Dagstuhl-Seminar

Algorithms and Complexity for Continuous Problems

Organized by:

Peter Mathé (WIAS - Berlin)

Klaus Ritter (Universität Passau)

Joseph F. Traub (Columbia University)

May 18–22, 1998

Overview

The Dagstuhl-Seminar on Algorithms and Complexity for Continuous Problems was attended by 40 scientists from 12 countries. We express our gratitude to the staff of Schloß Dagstuhl for providing a great atmosphere.
Continuous algorithms and complexity is a very rich and rapidly changing area. It relates to many other areas of computer science and mathematics, of both theoretical and practical orientation. This seminar has emphasized some of the most dynamic topics like:

- Tractability of high dimensional problems, numerical integration
- Low discrepancy sequences with applications in computational finance
- Operator equations, wavelet algorithms
- Zero finding, computation of fixed points, and global optimization
- Probabilistic analysis of algorithms
- Monte Carlo methods
- Ordinary and stochastic differential equations
- Path integration.

This Seminar-Report contains the abstracts of 37 lectures in alphabetical order.

Abstracts

A Lower Bound on the Average Number of Pivot Steps Required for Solving an LP – Valid for all Simplex-Variants

Karl Heinz Borgwardt
Universität Augsburg
(joint work with P. Huhn)

Consider the arithmetical effort to solve Linear Programming Problems

$$\begin{array}{ll} \text{maximize} & v^T x \\ \text{subject to} & a_1^T x \ \leq \ 1, \ldots, a_m^T x \ \leq \ 1 \\ \text{where} & x, v, a_1, \ldots, a_m \in \mathbb{R}^n \text{ and } m \geq n. \end{array}$$

The lower dimension n is the number of variables, and the larger dimension m gives the number of restrictions defining the feasible polyhedron

$$X = \{x \mid a_1^T x \ \leq \ 1, \ldots, a_m^T x \ \leq \ 1\}.$$

We are interested in the expected numbers of iteration steps of certain solution algorithms (namely Simplex-variants) for the LP, when the problem-instances are distributed according to the so-called Rotation-Symmetry-Model (RSM):

> $a_1, \ldots, a_m, v$ are distributed on $\mathbb{R}^n \setminus \{0\}$ identically, independently and symmetrically under rotations.

The general description of the Simplex-Method, is given as follows:

> Calculate a vertex x_0 of X and construct a sequence of vertices $x_0, \ldots, x_s \in X$, such that for all i the vertices x_i and x_{i+1} are adjacent and $v^T x_i < v^T x_{i+1}$. The final vertex x_s is either the optimal vertex or a vertex where the nonexistence of an optimum becomes obvious.

A complete and detailed rule for making the choice of the successor vertex will determine a variant of the Simplex algorithm. We show that a stochastic evaluation leads to the following *lower* bound on the average number of steps, which is valid for all variants.:
If the calculation of the vertex x_0 has been done without knowledge of v, then we know that

$$E_{m,n}(s) \geq m^{\frac{1}{n-1}} \cdot n^0 \cdot Const_1.$$

This insight results from an intensive evaluation of the stochastic properties of the random polyhedron X, as the distribution of the length of its edges.

For a special variant, the so called Shadow-Vertex-Algorithm, we know a (sharp) *upper* bound on this number in the form

$$E_{m,n}(s) \leq m^{\frac{1}{n-1}} \cdot n^2 \cdot Const_2.$$

The comparison of both results shows that the investigations on the Shadow-Vertex-Algorithm had been quite representative for general variants, and that no variant can do substantially better.

The P Algorithm for Global Optimization

Jim Calvin
Department of Computer Science, New Jersey Institute of Technology

The P algorithm is an adaptive algorithm for approximating the global minimum of a continuous function on a compact set, motivated by viewing the function as a sample path of a Gaussian process. In this talk we analyze the convergence of the P algorithm under the assumption of Wiener measure on the objective functions.

The P algorithm operates by choosing, at time $n+1$, to observe at the location that maximizes the probability that the function value is less than the current minimum observed, minus a positive number c_n. We consider three cases: constant $c_n = c$, $c_n \downarrow 0$ non-adaptively, and $c_n \downarrow 0$ adaptively. While it is possible to make the error converge to 0 faster than any polynomial in $1/n$, an exponential convergence rate is not attained.

Smolyak's Construction of Cubature Formulas of Polynomial Degree

Ronald Cools
Dept. of Computer Science, K.U.Leuven, Belgium
(joint work with E. Novak and K. Ritter)

We consider cubature formulas constructed using Smolyak's technique to approximate multivariate integrals. We look at these formulas using their polynomial degree of exactness as quality measure, both for the periodic case (trigonometric degree) and the non-periodic case (algebraic degree). We show that there is no advantage, from this point of view, of using 1-dimensional quadrature formulas of high degree in the beginning of the sequence of Smolyak rules. We also show that for a fixed algebraic degree the number of points depends on the dimensions in an order-optimal way, i.e., comparable with the theoretical lower bound. For the trigonometric degree, this is totally different however.

Delayed Curse of Dimension for Gaussian Integration

Francisco Curbera
Columbia University

We study weighted integration over $\mathbb{R}^d$ of scalar functions satisfying a Lipschitz condition, $|f(x) - f(y)| \leq \|x - y\|_p$, $p \in [1, \infty)$, for Gaussian weights with a diagonal covariance matrix. Let $\sigma_1 \geq \ldots \geq \sigma_d > 0$, represent the diagonal elements of this matrix. Define s_t and $\overline{s}_t$ to be the p-norm of the vectors $(\sqrt{\sigma_1}, \ldots, \sqrt{\sigma_t})$, $(\sqrt{\sigma_{t+1}}, \ldots, \sqrt{\sigma_d})$, and c the cost of one function evaluation. We prove that there are constants C_L and C_U independent of d such that, if

$$t_u = \min \{i : \ 2C_U \, \overline{s}_i \ \leq \ \varepsilon\},$$

and

$$t_l = \begin{cases} \max \{i : \ C_L \, \overline{s}_i \geq \varepsilon\}, & \text{if } \exists \text{ such } i, \\ 1 & \text{otherwise,} \end{cases}$$

then, the complexity of the integration problem is bounded as

$$c \left(C_L \frac{s_{t_l}}{\varepsilon} \right)^{t_l} \leq \mathrm{comp}(\varepsilon, \sigma) \leq (c+2) \left(C_U \frac{s_{t_u}}{\varepsilon} \right)^{t_u}.$$

This results shows that the effective dimension of the problem reduces from d down to some number between t_l and t_u, and that this reduction is a function of ε and the $p/2$-norm of a vector of diagonal elements of the covariance matrix. As ε approaches zero, this effective dimension increases, reaching d in the limit.

Adaptive Wavelet Methods for Elliptic Problems - Convergence Rates

Wolfgang Dahmen
RWTH Aachen
(joint work with A. Cohen and R. DeVore)

Linear operator equations $Au = f$ are considered where A is selfadjoint and boundedly invertible as an operator from some Sobolev space into its dual. This covers boundary value problems for elliptic partial differential equations as well as many classical singular integral operators such as the single layer potential, double layer potential and hypersingular operator. The efficient numerical treatment of such problems is obstructed by several factors such as the size of the resulting discrete problems, a possibly growing ill conditioning when the operator has an order different from zero or by the fact that densely populated matrices arise in connection with integral operators. An adaptive wavelet scheme is outlined that aims at determining in the course of the solution process a possibly small set of wavelets needed to recover the solution within some desired error tolerance. It is based on a-posteriori error estimates for the current approximate solution in terms of residuals. The main

result is its asymptotic optimality in the sense that (within a certain range of Besov regularity) the convergence rate of best N-term approximation is achieved at a computational expense which stays proportional to the number N of degrees of freedom provided that full information on the given data is available. As a consequence one observes an asymptotic gain in accuracy and efficiency over a-priorily given uniform refinements if the solution lacks the order of Sobolev regularity that corresponds to the order of the discretization error for uniform refinements. A crucial step in the analysis is to prove the uniform boundedness of the residuals evolving during the adaptive process with respect to discrete norms for certain Lorentz sequence spaces. The main ingredients of the analysis are norm equivalences for Sobolev and Besov spaces induced by wavelet expansions, related preconditioning effects, the near sparseness of the wavelet representations of the operators under consideration and the elements of Besov spaces, new fast approximate matrix-vector multiplication schemes suggested by the analysis and a judicious use of intermediate thresholding of current approximate solutions.

The Fourier Transform of the Characteristic Function and Estimates for the Discrepancy

Michael Drmota
TU-Wien

Recently Woźniakowski showed that Roth's lower bound for the L^2-discrepancy can be used to obtain the order of magnitude of the average case complexity of numerical integration of continuous functions with respect to the Wiener measure.
The aim of this talk is to show that estimates for the Fourier transform of characteristic functions are very useful to obtain estimates of the L^2-discrepancy of Roth-type. Firstly, the L^2-discrepancy of N points in the k-dimensional unit cube with respect to homothetic copies of a (fixed) convex polytope is bounded below by $c(logN)^{(k-1)/2}/N$. Secondly, the L^2-discrepancy of N points in the k-dimensional unit cube with respect to homothetic copies of a (fixed) sufficiently smooth convex body is bounded below by $cN^{-1/2-1/(2k)}$.
Both lower bounds are best possible. It is surprising that these two orders of magnitude are completely different. It is not known which property of the convex body dictates the order of magnitude. Only in the two-dimensional case there is a partial answer due to J. Beck.

Monte Carlo Complexity of Global Solution of Integral Equations

Stefan Heinrich
Universität Kaiserslautern

We study the problem of global solution of Fredholm integral equations. This means that we seek to approximate the full solution function (as opposed to the local

problem, where only the value of the solution in a single point or a functional of the solution is sought). We analyze the Monte Carlo complexity, i. e. the complexity of stochastic solution of this problem. The framework for this analysis is provided by information-based complexity theory. The results show that even in the global case Monte Carlo algorithms can perform better than deterministic ones, although the difference is not as large as in the local case.

On the Number of Tests Needed for Zero Finding

Peter Hertling
Department of Computer Science, University of Auckland

We show that the topological complexity of the following problem is one for any fixed $\varepsilon \in (0, \frac{1}{2})$: given an increasing continuous function f on the unit interval with $f(0) \cdot f(1) < 0$, compute an ε–approximation of the zero of f by using function evaluations, comparisons, and the basic arithmetic operations $+, -, *, /$. This is in contrast to the known facts that the topological complexity of the same problem drops to zero if one may additionally use either the functions log and exp (Novak and Woźniakowski 1996) or the absolute value $|\cdot|$ (the author 1996) as arithmetic operations.

Numerical Analysis in Lie Groups: Complexity Issues

Arieh Iserles
University of Cambridge

Many differential equations evolve on Lie groups or on homogeneous spaces and it is often important to retain this feature under discretization. This is among the main goals of the theory of *geometric integration.* In this talk we have briefly surveyed the subject, commencing from the definition of a Lie group and a Lie algebra and demonstrating that a natural avenue to a discretization that respects Lie groups is converting the underlying equation so that it evolves on a Lie algebra. This leads to a number of effective methods, e.g. Runge–Kutta/Munthe-Kaas schemes and Magnus expansions. On the face of it, such methods are exceedingly expensive, but it is possible to make them drastically more efficient (in fact, often cheaper than classical time-stepping methods) by using the theory of graded Lie algebras, in tandem with the exploitation of special features in multivariate integration of commutators over simplices. A lesson to the complexity community is that the framework changes completely when the *equations* evolve on a submanifold of the Euclidean space (this is different from merely choosing a different class of *functions!*) and the importance of differential topology and the theory of Lie algebras in making sense of this new situation.

Instant Radiosity

Alexander Keller
Universität Kaiserslautern

We present a fundamental procedure for instant rendering from the radiance equation. Operating directly on the textured scene description, the very efficient and simple algorithm produces photorealistic images without any finite element kernel or solution discretization of the under lying integral equation. Rendering rates of a few seconds are obtained by exploiting graphics hardware, the deterministic technique of the quasi-random walk for the solution of the global illumination problem, and the new method of jittered low discrepancy sampling.

Informational Complexity of Neural Networks

Mark A. Kon
Boston University and Warsaw University
(joint work with L. Plaskota)

Learning problems in neural network theory are essentially partial information issues. That is, we wish to reconstruct a desired input-output (i-o) function from information consisting of examples (i.e., individual function evaluations). A complexity theory for neural networks from the standpoint of information has had some beginnings in the work of Girosi and Poggio and others, but this theory has not yet been fully developed.

The theory of neural complexity contrasts itself from that of informational complexity in that here we deal with numbers of neurons in the hidden layer of a neural network which are necessary for the computation of a given i-o function, generally assuming full information about that function. This theory has seen some extensive and successful development in recent years, particularly in the work, e.g., of Mhaskar and Micchelli and Barron.

This work addresses what we consider to be the "second half" of neural complexity theory, that which deals with informational issues, and numbers of examples needed to code given tasks into neural networks. Specifically, issues of neural complexity (i.e. number of neurons needed to compute an i-o function within a given tolerance), have been studied by the above-mentioned authors. The equally important issue of informational complexity, i.e., the number of examples of an i-o function needed to approximate it within a given tolerance, has not been dealt with. We consider here this second complexity theory, which we believe is open to a good deal of development in the future. We will in fact consider the interaction of these two complexity theories for feed-forward networks, a specialization of which will be the neural complexity mentioned above. In particular we introduce the parameter k, which is the number of examples available to us in the construction of an RBF neural network. We will show that, just as neural complexity theory has many of

its roots in classical approximation theory, the theory of informational complexity is closely related to analytic complexity theory as it is currently practiced.
Our general aim here is to develop algorithms which optimize the amount of information (i.e., number of examples) and the number of neurons necessary to compute i-o functions in given classes, to show these algorithms are optimal or almost-optimal, and to show that they apply to examples of interest. It is a somewhat comforting conclusion that the so-called radial basis function (RBF) algorithms satisfy optimality criteria, and are indeed best-possible in a very strong sense. The intent of this work to show that this is true both in a theoretical and practical context.

Open Problems Concerning Low-Discrepancy Point Sets

Gerhard Larcher
Universität Salzburg

Low-discrepancy point sets are the basis for essentially all quasi-Monte Carlo methods. There exist powerful construction methods for low-discrepancy point sets, for example the general concept of digital nets. Moreover there exist discrepancy estimates which show that these methods provide point sets with the smallest possible order of magnitude for their discrepancy. However these discrepancy estimates are valid only if the number of points is large enough, and, in general, too large for concrete applications. It is pointed out, that for reasonably large dimension for no reasonably small point set strong discrepancy estimates are known. So for example it is not known if there is a constant $C < 1$ such that for every dimension s there exists a point set of 2^s points in the s-dimensional unit cube whose star-discrepancy is less than C. Further so-called shift-nets are introduced. These shift nets are point sets of 2^s points in the s-dimensional unit cube with excellent distribution properties. An average type estimate for the discrepancy of these shift nets is presented. It is estimated that these shift nets could satisfy the discrepancy property stated in the first problem.

Quasi-Monte Carlo Simulation of Diffusion Equations

Christian Lécot
Laboratoire de Mathématiques, Université de Savoie

The most famous particle method used to solve diffusive problems is the Monte Carlo method: the diffusive term is modeled by random motions of the particles according to a suitable probability law. However the random choices introduce a large amount of noise leading to inaccurate computations. A step toward improving the accuracy is to replace the pseudorandom numbers by low-discrepancy sequences. Quasi-Monte Carlo methods are deterministic versions of Monte Carlo methods. In the last ten years their appeal has broadened significantly since it was found that in

certain type of computational problems they significantly outperform Monte Carlo methods.
In the present work, we develop a particle method for solving a diffusion equation

$$\frac{\partial c}{\partial t}(x,t) = \nabla \cdot (D\nabla c)(x,t), \quad x \in R^s, t > 0,$$
$$c(x,0) = c_0(x), \quad x \in R^s.$$

We discretize time into intervals of length Δt. Discretizing c as a sum of δ-measures, we have

$$c^{(n)}(x) = \frac{1}{N} \sum_{0 \le j < N} \delta(x - x_j^{(n)}),$$

where $x_j^{(n)},\ 0 \le j < N$ represent the location of N particles at time $t_n = n\Delta t$. At every time step, each particle is moved by a quasirandom displacement.
Quasirandom points, unlike pseudorandom points, are highly correlated with one another by design. It means that they cannot be blindly used in place of pseudorandom points without the risk of introducing biases into the results. We have found that for simulation of diffusion we can use quasirandom sequences by making use of a recently developed technique involving renumbering the particles according to their position after each time step. With quasirandom points and renumbering, Monte Carlo results can be significantly improved in terms of error versus number of particles.
We state and prove a convergence theorem for this method. Pseudorandom and quasirandom sequences are compared in computational experiments, for some simple demonstration problems whose solutions can be found analytically. The results show that quasirandom points can produce more accurate results than pseudorandom points.

Numerical Computation of Characteristic Polynomials of Boolean Functions and its Applications

David Lee
Bell Labs, Lucent Technologies

We study the problem of evaluation of characteristic polynomials of Boolean functions with applications to combinational circuit verification. Two Boolean functions are equivalent if and only if their corresponding characteristic polynomials are identical. However, to verify the equivalence of two Boolean functions it is often impractical to construct the corresponding characteristic polynomials due to a possible exponential blow-up of the terms of the polynomials. Instead, we compare their values at a sample point without explicitly constructing the characteristic polynomials. Specifically, we sample uniformly at random in a unit cube and determine whether two characteristic polynomials are identical by their evaluations at the sample point;

the error probability is zero when there are no round-off errors. In the presence of round-off errors, we sample on regular grids and analyze the error probability. We discuss in detail the Shannon expansion for characteristic polynomial evaluation.

Efficiency of Monte Carlo Integration Using Markov Chains

Peter Mathé
WIAS Berlin

We investigate the efficiency of Monte Carlo integration when direct sampling is not adequate. It appears that the spectral properties of the underlying Markov chain are important, esp. the second largest eigenvalue β. If $\tau = \frac{1+\beta}{1-\beta}$ denotes the autocorrelation time, then

$$\lim NE \left| \int f \, d\mu - \frac{1}{N} \sum f(x_j) \right|^2 \leq \tau \mathrm{Var}(f).$$

Thus $N_{\mathrm{eff}} = \frac{N}{\tau}$ is the effective sample size.
Applications to Metropolis Sampling are indicated.

On the Computational Structure of the Connected Components of Hard Sets

Klaus Meer
RWTH Aachen
(joint work with M. Matamala)

The study of sparse sets has tremendous importance in Turing complexity theory. Thus it is a natural task to work out a related notion for real number models of computation. This has not fully been done so far. In this talk we suggest such a notion which turns out to share specific aspects of classically sparse sets. We call a set well structured (ws) if it has few connected components which additionally can be handled computationally in a specific sense. We are going to analyze the (non-)existence of ws sets which are complete within the Blum-Shub-Smale model of computation over the reals with linear operations and equality resp. inequality branches.

Optimal Approximation of Stochastic Differential Equations with Adaptive Step-Size Control, Part I

Thomas Müller-Gronbach
Freie Universität Berlin
(joint work with N. Hofmann and K. Ritter)

We study the pathwise strong approximation of scalar stochastic differential equations on the unit interval with respect to the global error in the L_2-norm. Usual

methods use a fixed number of observations of the driving Brownian motion which are taken at equally spaced discretization points. Here we introduce an adaptive step-size control for the Euler scheme. For equations with additive noise, we show that the new method is asymptotically more efficient than an Euler scheme that is based on an equidistant discretization. This superiority is confirmed in simulation experiments for equations with additive noise as well as for general scalar equations.

Intractability Results for Positive Quadrature Formulas

Erich Novak
Universität Erlangen-Nürnberg

Lower bounds for quadrature formulas with positive weights are proved. We get intractability results for quasi-Monte Carlo methods and, more generally, for positive formulas. We consider Hilbert spaces of functions with bounded mixed derivatives. For these Hilbert spaces, the known (Frolov, Bykovskii) optimal order of convergence is

$$n^{-k} \cdot (\log n)^{(d-1)/2},$$

i.e., it is independent of the dimension d. Nevertheless we prove a lower bound

$$e(Q_n)^2 \geq 1 - n\, c_k^d$$

for the error $e(Q_n)$ of any quadrature formula Q_n with n weights, all positive. Here $c_k < 1$ is explicitly known. To prove such a result, we first prove lower bounds for relatively small classes of trigonometric polynomials, where we have explicit formulas for the kernel. We also conjecture that similar lower bounds hold for arbitrary quadrature formulas and state different equivalent conjectures concerning positive definiteness of certain matrices and certain extremal problems for trigonometric polynomials.

We also study classes of functions with weighted norms where some variables are "more important" than others. We prove the following. Positive quadrature formulas are then tractable iff the sum of the weights is bounded. The proof of the upper bound is nonconstructive, we use a technique from a recent paper of Sloan and Woźniakowski.

Beating Monte Carlo

A. Papageorgiou and J. F. Traub
Department of Computer Science, Columbia University

Testing of a variety of financial derivatives has continued. Quasi-Monte Carlo (QMC) with generalized Faure Points has been consistently superior to Monte Carlo (MC)

by one to three orders of magnitude depending on the accuracy. These problems are non-isotropic and a possible explanation may be found in the paper by Sloan and Woźniakowski (J. Complexity, March, 1998).
Testing of Value at Risk (VaR) by Papageorgiou and Paskov has commenced. Preliminary results indicate that QMC is again superior to MC.
An isotropic model problem from physics has also been tested. QMC is vastly superior to both MC and classical numerical methods. An analysis by Papageorgiou (paper in progress) explains the empirical results.
There are two classes of integrands for which QMC has been shown superior to MC. An open question is to characterize for which classes QMC beats MC.

Brakhage's Method and Complexity of Equations with Operators having Closed Range

Sergei V. Pereverzev
Ukrainian Academy of Sciences, Kiev
(joint work with E. Schock)

We investigate the information complexity of finding the minimal norm solutions T^+y of equations

$$Tx = y \tag{1}$$

with operators T having closed range. The equations (1) are considered in the Hilbert space X.
Let $e_1, e_2, \dots, e_m, \dots$ be some orthonormal basis of the Hilbert space X and P_m be the orthogonal projector on $span\,\{e_1, e_2, \dots, e_m\}$. For $r \in (0, \infty)$ we let X^r denote a linear subspace of X which is equipped with the norm

$$\|u\|_{X^r} := \|u\|_X + \|D_r u\|_X,$$

where D_r is some linear (non-bounded) operator acting from X^r to X, and for any $m = 1, 2, \dots$

$$\|I - P_m\|_{X^r \to X} \le c_r m^{-r}.$$

We assume that operators T of equations (1) have some special structure. Namely,

$$T = B + A$$

where B is some fixed operator such that $B, B^* \in \mathcal{L}(X, X) \cap \mathcal{L}(X^r, X^r)$ and

$$A \in \mathcal{H}^r_\gamma := \{A : \|A\|_{X \to X^r} \le \gamma_1,\ \|A^*\|_{X \to X^r} \le \gamma_2,\ \|(D_r A)^*\|_{X \to X^r} \le \gamma_3\},$$

$$\gamma = (\gamma_1, \gamma_2, \gamma_3).$$

For fixed B such that $B, B^* \in \mathcal{L}(X, X) \cap \mathcal{L}(X^r, X^r)$ we denote by $\mathcal{K}^r_{B,\gamma}$ the set of operators $T \in B \oplus \mathcal{H}^r_\gamma$. If $Range(T) \ne X$ or $Ker(T) \ne 0$ then the problem (1) is

not well posed in the sense of Hadamard and the crux of the difficulty is that only an approximation $y_\delta \in X$ to $y \in Range(T)$ is available such that $\|y - y_\delta\|_X \leq \delta$, where δ is a known error bound.
Let $N_\delta(\mathcal{K}^r_{B,\gamma}, X^r_\rho)$ be the minimal amount of discrete information which allows to obtain the best possible order of accuracy $O(\delta)$ of recovering solutions T^+y of equations (1) with $T \in \mathcal{K}^r_{B,\gamma}$, $y \in Range(T) \cap X^r_\rho$ from perturbed equations $Tx = y_\delta$, where $X^r_\rho = \{g : \ \|g\|_{X^r} \leq \rho\}$.
Theorem. If $dim\, Ker(B) < \infty$, $Range(B)$ is closed and for the pretabulated width of the ball X^r_μ we have the estimate

$$\Delta_n(X^r_\mu, X) \geq c\, n^{-r}, \quad n = 1, 2, \ldots,$$

then

$$c_1\, \delta^{-1/r} \leq N_\delta(\mathcal{K}^r_{B,\gamma}, X^r_\rho) \leq c\, \delta^{-1/r} \log^{1+1/r} \frac{1}{\delta}.$$

It is also shown that the algorithm consisting in combination of Brakhage's method with some hyperbolic cross discretization scheme is order optimal in the sense of quantity $N_\delta(\mathcal{K}^r_{B,\gamma}, X^r_\rho)$.

Numerical Integration and Approximation of Piecewise Analytic Functions

Knut Petras
University of Munich

We consider a set of piecewise analytic functions on $[0, 1]$, where the regions of analyticity are not known, i.e., we define the class

$$\begin{aligned} F = \{f : [0,1] \to R \ \mid\ & \text{There exist } 0 = z_0 < z_1 < \ldots < z_{k-1} < z_k = 1 \\ & \text{and an open set } \Omega = \Omega(f) \subset C \text{ such that} \\ & \bigcup_{i=1}^{k} (z_{i-1}, z_i) \subset \Omega, \quad f \text{ analytic and bounded on } \Omega\} \end{aligned}$$

Let first function evaluations be the only type of admitted observations and let Ω consist of circles with midpoints $(z_i + z_{i-1})/2$ and radii $(z_i - z_{i-1})/2$, $i = 1, \ldots, k$. If the singularities $z_0, \ldots z_k$ and a bound for f on Ω are known, the complexity of the problem of integrating or approximating uniformly functions of that kind with an accuracy ε is $\asymp |\ln \varepsilon|^2$.
We now admit a certain kind of information on bounds and on the analyticity of f in a (complex) neighborhood of subintervals of $[0, 1]$. Then, it can be shown that we still obtain the complexity $\asymp |\ln \varepsilon|^2$ also for functions with unknown singularities, unknown bound on f and smaller region of analyticity. We argue that we may often obtain such type of information automatically on a computer within a time comparable to one function evaluation.

A New Algorithm and Worst Case Complexity for Feynman-Kac Path Integration

Leszek Plaskota
Institute of Applied Mathematics and Mechanics, University of Warsaw
(joint work with G.W. Wasilkowski and H. Woźniakowski)

We study algorithms for computation of Feynman-Kac path integrals, i.e., infinite dimensional integrals of the form

$$\int_{C([0,1])} v(x(t)) H\left(\int_0^t V(x(s))ds\right) w(dx),$$

where w is the classical Wiener measure on the space of continuous functions $C([0,1])$. A typical approach is to replace the function x by its finite dimensional approximation and to reduce the problem to a multivariate integral over R^n. The latter can be solved, for instance, by Monte Carlo.

We propose a different approach. We consider the worst case (instead of Monte Carlo) setting and construct a new algorithm which uses function values of the initial condition v and potential function V. Explicit bounds on the cost of the new algorithm to compute an ε-approximation to the Feynman-Kac path integral are presented.

We also establish bounds on the worst case complexity of Feynman-Kac path integration. The upper bound is equal to the cost of the new algorithm, and is given in terms of the complexity of a certain weighted function approximation problem. The lower bound is given in terms of the complexity of a certain weighted integration problem. For some classes of functions, these two bounds coincide. In this case, the new algorithm is almost optimal.

The new algorithm requires precomputation of certain real coefficients that are combinations of some weighted multivariate integrals with special weights. In general, the precomputation with high precision is a difficult task and probably requires the design of a special method. Unless we have such an efficient method, the practical use of the new algorithm will be most likely restricted to moderate accuracy ε.

We report the results of the precomputation for some specific cases.

Optimal Approximation of Stochastic Differential Equations with Adaptive Step-Size Control, Part II

Klaus Ritter
Universität Passau
(joint work with N. Hofmann and T. Müller-Gronbach)

We present lower bounds for the pathwise solution of scalar stochastic differential equations $dX(t) = a(t, X(t))\,dt + \sigma(t)\,dW(t)$ with additive noise. We study the class of methods that uses values of the driving Brownian motion W at n adaptively

chosen points. The minimum over the errors $\left(E\|X-\overline{X}_n\|_2^2\right)^{1/2}$ of arbitrary methods $\overline{X}_n$ of the above type tends to zero like $1/\sqrt{6}\cdot\|\sigma\|_1\cdot n^{-1/2}$. We conclude that the Euler method with adaptive step-size control is asymptotically optimal. Finally we explain why 'higher-order methods' do not exist if the error is defined globally with respect to some norm.

Stochastic Models and Algorithms for Solving Smoluchovski Equation Governing Coagulation of Particles in Turbulent Flows

Karl Sabelfeld
WIAS Berlin

Stochastic models and algorithms for solving equations with random parameters are suggested. The case of Smoluchovski equation governing aggregation-disaggregation process in turbulent collision regime is treated in details. We suggest a stochastic Lagrangian approach which enables to reduce an inhomogeneous Smoluchovski equation to homogeneous case for which an efficient Monte Carlo procedure is developed. Along Lagrangian trajectories, a random velocity process is simulated by solving a system of stochastic differential equations.
Two main theorems are proved: the first theorem concerns the convergence in probability of the stochastic algorithm as N, the number of interacting particles, goes to infinity, to the solution of a finite-difference approximation of the Smoluchovski equation. Second theorem proves the convergence of this approximation to the exact solution. The cost of the algorithm is estimated by $\mathcal{O}(N^2)$.

Condition Number of the Bernstein-Bezier-Basis

Karl Scherer
Institut für Angewandte Mathematik, Bonn
(joint work with T. Lyche)

The condition number of the triangular Bernstein-Bezier basis in L^p- sense is defined by

$$\kappa_{n,p}(S^d)=\sup_{\mathrm{c}\neq 0}\frac{||\mathbf{c}||_\infty}{\left\|\sum_{|\mathrm{j}|=n}c_{\mathrm{j}}B_{\mathrm{j}}^n(\mathbf{x})\right\|_{L^p(S^d)}}.$$

where S^d is the simplex

$$S^d:=\{\mathbf{x}\in R^d: x_i\geq 0, x_1+...+x_d:=|\mathbf{x}|\leq 1\}$$

and

$$B_{\mathrm{j}}^n(\mathbf{x}):=\frac{n!}{j_1!\cdots j_d!}x_1^{j_1}\cdots x_d^{j_d}(1-|\mathbf{x}|)^{n-|\mathrm{j}|}\qquad ,|j|:=j_1+\cdots+j_d$$

are the triangular Bernstein-polynomials of total degree at most n in d variables. We give an upper bound for the L^∞ condition number which grows like $(d+1)^n$ when d is fixed and n tends to infinity. Moreover the upper bound is independent of d for $d \geq n-1$. In the L^2- case the condition number $\kappa_{n,2}(S^d)$ is determined exactly and equal to $\sqrt{\binom{2n+d}{n}}$. The extremal polynomials (with worst condition number) are the Legendre -polynomials for the simplex S^d. From this there follows an improved upper bound for $\kappa_{n,\infty}(S^d) = O(2^n)$ when d is fixed and n tends to infinity. However for fixed n and d tending to infinity $\kappa_{n,2}(S^d)$ does not remain bounded .

Approximating Fixed Points of Mildly Contractive Functions

K. Sikorski
University of Utah
(joint work with Z. Huang and L. Khachiyan)

We derive new upper bounds on the complexity of interior ellipsoid, centroid and circumscribed ellipsoid algorithms for approximating fixed points of mildly contractive functions. To solve the problem in the residual sense $\|f(x)-x\| \leq \delta$ we need at most $\mathcal{O}(d \log \frac{1}{\delta})$ function evaluations, where d is the dimension, and the contractive factor $q \leq 1$. This bound is achieved by the interior ellipsoid and centroid algorithms. As a consequence we need at most $\mathcal{O}(d(\log \frac{1}{\epsilon} + \log \frac{1}{1-q}))$ function evaluations to solve the problem in the absolute sense $\|x-\alpha\| \leq \epsilon$, $f(\alpha) = \alpha$, for $q < 1$. We obtain corresponding bounds $\mathcal{O}(d^2 \log \frac{1}{\delta})$ and $\mathcal{O}(d^2(\log \frac{1}{\epsilon} + \log \frac{1}{1-q}))$ for the circumscribed ellipsoid algorithm.

Monte Carlo Complexity of Parametric Integration

Eugène Sindambiwe
Universität Kaiserslautern
(joint work with S. Heinrich)

The Monte Carlo complexity of computing integrals depending on a parameter is analyzed for smooth integrands. An optimal algorithm is developed on the basis of a multi-grid variance reduction technique. The complexity analysis implies that our algorithm attains a higher convergence rate than any deterministic algorithm. Moreover, because of savings due to computation on multiple grids, this rate is also higher than that of previously developed Monte Carlo algorithms for parametric integration.

Nearly Optimal Polynomial Projections on the Sphere

Ian H. Sloan
University of New South Wales

The paper considers uniform approximation of continuous functions on the unit sphere $S^{r-1} \subseteq R^r$ by spherical polynomials. The approximations are of the form $T_n f \in P_n^{(r)}$, where $P_n^{(r)}$ is the set of spherical polynomials on S^{r-1} of degree $\leq n$, and T_n is a linear projection. Recently it has been shown that the so-called hyperinterpolation approximation $L_n f$ has, under a mild regularity assumption on the underlying quadrature rule, the optimal rate of growth, namely $\Theta(n^{1/2})$, for its norm $\|L_n\|$. This and related issues are discussed, and illustrated by numerical examples.

Optimization of Projection Methods for Solving Linear Ill-Posed Problems

Sergei G. Solodky
Ukrainian Acad. of Sciences, Kiev

For the integral equation of the first kind

$$\int_0^1 a(t,\tau)\, x(\tau)\, d\tau = f(t) \tag{1}$$

we study the problem of optimal finite-dimensional approximation to the minimal norm solution. At present, such optimization problems are considered in the context of information-based complexity. To discretize (1) we construct a new projection scheme using the idea of hyperbolic cross. It is shown that the combination of the indicated scheme and Tikhonov regularization method with the parameter selection according to discrepancy principle allows to attain the optimal order of accuracy using the minimal amount of Galerkin functionals and arithmetic operations on their values. In the periodic case we establish, that the application of the Smolyak grid allows to keep the main order of the complexity even when the discrete information is given only in the standard form.

Interpolation on Sparse Grids and Nikol'skij–Besov Spaces of Dominating Mixed Smoothness

Frauke Sprengel
University of Rostock
(joint work with W. Sickel)

In this paper, we deal with the error of approximation of periodic functions obtained by interpolation on sparse grids. The method of j-th order blending itself is well–known. What is new here is the choice of the underlying function spaces. Former

papers (e.g. Delvos and Schempp (1989), Baszenski and Delvos (1989) or Pöplau and Sprengel (1997)) dealt with tensor products of spaces defined by certain decay properties of the Fourier coefficients (Korobov spaces or potential spaces built on L_2, respectively). Based on a recent paper by Sickel (1998), we are able to investigate the problem in more appropriate spaces, namely on Nikol'skij–Besov spaces and its tensor products, the classes $S^{s_1,s_2}_{p,\infty}B(T^2)$.
Interpolation on sparse grids can be reduced to error estimates of corresponding interpolation processes in the one–dimensional situation. As a result we obtain estimates like

$$\| f - B_j f \,|L_p(T^2)\| \le C\, j\, 2^{-js}\,, \tag{1}$$

where B_j denotes the interpolation operator with respect to the sparse grid (having $j\,2^j$ knots approximately). Comparing this with interpolation on the full grid (having 2^{2j} knots approximately), we would end with an estimate

$$\| f - I_j f \,|L_p(T^2)\| \le C\; 2^{-js}\,. \tag{2}$$

So, except a logarithmic term the error is of the same order. The price we have to pay consists in the following: whereas (2) is true for all functions taken from the Nikol'skij–Besov space $B^s_{p,\infty}(T^2)$ inequality (1) holds true for elements of the space of dominating mixed smoothness $S^{s,s}_{p,\infty}B(T^2)$ only.
We restrict ourselves to the bivariate situation mostly for transparency. All can be done also in the higher dimensional situation.

Quadrature Formulas for Wiener Integrals

Achim Steinbauer
Universität Erlangen-Nürnberg
(joint work with E. Novak and K. Ritter)

We present a sequence of quadrature formulas for integrals $I_\infty(f) = \int_C f(x)\,dw(x)$ with respect to the Wiener measure w on $C = C[0,1]$.
The integrals $I_\infty(f)$ are approximated by d-dimensional integrals over the space of piecewise linear functions with d equidistant breakpoints on [0,1]. We use a multiscale decomposition of the Wiener measure, due to Lévy and Ciesielski to obtain a problem which can be effectively treated by a modified version of the well known Smolyak algorithm.
The product formulas $U^{i_1} \otimes \cdots \otimes U^{i_d}$ of the original algorithm

$$A(q,d) = \sum_{q-d+1 \le |i| \le q} (-1)^{q-|i|} \binom{d-1}{q-|i|} (U^{i_1} \otimes \cdots \otimes U^{i_d}),$$

with $q \ge d$, $i \in N^d$, $|i| = i_1 + \ldots + i_d$, are replaced by formulas $U^{i_1-\beta_1} \otimes \cdots \otimes U^{i_d-\beta_d}$, where $U^j = U^1$ for $j < 1$.

The entries of β are chosen according to the importance of the respective coordinates in the d-dimensional integral. Numerical examples indicate, that the new method works well, if the integrand is smooth.

Numerical Experiments on Quasi-Monte Carlo for Derivative Pricing

Shu Tezuka
IBM Tokyo Research Laboratory

There are two different types of financial derivatives: fixed-income and equity.
Integrands associated with pricing fixed-income derivatives are classified as weighted class of integrands, for which the importance of the variables changes as the index goes up.
In this talk, we focus on equity derivative pricing, for which the importance of the variables of the integrand is equal.
We discuss integration problems associated with pricing equity derivatives with the Black-Scholes model and show several numerical results on the performance of quasi-Monte Carlo methods compared with ordinary Monte Carlo results. Future issues are also discussed.

Low Discrepancy Sequences and Financial Modeling

Robert Tichy
Technische Universität Graz

In the first part some stochastic models from ruin theory are discussed. In particular we consider the classical ruin model with linear dividend barrier and deterministic and stochastic interest rates. Furthermore stochastic parameter changes are studied as well as option payments in the case of high damages. The models are solved analytically, numerically and by means of simulation methods.The simulation methods make use of low discrepancy sequences. We present a comparison of different kinds of such sequences, mainly of so called net sequences.

"Smooth" Discrete-Stochastic Procedures for Global Estimation of Functions

A.V. Voytishek
Novosibirsk, Russia

The question is about construction, convergence and optimization of discrete-stochastic numerical procedures (DSNP) of global approximation of a function $\varphi(.)$ presented in an integral form on a compact set $D \subset \mathbf{R}^l$. An integral depended on a parameter and a solution of integral equation of the second kind can be considered as examples of the function $\varphi(.)$.

The construction of DSNP is provided as follows. Introduce a grid $\{\hat{x}_j\}$ in $\mathbf{R}^l$ such that the finite number M of it's nodes $\{x_i,\ i = 1, \ldots, M\}$ belongs to D. Implement Monte Carlo algorithms (with numbers of realizations $\{n_i\}$) for estimation of integral representation of values $\{\varphi(x_i)\}$. Then using a transformation $\Gamma(.)$ and the obtained values at the grid nodes $\{\tilde{\varphi}_{n_i}(x_i),\ i = 1, \ldots, M\}$ approximate the function $\varphi(.)$:

$$J_{(n_1,\ldots,n_M)}(x) = \Gamma(\tilde{\varphi}_{n_1}(x_1), \ldots, \tilde{\varphi}_{n_M}(x_M);\, x), \quad \tilde{\varphi}_{n_i}(x_i) = \frac{1}{n_i} \sum_{q=1}^{n_i} \eta_q^{(i)}\,, \tag{1}$$

where $\eta_q^{(i)}$ is the q-th realization of $\eta^{(i)}$ and $\eta^{(i)}$ is a unbiased stochastic estimator of $\varphi(x_i)$ used in the Monte Carlo method, $i = 1, \ldots, M$.
In previous works of A.Voytishek and E.Shkarupa the Strang-Fix approximation was considered as $\Gamma(.)$ from (1), but statements on convergence and optimization of DSNP were formulated only for the simplest particular case of multilinear approximation. Choosing of the Strang-Fix approximation is based on the simplicity of construction and possibility for error estimation and for calculation of optimal parameters of DSNP. The most useful property of the Strang-Fix approximation is the "error concentration at the grid nodes", which was proved earlier only for the multilinear case.
It is possible to demonstrate that the property of the "error concentration at the grid nodes" is true for approximations based on decomposition of the function $\varphi(.)$ of the form

$$L\hat{\varphi}(x) = \sum_j \hat{\varphi}(\hat{x}_j)\, \chi_j(x),$$

where $\{\chi_j(.)\}$ is a basis of functions, for which the condition

$$\sup_{x \in D} \sum_j |\chi_j(x)| \le C, \quad C = \text{const}$$

is true. We also prove that the Strang-Fix approximation with the B-spline "generating" function is the interpolation of this type (in this case $C = 1$).
This work was carried out with the support of the Russian Fund of Fundamental Research (project 97-01-00855) and the program "Universities of Russia – Fundamental Research" (project 3759).

Weighted Tensor-Product Algorithms for Linear Multivariate Problems

G. W. Wasilkowski
Department of Computer Science, University of Kentucky
(joint work with H. Woźniakowski)

We deal with an ε-approximation of multivariate problems defined over weighted tensor-product Hilbert spaces of functions of d variables. We define a class of weighted tensor-product (WTP) algorithms which use evaluations about the function given

by arbitrary linear functionals or function values. These algorithms depend on a number of parameters. We study tractability of multivariate problems by checking when WTP algorithms are polynomial time algorithms, i.e., when the minimal number of evaluations is polynomial in $1/\varepsilon$ and d. For the class of arbitrary linear functionals, we select a WTP algorithm which is optimal and provide a necessary and sufficient condition for tractability in terms of the sequence of weights and the sequence of singular values for $d = 1$. For the class of function values, we select a WTP algorithm which is only optimal for some sequences of weights.

Where does Smoothness Count the most for Two-Point Boundary-Value Problems?

Arthur G. Werschulz
Fordham University and Columbia University, New York

We study the complexity of scalar $2m$th order elliptic two-point boundary-value problems $Lu = f$. Previous work on the complexity of these problems has generally assumed that we had partial information about the right-hand side f and complete information about the coefficients of L. In this paper, we study the complexity of such problems when, in addition to partial information about f, we have only partial information about the coefficients of L. More precisely, we suppose that f has r derivatives in the L_p-sense, with $r \geq -m$ and $p \in [2, \infty]$, and that L has the usual divergence form

$$Lv = \sum_{0 \leq i,j \leq m} (-1)^i D^i (a_{i,j} D^j v),$$

with $a_{i,j}$ being $r_{i,j}$-times continuously differentiable, where $r_{i,j} \geq 0$. We first suppose that continuous linear information is available. Let

$$\tilde{r} = \min\{\, r, \min_{0 \leq i,j \leq m} \{r_{i,j} - i\} \,\}.$$

If $\tilde{r} = -m$, the problem is unsolvable; for $\tilde{r} > -m$, we find that the ε-complexity is proportional to $(1/\varepsilon)^{1/(\tilde{r}+m)}$, and we show that a finite element method (FEM) is optimal. We next suppose that only standard information (consisting of function and/or derivative evaluations) is available. Let

$$r_{\min} = \min\{\, r, \min_{0 \leq i,j \leq m} \{r_{i,j}\} \,\}.$$

If $r_{\min} = 0$, the problem is unsolvable; for $r_{\min} > 0$, we find that the ε-complexity is proportional to $(1/\varepsilon)^{1/r_{\min}}$, and we show that a modified FEM (which uses only function evaluations, and not derivatives) is optimal.

Efficiency of Quasi-Monte Carlo Algorithms for High Dimensions

Henryk Woźniakowski
Columbia University and University of Warsaw

Recently quasi-Monte Carlo algorithms have been successfully used for multivariate integration of high dimension d, and were significantly more efficient than Monte Carlo algorithms. The existing theory of the worst case error bounds of quasi-Monte Carlo algorithms does not explain this phenomenon.

In a paper with Ian H. Sloan we present a partial answer to why quasi-Monte Carlo algorithms can work well for arbitrarily large d. It is done by identifying classes of functions for which the effect of the dimension d is negligible. These are *weighted* classes in which the behavior in the successive dimensions is moderated by a sequence of weights γ_j. We prove that the minimal *worst case* error of quasi-Monte Carlo algorithms does not depend on the dimension d iff $\sum_{j=1}^{\infty} \gamma_j < \infty$. We also prove that the minimal number of function values in the worst case setting needed to reduce the initial error by ε is bounded by $C\varepsilon^{-p}$, where the exponent $p \in [1,2]$, and C depends exponentially on the sum of weights. Hence, the relatively small sum of the weights makes some quasi-Monte Carlo algorithms efficient for all d.

We show in a non-constructive way that many quasi-Monte Carlo algorithms are strongly tractable. Even random selection of sample points (done once for the whole weighted class of functions and then the worst case error is established for that particular selection, in contrast to Monte Carlo where random selection of sample points is carried out for a fixed function) leads to efficient quasi-Monte Carlo algorithms. In this case the minimal number of function values in the *worst case* setting is of order ε^{-p} with the exponent $p = 2$.

We also present some results for more general algorithms than quasi-Monte Carlo algorithms. In a paper with F. J. Hickernell we prove in a *non-constructive* way that there exist algorithms for which the exponent $p = 1$ whenever $\sum_{j=1}^{\infty} \gamma_j^{1/2} < \infty$. In a paper with G. W. Wasilkowski we show in a *constructive* way that $p = 1$ whenever $\sum_{j=1}^{\infty} \gamma_j^{1/3} < \infty$. This is achieved by a *weighted tensor product* algorithm which can be also applied for more general linear problems than multivariate integration.