 2001. [Pagination identical with First Edition, New Haven CT 1982.]

[2] Drton, M. / Schwingenschlögl, U.: *Surface Volumes of Rounding Polytopes.* Linear Algebra and its Applications **378** (2004), 71–91.

[3] Drton, M. / Schwingenschlögl, U.: *Seat allocation distributions and seat bias formulas of stationary divisor methods for proportional representation.* Under preparation (2004).

[4] Kopfermann, K.: *Mathematische Aspekte der Wahlverfahren — Mandatsverteilung bei Abstimmungen.* Mannheim 1991.

[5] Pólya, G.: *Proportionalwahl und Wahrscheinlichkeitsrechnung.* Zeitschrift für die gesamte Staatswissenschaft **74** (1919), 297–322.

[6] Schuster, K. / Pukelsheim, F. / Drton, M. / Draper, N.R.: *Seat biases of apportionment methods for proportional representation.* Electoral Studies **22** (2003), 651–676.

Negative Weights of Votes and Overhang Seats in the German Federal Electoral Law

Martin Fehndrich

In Elections to the German Bundestag, internal overhang seats cause an effect — negative weight of votes — where a party can get more seats if loosing some votes, or loose seats because it wins some additional votes [1],[2]. This effect is demonstrated in the federal German election 2002, where 1000 votes less for the

SPD in one federal state would have caused an additional seat for this party. In the talk, an overview over the German electoral system is given.

The reasons for Overhang Seats in general are traced back to two mechanisms: many won constituency seats and few party votes. These two mechanisms allow to describe the effect of every parameter of an electoral system on overhang seats. The possible treatments of overhang seats are presented with a view of their effect on disproportionality and additional seats. To prevent disproportionality and an increase of the house size, respectively, a rule must be defined of not awarding some of the overhanging constituency seats. Awarding all won constituency seats, one has to make tradeoffs between disproportionality and increasing house size. The biggest increase of house size with no or only a small disproporionality would be reached by awarding additional balance seats (as done in most German federal states), the biggest disproportionality but no increase of parliament by reducing the number of seats for the not overhanging parties (as in the Scottish parliamentary elections), while just awarding the overhang like in the German Bundestag stays somewhere in the middle. An additional possibility is given in systems with internal overhang seats, like the German system, where a party can have overhang seats in one federal state, but still list seats in other federal states. In this case an internal compensation could be used, where proportional seats are at first awarded to justify the constituency seats and than are awarded to a partys lists.

Negative votes are votes in a party election, without ranking, only one ballot and no second ballot.

One simple example for an electoral system allowing votes with a negative weight of votes is the quota system with largest remainder (named after Hamilton or Hare-Niemeyer), with a 5%-barring clause and 21 Seats. In an 4-party example with A, B 4400 votes, respectively, C 700 votes and D 500 votes, an additional vote for C (coming from nonvoters or D), would actually reduce the number of seats for C. Another example for negative votes is the house monotone quota system, described by Balinski and Young [3, Table A7.1/A7.2 p. 140].

A more serious problem with negative votes occurs in the German Bundestag elections. Here a reduction of the votes for the SPD in the federal state of Brandenburg by 1000 votes in the 2002 election would have caused an additional seat for this party. The effect is connected with the occurrence of internal overhang seats. Loosing votes in Brandenburg will cause a shift in the proportional seats within the party's federals state lists. Brandenburg would lose a seat in favour of Bremen. But since in Brandenburg there are enough constituency seats, this does not hurt Brandenburg's SPD-list, where then an overhang seat occurs, and in the end there is an additional seat for the SPD. The effect is independent from the rounding rule and can occur with Hamilton, Jefferson, Webster or other methods. It occurred in the elections with Jefferson until 1983, before the change to the Hamilton system. Even if we think about fractional seats, a vote for an overhanging federal list would cause the loss of a fractional part of a seat. The effect is sometimes that repeating and predictable that it becomes the best strategy under game theoretical aspects to vote for the disfavoured and overhanging party rather than voting for the favoured party. Even in other cases it is a better strategy to vote for a second choice party and not for the probably overhanging favoured list. The occurrence of this effect in an electoral system is critiqued, because it is against the rule of a direct election and some seats are justified by not given votes rather than given votes. There is a qualitative change exceeding the point of disproportionalty, if a votes weight is not just lower than others, but becoming smaller than zero. An election under this circumstances seems more a case for game theorists than an election. There is no reason in sight which could justify this effect as a trade-off against other favourable properties of an electoral system (as opposed to social choice, where a voter can rank or give more than one vote, allowing similar effects like the no show paradox).

As a solution for the German Electoral System an internal compensation rule is recommended, which prevents internal overhang seats and with that negative votes. To reduce some paradoxes one should also change from Hamilton to the Webster (Sainte-Laguë) system in the party distribution and sub-distribution.

References

[1] M. Fehndrich: *Paradoxien des Bundestags-Wahlsystems*, Spektrum der Wissenschaft **2** (1999) 70.

[2] http://www.wahlrecht.de

[3] M. Balinski and P. Young: *Fair Representation*, Washington 2001.

The Role of the Mean and the Median in Social Choice Theory

William Zwicker

A *center* is a function C that assigns, to each finite set S of points of $\mathbb{R}^n$, a central point $C(S)$ of the distribution. The mean is the most familiar center, but there are others. In particular, the *mediancentre* (the point minimizing the sum of the distances to members of S) seems attractive; it is one of several generalizations of the median to the multivariable context.

Extending work of Saari and Merlin, we show that many familiar voting systems — including Borda count, Condorcet's method of pairwise majorities, and the Kemeny Rule — have alternate descriptions as follows:

(1) Plot the vote v of each voter as a point $A(v)$ in n-space (where the choice of plotting function A depends on the particular voting system at hand).
(2) Take the mean location q of all points $A(v)$ (counting multiplicity).
(3) The outcome is the vote v_0 for which $A(v_0)$ is closest (in the l_2-metric) to q.

In particular, the plot function for the Borda count places rankings at vertices of the permutation polytope, or "permutahedron," while the Condorcet procedure and Kemeny rule each use the "pairwise comparison cube" discussed by Saari. The result for the Kemeny rule is particularly surprising, as the original description employs a type of median based on the Hamming distance between rankings, whereas the new characterization uses the mean on standard, Euclidean distance.

Several properties shared by these voting systems can now be traced to their common dependence on the mean.

If we replace the mean with the mediancentre in step (2) of any system, the result is typically a new system. For example, the *Mediancentre Borda* seems interesting; while it fails to have the consistency property, it is less manipulable than the standard Borda count, and has the interesting property that when a majority of the voters rank candidates similarly, their favorite will win. These differences can largely be explained by axiomatic differences between the mean and the mediancentre. In particular, the mean satisfies the property that

$$C(S+T) = C(S+kC(T)),$$

where S and T are multisets of points in $\mathbb{R}^n$ (several points may have the same spatial location), $S+T$ is the union counting multiplicity, T has k points counting multiplicity, and $kC(T)$ is the multiset having k points, each located at $C(T)$. In fact the mean is characterized by this property together with some symmetry and the requirement that $C(S)$ is uniquely defined for all nonempty multisets S of points of $\mathbb{R}^n$.

The corresponding axiom for the mediancentre seems to be

$$C(S+\{p\}) = C(S+\{p'\}),$$

where p is any point not located at $C(S+\{p\})$, and p' is any point on the one-sidedly infinite ray from $C(S+\{p\})$ through p (with $p' = C(S+\{p\})$ allowed). This property, together with some symmetry and the requirement that $C(S)$ be uniquely defined for all multisets S of points of $\mathbb{R}^n$, except for multisets S containing an even number of collinear points, implies a spatial majority rule property: $C(S) = p$ whenever either a strict majority of points are located at p, or exactly half the points are at p and the other half are not all located at some common different location. These same three axioms characterize the median in $\mathbb{R}^1$, but we do not know whether the same is true for $\mathbb{R}^n$.

Formal Analysis of the Results of Elections

Fuad Aleskerov

Four main issues are presented in the paper:

(1) Patterning of electoral outcomes,
(2) Polarization of electoral outcomes,
(3) Disproportionality of a parliament,
(4) Power distribution in Russian parliament during 1994–2003.

In the first issue I deal with the following problem: is it possible to find a similarity of electoral outcomes over several elections, and can we describe the notion of stability of electoral behavior being based on such similarity?

The approach uses the clustering algorithm applied to all data available on election outcomes. An important new feature of the algorithm (which is called a clustering of curves algorithm) is that it uses the relations among outcomes, not the numerical values themselves. The obtained clusters are called patterns, and one can analyze how the districts change their patterns over years. Then one can call the electoral behavior in a district as a stable one if there are no changes of patterns over years.

Using this very approach, Prof. Hannu Nurmi and I have studied the patterns of party competition in British general elections in 1992, 1997 and 2001 over 529 constituencies in England, 70 constituencies in Scotland, and 40 constituencies in Wales. Only 13 patterns of support distribution are obtained for English constituencies, and only 6 of them are sufficient to describe the electoral preferences distribution in more than 90% of the constituencies. Concerning the stability of electoral outcomes, it has been shown that almost 38% of constituencies have not changed their preferences over those three general elections. Almost 48% of constituencies changed their preference after 1992 elections and then kept stable. In other words, almost 86% of constituencies can be called stable or semi-stable in terms of their electoral outcomes. Approximately the same results are observed for Scotland and Wales. Next we have studied the stability of electoral outcomes during last seven municipal elections from 1976 to 2000 in Finland over 452 constituences. Naturally, the deviation from the stability is much higher when such long period is studied. However, 14% of constituencies are absolutely stable since they have not changed their electoral patterns during those 25 years. 51% of constituencies can be called semi-stable since they have experienced not more than one or two changes of patterns over this period, and only 1% of constituencies are completely unstable, i.e., they have experienced seven changes of patterns over these elections. These results are very illustrative for the use of this very powerful method of patterning electoral outcomes.

In the political studies literature one can find very few attempts to study a polarization of society on the basis of electoral outcomes. Such attempt was made by my B.S. student M. Golubenko and myself. We construct a polarization index using an analogy from physics which is called central momentum of forces with respect to the center of gravity. We consider the parties being positioned over the left-right position axes, and in each position the mass (percentage of votes for that party) is concentrated. Then by evaluating the polarization index one can conclude to which extent the electoral preferences are polarized. If there are only two parties with 50% of votes given to each of them, and these parties are located in the extreme opposite positions of the left-right spectrum, then the polarization is maximal and equal to 1. On the other hand, if there are several parties positioned at the same place on the left-right scale, never mind where this place is, the value of polarization index is equal to 0. We have evaluated the distribution of polarization over the regions of Russia using electoral outcomes of 1995, 1999 and 2003 general elections.

There are several well-known indices to evaluate the disproportionality of a parliament, e.g., Maximum Deviation index, Rae index, Gallagher index, Loosemore-Hanby index, etc. However, none of them take into account the turnout of elections and the percentage of votes "against all", which is allowed in Russia. My M.S.

student V. Platonov and I have proposed a disproportionality index which is a modification of Loosemore-Hanby's index and takes into account these additions. We have introduced a new index of disproportionality, that of relative representation. The index shows a percentage of seats in a parliament which a party receives for 1% of votes. The evaluation made for several countries (Russia, Finland, Sweden, Ukraine, Lithuania, Turkey) show that the countries of the former Soviet block are characterized with higher degree of disproportionality.

The last topic in my paper deals with the study of power distribution in the Russian parliament from 1994 to 2003. We studied Banzhaf and Shapley-Shubik indices on a monthly basis using the MPs' voting data. The indices have been evaluated for different scenarios of coalition formation. The model of coalition formation uses the index of groups positions consistency showing to which extent two groups (fractions) of MPs vote similarly. In the first scenario all evident opponents are excluded from coalitions, in the second scenario all evident and potential opponents are excluded, and in the third scenario coalitions only with evident allies are allowed. The first scenario is most close to the real coalition formation in the Russian parliament. The analysis shows, in particular, that due to the absence of intention to coalesce, the Communist Party during almost all period under study has had power near to 0, although there were periods when this party controlled more than 30% of seats. The dependence in the changes of the power indices distribution is compared with respect to political events during this period.

Procedure-Dependence of Electoral Outcomes

Hannu Nurmi

The theoretical literature abounds examples in which the voting outcomes — winners or the ranking of candidates — depends not only on the revealed preferences of the voters but also on the method used in determining the result. From the late 18th century, two main intuitive notions have played a prominent role in the literature, *viz.* one which maintains that in order to qualify as the winner, a candidate has to defeat, in pairwise comparisons, all other candidates, and the other which looks for the winner among those candidates that are placed highest on the voters' preference rankings. It is well-known that these two intuitive notions are not equivalent: the candidate that defeats all others in pairwise contests may not be best in terms of positions in the voters' preference rankings. But how often do these two notions conflict in real world elections?

The British parliamentary elections were studied by Colman and Pountney (1978) from the view point of estimating the probability of the Borda effect. This effect occurs whenever the elected candidate would be defeated by some other candidate in a pairwise comparison by a majority of votes. The British first-past-the-post (FPTP) system makes it possible that such instances occur. The problem is to know how often. Colman and Pountney used the interview data collected by the British polling organization MORI to construct preference profiles for the entire electorate. From these they then computed the likelihood of instances of the Borda effect. This paper replicates Colman and Pountney's study using the data on the 2001 British parliamentary elections. To get a wider perspective on the variability of electoral outcomes, we used Saari's (1995) geometric methodology to determine the range of all positional voting outcomes in the 2001 elections in all British constituencies. It turns out that — under the same assumptions as those made in the Colman and Pountney's study — in 12 constituencies the ranking of candidates could have been completely reversed depending on the voting rule used. Much more numerous were constituencies, 68 in number, where the actual winner would have been ranked last by another positional voting procedure. The first and second ranked candidates would have been different depending on voting rule in 49 constituencies.

The second aim of the study is to determine the pattern of party competition prevailing in British constituencies. In a study conducted together with Aleskerov we found that the optimal number of party support patterns needed to characterize the 500+ English constituencies over three most recent parliamentary elections is just 13. Moreover, about one-third of the constituencies were characterized by the same support patter over the period of three elections. Less than 10% of the constituencies were completely volatile in the sense of moving from one pattern to another in each election. In Scotland, nearly two-thirds of the constituencies experience no change in support pattern in the three elections. Similar study was conducted on Welsh constituencies. It shows that in terms of support stability, Wales is located between England and Scotland.

References

[1] Colman, A. and Pountney, I.: *Borda's voting paradox: Theoretical likelihood and electoral occurrences.* Behavioral Science **23** (1978), 15–20.

[2] Saari, D. G.: *Basic Geometry of Voting.* Springer 1995.

The Mathematical Source of Voting Paradoxes

Donald G. Saari

The social choice literature has many articles describing certain properties of decision rules: often these properties are obtained via the so-called "axiomatic approach." The thrust of

this talk was to 1) show why the way the "axiomatic approach" is used in the social choice literature often has very little, if anything, to do with "axioms" or the "axiomatic approach," 2)

explain a way, motivated by the mathematics of "chaotic dynamics," to identify all possible consistency properties and paradoxes — both positive and negative — of positional voting methods (and all other rules based on these methods), and 3) identify the source of all possible properties of these voting rules. I had intended to also discuss how to find all possible strategic settings, who can be strategic, and the possible strategies, for any specified voting rule, but time ran out.

As for the axiomatic approach, I gave some examples to show how the properties called "axioms" often are merely particular properties that happen to uniquely identify a particular decision rule. "Uniquely identifying" and "characterizing via axioms" are very different. As an illustration, the two properties "Finnish-American heritage" and a particular "DNA structure" uniquely identify me, but they are not "axioms," they do not characterize me, and they do not tell you "what you are getting," which is the usual claim for the axiomatic approach.

The second part described a way to characterize all possible outcomes. This work was motivated by the clever paradoxical example found by, for example, Brams, Fishburn, Nurmi and many others. The point is that a "paradox" identifies an unexpected property of a voting rule. For example, the profile where 6 prefer ACB, 5 prefer BCA, and 4 prefer CBA leads to the plurality ranking of ABC, and the conflicting pairwise rankings of CA, BA, CB. These rankings define the *plurality word* (ABC, BA, CA, CB), and the word identifies the plurality property that the plurality winner can be the Condorcet loser, while the plurality loser can be the Condorcet winner. In other words, each list of rankings — each word — that CAN occur defines a property of the voting rule. On the other hand, it turns out that this same list (ABC, BA, CA, CB) can never occur with the Borda Count; it can never be a Borda word. This means that a Borda property is that the Condorcet winner can never be Borda bottom ranked and the Condorcet loser cannot be Borda top ranked. Namely a listing that cannot occur — that cannot be a word — also defines a property of a voting rule. Consequently, to find all possible ranking properties of all possible positional methods over all possible subsets of candidates, we want to find all possible listings of rankings that could occur over all possible profiles; we want to find all possible words. Doing so directly may be impossible, but by use of notions from chaotic dynamics, this has been done, and the results are discouraging; e.g., for most collections of voting rules (one for each subset of candidates), anything can happen. Namely, any listing is a word. The unique voting rule that minimizes (significantly!) the number and kinds of listings that can be words is the Borda Count. Thus, this rule has the largest number (significantly so) of positive ranking properties.

The third topic showed how to construct all possible examples that can occur with a voting procedure, how to explain all of the "paradoxes", etc. The way this is done is to emphasize the profiles rather than the voting outcomes. This is done by finding configurations of preferences where it is arguable that the outcome is a tie. The conjecture, which turned out to be true, is that all possible differences among voting rules can be explained (and examples constructed) simply by knowing these configurations of preferences where procedures do, or do not, have a complete tie. As an illustration, all possible

properties, differences in outcomes, etc. among three candidate positional voting occur because of the different ways voting rules handle the "reversal configurations" such as (ABC, CBA). Here, only the Borda count gives a tie: all other positional methods either favor $A = C$ over B, or B over $A = C$. Indeed, the above example was created by starting with 1 person preferring ACB and 4

preferring CBA, where the CBA outcome holds for all positional pairwise outcomes. To create the paradox, 5 units of (ACB, BCA) were added: this adding of the reversal components is what caused

the plurality outcome to differ from the pairwise outcomes. Similarly, all possible differences in procedures using pairwise outcomes arise because of "Condorcet profile components" of the (ABC, BCA, CAB) type. Positional rankings are not affected, but these components change the pairwise tallies: for any number of candidates, it causes all problems with tournaments, agendas,

problems with methods using pairwise outcomes such as the Borda Count and the Kemeny method, etc... The two configurations of preferences completely describe all possible differences among three candidate decision rules that use pairwise and/or positional methods; e.g., it explains all possible differences between the Condorcet and Borda winners. Comments were made about results for $n > 3$ candidates.

On the Closeness Aspect of Three Voting Rules: Borda, Copeland and Maximin

Christian Klamler

The purpose of this paper is to provide a comparison of three different voting rules, Borda's rule, Copeland's rule and the maximin rule. Borda (1784) suggested assigning points to the m alternatives in the individual preferences, namely *m-1* points for the top ranked alternative, *m-2* points for the second ranked alternative, down to *0* points for the bottom ranked alternative. Then, for every alternative, one adds up those points over all individuals. The more points an alternative receives the higher ranked it is in the social preference. Copeland (1951) suggested calculating for each alternative the difference between the number of alternatives it beats and the number of alternatives it looses against. Again, the larger the derived number the higher ranked is the alternative in the social preference. Finally the maximin rule is based on the idea that alternatives should be ranked higher in the social preference the more minimal support they enjoy, i.e. the higher the minimal support over every other alternative.

Usual comparisons of such voting rules focus on non-binary aspects (Laffond et al., 1995), e.g. comparing the actual choices of such voting rules for different preference profiles, or calculating the probabilities of voting rules leading to the same choices (e.g. Gehrlein and Fishburn, 1978, and Tataru and Merlin, 1997). Nurmi (1988, p. 207) provides a possible interpretation of such results by stating that "the estimates concerning the probabilities that two procedures result in different choice sets can be viewed as distances between the intuitions." Moreover he adds that "...the fact that the Condorcet extension methods (Copeland's and the max-min method) are pretty close to each other was to be expected." "Closeness" in this sense means the probability of two voting rules choosing the same winner at the same preference profile. In contrast, "closeness" could also be reasonably interpreted with respect to the distance between the outcomes of the different voting rules, i.e. the difference between the rankings derived from two voting rules. To be more precise, assume a set of alternatives X and two social preferences $\succeq, \succeq'$ on X. We will consider two social preferences $\succeq, \succeq'$ as opposed if for all $x, y \in X$, $x \succeq y \Leftrightarrow y \succeq' x$ and for some $x, y \in X$, $x \succ y \Leftrightarrow y \succ' x$. I.e. opposed social preferences are exactly opposite to each other. This paper shows, that in contrast to the conclusions drawn from using a probabilistic approach, "closeness" in the sense of comparing social preferences is neither guaranteed for Copeland's and the maximin method nor for the Borda and the maximin method. It is proved that there exist preference profiles for which the Copeland ranking and the Borda ranking are exactly the opposite of the maximin ranking. That the Copeland ranking and the Borda ranking are opposed has been shown by Saari and Merlin (1996). Similar comparisons exist for Borda's rule and simple majority rule. It is well known that the Condorcet winner (the alternative that beats every other alternative by a simple majority) is never bottom ranked in the Borda ranking and the Condorcet loser (the alternative beaten by every other alternative) is never top ranked in the Borda ranking (Saari, 1995). Hence, even in cases where the winning alternatives are different, we can ensure a minimal degree of consistency between the rules. However, several recent results (e.g. Ratliff, 2001, 2002 and Klamler 2002) show that such a relationship does not exist for many other pairs of voting rules.

References

[1] Borda, J.C.: *Memoire sur les Elections au Scrutin.* In: Histoire de L'Academie Royale des Sciences (1784).

[2] Copeland, A.H.: *A "reasonable" social welfare function.* Notes from a seminar on applications of mathematics to the social sciences, Unviersity of Michigan (1951).

[3] Gehrlein, W.V., Fishburn, P.C.: *Coincidence Probabilities for Simple Majority and Positional Voting Rules.* Social Science Research **7** (1978), 272–283.

[4] Klamler, C.: *The Dodgson ranking and its relation to Kemeny's method and Slater's rule.* Social Choice and Welfare, forthcoming (2002).

[5] Laffond, G. et. al.: *Condorcet choice correspondences: a set-theoretical comparison.* Mathematical Social Sciences **30** (1995), 23–35.

[6] Nurmi, H.: *Discrepancies in the outcomes resulting from different voting schemes.* Theory and Decision **25** (1988), 193–208.

[7] Ratliff, T.C.: *A comparison of Dodgson's method and Kemeny's rule.* Social Choice and Welfare **18** (2001), 79–90.

[8] Ratliff, T.C.: *A comparison of Dodgson's method and the Borda count.* Economic Theory **20** (2002), 357–372.

[9] Saari, D.G.: *Basic Geometry of Voting.* Springer 1995.

[10] Saari, D.G., Merlin, V.R.: *The Copeland method I: relationships and the dictionary.* Economic Theory **8** (1996), 51-76.

[11] Tataru, M., Merlin, V.R.: *On the relationship of the Condorcet winner and positional voting rules.* Mathematical Social Sciences **34** (1997), 81–90.

Selecting Committees Without Complete Preferences

Thomas Ratliff

In many ways, the Condorcet criterion is the most natural way to compare candidates: if one candidate is preferred to every other candidate in head-to-head elections, then it is plausible to argue that this candidate should be the winner. When choosing a committee of size m, we can apply a similar criterion.

Definition 1. Given a profile with n candidates $A_1, A_2, \ldots, A_n$, define the *Condorcet committee* of size m to be the set M of m candidates such that A_i is preferred to A_j in pairwise elections for all $A_i \in M$ and all $A_j \notin M$.

As we know very well, the Condorcet winner may not exist since there may be a cycle among the top-ranked candidates, and a cycle involving all candidates would preclude the existence of a Condorcet committee. Notice that we are merely partitioning the candidates into two disjoint groups: those on the committee and those off. We do not care whether we have cycles within the disjoint groups, but only that those on the committee are preferred to those not on the committee.

When there is no Condorcet winner, Charles Dodgson (aka Lewis Carroll) proposed in 1874 picking the candidate that is "closest" to being a Condorcet winner by choosing the candidate that requires the fewest adjacent switches in the voters' preferences to become the Condorcet winner. Since he is selecting a single winner, Dodgson does not care if there is a cycle among the remaining candidates; requiring a complete transitive ranking forces more structure than Dodgson views as necessary. We can easily adapt Dodgson's method to measure how far a set of m candidates is from being the Condorcet committee.

Definition 2. In an election with n candidates, define the *Dodgson Committee*, denoted $\mathcal{DC}_m$, to be the set of size m that requires the fewest adjacency switches so that A_i is preferred to A_j in pairwise elections for all $A_i \in \mathcal{DC}_m$ and all $A_j \notin \mathcal{DC}_m$.

There are, however, several anomalous results that can arise:

- The Condorcet winner may be excluded from $\mathcal{DC}_m$.
- If $j \neq k$, then $\mathcal{DC}_j$ and $\mathcal{DC}_k$ may be disjoint or may have any number of candidates in common.

These results can be found in "Some startling inconsistencies when electing committees", T. Ratliff, *Social Choice and Welfare* **21** 3 (2003), 433–454.

In addition to these inconsistencies, a fundamental objection to selecting a committee based on the rankings of individual candidates is that this may not actually capture the voters' preferences. Voters are often concerned with the overall composition of the committee and consider how the individual members will interact. For example, a voter may prefer two candidates in their top-ranked committee because they represent contrasting viewpoints, but would not want one candidate on the committee without the other. A strict listing of the individual candidates could not detect such a preference without additional information.

The motivation for considering this issue arose in the spring of 2003 at Wheaton College in Massachusetts during the selection of three faculty to serve on the search committee for the next president of the college. When Wheaton had last conducted a presidential search in 1992, three men were elected as the faculty representatives on the committee, which was very controversial on the campus. Wheaton has a long standing commitment to gender balance and awareness, partially based upon its history as a women's college (Wheaton began admitting men in 1988). The faculty was almost evenly divided between women and men, and the election of three men was acceptable to almost no one, including those who were selected to serve on the committee. The selection was a result of a process that only considered voters' preferences for individual candidates and not their preferences for the overall composition of the representatives.

The goal was to select one faculty representative from each of the three academic divisions of the college. An initial ballot used approval voting to reduce the field of possible candidates to six, two from each of the divisions, and the final ballot allowed the faculty to select their preferred candidate in each division. This approach seems very reasonable on the surface. However, by decomposing the voters' preferences of the overall composition of the committee into choices on individual candidates, the procedure selected candidates that were individually preferred by a majority, but the overall composition was nearly unanimously unacceptable.

We should not divorce the voters' opinions of the overall group into opinions of individual candidates. This can be viewed as analogous to some of the objections that are raised to the binary independence axiom in Arrow's Theorem: If complete transitive rankings of candidates are broken down into comparisons on pairs and then reassembled to gain an overall ranking, then vital information is lost.

Because of the experience with the selection process in 1992, the faculty at Wheaton were open to adopting another voting method in 2003. The faculty committee responsible for all faculty elections (of which the author is a member) proposed a different method for the final ballot. An approval voting nominating ballot was used as in 1992 to reduce the field to two faculty members from each of the three divisions. Since the requirement was that there be one faculty member from each division selected, this left a total of eight possible groups of faculty representatives. The final ballot asked the voters to rank the eight possible groups, and the Borda Count was used to select the winning group.

There are several interesting observations in this election.

- Of the 71 ballots received, only three were disqualified because the voter failed to rank all eight groups.
- The group selected by the Borda Count was also the Condorcet winner.
- The voters' rankings indicate that their preferences are more complex than could be detected by a simple listing of the candidates or by simple yes/no

votes on the individual candidates. For approximately half of the voters (35 out of 68), their first place and last place committees were not disjoint. For seven of these voters, their first and last place committees differed by a single candidate.
- There are very few rankings that appear more than once; there are 64 distinct rankings from the 68 voters. Even if we restrict to the top three groupings in each ranking, there are still 45 distinct rankings, and the largest duplicate ranking had only five voters.

Overall, the Wheaton faculty were very pleased with the process and the outcome. However, several faculty commented that they would have had a difficult time ranking more than eight options. In general, it will often be impractical to expect the voters to rank all possible committees since the number of possible committees can be extremely large even for a small number of candidates. For example, there are 210 possibilities when selecting a committee of size four from a group of ten candidates.

We define an intermediate approach for selecting a committee that is based upon each voter ranking their top k committees, for some fixed value of k. From this partial ranking, we want to detect overlap within the ranked committees and to extract groups of candidates that the voters believe would work well together.

Definition 3. Assume that there are n possible candidates for a committee of size m and that each of the N voters ranks their top k committees.

Build a weighted graph G with n vertices corresponding to the n candidates. We form a complete graph with edges connecting every pair of vertices, and also include n loops, one for each vertex. Initially assign a weight of zero to every edge in G, and then determine the weights of the edges by examining the rankings of each of the N voters as follows:

- For a voter's top ranked committee, add k to each edge connecting candidates listed in the committee, including the loop that connects each candidate to itself.
- Apply the same technique to the second ranked committee, except in this case we add $k-1$ to each edge.
- In general, for the jth ranked committee, add $k-j+1$ to the edges corresponding to this committee.

The (not necessarily unique) winning committee $\mathcal{C}_m$ is the subgraph of G with m vertices of maximal weight.

Note that the reason for including the loops is to recognize overlaps in voters' preferences for single candidates as well the overlap in groups of candidates. Also notice that we can easily represent G as a symmetric $n \times n$ matrix M where the (i, j) entry corresponds to the weight of the edge connecting candidates i and j.

An objection to this approach is that it only detects an overlap in voters' preferences of single candidates or of pairs of candidates but places no additional weight on triples or quartets of candidates. A natural extension would be to form a hypergraph that includes all subsets up to size m and weight the hypergraph similarly.

A Question for Mathematicians: Would Disputed Elections Be (Sufficiently) Less Probable If U.S. Presidents Were Directly Elected?

Jack H. Nagel

In the 2000 U.S. presidential election, the very close vote in the pivotal state of Florida led to an agonizing recount that was ended by a highly controversial decision of the U.S. Supreme Court. The debacle provoked renewed calls for abolition of the Electoral College (E.C.). However, defenders of that institution countered that the E.C. system, by confining disputed results to just one state (or a small set of states), renders the problems of disputed outcomes much less severe than it would be if a national direct vote resulted in an extremely close vote, thus touching off a nightmarish Florida-style recount nationwide.

While conceding that a national recount would be worse than one confined to a single state, I conjecture that the E.C. structure makes the *probability* of such a dispute substantially higher than it would be with a national direct vote. As I am a political scientist and not at all a mathematician, I pose the following problem to my mathematical colleagues: Is it the case that

$$\mathrm{Prob}[(V_{1,N} - V_{2,N}) < T_N] \quad \ll \quad \begin{array}{l} \mathrm{Prob}[\{s_j\} : \{s_j\} \text{ is critical to a winning} \\ \text{candidate's Electoral College victory and} \\ (V_{1,j} - V_{2,j}) < T_j \text{ for all } s_j], \end{array}$$

where $V_{1,N}$ and $V_{2,N}$ are the popular vote totals nationwide of the leading candidate and the runner-up; $\{s_j\}$ is a set of states with one or more members; $V_{1,j}$ and $V_{2,j}$ are the candidates popular vote totals in state j; and T_N and T_j are vote margins that would "trigger" (either in a mandatory or in a permissive sense) recounts nationally and in state j, respectively.

Currently, 28 states specify recount triggers — 15 for mandatory (automatic) recounts and 13 for margins below which candidates are permitted to request recounts. The U.S. has never held a nationwide direct election, so T_N must be set somewhat arbitrarily. The analysis might be carried out with several possible values (e.g., 10K, 50K, 100K, and 250K). A key premise, however, is that although recount triggers as absolute numbers may grow with the size of the electorate, as a percentage of the total vote, they decrease. That principal already exists in the laws of 10 states, which specify that recounts in statewide elections require a margin that is a lower percentage of the vote than recounts in smaller election districts. In addition, when comparing across states, numerical triggers are only modestly associated with state populations, and they rise at a decreasing rate. There are, however, two exceptions, both extreme outliers with remarkably liberal triggers for recount requests (Texas and Illinois). Besides examining legal requirements, I am also attempting to find out the vote margins in elections where recounts were actually held, especially in the larger states and particularly in Texas and Illinois.

In addition to seeking the help of a mathematician or statistician for a comparison of the two probabilities based on a priori assumptions, I also plan a historical analysis based on actual nationwide and state-by-state popular vote margins.

Postscript: I am very pleased to report that the Workshop will result in just the sort of collaborative effort for which I hoped. Another participant, Professor Vincent Merlin of the University of Caen, has already conducted closely related analyses concerning the likelihood of "wrong-winner" elections (also known as the majority paradox or the referendum paradox) under district-based election structures like the Electoral College. He has taken an interest in the problem I pose, and we plan to work together on it in June 2004, during a visit he has scheduled to the U.S.

Probability Models for the Analysis of Voting Rules in a Federal Union
Vincent Merlin

In an election between two parties (A and B, Left and Right, Yes and No) it might be that a party wins in a majority of districts (or states, constituencies, etc...) while it gets less votes than its opponent in the whole country. In Social Choice Theory, this situation is known as the compound majority paradox, or the referendum paradox. Although occurrences of such paradoxical results have been observed worldwide in political elections (e.g. United States, United Kingdom, France), no study evaluates theoretically the likelihood of such situations. We propose three probability models in order to tackle this issue. The first two models have been used for a long time in social choice theory to compute the theoretical likelihood of discrepancies among voting rules in the three candidate case. The Impartial Culture (IC) assumption states that each voter picks randomly and independently his party affiliation with probability one half. The Impartial Anonymous Culture (IAC) assume that in each district every result is equally likely: Party A has the same probability to get 45%, 65% or 100% of the vote in a given constituency. However, if the number of districts is large enough, the distribution of the votes in favor of A in the country will follow a normal distribution centered around the point 50%. Thus, both IC and IAC models assume that the competition between A and B is close in the whole country. The third model introduce a bias (or shift) in favor of a party. The Biased and Rescaled Impartial Anonymous Culture (BRIAC) assumes that the percentage of votes for A in a given district is drawn uniformly on the interval $[\frac{1}{2} - D + E, \frac{1}{2} + D + E]$. The bias in favor of A is measured by E and the dispersion by D. In fact, the only key parameter of this model is $p = E/D$. The value $E = 0$ gives back the IAC model, up to scaling factor.

For the case where each district has the same (large) population and the IC model, our results prove that the likelihood of this paradox is 16.2% in the three-district case and computer simulations show that it rapidly tends to 20,5% when the number of districts increases. The same pattern is observed under the IAC assumption: The probability of the paradox is 12.5 % with three districts and we estimates that it tends to 16.5% as the number of districts increases. This probability decreases with the number of states when a candidate receives significatively more vote than his opponent over the whole country (parameter p of the BRIAC model). However, p needs to be larger than 0.1 (e.g. 1% bias for $\pm$ 10% dispersion) to get a significative result.

In the case of unequal population state, a new question arises : what is the apportionment method which minimizes the probability of the paradox under a given probability model? Let $m = (m_1, \ldots, m_i, \ldots m_N)$ be the vector of the distribution of the population on the N districts, and a_i be the number of mandates for district i. More precisely, we assume that $a_i = m_i^\beta$ for the IC case and $a_i = m_i^\alpha$ for the IAC case. We then run several computer simulations for different values of the vector m in order to find the optimal values of α and β. In each case, we find out that the minimal value for the paradox has been obtained around $\alpha = 1$ for the IAC model, and around $\beta = 0.5$ for the IC model. This last result with the IC case is consistent with some previous results by Felsenthal and Machover on the value of β that minimizes the mean majority deficit. Computers simulations also tend to show that the probability of the paradox slightly increases as the inequalities among the states in term of population increase.

References

[1] S. Barbera and M. Jackson: *On the Weights of Nations: Assigning Voting Power to Heterogeneous Voters*. mimeo, August 2003.

[2] D. Felsenthal and M. Machover: *Minimizing the mean majority deficit: The second square root rule*. Mathematical Social Sciences **37** (1999), 25–37.

[3] V. Merlin, M. Tataru and F. Valognes: *On the Probability that all Decision Rules Select the Same Winner*. Journal of Mathematical Economics **33** (2000), 183–208.

[4] M. Feix, D. Lepelley, V. Merlin and J.L. Rouet: *The Probability of Conflicts in a U.S. Presidential Type Election*. Economic Theory (2004).

[5] M. Feix, D. Lepelley, V. Merlin and J.L. Rouet: *Fair and Efficient Representation of the Citizens in a Federal Union*. mimeo, 2003.

Foundations of Behavioral Social Choice Research

Michel Regenwetter

This presentation consists of two parts. In the first part, I provide an overview of a forthcoming book with the same title, co-authored with Bernard Grofman (University of California at Irvine), A.A.J. Marley (University of Victoria) and Ilia Tsetlin (INSEAD). This book is a synergetic summary of several underlying journal articles [1, 2, 3, 4, 5, 6, 8, 9, 7]. We provide a mathematical modeling and statistical inference framework

that is tailored towards developing descriptive (as opposed to normative) theories of social choice behavior and towards testing them against empirical data. We believe that our work provides a first systematic attempt towards a formal behavioral theory of social choice behavior, in the spirit of behavioral economics and of behavioral decision theory (a la Kahneman and Tversky). Our empirical work on majority rule decision making demonstrates that some influential strands of theoretical research (the impartial culture assumption, domain restriction conditions such as Sen's value restriction and Black's single peakedness) are descriptively invalid. I now highlight our six most important contributions.

(1) We argue for the limited theoretical relevance and demonstrate the lack of empirical evidence for cycles in mass electorates by replacing "value restriction" and similar classic domain restriction conditions, as well as the "impartial culture" assumption, with more realistic assumptions about preference distributions. We show that our behaviorally plausible conditions, which we validate on empirical data, predict that majority rule decision making is extremely unlikely to generate cycles (among sincere preferences) for realistic distributions in mass electorates. A major implication is that majority rule provides a 'solution' (in practice) to Arrow's impossibility theorem.
(2) In order to better integrate social choice research with the other decision sciences, we expand the classical domains of permissible preference states by allowing for more general binary preference relations than linear or weak orders and by considering probabilistic representations of preference and utility, including a broad range of random utility models.
(3) We develop methodologies to (re)construct preference distributions from incomplete data, i.e., data which do not provide either complete rankings or complete sets of pairwise comparisons.
(4) We highlight the dependence of social choice results on assumed models of preference or utility.
(5) We develop a statistical sampling and Bayesian inference framework that usually places tight upper and lower bounds on the probability of any majority preference relation (cycle or not). We also discuss how such statistical considerations of social choice processes dramatically alter the focus of what are important research questions: For instance, finding the correct winner is often more important than worrying about cycles. Statistical and empirical considerations can also reverse some famous policy implications: For instance, high turnout, not low turnout, as often argued, is desirable when using majority rule.
(6) We demonstrate that in situations where sampling may be involved, misrepresentation (i.e., erroneous evaluations) of the majority preferences is a far greater (and much more probable) threat to democratic decision making than majority cycles.

In the second part of the presentation, I give an overview of results from a systematic analysis of American Psychological Association election data under the single transferable vote (STV), for four elections, each with five candidates and nearly 20,000 voters. This is collaborative work with graduate students Arthur

Kantor (University of Illinois at Urbana-Champaign) and Aeri Kim (University of Illinois at Urbana-Champaign). A full report on this work will be submitted for publication in a major research journal. STV is a particularly interesting paradigm for behavioral social choice research because the ballots provide partial or full preference rankings from the voters. To summarize our main findings: We use several methods to infer majority, Borda, plurality, STV, and other social welfare orders from the ballot data. We can report with high statistical confidence that there were no majority cycles, that the social welfare orders under majority rule and Borda were essentially identical, and that STV generates outcomes that are consistent with both of these classical criteria. Our findings are robust across multiple methods of data analysis. We also discuss the fact that some real world STV elections are tallied in a probabilistic fashion and we compare the probabilistic tally to the deterministic 'genuine STV' tally as defined by the *British Electoral Reform Society.*

References

[1] Doignon, J.-P. and Regenwetter, M.: *An approval-voting polytope for linear orders.* Journal of Mathematical Psychology **41** (1997), 171–188.

[2] Regenwetter, M., Adams, J., and Grofman, B.: *On the (sample) Condorcet efficiency of majority rule: An alternative view of majority cycles and social homogeneity.* Theory and Decision **53** (2002), 153–186.

[3] Regenwetter, M. and Grofman, B.: *Approval voting, Borda winners and Condorcet winners: Evidence from seven elections.* Management Science **44** (1998), 520–533.

[4] Regenwetter, M. and Grofman, B.: *Choosing subsets: A size-independent probabilistic model and the quest for a social welfare ordering.* Social Choice and Welfare **15** (1998), 423–443. bibitemRegenwetter:Grofman:Marley:Tsetlin:2003 Regenwetter, M., Grofman, B., Marley, A., and Tsetlin, I.: *Foundations of behavioral social choice research.* Cambridge University Press (Political Science Series), forthcoming.

[5] Regenwetter, M., Grofman, B., and Marley, A. A. J.: *On the model dependence of majority preferences reconstructed from ballot or survey data.* Mathematical Social Sciences: special issue on random utility theory and probabilistic measurement theory **43** (2002), 453–468.

[6] Regenwetter, M., Marley, A. A. J., and Grofman, B.: *A general concept of majority rule.* Mathematical Social Sciences: special issue on random utility theory and probabilistic measurement theory **43** (2002), 407–430.

[7] Regenwetter, M., Marley, A. A. J., and Grofman, B.: *General concepts of value restriction and preference majority.* Social Choice and Welfare **21** (2003), 149–173.

[8] Tsetlin, I. and Regenwetter, M.: *On the probability of correct or incorrect majority preference relations.* Social Choice and Welfare **20** (2003), 283–306.

[9] Tsetlin, I., Regenwetter, M., and Grofman, B.: *The impartial culture maximizes the probability of majority cycles.* Social Choice and Welfare **21** (2003), 387–398.

Matchings and Allocations
Michel Balinski
(joint work with Mourad Baïou)

There are two distinct finite sets of agents, the *row-agents* I ("employees") and the *column-agents* J ("employers"). Each agent has a strict preference order over the agents of the opposite set. The preferences are collectively called Γ. Each employee $i \in I$ has $s(i)$ units of work to offer, each employer $j \in J$ seeks to obtain $d(j)$ units of work, and $\pi(i,j) \geq 0$ is the maximum number of units that $i \in I$ may contract with $j \in J$. Accordingly, a *stable allocation problem* [3] is specified by a quadruple (Γ, s, d, π) where Γ is a set of preferences, $s > 0$ a vector of $|I|$ reals, $d > 0$ a vector of $|J|$ reals, and $\pi \geq 0$ an $|I|$ by $|J|$ matrix of reals.

Notation. $i' >_j i$ means that agent $j \in J$ prefers i' to i in I, and similarly, $j' >_i j$ means that agent $i \in I$ prefers j' to j in J. If either $i \in I$ or $j \in J$ refuses to work with the other, then $\pi(i,j) = 0$. The set $(i, j^{>}) \stackrel{\text{def}}{=} \{(i,l) : l >_i j\}$ identifies all agents $l \in J$ that are strictly preferred by row-agent i to column-agent j; and $(i, j^{\geq}) \stackrel{\text{def}}{=} \{(i,l) : l \geq_i j\}$ all that are strictly preferred as well as j itself. The sets $(i^{>}, j)$ and $(i^{\geq}, j)$ are defined similarly. In general, if S is a set, $(r, S) \stackrel{\text{def}}{=} \{(r,s) : s \in S\}$, and similarly for (S, r); moreover, if $y(s), s \in S$, is a real number, then $y(S) \stackrel{\text{def}}{=} \sum_{s \in S} y(s)$.

An *allocation* $x = \big(x(i,j)\big)$ of a problem (Γ, s, d, π) is a set of real-valued numbers satisfying

$$x(i,J) \leq s(i), \text{ all } i \in I,$$
$$x(I,j) \leq d(j), \text{ all } j \in J,$$
$$0 \leq x(i,j) \leq \pi(i,j), \text{ all } (i,j) \in \Gamma,$$

called, respectively, the *row*, the *column* and the *entry* constraints. It may be assumed that $\pi(i,j) \leq \min\big\{s(i), d(j)\big\}$. An allocation x is *stable* if for every $(i,j) \in \Gamma$,

$$x(i,j) < \pi(i,j) \quad \text{implies} \quad x(i, j^{\geq}) = s(i) \quad \text{or} \quad x(i^{\geq}, j) = d(j).$$

The *recruitment* or *university admissions problem* is an allocation problem where the $\pi(i,j) = 0$ or 1, the $s(i)$ are positive integers, and the $d(j) = 1$; and the *stable marriage problem* is a recruitment problem where in addition the $s(i) = 1$. There is an extensive literature on these problems (see in particular the book [4]).

In general the set of stable allocations form a nonempty distributive lattice, and the cardinality of the set may be exponential. However, generically, if the reals s, d, and π are chosen at random, then there exists exactly one stable allocation.

In the presence of many stable solutions it is of interest to determine a specific rule for choosing one. A rule is *I-monotonic* if when some agent $i \in I$ goes up in the rankings of one or several of the agents J then i may only receive a better allocation. A rule is *I-strategy-proof* if no subset of agents of I can alter their preferences (that is, falsify them) and thereby obtain better allocations for themselves. Exactly one and the same rule is characterized by either of these two properties [2]. These characterizations have practical applications to recruitment and admissions problems (for an expository account see [1]).

References

[1] Baïou, Mourad and Michel Balinski: *Admissions and recruitment.* American Mathematical Monthly **110** (2003), 386–399.
[2] Baïou, Mourad and Michel Balinski: *Stable allocation mechanisms.* Cahier du Laboratoire d'Econométrie de l'Ecole Polytechnique, 2002-9.
[3] Baïou, Mourad and Michel Balinski: *The stable allocation (or ordinal transportation) problem.* Mathematics of Operations Research **27** (2002), 485–503. Corrected version: **27** (2002), 662–680.
[4] Roth, A. E., and M. Sotomayor: *Two-Sided Matching: A Study in Game Theoretic Modeling and Analysis.* Cambridge University Press, London/New York (1990).

Analysis of QM Rules in the Draft Constitution for Europe Proposed by the European Convention, 2003

Moshé Machover

We analyse and evaluate the qualified majority (QM) decision rules for the Council of Ministers of the EU that are included in the Draft Constitution for Europe proposed by the European Convention. We use a method similar to the one we used in our paper on the Nice Treaty (Felsenthal and Machover 2001). However, we put a special stress on the power of a voter — in this case a minister representing a Member State on the CM — to block a proposed bill (Colemans "power to prevent action").

We make a detailed comparison between the decision rule proposed by the Draft Constitution and that included in the Nice Treaty. We show that the former is much less equitable than the latter. On the other hand, the former achieves a radical — perhaps too radical — increase in effectiveness by means of a great (but uneven) reduction in blocking powers.

The criteria we use in our evaluation are grouped under two main headings: *democratic legitimacy* and *effectiveness*.

Democratic legitimacy. Here we view the CM as the upper tier of a composite two-tier decision-making system. Assuming that each minister votes at the CM according to the majority opinion in his/her country, the indirect voters of this composite system are the citizens of the EU, acting via their respective ministers. We use two main criteria for assessing democratic legitimacy.

First, *equitability.* (Slogan: *One Person, One Vote!*) According to Penroses Square-Root Rule (PSQRR), all EU citizens have equal (indirect) voting power iff the voting powers of the Member States at the CM are proportional to the square root of the size their respective electorates.[3]

[3]For a proof, see Felsenthal and Machover (1998, pp. 667). Throughout, by *voting power* we mean voting power as quantified by the Penrose measure, aka "the absolute Banzhaf index". We take population size as proxy for the size of the electorate.

We measure global deviation of a given QM rule from PSQRR using the index of distortion D — variously attributed to Loosemore and Hanby (1971) or to Duncan and Duncan (1955) — between the distribution of the relative voting powers (measured by the normalized Banzhaf index) under the given rule, and the distribution prescribed by PSQRR. Individual deviations are measured by $d := q - 1$, where q is the ratio obtained by dividing the relative voting power of a given Member State under a given rule by the relative power it ought to have under PSQRR. In the table below, D as well as $\max |d|$ (the maximal value of d) and $ran(d)$ (the range of d, i.e., the difference between the greatest and smallest values of d) are given in percentage terms.

Second, *majoritarianism.* (Slogan: *Majority Rule!*) In any non-trivial two-tier system such as the one under consideration, it is possible that the decision at the upper tier may go against the majority of the indirect voters at large. When this happens, the margin by which the majority of citizens opposing the decision exceeds the minority supporting it is the *majority deficit* of the decision. (If the decision is not opposed by a majority of the citizens, the majority deficit is 0.) Assuming random voting (independent flipping of true coins), the majority deficit is a non-negative random variable. Its expected (mean) value Δ — the *mean majority deficit* — is a measure of the deviation of the given QM rule from majoritarianism.

A third putative criterion of democratic legitimacy — maximization of the sum of the citizens voting powers (slogan: *Power to the People!*) — turns out to be redundant. This is because this sum, denoted by Σ, satisfies the identity $\Sigma = \Sigma_{\max} 2\Delta$, where $\Sigma_{\max}$ is the maximal value of Σ, obtained under majority rule EU-wide direct referendum.[4] In fact, our calculation of Δ uses this very identity.

Effectiveness. Here we view the CM as a decision-making body in its own right, ignoring its role in the two-tier system. The main measure of effectiveness (or compliance) of a decision rule is Colemans index A ("ability of the collectivity to act").[5] A is the a priori probability that a bill will be approved (rather than blocked) by the CM. It is given by $A := \omega/2^n$, where ω is the number of divisions whose outcome is positive (= the number of so-called winning coalitions) and n is the number of voters — in our case, Member States. Equivalently, but more suggestively, we measure resistance to approving a bill in terms of a priori betting odds against a bill being approved.

Results. Some of our results are summarized in the following table.

As we can see, C_{27} is much less equitable than N_{27} and even than the present rule. The present rule has two egregious individual deviations: Germany with 20.1% less than its equitable share of voting power, and Luxembourg with 124.1% more than it "deserves; but the overall deviation from equitability, as measured by D, is much worse in the case of C_{27}. The latter over-endows the four largest and six smallest Member States and under-endows all the rest. The two most egregious cases are Malta (118.2% too much) and Greece (20.8% too little). On the other

[4]For a proof, see Felsenthal and Machover (1998, pp. 6061).

[5]Coleman (1971).

Rule	D	max $\lvert d \rvert$	$ran(d)$	Δ	A	Odds
Present	5.1903	124.1	144.2	5519	0.078	12:1
N_{27}	4.8227	77.6	99.7	7937	0.020	49:1
C_{27}	8.7090	118.2	139.0	3761	0.219	7:2
Rule B	0.2490	1.2	2.1	3882	0.198	4:1

Table 1. In this table, "Present" denotes the current QM rule, for the present 15-member CM. N_{27} is the QM rule prescribed in the Nice Treaty (signed 26 February 2001) for a 27-member CM (the existing 15 members plus the ten scheduled to join the EU in May 2004, plus Romania and Bulgaria). C_{27} is the QM rule (for the same 27 members) included in the Draft Constitution for Europe proposed by the European Convention. Rule B is a benchmark rule which we regard as optimal: it is a weighted rule in which weights are proportional to the square root of **population** sizes and the quota is 60% of the total weight. For the meaning of the column headings, see text above. The odds given in the last column are approximate.

hand, the Nice rule has a dangerously high resistance to passing a bill: the a priori odds against it are 49:1 as compared with the present 12:1 (and 9:1 in the previous period, before the 1995 enlargement). This threatens to paralyse the CM. The proposed rule C_{27} goes in the opposite direction and reduces the resistance dramatically. In our view, it goes a bit too far, as there are good arguments for privileging the status quo to some extent against proposed changes. The detailed figures (not reproduced here) show that C_{27} achieves this greater compliance at the cost of considerable reduction in the blocking powers of the Member States as compared with the Nice rule. Moreover this reduction is very uneven: the four largest and six smallest Member States stand to lose relatively little blocking power, while the others lose a substantial amount. The most egregious cases are Germany (14.5% loss) and Portugal and Belgium (62.4% loss).

References

[1] Coleman J. S.: *Control of collectivities and the power of a collectivity to act.* In: B. Lieberman (ed) Social Choice; New York: Gordon and Breach 1971 (pp. 269–300).

[2] Duncan, O. D. and B. Duncan: *A methodical analysis of segregation indexes.* American Sociological Review **20** (1955), 210–217.

[3] Felsenthal, D. S. and M. Machover: *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes.* Cheltenham: Edward Elgar 1998.

[4] Felsenthal, D. S. and M. Machover: *The Treaty of Nice and qualified majority voting.* Social Choice and Welfare **18** (2001), 431–464.

[5] Loosemore, J. and V. Hanby: *The theoretical limits of maximal distortion: some analytical expressions for electoral systems.* British Journal of Political Science **1** (1971), 467–477.

The Treaty of Nice and the Council of Ministers: A Mathematical Analysis of the Distribution of Power

Werner Kirsch

The treaty of Nice (EU-summit, December 2000) contains a complicated three-step-procedure for decision making in the Council of Ministers of the EU. For a proposal to pass, the first step requires a majority of countries (8 out of 15 states, resp. 13 out of 25). The second step consists of a weighted voting with qualified majority rule. This means that the states are assigned a certain number of votes (voting weights), which were the result of negotiations on the summit. For example, the four biggest states (Germany, France, United Kingdom and Italy) got 29 votes each, Spain and Poland, the next biggest countries, 27 votes each. To approve a proposal, more than (about) 70% of the votes are required, the exact threshold depending on the current number of members of the Union. While these weights are monotone in the countries population, they are pretty arbitrary otherwise, as Germany has more than 82 million citizens, France slightly less than 60 million and Poland about 38 million. Mainly to appease Germany, the Nice treaty contains a third voting step for the Council in which each state has a number of votes proportional to its population. In this step, a support of 62% of the population (as represented by their governments) is required. With its three steps, a far from transparent way to assign voting weights to the member states and strange looking thresholds, the Nice procedure is certainly one of the most complicated voting systems in history. For example, without a careful analysis, it is not at all clear to which extent the third step ("population voting") affects the power of the members in the Council. It is, however, easy to see that the first step of the voting is completely redundant as a qualified majority in the second step can only be achieved with a majority of countries. We use the *Banzhaf index* to quantify the power distribution in the Council after Nice. If we neglect the "population voting", the four big states obviously have the same power in the council. It turns out that the third step has virtually no effect on the power of the big states relative to each other. For example, the Banzhaf index of Germany for the three-step-voting is only by 10^{-7} bigger than the one of France, although the population of Germany is by more than one third bigger than the one of France. In other words: on average Germany will take advantage of its bigger population in one out of ten million votes in the Council! The voting system for the Council as provided by the Nice treaty seems to be very complicated and based on "smoky backroom negotiations" rather than on rational criteria. Moreover, with a 25 or 27 member Union the threshold is so high, that an effective work of the Council seems impossible.

Assignments of Seats as a Modelling Example in the Classroom of Upper Secondary Schools

Thomas Jahnke

In schools and even — to a certain extent — in universities, mathematics is often taught in a bureaucratic manner: getting and exercising and remembering standard procedures and algorithms. Teaching and learning mathematics should be based on principles like

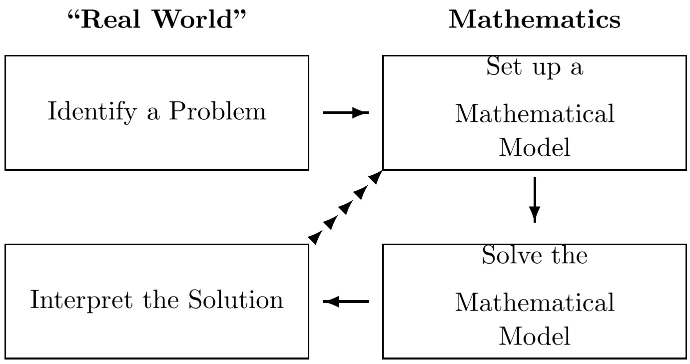
- posing questions instead of answering them,
- active investigations and
- exploring and discovering.

Teaching and learning modelling could enrich the usual syllabus by its contents as well as its methods. While application could be seen as looking from mathematics to the "real world," science and technology modelling emphasises the other direction: looking from the real world towards mathematics in order to solve problems by the use of mathematical knowledge and methods. We cannot define modelling but we can characterise this activity by some essential points [2]:

- Mathematical modelling consists of applying your mathematical skills to obtain useful answers to real problems.
- Learning to apply mathematical skills is very different from learning mathematics itself.
- Models are used in a very wide range of applications, some of which initially do not appear to be mathematical by nature.
- Models often allow quick and cheap evaluations of alternatives, leading to optimal solutions which are not otherwise obvious.
- There are no precise rules in mathematical modelling and no 'correct'answers.
- Modelling can be learned only by doing.

While applications of mathematics often turn out to be a very straightforward approach from a problem to its unique solution, mathematical modelling is done step by step in a specific circle.

The Process or Circle of Modelling



Not all modelling examples are suitable for a classroom. A good modelling example should be

- relevant (nor only for mathematicians),
- realistic,
- motivating (not only for mathematicians),
- rich (in its mathematical aspects),
- allowing different approaches,
- enlightening,
- accessible,
- not too open and too closed and
- mathematically dense.

The problem of the assignment of parliament seats after elections in a representative democracy satisfies these demands. In the German constitution, there are no special rules or procedures for the assignments of seats stated. This is provided in special election laws. The following methods are used:

- *Hare/Niemeyer* (i.e. Hamilton) in elections for the Bundestag and the elections for the Landtag in Bayern, Berlin, Brandenburg, Bremen, Hamburg, Hessen, Mecklenburg-Vorpommern, Nordrhein-Westfalen, Rheinland-Pfalz, Sachsen-Anhalt and Thüringen;
- *D'Hondt* (i.e. Jefferson) in elections for the Landtag in Baden-Wrtemberg, Niedersachsen, Saarland, Sachsen and Schleswig-Holstein;
- *Sainte-Laguë* (i.e. Webster) in elections for the Landtag in Bremen and to choose the heads and the members for the differently sized committees of the Bundestag.

In the classroom, we start by providing the students with the results of a Bundestag election and the texts of the election laws. Later we give them data to discover the Alabama paradox and to construct the majority paradox. The discussion of the concept of the *success value* of a vote leads from the Hare/Niemeyer method to the DHondt method, whose principle is to make the seats as expensive as possible. On the other hand, DHondt violates the upper quota condition and is biased in favour of the bigger parties.

So far, the students have worked on elections laws, procedures, paradox results and principles. The floor is now open to see the assignment problem in a more general way:

> Given votes $v_1 + v_2 + v_3 + \ldots + v_n = v$ and seats $s_1 + s_2 + s_3 + \ldots + s_n = s$, make the set $(v_1; v_2; v_3; \ldots; v_n)$ as similar as possible to the set $(s_1; s_2; s_3; \ldots; s_n)$.

Now, the students are prepared and able to set up their own research program:

(1) Which principles are realising justice the best?
(2) Test the principles and discuss their consequences.
(3) Find for every principle one or several algorithms to realise it.
(4) Construct a procedure fulfilling positive or negative conditions.
(5) Find visualisations for the different procedures.

Beside more formal questions like

> assign the seats to the parties in a way that
> - $\min\{v_i/s_i\}$ is maximal,
> - $|v_i/s_i - v/s|$ is minimal,
> - $|v_i/s_i - v_j/s_j|$ is minimal,

the students discuss the whole modelling process and are asked to write a final report presenting their results. They learn modelling and get deep insights about the problem as well as mathematics itself.

References

[1] T. Jahnke (ed.): *Wahlen.* Mathematik Lehren **88** (1998).
[2] D. Edwards, M. Hamson: *Guide to mathematical modelling.* London 1989.

A Weighted Voronoi Diagram Approach to Political Districting

Bruno Simeone
(joint work with Isabella Lari and Frederica Ricca)

Soon after modern democracies were established, gerrymandering practices, consisting of partisan manipulation of electoral district boundaries, began to occur in several states and countries. In order to oppose such practices, researchers started thinking of automatic procedures for political districting, designed so as to be as neutral as possible. Commonly adopted criteria are:

- *Integrity:* The territory to be subdivided into districts consists of territorial units (wards, townships, counties, etc.) and each unit cannot be split between two or more districts.
- *Contiguity:* The units of each district should be geographically contiguous, that is, one can walk from any point in the district to any other point of it without ever leaving the district.

- *Population equality* (or *population balance*): Under the assumption that the electoral system is majoritarian with single-member districts, all districts should have roughly the same population (*one man – one vote* principle).
- *Compactness:* Each district should be compact, that is, "closely and neatly packed together" (Oxford Dictionary). Thus, a round-shaped district is deemed to be acceptable, while an octopus- or an eel-like one is not.

A broad survey of political districting algorithms is given in (Grilli di Cortona et al., 1999). Later work focuses on local search (e.g., Ricca and Simeone, 2000; Bozkaya, Erkut, and Laporte, 2003). It is also worth mentioning the branch-and-price approach in (Mehrotra, Johnson, and Nemhauser, 1998). Here we propose a novel approach based on weighted Voronoi regions (or diagrams). This notion is not new in the literature, especially in the area of computational geometry (see, e.g., Aurenhammer and Edelsbrunner, 1984). What we believe to be new, besides the specific application to political districting, is our iterative updating of node weights to achieve population balance.

The input to our procedure is the following:

- a *contiguity graph* $G = (V, E)$, whose nodes represent the territorial units and there is an edge between two nodes if the two corresponding units are neighboring;
- a positive integer r, the number of districts;
- a subset $S \subset V$ of r nodes, called *centers* (all remaining nodes will be called *sites*);
- positive integral node weights $p_i, i \in V$, representing territorial unit *populations*;
- positive real distances $d_{i,s}$ for all sites i and all centers s.

We denote by $\bar{P}$ the mean district population (= (total population)/r).

The integrity criterion dictates that a district must be a subset of nodes; according to the contiguity criterion, such a subset must be connected.

A *district map* is a partition of V into r connected subsets (the districts), each containing exactly one center. Given any district map, we denote by D_s the unique district containing center s. We look for a district map such that, informally speaking, the district population imbalance is small and the districts are compact enough.

If one takes as districts the ordinary Voronoi regions w.r.t. the distances $d_{i,s}$, a good compactness is usually achieved, but a poor population balance might ensue. In order to re-balance district populations, one would like to promote site migration out of "heavier" districts (populationwise) and into lighter ones. Then the basic idea is to consider weighted distances $d'_{i,s} = w_s \cdot d_{i,s}$, where each weight w_s is proportional to P_s, the population of district D_s; and to perform a Voronoi iteration w.r.t. the biased distances $d'_{i,s}$. Do this iteratively: at iteration k, $k = 1, 2, \ldots$, two different recursions may be taken into consideration, namely, a *static* one,

$$d^k_{i,s} = \frac{P^{k-1}_s}{\bar{P}} d^0_{i,s} \quad , i \in V \setminus S, s \in S$$

and a *dynamic* one,

$$d_{i,s}^{k} = \frac{P_s^{k-1}}{\bar{P}} d_{i,s}^{k-1} \quad , i \in V \setminus S, s \in S \ ,$$

where, in both cases, $d_{i,s}^0 = d_{i,s}$, P_s^0 is the population of the (ordinary) Voronoi region containing center s, and P_s^k is the population of D_s after iteration $k = 1, 2, \ldots$; stop as soon as the districts become stable.

The above sketched algorithm will be called a *full transfer* one. One may also consider a *single transfer* version of it, by letting sites migrate from one district to another one at a time. Here too, one may adopt either the static or the dynamic recursion defined above. So one gets altogether four variants of the weighted Voronoi algorithm (static/dynamic recursion; full/single transfer).

One possible implementation of the single transfer algorithm is the following: at iteration k, site i is a candidate for migrating from D_q to D_t if:

(1) $P_t^{k-1} = \min\{P_s^{k-1} \, : \, s = 1, \ldots, r\}$
(2) $d_{i,t}^k = \min\{d_{j,t}^k \, : \, j \notin D_t\}$
(3) $d_{i,t}^k < d_{i,q}^k$
(4) $P_t^k < P_t^{k-1}$

The algorithm stops when the set of candidates is empty.

Next, we define four desirable properties to be met by weighted Voronoi algoritms — or at least by some variants of them.

(i) *Order invariance:* $d_{i,s}^k < d_{j,s}^k \iff d_{i,s} < d_{j,s}, \ \ s \in S; i, j \in V \setminus S.$
(ii) *Re-balancing:* At iteration $k = 1, 2, \ldots$, site i migrates from D_q to D_t only if $P_q^{k-1} > P_t^{k-1}$.

Definition 1. Given a graph G and any two nodes i, j of G, a *geodesic* between i and j is any shortest path between i and j in G when all edge-lengths are equal to 1. The *geometric distance* between i and j in G is the number of edges in any geodesic between i and j.

(iii) *Geodesic consistency:* At any iteration, if node j belongs to district D_s and node i lies on any geodesic between j and s, then i also belongs to D_s.

One can show that geodesic consistency implies contiguity, but the converse does not necessarily hold. Moreover, geodesic consistency holds when the input distances are geometric distances on the contiguity graph G. On the other hand, some counterexamples show that for arbitrary input distances contiguity might not hold.

(iv) *Finite termination.*

Finite termination in general is not guaranteed, as shown by counterexamples. However, one can prove that the single transfer dynamic weighted Voronoi algorithm, under conditions (1) – (4) above, enjoys finite termination.

The following table shows the results we have obtained so far.

Property	Static		Dynamic	
	Full transfer	**Single transfer**	**Full transfer**	**Single transfer**
Order invariance	$\surd$	$\surd$	$\surd$	$\surd$
Re-balancing			$\surd$	$\surd$
Geodesic consistency				$\surd^1$
Finite termination				$\surd^2$

References

[1] F. Aurenhammer and H. Edelsbrunner: *An optimal algorithm for constructing the weighted Voronoi diagram in the plane.* Pattern Recognition **17** (1984), 251–257.

[2] B. Bozkaya, E. Erkut, and G. Laporte: *A tabu search heuristic and adaptive memory procedure for political districting.* European Journal of Operational Research **144** (2003), 12–26.

[3] P. Grilli di Cortona, C. Manzi, A. Pennisi, F. Ricca, B. Simeone: *Evaluation and Optimization of Electoral Systems.* SIAM Monographs in Discrete Mathematics, SIAM, Philadelphia 1999.

[4] A. Mehrotra, E.L. Johnson, G.L. Nemhauser: *An optimization based heuristic for political districting.* Management Science **44** (1998), 1100–1114.

[5] F. Ricca, B. Simeone: *Local search heuristics for political districting.* Technical Report, Dipartimento di Statistica, Probabilitá e Statistiche Applicate, Universitá "La Sapienza", Roma, Ser. A, n. 11–2000 (2000), submitted for publication.

Voting in Social Choice Theory

Maurice Salles

My purpose is to outline the difference between aggregation functions directly based on individual preferences and those based on voting games. In doing this, the role of individual indifferences is shown to be crucial. The most studied practical aggregation function is majority rule. An option x is socially preferred to an option y if the number of individuals who prefer x to y is greater than the number of individuals who prefer y to x. In this case, when individual preferences are conveniently restricted, we know since Duncan Black that the social preference is transitive. Taking as a basis a voting game where "powerful" coalitions are a priori defined, Dummett and Farquharson have demonstrated that Black's type of conditions could be extended. When there are no restrictions on individual preferences (supposed to be complete preorders on a finite set of options), Nakamura provided an existence theorem for the core, given a list of individual preferences, based on a comparison between the number of options and a number given by the structure of the voting game. Unfortunately this comparison is very restrictive and accordingly other solution concepts have been studied. The stability set due to Rubinstein is never empty when individual preferences are linear orders. However,

[1]under the assumption that distances are geometric ones in G

[2]under conditions (1)–(4) above

this property does not hold when individual preferences are complete preorders. Le Breton and Salles then obtained results using the same number as the

number used by Nakamura. Some other solution concepts are based on binary relations that are transitive so that obtaining maximal elements is not a problem when the set of options is finite. The interesting mathematical aspects in this case are due to the consideration of sets of options having a geometrical or topological structure (the problems we are then facing are related to the absence of continuity properties of the relevant social preference

relation). When the space of options is a compact part of the Euclidean space, the core is non-empty when the dimension of the Euclidean space is conveniently restricted (this was shown by Greenberg, and extended by Saari).

Dividing the Indivisible: Procedures for Allocating Cabinet Ministries to Political Parties in a Parliamentary System

Steven J. Brams
(joint work with Todd R. Kaplan)

How coalition governments in parliamentary democracies form and allocate cabinet ministries to political parties is the subject of a substantial empirical and theoretical literature. By and large, a rule of proportionality, whereby parties are given more ministries or more prestigious ministries (e.g., finance, foreign affairs, or defense) in proportion to their size, is followed. However, small centrist parties that are pivotal in coalitions (e.g., the Free Democrats in Germany) have successfully bargained for larger-than-proportional allocations.

This task is complicated when less-than-compatible parties, like the Christian Democrats and the Greens, join the same coalition. While fiscal conservatism and protecting the environment are often at odds, these parties may still be accommodated if, for instance, the Christian Democrats are given the finance ministry and the Greens the environmental-protection ministry, and each has major influence over policies in its area.

To facilitate the allocation of cabinet ministries to political parties, we propose procedures that take into account both party interests and party size. This mechanism shifts the burden of making cabinet choices from the prime minister designate, or *formateur*, who is usually the leader of the largest party in a coalition government, to party leaders that join the government. Thereby these procedures give party leaders primary responsibility for the make-up of the coalition government.

We assess the fairness of this procedure, based on different criteria of fairness. Our analysis is inspired by an apportionment method used in Northern Ireland in 1999 to determine the *sequence* in which parties made ministry choices (it has also been used in Danish cities and the German Bundestag). This method works such that the largest party in a coalition gets first choice; presumably, it would choose the position of prime minister. After that, the apportionment method determines the order of choice.

For example, suppose there are three parties, ordered by size $A > B > C$, and there are six ministries to be allocated. If the sequence is $ABACBA$, A will receive three ministries, B two ministries, and C one ministry. But beyond these numbers, the sequence says that A is entitled to a second choice before C gets a first choice, and C gets a first choice before B gets a second choice.

If parties have complete information about each others preferences, we show that it may not be rational for them to choose *sincerely* — that is, to select their most-preferred ministry from those not yet chosen. Rather, a party (e.g., A) may do better postponing a sincere choice and, instead, selecting a less-preferred ministry if (i) that ministry might be the next choice of a party that follows it in the sequence (e.g., B or C) and (ii) As sincere choice is not in danger of being selected by B or C before As turn comes up again. Such *sophisticated choices*, which take into account what other parties desire, can lead to very different allocations from sincere ones.

If there are only two parties, sophisticated choices and sincere choices both yield Pareto-optimal allocations: No parties, by trading ministries, can do better, based on their ordinal rankings of ministries. However, this is not true if there are three or more parties that make sophisticated choices, which was first demonstrated for sequential choices made in professional sports.

What we show here for the first time is the problem of *nonmonotonicity*: A political party may do worse by choosing earlier in a sequence, independent of the Pareto-optimality of the sophisticated choices. Hence, the apparent advantage that a partys size gives it by placing it early in a sequence can, paradoxically, work to its disadvantage — it may actually get more preferred choices by going later.

Like Pareto-nonoptimal allocations, nonmonotonicity *cannot* occur if parties are sincere. Thus, we are led to ask how sincere choices might be recoveredor induced in the first place if the parties know they "cannot get away with" insincere choices. While there is an allocation mechanism that makes sincerity optimal for two parties, there are difficulties in extending it to more than two parties.

By putting the choice of ministries in the hands of party leaders, these leaders are made responsible for their actions. Ultimately, we believe, party leaders will be more satisfied making their own choices rather than having to bargain for them. Moreover, this greater satisfaction should translate into more stable coalition governments, which is a subject that has been extensively studied by a many scholars.

The allocation procedures we analyze could go a long way toward minimizing the horse trading that typically ensues when a *formateur* bargains with party leaders over the ministries they will be offered. By cutting down on the rents extracted in the bargaining process, a coalition government is likely to form more expeditiously and be less costly to maintain.

This is not to say that the procedures we discuss solve all problems. Because ministries are indivisible, there will not generally be a perfect match of the claims of each party and its allocation. Furthermore, there are certain problems that are ineradicable, whatever allocation procedure is used. For example, it may not be possible to eliminate envy among equally entitled parties. Nevertheless, we believe the procedures that we discuss offer a promising start to attenuating conflicts that have plagued the formation of coalition governments and, not infrequently, led to their downfall.

Reporter: Johannes Rückert

Participants

Prof. Dr. Fuad Aleskerov
alesk@ipu.rssi.ru
Institute for Control Sciences
Russian Academy of Sciences
65 Profsoyuznaya
Moscow, GSP-7 117806 – Russia

Prof. Dr. Michel L. Balinski
Balinski@Poly.Polytechnique.Fr
Laboratoire d'Econometrie de
l'Ecole Polytechnique
1, rue Descartes
F-75005 Paris Cedex 05

Prof. Dr. Steven J. Brams
Steven.Brams@NYU.Edu
Department of Politics
New York University
7th Floor
726 Broadway
New York NY 10003 – USA

Dipl.-Math.oec. Mathias Drton
drton@stat.washington.edu
Department of Statistics
University of Washington
Box 35 43 22
Seattle, WA 98195-4322 – USA

Prof. Dr. Paul H. Edelman
paul.edelman@Vanderbilt.Edu
Vanderbilt University Law School
131 21st Avenue South
Nashville TN 37203-1181 – USA

Dr. Martin Fehndrich
fehndrich@wahlrecht.de
Schloßstraße 90
D–49080 Osnabrück

Prof. Dr. Thomas Jahnke
jahnke@rz.uni-potsdam.de
Institut für Mathematik
Universität Potsdam
Postfach 601553
D–14415 Potsdam

Prof. Dr. Marc Kilgour
mkilgour@wlu.ca
Department of Mathematics
Wilfrid Laurier University
Waterloo Ontario N2L 3C5 – Canada

Prof. Dr. Werner Kirsch
werner.kirsch@mathphys.ruhr-uni-bochum.de
werner.kirsch@ruhr-uni-bochum.de
Mathematisches Institut
Ruhr Universität Bochum
Universitätsstraße 150
D–44780 Bochum

Mag. Christian Klamler
christian.klamler@uni-graz.at
Institut für Finanzwissenschaft
und Öffentliche Wirtschaft
Karl-Franzens-Universität
Universitätsstr. 15/E4
A-8010 Graz

Prof. Dr. Moshe Machover
moshe.machover@kcl.ac.uk
5, Milman Road
Queen's Park
GB-London NW6 6EN

Prof. Dr. Vincent R. Merlin
merlin@econ.unicaen.fr
Charge de recherche du CNRS
University Caen
PB 5186
F-14032 Caen Cedex

Prof. Dr. Jack H. Nagel
nageljh@sas.upenn.edu
Department of Political Science
University of Pennsylvania
208 S. 37th Street
Philadelphia PA 19104-6215 – USA

Prof. Dr. Hannu Nurmi
hannu.nurmi@utu.fi
Department of Political Science
University of Turku
20014 Turun Yliopisto – Finland

Prof. Dr. Friedrich Pukelsheim
Pukelsheim@Math.Uni-Augsburg.DE
Lehrstuhl für Stochastik und ihre Anwendungen
Institut für Mathematik der Univ.
Universitätsstr. 14
D–86135 Augsburg

Prof. Dr. Victoriano Ramirez
vramirez@ugr.es
Department of Applied Mathematics
University of Granada
Edificio Politecnico, E.T.S.Caminos
Campus de Fuente Nueva
E-18071 Granada

Prof. Dr. Thomas Ratliff
tratliff@wheatonma.edu
Department of Mathematics
Science Center 101
Wheaton College
E. Main Street
Norton MA 02766 – USA

Prof. Dr. Michel Regenwetter
regenwet@uiuc.edu
Department of Psychology
University of Illinois at Urbana-Ch
435 Psychology Building, MC 716
603 East Daniel Street
Champaign IL 61820 – USA

Johannes Rückert
j.rueckert@iu-bremen.de
School of Engineering and Science
International University Bremen
Campus Ring 12
D–28759 Bremen

Prof. Dr. Donald G. Saari
dsaari@uci.edu
Institute for Mathematical
Behavioral Science
University of California
2119 Social Science Plaza
Irvine CA 92697-5100 – USA

Prof. Dr. Maurice Salles
salles@econ.unicaen.fr
Institut SCW
MRSH
Universite de Caen
F-14032 Caen Cedex

Prof. Dr. Remzi Sanver
Sanver@Bilgi.Edu.Tr
Istanbul Bilgi University
Inonu Cad. No. 28
Kustepe
Istanbul 80310 – Turkey

Prof. Dr. Bruno Simeone
bruno.simeone@uniroma1.it
Dipart. di Statistica, Probabilita
e Statistiche Applicate;
Universita degli Studi
di Roma"La Sapienza"
Piazzale Aldo Moro, 5
I-00185 Roma

Prof. Dr. William S. Zwicker
zwickerw@union.edu
Department of Mathematics
Union College
Schenectady, NY 12308 2311 – USA

Author Index