

Dagstuhl–Seminar
Algorithms and Complexity for Continuous Problems

ORGANIZED BY:

STEFAN HEINRICH (UNIVERSITÄT KAISERSLAUTERN)
JOSEPH F. TRAUB (COLUMBIA UNIVERSITY, NEW YORK)
HENRYK WOŹNIAKOWSKI (COLUMBIA UNIVERSITY, NEW YORK)

OCTOBER 17–21, 1994

Overview

The Seminar was attended by 39 scientists from 11 countries. The main subjects of the seminar were: Complexity of the approximate solution of continuous problems, multivariate problems and algorithms for their solution, randomized algorithms, path integrals, low discrepancy sequences, approximation and n -widths, average case analysis, combinatorial continuous complexity. Contributions spanned the whole domain from theoretical foundation up to concrete applications such as finance, geology, gas dynamics, and computer graphics. The seminar provided a forum of exchange of ideas between these different topics, based on the general unifying idea of complexity. In addition to the lecture program we had two sessions on open problems and directions of research, one about approximation, the other about complexity and algorithms for multivariate problems.

The Dagstuhl center provided us with excellent working and living conditions, we profited very much from the efficiency of the staff of Schloss Dagstuhl as well as of the office in Saarbruecken. Our thanks go to these people for their considerable help in making the seminar a success.

Abstracts

On Optimization of Quadrature Formulae

Vladislav F. Babenko
Dnepropetrovsk State University, Ukraine

We want to compute $I_\mu(x) = \int_0^1 x(t) \mu(dt)$ (μ is a given measure on $[0, 1]$ which is absolutely continuous with respect to the Lebesgue measure) using Monte Carlo rules of the form $q(x) = \sum_{i=1}^n c_i x(t_i)$.

Let P_1 be some probability on $[0, 1]^n$ and P_2^1 be some transitional probability on $[0, 1]^n \times B^n$, where B^n is the σ -field of Borel subsets of \mathbb{R}^n . We consider a pair (P_1, P_2^1) as a model of a given Monte Carlo method and obtain the exact solution for some problem of optimal Monte Carlo integration and optimal Monte Carlo recovery of some more general functionals.

For example let $g \in L_1[0, 1]$ and $f_g(x) = \int_0^1 x(t) g(t) dt$. Denote by MC_{ac} the class of Monte Carlo methods (P_1, P_2^1) such that P_1 is absolutely continuous with respect to the Lebesgue measure on $[0, 1]^n$. Then we have for the optimal error $R_n(B_\infty, f_g, MC_{ac})$ of recovery of f_g on the unit ball $B_\infty \subset L_\infty[0, 1]$ by methods from MC_{ac}

$$R_n(B_\infty, f_g, MC_{ac}) = \frac{\|g\|_1}{1 + \sqrt{n}}.$$

Improving the Upper Bound for the Expected Number of Shadow Vertices

Karl-Heinz Borgwardt
Universität Augsburg

The talk deals with the number of vertices which are visited by a variant of the Simplex method, the shadow-vertex algorithm, during the solution process of linear programming problems of the type maximize $v^T x$ subject to $a_1^T x \leq 1, \dots, a_m^T x \leq 1$; $v, x, a_1, \dots, a_m \in \mathbb{R}^n$. It had been shown that there are the shadow vertices, which keep the vertex property even when the feasible region is (orthogonally) projected on $\text{span}(u, v)$. In 1982 we had derived an upper bound for the expected number of such shadow vertices under the stochastic assumption that a_1, \dots, a_m, v (and u) are distributed identically, independently and symmetrically under rotations on $\mathbb{R}^n \setminus \{0\}$. That bound had the form $E_{m,n}(S) \leq \text{const} \cdot m^{1/(n-1)} n^3$. Recently, we

managed to improve that bound. We explain, how and why a refined calculation of space-angles, a better approximation of the marginal density function of the uniform distribution on the unit sphere and a full reorganization of the integration process for the appearing convolution integrals leads to a sharper bound

$$E_{m,n}(S) \leq \text{const} \cdot m^{1/(n-1)} n^{5/2}$$

On the power of real Turing machines over binary inputs

Felipe Cucker

(joint work with D.Grigoriev)

In recent years, attention has been paid to comparing the computational power of real Turing machines and classical Turing machines. This is done by feeding the formers with binary inputs and characterizing the class of sets thus obtained under a certain restriction of the machines resources. In this talk we show that the class of binary sets decided by parallel real Turing machines working in polynomial time is exactly $\text{PSPACE}/\text{poly}$.

Directions in Approximation Theory

Ronald A.DeVore

University of South Carolina

In this talk, we survey some of the recent trends in approximation theory. Many of these are spurred by applications. Among the topics to be discussed are approximation by wavelets and neural networks. Distinctions will be made between linear and nonlinear approximation as well as the status of univariate versus multivariate problems.

Wavelets have been well studied as an approximation tool during the last several years. Univariate linear approximation is well understood and coincides with classical methods such as polynomial and spline approximation. Nonlinear approximation by m -terms in the wavelet decomposition is also well understood including the advantage of such approximation (the requirement of less smoothness in the function to be approximated) over linear methods. New directions in nonlinear approximation include the choice of an m -term approximation from an optimally chosen basis from a library of bases. The latter problem is now being promoted as a tool in signal and image processing but there is a lack of theorems showing the advantages of this approach, in particular theorems that take into consideration the complexity of the basis selection.

Multivariate wavelet approximation for the standard basis is well understood in the linear and nonlinear cases. Results similar to the univariate cases are available. Approximation by hyperbolic bases (tensor products of univariate bases) have now been studied in the case of linear approximation with a theory that compares with Fourier approximation with frequencies from the hyperbolic cross. There is not yet a complete theory understanding of nonlinear approximation in the hyperbolic case save for the special case where approximation takes place in the mean square norm.

Neural networks which are receiving ever more attention in applications is from the approximation viewpoint in its infancy. Only recently have we been able to show that approximation by linear combinations of sigmoidal functions in several variables has the same approximation efficiency as classical methods for standard smoothness classes. much remains to be done in the linear case to understand if there are other classes of functions for which neural network approximation has distinct advantages over classical methods. There are now available several theorems for nonlinear approximation by sigmoidal functions which are counterparts of m -term approximation by complex exponentials.

On best nonlinear approximations

Dinh Dung

Hanoi Institute of Information Technology, Vietnam

In nonlinear approximations such as rational approximation, approximation by splines with free knots, wavelet compression etc., functions are approximated by elements of nonlinear manifolds parametrized continuously by \mathbb{R}^n . The nonlinear n -width of the compact subset W in the normed linear space X

$$\delta_n(W, X) := \inf_{\substack{F \in C(W, \mathbb{R}^n) \\ M \in C(\mathbb{R}^n, X)}} \sup_{x \in W} \|x - M(F(x))\|$$

is a characterization of the best method of such approximations. We have

$$\delta_{2n+1}(W, X) \leq a_n(W, X) \leq \delta_n(W, X),$$

where a_n is the well-known Alexandrov width. However, an example shows that these quantities are strictly different.

Let U_p^α and $V_{p,\theta}^\alpha$ be the unit balls of the Sobolev space $W_p^\alpha(\mathbb{T}^d)$ and the Besov space $B_{p,\theta}^\alpha(\mathbb{T}^d)$ respectively. Then, for $1 \leq p, q, \theta \leq \infty$ and $\alpha > \left(\frac{1}{p} - \frac{1}{q}\right)_+$, we have

$$\begin{aligned} \delta_n(U_p^\alpha, L_q(\mathbb{T}^d)) &\approx n^{-\alpha/d} \\ \delta_n(V_{p,\theta}^\alpha, L_q(\mathbb{T}^d)) &\approx a_n(V_{p,\theta}^\alpha, L_q(\mathbb{T}^d)) \approx n^{-\alpha/d}. \end{aligned}$$

The above formulated results are obtained by Dinh Dung and Vu Quoc Thanh.

On the complexity of local solution of multivariate integral equations

Karin Frank
Universität Kaiserslautern

We study multivariate Fredholm integral equations of the second kind. Instead of computing the solution function on the whole domain, we are interested only in the value of one single functional $\chi \in L_2([0, 1]^d)$ applied to it. For two Sobolev classes of input data the exact order of the complexity for the problem of local solution with linear information is derived. The lower bounds are shown using a Gelfand number technique developed by Heinrich (1993), the upper bounds are shown in a constructive way providing algorithms of optimal order. These algorithms are based on a hyperbolic–cross type approximation of the kernel function. Numerical experiments are presented comparing both algorithms with a standard Galerkin method.

Variance Reduction for Integral Equations and Applications to Computer Graphics

Stefan Heinrich
Universität Kaiserslautern

The global illumination problem of computer graphics can be described by a Fredholm integral equation of the second kind – the radiance equation. Due to the high dimensionality of the problem and the discontinuity of the kernel, Monte Carlo methods are often used.

The crucial efficiency figure for Monte Carlo methods is variance. On the basis of previous complexity analysis due to P. Mathé and the author, a variance reduction method is developed which combines deterministic and stochastic elements to reduce variance in an optimal way. Numerical computations for simple test scenes show considerable gains in efficiency.

Quasi-Monte Carlo Methods in Computer Graphics

Alexander Keller
Universität Kaiserslautern

Realistic image generation in Computer Graphics is concerned with the solution of the radiance equation, which is a Fredholm integral equation of the second kind. Due to the complexity of the kernel of this integral equation, only Monte Carlo methods

seem to allow to model a solution close to reality. In our work we compare (pseudo-) random to quasi-random approaches to solve the global illumination problem. Experiments show that the application of high dimensional low discrepancy point sets, as used in deterministic quasi-Monte Carlo methods, are superior to the use of pseudo-random numbers. This superiority is characterized by a smoother and faster convergence and a deterministic upper error bound, although complete mathematical theory is not yet available. We also investigate the combination of random and quasi-random techniques so as to acquire even faster algorithms.

Recovery of Functions from their Wavelet Transform

Mark Kon
Boston University

Consider a function $\psi(x)$ with localized support, e.g. $\psi(x) = \frac{d^2}{dx^2}e^{-x^2}$ or $\psi(x) = \frac{d^2}{dx^2}O(x)$, where $O(x)$ is a cubic B -spline. If $f(x)$ is a function (on \mathbb{R} or \mathbb{R}^d) with localized support (say compact support), f may be interpreted as a physical signal. Its wavelet transform is defined by

$$f(x, j) = f * \psi_j(x),$$

where $*$ denotes convolution and $\psi_j(x) = 2^j\psi(2^jx)$ is an L^1 -rescaling of ψ . This transform $f(x, j)$ is defined for all scales 2^j , $-\infty < j < \infty$, with j an integer. The information stored in $f(x, j)$ is quite redundant, and Marr (for $\psi(x) = \frac{d^2}{dx^2}e^{-x^2}$) and Mallat (for $\psi(x) = \frac{d^2}{dx^2}O(x)$ as above) studied (in the context of vision and image compression questions) whether keeping just the zeroes of $f(x, j)$ for each value of j was sufficient to reconstruct $f(x)$. If this information is sufficient, there is the further question of the complexity of reconstructing f within ε (in same norm) from this information. We study both the Marr and the Mallat questions in the context of whether the zero crossing information is sufficient, and in the case where it is sufficient, what the informational complexity is of recovery.

On the informativeness of functionals

Nikolay P. Korneichuk
Institute of Mathematics, Ukrainian Academy of Sciences, Kiev

Let (X, ρ) be a metric space with a distance $\rho = \rho(x, y)$, X' is a set of given continuous on X functionals. If Q is a bounded set in X then we put

$$D(Q, \rho) = \sup\{\rho(x, y) : x, y \in Q\}$$

and for $\mu \in X'$

$$D(\mu, Q, \rho) = \sup\{\rho(x, y) : x, y \in Q, \mu(x) = \mu(y)\}.$$

We shall call the value

$$\text{In}(\mu, Q, \rho) = D(Q, \rho) - D(\mu, Q, \rho)$$

the informativeness of the functional μ relatively of the set Q and the distance ρ . The problem is considered to find $\sup\{\text{In}(\mu, Q, \rho) : \mu \in X'\}$ and to indicate the functional μ which realizes the supremum.

We obtained the solution of this problem for some set in the space $C = C[a, b]$. Let C^* be the conjugate space and C_τ^* be the subset of functionals $\mu \in C^*$ of the form $\mu(x) = \mu_\tau(x) = x(\tau)$, $a \leq \tau \leq b$. For example, the following statement is true. Let

$$KH_0^1 = \{x(t) : x(t) \in C, |x(t') - x(t'')| \leq K|t' - t''|, x(a) = x(b) = 0\},$$

$\psi(t)$ be a fixed function in KH_0^1 , $\psi(t) \geq 0$, $a \leq t \leq b$,

$$Q = \{x(t) : x(t) \in KH_0^1, |x(t)| \leq \psi(t), a \leq t \leq b\},$$

and ρ is the C -metric or L_p -metric, $1 \leq p \leq \infty$. For any $\mu \in C^*$ exists a functional $\mu_\tau \in C_\tau^*$ such that $\text{In}(\mu_\tau, Q, \rho) \geq \text{In}(\mu, Q, \rho)$ and therefore

$$\sup\{\text{In}(\mu, Q, \rho) : \mu \in C^*\} = \sup\{\text{In}(\mu_\tau, Q, \rho) : \mu_\tau \in C_\tau^*\}.$$

An analogous result is obtained for the set $M_N = \{\mu_1, \mu_2, \dots, \mu_N\}$ of functionals from C^* . Some applications are considered.

Recovering multivariate signals

Marek Kowalski

University of Warsaw, Poland

We consider multivariate band- and energy-limited signals

$$f \in J_d = \left\{ \int_{\Omega} x(\omega) e^{i(\omega, t) d\omega} : x \in L_2(\Omega), \|x\| \leq 1 \right\}$$

where $\omega, t \in \mathbb{R}^d$ and Ω is a compact subset of \mathbb{R}^d such that $\text{Int } \Omega \neq \emptyset$. Our knowledge about each $f \in J_d$ is $N(f) = [f(t_1), f(t_2), \dots, f(t_n)]$ and we are interested in minimal

complexity recovery of f in the worst case setting. The error of an algorithm ϕ recovering f from $N(f)$ is defined as

$$e(\phi) = \sup_{f \in J_d} \|f - \phi(N(f))\|_p,$$

where $1 \leq p \leq \infty$ and $\|\cdot\|_p$ is the p -norm on the domain of recovery I (I – compact, $\text{Int } I \neq \emptyset$). Let $m(\varepsilon)$ denote the minimal number of samples of f in I required to find an algorithm ϕ^* using these samples such that $e(\phi^*) \leq \varepsilon$, $\varepsilon > 0$. We assume that the cost of arithmetic operations ($+$, $-$, $*$, $/$) and the cost of signal evaluation are taken as unity and c respectively. We define $\text{comp}(\varepsilon)$ as the minimal cost of obtaining an ε -approximation.

The main results presented in the talk are:

- $\lim_{\varepsilon \rightarrow 0^+} m(\varepsilon) \left(\frac{\log \log 1/\varepsilon}{\log 1/\varepsilon} \right)^d = 1;$
 $\text{comp}(\varepsilon) = \Theta \left(\left(\frac{\log 1/\varepsilon}{\log \log 1/\varepsilon} \right)^d \right), \text{ as } \varepsilon \rightarrow 0^+;$
- For sufficiently small $\varepsilon > 0$, Lagrangian interpolation using $m(\varepsilon)(1 + o(1))$ arbitrary grid nodes yields an ε -approximation with almost minimal cost.

Best Approximation of Some Functional Classes on Groups in L_2

G. Magaril–Il’yaev
Moscow State University
(joint work with V. Tikhomirov)

The exact values of the Kolmogorov n -widths and the Kolmogorov average ν -widths of generalized Sobolev-classes on locally compact groups in the L_2 -metric are discussed in the lecture.

Let G be an arbitrary compact group, $K(\cdot) \in L_2(G)$, $1 \leq p \leq \infty$ and

$$W_p^K(G) := \left\{ x(\cdot) \mid x(\cdot) = (K * u)(\cdot), \|u(\cdot)\|_{L_p(G)} \leq 1 \right\}.$$

Then

$$d_n \left(W_p^K(G), L_2(G) \right) = \begin{cases} \left(\sum_{m \geq n} |\hat{K}^+(m)|^2 \right)^{\frac{1}{2}}, & p = 1 \\ \hat{K}^+(n) & , p = 2. \end{cases}$$

Here $\{\hat{K}^+(\cdot)\}$ is the nonincreasing rearrangement of the sequence $\{|\hat{K}(\cdot)|\}$ where $\{\hat{K}(\cdot)\}$ is the sequence of the Fourier coefficients of $K(\cdot)$ with respect to the complete orthonormal system in $L_2(G)$ consisting of the matrix elements of all representations of G .

An analogous result was obtained for the Kolmogorov average ν -width of $W_p^K(G)$, $p = 1, 2$, where G is a noncompact Abelian group. These theorems generalize many well-known results about exact solutions of the widths-problem.

Width and distributions of values of the approximation functional on Sobolev spaces

V. E. Maiorov
Technion, Haifa, Israel

Consider the Sobolev space $W_2^r = W$ equipped with a Gaussian measure μ whose mean is zero and whose correlation operator C_μ has eigenfunctions $e^{in(\cdot)}$ and eigenvalues $\lambda_n = |n|^{-s}$, $n \in \mathbf{Z}$, $s > 1$. Let $Y^n = W \times \dots \times W$ be the product of n copies of the space W and $\nu = \mu \times \dots \times \mu$ be the Gaussian measure on Y^n .

Consider the approximation functional on $W \times Y^n$

$$e(x, y)_q = \inf \{ \|x - z\|_q : z \in \text{span}\{y_1, \dots, y_n\} \}.$$

We call the following values

$$d_{n,\varepsilon,\delta}(W, \mu, \nu, L_q) = \sup_{\nu(G) \leq \varepsilon} \inf_{y \in Y^n \setminus G} \left[\inf_{\mu(Q) \leq \delta} \sup_{x \in W \setminus Q} e(x, y)_q \right],$$

$$d_{n,\varepsilon,\delta}^{(ad)}(W, \mu, \nu, L_q) = \inf_{\mu(Q) \leq \delta} \sup_{x \in W \setminus Q} \left[\sup_{\nu(G) \leq \varepsilon} \inf_{y \in Y^n \setminus G} e(x, y)_q \right]$$

respectively the Kolmogorov and the adaptive (n, ε, δ) -widths of the Sobolev space W in the L_q -norm.

Theorem: If $r > \frac{1}{2}$, $1 < q < \infty$, $\varepsilon \in [2^{-n}, 1 - 2^{-n}]$, $\delta \in (0, \frac{1}{2}]$, then the asymptotic

$$d_{n,\varepsilon,\delta}(W, \mu, \nu, L_q) \asymp d_{n,\varepsilon,\delta}^{(ad)}(W, \mu, \nu, L_q) \asymp \frac{\sqrt{1 + \frac{1}{n} \ln \frac{1}{\delta}}}{n^{r+(s-1)/2}}$$

holds.

The optimal error of Monte Carlo Integration

Peter Mathé
WIAS Berlin

We study the problem of optimal Monte Carlo integration in the following sense:

Let $[T, \mathcal{F}, \mu]$ be a probability space with a polish space T and let $B_\infty(T)$ denote the class of continuous functions on T uniformly bounded in absolute value by 1. We are interested in stochastic quadrature rules, say ϑ_ω for approximating $\int_T f(t)d\mu(t)$. If we agree to measure the error of such a Monte Carlo quadrature formula in the mean square, then the following result is true.

The following is equivalent

- The measure μ is non atomic.
- We have

$$\inf_{\text{card}(\vartheta) \leq n} \sup_{f \in B_\infty(T)} \left(\mathbf{E} \left| \int f d\mu - \vartheta(f) \right|^2 \right)^{1/2} = \frac{1}{1 + \sqrt{n}}.$$

Moreover, the error will not decrease neither by admitting nonlinear quadrature rules nor by varying cardinality.

We demonstrate the close relation of the integration problem to a problem in the theory of point estimation, which is then proved explicitly.

A non-completeness result for quadratic programming

Klaus Meer
RWTH Aachen

The complexity of linearly constrained (non-convex) Quadratic Programming is analyzed within the framework of real number models, namely the one of Blum, Shub and Smale and its modification recently introduced by Koiran (“weak BSS-model”). In particular we show that this problem is not NP -complete in the Koiran setting. Applications to the (full) BSS-model are discussed.

Approximating the Path of a Stochastic Process Using Derivatives

Thomas Müller–Gronbach
FU Berlin

We consider a stochastic process $Y = \{Y(t) : 0 \leq t \leq 1\}$ given by

$$Y(t) = \int_0^t X(s) \frac{(t-s)^{k-1}}{(k-1)!} ds$$

where X is a centered stochastic process which satisfies the Sacks–Ylvisaker regularity conditions of order zero. For a design $d = (t_1, \dots, t_n)$ of observation sites in $[0, 1]$ each $Y(t)$ is approximated by its orthogonal projection $\text{opr}_{k,d} Y(t)$ onto the subspace spanned by $Y^{(j)}(t_i)$, $1 \leq i \leq n$, $0 \leq j \leq k$, where $Y^{(j)}$ denotes the j -th mean square derivative of Y . The quality of this approximation is measured by the corresponding integrated mean squared error

$$\int_0^1 E (Y(t) - \text{opr}_{k,d} Y(t))^2 dt.$$

We construct sequences of designs which perform asymptotically optimal with respect to this criterion. Furthermore we study the problem of approximating Y without using derivatives in the particular case that $Y(t) = \int_0^t B(s) ds$, where B denotes a Brownian motion on $[0, 1]$.

A Combinatorial Relaxation Algorithm for Finding the Maximum Degree of Subdeterminants

Kazuo Murota
Research Institute for Mathematical Sciences, Kyoto University

Let $A(x) = (A_{ij}(x))$ be a matrix with $A_{ij}(x)$ being a polynomial or rational function in x . We consider various algorithms for computing the highest degree $\delta_k(A)$ of a minor of $A(x)$ of a specified order k . Such algorithms can be used to compute the Smith-McMillan form of a rational function matrix at infinity, as well as the structure of the Kronecker form of a matrix pencil.

Among the algorithms we put emphasis on a “combinatorial relaxation” type algorithm that is based on a combinatorial upper bound $\hat{\delta}_k(A)$ on $\delta_k(A)$; $\hat{\delta}_k(A)$ is defined as the maximum weight of a matching of size k in a bipartite graph associated with A . The algorithm is efficient, making full use of the fast algorithms for weighted

matchings. It is combinatorial in almost all cases (or generically) and invokes algebraic elimination routines only when accidental numerical cancellations occur. The validity relies on the integrality of bipartite matching polytopes.

Particle Methods for the Boltzmann Equation: Mathematical and Practical Aspects

Helmut Neunzert
Universität Kaiserslautern

The talk describes a particle ("low discrepancy" or "Quasi Monte Carlo" or "Finite Pointset") method for solving kinetic equation; these equations are of practical relevance in space research, semiconductor design and vacuum technology.

The method uses low discrepancy methods for typically 6–10 dimensional functions, but the main problem is the time evolution of the particle sets. In simulating collisions one has to solve optimization problems for special integration tasks, some stochastic elements are still used in these processes and the question posed is, whether information-based complexity can give some information about whether this is advantageous or not.

The Real Number Model in Numerical Analysis

Erich Novak
Erlangen (Augsburg)

We assume that arithmetic operations with real numbers and comparisons can be done with unit cost, also copy instructions (with indirect addressing) are allowed. First we compare our model with the BSS model and show that the sets P and NP coincide with respect to both cost functions (and construction sets). In numerical analysis and information-based complexity we study problems where the information is only partial. We describe this model with oracle more carefully. We also define two variants of stochastic or randomized algorithms.

As an example we present a uniform problem: construct one algorithm to compute the integral of $f \in C^k([0, 1])$ up to ε , if a bound on $\|f^{(k)}\|_\infty$ is known. Here ε and k are inputs of the algorithm. A complexity bound can be proved using compound Clenshaw–Curtis formulae.

Reduction Theory for Statistical Ill-posed Problems

M. Nussbaum

Weierstrass Institute, Berlin

We consider an ill-posed problem of recovery of a function f from *noisy information* $y = Af + \epsilon \xi$ where $\epsilon \xi$ is noise and ϵ is assumed small. Here A is a compact linear operator between Banach spaces of functions on $[0, 1]$, and there is additional a priori information: $f \in \Sigma$ (regularization). For algorithms we consider the worst case setting with respect to f ; the error criterion depends on the character of the noise. In the first case ξ is assumed deterministic and bounded, this is known as "optimal recovery". Secondly we consider random noise, especially Gaussian white noise (GWN) which is a random generalized function. In the third class of problems, independent identically distributed random variables Y_1, \dots, Y_n are observed having a probability density $g = Af$ on the unit interval (statistical ill-posed problem or density model). This does not have an additive noise structure, but heuristically it is similar to the Gaussian white noise model on the basis of a central limit theorem. The statistical literature extensively treats complexity issues for the GWN and density models, demonstrating their analogy case by case. We give a result which establishes an *asymptotic complexity equivalence* of both models. It is based on Le Cam's theory of statistical experiments, utilizing the concept of deficiency distance. Having thus performed a reduction of the density model to the GWN model, we discuss the relationship of the latter to optimal recovery (deterministic noise) in the context of pertaining recent results of Donoho et al.. We remark that the established universal rate optimality of wavelet thresholding for GWN can be reproduced for bounded noise in function space.

Optimal recovery of analytic functions from their Fourier coefficients given with an error

K. Yu. Osipenko

Moscow State University of Aviation Technology

Let X be a Hilbert space and $\{e_j\}$ a complete orthonormal system in X . We consider the problem of optimal recovery of the linear functional (x, f) , $f \in H$, from approximate values of Fourier coefficients $x_j = (x, e_j)$.

Put

$$e_n(f, \delta) := \inf_{\varphi: C^n \rightarrow C} \sup_{\substack{x \in H \\ \|h\| \leq 1}} \sup_{\substack{\tilde{x}_j, j=1, \dots, n \\ |\tilde{x}_j - x_j| \leq \delta_j}} |(x, f) - \varphi(\tilde{x}_1, \dots, \tilde{x}_n)| .$$

Set

$$a_+ := \begin{cases} a & , a \geq 0 \\ 0 & , a < 0. \end{cases}$$

Theorem: The method

$$(x, f) \approx \sum_{j=1}^n \left(1 - \lambda \delta_j |f_j|^{-1}\right)_+ \bar{f}_j \widetilde{x}_j$$

is an optimal method of recovery and

$$e_n(f, \delta) = \lambda + \sum_{j=1}^n \delta_j (|f_j| - \lambda \delta_j)_+ ,$$

where $\lambda \in (0, \|f\|]$ is a solution of the equation

$$\|f\|^2 - \sum_{j=1}^n \left(|f_j|^2 - \lambda^2 \delta_j^2\right)_+ - \lambda^2 = 0 .$$

Using this Theorem we construct optimal methods of recovery of 2π -periodic and analytic in a strip functions and its derivatives from approximate values of their Fourier coefficients.

Some estimates of complexity of ill-posed problems

Sergei Pereverzev
Universität Kaiserslautern (Kiev)

Using the idea of telescopic system of rectangles we propose a modification of the projection scheme for solving ill-posed problems. We show that this modification allows to obtain the best possible order of accuracy of Tikhonov regularization using an amount of information which is far less than for the standard projection technique.

Numerical Integration of Univariate and Multivariate Convex Functions

Knut Petras

Technische Universität Braunschweig
(joint work with Carsten Katscher and Erich Novak)

We consider worst-case error bounds for different types of n -point quadrature formulae applied to convex functions. The most investigated case is that of linear deterministic methods in the univariate case. We argue that it is advantageous to define classes of convex functions without using boundary derivatives, and that it is useful to apply Gaussian instead of optimal formulae.

Then, we compare the different types of optimal formulae with respect to worst-case error bounds. In the univariate case, we obtain the orders

- linear deterministic methods: $\asymp n^{-2}$
- adaptive deterministic methods: $\asymp n^{-2}$
- stochastic non-adaptive methods: $\asymp n^{-2}$
- stochastic adaptive methods: $\asymp n^{-5/2}$.

For convex functions on $[0, 1]^d$, the optimal worst-case orders are

- linear deterministic methods: $\asymp n^{-2/d}$
- adaptive deterministic methods: $\asymp n^{-2/d}$
- stochastic non-adaptive methods: $\leq \text{const} \cdot n^{-1/2-3/(2d)}$
- stochastic adaptive methods: $\asymp n^{-1/2-2/d}$.

Approximation of functionals from noisy information

Leszek Plaskota

University of Warsaw, Poland

We study approximation of linear functionals when information is corrupted by some noise. We consider four different settings which are determined by assumptions put on problem elements and noise: worst case, average case, and mixed settings. It turns out that in all four settings optimal affine (or linear) algorithms belong to the same class of algorithms, and that it is impossible to essentially reduce the error using nonaffine (nonlinear) algorithms. Moreover, the minimal errors are also comparable.

These results do not hold any longer when instead of approximating a functional we approximate the whole object.

Differentiation of Random Functions using Noisy Data

Klaus Ritter
Erlangen

We study differentiation of random functions $f : [0, 1] \rightarrow \mathbb{R}$ based on noisy data $f(t_i) + \varepsilon_i$. The noise is assumed to be uncorrelated with zero mean and common variance $\sigma^2 > 0$; the smoothness of f in mean square sense is specified by $H(R) = W_2^{r+1}([0, 1])$. Here $H(R)$ denotes the reproducing kernel Hilbert space associated with the covariance kernel R of f . We study linear estimates $S_n(f, \varepsilon)$, and their error is defined as

$$\left(E(\|f^{(k)} - S_n(f, \varepsilon)\|_{L_2}^2) \right)^{\frac{1}{2}} .$$

The following orders are obtained for the minimal errors of estimates using n samples:

$$\begin{aligned} & n^{-(r-k+\frac{1}{2})} \quad \text{if } \sigma^2 \leq n^{-(2r+1)} \\ \left(\frac{\sigma^2}{n} \right)^{\frac{1}{2} - \frac{k+1/2}{2r+2}} & \quad \text{if } n^{-(2r+1)} \leq \sigma^2 \leq n \\ & 1 \quad \text{if } \sigma^2 \geq n \end{aligned}$$

The upper bounds are achieved by smoothing splines using equidistant knots.

On Sard's Theorem for Nonsmooth Functions

Andrea Rohde
RWTH Aachen

Complexities of certain classes of nonsmooth functions and concepts of generalized derivatives are compared from a metric point of view. Here the density of generalized critical values serves as a measure of complexity. We study continuous selections either of a finite number of differentiable functions g_1, \dots, g_s (with $g_i \in C^k([0, 1]^n, \mathbb{R}^m)$, $m \in \{1, n\}$) or componentwise of maximum functions of the form

$$g_i(x) = \max_{y \in [0, 1]^q} h_i(x, y), \quad h_i \in C^k([0, 1]^{n+q}, \mathbb{R}) .$$

The entropy dimensions of the set of generalized critical values can be bounded in terms of u, m, q and k only. In particular, it turns out that the continuous selection does not increase the complexity compared with Sard's Theorem for C^k -mappings. For vectorvalued functions $f : [0, 1]^n \rightarrow \mathbb{R}^n$ a definition of generalized

critical points/values based on a complete characterization of local Lipschitz invertibility by Kummer/Thibault yields much better bounds than a definition based on Clark's generalized Jacobians.

A Quasi-Monte Carlo approach to Efficient 3-D Seismic Image Reconstruction

Krzysztof Sikorski

Computer Science Dep., University of Utah

(joint work with G.Schuster, Y.Sun, Geophysics Dep., Utah)

We study the feasibility of applying quasi Monte Carlo methods to seismic imaging by 3-D pre-stack Kirchhoff migration. This earth imaging technique involves computing a large ($\geq 10^9$) number of 3- or 4-dimensional integrals. Our numerical studies show that nearly optimal quasi-Monte Carlo migration can produce the same or better quality earth images using only a small fraction (one fifth or less of the data required by a conventional Kirchhoff migration).

Lattice Methods for Multiple Integration

Ian H.Sloan

University of New South Wales, Australia

Lattice methods are a generalization of the method of good lattice points, introduced in the late 1950s by Bakhvalov, Korobov, Hlawka, and others. This talk will review lattice methods, beginning with an introduction to the now accepted classification of lattice rules according to 'rank'. The traditional rules of the method of good lattice point have rank 1, but all ranks up to the dimensionality of the integrand are possible. Particular attention will be paid to the rapid convergence rates (and correspondingly low costs) that can prevail when the integrands are required to be periodic functions with some smoothness. It is shown that lattice rules with maximal rank can be competitive with rules of rank 1 with respect to order, and have many practical advantages, including a lower pre-computational cost of finding good rules, and a low-cost error estimate.

Infeasible-interior-point algorithms for solving geometrical linear complementarity problems

J.Stoer

Würzburg

(joint work with S. Misuko, Tokyo and F. Jarre, Würzburg)

There exist many interior–point algorithms for linear programming (LP), convex quadratic programming (QP), and more generally for linear complementarity problems (LCP) satisfying a monotonicity condition. These LCP’s can be formulated geometrically in a unified way as problems of finding a point $z = (x, y) \geq 0$ in a n -dimensional linear submanifold of k^{2n} with $Xy = 0$, $X := \text{diag}(x)$. Some basic properties are derived for these geometrical LCP’s, which in particular permit to define and analyze an infeasible–interior–point algorithm for solving them. It features global convergence, polynomial time convergence (if there exists a solution that is “small” than the initial point), and quadratic convergence, if there exists a strictly complementary solution.

Nonlinear approximation

Vladimir Temlyakov

University of South Carolina

This talk is based on two joint papers, the first with R.DeVore and the second with B.Kashin. We consider m -term trigonometric approximation in the multivariate case

$$\sigma_m(f, T)_p = \inf_{K^j, c_{K^j}} \left\| f(x) - \sum_{j=1}^m c_{K^j} e^{i(K^j, x)} \right\|_p.$$

We find the exact order of $\sigma_m(F, T)_p = \sup_{f \in F} \sigma_m(f, T)_p$ for Besov classes $B_s^r(L_q)$ for all $0 < s, q \leq \infty$, $1 \leq p \leq \infty$. We created new methods to obtain the upper estimate for $p = \infty$ and the lower estimate for $q = \infty$, $p = 1$. Both these methods are based on deep results from finite geometry.

Harmonics and Splines on Groups and Homogeneous Manifolds Some Problems of Approximation and Recovery

V. M. Tikhomirov
Moscow State University

In the lecture definitions of generalized spaces of harmonics ($\mathcal{B}_A(G), \mathcal{B}_\sigma(M)$) and splines ($S_H^\varphi(G)$) were introduced. The following results were discussed:

1. Generalized Whittacker–Sharmon–Kotelnikov formula for $x(\cdot) \in \mathcal{B}_A(G)$ (G – locally compact abelian group (LCAG)), and generalized Shoenberg formula for splines $S_H^\varphi(G)$ (G – LCAG, H – closed discrete subgroup).
2. Optimality and asymptotical optimality of the spaces $\mathcal{B}_A(G)$ in approximation of generalized Sobolev classes $W_p^k(G)$, $p = 1, 2$ in $L_2(G)$ and $W_p^k(G)$ in $L_q(G)$.
3. Generalized Riesz interpolation formula for spaces $\mathcal{B}_\sigma(M)$ (M – homogeneous manifold of range 1).
4. Approximation of harmonics by spline ($S_H^{\varphi_n}(G) \rightarrow B_H(G)$, $\varphi_n = \varphi * \dots * \varphi$).
5. Generalization of Ligun Theorem

$$K \in ETP(\mathbb{T}) \Rightarrow d_n(W_\infty^k(\mathbb{T}), C(\mathbb{T})) = d(W_\infty^k(\mathbb{T}), S_{Z_{2n}}^{K_m}, C(\mathbb{T})), m \geq 1.$$

From Finance to Physics

Joseph F. Traub
Columbia University

Two topics are presented. The first is

High Dimensional Integration With Application to Finance

Recent theoretical results on the average case complexity of high dimensional integration are presented. Theory suggests that low discrepancy methods provide a powerful alternative to Monte Carlo. However, a number of researchers who have tested low discrepancy methods report that this theoretical advantage is not seen in testing.

The valuation of a financial derivative given to us by Goldman/Sache is used as a test example. It is a CMO whose valuation leads to an integral in 360 dimensions. This testing has been carried out by Spassimir Paskov. The conclusions for this

CMO is that the Sobol algorithm is superior to Halton, Monte Carlo, and antithetic variables algorithms. The tests are run on a network of workstations.

The second topic is

Are There Proveable Limits to Scientific Knowledge?

Some of the most influential results of mathematics of the last 60 years concern limitations of mathematics and mathematically posed problems. These include incompleteness, non-computability, undecidability, and intractability. Analogously, can limitations to scientific knowledge be made part of science?

Explicit Cost Bounds of Algorithms for Multivariate Tensor Product Problems

Grzegorz W. Wasilkowski
University of Kentucky

(joint work with Henryk Woźniakowski, Columbia University)

We discuss multivariate tensor product problems in the worst case and the average case settings. They are defined as functions of d variables. For arbitrary d , we provide explicit upper bounds on the costs of algorithms which compute an ε -approximation to the solution. The cost bounds are of the form

$$(c(d) + 2)\beta_1 \left(\beta_2 + \beta_3 \frac{\ln 1/\varepsilon}{d-1} \right)^{\beta_4(d-1)} \left(\frac{1}{\varepsilon} \right)^{\beta_5}.$$

Here $c(d)$ is the cost of one function evaluation (or one linear functional evaluation), and β_i 's do not depend on d ; they are determined by the properties of the problem for $d = 1$. For certain tensor product problems, these cost bounds do not exceed $c(d) K \varepsilon^{-p}$ for some numbers K and p , both independent of d .

We apply these general estimates to certain integration and approximation problems in the worst case and the average case settings.

We also obtain an upper bound, which is independent of d , for the number $n(\varepsilon, d)$ of points for which discrepancy (with unequal weights) is at most ε ,

$$n(\varepsilon, d) \leq 7.26\varepsilon^{-2.454}, \quad \forall d, \varepsilon \leq 1.$$

Computation of Path Integrals

Henryk Woźniakowski

Columbia University & University of Warsaw

(joint work with G.W.Wasilkowski, University of Kentucky)

We consider an approximate computation of infinite dimensional integrals with respect to a Gaussian measure and for a class F of integrands.

We consider two classes of F . The first one is defined by integrands with smoothness r . If r is finite then the path integral problem is intractable. If $r = +\infty$ and F consists of entire functions then we have tractability with exponent depending on the Gaussian measure.

The second class F is given by the Feynman–Kac formula for solving heat equations with a potential function v . We prove that if v is r -times differentiable and has a finite support then the complexity of the path integral in the worst case setting is of order $\varepsilon^{-1/r}$ (roughly). The upper bound is constructive and is achieved by a repetitive use of Smolyaks algorithm for tensor product problems.