92. Band Heft 2
ausgegeben am 8. 5. 1990

DMV

Jahresbericht
der Deutschen Mathematiker-Vereinigung

Herausgegeben von W.-D. Geyer
unter Mitwirkung von
P. L. Butzer, U. Felgner,
K.-H. Hoffmann, H. Kurzweil

B. G. Teubner Stuttgart 1990

Jahresbericht
der Deutschen Mathematiker-Vereinigung

Der „Jahresbericht" ist das offizielle Veröffentlichungsorgan der Deutschen Mathematiker-Vereinigung, für dessen inhaltliche Gestaltung im Auftrag des Präsidiums der jeweilige Herausgeber zuständig ist. Im „Jahresbericht" sollen vornehmlich Überblicksartikel über Teilgebiete der reinen und angewandten Mathematik, Nachrufe sowie historische Artikel und Buchbesprechungen veröffentlicht werden.

Manuskripte

Alle für die Schriftleitung des Jahresberichts bestimmten Sendungen (Briefe, Manuskripte, Ankündigung von neuerscheinenden Büchern) sind an Prof. Dr. W.-D. Geyer zu richten. Bücher, von denen eine Besprechung erfolgen soll, werden bei den Verlagen angefordert. Die Autoren werden gebeten, bei der Vorbereitung ihrer Manuskripte die „Hinweise für Autoren" zu beachten.

Grundsätzlich sollen nur solche Manuskripte eingereicht werden, die nicht gleichzeitig an anderer Stelle zur Veröffentlichung eingereicht oder bereits veröffentlicht worden sind. Mit der Annahme zur Veröffentlichung erwirbt der Verlag das Verlagsrecht zur Vervielfältigung und Verbreitung sowie das Recht der Übersetzung in andere Sprachen.

Bezugsinformationen

Jährlich wird ein Band veröffentlicht, bestehend aus 4 Heften, die vierteljährlich erscheinen. Der Bezug ist nur bandweise möglich.

Der im voraus zahlbare Bezugspreis pro Band beträgt DM 118,– einschließlich Versand. Bestellungen nehmen jede Buchhandlung oder der Verlag entgegen.

Für Mitglieder der DMV ist der Bezug des Jahresberichts im Mitgliedsbeitrag enthalten.

Verlag:
B. G. Teubner, Industriestr. 15, Postfach 801069
D-7000 Stuttgart 80, Tel. (0711) 78901-0
Verantwortlich für den Anzeigenteil: Berthold Gaupp

Printed in Germany – ISSN 0012-0456
Satz: Elsner & Behrens GmbH, D-6836 Oftersheim
Druck: Krebs-Gehlen Druckerei GmbH & Co KG, D-6944 Hemsbach

Jahresbericht
der Deutschen Mathematiker-Vereinigung

Inhalt Band 92, Heft 2

1. Abteilung

2. Abteilung

In den nächsten Heften erscheinende Arbeiten:

R. Frankl, R. L. Graham, V. Rödl: Quantitative Versionen von kombinatorischen Sätzen
L.-Ch. Kappe, H.-P. Schlickewei, W. Schwarz: Theodor Schneider zum Gedächtnis
P. Ullrich: Wie man beim Weierstraßschen Aufbau der Funktionentheorie das Cauchysche Integral vermeidet

---

Bezugshinweis

Früher erschienene Bände (ab Band 68) des „Jahresberichts der Deutschen Mathematiker-Vereinigung" können durch den Buchhandel oder den Verlag bezogen weren.

Nachdruck der Bände 1 bis 40 liefert: Johnson Reprint Corp., 111 Fifth Avenue, New York, N.Y. 10003

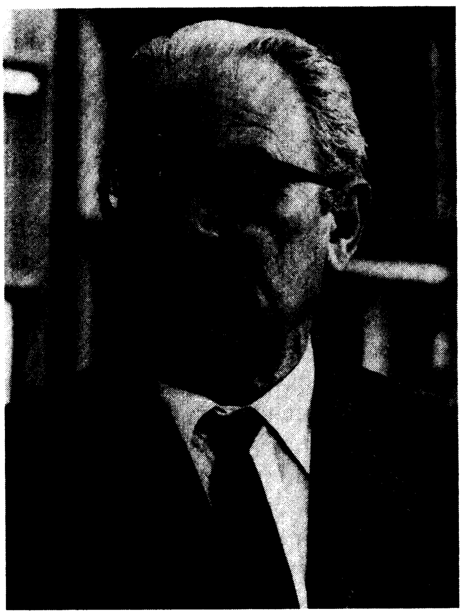
Nachdruck der Bände 41 bis 67 liefert: Swets & Zeitlinger, Heereweg 347b, POB 810, NL-2160 SZ Lisse/Holland

Jber. d. Dt. Math.-Verein.
92 (1990) 47–57

Eberhard Hopf

04–17–1902 to 07–24–1983

M. Denker, Göttingen



Eberhard Hopf died on July 24, 1983 in Bloomington, Indiana, following a long illness preventing him from any scientific activity for more than a decade.

He was born in Salzburg on April 17, 1902. When his father was appointed civil servant in Berlin, he attended the Gymnasium in Berlin-Friedenau and, after finishing school in 1920 he studied at the universities of Berlin and Tübingen. In 1926 (February 9; the examination took place in July 1925) he received his doctorate with a thesis on real analysis [1] written under the supervision of E. Schmidt. The second examiner was I. Schur.

During the years 1926–1930 he started his academic career as an assistant at the Astronomische Rechenzentrum of the University of Berlin and after his habilitation in 1929 he became Privatdozent at this institution. Eberhard Hopf visited the United States for the first time in 1930 as an International Fellow of the Rockefeller Foundation. He worked until 1932 at the Harvard Astronomical Observatory when he was appointed Assistant Professor of mathematics at the Massachusetts Institute of Technology. In 1928 he married Ilse Wolf and their only daugther was born shortly before Eberhard Hopf returned to Leipzig in 1936 to succeed Lichtenstein. In 1942 he was drafted to serve at the Luftfahrtforschungsinstitut in Einring near Munich. Subsequently he accepted a professorship at the University of Munich (the chair of Carathéodory) in 1944.

By special arrangements (R. Courant) he was brought to the United States in 1947 to serve as a visiting professor at the Courant Institute ("imported ... as a paper clip scientist for the U.S. Army" he said) and stayed in the United States "because in the U.S.A., there's more time for research". In 1948 he took a teaching assignment with the U.S. Navy and then joined the Institute of Mathematics at Indiana University as professor, in order to collaborate with the physicist Truesdell. He became research professor in 1962 and retired in 1972.

Eberhard Hopf was a member of the Deutsche Mathematiker-Vereinigung, the American Mathematical Society and elected member of the Bayerische and Sächsische Akademie der Wissenschaften. He was co-founder of the journal of Mathematics & Mechanics (now Indiana University Mathematics Journal), and served for many years as a co-editor of this journal and of the Grundlehren der Mathematischen Wissenschaften.

Eberhard Hopf received many honors during his career. He gave the Josiah Willard Gibbs Lecture in 1971 ([79]) and invited addresses at numerous International Congresses and Symposia. In 1981 he received the Steele Prize for work of "fundamental or lasting importance" ([46], [53] and [56]). At that time he was no longer able to accept the award in person.

Hopf's contributions to mathematics have been both significant to the growth of knowledge and the direction of contemporary research. Many of his mathematical discoveries are now in the form of basic principles and their importance will probably remain undiminished far into the future. His name is attached to several theorems from different branches of pure and applied mathematics, e.g.: The Hopf Maximum Principle in the theory of elliptic differential equations, the Wiener-Hopf technique in the theory of integral equations ([17]), the Hopf bifurcation in the theory of differential equations and dynamical systems ([46]), the Hopf decomposition and Hopf's maximal ergodic theorem in ergodic theory. Besides this, his contributions influenced research in the calculus of variations, hydrodynamics, theoretical astrophysics, Navier-Stokes equations, differentiable dynamics and probability.

Eberhard Hopf started his scientific career in the field of differential equations, in particular of elliptic partial differential equations. Many of his early results have become classical and are well represented in modern textbooks on this topic. His interest in hydrodynamics and turbulence, to be discussed below, is based on his deep understanding of differential equations. His early work (during the years 1927–1932) on the maximum principle and the regularity and analyticity of solutions of elliptic partial differential equations $\mathscr{M}u = f$ had great impact on many of his successors in the area ([5], [7], [14], [18], [24]). The papers [5] and [7] contain a proof of the maximum principle for elliptic differential equations, while [14] provides a procedure to establish a priori majorization formulas for the solutions of Dirichlet's problem and it was also used to prove new existence theorems. In particular, he proved Hölder continuity of the second derivatives of the solutions of elliptic equations in the region of definition provided the coefficients are Hölder continuous themselves, with the same Hölder exponents. Finally, in [24], he showed, among other things, that every solution of the elliptic equation $F(p_{ik}, p_i, u, x) = 0$ of class $C^{m,\lambda}$ is analytic provided F is.

In later years, Hopf published four smaller notes on elliptic differential equations and related equations for minimal surfaces ([59], [69], [61], [63]), where he deals with similar types of problems.

At the time of his employment as an assistant in Berlin, Eberhard Hopf started his investigation of integral equations.*) During the years from 1927 to 1937, he made important contributions to the study of integral equations, particularly those which pertain to radiative transfer in stellar atmospheres. This work is summarized in his Cambridge Tract, Mathematical Problems of Radiative Equilibrium [32]. In the introduction, Hopf wrote: "The astrophysicists mostly contented themselves with an approximate solution of these problems. Owing to a certain inherent beauty, however, they aroused also as the interest of a rigorous mathematician. It is the purpose of this tract to attempt a coherent representation of all that has been achieved in the direction of a rigorous treatment of these standard problems". Hopf went far beyond this rather modest goal. He not only gave a rigorous treatment of radiative transfer problems, with many new results, but he also extended his analysis to broad classes of integral equations with similar features.

As a prototype, the Milne problem for the intensity of radiation in an idealized, stratified atmosphere, with infinite optical depth an isotropic scattering, has the form

$$f(s) = \int_0^\infty k(s-t) f(t)\, \mathrm{d}t \qquad (s \geqslant 0),$$

with $k(u) = E_1(|u|)$, the exponential integral function. In 1928, Hopf ([8], [9]) proved that the Milne problem has a unique normalized positive solution, which is asymptotically linear. In [10], he extended his result to a general class of integral equations with positive kernels. Hopf's analysis is noteworthy. It represents an early application of positive operator techniques to integral equations.

The Milne problem also led to the Wiener-Hopf technique, presented in [17], for using Fourier transforms to solve integral equations on the half-line with convolution kernels. In [32], pg. 103, Hopf wrote: "... this method, though solving [the problem] explicitly, has but a limited applicability, as concerns many other problems of radiative equilibrium. The comparison methods, resting on the positivity of the kernel, furnish a simpler and more general access to a physically

---

*) The following appreciation of Hopf's contributions to this field was communicated to me by P. M. Anselone and this passage is due to him.

satisfactory (though not explicit) solution." Apparently, Hopf did not forsee the far-reaching impact that the Wiener-Hopf method and its extensions would have on future developments in analysis and applications to many fields, such as mechanics and electromagnetic theory.

After many years, Hopf returned in 1963 to the field of integral equations. In [74], [75] he derived the inequality,

$$|\lambda| \leqslant \frac{\varkappa - 1}{\varkappa + 1} \lambda_0,$$

for the secondary eigenvalues λ of a positive integral operator K with the cross-ratio property

$$\frac{K\phi(x')}{K\phi(x)} \Big/ \frac{K\psi(x')}{K\psi(x)} \leqslant \varkappa^2, \quad (\phi \geqslant 0, \psi \geqslant 0).$$

In this analysis, as throughout his research, Eberhard Hopf displayed a keen grasp of essentials, command of technique, and elegance of presentation.

While still at the beginning of his mathematical carreer Hopf took up another subject, that of ergodic theory, smooth dynamical systems and questions connected with the global behavior of solutions of differential equations. Undoubtedly, two of his contributions are of lasting influence in ergodic theory. The first one is [65] where he introduced the ergodic theory for L^1-contractions, especially for Markov operators. Secondly, in a series of papers ([37], [40], [42] and [43]) he proved ergodicity for the geodesic flow on manifolds of negative curvature. The lecture notes [72] give an excellent survey on both topics.

As early as 1930 ([16]) Hopf developed his interest in ergodic theory, a subject which came to life two years later with Birkhoff's and v. Neumann's theorems, and which was greatly influenced by Hopf's ideas and inspiration. His first paper contains a strengthening of Poincaré's recurrence theorem (of 1890 and 1899) and an extension of Carathéodory's proof of these. He shows that for a measure-preserving flow, either a.e. point is a cluster point of its forward orbit, or its forward orbit has no cluster point at all. In the case of a finite measure only the first possibility can occur (Poincaré's theorem). This is the first form of the celebrated "Hopf decomposition". The monograph [39] contains this decomposition into conservative and dissipative parts explicitly. For null-preserving transformations T (i.e. Tm is absolutely continuous with respect to m) this decomposition can be found in [65].

Koopmann's association of a unitary group to a flow and the ergodic theorems of v. Neumann and Birkhoff certainly inspired Hopf to work on the foundations of this theory. His interest has its origin from physics, classical Hamiltonian dynamical systems and the ergodic hypothesis in the Maxwell-Boltzmann gas theory and Gibbs statistical mechanics. This motivation can be found in many of his papers. He certainly noticed the importance of the ergodic theoretic and functional theoretic approach for an asymptotic global understanding of dynamical systems. Especially he expected that these methods would help to solve the problem of turbulence on the basis of classical hydrodynamics, see e.g. [53].

Some of Hopf's early theorems in this field can be found in every introductory book on ergodic theory. v. Neumann showed that complete transitivity is equivalent to ergodicity. In [29], [30] and [33] Hopf found some of the properties which are equivalent to weak mixing, for example the equivalent condition that the flow has a simple spectrum with multiplicity 1. (In fact, this last result was discovered independently by him and Koopmann/v. Neumann.) His proofs are now classical, as it was one of Hopf's great abilities to give most elegant and natural proofs. The proofs of the classical ergodic theorems in [28] show this also, and they may have influenced Kakutani and Yoshida to obtain in 1939 the maximal ergodic theorem for point transformations (see also [47]). Later, in [65], he extended this theorem to positive L^1-contractions T: If $E_n = \left\{ \max_{1 \leqslant i < n} \sum_{j=0}^{i-1} T^j f > 0 \right\}$ and $E_\infty = \bigcup_{n \geqslant 1} E_n$, then for all $n = 1, 2, \ldots, \infty$

$$\int_{E_n} f \, \mathrm{d}m \geqslant 0.$$

An extension of the statistical ergodic theorem appears in [25]: Every group of unitary operators (U_t) admits a decomposition into an almost periodic part (V_t) such that the ergodic averages of $(U_t - V_t)$ converge to zero. Another ergodic theorem is obtained in [33] (see also [41]), but the best known ergodic theorem of Hopf appears in [65]. This paper is a breakthrough in the development, since it opened a wide field of research with far reaching applications: Let T be a positive contraction on $L^1(m)$ such that $T1 = 1$ and such that m is a finite measure. Then for every integrable function f, $\frac{1}{n} \sum_{j=0}^{n-1} T^j f$ converges a.e. The condition $T1 = 1$ was removed later by Dunford and Schwarz in 1956. In the same paper Hopf conjectured the ratio ergodic theorem, that $\sum_{j=0}^{n} T^j f \Big/ \sum_{j=0}^{n} T^j p$ converges a.e. on $\left\{ \sum_{j=0}^{\infty} T^j p > 0 \right\}$ for integrable functions f and nonnegative p. This theorem was obtained by himself for point transformations much earlier ([39, p. 49]). Chacon and Ornstein obtained the general result, an elegant proof has been given by Hopf in [71] (see also [72]).

Ergodic theorems may be considered as almost sure convergence results in probability theory, and in fact, they are extensions of Kolmogorov's strong law of large numbers for independent, identically distributed random variables. Moreover, a Markov kernel may be regarded as an operator acting on some domain of measurable functions. [65] gives an excellent introduction into the subject and the relation between Markov operators and Markov processes. Earlier he showed his interest in probability in [33], [34], [36] and [41], where he presented a modern treatment of the foundations of measure and probability theory. He was aware that deterministic dynamical systems can created randomness.

The existence of equivalent invariant measures is one of the important problems in dynamical systems. In [31], Hopf extended a theorem of Birkhoff and

Smith, which is now contained in the Hajian-Kakutani theorem: If a transformation is null-preserving, then there exists an equivalent finite invariant measure if and only if there is no weakly wandering set.

It remains to discuss Hopf's contribution to ergodic theory in connection with smooth dynamical systems. In [36] there is a remark concerning the billiard problem, one of the most stimulating problems in recent years. Inspired by a work of Artin (1924), he proved ergodicity for a free motion in a certain geodesic triangle on the surface of revolution of a tractrix. His main contribution remains the work on the ergodicity of the geodesic flow on manifolds of negative curvature. His work influenced the Russian school around Anosov and Sinai in the sixties. The topic became of much wider interest especially recently; Hopf's ideas, however, remain valuable.

The well known correspondance of Riemannian surfaces S and Fuchsian groups G enables to study ergodicity of the geodesic flow on S by the action of G as Moebius transformations on the Poincaré model of the hyperbolic plane. The case of constant negative curvature was treated first in [37]. If the volume of S is finite, it is of the first kind (i.e. no arc of the fundamental domain is contained in S^1) and Hopf showed ergodicity of the geodesic flow (see also the monograph [39]). In the other case, almost every geodesic disappears in a funnel. (In the finite volume case Artin had proven topological transitivity, Myrberg extended this result and Hedlund in 1934/35 proved ergodicity in special cases.) The proof uses potential theoretic methods (harmonic functions). In [40], Hopf even obtained weak mixing.

Hedlund continued in 1939 Hopf's investigations (finite volume case) and showed weak mixing by introducing the hyrocycle flow (see Hopf's discussion in Zentralblatt der Mathematik **20** (1939) 403). The extension to n-dimensional Riemannian manifolds S of constant negative curvature and possibly infinite volume appears in his remarkable work [42] (cf. A. Weil's remark in [W], p. 368), where he proved that either the geodesic flow is conservative ergodic or a.e. geodesic tends to infinity: S is called of the first class if the set of points whose geodesic tends to infinity has measure zero. Otherwise S is called of the second class. If the associated discontinuous group is of the second kind, S is of the second class. In the other case both classes of S can occur. Hopf showed that the associated geodesic flow is necessarily dissipative if S is of the second class. For the first class the geodesic flow is ergodic and even weak mixing in the finite volume case. He also proved these results in the case of variable negative curvature in dimension 2. Finally, in [43], the general result (non-constant negative curvature) is achieved by introducing asymptotic geodesics (geodesics having the same endpoint at infinity). This furnished the basic tool for investigations and may be seen as one of the origins of the foliation theory in differentiable dynamics.

As much as his research, the book [39] on ergodic theory influenced a whole generation of mathematicians in the forties and fifties and it is still a symbol of Hopf's lasting importance to the field.

His interest in differential geometry of surfaces apparently developed out of his investigations of the geodesic flow on manifolds of negative curvature. A few contributions arose from it, devoted to different topics. In [52] he shows that on compact surfaces S of class C^3 the (total) curvature is non-positive if no geodesic on S has on it two mutually conjugate points. If the curvature is zero, the Gaussian curvature must vanish identically. The Hopf conjecture that the n-dimensional torus without conjugate points has curvature zero has its origin in this work, but it certainly was never conjectured by himself.

In 1943, Hopf published his famous article [46] on the bifurcation for stationary, time-periodic solutions of a system of differential equations. It established the foundations of the Hopf bifurcation theory. His work was extended to partial differential (and abstract) equations and has been applied to problems in fluid mechanics, solid mechanics and chemical reactor theory. He considered the first order equation $\dot{x}=F(x,\mu)$ where F is analytic in $x=x(t)$ and the real parameter μ, and he showed the existence of a family of real periodic solutions $x(t,\varepsilon)$, $\mu=\mu(\varepsilon)$ which coincide with $x(0)$ for $\varepsilon=0$ but not for $\varepsilon\neq 0$.

It remains to discuss Eberhard Hopf's contributions to the theory of hydrodynamics, to which he devoted a great deal of his research in the fifties and sixties. The lecture notes [65] present an excellent survey on his own contributions to the theory of non-linear partial differential equations. His contributions in this area are fundamental and determined the shape and direction of research for a whole mathematical generation.

The paper [45], in 1941, seems to be his first contribution in this area, where he considered a viscous incompressible fluid filling a region G bounded by a finite number of surfaces each of which moving as a rigid body, and where he proved limitations for the kinetic energy. Still it relies to some extend on Leray's fundamental work in the early thirties; however, his work [57], in 1951, goes far beyond Leray's ideas by introducing the concept of weak solutions for treating Navier-Stokes equations and thus introducing linear methods of functional analysis into this non-linear problem. The basic idea is to replace the Navier-Stokes equation by the equation

$$\int_{\hat{G}} \left[\frac{\partial a_i}{\partial t} u_i + \frac{\partial a_i}{\partial x_\alpha} u_\alpha u_i + \mu \frac{\partial^2 a_i}{\partial x_\alpha \partial x_\beta} u_i \right] \mathrm{d}x \, \mathrm{d}t = 0$$

where $\hat{G}$ denotes an open domain in 4-dimensional space (i.e. introducing time t as a further coordinate). Then he proves the existence of solutions $u(x,t)$ with compact support and satisfying this equation.

The work in [56] laid the foundations of the viscosity method for treating shocks in fluid dynamics. Hopf analyzed the Burger's partial differential equation $u_t+uu_x=\mu_{xx'}$, which models the Navier-Stokes equation of fluid dynamics. Especially, he considers the behavior of the solutions in the limit $t\to\infty$ for finite μ and as $\mu\to 0$ (i.e. viscosity and conductivity approach zero). The limit function with $\mu\to 0$ is analogous to the solution for inviscid compressible fluid with shocks. He (and others) continued these investigations on Cauchy problems for first order equations. In [76] Conway and Hopf derived formulas representing the general solution, and Hopf gives a method of construction in [77] and a characterization in [78].

Eberhard Hopf studied the phenomena of turbulence in a series of papers ([53], [58], [62], [67], [68], [73]). In 1948, [53], he gave an example of a system, parameterized by μ, say, modelling a viscous fluid whose solutions exhibit turbulent behavior. The paper [67] provides a considerably simpler example of this nature. In particular, he showed that there is an infinite number of critical values $\mu_1 > \mu_2 \ldots \to 0$ such that for $\mu > \mu_1$ there is a stationary laminar solution which will be approached by any other solution for the same value of μ. For $\mu_n > \mu > \mu_{n+1}$ there is an n-dimensional manifold of turbulent solutions such that almost all other solutions will approach one of the turbulent solutions for the same value of μ. This is clearly a statistical treatment as it is reflected by one of Hopf's basic ideas discussed above. The work in [58], [62], [68] and [73] contain an attempt to introduce Fourier analysis into the subject, posing a variety of open problems.

References

[A] Anselone, P.M.: In honor of Professor Eberhard Hopf on the occasion of his seventieth birthday. Applicable Analysis **3** (1973) 1–5

[W] Weil, A.: Oeuvres scientifiques. Vol. 1 (1926–1951)

Publications of Eberhard Hopf

[1] Über die Zusammenhänge zwischen gewissen höheren Differenzenquotienten reeller Funktionen einer reellen Variablen und deren Differenzierbarkeitseigenschaften. Dissertation, Berlin 1926

[2] Bemerkungen zum ersten Randwertproblem der Potentialtheorie im Raume. Sitzungsberichte Berliner Math. Ges. (1926) 43–48

[3] (with E. Freundlich and U. Wegner) On the integral equation for radiative equilibrium. Roy. Astr. Soc. Monthly Notices **88** (1927) 139–142

[4] Eine Bemerkung zur Theorie der partiellen Differentialgleichungen zweiter Ordnung vom elliptischen Typus. Sitzungsberichte Berliner Math. Ges. (1927) 111–112

[5] Elementare Bemerkungen über die Lösungen partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus. Sitzungsberichte Preußische Akad. Wiss. Berlin, Math.-Nat. Klasse (1927) 147–152

[6] Ein Analogon zu einem Mittelwertsatze von H. A. Schwarz bei komplexen Polynomen. Jber. Dt. Math.-Verein. **37** (1928) 249–251

[7] Bemerkungen zu einem Satz von S. Bernstein aus der Theorie der elliptischen Differentialgleichungen. Math. Z. **29** (1928) 744–745

[8] Zum Problem des Strahlungsgleichgewichts in den äußeren Schichten der Sterne. Z. Physik **46** (1928) 374–382

[9] Über Strahlungsgleichgewicht in den äußeren Schichten der Sterne II. Z. Physik **49** (1928) 155–161

[10] Über lineare Integralgleichungen mit positivem Kern. Sitzungsberichte Preußische Akad. Wiss. Berlin, Math.-Nat. Klasse (1928) 233–245

[11] Über die geschlossenen Bahnen in der Mondtheorie. Sitzungsberichte Preußische Akad. Wiss. Berlin, Math.-Nat. Klasse (1929) 401–413

[12] Zum Strahlungsgleichgewicht der Sternatmosphären. Astr. Nachr. **235** (1929) 97–104

[13] Über eine Klasse singulärer Integralgleichungen; Bemerkungen zur Methode von Hardy und Titchmarsch. J. London Math. Soc. **4** (1929) 23–27

[14] Zum analytischen Charakter der Lösungen regulärer zweidimensionaler Variationsprobleme. Math. Z. **30** (1929) 404–413

[15] Remarks on the Schwarzschild-Milne model of the outer layers of a star. Roy. Astr. Soc. Monthly Notices **90** (1930) 287–293

[16] Zwei Sätze über den wahrscheinlichen Verlauf der Bewegungen dynamischer Systeme. Math. Ann. **103** (1930) 710–719

[17] (with N. Wiener) Über eine Klasse singulärer Integralgleichungen. Sitzungsberichte Preußische Akad. Wiss. Berlin, Math.-Nat. Klasse (1931) 696–706

[18] Kleine Bemerkung zur Theorie der elliptischen Differentialgleichungen. J. Angew. Math. **165** (1931) 50–51
[19] Über den Polytropenindex eines Sternmodells I. Z. Astrophys. **3** (1931) 67–70
[20] Über den Polytropenindex eines Sternmodells II. Z. Astrophys. **3** (1931) 71–76
[21] Über den Polytropenindex eines Sternmodells III. Z. Astrophys. **3** (1931) 108–115
[22] On Emden's differential equation. Roy. Astr. Soc. Monthly Notices **91** (1931) 653–663
[23] Mathematisches zur Strahlungsgleichgewichtstheorie der Fixsternatmosphäre. Math. Z. **33** (1931) 109–131
[24] Über den funktionalen, insbesondere den analytischen Charakter der Lösungen eliptischer Differentialgleichungen zweiter Ordnung. Math. Z. **34** (1932) 194–233
[25] Über lineare Gruppen unitärer Operatoren im Zusammenhange mit den Bewegungen dynamischer Systeme. Sitzungsberichte Preußische Akad. Wiss. Berlin, Math.-Nat. Klasse (1932) 182–190
[26] Remarks on the Schwarzschild-Milne model of the outer layers of a star II. Roy. Astr. Soc. Monthly Notices **92** (1932) 863–868
[27] On certain integral equations. Quart. J. Math. **3** (1932) 269–272
[28] On the time average theorem in dynamics. Proc. Nat. Acad. Sci. U.S.A. **18** (1932) 93–100
[29] Complete transitivity and the ergodic principle. Proc. Nat. Acad. Sci. U.S.A. **18** (1932) 204–209
[30] Proof of Gibbs' hypothesis on the tendency toward statistical equilibrium. Proc. Nat. Acad. Sci. U.S.A. **18** (1932) 333–340
[31] Theory of measures and invariant integrals. Trans. Amer. Math. Soc. **34** (1932) 373–393
[32] Mathematical Problems of Radiative Equilibrium. Cambridge Univ. Press (Cambridge Tracts in Mathematical Physics No. 31), 1934. Reprint 1964 by Stechert-Hafner, New York
[33] On causality, statistics and probability. J. Math. Phys. **13** (1934) 51–102
[34] Remarks on causality and probability. J. Math. Phys. **14** (1935) 4–9
[35] Absorption lines and the integral equation of radiative equilibrium. Roy. Astr. Soc. Monthly Notices **96** (1936) 522–533
[36] Über die Bedeutung der willkürlichen Funktionen für die Wahrscheinlichkeitstheorie. Jber. Dt. Math.-Verein. **46** (1936) 179–195
[37] Fuchsian groups and ergodic theory. Trans. Amer. Math. Soc. **39** (1936) 299–314
[38] Ein Verteilungsproblem bei dissipativen dynamischen Systemen. Math. Ann. **114** (1937) 161–186
[39] Ergodentheorie. Ergebnisse der Mathematik und ihrer Grenzgebiete 5, No. 2. Berlin: Springer 1937. Reprint 1948 by Chelsea Publ. Comp., New York. Reprint 1970 by Springer. Chapters 1–4 translated into Russian: Uspehi Matem. Nauk (N.S.) 4, No. 1 (1949) 113–182
[40] Beweis des Mischungscharakters der geodätischen Strömung auf Flächen der Krümmung minus Eins und endlicher Oberfläche. Sitzungsberichte Preußische Akad. Wiss. Berlin, Math.-Nat. Klasse (1938) 333–344
[41] Statistische Probleme und Ergebnisse in der klassischen Mechanik. Contribution to "Le principe ergodique et les probabilités en chaines". Actualités Sci. Indust. **737** (1938) 5–16
[42] Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung. Ber. Verh. Sächsische Akad. Math.-Nat. Klasse **91** (1939) 261–304. Translated into Russian: Uspehi Matem. Nauk (N.S.) 4, No. 2 (1949) 129–170
[43] Statistik der Lösungen geodätischer Probleme vom unstabilen Typus II. Math. Ann. **117** (1940) 590–608
[44] Randbemerkungen zu einigen Existenzsätzen der Differentialgeometrie. Jber. Dt. Math.-Verein. **49** (1940) 253–255
[45] Ein allgemeiner Endlichkeitssatz der Hydrodynamik. Math. Ann. **117** (1941) 764–775
[46] Abzweigung einer periodischen Lösung von einer stationären Lösung eines Differentialsystems. Ber. Verh. Sächsische Akad., Wiss. Math.–Nat. Klasse **95** (1943) 3–22
[47] Über eine Ungleichung der Ergodentheorie. Sitzungsberichte Math.-Nat. Abt., Bayer. Akad. Wiss. 1944 (1947) 171–176
[48] Kennzeichnung der durch Punkttransformation erzeugten linearen Funktionaloperatoren. Sitzungsberichte Math.-Nat. Abt., Bayer. Akad. Wiss. 1944 (1947) 233–236
[49] Über die Funktionalgleichungen der trignometrischen und hyperbolischen Funktionen. Sitzungsberichte Math.-Nat. Abt., Bayer. Akad. Wiss. 1945/46 (1947) 167–173

[50] Gewöhnliche Differentialgleichungen nach Prof. Dr. E. Hopf. München: Seelig & Co. Skriptenreihe Hsg. Studentenwerk München 1946
[51] (with W. Damköhler) Über einige Eigenschaften von Kurvenintegralen und über die Äquivalenz von indefiniten mit definiten Variationsproblemen. Math. Ann. **120** (1947) 12–20
[52] Closed surfaces without conjugate points. Proc. Nat. Acad. Sci. U.S.A. **34** (1948) 47–51
[53] A mathematical example displaying features of turbulence. Comm. on Appl. Math. **1** (1948) 303–322
[54] A theorem on the accessibility of boundary parts of an open point set. Proc. Amer. Math. Soc. **1** (1950) 76–79
[55] On S. Berstein's theorem on surfaces $z(x,y)$ of nonpositive curvature. Proc. Amer. Math. Soc. **1** (1950) 80–85
[56] The partial differential equation $u+uu_x=\mu u_{xx'}$. Comm. Pure Appl. Math. **3** (1950) 201–230
[57] Über die Anfangswertaufgabe für die hydrodynamischen Grundgleichungen. Math. Nachr. **4** (1950/51) 213–231
[58] Statistical hydromechanics and functional calculus. J. Rational Mech. Anal. **1** (1952) 87–123
[59] A remark on linear elliptic differential equations of second order. Proc. Acad. Amer. Math. Soc. **3** (1952) 791–793
[60] Remarks on the preceding paper by D. Gilbarg. J. Rational Mech. Anal. **1** (1952) 419–424
[61] On a inequality for minimal surfaces $z=z(x,y)$. J. Rational Mech. Anal. **2** (1953) 519–522
[62] (with E. W. Titt) On certain special solutions of the Φ-equation of statistical hydrodynamics. J. Rational Mech. Anal. **2** (1953) 587–591
[63] Correction to the paper "On an inequality for minimal surfaces $z=z(x, y)$". J. Rational Mech. Anal. **2** (1953) 801–802
[64] Five lectures on non-linear partial differential equations. Conf. Partial Differential Equations, University of California, Berkeley. Lecture Series of the Symposium on Partial Differential Equations. The University of Kansas Press 1955
[65] The general temporally discrete Markoff process. J. Rational Mech. Anal. **3** (1954) 13–45
[66] Die Harnacksche Ungleichung für positive harmonische Funktionen. Math. Z. **63** (1955/56) 156–157
[67] Repeated branching through loss of stability. An example. Proc. Conference on Differential Equations. College Park: University of Maryland Book Store 1956, 49–56
[68] On the application of functional calculus to the statistical theory of turbulence. Applied Probability. Proc. of Symposia in Applied Math. New York – Toronto – London: McGraw-Hill Book Co., for the Amer. Math. Soc., Providence **7** (1957) 41–50
[69] The temporal behaviour of a wave packet. Duke Math. J. **24** (1957) 477–480
[70] Zur Kennzeichnung der euklidischen Norm. Math. Z. **72** (1959/60) 76–81
[71] On the ergodic theorem for positive linear operators. J. Reine Angew. Math. **205** (1960) 101–106
[72] Some topics of ergodic theory. Centro Internazionale Matematico Estivo. Rome 1960
[73] Remarks on the function-analytic approach to turbulence. Proc. Sympos. in Applied Mathematics: Hydrodynamic Instability. Amer. Math. Soc., Providence **13** (1962) 157–163
[74] An inequality for positive linear integral operators. J. Math. Mech. **12** (1963) 683–692
[75] Remarks on my paper "An inequality for positive linear integral operations". J. Math. Mech. **12** (1963) 889–892
[76] (with E. D. Conway) Hamilton's theory and generalized solutions of the Hamilton-Jacobi equation. J. Math. Mech. **13** (1964) 929–986
[77] Generalized solutions of non-linear equations of first order. J. Math. Mech. **14** (1965) 951–973
[78] On the right weak solution of the Cauchy problem for a quasi-linear equation of the first order. J. Math. Mech. **19** (1969/70) 483–487
[79] Ergodic theory and the geodesic flow on surfaces of constant negative curvature. Bull. Amer. Math. Soc. **77** (1971) 863–877

PhD students of Eberhard Hopf

Dennis Mercer Nead (1952): A theorem on the limit passage viscosity $\mu \to 0$, heat conduction $\varkappa \to 0$, in the solutions of the equations of fluid flow

Robert Weiller (1955): Part I: On the differentiability of solutions of 2-dimensional regular variational problems. Part II: The Harnack-inequality for solutions of linear partial differential equations of elliptic type.

Dietrich Morgenstern (1955): Analytic studies related to the Maxwell-Boltzmann equation.

Terence Butler (1958): A mathematical example displaying features of turbulence.

Kasturi Lal Arora (1962): A boundedness theorem for the flow of a homogeneous incompressible and viscous fluid through an infinite cyclindrical pipe.

E. D. Conway III (1963): Hamilton-Jacobi theory and generalized solutions of partial differential equations.

Manfred Denker
Institut für Mathematische Stochastik
Lotzestr. 13
3400 Göttingen

(Eingegangen 25. 10. 1988)

Jber. d. Dt. Math.-Verein.
92 (1990) 58–76
AMS subject classification: 11 D 25

Elliptische Kurven: Fortschritte und Anwendungen

D. Zagier, Bonn and College Park

§ 1 Elliptische Kurven

Ich fange an mit einer Definition, die zunächst weder einleuchtend noch besonders vielversprechend aussieht: Eine *elliptische Kurve* ist eine Kurve, die durch eine Gleichung

(1) $\quad y^2 = x^3 + ax + b \quad (a, b \text{ Konstanten}, 4a^3 + 27b^2 \neq 0)$

in der Ebene gegeben werden kann.

Wegen ihrer Tiefe und der Vielfalt ihrer Zusammenhänge mit anderen Gebieten gehört die Theorie der elliptischen Kurven zu den schönsten in der Mathematik. Sie erscheint in der *Funktionentheorie* in der Gestalt der Theorie der elliptischen Funktionen, deren Entwicklung durch Gauß, Abel und Jacobi zu den Höhepunkten der Mathematik des 19ten Jahrhunderts gehört. Die elliptischen Funktionen sind die doppelt-periodischen Funktionen, die man findet, wenn man elliptische Integrale (die Funktionen, die die Bogenlänge auf Ellipsen messen – daher der Name) umkehrt; für beliebige $a, b \in \mathbb{C}$ läßt sich die Kurve (1) durch solche Funktionen parametrisieren, d. h. es gibt Perioden $\omega_1, \omega_2 \in \mathbb{C}$ und komplexwertige Funktionen $X(z)$, $Y(z)$ auf $\mathbb{C}/\mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$, die identisch die Gleichung $Y(z)^2 = X(z)^3 + aX(z) + b$ erfüllen. In der *algebraischen Geometrie* sind die elliptischen Kurven gleichwohl die einfachsten interessanten Kurven (d. h. Kurven von positivem Geschlecht) und die einfachsten Beispiele von abelschen Varietäten (Varietäten mit einer abelschen Gruppenstruktur, die durch polynomiale Funktionen gegeben wird) und dienen somit als ideale Einführung in die allgemeine Theorie. Insbesondere liefern sie schöne Anwendungen des Satzes von Riemann-Roch, der Theorie der Zetafunktionen von Varietäten über endlichen Körpern und des Begriffs einer formalen Gruppe. Der für den heutigen Vortrag wichtige Aspekt von elliptischen Kurven ist aber ihr Auftreten in der *Zahlentheorie*. Vor allem werden wir uns für die rationalen Lösungen der Gleichung (1) interessieren unter der Annahme, daß die Koeffizienten a und b rationale oder ganze Zahlen sind. Diese harmlos klingende Frage führt zu unerwartet tiefen Erkenntnissen in der Zahlentheorie. Ich will heute einen Überblick über diesen Aspekt der Theorie geben und erzählen, welche Fragen uns gegenwärtig am meisten beschäftigen, welche Fortschritte in letzter Zeit erzielt worden sind und welche Anwendungen inner- und außerhalb der Zahlentheorie daraus entstanden sind. Um die modernen Entwicklungen voll genießen zu können, muß man aber etwas über die Vorgeschichte der elliptischen Kurven wissen, die ich jetzt kurz schildern möchte. Dazu fange ich ganz früh an, mit Diophant, einem der größten (und meiner Meinung nach sehr unterschätzten) Mathematiker der Antike.

Die Methode von Diophant

Diophant, der in Alexandria vermutlich im dritten Jahrhundert lebte, war der erste Mathematiker, der systematisch mit negativen und rationalen Zahlen arbeitete, der erste, der Produkte von Zahlen als reine Zahlen betrachtete (statt – wie seine Vorgänger – als Flächeninhalte oder Volumina) und daher mit Potenzen größer als der dritten arbeiten konnte, und der erste, der einen kompletten algebraischen Symbolismus einführte und somit Gleichungen wie etwa $3x^3-2x^2=4$ (was bei ihm so ausgesehen hätte: $\bar{\gamma}K^y ⋔ \bar{\beta}\Delta^y\iota\bar{\delta}\mathring{M}$) schreiben konnte. Seine aus dreizehn Büchern bestehende *Arithmetika* ging nach der Vernichtung der alexandrinischen Bibliothek verloren. Erst im Jahre 1464 entdeckte der Astronom und Mathematiker Johannes Müller („Regiomontanus“) sechs der Bücher in Venedig; die Veröffentlichung (in Latein) erfolgte erst 1575. Diophant war seiner Zeit so weit voraus gewesen, daß es auch dann noch 50 Jahre dauerte, bis der erste europäische Mathematiker, und zwar Pierre de Fermat, sein Werk voll verstehen konnte.

Diophant hatte eine systematische Methode, um quadratische Gleichungen in zwei Veränderlichen zu lösen und kubische anzugreifen. Bei quadratischen ist seine Methode ganz leicht: man fängt mit einer schon bekannten Lösung an, schreibt die Gleichung einer beliebigen Geraden durch diesen Punkt auf und erhält damit eine quadratische Gleichung in einer Veränderlichen, von der man eine rationale Lösung schon kennt; die zweite Lösung ist dann ebenfalls rational. Um z. B. die Lösungen von $x^2+y^2=1$ zu finden, fängt man mit der Lösung $(x,y)=(0,1)$ an, schreibt die Gleichung einer durch diesen Punkt gehenden Geraden als $y=1-tx$ (t rational) auf und rechnet dann:

$$x^2+(1-tx)^2=1, \qquad x^2-2tx+t^2x^2=0,$$
$$x=\frac{2t}{1+t^2}, \qquad y=1-tx=\frac{1-t^2}{1+t^2}.$$

Somit erhält man die allgemeine ganzzahlige Lösung $a=2pqr$, $b=(p^2-q^2)r$, $c=(p^2+q^2)r$ $(p, q, r \in \mathbb{Z})$ der pythagoräischen Gleichung $a^2+b^2=c^2$ $\left(\text{setze } x=\frac{a}{c}, y=\frac{b}{c}, t=\frac{q}{p}\right)$, die allerdings lange vor Diophant bekannt war und insbesondere schon bei Euklid steht.

Bei kubischen Gleichungen versagt Diophants Methode zunächst, weil die Gerade *zwei* weitere Schnittpunkte mit der Kurve besitzt, aber Diophant entdeckte, daß man bei passender Wahl der Steigung der Geraden nur *eine* weitere Lösung, und zwar eine rationale, findet. Als Beispiel betrachten wir die Gleichungen $y^2=x^3-3x+1$. Wie vorher setzen wir die Gleichung $y=1-tx$ der allgemeinen Geraden durch die bekannte Lösung $(x,y)=(0,1)$ ein und erhalten eine polynomiale Gleichung in x, deren konstantes Glied wegfällt:

$$(1-tx)^2=x^3-3x+1, \qquad x^3-t^2x^2+(2t-3)x=0.$$

Da diesmal aber die Gleichung kubisch statt quadratisch ist, haben wir auch nach Kürzung eines Faktors x immer noch eine quadratische Gleichung, deren Lösungen nicht rational zu sein brauchen. Wählen wir aber speziell $t=3/2$, so fällt auch das lineare Glied unserer kubischen Gleichung weg und wir können einen Faktor x^2 kürzen, um eine lineare Gleichung zu erhalten, die sofort gelöst werden kann:

$$y=1-\frac{3}{2}x, \qquad x^3-\frac{9}{4}x^2=0, \qquad x=\frac{9}{4}, \qquad y=1-\frac{3}{2}\cdot\frac{9}{4}=-\frac{19}{8}.$$

Der geometrische Inhalt der Methode ist klar: während wir im Falle eines Kegelschnittes durch einen gegebenen Punkt eine *beliebige* Gerade ziehen und genau einen weiteren Schnittpunkt erhalten konnten, müssen wir im kubischen Falle die *Tangente* wählen, damit nicht zwei, sondern nur ein weiterer Schnittpunkt entsteht. Die geometrische Interpretation, die sehr wichtig ist – wir werden in Zukunft gar nicht zwischen den *Lösungen* der Gleichung (1) und den *Punkten* auf der von ihr beschriebenen Kurve E unterscheiden – stand natürlich Diophant noch nicht zur Verfügung. Sie wurde erst von Isaac Newton bemerkt.

Fermat und die Methode des unendlichen Abstiegs

Diophants Methode wurde von Fermat zu seiner berühmten Methode der „descente infinie" oder „unendlicher Abstieg" weiterentwickelt. Diese Methode sieht wie folgt aus. Die Diophantische Konstruktion assoziiert zu jeder Lösung $P=(x,y)$ von (1) eine zweite Lösung $P'=(x',y')$ (= dritter Schnittpunkt der Tangenten durch P mit der Kurve), die im allgemeinen kompliziertere Koordinaten hat, d. h., die Zähler und Nenner von x' und y' sind meistens viel größer als die von x und y. Unter geeigneten Umständen kann man zeigen, daß jede Lösung von (1) auf diese Weise entsteht, d. h. sie hat die Gestalt P' für eine i. a. viel einfachere Lösung P. Wiederholt man dieses Verfahren, so kommt man zu immer kleineren Lösungen von (1) und somit zu einem Widerspruch, falls (1) keine ganz kleinen Lösungen hat, und zu einer Beschreibung aller Lösungen, falls es eine kleine Lösung gibt. Die Methode hat sowohl negative wie auch positive Anwendungen. Z. B. hat Fermat selbst sie benutzt, um zu zeigen, daß die Gleichungen $x^4+y^4=z^2$ und $x^4+y^2=z^4$ (also erst recht die Gleichung $x^4+y^4=z^4$) keine Lösungen in positiven ganzen Zahlen besitzen, aber auch, um das einfachste ganzzahlige pythagoräische Dreieck zu finden, wovon sowohl die Hypotenuse wie auch die Summe der Katheten Quadrate sind (Brief an Mersenne, 1643):

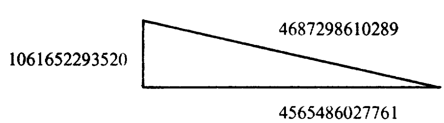


Fig. 1

Wir illustrieren Fermats Methode anhand eines sehr klassischen Problems, das über 1000 Jahre alt ist (es taucht in einem vor 972 datierten anonymen arabischen Manuskript auf) und dem Leonardo Pisano („Fibonacci") ein ganzes Buch, sein *liber quadratorum*, widmete: das Problem der sogenannten kongruenten Zahlen. Die Aufgabe besteht darin, eine gegebene Zahl n als Flächeninhalt eines rechtwinkligen Dreiecks mit rationalen Seiten darzustellen oder (äquivalenterweise: Übungsaufgabe!) drei rationale Quadrate in arithmetischer Progression mit gemeinsamer Differenz n zu finden. Falls das Dreieck bzw. die Quadrate existieren, heißt n „kongruent". Für $n=5$ lautet die bereits in den erwähnten Büchern gegebene Lösung

$$\frac{3}{2}, \frac{20}{3}, \frac{41}{6}$$

für die Seiten des Dreiecks bzw.

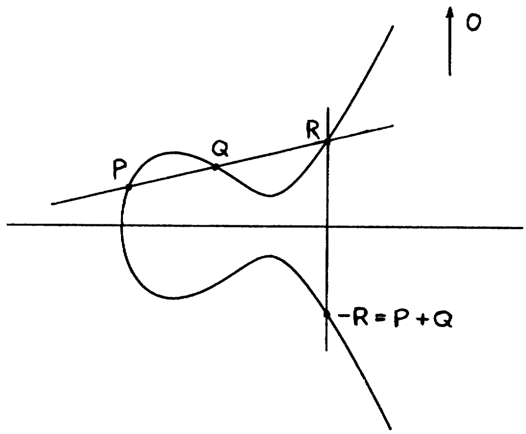
$$6\frac{97}{144}=\left(2\frac{7}{12}\right)^2, 11\frac{97}{144}=\left(3\frac{5}{12}\right)^2, 16\frac{97}{144}=\left(4\frac{1}{12}\right)^2$$

für die Quadrate in arithmetischer Progression.

Um das Problem anzugehen, betrachten wir die durch die Gleichung $y^2=x(x+n)(x-n)$ gegebene elliptische Kurve. Ist $P=(x,y)$ eine beliebige nichttriviale Lösung (d. h. $y\neq 0$) dieser Gleichung, so hat der durch Diophants Tangentenmethode konstruierte Punkt $P'=(x',y')$ die Eigenschaft, daß nicht nur das Produkt $x'(x'+n)(x'-n)\,(=y'^2)$, sondern alle drei Faktoren $x', x'+n, x'-n$ Quadratzahlen sind, d. h., n ist kongruent. Für die Ausgangslösung P brauchen x und $x\pm n$ nicht Quadrate zu sein, aber sie sind bis auf quadratische Faktoren sehr eingeschränkt. Ist z. B. n prim und $\equiv 5 \pmod 8$, so zeigt man leicht, daß jede von diesen drei Zahlen eine der Gestalten $\pm\square$, $\pm 2\cdot\square$, $\pm n\cdot\square$ oder $\pm 2n\cdot\square$ ($\square=$ rationale Quadratzahl) haben muß. Dies führt zur Unterscheidung endlich vieler Fälle, die einzeln untersucht werden. Ist z. B. $x=-A^2$, $x+n=B^2$, $x-n=-C^2$ (und man kann in unserem Falle n prim, $n\equiv 5 \pmod 8$ leicht zeigen, daß es, wenn überhaupt, stets eine solche Lösung gibt), so müssen wir das Gleichungssystem $C^2-B^2=2A^2$, $C^2-A^2=n$ lösen. Die erste Gleichung läßt sich nach Diophants Methode sofort lösen: es muß gelten $A=2RS/M$, $B=(R^2-2S^2)/M$, $C=(R^2+2S^2)/M$ für passende ganze Zahlen R, S, M. Damit ist das Problem auf die Lösbarkeit von $M^2n=R^4+4S^4$ reduziert worden. Für $n=5$ ist jetzt die Lösung $M=R=S=1$ $\left(\Rightarrow x=-4, y=6, x'=6\frac{97}{144}\right)$ evident. Für andere Primzahlen n muß man unter Umständen den Abstieg ein- oder mehrmals wiederholen, wobei man als ersten Schritt die Diophantische Methode auf die quadratische Gleichung $n=U^2+4V^2$ anwendet und eine Lösung mit $UV=\square$ ($\Rightarrow U=R^2/M$, $V=S^2/M$) zu finden versucht. Für $n=157$ etwa führt diese Methode nach einigen Schritten zu einer Lösung. Diese ist, wie Fermat vielleicht gesagt haben würde, ganz wunderbar, aber die Seite ist leider zu schmal, um sie zu enthalten: die drei rationalen Quadratzahlen haben im Zähler und Nenner jeweils fast 100 Dezimalziffern.

Gruppenstruktur. Satz von Mordell

Was Diophant und Fermat entdeckt hatten, war die erste Andeutung des Additionsgesetzes auf elliptischen Kurven: Statt der Tangenten durch *einen* Punkt P kann man die Sehne durch *zwei* Punkte P, Q betrachten, die einen dritten Schnittpunkt mit der Kurve besitzt. Schreibt man diese als $-(P+Q)$, so ist „+", wie sich herausstellt, eine kommutative, assoziative binäre Operation auf der Menge der Lösungen, die (unter Hinzunahme des unendlichen fernen Punkts O als Nullelement) damit zu einer *abelschen Gruppe* wird (Fig. 2). Es ist die Existenz dieser Zusatzstruktur, die für den Reichtum der Theorie der elliptischen Kurven verantwortlich ist.



Gruppengesetz auf $E(\mathbb{Q})$
Addition: $P+Q+R=0 \iff P, Q, R$ kollinear
Inverses: $P=(x,y) \implies -P=(x,-y)$
Nullelement: $O=$ Punkt im Unendlichen

Fig. 2

Natürlich kam diese Formulierung in der Sprache der Gruppentheorie erst viel später, und zwar mit Poincaré, um 1900. Poincaré hat vermutet, daß die Gruppe der rationalen Lösungen einer elliptischen Kurve mit rationalen Koeffizienten stets *endlich erzeugt* sei, d. h. daß man sämtliche Lösungen der Gleichung aus endlich vielen durch wiederholte Anwendung der Sehnenkonstruktion erhalten könne. Seine Vermutung wurde 1922 von dem englischen Mathematiker L. J. Mordell bewiesen. Mordells Satz wurde 1940 von A. Weil verallgemeinert, indem er beliebige algebraische Zahlkörper (statt $\mathbb{Q}$) und beliebige abelsche Varietäten (statt nur elliptische Kurven) zuließ. Die Gruppe der rationalen Punkte auf einer elliptischen Kurve E (einschließlich des Punktes O) wird die *Mordell-Weil-Gruppe* genannt und mit $E(\mathbb{Q})$ bezeichnet. Unsere Hauptfragen betreffen die Struktur dieser Gruppe.

Struktur der Mordell-Weil-Gruppe

Nach dem Struktursatz für endlich erzeugte abelsche Gruppen hat die Mordell-Weil-Gruppe die Gestalt

(2) $$E(\mathbb{Q}) \simeq \underbrace{\mathbb{Z} \oplus \mathbb{Z} \oplus \ldots \oplus \mathbb{Z}}_{r} \oplus T$$

für eine bestimmte Zahl $r \geqslant 0$ (der *Rang* von E) und eine endliche abelsche Gruppe T (die *Torsionsgruppe* von E). Die Zahl r, die die wichtigste diophatische Invariante der Kurve E ist, hat eine schöne Interpretation, die ohne die Gruppenstruktur auf $E(\mathbb{Q})$ erklärt werden kann: bezeichnet $N(B)$ die Anzahl der Lösungen von Gleichung (1), für die der Zähler von x kleiner oder gleich einer positiven Zahl B ist, so gilt

(3) $$N(B)+1 \sim C(\log B)^{r/2} \quad (B \to \infty)$$

mit einer bestimmten, nur von der Kurve E abhängigen reellen Konstanten C. (Die Zahl „$+1$" auf der linken Seite von (3) wird gebraucht, weil man beim Zählen der rationalen Lösungen auch den Punkt im Unendlichen mitzählen muß.) Der Beweis von (3) ist ein Korollar des Beweises des Mordellschen Satzes.

In Anbetracht der Formel (2) können wir die Frage nach der Struktur von $E(\mathbb{Q})$ wie folgt präzisieren:

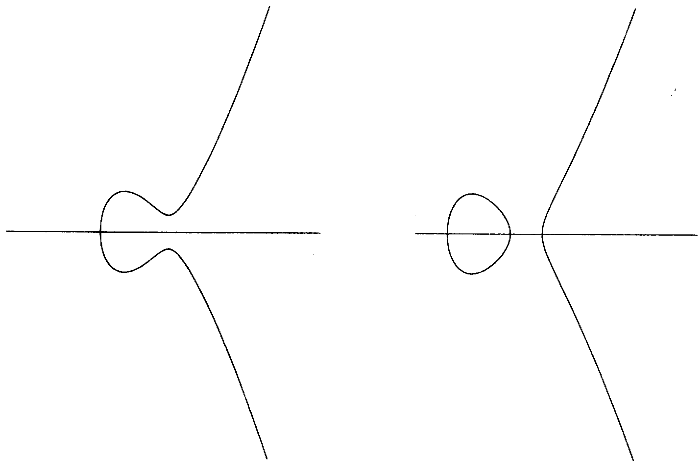
i) Wie bestimmt man die Torsionsgruppe T einer gegebenen elliptischen Kurve?
ii) Was sind die Möglichkeiten für T? Wie häufig kommen sie jeweils vor?
iii) Wie bestimmt man den Rang r einer gegebenen elliptischen Kurve?
iv) Was sind die Möglichkeiten für r? Wie häufig kommen sie jeweils vor?

Die Antwort zur ersten Frage ist ein elementarer Satz, die zur zweiten Frage ein sehr schwieriger Satz, die zur dritten Frage nur eine Vermutung, und die zur vierten Frage noch unbekannt. Genauer:

Zu i): Seien o.B.d.A. die Zahlen a, b in (1) ganz (da wir sie vermöge $x \mapsto k^2x, y \mapsto k^3y$ durch k^4a, k^6b ersetzen können). Dann gilt der relativ elementare Satz von Nagell-Lutz: Jeder Torsionspunkt $P \neq 0$ in $E(\mathbb{Q})$ hat ganze Koordinaten und eine y-Koordinate, die entweder gleich 0 oder ein Teiler von $4a^3+27b^2$ ist. (Natürlich ist $y=0$ mit $P=-P$ oder $2P=O$ äquivalent, da $-(x,y)=(x,-y)$.) Damit hat man für eine gegebene Kurve E nur endlich viele Kandidaten für Torsionspunkte und kann sie alle durchprobieren.

Zu ii): Die Gruppe $E(\mathbb{Q})$ ist eine Untergruppe von $E(\mathbb{R})$, der Gruppe der *reellen* Lösungen von (1) (zusammen mit dem unendlich fernen Punkt O). Diese Gruppe ist isomorph zur Kreislinie $S^1 = \mathbb{R}/\mathbb{Z}$ oder zum Produkt $S^1 \times \mathbb{Z}/2\mathbb{Z}$, je nachdem, ob $4a^3+27b^2$ positiv oder negativ ist (Fig. 3).

Somit muß der Torsionsanteil T von $E(\mathbb{Q})$ entweder zyklisch oder das Produkt einer zyklischen Gruppe mit $\mathbb{Z}/2\mathbb{Z}$ sein. Da die 2-Torsion von $E(\mathbb{Q})$ gerade die



$4a^3+27b^2>0$ $4a^3+27b^2<0$

Fig. 3

Menge der Punkte $(x, 0)$ mit $x \in \mathbb{Q}$, $x^3+ax+b=0$, ist, haben wir den Struktursatz:

x^3+ax+b über $\mathbb{Q}$ irreduzibel $\Rightarrow$ $T \simeq \mathbb{Z}/n\mathbb{Z}$, n ungerade,
$x^3+ax+b=$ (linear) $\times$ (quadratisch) über $\mathbb{Q}$ $\Rightarrow$ $T \simeq \mathbb{Z}/n\mathbb{Z}$, n gerade,
x^3+ax+b zerfällt in lineare Faktoren über $\mathbb{Q}$ $\Rightarrow$ $T \simeq \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, n gerade.

So weit ist alles elementar. Alles andere als elementar ist der 1977 von B. Mazur bewiesene Satz, daß der Wert von n in den drei Fällen durch 9 bzw. 12 bzw. 8 beschränkt ist. Hiernach gibt es genau 15 Möglichkeiten für die Torsionsgruppe T. Alle kommen für unendlich viele Kurven E vor. Dabei ist allerdings $|T|=1$ der generische Fall, und die anderen Möglichkeiten treten nur für spezielle Werte von b^2/a^3 auf, z. B. kann $|T|$ nur gerade sein, wenn b^2/a^3 die Gestalt $c(c+1)^2$ mit $c \in \mathbb{Q}$ hat (nämlich $c=x^2/a$, $x \in \mathbb{Q}$, $x^3+ax+b=0$). Mazurs Satz beruht auf einer sehr tiefliegenden Analyse von Modulkurven vom Standpunkt der arithmetischen algebraischen Geometrie.

Zu iii): Eine obere Abschätzung für den Rang einer gegebenen elliptischen Kurve E wird durch die Fermatsche Methode des unendlichen Abstiegs geliefert. Z. B. hat die Kurve $y^2=x^3-n^2x$, die wir im Zusammenhang mit dem Problem der kongruenten Zahlen betrachteten, den Rang 0, falls n eine Primzahl $\equiv 3 \pmod 8$ ist, Rang $\leqslant 1$, falls n prim und $n \equiv 5$ oder $7 \pmod 8$, und Rang $\leqslant 2$, falls n prim und $n \equiv 1 \pmod 8$. Eine untere Abschätzung kann natürlich dadurch geschehen, daß man eine bestimmte Anzahl von linear unabhängigen Lösungen hinschreibt. Z. B. hat die Kurve $y^2=x^3-25x$ Rang mindestens 1 (also tatsächlich gleich 1), weil $(-4, 6)$ eine Lösung unendlicher Ordnung ist. Häufig gibt es aber eine Kluft zwischen der oberen und der unteren Abschätzung – entweder, weil Fermats Methode zu grob ist (z. B. ist im Falle $n \equiv 1 \pmod 8$ oben der wahre Wert des Ranges häufig 0 statt 2) oder, weil die Lösungen außerhalb des Suchbereichs liegen (etwa die fast 100-stellige kleinste Lösung für $n=157$). Eine einigermaßen befriedigende Antwort auf die Frage, wie man den Rang bestimmt, wird erst durch die Birch-Swinnerton-Dyer-Vermutung gegeben, die wir gleich besprechen wollen.

Zu iv): Hier ist, wie bereits gesagt, nichts bekannt. Der Weltrekord für den größten bekannten Rang stand 1948 bei 4, 1974 bei 6, und ist jetzt 14 (Mestre, 1984), nachdem Néron 1954 abstrakt (d. h., ohne Beispiele anzugeben) bewiesen hatte, daß es unendlich viele elliptische Kurven mit Rang $\geqslant 11$ geben muß. Es wird allgemein vermutet, daß es elliptische Kurven mit beliebig hohem Rang gibt. Kurven von hohem Rang sind aber nicht leicht zu finden: Mestres Beispiel hat (in der Schreibweise (1)) die Koeffizienten $a=-35971713708112$, $b=85086213848298394000$.

Über die Frage, wie die verschiedenen Werte von r verteilt sind, ist ebenfalls nichts bekannt. Immerhin liefert die Vermutung von Birch und Swinnerton-Dyer eine Vorhersage für die Parität von r (siehe unten), die ungefähr gleich häufig gerade und ungerade ausfällt. Bis vor kurzer Zeit glaubten die meisten Spezialisten, daß für fast alle Kurven der Rang den minimalen Wert annimmt, der nach dieser Einschränkung der Parität zulässig ist, d. h., r ist fast immer gleich 0, wenn er gerade sein muß, und gleich 1, wenn er ungerade sein muß. Aber numerische Rechnungen für die Familie der elliptischen Kurven $x^3+y^3=m$ ($m\leqslant 70000$, m kubenfrei), die vor kurzem ausgeführt worden sind, ergaben für $r=0$, $r=1$, $r\geqslant 2$ die Dichten 38,3%, 48,9% und 12,8% anstelle der erwarteten Werte 50%, 50% und 0%. Die Situation hier ist also zur Zeit unklar.

Die Vermutung von Birch und Swinnerton-Dyer. L-Reihen. Modulare elliptische Kurven

Die entscheidende Idee zu der Frage, wie man den Rang einer elliptischen Kurve finden kann, ist in einer Vermutung enthalten, die Anfang der 60er Jahre von B. Birch und H. P. S. Swinnerton-Dyer aufgestellt wurde. Diese Idee ist, daß eine Gleichung mit besonders vielen rationalen Lösungen (d. h. eine Kurve mit besonders hohem Rang) auch besonders viele Lösungen in dem Körper der ganzen Zahlen modulo einer Primzahl p haben soll – mindestens „im Durchschnitt", wenn p variiert. Um dies quantitativ zu formulieren, definiert man für jede Primzahl p

$$N(p)=\text{Anzahl der Paare } (x,y) \pmod{p} \text{ mit } y^2\equiv x^3+ax+b.$$

Man würde naiv erwarten, daß $N(p)$ ungefähr gleich p ist, weil es p^2 Paare (x,y) $(\bmod p)$ gibt und für jedes Paar die Wahrscheinlichkeit, daß die Kongruenz gilt, gleich $1/p$ ist, oder weil es p Werte für x modulo p gibt und für jedes x mit $x^3+ax+b\not\equiv 0$ die Anzahl der Quadratwurzeln von x^3+ax+b mit gleicher Wahrscheinlichkeit gleich 0 oder 2 ist. Diese naive Erwartung wird bestätigt durch den von H. Hasse 1933 bewiesenen Satz

$$(4) \qquad p-2\sqrt{p}<N(p)<p+2\sqrt{p}$$

(„Riemannsche Vermutung für elliptische Kurven"). Die Vermutung von Birch und Swinnerton-Dyer („BSD-Vermutung") besagt nun in der einfachsten Formulierung, daß man die asymptotische Formel

$$(5) \qquad \prod_{p<x} \frac{N(p)+1}{p} \sim C_1 (\log x)^r \quad (x \to \infty)$$

hat, wobei C_1 eine positive reelle Zahl ist, die nur von der elliptischen Kurve abhängt, und r den Rang der Kurve bezeichnet. (Man nimmt hier $N(p)+1$ statt $N(p)$ aus demselben Grund wie in (3).) Daß dieselbe Zahl r in den asymptotischen Formeln (3) und (5) vorkommt, stellt eine wunderbare und tiefliegende Beziehung zwischen den Lösungen von (1) in den Körpern $\mathbb{Q}$ und $\mathbb{Z}/p\mathbb{Z}$ dar.

Allerdings ist die Formel (5) wegen der sehr langsamen Divergenz des Produktes sowohl für theoretische wie auch für numerische Überlegungen etwas ungeeignet. Eine viel bessere Formulierung der BSD-Vermutung, die ebenfalls von Birch und Swinnerton-Dyer angegeben wurde, involviert die *L-Reihe* von E. Diese ist eine Funktion einer komplexen Variablen s, die für s genügend groß durch die Formel

$$(6) \qquad L(s) \approx \prod_p \frac{1}{1+(N(p)-p)p^{-s}+p^{1-2s}}$$

definiert wird, wobei das Produkt über alle Primzahlen p läuft und das Zeichen „$\approx$" bedeutet, daß die Definition des Faktors für p geändert werden muß, falls p die Diskriminante $4a^3+27b^2$ teilt (die genaue Definition in diesem Falle spielt für uns keine Rolle). Wegen (4) konvergiert dieses Produkt, falls s (oder der Realteil von s) größer als 3/2 ist. Die präzisere Formulierung der BSD-Vermutung besagt, daß die Funktion $L(s)$ in s analytisch fortsetzbar ist und eine Nullstelle genau r-ter Ordnung im Punkt $s=1$ besitzt. (Man sieht, daß dies eine Art Präzisierung von (5) ist, weil bei $s=1$ das unendliche Produkt in (6) formal zu $\prod \frac{p}{N(p)+1}$ wird.) Ferner soll der Anfang der Taylor-Entwicklung von $L(s)$ bei $s=1$ durch die Formel

$$(7) \qquad L(s) \sim C_0 (s-1)^r \quad (s \to 1), \qquad C_0 = \pi^{-r}\Gamma\left(\frac{r}{2}+1\right)^2 C^{-2}\,\Omega\,|\text{Ш}|$$

gegeben werden, wobei C die Konstante in (3) bezeichnet, Ω ein einfaches Vielfaches des elliptischen Integrals $\int_\alpha^\infty (x^3+ax+b)^{-1/2}dx$ (α = größte reelle Wurzel von $x^3+ax+b=0$), und $|\text{Ш}|$ die Ordnung der sogenannten *Tate-Schafarewitsch-Gruppe* Ш von E, einer bestimmten abelschen Gruppe, die mit der Lösbarkeit in rationalen Zahlen von gewissen mit (1) verwandten diophantischen Gleichungen zu tun hat (die genaue Definition von Ш werden wir nicht geben). Die Vermutung (7) ist besonders kühn, weil man im allgemeinen nicht nur nicht weiß, ob die Funktion $L(s)$ in der Nähe von $s=1$ erklärt ist, sondern auch nicht, ob die Gruppe Ш überhaupt endlich ist!

Die Gestalt der Euler-Faktoren in (6) ist zunächst keineswegs einleuchtend. Daß man genau diese Definition der L-Funktion nehmen muß, liegt daran, daß es eine sehr große Klasse von elliptischen Kurven gibt, für die man weiß, daß $L(s)$ gute analytische Eigenschaften hat. Das sind die sogenannten ***modularen elliptischen Kurven***. Eine elliptische Kurve E heißt ***modular***, wenn sie durch ***Modulfunktionen*** parametrisiert werden kann, d. h., es gibt eine Zahl N und zwei Funktionen $\xi(\tau)$ und $\eta(\tau)$ $(\tau \in \mathbb{C}, \operatorname{Im}(\tau) > 0)$ mit den folgenden Eigenschaften:

(a) $\xi(\tau)$ und $\eta(\tau)$ sind invariant unter $\tau \mapsto \dfrac{A\tau + B}{NC\tau + D}$ $(A, B, C, D \in \mathbb{Z}, AD - NBC = 1)$;
(b) $\xi(\tau)$ und $\eta(\tau)$ haben Fourier-Entwicklungen in $e^{2\pi i\tau}$ mit Koeffizienten in $\mathbb{Q}$;
(c) Es gilt $\eta(\tau)^2 = \xi(\tau)^3 + a\xi(\tau) + b$ identisch in τ.

Unter diesen Voraussetzungen weiß man, daß die zwei folgenden Behauptungen gelten:

(d) Ist $L(s) = \sum_{n=1}^{\infty} a_n n^{-s}$, so ist die assoziierte Funktion $f(\tau) = \sum_{n=1}^{\infty} a_n e^{2\pi i n\tau}$ eine *Modulform der Stufe N*: es gilt $f\left(\dfrac{A\tau + B}{NC\tau + D}\right) = (NC\tau + D)^2 f(\tau)$ für $A, B, C, D \in \mathbb{Z}, AD - NBC = 1$.
(e) Die Funktion $L(s)$ ist eine ganze Funktion von s und erfüllt die Funktionalgleichung $(2\pi)^{-s}N^{s/2}\Gamma(s)L(s) = \pm(2\pi)^{s-2}N^{1-s/2}\Gamma(2-s)L(2-s)$. Dabei läßt sich das Vorzeichen explizit bestimmen, was zusammen mit der Formel (7) eine explizite Vermutung über die Parität des Ranges liefert (r soll gerade oder ungerade sein, je nachdem, ob ein Plus- oder Minuszeichen in der Funktionalgleichung vorkommt).

Es wird allgemein vermutet, daß *jede* elliptische Kurve über $\mathbb{Q}$ in der Tat modular ist. Diese tiefe und aufregende Vermutung, die auch außerhalb der Theorie der elliptischen Kurven sehr wichtige Konsequenzen haben würde (siehe „Anwendungen"), heißt meistens die *Taniyama-Weil-Vermutung*, weil Taniyama 1955 die Frage aufgeworfen hat, ob die L-Funktionen von elliptischen Kurven mit den L-Reihen von Modulformen zusammenhängen, und Weil 1967 bewiesen hat, daß elliptische Kurven über $\mathbb{Q}$, deren assoziierte L-Funktionen Funktionalgleichungen wie in (e) besitzen (was nach den schon damals existierenden Vermutungen über Zetafunktionen von algebraischen Varietäten immer zu erwarten gewesen wäre), tatsächlich modular sind. In der Literatur findet man häufig die Terminologie „Weil-Kurve" statt „modulare elliptische Kurve".

§ 2 Fortschritte

Im letzten Abschnitt haben wir einen Teil der historischen Entwicklung der Theorie der elliptischen Kurven beschrieben und dabei auch über einige wichtige Sätze berichtet, die in neuerer Zeit bewiesen worden sind (z. B. Hasses Satz über die Anzahl der Punkte modulo einer Primzahl und Mazurs Satz über die möglichen Torsionsuntergruppen einer elliptischen Kurve über $\mathbb{Q}$). Wir wollten jetzt systematischer über neuere Ergebnisse berichten, wobei wir uns nur auf solche konzentrieren werden, die mit der Vermutung von Birch und Swinnerton-Dyer zusammenhängen. Natürlich hat es viele wichtige und schöne Resultate gegeben, die andere Aspekte der Theorie betreffen, etwa den Satz von Silverman (1986), wonach – grob formuliert – die Anzahl der ganzzahligen Lösungen von (1) durch $C^{(r+1)(s+1)}$

(r = Rang der elliptischen Kurve, s = Anzahl der Primfaktoren von $4a^3+27b^2$, C eine effektiv bestimmbare universelle Konstante) abgeschätzt werden kann (daß diese Anzahl endlich ist, wurde schon 1929 von C. L. Siegel bewiesen), oder der Satz von Elkies (1987), wonach eine beliebige elliptische Kurve E unendlich viele supersinguläre Primzahlen p (das sind solche mit $N(p)=p$) besitzt. Aus naheliegenden Gründen müssen wir uns aber auf nur einen Teil der Theorie beschränken.

Um einige der Ergebnisse zu formulieren, müssen wir noch eine wichtige Klasse von elliptischen Kurven einführen. Das sind die sog. Kurven mit *komplexer Multiplikation* oder (nach den englischen Anfangsbuchstaben) *CM-Kurven*. Eine elliptische Kurve E hat genau dann komplexe Multiplikation, wenn sie eine nichttriviale algebraische Abbildung in sich selbst zuläßt, die die Gruppenstruktur erhält. Zum Beispiel ist

$$y^2 = x^3 - 35x + 98 = (x+7)(x-\alpha)(x-\bar{\alpha}) \qquad \left(\alpha = \frac{7+i\sqrt{7}}{2}\right)$$

eine CM-Kurve, weil mit (x, y) auch

$$\left(x' = \frac{(x+7)(x-\bar{\alpha})}{\lambda^2(x-\alpha)} + \bar{\alpha},\ y' = \frac{y(x+i\sqrt{7})(x-7-2i\sqrt{7})}{\lambda^3(x-\alpha)^2}\right)$$

$$\left(\lambda = \frac{-1+i\sqrt{7}}{2}\right)$$

eine Lösung ist und die Abbildung $(x,y)\mapsto(x',y')$ ein Gruppenhomomorphismus ist. Die Kurven mit komplexer Multiplikation sind relativ selten: man weiß, daß die Gleichung (1) (mit $a,b\in\mathbb{Q}$) genau dann eine CM-Kurve definiert, wenn entweder a oder b gleich Null ist oder die rationale Zahl a^3/b^2 einen von 11 speziellen Werten annimmt, wovon der einfachste $-\frac{125}{28}$ und der komplizierteste $-\frac{37982843264000}{5627087890963}$ ist. Eine Besonderheit von CM-Kurven, deren Entdeckung im Prinzip auf Gauß zurückgeht, ist, daß die Anzahl $N(p)$ der Lösungen modulo einer Primzahl p durch eine einfache Formel gegeben werden kann, z. B. ist diese Anzahl für die Kurve $y^2=x^3-35x+98$ gleich $p\pm 2M$, falls p eine Darstellung als M^2+7N^2 zuläßt, und gleich p sonst. Nach Arbeiten von Max Deuring weiß man, daß jede CM-Kurve auch modular ist. Die bisher eingeführten

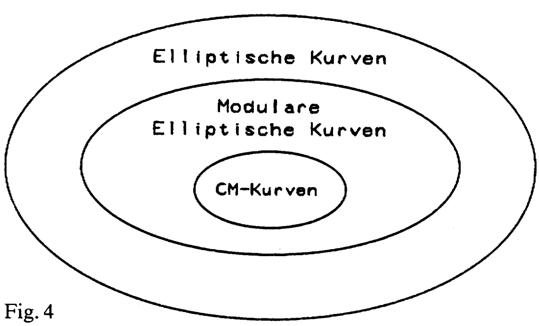


Fig. 4

Klassen von elliptischen Kurven lassen sich also durch ein Venn-Diagramm veranschaulichen, wobei die kleinste Teilmenge klein und die mittlere Teilmenge vermutlich gleich der gesamten Menge ist.

Wir bezeichnen den Rang einer elliptischen Kurve E mit r und, falls E modular ist, die Ordnung der L-Reihe von E bei $s=1$ (die manchmal „analytischer Rang von E“ genannt wird, weil sie nach der BSD-Vermutung mit r übereinstimmen soll) mit r'. Die Sätze, die wir nennen wollen, lauten nun wie folgt:

Coates-Wiles (1977): Ist E eine CM-Kurve mit $r'=0$, so ist auch $r=0$.

Der Satz besagt also, daß die Mordell-Weil-Gruppe einer CM-Kurve endlich ist, falls $L(1)$ nicht verschwindet. Der sehr raffinierte Beweis läuft nicht über die komplexe L-Reihe, sondern über ein p-adisches Analogon $L_p(s)$, dessen Wert bei $s=1$ mit dem entsprechenden Wert der komplexen L-Funktion zusammenhängt.

Greenberg (1983): Ist E eine CM-Kurve mit r' ungerade, so ist entweder $r \geqslant 1$ oder Ш, die Tate-Schafarewitsch-Gruppe von E, ist unendlich.

Zu diesem Satz können wir, da wir Ш nicht genau definiert haben, nur bemerken, daß der Beweis wieder p-adische L-Funktionen benutzt.

Gross-Zagier (1983): Ist E eine modulare elliptische Kurve mit $r'=1$, so ist $r \geqslant 1$.

Mit anderen Worten, man hat hier im Gegensatz zu dem Satz von Coates und Wiles ein Kriterium dafür, daß die Mordell-Weil-Gruppe *unendlich* ist: das ist bestimmt der Fall, wenn $L(1)=0$ und $L'(1) \neq 0$. Der Beweis beruht auf der Theorie der *Heegner-Punkte*. Das sind gewisse rationale Punkte auf E, die man konstruieren kann, wenn E modular ist und die L-Reihe im Punkt $s=1$ verschwindet. Diese Punkte erhält man auf die folgende Weise. Setzt man in der modularen Parametrisierung $\tau \mapsto P(\tau)=(\xi(\tau), \eta(\tau))$ einen imaginär-quadratischen Wert τ_0 ein, so hat der Punkt $P(\tau_0) \in E$ algebraische Koordinaten, und die Summe P_0 von allen zu $P(\tau_0)$ konjugierten Punkten gehört zu $E(\mathbb{Q})$. In der Arbeit von Gross-Zagier wurde gezeigt, daß die *Höhe* von P_0 (das ist ein bestimmtes Maß für die Größe der Zähler und Nenner der Koordinaten von P_0) proportional zu $L'(1)$ ist. Insbesondere ist (für eine geeignete Wahl von τ_0) P_0 ein Element unendlicher Ordnung von $E(\mathbb{Q})$, falls $L'(1) \neq 0$, womit man die oben angegebene Version des Satzes erhält.

Rubin (1987): Ist E eine CM-Kurve mit $r'=0$, so ist Ш endlich.

Mit anderen Worten, unter denselben Voraussetzungen wie im Satz von Coates-Wiles (CM-Kurve, $L(1) \neq 0$) läßt sich beweisen, daß die Tate-Schafarewitsch-Gruppe von E endlich ist; außerdem liefert Rubins Beweis einen Teil der Behauptung über die Ordnung $|Ш|$ in der BSD-Vermutung. Da wir die Definition von Ш nicht gegeben haben, können wir schwerlich eine Vorstellung von der Rubinschen Beweisidee geben, außer zu sagen, daß Ш eine Art Analogon der Klassengruppe eines algebraischen Zahlkörpers ist und daß Rubins Beweis durch eine neue Methode von F. Thaine zur Abschätzung solcher Klassengruppen inspiriert wurde. Bevor dieser Satz bewiesen wurde, gab es keine einzige elliptische Kurve, für die die Endlichkeit der Tate-Schafarewitsch-Gruppe nachgewiesen worden war!

Kolyvagin (1988): Ist E eine modulare elliptische Kurve mit $r'=0$, so sind $E(\mathbb{Q})$ und Ш endlich (und die BSD-Formel für $L(1)$ ist fast wahr).

Mit anderen Worten, die sehr einschränkende Annahme der komplexen Multiplikation in den Sätzen von Coates-Wiles und Rubin läßt sich durch die viel schwächere Annahme der Modularität ersetzen. Eigentlich hat Kolyvagin den Satz nicht ganz wie oben formuliert bewiesen, sondern er benötigte zusätzlich noch eine gewisse analytische Voraussetzung (nämlich, daß es unter den Kurven, die man erhält, indem man die linke Seite von (1) mit einer rationalen Zahl $d \neq 0$ multipliziert, mindestens eine mit analytischem Rang $r' = 1$ gibt). Es war aber sehr stark anzunehmen, daß diese zusätzliche Bedingung für jede modulare elliptische Kurve erfüllt ist, und diese Annahme ist gerade von Bump-Friedberg-Hoffstein und unabhängig von K. und R. Murty auf analytische Weise bewiesen worden. Kolyvagins genialer Beweis beruht auf kohomologischen Methoden und ist in einem gewissen Sinne erstaunlich einfach.

§ 3 Anwendungen

Wir werden in diesem letzten Abschnitt einige Anwendungen der Theorie der elliptischen Kurven auf klassische Probleme der diophantischen Analysis beschreiben. Natürlich bestehen auch andere Arten von Anwendungen. Insbesondere hat die Theorie der elliptischen Kurven über endlichen Körpern (speziell: die Abschätzung (4) und die Tatsache, daß die Gruppe der Lösungen modulo einer Primzahl p immer entweder zyklisch oder das Produkt zweier zyklischer Gruppen ist) in den letzten Jahren mehrere Anwendungen in der algorithmischen Zahlentheorie gefunden, z. B. die deterministische Bestimmung der Darstellung einer Primzahl $p \equiv 1 \pmod 4$ als Summe zweier Quadratzahlen in polynomialer Zeit (Schoof, 1980), ein nicht-deterministischer, aber häufig sehr schneller Algorithmus zur Faktorisierung großer Zahlen (Lenstra, 1986) und verschiedene hiermit verwandte Algorithmen zur Nachprüfung oder zur „Zertifikation" von großen Primzahlen (Goldwasser-Kilian, 1986, Pomerance, 1987). In einer völlig anderen Richtung liefert die Theorie der ganzzahligen Punkte auf elliptischen Kurven die Aussage, daß 24 die einzige Zahl $n > 1$ ist, für die $1^2 + \ldots + n^2$ ein Quadrat ist; diese Eigenschaft von 24 hängt nach Conway-Sloane mit der Existenz des 24-dimensionalen Leechschen Gitters und somit mit der Erscheinung der Zahl 26 in der Stringtheorie („no-ghost theorem") und vielleicht mit der Dimension des Universums zusammen. Wie gesagt wollen wir uns aber lieber auf rein diophantische Fragen beschränken!

Anwendung 1: Die Fermatsche Tripelgleichung

Eine hübsche Anwendung des Mazurschen Satzes über Torsionspunkte auf elliptischen Kurven ist eine Aussage über die sogenannte Fermatsche Tripelgleichung

$$1 + Ax = \square, \quad 1 + Bx = \square, \quad 1 + Cx = \square \quad (\square = \textit{rationale Quadratzahlen}),$$

wobei A, B, C gegebene, voneinander und von Null verschiedene rationale Zahlen bezeichnen:

Theorem. *Die Tripelgleichung hat stets unendlich viele rationale Lösungen, außer wenn A, B, C* (eventuell nach Vertauschung) *einer der folgenden vier algebraischen Gleichungen genügen:*

(i) $\sqrt{\dfrac{A}{C}}+\sqrt{\dfrac{B}{C}}=1,$ (ii) $\dfrac{A}{C}+\dfrac{B}{C}=1,$

(iii) $\sqrt{\dfrac{C-A}{C}}+\sqrt{\dfrac{C-B}{C}}=1,$

(iv) $\left(1+\sqrt{\dfrac{C-B}{A}}\right)\left(1+\sqrt{\dfrac{C-A}{B}}\right)=2.$

(Bemerkung. Die Umkehrung des Satzes gilt nicht: auch wenn A, B, C einer der vier Gleichungen (i) bis (iv) genügen, kann es vorkommen, daß die Tripelgleichung unendlich viele Lösungen besitzt; so z. B. bei $(A, B, C)=(5, 16, 21)$ oder $(9, 25, 64)$.)

Um das Theorem zu erhalten, betrachtet man die durch $y^2=(Ax+1)(Bx+1)(Cx+1)$ gegebene elliptische Kurve E und den rationalen Punkt $P=(0,1)$ auf E. Man verifiziert direkt, daß dann jeder Punkt der Gestalt $P+2Q$ ($Q\in E(\mathbb{Q})$ beliebig) eine x-Koordinate hat, die die Fermatsche Tripelgleichung erfüllt. Insbesondere bekommt man unendliche viele Lösungen, falls P unendliche Ordnung hat, nämlich die x-Koordinaten von $3P, 5P, 7P, \dots$. Da $E(\mathbb{Q})$ aber mindestens die drei 2-Torsionspunkte $\left(-\dfrac{1}{A}, 0\right), \left(-\dfrac{1}{B}, 0\right), \left(-\dfrac{1}{C}, 0\right)$ besitzt, kann nach dem Mazurschen Satz die Ordnung eines weiteren Torsionspunkts auf E nur 3, 4, 6 oder 8 sein. Diese vier Fälle entsprechen genau den vier im Satz angegebenen speziellen Gestalten des Ausgangstripels (A, B, C). Übrigens hatte schon Fermat die generische Lösung $3P$ entdeckt und bemerkt, daß diese nicht mehr existiert, wenn $A+B=C$.

Anwendung 2: Teillösung des Sylvesterschen Problems

Eine alte und naheliegende Aufgabe der diophantischen Analysis besteht darin, zu entscheiden, ob eine gegebene Zahl n als Summe zweier Kubikzahlen dargestellt werden kann. Dieses Problem ist deswegen schwierig, weil man keine a-priori-Abschätzung für die Größe der dabei auftretenden Zähler und Nenner hat: z. B. ist $n=13$ ganz einfach als $\dfrac{343}{27}+\dfrac{8}{27}$ darstellbar, während die Darstellung von $n=382$ Zahlen von fast 50 Dezimalstellen involviert. Im 19ten Jahrhundert hat J. J. Sylvester vermutet, daß n stets darstellbar ist, falls n prim und $n\equiv 4 \pmod 9$. Diese Behauptung, die aus der Birch-Swinnerton-Dyer-Vermutung folgen würde, ist noch unbewiesen, aber Ph. Satgé konnte 1986 einen ganz ähnlichen Satz beweisen:

Theorem. *Falls* $n/2$ prim *und* $n\equiv 4 \pmod 9$, *so ist* n *als* x^3+y^3 $(x, y\in\mathbb{Q})$ *darstellbar.*

Zum Beispiel ist $22=\left(\frac{17299}{9954}\right)^3+\left(\frac{25469}{9954}\right)^3$. Der Beweis des Theorems beruht auf der Theorie der Modulformen und der Heegner-Punkte: man weiß, daß die elliptische Kurve $X^3+Y^3=2p$ eine Parametrisierung durch Modulformen $X=\xi(\tau)$, $Y=\eta(\tau)$ zuläßt und daß für geeignete quadratische Irrationalitäten τ die Werte dieser Modulformen Zahlen in einem bestimmten algebraischen Zahlkörper („Ringklassenkörper vom Führer p zu $\mathbb{Q}(\sqrt{-3})$") sind; Satgé konnte beweisen, daß man eine nichttriviale rationale Lösung erhält, indem man diese algebraischen Lösungen im Sinne des Additionsgesetzes der Kurve aufsummiert.

Anwendung 3: Bedingte Lösung des Problems der kongruenten Zahlen

Wenn die Vermutung von Birch und Swinnerton-Dyer wahr ist, haben wir die folgende elementare und überraschende Antwort auf die Frage, wann eine vorgegebene Zahl n kongruent ist. Sei n positiv und quadratfrei (o.B.d.A., da mit n auch k^2n für jedes $k\in\mathbb{Q}$ kongruent ist). Ist n ungerade, so definieren wir $A_+(n)$ bzw. $A_-(n)$ als die Anzahl der Tripel $(x,y,z)\in\mathbb{Z}^3$ mit $x^2+2y^2+8z^2=n$ und z gerade bzw. ungerade. Ist n gerade, so definieren wir $A_\pm(n)$ genauso, aber mit $2x^2+2y^2+16z^2$ statt $x^2+2y^2+8z^2$. Dann gilt:

Theorem (unter Annahme der BSD-Vermutung). *Die Zahl n ist genau dann kongruent, wenn $A_+(n)=A_-(n)$.*

Hierbei ist die Richtung „$\Rightarrow$" eine Konsequenz des Satzes von Coates und Wiles; die BSD-Vermutung braucht man nur für die Umkehrung. Diese schöne Anwendung der Theorie der elliptischen Kurven wurde von J. Tunnell 1985 bemerkt. Sie geht wie folgt. Wir haben oben schon gesehen, daß die Zahl n genau dann kongruent ist, wenn die elliptische Kurve $E_n: y^2=x^3-n^2x$ eine nichttriviale Lösung hat. Nach der BSD-Vermutung ist dies genau dann der Fall, wenn $L(E_n,1)=0$, wobei die Richtung „$L(E_n,1)\neq 0 \Rightarrow n$ nicht kongruent" schon aus dem Satz von Coates-Wiles folgt, da E_n eine CM-Kurve ist. Ein Satz von Waldspurger (1984) liefert aber für jede modulare elliptische Kurve E eine Formel für $L(E,1)$ als im wesentlichen das Quadrat eines Fourier-Koeffizienten einer bestimmten Modulform. In unserem Fall ist dieser Koeffizient gleich $A_+(n)-A_-(n)$, und die Behauptung folgt.

Auch ohne Vermutung kann man häufig entscheiden, ob eine Zahl kongruent ist. Ist $A_+(n)\neq A_-(n)$, so ist nach dem oben gesagten n nicht kongruent. Ist dagegen $A_+(n)=A_-(n)$ (also $L(E_n,1)=0$), aber die Ableitung $L'(E_n,1)\neq 0$ – was man ja numerisch nachprüfen kann – so ist n nach dem Satz von Gross-Zagier kongruent. Für die übrigen n findet man in der Praxis einen Punkt auf E_n mit verhältnismäßig kleinen Koordinaten, da nach der BSD-Vermutung der Rang von E_n in diesen Fällen mindestens 2 ist. Auf diese Weise hat man die Kongruenz oder Nichtkongruenz von allen Zahlen bis 2000 in relativ kurzer Rechenzeit bestimmen können, während früher jede einzelne Zahl n eine schwierige Aufgabe darstellte.

Anwendung 4: Lösung des Klassenzahlproblems von Gauß

In den *Disquisitiones Arithmeticae* von Gauß wird jeder negativen Zahl d eine Zahl $h(d)$, die Klassenzahl von d, zugeordnet. Diese Zahl mißt die Abweichung von eindeutiger Primzahlzerlegung in dem imaginärquadratischen Zahlkörper $\mathbb{Q}(\sqrt{d})$. Sie läßt sich auch elementar interpretieren, indem man sagt, daß die Dichte der Primzahlen p, für die $4p$ als $x^2 + dy^2$ darstellbar ist, $\frac{1}{2h(d)}$ beträgt. Gauß stellte die berühmte Vermutung auf, daß die Klassenzahl zusammen mit $|d|$ nach Unendlich strebt, eine Behauptung, die 1934 von Heilbronn (nach Vorarbeiten von Deuring und Mordell) bewiesen wurde. Der Beweis war aber *ineffektiv*: das heißt, er gab nicht einmal im Prinzip die Möglichkeit, die Liste aller Zahlen d mit einem gegebenen Wert der Klassenzahl $h(d)$ zu bestimmen. Die Spezialfälle $h(d)=1 \Rightarrow |d| \leqslant 163$, $h(d)=2 \Rightarrow |d| \leqslant 427$ wurden durch Arbeiten von Heegner (1952), Baker (1969) und Stark (1969) bewiesen, aber der allgemeine Fall blieb offen. Im Jahre 1975 bewies D. Goldfeld durch ein ingeniöses analytisches Argument eine effektive untere Abschätzung für $h(d)$, die mit $|d|$ nach Unendlich strebt. Er brauchte aber hierfür die Existenz einer elliptischen Kurve, deren L-Reihe bei $s=1$ eine Nullstelle mindestens dritter Ordnung hat (und noch eine weitere Eigenschaft besitzt). Erst 1983 konnte man mit Hilfe des Satzes von Gross-Zagier beweisen, daß etwa die elliptische Kurve $-139y^2 = x^3 + 10x^2 - 20x + 8$ die erforderlichen Eigenschaften besitzt: einerseits ist diese Kurve modular und hat einen ungeraden analytischen Rang r', andererseits verschwindet der Heegner-Punkt, der nach dem Satz unendliche Ordnung haben müßte, wenn r' gleich 1 wäre; also ist $r' \geqslant 3$. Somit haben die elliptischen Kurven den letzten Schritt zur Lösung eines klassischen zahlentheoretischen Problems beitragen können.

Der Spezialfall „$h(d)=4 \Rightarrow |d| \leqslant 1555$“ hat übrigens eine schöne Anwendung, die von E. Grosswald bemerkt wurde:

Theorem. *Die einzigen Zahlen, die eine eindeutige Darstellung als Summe dreier Quadratzahlen zulassen, sind* 1, 2, 3, 5, 6, 10, 11, 13, 14, 19, 21, 22, 30, 35, 37, 42, 43, 46, 58, 67, 70, 78, 91, 93, 115, 133, 142, 163, 190, 235, 253, 403 *und* 427 *und die Produkte von diesen Zahlen mit Potenzen von* 4.

Die Beziehung zwischen den beiden Fragestellungen rührt von dem Satz von Gauß-Hermite her, welcher die Anzahl der Darstellungen einer beliebigen natürlichen Zahl als Summe dreier Quadratzahlen mittels Klassenzahlen ausdrückt.

Anwendung 5: Bedingter Beweis des großen Theorems von Fermat

Wie wir gesehen haben, ist die Taniyama-Weil-Vermutung, daß jede elliptische Kurve über $\mathbb{Q}$ modular ist, eine Aussage, die auf einigermaßen überzeugender theoretischer und numerischer Evidenz basiert. 1986 bewies K. Ribet:

Theorem. *Gilt die Vermutung von Taniyama-Weil, so gilt auch der „letzte Satz von Fermat“: es gibt keine natürliche Zahlen A, B, C und $n > 2$ mit $A^n + B^n = C^n$.*

Um dieses überraschende Resultat zu erhalten, nimmt man an, es gäbe solche Zahlen, und betrachtet die spezielle elliptische Kurve $E_F: y^2 = x(x-A^n)(x-C^n)$, die für einen anderen Zweck von Y. Hellegouarch und in diesem Zusammenhang von G. Frey eingeführt wurde. Ribet konnte einen Teil einer Vermutung von J-P. Serre beweisen, wonach jede modulare elliptische Kurve eine bestimmte Eigenschaft haben muß, die die Kurve E_F nachweislich nicht besitzt. Somit wäre die Kurve E_F ein Gegenbeispiel zu der Vermutung von Taniyama und Weil. Wir können die Situation mit Hilfe unseres Venn-Diagramms auf folgende Weise veranschaulichen:

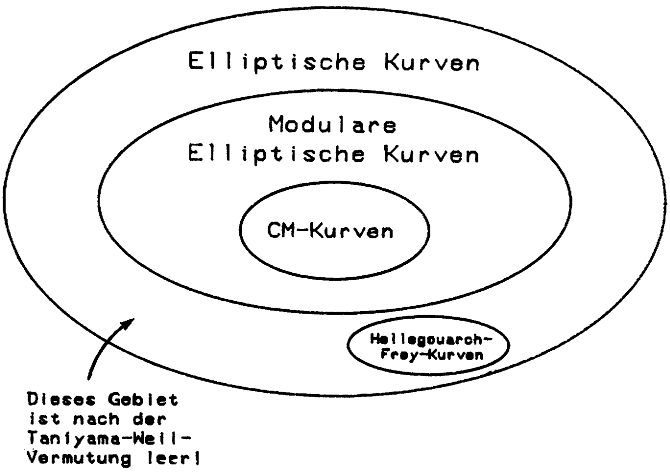


Fig. 5

Literatur

Für den Leser, der näheres zu den hier angeschnittenen Themen erfahren möchte, gebe ich einige Hinweise auf die Literatur. Außerdem werden die Arbeiten angegeben, deren Ergebnisse im Text vorkommen, auch wenn ihr Niveau weit über das dieser Einführung hinausgeht.

Zu § 1: Die Lehrbücher von J. Silverman (*The Arithmetic of Elliptic Curves*, Graduate Texts in Math. 106, Springer, 1985), D. Husemöller (*Elliptic Curves*, Graduate Texts in Math. 111, Springer, 1987) und J. Chahal (*Topics in Number Theory*, Univ. Series in Math., Plenum, 1988) sind alle drei als allgemeine Einführung in die Theorie der elliptischen Kurven sehr zu empfehlen. Über die Beiträge von Diophant kann man in dem schönen Büchlein *Diophant und diophantische Gleichungen* von I. G. Bashmakova (DVW, 1974) oder in B. L. van der Waerden, *Geometry and Algebra in Ancient Civilizations* (Springer, 1983) lesen. Zu der Geschichte spezieller diophantischer Gleichungen vor 1920 siehe Band II von L. E. Dickson, *History of the Theory of Numbers* (Stechert, 1934, und Chelsea, 1952); z. B. enthält Dickson viel Auskunft über die Werke von Fermat und Diophant und einen ganzen Abschnitt über das Problem der kongruenten Zahlen.

Mazurs Satz über die möglichen Torsionsgruppen von Mordell-Weil-Gruppen über $\mathbb{Q}$ wird in seinen (sehr schwierigen) Arbeiten „Modular curves and the Eisenstein ideal," IHES Publ. Math. **47** (1977) 33–186, und „Rational points on modular curves," *Modular Functions of One Variable V*, Lecture Notes in Math. **601** (1977) 107–148, bewiesen. Ein Beispiel einer elliptischen Kurve mit Rang (mindestens, und nach der BSD-Vermutung genau) 14 wird in J.-F. Mestre, „Formules explicites et minorations de conducteurs de variétés algébriques," Compositio Mathematica **58** (1986) 209–232, vorgestellt. Die Rechnungen zum Rang der Kurven $x^3+y^3=m$ befinden sich in D. Zagier and G. Kramarz, „Numerical investigations related to the L-series of certain elliptic curves," J. Ind. Math. Soc. **52** (1987) 1–19.

Zu § 2: Die – durchgehend schwierigen – Originalarbeiten zu diesem Abschnitt sind wie folgt: J. Silverman, „A quantitative version of Siegel's theorem: integral points on elliptic curves and Catalan curves," J. Reine Angew. Math. **378** (1987) 60–100; N. Elkies, „The existence of infinitely many supersingular primes for any elliptic curve over $\mathbb{Q}$," Invent. Math. **89** (1987) 561–567; J. Coates and A. Wiles, „On the conjecture of Birch and Swinnerton-Dyer," Invent. Math. **39** (1977) 223–251; R. Greenberg, „On the Birch and Swinnterton-Dyer conjecture," Invent. Math. **72** (1983) 241–265; B. Gross and D. Zagier, „Heegner points and derivatives of L-series," Invent. Math. **84** (1986) 225–320; K. Rubin, „Tate-Shafarevich groups and L-functions of elliptic curves with complex multiplication," Invent. Math. **89** (1987) 527–560; B. A. Колывагин, Конечность $E(\mathbb{Q})$ и $Ш(E,\mathbb{Q})$ для подкласса кривых Вейля, Изв. Акад. Наук СССР **52** (1988) 522–540. Ein Überblick über einige dieser Ergebnisse, und speziell über die Sätze von Rubin und Kolyvagin, wird von L. Washington in „Number fields and elliptic curves" (in: *Number Theory and Applications*, Proc. NATO Advanced Study Institute, Banff 1988, NATO ASI series C 265 (1989) 245–278) gegeben. Eine Exposition des Satzes von Kolyvagin gibt K. Rubin in „The work of Kolyvagin on the arithmetic of elliptic curves" (in *Arithmetic of Complex Manifolds, Proceedings* 1988, Lecture Notes in Math. 1399, Springer, 1989).

Zu § 3: Die Anwendung der Theorie der elliptischen Kurven über endlichen Körpern auf die Darstellung einer Primzahl als Summe zweier Quadrate ist in R. Schoof, „Elliptic curves over finite fields and the computation of square roots mod p," Math. Comp. **44** (1985) 483–494, enthalten. Für die Anwendungen auf Primalitätstests und Faktorisierungsalgorithmen siehe den Übersichtsartikel von H. Lenstra, „Elliptic curves and number-theoretic algorithms," Proc. ICM Berkeley 1986, 99–120.

Zur Fermatschen Tripelgleichung siehe J. Leech, „Four integers whose twelve quotients sum to squares," Can. J. Math. **38** (1986) 1261–1280. Leech macht auch die Bemerkung, daß es außer in Ausnahmefällen immer unendlich viele Lösungen der Fermatschen Tripelgleichung gibt, begründet dies aber unzureichend und ohne den Fall (iv) zu erwähnen.

Der Satz von Ph. Satgé über die Darstellbarkeit von $2p$ als Summe zweier Kubikzahlen ist enthalten in seiner Arbeit „Un analogue du calcul de Heegner," Invent. Math. **87** (1987) 425–439.

Der Satz von J.-L. Waldspurger über Werte von $L(1)$ wird in seiner Arbeit „Sur les coefficients de Fourier des formes modulaires de poids demi-entier," J. Math. Pures Appl. **60** (1981) 375–484 bewiesen. Die Arbeit von J. Tunnell, in der er diesen Satz auf das Problem der kongruenten Zahlen anwendet, ist „A classical Diophantine problem and modular forms of weigth 3/2," Invent. Math. **72** (1983) 323–334. Eine Einführung in die Theorie der elliptischen Kurven und Modulformen, die diese Anwendung als Endziel setzt, ist das Buch *Introduction to Elliptic Curves and Modular Forms* von N. Koblitz, Graduate Texts 97, Springer, 1984. Die im Text erwähnten numerischen Rechnungen bis $n=2000$ sind in G. Kramarz, „All congruent numbers less than 2000," Math. Ann. **273** (1986) 337–340 enthalten.

Die ursprüngliche Arbeit von D. Goldfeld zum Klassenzahlproblem ist „The class numbers of quadratic fields and the conjectures of Birch and Swinnerton-Dyer," Ann. Scuola Norm. Sup. Pisa **3** (1976) 623–663. Eine Exposition seiner Methode und der Anwendung des Gross-Zagierschen Satzes gibt J. Oesterlé in „Nombres de classes des corps quadratiques imaginaires," Sém. Bourbaki 1983–1984, Exposé 631, Astérisque **121–122** (1985) 309–323, oder (weniger detailliert und weniger technisch) D. Zagier in „L-series of elliptic curves, the Birch-Swinnterton-Dyer conjecture, and the class number problem of Gauss," Notices A.M.S. **31** (1984) 739–743. Die Bemerkung, daß man mit der Bestimmung aller Diskriminanten mit Klassenzahlen 1, 2 oder 4 auch die Liste aller eindeutig als Summe dreier Quadratzahlen darstellbaren Zahlen kennt, steht bei P. T. Bateman und E. Grosswald, „Positive integers expressible as a sum of three squares in essentially only one way," J. Number Theory **19** (1984) 301–308 und in Grosswalds schönem Buch *Representations of Integers as Sums of Squares* (Springer, 1987). Der Nachweis, daß -1555 die letzte Diskriminante mit der Klassenzahl 4 ist, wurde in der Doktorarbeit „Class Number 4" von S. Arno (Stanford, 1986) erbracht.

Für die Eigenschaften der Hellegouarch-Frey-Kurven und ihre Anwendung auf den letzten Satz von Fermat siehe G. Frey, „Links between stable elliptic curves and certain diophantine equations," Ann. Univ. Saraviensis **1** (1986) 1–40. Der im Text zitierte Satz von Ribet wird in seiner (schweren) Arbeit „On modular representations of Gal($\bar{\mathbb{Q}}/\mathbb{Q}$) arising from modular forms" (erscheint in Invent. Math., 1990) bewiesen.

Don Zagier
University of Maryland
College Park, MD 20742
USA
Max-Planck-Institut für Mathematik
Gottfried-Claren-Straße 26
5300 Bonn 3

(*Eingegangen 10. 1. 1989, leicht revidiert 26. 4. 1989*)

Jber. d. Dt. Math.-Verein. 92 (1990) 77–88
AMS subject classification: 73–02, 65 N 30

Theoretical and Practical Aspects of Finite Elements in the Context of Some Problems of Solid Mechanics

J. R. Whiteman*), A. E. Beagles and M. K. Warby, Uxbridge

1 Introduction and Use of Finite Element Methods in Solid Mechanics

This paper is concerned with the use of computational modelling to produce approximations to the solutions of certain linear and nonlinear problems of solid mechanics. This type of computational modelling consists of two distinct steps, the first of which involves the production, using the laws of physics, of a mathematical model representing the physical situation. This is followed by the second step in which the mathematical model is discretised using a numerical technique, so that an algorithm can be developed and numerical solutions calculated.

The range of problems from solid mechanics considered here is necessarily limited but does include some linear and nonlinear steady state problems as well as some time dependent problems. There is a vast difference between the state-of-the-art mathematical theory available for linear and for nonlinear problems, which correspondingly has an effect on the range of theoretical error estimates which can be derived for the numerical solutions of the respective classes of problems. The discussion covers some problems of linear elasticity (small deformation, small strain), elasto-plasticity (nonlinear material behaviour, small deformation), visco-elasticity (time dependent material behaviour, small deformation). In each case the mathematical model is based on a system of equilibrium equations, together with constitutive equations involving material parameters which describe the material behaviour. The numerical technique employed here is exclusively that of finite elements, as this is now the most widely used numerical technique in computational mechanics. An immense literature exists on the theory and application of the method throughout continuum mechanics and many textbooks have been written on the subject; see e.g. [2], [8], [10], [23], [29], [32].

The aims of the paper are as follows: to describe the forms of the mathematical models of the problems, indicating the major differences between

*) Invited plenary lecturer, 1988 Annual Meeting of the Deutsche Mathematiker-Vereinigung (DMV), Regensburg, September 1988

the various contexts and explaining what assumptions have been made in these concerning the physics; to describe the application of the finite element method using Galerkin techniques for each problem and to give some theoretical finite element error estimates where this in possible. In the context of error estimates the phenomenon of superconvergence will be introduced.

2 Equilibrium Problem and Galerkin Approximation

We consider first the deformation of a solid body, defined in a region $\Omega \subset \mathbb{R}^3$ with boundary $\partial\Omega$, under the action of external forces. The displacement at any point $\boldsymbol{x} \equiv (x_1, x_2, x_3)^{\mathrm{T}}$ of Ω (the reference configuration) is denoted by $\boldsymbol{u} \equiv (u_1, u_2, u_3)^{\mathrm{T}}$, and the stress and strain tensor components are denoted respectively by σ_{ij} and ε_{ij}, $1 \leqslant i,j \leqslant 3$. (Whilst the description of this section is given for the three-space dimension case, the simplification to two space dimensions (plane strain) is obvious; the discussion of superconvergence will be in a two-dimensional context.)

Based on the law of conservation of momentum for the steady state case (when there is no acceleration), see e.g. [14], [28], the deformation of the body under the action of external forces $\boldsymbol{f} \equiv (f_1, f_2, f_3)^{\mathrm{T}}$ and boundary tractions $\boldsymbol{g} \equiv (g_1, g_2, g_3)^{\mathrm{T}}$ satisfies the equilibrium equation system

$$(2.1) \qquad \sum_{j=1}^{3} \frac{\partial \sigma_{ij}(\boldsymbol{x})}{\partial x_j} + f_i(\boldsymbol{x}) = 0, \qquad \boldsymbol{x} \in \Omega, \quad i = 1, 2, 3,$$

together with boundary conditions

$$(2.2) \qquad \boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{0}, \qquad \boldsymbol{x} \in \partial\Omega_c$$

$$(2.3) \qquad \sum_{j=1}^{3} \sigma_{ij}(\boldsymbol{x}) \cdot n_j = g_i(\boldsymbol{x}), \qquad \boldsymbol{x} \in \partial\Omega_T, \quad i = 1, 2, 3,$$

where $\partial\Omega \equiv \partial\Omega_c \cup \partial\Omega_T$ is a partition of the boundary $\partial\Omega$.

The finite element method is applied to problem (2.1)–(2.3) via a weak formulation and the Galerkin technique. This is obtained by first taking the scalar product of (2.1) with a test vector function $\boldsymbol{v} \in V$, where

$$V \equiv \{\boldsymbol{v} : \boldsymbol{v} \in (H^1(\Omega))^3,\ v_i|_{\partial\Omega_c} = 0,\ i = 1, 2, 3\}$$

is the space of admissible displacement vectors, and then integrating by parts. Thus in the weak problem we seek $\boldsymbol{u} \in V$ such that

$$(2.4) \qquad \int_\Omega \boldsymbol{\sigma}^{\mathrm{T}} \cdot \boldsymbol{\varepsilon}(\boldsymbol{v}) \mathrm{d}\boldsymbol{x} - \int_\Omega \boldsymbol{f}^{\mathrm{T}} \cdot \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} - \int_{\partial\Omega_T} \boldsymbol{g}^{\mathrm{T}} \cdot \boldsymbol{v} \, \mathrm{d}s = 0, \qquad \forall\, \boldsymbol{v} \in V,$$

where in (2.4) the tensor components ε_{ij} are defined by

$$(2.5) \qquad \varepsilon_{ij}(\boldsymbol{v}) \equiv \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right), \qquad i, j = 1, 2, 3,$$

the vectors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$ are given by $\boldsymbol{\varepsilon}=(\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}, 2\varepsilon_{12}, 2\varepsilon_{23}, 2\varepsilon_{31})^{\mathrm{T}}$, $\boldsymbol{\sigma}=(\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{23}, \sigma_{31})^{\mathrm{T}}$ and the displacement $\boldsymbol{u}$ is involved in (2.4) through $\boldsymbol{\sigma}$ via an appropriate constitutive relation.

For the application of the finite element method the region Ω is partitioned into elements $\Omega = \bigcup_e \Omega^e$, with generic mesh size h. A finite dimensional space $S^h \subset V$ consisting of piecewise polynomial functions defined over the partition is set up and the Galerkin problem approximating (2.4) is that of finding $\boldsymbol{u}_h \in S^h$ such that

$$(2.6) \qquad \int_\Omega \boldsymbol{\sigma}_h^{\mathrm{T}} \cdot \boldsymbol{\varepsilon}(\boldsymbol{v}_h)\mathrm{d}x - \int_\Omega \boldsymbol{f}^{\mathrm{T}} \cdot \boldsymbol{v}_h\, dx - \int_{\partial\Omega_T} \boldsymbol{g}^{\mathrm{T}} \cdot \boldsymbol{v}_h\, \mathrm{d}s = 0 \qquad \forall\, \boldsymbol{v}_h \in S^h,$$

where, in a similar manner to (2.4), the approximation $\boldsymbol{u}_h$ to the displacement $\boldsymbol{u}$ is involved in (2.6) through $\boldsymbol{\sigma}_h$ via an approximation to an appropriate constitutive relation.

Each component of the approximating vector $\boldsymbol{u}_h(\boldsymbol{x})$ is defined in terms of basis functions $N_k(\boldsymbol{x})$ for the n nodes of Ω so that, in terms of point evaluations $\boldsymbol{U}_k$ of $\boldsymbol{u}_h(\boldsymbol{x})$ at the nodes $k=1,2,\ldots,n$,

$$(2.7) \qquad \boldsymbol{u}_h(\boldsymbol{x})=\boldsymbol{N}(\boldsymbol{x})\boldsymbol{U},$$

where $\boldsymbol{N}(\boldsymbol{x})=[\boldsymbol{N}_1, \boldsymbol{N}_2(\boldsymbol{x}),\ldots,\boldsymbol{N}_n(\boldsymbol{x})]$ with $\boldsymbol{N}_k(\boldsymbol{x})=N_k(\boldsymbol{x})\boldsymbol{I}_3$ and $\boldsymbol{I}_3$ is the 3×3 unit matrix. If we define the approximate strain

$$(2.8) \qquad \boldsymbol{\varepsilon}_h \equiv \boldsymbol{\varepsilon}(\boldsymbol{u}_h(\boldsymbol{x})) \equiv \boldsymbol{B}\,\boldsymbol{U},$$

in the usual way, see e.g. Zienkiewicz [32], we now require a constitutive relation between $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ which is of course dependent on the mechanical behaviour of the material of the solid.

3 Linear Elasticity

The behaviour of most solids conforms well to a linear elastic model under the following two restrictions; (a) at each point the stress, and hence the strain, is sufficiently small, and (b) time dependent effects are negligible. We consider now the linear elastic situation and then in two subsequent sections describe the case of plasticity, where (a) does not hold, and of viscoelasticity, where (b) does not hold.

3.1 Linear Elastic Constitutive Relation and Galerkin Technique

For a linear elastic material, by Hooke's law, the constitutive relation can be expressed in the form

$$(3.1) \qquad \boldsymbol{\sigma}(\boldsymbol{u}(\boldsymbol{x}))=\boldsymbol{D}\,\boldsymbol{\varepsilon}(\boldsymbol{x}),$$

where $\boldsymbol{D}$ is a 6×6 matrix. In the case of an isotropic material this relation takes the form

$$(3.2) \qquad \sigma_{ij}(\boldsymbol{u}) = \lambda\left(\sum_{k=1}^{3} \varepsilon_{kk}(\boldsymbol{u})\right)\sigma_{ij} + 2\mu\,\varepsilon_{ij}(\boldsymbol{u}), \qquad 1 \leqslant i,j \leqslant 3,$$

where $\lambda, \mu > 0$ are the Lamé coefficients. When these relations are substituted into the equilibrium equations (2.1) the well-known Lamé equations

$$(3.3) \qquad -\mu \Delta \boldsymbol{u} - (\lambda + \mu) \operatorname{grad} \operatorname{div} \boldsymbol{u} = \boldsymbol{f}$$

are obtained, whilst the weak problem (2.4) becomes that of finding $\boldsymbol{u} \in V$ such that

$$(3.4) \qquad a(\boldsymbol{u}, \boldsymbol{v}) = F(\boldsymbol{v}) \qquad \forall\, \boldsymbol{v} \in V,$$

where

$$(3.5) \qquad a(\boldsymbol{u}, \boldsymbol{v}) \equiv \int_{\Omega} \left\{ \lambda \operatorname{div} \boldsymbol{u} \operatorname{div} \boldsymbol{v} + 2\mu \sum_{i,j=1}^{3} \varepsilon_{ij}(\boldsymbol{u}) \varepsilon_{ij}(\boldsymbol{v}) \right\} \mathrm{d}\boldsymbol{x}$$

and

$$(3.6) \qquad F(\boldsymbol{v}) \equiv \int_{\Omega} \boldsymbol{f}^{\mathrm{T}} \cdot \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} + \int_{\partial\Omega_T} \boldsymbol{g}^{\mathrm{T}} \cdot \boldsymbol{v} \, \mathrm{d}s.$$

Application of the finite element method as in Section 2 produces from problem (3.4) the approximating problem of finding $\boldsymbol{u}_h \in S^h$ such that

$$(3.7) \qquad a(\boldsymbol{u}_h, \boldsymbol{v}_h) = F(\boldsymbol{v}_h) \qquad \forall\, \boldsymbol{v}_h \in S^h.$$

In (3.7) $\boldsymbol{u}_h$ has the form of (2.7), so that if $\boldsymbol{v}_h$ is taken in turn to be each column of $\boldsymbol{N}$ we obtain the linear equation system

$$\int_{\Omega} (\boldsymbol{B}^{\mathrm{T}} \boldsymbol{D} \, \boldsymbol{B} \, \mathrm{d}\boldsymbol{x}) \boldsymbol{U} = \int_{\Omega} \boldsymbol{f}^{\mathrm{T}} \cdot \boldsymbol{N} \, \mathrm{d}\boldsymbol{x} + \int_{\partial\Omega_T} \boldsymbol{g}^{\mathrm{T}} \cdot \boldsymbol{N} \, \mathrm{d}s,$$

which we write as

$$(3.8) \qquad \boldsymbol{K} \boldsymbol{U} = \boldsymbol{F}$$

and which when solved produces $\boldsymbol{U}$ and hence $\boldsymbol{u}_h(\boldsymbol{x})$. The system (3.8) is the global stiffness equation system and $\boldsymbol{K}$ is the global stiffness matrix. Direct methods based on some form of Gaussian elimination have in the past mainly been used for solving (3.8); [27], [32]. However, recently conjugate gradient methods have proved successful and, with the arrival of parallel computers, iterative methods are gaining in popularity.

3.2 Finite Element Error Analysis and Superconvergence

The bilinear form $a(\boldsymbol{u}, \boldsymbol{v})$ of (3.5) is continuous on V, whilst $F(\boldsymbol{v})$ of (3.6) is a linear functional on V. It has also been shown using Korn's inequality, see Fichera [11], and the fact that $\lambda, \mu > 0$ that $a(\boldsymbol{u}, \boldsymbol{v})$ is V-elliptic, see also Ciarlet [8].

It is worth noting that in this case an alternative formulation for the problem of linear elasticity can be defined in terms of potential energy. The well known theorem stating that the solution of the weak problem (3.4) is also the vector from V which minimises the potential energy functional

$$(3.9) \qquad I[\boldsymbol{v}] = \frac{1}{2} a(\boldsymbol{v}, \boldsymbol{v}) - F(\boldsymbol{v}),$$

is given e.g. in Whiteman [30].

The properties of the bilinear form $a(\boldsymbol{u}, \boldsymbol{v})$ and of the linear functional $F(\boldsymbol{v})$ have been very useful in the development of a comprehensive error analysis for the Galerkin approximation $\boldsymbol{u}_h$ to the solution $\boldsymbol{u}$ of (3.4), see Ciarlet [8] and Oden and Reddy [23].

Firstly it can be shown that $\boldsymbol{u}_h \in S^h$ is the best approximation to $\boldsymbol{u}$ in the sense that

$$(3.10)\quad \|\boldsymbol{u}-\boldsymbol{u}_h\|_{V,\Omega} \leqslant C\|\boldsymbol{u}-\boldsymbol{v}_h\|_{V,\Omega} \qquad \forall\, \boldsymbol{v}_h \in S^h,$$

where C is a constant; the constants in all inequalities will here be denoted by C. Since (3.10) holds for all $\boldsymbol{v}_h \in S_h$, this includes the interpolant $\boldsymbol{u}_I \in S^h$ to $\boldsymbol{u}$, so that estimates for $\|\boldsymbol{u}-\boldsymbol{u}_h\|_{.,\Omega}$ can be derived using approximation theory, see e.g. Ciarlet and Raviart [9]. Such estimates have the form

$$(3.11)\quad \|\boldsymbol{u} - \boldsymbol{u}_h\|_{V,\Omega} \leqslant C\frac{h^{\gamma}}{p^{k-1}}\,|\boldsymbol{u}|_{k,\Omega},$$

where γ is a positive constant depending on the order p of the polynomials used in vectors of S^h and on k. Whilst p can be chosen, the regularity of the solution $\boldsymbol{u}$ determines the largest value of k and thus controls γ. *A priori* estimates of the type (3.11) guarantee the convergence with decreasing mesh size of $\boldsymbol{u}_h$ to $\boldsymbol{u}$. The rate of convergence γ, decreases as $\boldsymbol{u}$ becomes less smooth; for example γ is small when the problem (3.8) contains a boundary singularity. The estimate (3.11) is in terms of the *energy* norm, $\|.\|_V$, and as such is a *global* estimate. By use of the Nitsche lift, see [20], an estimate of the L_2-norm $\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,\Omega}$ can be obtained from (3.11), and this global estimate has a higher rate of convergence. Maximum norm estimates have also been derived in this context, based on the original work of Natterer [18] and Nitsche [20], and have been localised onto interior domains by Nitsche and Schatz [21].

The energy norm estimate (3.11) involves gradients of $\boldsymbol{u}$ and $\boldsymbol{u}_h$. In recent years the phenomenon of *recovered gradient superconvergence* of piecewise linear finite element approximations to the solutions of planar second order elliptic boundary value problems has sprung to prominence. In this context the superconvergence effect arises from the fact that there exist in each element certain points at which the gradient, or a component of the gradient, of the finite element approximation has a higher rate of convergence to the gradient of the true solution than that found globally. Although the superconvergence effect has to date largely been analysed mathematically in the context of two-dimensional *scalar* problems, see the survey article of Krizek and Neittaanmäki [16] and the 200 references contained therein, it has been extended to planar linear elasticity by Goodsell and Whiteman [12], [31]. The first generalisation of three space dimensions, again for scalar problems, has been given by Andreev and Lazarov [1] using partitions of tetrahedral elements. With simplex elements (triangles, tetrahedra, ...) it is necessary to use retrieval techniques to post-process the calculated solutions in order to obtain gradient superconvergence. In the context of linear elasticity, although the primary unknowns (the displacements $\boldsymbol{u}$) are approximated in the main calculation, it is approximations to secondary variables (strains and stresses) which are often of main interest and it is to these variables that the (gradient) superconvergence effect applies.

Adaptive methods and *a posteriori* error estimates have come increasingly to the fore in recent years, to a large extent due to the pioneering work of Babuska and Rheinboldt, see e.g. [5]. In this type of work an error estimate is calculated in a suitable norm for each element and by amalgamation of these for all the elements a global estimate is derived. The local element estimates can be used to produce adaptive mesh refinement schemes under which local refinement is performed in subdomains where large errors occur until the estimated error in each element of the partition is less than a specified value, see [3], [4], [22]. The purpose of this type of error estimation is to try to answer the question "What is the actual error in the calculated solution to a problem on a specific mesh?" Adaptive techniques and some form of error estimation are becoming features of the finite element calculations of the present day.

4 Elasto-plasticity

Uniaxial tests on most metals and alloys indicate that, although they deform elastically under the action of small stresses, when they are subjected to stresses greater than a material dependent *yield stress* the materials experience irreversible (*plastic*) strains. It is this phenomenon of plasticity which we now wish to model mathematically and computationally.

4.1 Incremental plasticity and finite element algorithm

In order to provide a mathematical model for the case where the material of the solid exhibits an elasto-plastic response, we have to set up a model of the constitutive relationship between stress and strain appropriate to the nonlinear post-yield plastic case.

We adopt here he incremental (flow theory) of plasticity and apply the loading incrementally. Thus in (2.1)–(2.3) we consider increments $d\boldsymbol{\sigma}$, $d\boldsymbol{\varepsilon}$, $d\boldsymbol{u}$, respectively of stress, strain and displacement, which result from increments of loading $d\boldsymbol{f}$ and $d\boldsymbol{g}$. The displacement $\boldsymbol{u}$ is now a function not only of space but also of the current load. We therefore introduce a load factor t (fraction of total load), so that $\boldsymbol{u}=\boldsymbol{u}(\boldsymbol{x},t)$. In the usual manner, see e.g. Owen and Hinton [25], Harrison, Ward and Whiteman [13], the level of stress at which plastic deformation takes place is determined by a yield criterion, based on a yield function F,

$$(4.1)\quad F(\sigma,k)=f(\sigma)-k\leqslant 0,$$

where f is the equivalent stress function and k varies during plastic deformation so that $k=f(\sigma)$ and $F(\sigma,k)=0$. For any load increment, after initial yielding, it is assumed that the increment of strain can be written as the sum of elastic and plastic components so that

$$(4.2)\quad \mathrm{d}\varepsilon=\mathrm{d}\varepsilon_e+\mathrm{d}\varepsilon_p,$$

where $\mathrm{d}\varepsilon_e$ is related to $\mathrm{d}\boldsymbol{\sigma}$ by the $\boldsymbol{D}$ matrix of (3.1). The plastic flow of the material is governed by a flow rule which, for associative plasticity, relates the increment of plastic strain to the gradient of the yield function, so that

$$\text{(4.3)} \quad d\varepsilon_p = d\lambda \frac{\partial F}{\partial \boldsymbol{\sigma}},$$

where $d\lambda$ is the plastic multiplier. It is related to the k of (4.1) through a hardening rule, $d\lambda = A dk$. When a state of plastic flow exists stresses must remain on the yield surface so that

$$dF = \left(\frac{\partial F}{\partial \boldsymbol{\sigma}}\right)^{\mathrm{T}} d\boldsymbol{\sigma} + \frac{\partial F}{\partial k} dk = 0$$

and hence

$$\text{(4.4)} \quad \boldsymbol{a}^{\mathrm{T}} d\boldsymbol{\sigma} - A d\lambda = 0,$$

where the flow vector $\boldsymbol{a}$ is defined by

$$\boldsymbol{a} = \frac{\partial F}{\partial \boldsymbol{\sigma}}.$$

We thus obtain from (4.1)–(4.4) the relation

$$d\boldsymbol{\varepsilon} = \left(\boldsymbol{D}^{-1} + \frac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{A}\right) d\boldsymbol{\sigma},$$

from which we obtain

$$\text{(4.5)} \quad d\boldsymbol{\sigma} = \left(\boldsymbol{D} - \frac{\boldsymbol{D}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{D}}{A + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{a}}\right) d\boldsymbol{\varepsilon} \equiv \boldsymbol{D}_{\mathrm{ep}} d\boldsymbol{\varepsilon},$$

where $\boldsymbol{D}_{\mathrm{ep}} \equiv \boldsymbol{D}_{\mathrm{ep}}(\boldsymbol{\sigma}, k)$ is the elasto-plastic constitutive matrix. Thus, for the load increment in the post yield state, we have a nonlinear constitutive relation.

If the Galerkin technique is applied to the elasto-plastic problem in a specific incremental load step, the resulting (*nonlinear*) global equation system corresponding to (3.9) is

$$\text{(4.6)} \quad \int_{\Omega} (\boldsymbol{B}^{\mathrm{T}}\boldsymbol{D}_{\mathrm{p}}\boldsymbol{B}\, dx)\, d\boldsymbol{U} - \int_{\Omega} d\boldsymbol{f}^{\mathrm{T}} \cdot \boldsymbol{N}\, dx - \int_{\partial\Omega_T} d\boldsymbol{g}^{\mathrm{T}} \cdot \boldsymbol{N}\, ds = \boldsymbol{0},$$

where $\boldsymbol{D}_{\mathrm{p}} \equiv \boldsymbol{D}_{\mathrm{ep}}$ during yielding and $\boldsymbol{D}_{\mathrm{p}} \equiv \boldsymbol{D}$ otherwise. When yield takes place the system (4.6) is nonlinear and is solved iteratively within the load step; the two most used methods for this are the *initial stiffness* and the *tangent stiffness* methods. If we define

$$\text{(4.7)} \quad \boldsymbol{K}(\boldsymbol{\sigma}, k) \equiv \int_{\Omega} \boldsymbol{B}^{\mathrm{T}} \boldsymbol{D}_{\mathrm{p}} \boldsymbol{B}\, dx,$$

then the iteration for solving (4.6), with general iteration step i, is as follows:

Step 1. Set $i = 1$, $d\boldsymbol{U}_1 = \boldsymbol{0}$, and take $\boldsymbol{\sigma}_1$ and k_1 to be their final values from the previous load step. Define

$$\boldsymbol{R}_1^{\mathrm{T}} \equiv \int_{\Omega} d\boldsymbol{f}^{\mathrm{T}} \cdot \boldsymbol{N}\, dx + \int_{\partial\Omega^T} d\boldsymbol{g}^{\mathrm{T}} \cdot \boldsymbol{N}\, ds.$$

Step 2. Set $\mathrm{d}\boldsymbol{U}_{i+1} = \mathrm{d}\boldsymbol{U}_i + \boldsymbol{H}^{-1}\boldsymbol{R}_i$.
Calculate $\boldsymbol{\sigma}_{i+1}$, k_{i+1}, A_{i+1} and $\boldsymbol{R}_{i+1} \equiv \boldsymbol{R}_1 - \boldsymbol{K}(\boldsymbol{\sigma}_{i+1}, k_{i+1})\mathrm{d}\boldsymbol{U}_{i+1}$.

Step 3. For some tolerance ε if $\|\boldsymbol{R}_{i+1}\| > \varepsilon\|\boldsymbol{R}_1\|$ then set $i := i+1$ and repeat 2. Otherwise set $\mathrm{d}\boldsymbol{U} := \mathrm{d}\boldsymbol{U}_{i+1}$ and stop the iteration.

In Step 2 the matrix $\boldsymbol{H}$ can be taken as $\boldsymbol{K}(\boldsymbol{0}, 0)$ giving the initial stiffness method, or as the matrix $\boldsymbol{K}(\boldsymbol{\sigma}_i, k_i)$ from the previous iteration step giving the tangent stiffness method. The values of $\boldsymbol{\sigma}_{i+1}$ and k_{i+1} are calculated using (4.5), (2.5) and the hardening rule. Currently this initial value problem is integrated using the explicit forward Euler scheme. The value obtained is then scaled to lie on the yield surface $F(\boldsymbol{\sigma}, k) = 0$.

4.2 Other models and error analysis

In modelling problems of elasto-plastic deformation, which is a nonlinear effect, the situation is totally different to the linear elasticity context where we had a well defined model backed up by considerable theory. Now the modelling rests on the choice of assumptions, which are of course open to debate. The model most popular with mathematicians, unlike that of Section 4.1, has been that based on the deformation theory of plasticity. In its simplest form this model is in fact one of nonlinear elasticity and does not allow for unloading in the material, see Nečas and Hlavaček [19], and is thus rather unsatisfactory. Mathematical formulations of deformation theory plasticity take the form of variational inequalities, and various finite element techniques have been applied to these to produce numerical approximations, Nečas and Hlavaček [19] and Oden and Whiteman [24]. Using penalty techniques and under various regularity assumptions on the solution and the yield function the convergence of the approximate stress is proved in [24].

Unfortunately the situation is not nearly so well developed for models based on incremental (flow theory) plasticity in strain space, as used in Section 4.1. First attempts at a model, an algorithm and error analysis in this context have been made by Korn'eev and Langer [15] and Miyoshi [17], whilst another model based on a variational inequality has been formulated by Beagles, Hlavaček, Rosenberg and Whiteman [17]. This lack of general results is disappointing, as a majority of the leading nonlinear finite element codes adopt an incremental approach of the type of Section 4.1; see e.g. Owen and Hinton [25].

5 Viscoelasticity

Many solid materials of the polymeric type undergo viscoelastic deformation; that is they possess characteristics of elastic solids and of viscous fluids. For example if a stress free viscoelastic solid is subjected to a rapidly applied loading which is then held constant, then it will experience an initial (elastic) deformation followed by deformation which increases (creeps) with time up to a finite level. Alternatively, if a displacement is rapidly imposed on the solid and then held constant, the stress will initially increase rapidly (the elastic response) and will then decrease to a nonzero value (the stress relaxation effect). Each example illustrates the fact that the response at reference position $\boldsymbol{x}$ for time t depends on the history of the deformation at the point $\boldsymbol{x}$. In the case of *linear* viscoelasticity the dependence on the history is obtained by assuming a principle of linear superposition. Increments $\delta\varepsilon_{ij}$ of the strain components at time τ produce increments $\delta\sigma_{ij}$ in the stress components at subsequent time t, with the size of these increments depending on the lapsed time since the application of the strain increment, i.e.

$$(5.1)\qquad \delta\sigma_{ij}(t)=D_{ijkl}(t-\tau)\delta\varepsilon_{kl}(\tau),$$

where the usual tensor summation convention has been used. In the limit the principle of linear superposition then implies that the stress at time t is obtained by considering all the contributions (5.1) for $\tau<t$, to give

$$(5.2)\qquad \sigma_{ij}(t)=\int_{-\infty}^{t} D_{ijkl}(t-\tau)\frac{\mathrm{d}\varepsilon_{kl}}{\mathrm{d}\tau}\,\mathrm{d}\tau.$$

If we assume that there is no deformation for time $\tau<0$ and again use the vectors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$, as in Section 2, then (5.2) can be written in the form

$$(5.3)\qquad \boldsymbol{\sigma}(t)=\int_{-\infty}^{t} \boldsymbol{D}(t-\tau)\dot{\boldsymbol{\varepsilon}}(\tau)\mathrm{d}\tau,$$

where $\boldsymbol{D}(t-\tau)$ is a 6×6 matrix of material functions.

5.1 Finite Element Discretisation

For viscoelastic deformation the displacement $\boldsymbol{u}(\boldsymbol{x},t)$ is defined at all points $\boldsymbol{x}\in\bar{\Omega}$ for time t. The weak problem at each time t relating to the equilibrium equations (2.1)–(2.3) for a general viscoelastic material is then

$$(5.4)\qquad \int_{\Omega} \boldsymbol{\sigma}(\boldsymbol{u},(\boldsymbol{x},\tau);\tau\leqslant t)^{\mathrm{T}}\cdot\boldsymbol{\varepsilon}(\boldsymbol{v})\mathrm{d}\boldsymbol{x}-\int_{\Omega}\boldsymbol{f}(t)^{\mathrm{T}}\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}$$
$$-\int_{\partial\Omega_T}\boldsymbol{g}(t)^{\mathrm{T}}\cdot\boldsymbol{v}\,\mathrm{d}s=0,\qquad \forall\,\boldsymbol{v}\in V,$$

where the test space V remains the same as in Section 2, i.e. $\boldsymbol{v}\in V$ does not involve time. In discretising (5.4) at time t we approximate $\boldsymbol{u}(\boldsymbol{x},t)$ by $\boldsymbol{u}_h(\boldsymbol{x},t)$ where

$$(5.5)\qquad \boldsymbol{u}_h(\boldsymbol{x},t)=\boldsymbol{N}(\boldsymbol{x})\boldsymbol{U}(t)$$

with $\boldsymbol{U}(t)$ denoting the vector of displacement nodal variables on the partition of Ω at time t. Then, with the assumption of small strain, the approximate strain $\boldsymbol{\varepsilon}_h$ corresponding to (2.8) can be expressed as

$$(5.6)\qquad \boldsymbol{\varepsilon}_h\equiv\boldsymbol{\varepsilon}(\boldsymbol{u}_h(\boldsymbol{x},t))=\boldsymbol{B}(\boldsymbol{x})\boldsymbol{U}(t)$$

and the approximate stress corresponding to (5.3) is

$$(5.7)\qquad \boldsymbol{\sigma}_h\equiv\boldsymbol{\sigma}_h(\boldsymbol{x},t)=\int_{0}^{t}\boldsymbol{D}(t-\tau)\boldsymbol{B}(\boldsymbol{x})\dot{\boldsymbol{U}}(\tau)\mathrm{d}\tau.$$

The semidiscrete form of (5.4) is then obtained by substitution of (5.6) and (5.7) as the linear integro-differential equation system

$$(5.8)\quad \int_0^t \left(\int_\Omega \boldsymbol{B}^{\mathrm{T}} \boldsymbol{D}(t-\tau) \boldsymbol{B} \mathrm{d}\boldsymbol{x} \right) \dot{\boldsymbol{U}}(\tau) \mathrm{d}\tau - \int_\Omega \boldsymbol{f}(t)^{\mathrm{T}} \boldsymbol{N} \mathrm{d}\boldsymbol{x} - \int_{\partial\Omega_T} \boldsymbol{g}(t)^{\mathrm{T}} \boldsymbol{N} \mathrm{d}s = 0.$$

We discretise (5.8) in time by taking time levels $t_j, j=0,1,2,\dots$ and for $t_{j-1} < \tau < t_j$ approximating $\dot{\boldsymbol{U}}(\tau)$ by $(\boldsymbol{U}(t_j) - \boldsymbol{U}(t_{j-1}))/(t_j - t_{j-1})$. This gives

$$(5.9)\quad \left(\int_\Omega \boldsymbol{B}^{\mathrm{T}} \tilde{\boldsymbol{D}}\, \boldsymbol{B}\, \mathrm{d}\boldsymbol{x} \right) (\boldsymbol{U}(t_j) - \boldsymbol{U}(t_{j-1})) = \int_\Omega \boldsymbol{f}(t)^{\mathrm{T}} \boldsymbol{N}\, \mathrm{d}\boldsymbol{x} + \int_{\partial\Omega} \boldsymbol{g}(t)^{\mathrm{T}} \boldsymbol{N}\, \mathrm{d}s + \boldsymbol{P}(t_j),$$

where $\tilde{\boldsymbol{D}} = \left(\int_{t_{j-1}}^{t_j} \boldsymbol{D}(t_j - \tau) \mathrm{d}\tau \right) / (t_j - t_{j-1})$

and $\boldsymbol{P}(t_j) = \sum_{q=1}^{j-1} \int_\Omega \boldsymbol{B}^{\mathrm{T}} \left(\int_{t_{q-1}}^{t_q} \boldsymbol{D}(t_j - \tau) \mathrm{d}\tau / (t_q - t_{q-1}) \right) \boldsymbol{B}\, \mathrm{d}\boldsymbol{x} (\boldsymbol{U}(t_q) - \boldsymbol{U}(t_{q-1}))$.

For a general form of $\boldsymbol{D}$ the computation of $\boldsymbol{P}(t_j)$ involves the solution at all previous time steps. We however restrict attention to materials with constant Poisson's ratio and matrices $\boldsymbol{D}$ which can be expressed in terms of decaying exponentials, i.e.

$$(5.10)\quad \boldsymbol{D}(s) = \varphi(s) \boldsymbol{D}(0), \qquad \varphi(s) = \sum_{k=1}^{M} C_k \mathrm{e}^{-\alpha_k s}, \qquad \alpha_k \geqslant 0,$$

where φ is the stress relaxation function. In this way, at each time level t_j, $\boldsymbol{P}(t_j)$ can be computed from the M vectors $\gamma(\alpha_k, t_{j-1})$, $k = 1(1)M$ where $\gamma(.,.)$ is defined by

$$(5.11)\quad \gamma(\alpha, t_q) = \int_0^{t_q} \mathrm{e}^{-\alpha(t_q - \tau)} \dot{\boldsymbol{U}}(\tau) \mathrm{d}\tau.$$

The assumption of constant Poisson's ratio puts the problem in a form where correspondence principles can be used, see e.g. Schapery [26], and this enables an alternative solution strategy to be used. In this approach the displacement at time t_i is expressed as a convolution of a *creep* function and the displacements of related elastic problems for times $\tau \leqslant t_i$.

The above analysis has been produced under the assumption of linear viscoelasticity and an exponential decay form for the stress relaxation function. It is known however that at best this only approximately describes real viscoelastic materials and is inappropriate for many of the polymeric materials currently being produced. These can only be described by more complicated nonlinear models.

6 Concluding Remarks

The previous three sections have shown the applicability and state of knowledge in the use of finite element methods in the context of linear elasticity, elastoplasticity and viscoelasticity. The major differences between these result from the different constitutive relations required. In the simplest case of linear elasticity no time or history effects are involved and the theory is well developed. However with the more complicated materials the models may involve time and history effects. In the elastoplastic model of Section 4 the resulting equations are nonlinear. This model involves a "time" parameter as a result of the incremental loading procedure, as well as indirectly in the integration of the hardening rule. The history affects the value of the hardening parameters. In the viscoelastic model of Section 5 time and history effects are explicitly involved through the constitutive equations. As has been said, in both these cases the theory concerning error estimates is much less developed than for the case of linear elasticity.

References

[1] Andreev, A. B.; Lazarov, R. D.: Superconvergence of the gradient for quadratic triangular finite elements. Numer. Meth. Partial Differential Equations **4** (1988) 15–32

[2] Axelsson, O.; Baker, V. A.: Finite Element Solution of Boundary Value Problems. New York: Academic Press 1984

[3] Babuska, I.; Chandra, J.; Flaherty, J. E. (eds): Adaptive Computational Methods for Partial Differential Equations. Philadelphia: SIAM 1983

[4] Babuska, I.; Zienkiewicz, O. C.; Gago, J.; Oliveira, E. R. de A: Accuracy Estimates and Adaptive Refinements in Finite Element Computations. Chichester: Wiley 1986

[5] Babuska, I.; Rheinboldt, W.: Adaptive finite element processes in structural mechanics. In: Birkhoff, G.; Schoenstadt A. (eds.): Elliptic Problem Solvers, Vol. II. New York: Academic Press, 1984, pp. 345–378

[6] Bathe, K.-J.; Finite Element Procedures in Engineering Analysis. New Jersey: Prentice Hall, 1981

[7] Beagles, A. N.; Hlavaček, I.; Rosenberg, J.; Whiteman, J. R.: Variational inequality formulation in strain space of an elasto-plastic problem with hardening (to appear)

[8] Ciarlet, P. G.; The Finite Element Method for Elliptic Problems. Amsterdam: North Holland 1978

[9] Ciarlet, P. G.; Raviart, P.-A.: General Lagrange and Hermite interpolation in $\mathbb{R}^n$ with applications to finite element methods. Comp. Meth. Appl. Mech. Eng. **1** (1972) 217–249

[10] Desai, C.S.; Abel, J. F.: Introduction to the Finite Element Method. New York: Van Nostrand 1972

[11] Fichera, G.: Existence theorems in elasticity – boundary value problems of elasticity with unilateral constraints. In: Truesdell, C. (ed.): Mechanics of Solids II. Handbuch der Physik. Berlin: Springer-Verlag 1972, pp. 347–424

[12] Goodsell, G.; Whiteman, J. R.; Some gradient superconvergence results in the finite element method. In: Turner, P. (ed.): Numerical Analysis, Lancaster 1987. Lecture Notes in Mathematics, Berlin: Springer-Verlag (to appear)

[13] Harrison, D.; Ward, T. J. W.; Whiteman, J. R.: Finite element analysis of plates with nonlinear properties. Comp. Meth. Appl. Mech. Eng. **34** (1982) 1019–1034

[14] Hunter, S. C.: Mechanics of Continuous Media. Chichester: Ellis Horwood 1983

[15] Korn'eev, V. G.; Langer, U.: Approximate Solution of Plastic Flow Theory Problems. Leipzig: BSB Teubner 1984

[16] Křižek, M.; Neittaanmäki, P.: On superconvergence techniques. Acta Applicandae Mathematicae **9** (1987)

[17] Miyoshi, T.: Foundations of the Numerical Analysis of Plasticity. Amsterdam: North Holland 1985

[18] Natterer, F.: Über die punktweise Konvergenz finiter Elemente. Numer. Math. **25** (1975) 67–77

[19] Nečas, J.; Hlavaček, I.: Mathematical Theory of Elastic and Elasto-plastic Bodies. Amsterdam: Elsevier 1981

[20] Nitsche, J. A.: L_∞-convergence of finite element approximation. Proc. 2nd Conference on Finite Elements, Rennes 1975

[21] Nitsche, J. A.; Schatz, A. H.: Interior estimates for Ritz-Galerkin methods. Math. Comp. **28** (1974) 937–958

[22] Oden, J.T.; Demkowicz, L.; Strouboulis, T.; Devloo, P.: Adaptive methods for problems in solid and fluid mechanics. In: Babuska, I.; Zienkiewicz, O. C.; Gago, J.; Oliveira, E. R. de A (eds.): Accuracy Estimates and Adaptive Refinements in Finite Element Computations. Chichester: Wiley 1986, pp. 249–281

[23] Oden, J. T.; Reddy, J. N.: Introduction to the Mathematical Theory of Finite Elements. New York: Wiley-Interscience 1976

[24] Oden, J. T.; Whiteman, J. R.: Analysis of some finite element methods for a class of problems in elasto-plasticity. Int. J. Eng. Sci. **20** (1982) 977–988

[25] Owen, D. R. J.; Hinton, E.: Finite Elements in Plasticity: Theory and Practice. Swansea: Pineridge Press 1980

[26] Schapery, R. A.: Correspondence principle and a generalised J-integral for large deformation and fracture analysis of viscoelastic media. Int. J. Fracture **25** (1984) 195–223

[27] Schwarz, H. R.: Finite Element Methods. London: Academic Press 1988

[28] Spencer, A. J. T.: Continuum Mechanics. London: Longman 1980

[29] Whiteman, J. R. (ed.): The Mathematics of Finite Elements and Applications I–VI, MAFELAP 1972, 75, 78, 81, 84 and 87. London: Academic Press 1973, 1976, 1979, 1982, 1985, 1988

[30] Whiteman, J. R.: Some aspects of the mathematics of finite elements. In: Whiteman, J. R. (ed): The Mathematics of Finite Elements and Applications II, MAFELAP 1975. London: Academic Press 1976, pp. 25–42

[31] Whiteman, J. R.; Goodsell, G.: Superconvergent recovery for stresses from finite element approximations on subdomains for planar problems of linear elasticity. In: Whiteman, J. R. (ed.): The Mathematics of Finite Elements and Applications VI, MAFELAP 1987. London: Academic Press 1988, pp. 29–54

[32] Zienkiewicz, O. C.: The Finite Element Method, 3rd Edition. New York: McGraw Hill 1977

J. R. Whiteman
A. E. Beagles
M. K. Warby
Institute of Computational Mathematics
Department of Mathematics and Statistics
Brunel University
Uxbridge, UB8 3PH, England

(*Eingegangen 24. 11. 1988*)

Buchbesprechungen

Yaglom, I. M., Felix Klein and Sophus Lie, Evolution of the Idea of Symmetry in the Nineteenth Century, Boston – Basel: Birkhäuser-Verlag 1988, 237 S., DM 82,–

Die Bezeichnung „das goldenen Zeitalter der Geometrie" umreißt die Zeit von der Napoleonischen Ära bis zur Wende zum 20. Jahrhundert. Dieser Epoche der Geschichte der Geometrie gilt das vorliegende Buch. Sein Verfasser orientiert die Beschreibung neuer Entwicklungen in dieser Phase an verschiedenen Schwerpunkten. Dazu zählt gewiß die *projektive Geometrie.* Hierin sind zwei Schulen erkennbar: Die französische mit Monge, Carnot, Gergonne, Poncelet, Chasles und die deutsche mit Möbius, Steiner, von Staudt, Plücker.

Ein folgenreicher Schritt ist die Entdeckung der *nichteuklidischen Geometrien.* Man verdankt sie bekanntlich den fast gleichzeitig und unabhängig gewonnenen Einsichten von Bolyai, Gauss, Lobatschewski.

Schließlich verweist der Verfasser auf einen Zweig der Geometrie des 19. Jahrhunderts, dessen Verästelungen weit in die moderne Mathematik hineinreichen, wir sagen hier kurz: die *Grundlegung der linearen und multilinearen Algebra* durch Grassmann, Cayley, Hamilton.

Die von Cremona gegründete *italienische Schule der algebraischen Geometrie* sowie die *italienischen Differentialgeometer* Ricci, Levi-Civita, Bianchi werden zwar erwähnt, aber nur in den Fußnoten.

Die überragenden Figuren von Felix Klein (1849–1925) und Sophus Lie (1842–1899) leiten die Geschichte der Geometrie im letzten Drittel des 19. Jahrhunderts über in ein neues Zeitalter der Mathematikgeschichte. Die Leistungen dieser Mathematiker ist ohne die Entwicklungen der Geometrie, die ihnen vorausgegangen waren und mit denen sie heranwuchsen, nicht denkbar. Gewiß fußt die Geometrie des 19. Jahrhunderts auf dem bisher Erreichten. Wollte man jedoch ein Prinzip herausgreifen, durch welches im 19. Jahrhundert die Moderne in der Mathematikgeschichte beginnt, so müßte dies der Begriff der Symmetrie sein. Um die Wende vom 18. zum 19. Jahrhundert beginnt sich die Absicht durchzusetzen, das Prinzip der *Symmetrie* ernsthaft anzuwenden. Sie wird nun zum methodischen Werkzeug und weist nicht nur der Geometrie, sondern auch der Algebra neue Wege, die bis in die Mathematik der Gegenwart hineinführen. Umsomehr ist es gerechtfertigt, daß ein Buch über das Goldene Zeitalter der Geometrie die Symmetrie als Leitmotiv wählt. Das Verständnis der Symmetrie vertiefte sich im 19. Jahrhundert im selben Maß, wie sich der Begriff der Gruppe und Transformationsgruppe entwickelte. Insoweit wird verständlich, daß der Verfasser die geschichtlichen Fäden der Geometrie im 19. Jahrhundert auf die Gestalten von Felix Klein und Sophus Lie zulaufen sieht. Folgerichtig beginnt seine Darstellung mit den Mathematikern, die die Voraussetzungen für das gruppentheoretische Wirken dieser beiden schufen: Lagrange, Cauchy, Abel, Galois, Cayley, Jordan.

Yaglom untermauert seine Geschichte der Geometrie des 19. Jahrhunderts mit der Darstellung aller wesentlichen mathematischen Gesichtpunkte und Erkenntnisse. Auch dem an traditioneller Geometrie nicht Geschulten müßten diese Beschreibungen zugänglich sein. Der Verfasser meint, eine Gymnasialbildung in Mathematik reiche; etwas weiterreichende Erfahrung könnte nicht schaden. Die Mathematik, die mathematikgeschichtlichen Hintergründe, die Lebensschicksale sind packend geschildert. Der Verfasser gibt Hinweise auf die Fachliteratur bis hin zur allerneuesten. Damit der Fluß seiner Schilderung nicht unterbrochen wird, sind ergänzende Informationen mathematischer, historischer, biographischer, anekdotischer und wertender Natur in einem umfangreichen Fußnotenapparat auf den Seiten 138–237 zusammengefaßt, den gründlich mitzulesen sich unbedingt lohnt.

Das schiere Volumen des gesammelten Materials ist beeindruckend. Trotz der ausgebreiteten Fülle sind historische und biographische Ereignisse so farbig geschildert, daß die Lektüre nirgendwo Langeweile aufkommen läßt. Einfühlsam sind die Persönlichkeitsprofile der Protagonisten gezeichnet. Die Lektüre macht da und dort nachdenklich. Die geschichtliche Entwicklung der Mathematik durch die schöpferische Einzelleistung ist sehr wohl eine Sache, die Beziehung der Mathematiker untereinander auf fachlicher und zwischenmenschlicher Ebene indessen eine andere, und in mancherlei Hinsicht ist diese nicht minder wichtig als jene. Daher spielen neben intellektuellen Anlagen auch charakterliche historisch eine Rolle. Der Verfasser versteht diese Ambivalenz etwa im Falle von Gauss zu verdeutlichen. Auch diese oder jene Anekdote, wenn sie als solche in einer Fußnote mit einer gewissen persönlichen Note erzählt wird, trägt zum historischen Kolorit bei (z. B. S. 200 über Hamiltons Auffindung der Quaternionen).

Alle neuzeitlichen Darstellungen der neueren Mathematikgeschichte und biographische Sammlungen geben freigibig Auskunft über Felix Klein. In der Geschichtsschreibung jedoch steht Sophus Lie noch heute im Schatten Felix Kleins. Zum Teil erkärt sich dies sicher daraus, daß Klein durch sein unbestechliches Urteil, seinen unermüdlichen Fleiß und sein geniales Verwaltungstalent die Entwicklung der Mathematik nicht nur in der Forschung selbst, sondern auch auf organisatorischer und didaktischer Ebene dauerhaft beeinflußt hat. Auch der bekannte Umstand, daß er sich in der Mathematikgeschichte engagiert und auch die Sammlung seiner eigenen Arbeiten kommentierend mitbesorgt hat, ist nicht ohne Wirkung auf die späteren Generationen von Historikern geblieben. Erst nachdem sich die an Liegruppen und ihren Anwendungen interessierten Mathematiker in den letzten Jahrzehnten wiederholt mit den Lieschen Originalquellen und deren Interpretation auseinandergesetzt haben, gehen nun die Mathematikhistoriker mit Nachdruck daran, die Gestalt Lies und seine Beziehung zu Klein geschichtswissenschaftlich gründlich zu untersuchen (Bourbaki, N., Groupes et algèbres de Lie, Chap. 2 et 3, Paris, Hermann 1972, insbesondere S. 286ff.; Rowe, D. E., *Three letters from Sophus Lie to Felix Klein on Parisian Mathematics during the early 1880s,* Mathem. Intelligencer **7** (1985), 74–77; *Der Briefwechsel Sophus Lie – Felix Klein, eine Einsicht in ihre persönlichen und wissenschaftlichen Beziehungen,* NTM-Schriftenr. Gesch. Naturwiss. Technik Med. – Leipzig **25** (1988), 37–47; *On the Mathematical and Personal Relationship of Sophus Lie and Felix Klein with a Transcription of Klein's unpublished manuscript:* „Über Lie's und meine älteren geometrischen Arbeiten", in preparation). Inwieweit Lie zu seinen Lebzeiten schon im Schatten des gewiß extrovertierteren und breitenwirksameren Klein stand, wird wohl im Einzelnen aus den Quellen noch weiter erarbeitet werden; sicher ist es ihm selber zeitenweise so erschienen (vgl. S. 25).

Das vorliegende Buch ist in der Behandlung des „Klein-Lie-Komplexes" aus historischer Sicht sicher ein Fortschritt, weil sich sein Verfasser sichtlich um eine (fast!) gleichgewichtige Behandlung beider Gestalten bemüht. Insoweit wird es seinem Titel „Felix Klein und Sophus Lie" gerecht. Leser, die auf diesen Titel die Hoffnung begründet hätten, das Buch böte ihnen neue Einsichten und Fakten über das komplizierte Wechselverhältnis der beiden großen Mathematiker sowohl im mathematischen als auch im biographisch-persönlichen Bereich, werden allerdings enttäuscht. Diejenigen Leser hingegen, die den Untertitel „Evolution of the idea of symmetry in the nineteenth century" als Obertitel lesen, werden durch die Lektüre des Buches reich belohnt. Unter dieser Überschrift sind die Figuren von Klein und Lie die Kerne, an denen sich des Verfassers Darstellung der Geschichte der Geometrie im 19. Jahrhundert immer aufs Neue kristallisiert. Biographische Details dieser Protagonisten läßt der Verfasser wie den sprichwörtlichen roten Faden im ganzen Gewebe des Werkes immer wieder durchblicken.

Gelegentlich versucht er, seine Leser zu gewagten Bewertungen mitzureißen; nicht jeder wird ihm darin bereitwillig folgen. So legt er es darauf an, die Persönlichkeiten mindestens von Klein und Lie typologisch zu klassifizieren. Zunächst ist da die These, man müsse Lie rückwärts, Klein aber vorwärts gerichtet sehen (S. 24, S. 128, S. 134). Was Kleins und Lies Arbeitsweise oder ihre Einstellung zu Forschung und Lehre angeht, mag es einleuchten, Lie eher in das 19. Jahrhundert, aber Klein in das 20. Jahrhundert einzuordnen. Vom mathematischen Inhalt ihrer schöpferischen Leistung her gesehen freilich kann man Lie unter keinen Umständen seinen kaum zu überschätzenden Einfluß auf die Mathematik des 20. Jahrhunderts absprechen, den er bis in die gegenwärtige Forschung hinein ausübt. In zahlreichen mathematischen Spezialdisziplinen ist die Theorie der Liegruppen ein zentrales Werkzeug geworden (Dieudonné, J., Orientation générale des mathématiques pures en 1973, Gazette des Mathématiciens Octobre (1974), 73–79, insbesondere Seite 77, Absatz 3). Nicolas Bourbaki bezieht in den notes historiques zu dem oben zitierten Werk eine unzweideutige Stellung zu der Bedeutung Lies. (Übrigens wird Bourbaki vom Verfasser gern und oft zitiert.)

Eine noch kühnere Typologie soll zum Beispiel die Antipathie zwischen Weierstrass und Klein erklären. Dabei dürfen die verschiedenartigen Funktionen der beiden Hälften des menschlichen Gehirns für eine Typenklassifikation von Mathematikern herhalten (S. 25, 26). Die linke Gehirnhälfte kontrolliert analytische und logische Operationen, steuert das Rechnen und sorgt also alles in allem für die Algebra. Hingegen ist die rechte Hirnhälfte für eine bildhafte, synthetische Weltsicht zuständig, befördert also in der Mathematik die geometrische Anschauung, das Denken in Diagrammen und Bildern. Damit erklärt der Verfasser die zwei weitgehend gegensätzlichen Typen der „Algebraiker" auf der einen Seite und der „Geometer und Physiker" auf der anderen. Schon Hermann Weyl hat über eine solche Typenaufteilung nachgedacht (Weyl, H., Topologie und abstrakte Algebra als zwei Wege mathematischen Verständnisses, Unterrichtsblätter für Mathematik und Naturwissenschaften **38** (1933), 177–188 (Gesammelte Abhandlungen **3**, Springer Heidelberg 1968)). Nach seiner Meinung gehören Klein und Riemann zum zweiten, jedoch Weierstrass zum ersten Typ. Der Verfasser scheint offen zu lassen, wo Lie einzuordnen sei – wohl eher bei den Geometern; jedenfalls hat Lie sich zeitlebens als Geometer verstanden. Der Leser freilich schwankt zwischen der Bewunderung für eine phantasievolle Ausdeutung, die er wohl hier und dort in der eigenen Erfahrung bestätigen wird, auf der einen Seite und Zurückhaltung vor einer Überbewertung derartiger Spekulationen auf der anderen.

Das Buch wäre noch viel nützlicher gewesen, wenn ihm ein gut organisiertes alphabetisches Namenregister und ein ebensolches Sachregister beigegeben worden wäre. Auch ein systematisches Literaturverzeichnis fehlt. (Bei der Technologie der elektronischen Datenverarbeitung, die Autoren heute gemeinhin zur Verfügung steht, ist eine Forderung nach einem solchen mechanisch erstellbaren Apparat nicht unbillig.) Abgesehen hiervon ist die Dokumentation detailliert und reich. Gegenüber der lebendig und packend dargestellten Fülle der Fakten fallen kleinere Ungenauigkeiten nicht ins Gewicht. So wird auf S. 129 die in der Literatur eingewurzelte Darstellung fortgepflanzt, das Kleinsche Erlanger Programm (Klein, F., *Vergleichende Betrachtungen über neuere Geometrische Forschungen,* Programm zum Eintritt in die philosophische Fakultät und den Senat der K. Friedrich Alexander-Universität zu Erlangen, Erlangen, Deichert 1972) sei Gegenstand einer Antrittsvorlesung gewesen. Kleins Antrittsrede in Erlangen von 1872 befaßte sich mit dem Stand der Mathematikdidaktik an Gymnasien und Hochschulen und entwirft, wie diese zukünftig zu gestalten sei (D. E. Rowe, *Felix Klein's „Erlanger Antrittsrede". A transscription with English translation and Commentary,* Historia Math. **12** (1985), 123–141). Das später in verschiedenen Zeitschriften veröffentlichte „Erlanger Programm" ist, worauf D. E. Rowe *(A forgotten chapter in the history of Felix Klein's Erlanger Programm,* Historia Math. **10** (1983), 448–457) nachdrücklich hingewiesen hat, in einer Programmschrift vorgelegt und – den meisten Quellen zum Trotz – in einer Erlanger Antrittsvorlesung nicht vorgetragen worden.

Darmstadt K. H. Hofmann

Rucker, R., Der Ozean der Wahrheit oder die fünf Arten zu denken, Frankfurt am Main: S. Fischer Verlag GmbH 1988, 399 S., DM 39,80

Davis, Ph. J., Hersh, R., Descartes' Traum, Frankfurt am Main: S. Fischer Verlag GmbH 1988, 422 S., DM 48,–

Mit dem Eindringen mathematischer Ideen und Methoden in nahezu alle Bereiche menschlicher Existenz ist die Aufgabe, Verständnis für mathematisches Denken und Vorgehen zu vermitteln, für die forschenden und lehrenden Mathematiker dringender denn je geworden. Es genügt nicht mehr, einer Gemeinde von Hobbymathematikern reizvolle Aufgaben zu stellen, es kommt darauf an, nahezu jedem Bildungswilligen eine Art Kanon mathematischer Ideen samt deren „Sitz im Leben" zu vermitteln, und ihm deutlich zu machen, was es heute heißt, Mathematiker zu sein. Die beiden hier zu besprechenden Bücher widmen sich dieser Aufgabe auf verschiedene Weise.

Ruckers Buch „Der Ozean der Wahrheit" stellt dem Leser einen – m. E. trefflich gewählten – Fünfer-Kanon mathematischer Ideenkreise vor: Zahl, Raum, Logik, Unendlichkeit, Information. Die Entfaltung des damit jeweils Umfaßten konfrontiert den Leser im Endeffekt mit einem Ensemble mathematischer Schaustücke, das so gut wie alles umfaßt, was ein Gebildeter von heute gesehen haben sollte, bis hin zu Turing-Maschinen, Fraktalen, algebraischen Kurven und Penrose-Kacheln. Gelegentlich, so etwa im Abschnitt „Zahlenmuster", geht der Vorzeigetrieb mit dem Autor durch, aber welcher vom eigenen Fach Begeisterte möchte da den ersten Stein werfen? In seinem Drang, mathematischen Ideen einen „Sitz im Leben" anzuweisen, spielt der Autor das Klavier der Analogien und Assoziationen sozusagen ständig im Fortissimo mit durchgetretenem Pedal. Das macht sicher manchem Leser großen Spaß, führt aber den Laien sicher auch auf Holzwege. Und auch, wo nur Mathematik erläutert werden soll, führt das Bemühen, dem Leser nahe zu bleiben, gelegentlich weit weg vom Ziel zutreffender Information, m. E. besonders schlagend im Abschnitt „Hilbertraum". Dennoch möchte ich meinen, daß ein engagierter Leser dieses Buches hinterher bemerkenswert fähig zum Gespräch mit entgegenkommenden Mathematikern sein dürfte.

Der Essay-Band „Descartes' Traum" des durch „The Mathematical Experience" (1980, deutsch 1985) einem großen Publikum rühmlich bekannten Autorengespanns Davis/ Hersh läßt die sozialen und geistesgeschichtlichen Bezüge der Mathematikerschaft gegenüber systematischer Sachinformation über mathematische Inhalte in den Vordergrund treten. Beispielsweise verdeutlicht das Kapitel „Die soziale Tyrannei der Zahlen", wieviel Skepsis man manchem außermathematischen Auftreten von Mathematik(ern) besser entgegenbringen sollte. Interviews mit dem Informatiker C. M. Strauss treiben die in diesem Buch immer wieder angestrebte Aktualität auf eine – allerdings, wie ich fürchte, kurzlebige – Spitze und unterstreichen die Bezüge zum Computerwesen. Ein Abschnitt „Platonistische Mathematik und platonische Religionsphilosophie: eine ethische Metapher" wird manchen Mathematiker, der sich als „naiven Platoniker" bekennt, ein gutes Stück weiter bringen. Im Nachwort des Buches findet sich die sehr bedenkenswerte These

> Die Mathematik verlangt von ihren Ursprüngen und ihrer Entwicklung her nach einer vollständigen Verbindung mit allen Arten von geistigen und körperlichen Betätigungen des Menschen.

Mit diesen fragmentarischen Hinweisen auf die bunte Fülle der in diesem Buch angerissenen Themen möchte ich es hier bewenden lassen.

Beide Bücher entlassen wohl jeden Leser erheblich informierter, nachdenklicher und gebildeter als er vorher war. Dem Hochschulmathematiker wird die Lektüre vor allem dann dienlich sein, wenn er sich anschickt, mit Nicht-Mathematikern ins Gespräch zu kommen.

Erlangen K. Jacobs

Loewner, Ch., Collected Papers, Edited by Lipman Bers, Boston – Basel: Birkhäuser Verlag 1988, 536 S., DM 168,–

Karl Löwner ist den Analytikern vorwiegend bekannt durch seine fundamentale Methode zur Behandlung des Koeffizientenproblems in der Klasse S der schlichten Funktionen. Ist $f(z) = z + a_2z^2 + a_3z^3 + \ldots$ im Einheitskreis $\mathbb{D}$ schlicht [biholomorphe Abbildung von $\mathbb{D}$ auf $f(\mathbb{D})$], so hatte Bieberbach 1916 vermutet, daß $|a_n| \leqslant n$ gilt ($n = 2, 3, \ldots$). Der Fall $n = 2$ ist elementar, $|a_3| \leqslant 3$ zeigte 1923 Löwner mit seiner Methode, die auch wesentliches Element bei dem von DeBranges (1985) erbrachten allgemeinen Beweis der Bieberbachschen Vermutung ist.

Löwner bemerkt, daß es ausreicht, schlichte Abbildungen von $\mathbb{D}$ auf $\mathbb{D}\backslash C$ zu betrachten, wobei C ein stückweise analytischer Schlitz in $\mathbb{D}$ mit Schlitzende auf $\partial\mathbb{D}$ ist. Jede solche Abbildung läßt sich als Lösung der „Löwnerschen Differentialgleichung"

$$\frac{\partial w}{\partial t} = -w\frac{1 + \varkappa(t)w}{1 - \varkappa(t)w} \quad (0 \leqslant t \leqslant t_0)$$

darstellen: $w = f(z, t_0)$, wenn man $f(z, 0) = z$ verlangt; dabei ist $\varkappa$ in $[0, t_0]$ stetig und vom Betrag 1. Nach weiteren Zwischenrechnungen gelangt man zu $|a_3| \leqslant 3$; Löwner bewies auch schon $|b_n| \leqslant \dfrac{1}{n}\dbinom{2n}{n+1}$ für die Koeffizienten b_n der Umkehrfunktion $g = f^{-1}$ eines $f \in S$.

Die vorliegenden Gesammelten Abhandlungen (33 Arbeiten sind reproduziert) zeigen nun eindringlich, daß Löwner auch in anderen Bereichen der Analysis wesentliches geleistet hat, Die Hauptthemen stammen aus klassischer Funktionentheorie (Verzerrung, Interpolation), partiellen Differentialgleichungen, Strömungstheorie, Halbgruppen von Abbildungen. Ihre Impulse kommen aus der Prager Zeit und aus Löwners Beschäftigung mit technischen Problemen in den Kriegs- und Nachkriegsjahren. Besonders hervorzuheben sind eine tiefliegende und schöne Arbeit über monotone Matrixfunktionen (1934), welche mit Picks positiven Funktionen in Zusammenhang steht, ein Übersichtsartikel über „Semigroups in Analysis and Geometry" (1964), sowie mehrere Arbeiten über Erhaltungssätze bei kompressiblen und inkompressiblen Strömungen, die teilweise aus NACA-Reports stammen.

Die Herausgabe der Werke besorgte Lipman Bers, der eine eingehende und farbige Schilderung des teilweise schwierigen Lebens von Karl Löwner gibt. Löwner wurde 1893 bei Prag geboren und besuchte 1912–1917 die dortige Universität; sein Doktorvater war Georg Pick. 1922 wurde er Privatdozent an der Universität Berlin; er kehrte dann 1930 als Professor nach Prag zurück, von wo er 1939 nach den USA emigrierte. Nach mehreren Zwischenstationen war Löwner von 1951 bis zu seiner Emeritierung 1963 an der Stanford-Universität, wo er im Januar 1968 verstarb.

Giessen D. Gaier

Jungnickel, D., Graphen, Netzwerke und Algorithmen, Mannheim – Wien – Zürich: Bibliographisches Institut 1987, 405 pp., Hard cover, DM 68,–

Beim Schreiben eines Buches über kombinatorische Optimierung gibt es derzeit zwei Probleme. Zum einen ist dieses Gebiet in den letzten Jahren explosionsartig gewachsen, so daß es – selbst bei einem so jungen Teilbereich der Mathematik – kaum mehr möglich ist, alle relevanten Untergebiete und ihre Beziehungen zu anderen Disziplinen der Mathematik und zu den Anwendungen abzudecken, zum anderen werden gegenwärtig weltweit enorme Forschungsanstrengungen unternommen, so daß die Gefahr besteht, daß ein Buch sehr schnell veraltet. Dies sind sicherlich einige der Gründe dafür, daß zur Zeit so wenige allgemeine Lehrbücher über kombinatorische Optimierung geschrieben werden. Gerade deswegen besteht – insbesondere auf dem Markt deutschsprachiger Bücher – ein echter Bedarf an einführenden Lehrbüchern in dieses faszinierende Gebiet. Die Nachfrage gerade nach Literatur, die für Anfänger verständlich ist, ist stark wachsend, da einige Teilbereiche der kombinatorischen Optimierung vermehrt in Gebieten wie Informatik, Wirtschaftswissenschaften, Ingenieurwissenschaften benötigt und an Universitäten gelehrt werden. Insofern ist es höchst verdienstvoll, daß D. Jungnickel sich der Mühe unterzogen hat, ein Buch zu schreiben, das in einen wichtigen Teilbereich der kombinatorischen Optimierung (polynomial lösbare, auf Graphen formulierbare Probleme) einführt. Es füllt mit Sicherheit eine derzeit bestehende Lücke auf dem Buchmarkt.

Das Buch ist didaktisch geschickt aufbereitet. Die jeweiligen Kapitel werden häufig aus Anwendungsfragen heraus motiviert. Die Terminologie ist klar und übersichtlich. Die Beweise sind, sofern dies sinnvoll ist, algorithmisch formuliert, kurz aber doch gut verständlich. Wahrscheinlich werden auch mathematisch weniger geschulte Leser dem Stoff folgen können. Dennoch hat Jungnickel als Leser sicherlich Mathematik-Studenten bzw. mathematisch geschulte Informatiker als Zielgruppe vor Augen.

Wie bereits erwähnt, konzentriert sich das Buch von Jungnickel auf kombinatorische Optimierungsprobleme, die mit Hilfe von Graphen bzw. Netzwerken formuliert werden können. Innerhalb dieser Problemklasse wird das Hauptaugenmerk auf polynomial lösbare Probleme gelegt. Der Autor verzichtet aus Platzgründen bewußt auf die Darstellung von Algorithmen, die auf linearer Programmierung basieren und ebenso auf die Beschreibung polyedrischer Methoden. Im einzelnen ist der Inhalt wie folgt gegliedert.

Kapitel 1 ist eine hübsche, knappe Einführung in wichtige Bereiche der Graphentheorie mit schönen Beweisen durchaus nichttrivialer Sätze. In Kapitel 2 werden die Instrumente zur informellen Beschreibung und genaueren Formulierung und Analyse graphentheoretischer Algorithmen bereitgestellt. Dies ist ein didaktisch schwieriger Bereich, da bei zu formalistischem Vorgehen der Überblick verloren geht, insbesondere die grundlegenden (häufig einfachen) Ideen nicht mehr sichtbar sind. Bei zu viel „handwaving" übersieht man eine Reihe technischer Details und Finessen. Diese Probleme hat Jungnickel auf didaktisch ansprechende Weise gelöst – sicherlich auch in den Augen der Leser, die sich nicht mit Pseudoprogrammiersprachen anfreunden können.

Kapitel 3 behandelt Wegeprobleme und folgt im wesentlichen dem (z. B. im Buch von Lawler „Combinatorial Algorithms: Networks and Matroids" vorgezeichneten) Standardweg, der sich bewährt hat. Nichtstandard-Bereicherungen des üblichen Materials sind z. B. die Behandlung von endlichen metrischen Räume und von Wegalgebren, die Erwähnung der Beziehungen zu Hasse-Diagrammen.

Kapitel 4 zeigt, wie man maximale oder minimale gewichtete Bäume berechnen kann, und führt – durch kanonische Fortführung der dazu entwickelten Theorie – in die Matroidtheorie ein. Speziell wird gezeigt, daß der Greedy-Algorithmus für Matroide „funktioniert". Da Algorithmen vom Greedy-Typ in der Praxis sehr häufig auch für nichtmatroidale Probleme als Heuristiken verwendet werden, hätte ich mir hier noch Informationen zum Worst-Case-Verhalten des Greedy-Algorithmus (z. B. für Unabhängigkeitssysteme) gewünscht.

Kapitel 5 behandelt die „klassische" Netzwerkflußtheorie (Ford-Fulkerson-Algorithmus und die Edmonds-Karp-Version, Dinic- und MKM-Algorithmus). Gerade dieses Gebiet ist ein Beispiel dafür, wie rasant die Entwicklung in der kombinatorischen Optimierung abläuft. In den letzten drei Jahren – und daher noch nicht in Jungnickel's Buch enthalten – sind u. a. von Goldberg, Tarjan, Ahuja, Orlin, Tardos verschiedene auf neuen Ansätzen basierende Maximalflußalgorithmen entwickelt worden, die die klassischen Verfahren sowohl bezüglich theoretischer als auch praktischer Laufzeit schlagen. Analoges gilt für Kapitel 7 über Zirkulationen in Netzwerken. Die im Buch behandelten Algorithmen sind klar und übersichtlich dargestellt. Es ist auch heute noch notwendig, diese zu lehren und zu lernen, da sie auf wichtigen algorithmischen Ideen beruhen.

In Kapitel 6 stellt Jungnickel Beziehungen der Netzwerkfluß-Theorie zu anderen Bereichen der Kombinatorik dar (Menger's Sätze, bipartites Matching, Heiratssatz, Satz von Dilworth etc.). Dies ist eine schöne Zusammenstellung und enthält auch einige, für einführende Lehrbücher Nicht-Standardanwendungen (z. B. Satz von Baranyai, Gale-Ryser).

Netzwerk-Synthese wird in Kapitel 8 kurz behandelt, ein Gebiet, das wichtig, aber nicht allzu häufig in Büchern zu finden ist. Der bewußte Verzicht auf die lineare Optimierung führt jedoch leider dazu, daß der allgemeinste Fall nicht diskutiert werden kann.

In Kapitel 9 werden die algorithmischen Aspekte des Zusammenhangs von Graphen und Digraphen so gut wie vollständig abgehandelt.

Matchings (in Jungnickel's Buch Korrespondenzen genannt) werden in den Kapiteln 10 und 11 untersucht. Gerade bei der Behandlung von Verfahren zur Bestimmung gewichtsmaximaler Matchings muß dann doch ein kurzer Steilkurs in die Theorie der linearen und ganzzahligen Programmierung gemacht werden, ohne die die Edmonds'schen Resultate (und andere) schwer beschreibbar sind. Jungnickel hat sich hier durch einen lesenswerten Überblick (unter weitgehendem Verzicht auf Beweise) aus der „Affäre" gezogen. Eine tiefergehende Darstellung hätte mit Sicherheit den selbstgesteckten Rahmen gesprengt.

Die meisten in der Praxis auftretenden kombinatorischen Optimierungsprobleme sind (leider) $\mathcal{NP}$-vollständig. Es wäre also unredlich so zu tun, als wären die in Kapitel 1–11 untersuchten Probleme die wichtigsten Probleme im Anwendungsbereich der kombinatorischen Optimierung. Was diese Probleme auszeichnet, ist, daß für sie eine schöne und elegante Theorie existiert, die i. a. zu theoretisch und praktisch effizienten Lösungsverfahren führt. Ihre Bedeutung für die Praxis rührt daher, daß Kürzeste-Wege-Probleme, Netzwerkfluß-Probleme, die Berechnung minimaler aufspannender Bäume etc. dazu benutzt werden, um gute primale oder duale Heuristiken für praxisrelevante Probleme zu bekommen. Dies ist – aus anwendungsorientierter Sicht – die wesentliche Bedeutung der in diesem Buch behandelten polynomial lösbaren Fälle.

In Kapitel 12 behandelt nun Jungnickel – als exemplarisches Beispiel – das wohl bekannteste (schwere) kombinatorische Optimierungsproblem, nämlich das Travelling Salesman Problem und erwähnt, wie man einige der vorher besprochenen polynomialen Algorithmen zur „Lösung" des TSP benutzen kann. Er geht kurz auf die Theorie der $\mathcal{NP}$-Vollständigkeit ein, erläutert Relaxationen zur Bestimmung unterer Schranken und verschiedenste Heuristiken zur Berechnung oberer Schranken. Eine kurze Skizze von Branch & Bound wird ebenfalls gegeben.

Zusammenfassend möchte ich feststellen, daß das vorliegende Buch eine lesenswerte Einführung in die kombinatorische Optimierung ist, die ich als Lektüre empfehlen kann.

Kurz vor Drucklegung dieser Buchbesprechung ist bereits die zweite Auflage des Buches von Jungnickel erschienen. Der Autor hat im Haupttext einige kleinere Korrekturen vorgenommen und das Literaturverzeichnis auf den neuesten Stand gebracht (siehe z. B. obige Kommentare zu Netzwerkflußalgorithmen). Ein Anhang B ist hinzugekommen, in dem auf einige Themen des Hauptteils etwas vertiefter eingegangen wird und mehrere neue Problemkreise (z. B. Färbungen, Turnierpläne, eine Liste $\mathcal{NP}$-vollständiger Probleme) aufgenommen wurden.

Augsburg M. Grötschel

Budach, L., Graw, B., Meinel, Ch., Waack, S., Algebraic and Topological Properties of Finite Partially Ordered Sets (Teubner Texte zur Mathematik 109), Leipzig: BSB B. G. Teubner Verlagsgesellschaft 1988, 164 S., soft cover, 17,–

Finite partially ordered sets (posets) enter the stage of combinatorial mathematics from almost every direction, providing a wealth of examples,

- boolean algebras,
- distributive and modular lattices,
- subgroup lattices of finite groups,
- matroids and their lattices of flats,
- subspace lattices of finite vector spaces,
- set partitions (ordered by refinement),
- partitions of numbers (tableaux) under various partial orders,
- Coxeter groups with the "weak" and the Bruhat order,
- face lattices of convex polytopes and combinatorial spheres,

to name a few. They come in various disguises, and with a lot of interesting questions attached to them. With this arises the need for a comprehensive study and structure theory.

G.-C. Rota's fundamental paper [2] in 1964 observed that many fundamental enumeration problems can be formulated as inversion problems of sums over such partial orders, and that the Möbius function on these posets, defined recursively by

$$\mu(x, y) = \delta(x, y) - \sum_{x \leqslant z < y} \mu(x, z),$$

allows to solve these problems through the *Möbius inversion formula*. This leads to ask for techniques to compute Möbius functions, and for explanations how the structure of the posets considered determines the numerical invariants.

In fact, the Möbius function of a finite poset is the Euler characteristic of its simplicial complex of chains – this is Rota's interpretation of Hall's formula.

The book by Budach, Graw, Meinel und Waack reviewed here presents in its first chapter Rota's approach to the subject. It focuses on his famous theorem that posets related by a Galois connection have the same Möbius number.

In the following four chapters the book follows four alleys of research that have very successfully attempted to provide insight into the structure of posets that are related to Möbius inversion.

The first line of attack, through homological algebra, is due to K. Bacławski. It interprets posets as "small categories", associates with them "diagrams" (functors to a category of modules) and constructs a corresponding cohomology theory for which the Möbius function describes the Euler characteristic. The Galois connection theorem then follows from the Leray spectral sequence.

The second approach, through homotopy theory, was initiated by D. Quillen and shaped by A. Björner and J. Walker. It develops elegant techniques to determine the homotopy type of (the complex of chains of) a poset. These yield Quillen's "fiber theorem", which implies that posets related by a Galois connection are in fact homotopy equivalent, and because of this have the same Euler characteristic. The homotopy methods for the analysis of partial orders are in fact very powerful. We refer to Björner's [1] lucid survey for a demonstration of the scope of topological methods in various combinatorial settings.

The third line of investigation is algebraic, due to L. Solomon, C. Greene and K. Bacławski. It focuses on the matrix algebra ("incidence algebra") of functions on the intervals of a poset and its interesting structure. In this algebra, the Möbius function on the intervals arises as the matrix inverse of the characteristic function of the partial order. This setting provides very simple proofs for many important Möbius function identities.

The fourth perspective on posets is through commutative algebra. It relies on the observation that the Stanley-Reisner ring of a poset carries important structural information. Here the case where the ring is Cohen-Macaulay is of special interest. This very strong property is shared by many important examples – also by many of those mentioned above. It lies at the heart of Stanley's famous proof of the "Upper Bound Conjecture for spheres". Due to a combined effort of several brilliant researchers we have a topological characterization (by Reisner's theorem) and sophisticated verification methods (through Björner's lexicographic shelling technique) for this case. To readers who want to see more than a sketch we recommend Stanley [3] for an in-depth treatment of these connections between combinatorics and commutative algebra. S. Yuzvinsky's
recent work [5] connects it with the diagram cohomology theory mentioned above.

In the appendix Budach and his collaborators report on their own research. They apply technology developed in the previous chapters to problems of theoretical computer science. Here homological invariants of associated posets and complexes lead to lower bounds for the computational complexity of "classification problems".

"Algebraic and Topological Properties of Finite Partially Ordered Sets" is a journey through a wide and fascinating world of combinatorial mathematics that is certainly worthwhile. It is not going to be an easy journey, though.

One reason is that the reader is not "picked up at home": the book presupposes a basic understanding of the combinatorics of partially ordered sets. For this we recommend Stanley [4] as a brilliant exposition.

Furthermore some of the roads in this country are not quite smooth in the description Budach and his collaborators give us. The book has more typos and errors than might have been necessary, and at a few places it really leads the traveller astray.

Also the present book lacks concrete examples (Stanley's book provides them), which are important to illustrate the power and the scope of the technology developed. For example, even for Rota's basic Galois connection theorem it does not even describe a single application. This style at its worst instances gives the reader the feeling of reading, in his travel guide, the description of a map of the region.

To the experienced traveller the tour guide does not provide many new views of the landscape. For this the presentation is too close to the descriptions one might have read before, by explorers who originally discovered the area.

The country surveyed in this book is not flat, either. Its most beautiful (and famous) sights lie at considerable hight. To reach them requires a fair amount of technique (homological algebra, topology, commutative algebra). Thus the country itself makes the journey difficult. But its beauty partially lies in the way it links different areas (of mathematics) by strange, unexpected and fascinating connections. Its bridges and peaks surely are beautiful and certainly worth the journey.

[1] Björner A.: Topological methods, Handbook of Combinatorics, edited by R. Graham, M. Grötschel and L. Lovász. North Holland, to appear

[2] Rota G.-C.: On the foundations of combinatorial theory, I. Theory of Möbius functions. Z. Wahrscheinlichkeitstheorie **2** (1964) 340–368

[3] Stanley R. P.: Combinatorics and Commutative Algebra. Birkhäuser 1983

[4] Stanley R. P.: Enumerative Combinatorics, Vol. I. Wadsworth 1986

[5] Yuzvinsky S.: Cohen-Macaulay rings of sections. Advances in Mathematics **63** (1987) 172–195

Augsburg G. M. Ziegler

Rédei, L., Endliche p-Gruppen, Budapest: Académia Kiadó 1989, 304 S., DM 68,–

Endliche abelsche Gruppen erlauben bekanntlich die Wahl einer Basis derart, daß jedes Element genau eine Beschreibung als Produkt von Basiselementen hat. In dem vorliegenden Buch wird die Theorie der p-Gruppen entwickelt unter diesem Aspekt der eindeutigen Beschreibung der Elemente: Man geht von einer Normalreihe mit zyklischen Faktoren aus und beschreibt die Elemente als Produkte (fester Reihenfolge) von Potenzen der Erzeugenden dieser Faktoren. Dies Verfahren (mit Zentralfaktoren von Primzahlordnung) ist zur Beschreibung im Rechner seit langem in Gebrauch, vgl. Wamsley [4], Bayes, Kautsky, Wamsley [1], M. F. Newman [3], auch für auflösbare Gruppen gibt es entsprechende Verfahren, vgl. R. Laue [2]. Das Manuskript des vorliegenden Buches war auch schon vor dem Tode Rédeis 1980 beim Verlag und ist jetzt von seinen Schülern L. Márki und P. P. Pálfy herausgegeben worden.

Das Buch beschäftigt sich nach einer ausführlichen Einführung (95 Seiten) in die p-adischen Zahlen mit den oben beschriebenen Basen. Es ergeben sich Strukturkonstanten, die wegen der Assoziativität der Gruppe gewissen Einschränkungen unterliegen (II § 5).

Die im Rahmen dieser Betrachtungen auftretenden natürlichen Klassen sind nun die Klassen gleicher Minimallänge von Basen, die ersten zwei natürlichen Klassen sind also die zyklischen und die metazyklischen Gruppen.

Die Benutzung der Basen erlaubt die Angabe eines Erzeugendensystems für die absteigenden Glieder der Zentralreihe unter ausschließlicher Benutzung von Basiskommutatoren (Satz II.12.9). Neben der minimalen Basislänge l werden als weitere Unterteilung „Defektklassen" eingeführt, die durch Eigenschaften der zugehörigen Koeffizientenmatrix bestimmt werden.

Die Klassifizierung der metazyklischen Gruppen im dritten Teil des Buchs erntet, was zuvor gesät wurde, für den ersten nichttrivialen Fall (Satz III.6.1).

Ein ergänzender Lektürestoff für diejenigen, die sich näher mit p-Gruppen befassen, mit und ohne Rechner.

[1] Bayes, A. J., Kautsky, J., Wamsley, J.: Computation in nilpotent groups (Application). Proceedings Second International Conference on theTheory of Groups, Canberra 1973. Lecture Notes in Mathematics **372** (1974) 82–89

[2] Laue, R.: Zur Konstruktion und Klassifikation endlicher auflösbarer Gruppen. Bayreuther Mathem. Schriften, Heft 9, 1982, 304 S.

[3] Newmann, M. F.: Determination of groups of prime power order. Proceedings Miniconference on the Theory of Groups, Canberra 1975. Lecture Notes in Mathematics **573** (1977) 73–84

[4] Wamsley, J. W.: Computation in nilpotent groups (Theory). Proceedings Second International Conference on the Theory of Groups, Canberra 1973. Lecture Notes in Mathematics **372** (1974) 691–700

Würzburg H. Heineken

Bigalke, H. G., Heinrich Heesch, Kristallgeometrie, Parkettierungen, Vierfarbenforschung (Vita Mathematica) Basel: Birkhäuser Verlag, 1988, 319 S., DM 78,–

Man nimmt dieses Buch in die Hand und kann es nicht mehr weglegen. Eine Vita zieht an uns vorbei, die ihresgleichen sucht: Die endlos ferne Zeit der Jahre vor 1914, die Musik, die für Heesch vielerorten das gesellschaftliche Entrée bedeutete, die großen Lehrer, die Physiker Sommerfeld und Wien, Tietze, Perron und Carathéodory in München, Weyl, Speiser und Wentzel in Zürich. Sein Violinlehrer Berber, der Heeschs fehlende Sinnlichkeit im Spiel mit dem Satz kommentierte: „Ein Student hat eine Studentin ermordet. Da hab' ich gleich zu meiner Frau gesagt: Der Heesch ist das nicht gewesen!" Und dann Heeschs heute undenkbare Einstellung zur „Karriere". Z. B. Habilitation: „Ich betrachte die Habilitation als Hilfsschritt für die Lösung der finanziellen Frage..." (Seite 162); z. B. Publikation: „... der Tatsache, daß

Heesch als überaus kreativer Mathematiker und besessener Arbeiter keine Zeit dafür zu haben glaubte, seine Ergebnisse anderen ... näherzubringen." (Seite 178); z. B. Förderung: „Da die Verneinung durch den Hauptausschuß und die neuen Herren Gutachter nicht aus Kritik an meiner Arbeitsweise kommen, sondern aus der Meinung, mein Forschungsgebiet sei unbedeutend, begründet wird, bin ich ... nicht fähig, neue Argumente zu erbringen." (S. 223). Die Liste ließe sich verlängern.

Bigalkes Buch weist alle Vorzüge und Mängel einer mehr oder weniger mündlich überlieferten Historie auf. Er nimmt für seinen Helden deutlich Partei. Diese Subjektivität der Darstellung kommt der Lebendigkeit zugute, man kann nicht anders, als für das Schicksal von Heesch eingenommen zu sein, in seinen glücklichen und noch mehr in seinen tragischen Augenblicken. Der historischen Authenthizität ist diese Betrachtungsweise nicht immer förderlich, vor allem dort, wo (vergeblich) versucht wird, allgemeine Geschichte mit der Biographie Heeschs zu verknüpfen. Als Beispiel seien die Ereignisse 1933/34 in Göttingen angeführt, die mittlerweile ja gut dokumentiert vorliegen (vergleiche etwa Schappacher). Bigalkes Bericht z. B. über den Landau-Boykott (Seite 105ff.): „Als die Demonstrationen [gegen Landau] kein Ende nehmen wollten, erklärte Landau seinen Rücktritt..." ist Fiktion: Es gab *einen* Boykott, und der ließ Landau den Rücktritt einreichen. Oder „In der Folgezeit emigrierten immer mehr Wissenschaftler" ist falsch, weil der Landau-Boykott fast schon der Schlußstein der Vertreibungen aus Göttingen war. Auch die Chronologie gegen Hasse 1934 samt der berüchtigten verweigerten Schlüsselübergabe und der Rolle von Teichmüller ist nicht richtig wiedergegeben. Auch diese Liste ließe sich verlängern.

Heesch war ein Künstler: In der Musik, in der Mathematik und im Leben, an das er nur kargste Ansprüche geltend machte. Kommen wir also zur Mathematik; auch hier war der Übergang zur Kunst fließend. In der Kristallographie war es die Symmetrie, die ihn fesselte; seine regulären Parkettierungen erschienen als Fliesenwand im Erweiterungsbau des Stadthauses Göttingen (Seite 127) und als bemalte Kacheln; und wer den Ausschnitt der Liste der z-positiven Figuren beim 4-Farben-Problem auf Seite 203 ansieht, ist an chinesische Kalligraphien erinnert. Das Klischee vom tragischen Künstler wird für Heesch in seiner Arbeit zum 4-Farben-Problem Wirklichkeit. Heute steht außer Frage, daß die Lösung 1976 durch Appel und Haken ohne die Heesch'schen Vorarbeiten zur Reduktion und Unvermeidbarkeit nicht möglich gewesen wäre. Ob *er* es geschafft hätte, bei besserer und weiterer Unterstützung durch die mathematische Öffentlichkeit in Deutschland, läßt sich nicht beantworten. Daß seine Arbeit nicht gebührend anerkannt wurde, ist jedenfalls eine Tatsache, die nachdenklich macht über uns selbst und über unser Förderungssystem. So lesen wir z. B. auf Seite 222 in einem Gutachten der DFG: „Der wesentliche Punkt dieser „großen Probleme" [gemeint ist das 4-Farben-Problem] in der Wissenschaft sei nicht, sie zu lösen, sondern der Wert des Unternehmens bestehe in dem methodischen Durchbruch, der auf weite Gebiete der Wissenschaft auszustrahlen habe. Fehle ein solcher methodischer Durchbruch, dann sei die Tatsache allein, daß das Problem endlich gelöst werden könne, kaum etwas wert." Jeder Leser kann sich auf diese und ähnliche Sätze, die über Wohl und Wehe von Forschungsprojekten entscheiden, seinen eigenen Reim machen. Allerdings muß man fairerweise hinzufügen, daß Heesch es den Referenten auch nicht gerade leicht gemacht hat, da er sich beharrlich zierte, seine Ergebnisse vorzulegen: „Wie Sie richtig sagen, halte ich mit Publizieren der Reduktionen zurück ..." (Seite 245), in unserer Zeit des 'publish oder perish' wahrhaftig eine singuläre Einstellung.

Ja, Heinrich Heesch war und ist ein singulärer Mensch. Sein Lebenslauf mutet neben unseren wohlgeordneten und gleichförmigen Schicksalen wie ein Elementarereignis an. Wie gesagt, man kann das Buch, einmal begonnen, nicht mehr aus der Hand legen – und dafür gebührt seinem Biographen Bigalke unser aller Dank.

Berlin M. Aigner

Burago, Yu. D., Zalgaller, V. A., Geometric Inequalities (Grundlehren der mathematischen Wissenschaften. A Series of Comprehensive Studies in Mathematics, Band 285), Berlin u. a.: Springer-Verlag 1988, 43 figs., 331 pp., Hard cover, DM 184,–

Das in seiner Vielfalt fast unübersehbare Gebiet der geometrischen Ungleichungen ist sowohl von hohem ästhetischem Reiz – man denke nur an all die Kennzeichnungen besonders symmetrischer Objekte durch Extremaleigenschaften – als auch beeindruckend durch den Reichtum an Ergebnissen und Hilfsmitteln, die es für Anwendungen etwa in der Analysis bereitstellt. Das vorliegende Buch, eine überarbeitete und in zwei Paragraphen abgeänderte Übersetzung des 1980 erschienenen russischen Originals, hat sein Schwergewicht auf Ungleichungen aus der Differentialgeometrie und behandelt auch Ungleichungen der Konvexgeometrie und allgemeineren Punktmengengeometrie sowie der geometrischen Maßtheorie. Ausgespart bleiben also zum Beispiel elementargeometrische Ungleichungen, Extremalprobleme der mathematischen Physik und Abschätzungen aus der diskreten oder der kombinatorischen Geometrie. Das erste Kapitel gibt eine ausführliche Darstellung von Ungleichungen auf zweidimensionalen Flächen. In sehr allgemeinem Rahmen (Mannigfaltigkeiten beschränkter Krümmung im Sinne von A. D. Alexandrov) werden vorwiegend Flächenabschätzungen im Vergleich mit anderen Größen der inneren und äußeren Geometrie behandelt. Kapitel 2 ist der Brunn-Minkowskischen Ungleichung und der klassischen isoperimetrischen Ungleichung (für die Minkowski-Oberfläche) gewidmet. Auch hier ist der Rahmen allgemein; es werden kompakte und allgemeinere Mengen in den einfach zusammenhängenden Räumen konstanter Krümmung zugelassen. Die erfolgreichen Symmetrisierungsverfahren werden dargestellt, und die Gleichheitsfälle werden diskutiert. Auch andere klassische Volumenabschätzungen werden erwähnt. Das dritte Kapitel befaßt sich mit isoperimetrischen Ungleichungen unter Zugrundelegung verschiedener Oberflächenbegriffe. Nach Bereitstellung von Hilfsmitteln aus der geometrischen Maßtheorie (Eilenbergs Ungleichung, Federers Koflächenformel) werden Mengen beschränkter Oberfläche (im Sinne von de Giorgi), Ströme (currents) und Varifaltigkeiten (varifolds), Lebesguesche Oberfläche und integralgeometrische Oberflächenmaße behandelt und jeweils isoperimetrische Ungleichungen hierfür, ferner der Zusammenhang mit Einbettungssätzen. Kapitel 4 ist der Theorie der gemischten Volumina konvexer Körper gewidmet. Ein Kernstück sind hier die Alexandrov-Fenchelschen Ungleichungen und ihre Folgerungen. Für die Ungleichungen wird in einem von A. G. Khovanskii verfaßten Paragraphen einer der neuen Beweise unter Verwendung von Methoden der algebraischen Geometrie (Index-Ungleichung von Hodge) gegeben. So faszinierend dieser Zusammenhang ist, der Aufwand ist, verglichen mit Alexandrovs Beweisen, doch recht erheblich, und vom Standpunkt der Konvexgeometrie aus gesehen ist dieser Zugang wohl eher ein Umweg. Das Kapitel enthält viele weitere Ungleichungen über konvexe Körper und in Anhängen Material über Gegenstücke (zum Beispiel gemischte Diskriminanten) und Verallgemeinerungen gemischter Volumina, nicht alles unbedingt in Bezug zu Ungleichungen. Die beiden letzten Kapitel sind wieder differentialgeometrischer Natur. Das fünfte, über Immersionen in den $\mathbb{R}^n$, behandelt isoperimetrische Ungleichungen unter Einbezug der mittleren Krümmung, Resultate über totale Absolutkrümmung und Ungleichungen zwischen Größen der inneren Metrik und Einbettungsgrößen. Im sechsten Kapitel werden Ungleichungen der Riemannschen Geometrie untersucht. Hier findet man Vergleichssätze für Jacobi-Felder, Volumenabschätzungen von oben und unten, Ungleichungen für die mittlere Krümmung von Immersionen in Riemannsche Räume und vor allem die untere Abschätzung des Volumens einer geschlossenen Riemannschen Mannigfaltigkeit negativer Krümmung nach Margulis und Gromov.

Der Charakter des Werkes ist bewußt nicht einheitlich gehalten. Die Darstellung wechselt in kurzweiliger Weise zwischen Einführung, mehr oder weniger ausführlicher Darstellung einzelner tiefer liegender Beweise und aufzählender Übersicht mit historischen

Bemerkungen. Naturgemäß sind die Anforderungen, die an den Leser gestellt werden, in den einzelnen Teilen recht unterschiedlich. Die Literaturhinweise sind sehr zahlreich, wenn auch die Autoren keine enzyklopädische Vollständigkeit angestrebt haben.

Dieses Buch, zu dem es wohl kein vergleichbares gibt, besticht durch seine Vielseitigkeit und durch die Substanz der dargestellten Ergebnisse. Es umfaßt ein Stück bester und lebendiger Geometrie. Dem allgemeinen mathematischen Leser kann es einen guten Eindruck von einem zentralen Teil dieses Gebietes vermitteln. Die Spannweite der Thematik ist aber so groß, daß auch sicher jeder Experte vieles finden wird, das ihm neu ist, und er wird die Lektüre als anregend empfinden.

Freiburg R. Schneider

Stöcker, R., Zieschang, H., Algebraische Topologie. Eine Einführung (Mathematische Leitfäden), Stuttgart: B. G. Teubner 1988, 414 pp, softcover, DM 52,–

Wer hat wohl noch nicht gerätselt, wann und von wem ein solches Buch erscheinen wird. Aus dem Vorwort: „Dieser Text ist eine Einführung in die algebraische Topologie ... Mit den zentralen Begriffen ‚Homotopie‘ und ‚Homologie‘ werden tiefliegende Eigenschaften topologischer Räume beschrieben. Um das möglichst einfach deutlich zu machen, haben wir die Sätze nicht immer in ihrer vollen Allgemeinheit bewiesen. Ein Beispiel mehr war uns oft lieber als eine Voraussetzung weniger.“ Damit sind Umfang und Stil des Buches angedeutet. Ein Hinweis, daß es hier um anderes geht als bei Boto, findet sich auch gleich.

Der erste Teil des Buches „Geometrisch-topologische Vorbereitungen“ besteht aus vier Kapiteln, einem einführenden mit vielen Beispielen und einer Diskussion des Homöomorphieproblems: dann Kap. 2 „Homotopie“, Kap. 3 „Simplizialkomplexe und Polyeder“ und Kap. 4 „CW-Räume“. Auf Seite 52ff. werden z. B. Homotopieklassen von Selbstabbildungen von S^1 klassifiziert, die klassischen Anwendungen gegeben.

Teil II behandelt ausführlich in Kap. 5 „Die Fundamentalgruppe“ und in Kap. 6 „Überlagerungen“. Damit ist man Seite 175 bei Teil III „Homologietheorie“ angekommen. „Wir folgen der historischen Entwicklung und führen zuerst die geometrisch sehr anschaulichen Homologiegruppen von Simplizialkomplexen ein.“ In sechs Kapiteln wird die simpliziale und singuläre Homologietheorie von Grund auf entwickelt bis zum Satz von Eilenberg-Zilber mit den klassischen Anwendungen.

Der vierte Teil besteht aus drei Kapiteln: „Cohomologie“, „Dualität in Mannigfaltigkeiten“ (wobei Triangulierbarkeit vorausgesetzt wird) und „Der Cohomologiering“. Abschnitt 15.6 trägt den Titel „Cap-Produkt und Verschlingungszahlen.“

Fast überflüssig zu sagen, daß das Buch klar und übersichtlich verfaßt und gestaltet ist. Zu Beginn jeden Kapitels klärt eine kleine Vorbemerkung die Erwartungen, am Ende findet man jeweils interessante (historische) Notizen. Aufgaben sind reichlich vorhanden.

Ich werde das Buch z. B. so benutzen, daß ich nach einem einführenden Kurs nach K. Jänich's „Topologie“ im zweiten Semester sofort auf Seite 175 einsteigen werde.

Berlin (FU) H. Scheerer

Rotman, J. J., An Introduction to Algebraic Topology (Graduate Texts in Mathematics), New York etc.: Springer-Verlag 1988, 433 pp., hardcover, DM 108,–

Konzeption und Inhalt dieses Buches sind sehr ähnlich denen des oben besprochenen von R. Stöcker und H. Zieschang. Es ist ebenfalls sorgfältig gearbeitet. Statt eines Exkurses über Dualität enthält es einen über Homotopiegruppen mit einem Abschnitt „A

Glimpse Ahead". Die Fundamentalgruppe von S^1 wird auf Seite 50ff. berechnet. (Eine sorgfältige Einführung in die algebraische Topologie scheint nicht allzuviel Wahlen zu lassen.)

Es gibt zwölf Kapitel; die Überschriften lauten wie folgt: Introduction. Some Basic Topological Notions. Simplexes. The Fundamental group. Singular Homology. Long Exact Sequences. Excision and Applications. Simplicial Complexes. CW-complexes. Natural Transformations. Covering Spaces. Homotopy Groups. Cohomology.

Viele Übungsaufgaben gehören dazu. Die singuläre Homologietheorie kommt diesmal vor der simplizialen und auch sonst ist die Ordnung etwas anders als im oben besprochenen Buch.

„The basic outline of this book corresponds to the syllabus of a firstyear's course in algebraic topology designed by geometers and topologists at the University of Illinois, Urbana."

Nur die Höhe des Preises stört.

Berlin (FU) H. Scheerer

Goresky, M., MacPherson, R., Stratified Morse Theory (Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Band 14, A Series of Modern Surveys in Mathematics), Berlin u.a.: Springer-Verlag 1988, 272 pp., Hard cover, DM 148,–

Der Titel des Buches deckt sich mit demjenigen einer kurzen Arbeit, in der die Autoren bereits wesentliche Resultate angekündigt hatten [Proc. of the A.M.S. Summer Institute on Singularities (Arcata 1981), part 1, pp. 517–533]. Das zentrale Ergebnis ist eine durchgreifende Verallgemeinerung der klassischen Morsetheorie; der entsprechende Hauptsatz wird im 1. Teil des Buches bewiesen.

Bei der klassischen Morsetheorie gibt man sich eine eigentliche Funktion f vor, die auf einer differenzierbaren Mannigfaltigkeit M definiert ist und gewisse zusätzliche Eigenschaften erfüllt; f heißt dann eine Morsefunktion, und das Ziel ist der Vergleich der Teilräume f^{-1} ($]-\infty, a]$) für verschiedene Zahlen a. Die Verallgemeinerung der klassischen Morsetheorie erfolgt nun in drei Richtungen: es werden singuläre Räume zugelassen, genauer: M wird durch eine Whitney-stratifizierte Teilmenge X einer Mannigfaltigkeit ersetzt (daher der Name „stratifizierte Morsetheorie"); es wird eine relative Version der Morsetheorie betrachtet, bei der die Morsefunktion auf dem Wertebereich einer eigentlichen Abbildung definiert ist; und schließlich werden nichteigentliche Morsefunktionen zugelassen, die dadurch entstehen, daß gewisse Strata entfernt werden.

Um die Sache nicht von vornherein zu kompliziert zu machen, konzentrieren sich die Autoren zunächst auf die zuerst genannte Art der Verallgemeinerung und formulieren in Sect. 3.7 einen entsprechenden Hauptsatz: Sei S ein Stratum von X und p ein kritischer Punkt von $f|S$. Es sollen die Räume $X_{c-\varepsilon}$ und $X_{c+\varepsilon}$ verglichen werden; dabei sei $\varepsilon > 0$ genügend klein, $c = f(p)$, $X_a = f^{-1}(]-\infty, a])$. Zu diesem Zweck wird das „lokale Morsedatum" von f in p eingeführt; dabei handelt es sich um ein Raumpaar (A, B) derart, daß $X_{c+\varepsilon}$ bis auf Homöomorphie aus $X_{c-\varepsilon}$ durch Anfügen von A längs B entsteht. (Im klassischen Fall ist ein Henkel anzufügen.) Nun ist X längs S lokal homöomorph zu einem Produkt von S und N, der normalen Scheibe bezüglich S in X. Der Hauptsatz von Goresky und MacPherson besagt nun, daß man eine analoge Produktstruktur für das lokale Morsedatum hat: bezeichnet man das lokale Morsedatum von $f|S$ bzw. $f|N$ als tangentiales bzw. normales Morsedatum, so ist das lokale Morsedatum von f in p homöomorph zum Produkt von diesen beiden. Entsprechende Sätze werden später im relativen und nichteigentlichen Fall bewiesen (Sect. 9.5, 10.5, 11.5).

Der 2. Teil des Buches enthält Anwendungen auf komplexanalytische Varietäten. Eine Vereinfachung ergibt sich hierbei dadurch, daß das normale Morsedatum nicht mehr von der Morsefunktion abhängt. Bei den Anwendungen handelt es sich vor allem um Verallgemeinerungen bzw. Varianten des Lefschetz-Satzes über Hyperebenenschnitte und entsprechende duale Sätze. In diesem Zusammenhang gibt es in der Literatur verwandte Methoden und Ergebnisse; wichtig ist hier vor allem, daß die im 1. Teil entwickelte stratifizierte Morsetheorie eine Übertragung der klassischen Beweismethoden gestattet. Dabei genügt es, den Homotopietyp zu beachten – eine Vereinfachung, die bei der anschließenden Herleitung entsprechender Sätze für die Schnitthomologie nicht ohne weiteres möglich ist.

Der 3. Teil enthält schließlich eine Anwendung der stratifizierten Morsetheorie im reellen Fall. Es geht hierbei um die Homologie des Komplements von endlich vielen affinen Teilräumen von $\mathbb{R}^n$. Dabei ist es wichtig, daß bei der stratifizierten Morsetheorie auch nichteigentliche Morsefunktionen zugelassen sind.

Die Resultate des 2. und 3. Teils stellen eine eindrucksvolle Rechtfertigung für den Aufwand dar, der im 1. Teil bei der Entwicklung der stratifizierten Morsetheorie betrieben wurde. Für den Leser ist es ein großer Vorteil, daß der eingangs erwähnte Hauptsatz geometrisch gut faßbar ist. Weniger leicht ist es aus Dimensionsgründen, sich ein angemessenes Bild von der Ausgangssituation bei dem Beweis zu machen. Es sind Räume zu vergleichen, die mit Hilfe diverser Hilfsfunktionen definiert sind; diesem Zweck dient ein neues Hilfsmittel („Moving the wall").

Der zentralen Bedeutung des 1. Teils entsprechend haben die Autoren hier besonderen Wert auf die Vollständigkeit der Darstellung gelegt. Im 2. Teil werden Detailüberlegungen gelegentlich dem Leser überlassen; bei der Schnitthomologie wird verstärkt auf die Literatur verwiesen.

Es ist hervorzuheben, daß die Autoren auch relativ ausführlich auf die historische Entwicklung eingehen. Es wird auf eine Vielzahl einschlägiger Arbeiten verschiedener Autoren hingewiesen, dementsprechend umfangreich ist das Literaturverzeichnis.

Bei dem vorliegenden Buch handelt es sich im Kern um eine Originalarbeit, die in einer Weise vervollständigt wurde, daß sich eine thematisch in sich geschlossene Darstellung des behandelten Gebiets ergab. Das Buch enthält daher nicht zuletzt eine zuverlässige Darstellung grundlegender Sätze und Begriffe. Die Lektüre der Beweise stellt erhebliche Anforderungen an den Leser, vor allem, wenn dieser mit einschlägigen Techniken (z. B. aus der Differentialtopologie) nicht allzusehr vertraut ist. Eine wichtige Orientierungshilfe bilden die häufig vorangestellten Einleitungen.

Alles in allem handelt es sich wegen der wichtigen Thematik und der gründlichen Darstellung um ein Standardwerk, das nicht nur für Fachleute interessant ist und das daher in keiner Bibliothek fehlen darf.

Münster H. A. Hamm

Berger, M., Gostiaux, B., Differential Geometry: Manifolds, Curves, and Surfaces (Graduate Texts in Mathematics, Vol. 115), New York – Berlin – Heidelberg – London – Paris – Tokyo: Springer Verlag 1988, 249 fig. xii, 474 pp., Hard cover, DM 98,–

Das Buch besteht aus zwei ziemlich unterschiedlichen Teilen: Die Kapitel 0 bis 9 sind eine etwas erweitere und überarbeitete Fassung der „Géométrie Différentielle", erschienen 1972 in der Librairie Armand Colin.

Dabei waren die Flächen im 3-dimensionalen Raum zu kurz gekommen. Dieses ist nun in dem neu geschriebenen zweiten Teil, bestehend aus den Kapiteln 10 und 11, aufs beste behoben.

In der vorliegenden Form ist das Buch einzigartig: Die grundlegenden Fakten der Analysis auf Mannigfaltigkeiten sind in großer Allgemeinheit sorgfältig hergeleitet. Wenn es dann zu Anwendungen kommt oder zu tieferliegenden Resultaten, so wird auf detaillierte Beweise verzichtet. Stattdessen wird – meist mit gutem Erfolg – versucht, die Behauptungen plausibel zu machen; und stets werden sorgfältig Verweise auf die Literatur gegeben.

Insbesondere in dem neu hinzugekommenen Teil über Flächen findet sich auf diese Weise eine Fülle von hochinteressantem und ganz aktuellem Material, das nirgends sonst sich zusammengestellt findet. So ist dies Buch nicht nur für einen Studenten nützlich, sondern auch für den Experten, der hier findet, was er immer schon genau wissen wollte, sich aber scheute, danach in seiner Sammlung von Sonderdrucken zu suchen oder einen Kollegen zu fragen. Wer immer noch Zweifel hat, ob die Differentialgeometrie lebendig ist und interessant, der lasse sich durch ein Studium dieses Buches eines besseren belehren!

Bonn

W. Klingenberg

Anosov, D. V., Arnold, V. I. (Eds.), Dynamical Systems I – Differential Equations, Dynamical Systems (Encyclopaedia of Mathematical Sciences, Vol. 1), Berlin u. a.: Springer-Verlag 1988, 25 figs, approx. 250 pp, hard cover, DM 128,–

Dieses Buch ist jetzt als zweites der mehrbändigen Reihe über dynamische Systeme in der „Encyclopaedia of Mathematical Sciences" beim Springer-Verlag erschienen. Während im zuerst erschienenen Band III, der den Untertitel „Mathematical Aspects of Classical and Celestial Mechanics" trägt, das Gewicht auf praktischen Methoden zur Behandlung mechanischer Bewegungsgleichungen liegt, handelt der vorliegende Band I hauptsächlich von grundlegenden Fakten und allgemeinen Hintergründen der Theorie gewöhnlicher Differentialgleichungen. Tatsächlich ist das Buch in zwei unabhängige Teile unterteilt, nämlich in einen Artikel von V. I. Arnold und Yu. S. Il'yashenko über „Ordinary Differential Equations" und einen Artikel von D. V. Anosov, I. V. Bronshtein, S. Kh. Aranson und V. Z. Grines über „Smooth Dynamical Systems". Beide sind Übersichtsartikel: Beweise werden nur gelegentlich erwähnt.

Im ersten Artikel werden gewöhnliche Differentialgleichungen vom Standpunkt der lokalen, qualitativen Theorie aus behandelt. Sieht man von linearen Gleichungen und Gleichungen auf speziellen Flächen wie der Ebene, der 2-Sphäre und dem 2-Torus ab, so werden globale Eigenschaften ebensowenig behandelt wie die tatsächliche Integration einzelner Gleichungen. Wie schon in Band III ist es den Autoren wieder meisterhaft gelungen, älteren Stoff mit neueren Entwicklungen in einer einheitlichen Darstellung miteinander zu verknüpfen. Auch der Einfluß von Disziplinen wie Algebra, Algebraische Geometrie und Komplexe Analysis zur Lösung klassischer Probleme aus der Theorie gewöhnlicher Differentialgleichungen wird zum Ausdruck gebracht. In Kapitel 1 werden Grundbegriffe bis hin zum „Attraktor" sowie die wichtigsten allgemeinen Sätze der klassischen Theorie behandelt. Das eigentliche Kernstück des Artikels ist die Klassifikation von singulären Punkten und von Zykeln. Diesem Fragenkreis sind insgesamt vier Kapitel gewidmet. Davon beziehen sich drei auf singuläre Punkte, wobei zwischen höher-dimensionalen reellen und komplexen sowie ebenen Phasenräumen unterschieden wird. Die Klassifikation beruht auf gewissen Konjugationsbegriffen, zugehörigen Normalformen und Generizitätsannahmen. Verzweigungen werden außer Acht gelassen. Der Leser bekommt einen exzellenten Überblick über verschiedene Normalformtheorien. Die Theorie invarianter Mannigfaltigkeiten ist darin einbezogen. Erwähnt sei auch, daß das Endlichkeitsproblem für Grenzzykel ausführlich diskutiert wird. Der Artikel schließt mit einem Kapitel über die asymptotische Theorie gewöhnlicher Differentialgleichungen mit komplexer Zeit (Fuchssche Theorie).

Der zweite Artikel stellt eine ideale Ergänzung zum ersten dar. Hier werden abstrakte Flüsse auf allgemeinen Mannigfaltigkeiten behandelt. Dabei wird zwischen kontinuierlichen und diskreten Flüssen unterschieden. Letztere werden Kaskaden genannt. Es werden wie im ersten Artikel generische Phänomene und strukturell stabile Systeme, z. B. Morse-Smalesysteme, untersucht und klassifiziert. Die verwendeten Konzepte sind hier aber größtenteils topologischer Natur und stammen zum Teil aus der algebraischen Topologie und der Differentialtopologie. Einen breiten Raum nehmen verschiedene Indextheorien ein. Das reicht vom Morseindex für singuläre Punkte bis hin zum Conleyindex für isolierte invariante Mengen. Indizes sind nicht nur für die Klassifizierung, sondern wegen ihrer Homotopieinvarianz auch für Fragen der Fortsetzbarkeit dynamischer Eigenschaften von Bedeutung. Die allgemeine Theorie wird schließlich am Beispiel von Flüssen auf allgemeinen Flächen eindrucksvoll illustriert. Die hyperbolische Theorie glatter dynamischer Systeme, die sonst in der Literatur einen breiten Raum einnimmt, wird hier nicht behandelt.

Das Buch ist nicht nur ein hervorragendes Nachschlagewerk, sondern könnte auch als Leitfaden zur Stoffauswahl für einen modernen Kurs über gewöhnliche Differentialgleichungen dienen.

Hamburg J. Scheurle

Krylov, N. V., Nonlinear elliptic and parabolic operators of the second order (Mathematics and Applications Soviet Series), Dordrecht – Boston – Lancaster – Tokyo: D. Reidel Publishing Co. 1987, 484 pp., Dfl. 245.00

Wenige Jahre nach dem Erscheinen der richtungsweisenden Arbeiten von Evans und Krylov über die Lösung vollständig nichtlinearer elliptischer und parabolischer Gleichungen liegt Krylovs systematische, auf große Allgemeinheit angelegte Darstellung der Theorie derartiger Differentialgleichungen nun in einer englischen Übersetzung vor. Damit erschließt sich zugleich für viele Analytiker im Westen erstmalig Krylovs Werk auf dem Gebiet der partiellen Differentialgleichungen, das zuvor nahezu ausschließlich in russischer Sprache erschienen war.

Krylov ist bislang im Westen hauptsächlich auf dem Gebiet der Stochastischen Kontrolltheorie hervorgetreten. Seine fundamentalen Beiträge zur Theorie der partiellen Differentialgleichungen belegen (wie auch die Arbeiten von Evans, Friedman, Lions u. a.) die enge Beziehung zwischen Stochastik und Analysis. Dennoch wirkt die Sprache der Stochastik in der Analysis oft fremd. Von besonderer Bedeutung ist daher, daß es Krylov im vorliegenden Buch geglückt ist, für fast alle Sätze analytische Beweise anzugeben und die ursprünglich stochastischen Methoden zu ersetzen.

Bereits beim Durchblättern des Inhaltsverzeichnis fällt die programmatische Ähnlichkeit mit den Büchern von Ladyženskaja, Solonnikov und Ural'ceva auf, die in den 60er Jahren aus den bahnbrechenden Arbeiten von De Giorgi und Nash Universalwerkzeuge zur Lösung quasilinearer elliptischer und parabolischer Differentialgleichungen in Divergenzform schufen.

Kapitel 1 enthält auf gut 20 Seiten die nötigsten Notationen und Definitionen.

In Kapitel 2 werden die Konstruktion des Lösungsoperators für die Hamilton-Jacobi-Bellman-Gleichung im $\mathbb{R}^n$ dargestellt und Abschätzungen entwickelt, die für nichtlineare Differentialgleichungen dieselbe Rolle spielen wie Abschätzungen für die Greensche Funktion im linearen Fall.

Diese Abschätzungen werden in Kapitel 3 ausgenützt, um a-priori-Abschätzungen (Alexandrov-Bakelman-Maximumprinzip) des Supremums der Lösungen nichtlinearer

Gleichungen herzuleiten und um daraus Vergleichprinzipien, Eindeutigkeits- und Konvergenzsätze zu gewinnen. Anders als Alexandrov, Bakelman oder Pogorelov in ihren Arbeiten zur Monge-Ampère-Gleichung erhält Krylov diese a-priori-Schranken nicht auf geometrischem Wege, sondern indem er eine Korrespondenz zwischen der Hamilton-Jacobi-Bellman-Gleichung und der Monge-Ampère-Gleichung ausnutzt, die auf der Ungleichung zwischen geometrischem und arithmetischem Mittel beruht.

Kapitel 4 enthält den Beweis der Harnack-Ungleichung für Lösungen linearer elliptischer Differentialgleichungen ohne Divergenzstruktur mit lediglich meßbaren Koeffizienten. Diese fundamentale Abschätzung wurde erstmalig von Krylov und Safonov bewiesen. Zugleich bereitet der Abschnitt auf die $C^{2+\alpha}$-a-priori-Abschätzungen und Existenzsätze der folgenden Kapitel 5 und 6 vor. Dabei soll auf die kleine Beispielsammlung (S. 301–305) in Kapitel 6 hingewiesen sein, die die Spannbreite der Methoden und Resultate erahnen läßt.

Die dargestellten Methoden sind auch auf entartet elliptische oder parabolische Probleme anwendbar. Dieser Problemkreis wird in Kapitel 7 erörtert.

Abschließend werden in Kapitel 8 Existenzsätze für (möglicherweise entartete) Differentialgleichungen insbesondere vom Monge-Ampère-Typ angegeben.

Das Werk beeindruckt durch die Fülle und Tiefe des dargestellten Materials. Die für den Spezialisten außerordentlich lohnende Lektüre ist aber alles andere als leicht. Abgesehen von gewissen (aus der Sicht des Referenten) unglücklich gewählten Notationen trägt dazu der unübersichtliche Druck des Werkes bei. So bedarf es z. B. bereits einiger Mühe, eine angegebene Referenz im Literaturverzeichnis aufzuspüren. Die wenigen unauffällig eingestreuten Beispiele sind darüber hinaus kaum geeignet, den Reichtum der zu erwartenden Phänomene adäquat zu illustrieren oder die Grenzen der gewonnenen Resultate aufzuzeigen. In dieser Beziehung hält Krylovs Buch dem Vergleich mit den eingangs zitierten Werken von Ladyženskaja et al. kaum stand.

Als Einführung in den Problemkreis der vollständig nichtlinearen elliptischen Gleichungen würde der Referent jedenfalls die Darstellung von Gilbarg-Trudinger (Grundlehren 224, Springer, 2. Ausgabe (1983)) vorziehen, obwohl man dort z. B. die von Krylov bewiesenen inneren $C^{2+\alpha}$-Abschätzungen für gleichmäßig elliptische Gleichungen noch nicht finden kann.

Zürich M. Struwe

Gangolli, R. A., Varadarajan, V. S., Harmonic Analysis of Spherical Functions on Real Reductive Groups (Ergebnisse der Mathematik und ihrer Grenzgebiete. 2. Folge, Vol. 101), Berlin u.a.: Springer-Verlag 1988, 362 pp., Hard cover, DM 198,–

Die harmonische Analyse der sphärischen Funktionen ist ein Treffpunkt zwischen der Theorie der Gruppendarstellungen und der Analyse der Entwicklungen durch Eigenfunktionen von Differentialoperatoren. Man kann annehmen, daß die moderne Geschichte der sphärischen Funktionen in den dreißiger Jahren beginnt und zwar mit den Arbeiten von Hermann Weyl über die kompakten Gruppen und von Elie Cartan über die kompakten symmetrischen Räume. Die Rolle der sphärischen Funktionen in der harmonischen Analyse der halbeinfachen Lie'schen Gruppen wurde von Gel'fand und Godement in den fünfziger Jahren klar hervorgehoben. Diese sind im Mittelpunkt des monumentalen Werkes von Harish Chandra.

Die Theorie der sphärischen Funktionen wurde schon in mehreren Monographien dargelegt, insbesondere in dem Buch von Helgason, *Groups and geometric analysis* (Academic Press, 1984). Das Buch von Gangolli und Varadarajan gibt eine sehr vollständige

Darstellung der Arbeiten von Harish Chandra über die sphärischen Funktionen der reduktiven Gruppen und der Fortsetzungen, die wir den Autoren und ebenfalls Helgason, Rosenberg, Trombi und anderen verdanken.

Die Theorie der Gel'fand Paare, die die abstrakte Plancherel Formel einschließt, wird im ersten Kapitel dargestellt. Im zweiten Kapitel werden die Gruppen der Klasse $\mathscr{H}$ von Harish Chandra eingeführt. Selbst wenn das Ziel ist, die sphärischen Funktionen über den halbeinfachen Gruppen zu studieren, verlangt die „descent method" von Harish Chandra die Einführung der Klasse $\mathscr{H}$. Im dritten Kapitel erreicht man den Kernpunkt des Themas: die Beziehung zwischen den sphärischen Funktionen und den Darstellungen der Hauptserie. Im Kapitel 4 wird die Reihenentwicklung der sphärischen Funktionen nachgewiesen. Es ist dies die Entsprechung in mehreren Veränderlichen der Reihenentwicklung einer Lösung einer Differentialgleichung in der Nähe eines regulären singulären Punktes. Im Kapitel 5 wird das asymptotische Verhalten der sphärischen Funktionen untersucht, was ermöglicht, im Kapitel 6 nachzuweisen, daß die sphärische Fourier'sche Transformation, oder die Harish Chandra Transformation, ein Isomorphismus zwischen dem Schwartz'schen Raum $\mathscr{C}(G//K)$ und einem üblichen euklidischen Schwartz'schen Raum ist. Zwei Methoden werden angeboten: die urspüngliche Methode von Harish Chandra, und die von Gangolli-Helgason-Rosenberg. Kapitel 7 enthält eine Darlegung der L^p Theorie der sphärischen Fourier'schen Analyse von Trombi und Varadarajan.

Obwohl das Buch „selfcontained" ist, zumindest nach der Meinung der Autoren, werden beim Leser solide Kenntnisse der nichtkommutativen harmonischen Analyse vorausgesetzt. Andererseits ist das Buch reichhaltig und enthält den Grundstoff für den Forscher, der seine Kenntnisse der Theorie der sphärischen Funktionen auf halbeinfachen Gruppen vertiefen möchte. Die Anmerkungen am Ende jedes Kapitels bringen nützliche Informationen.

Straßburg J. Faraut

Nikiforov, A. F., Uvarov, V. B., Special Functions of Mathematical Physics, Basel – Boston: Birkhäuser Verlag 1987, xviii und 428 pp., DM 98,–

Das vorliegende Lehrbuch über Spezielle Funktionen der Mathematischen Physik ist eine Übertragung (und teilweise Überarbeitung) des russischen Originals von 1978. Es trägt den Untertitel „A Unified Introduction with Applications" und wendet sich laut Vorwort vor allem an Physiker.

Hauptgegenstand des (für eine Einführung relativ umfangreichen) Buches ist die Klasse der („einfachen") Speziellen Funktionen der Mathematischen Physik, welche als Lösungen der allgemeinen Riemannschen Differentialgleichung und ihrer konfluenten Formen auftreten. Insbesondere werden klassische orthogonale Polynome, darunter vor allem auch die entsprechenden orthogonalen Polynome zu diskreten Skalarprodukten, Zylinderfunktionen und hypergeometrische sowie konfluente hypergeometrische Funktionen betrachtet. In einem Anwendungskapitel werden verschiedene Probleme der Quantenmechanik mittels dieser Funktionen unter Zuhilfenahme der Theorie der Sturm-Liouvilleschen Randeigenwertprobleme behandelt.

Zentrales Anliegen der Verfasser ist die Entwicklung der Theorie der Speziellen Funktionen aus einem einheitlichen Prinzip heraus. Ihr Ausgangspunkt ist die Klasse der o.g. Differentialgleichungen, welche sie als Differentialgleichungen vom „hypergeometrischen Typ" bezeichnen und für deren Lösungen sie spezielle Integraldarstellungen gewinnen. Diese Integraldarstellungen, welche als verallgemeinerte Rodrigues-Formeln aufgefaßt werden können, dienen als die Grundlage der gesamten folgenden Theorie. Durch dieses Konzept ist die Klasse der zu behandelnden Funktionen auf die o.g. Fälle beschränkt. Die Behandlung „höherer“ Spezieller Funktionen wie z. B. Mathieufunktionen, Sphäroidfunktionen, Lamefunktionen, etc. erfordert allgemeinere Prinzipien und Hilfsmittel. Ebenso können in diesem Rahmen auch weitergehende Resultate wie allgemeine Entwicklungssätze z. B. nach hypergeometrischen Funktionen keine Berücksichtigung finden. Leider führt die Verwendung der hier betrachteten speziellen Integraldarstellungen nicht immer zum einfachsten Zugang zu diesen Funktionen. Gelegentlich sind andere Kerne zweckmäßiger. Zum anderen sind die betrachteten Integraldarstellungen auch hinsichtlich der Integrationswege nicht allgemein genug gewählt. So würde z. B. die Verwendung von geschlossenen Wegen in der komplexen Ebene oft unnötige Einschränkungen im Parameterbereich vermeiden.

Positiv hervorzuheben ist die Theorie der orthogonalen Polynome einer diskreten Variablen, die hier wohl zum ersten Mal in einheitlicher Darstellung präsentiert wird. Ca. ein Viertel des Buches ist diesem Stoff gewidmet. Der übrige Stoff (auch in Hinblick auf die Anwendungen) bewegt sich ganz im üblichen Rahmen.

Aus mathematischer Sicht ist an dem vorliegenden Buch in verschiedener Hinsicht Kritik anzumelden. Einmal mangelt es unseres Erachtens häufig an der nötigen Präzision sowohl in der Darstellung als auch in der Schlußweise. Zum anderen sind die Darstellung und Beweisführung an vielen Stellen unübersichtlich. Vieles wäre wesentlich besser lesbar und nachvollziehbar unter Zuhilfenahme einfacher Begriffe und Resultate aus der linearen Algebra, der Theorie der gewöhnlichen linearen Differentialgleichungen mit singulären Stellen und elementarer Hilbertraumtheorie.

Bezüglich des Literaturverzeichnisses ist das Fehlen sämtlicher einschlägiger deutsch- und französischsprachiger sowie auch einiger wichtiger englischsprachiger Monographien zu bemängeln.

Abschließend können wir das vorliegende Lehrbuch mathematisch interessierten Lesern nur bedingt empfehlen.

Essen D. Schmidt, G. Wolf

Rozanov, Y. A., Introduction to Random Processes (Springer Series in Soviet Mathematics) (Übersetzung: B. Rothinger), Berlin – Heidelberg – New York: Springer-Verlag 1987, 117 Seiten, DM 54,–

Das vorliegende Buch ist die Übersetzung des russischen Titels „Vvedenie v teoriiu sluchainykh protsessov“, und soll – so der Verfasser – eine kurze Einführung in die Theorie der stochastischen Prozesse geben. Das ist natürlich nicht einfach, und so beschränkt sich der Autor darauf, anhand einer Reihe von mathematisch einfacheren Prozessen, grundlegende mathematische Fakten darzustellen und die Anwendungsbreite dieser Modelle anzuprechen. Dementsprechend treten in der Darstellung mathematische Probleme in den Hintergrund und werden oft in Form von Problemen angesprochen. Vom Leser wird eine gehörige Portion wahrscheinlichkeitstheoretischer Intuition und Denkweise vorausgesetzt. Maßtheoretische Grundlagen sind zudem unabdingbar.

Homogene Markoffprozesse mit abzählbarem Zustandsraum werden zunächst mit ihren grundlegendsten Eigenschaften besprochen. Verzweigungsprozesse bilden ein einfaches, intuitiv sehr einleuchtendes Beispiel. Ebenso sind die Prozesse in Bedienersystemen als klassisches Beispiel zu nennen.

Natürlich werden in dem Buch auch die elementaren Eigenschaften der Brownschen Bewegung dargestellt. Es folgen anschließend drei Abschnitte über stochastische Integration und stochastische Differentialgleichungen. Der Autor lehnt sich hier stärker an andere Darstellungen an. Im Gegensatz zu den anderen Teilen des Buches fehlt es in dieser Passage an Motivation: es wird auch kaum deutlich, daß das besprochene stochastische Integral auf dem Ito Kalkül aufbaut.

Ähnliches kann über das elfte Kapitel, das Diffusionsprozesse behandelt, gesagt werden. Mit den sich anschließenden Kapiteln über lineare und stationäre Prozesse wird der motivierende Stil des ersten Teiles des Buches wieder aufgenommen. In den abschließenden beiden Kapiteln wird schließlich ein Schätzproblem bei stochastischen Differentialgleichungen und der Kalman-Bucy Filter besprochen.

Bei der Kürze des Buches und der Vielzahl von angesprochenen Themenkreisen ist sicherlich nicht zu erwarten, einen einigermaßen guten Einblick in die Theorie zu gewinnen. Das war wohl auch nicht die Absicht des Verfasser. Trotzdem meine ich, daß das Buch gut geeignet ist, sich einen Eindruck über einige Probleme zu verschaffen, mit denen die Theorie stochastischer Prozesse befaßt ist.

Göttingen M. Denker

Daley, D., Vere-Jones, D., An Introduction to the Theory of Point Processes (Springer Series in Statistics), Berlin u.a.: Springer-Verlag 1988, 1 fig., 702 + XXI S., Hard cover, DM 145,–

Die Theorie der Punktprozesse, die sich allgemein gesprochen mit zufällig in einem metrischen Raum verteilten Punkten befaßt, ist ein relativ junger Zweig der Stochastik. Isolierte Probleme (wie die Verteilung der Fixsterne am Firmament) hat man schon lange erörtert. Systematische Bemühungen um den Aufbau einer Theorie gibt es jedoch erst seit den 60er Jahren, nachdem man Poisson-Prozesse ausreichend beherrschte. (Bei einem Poisson-Prozeß werden auf zufällige Weise abzählbar viele Punkte derartig in einem Grundraum verteilt, daß die Anzahlen der Punkte in disjunkten Teilmengen stochastisch unabhängig sind.) Wichtige Impulse kamen aus der sich gleichzeitig entwickelnden Theorie von Maßen auf metrischen Räumen, von ähnlich fundamentaler Bedeutung wurde später die moderne Martingaltheorie. Inzwischen kann man sich über die wesentlichen theoretischen wie anwendungsorientierten Aspekte von Punktprozessen in einer Reihe von beachtlichen Monographien informieren.

Eine weitere Einführung legen nun D. J. Daley und D. Vere-Jones vor, beides Autoren, die zur Entwicklung der Theorie und ihrer Anwendungen beigetragen haben. Beim ersten Durchblättern fällt zunächst der enorme Umfang dieses Werkes auf: es umfaßt über 700 eng bedruckte Seiten, die sicherlich nützliche Bibliographie enthält 438 Einzeleintragungen. Die Erwartung, daß in diesem Buch das Gebiet in seiner gesamten Breite dargelegt wird, erfüllt sich jedoch nur zum Teil. Nicht an erster Stelle würde ich den Text denjenigen empfehlen wollen, die die vielfältigen Anwendungsbezüge der Theorie kennenlernen möchten. Über Zusammenhänge mit Warteschlangen, stochastischer Geometrie und Stereologie oder statistischen Fragestellungen (Prediktion und Filtrieren) informiert man sich günstiger an anderer Stelle. Es handelt sich um eine Einführung in die *Theorie*, wie dies auch die Autoren in der Einleitung betonen. Damit ist ein gewisser Verlust verbunden, denn die Entwicklung gerade der Theorie der Punktprozesse wurde wesentlich auch durch ihre Anwendungsbezüge mitbestimmt.

Die Intention der Autoren ist es, den Teil der Theorie, den man als Verteilungstheorie von Punktprozessen auf metrischen Räumen bezeichnen könnte, ausführlich darzustellen. Es geht um folgendes: Abzählbare Mengen $\{z_1, z_2, \ldots\}$ von (ununterscheidbaren) Punkten aus einer Grundmenge E beschreibt man günstig durch ihre Zählmaße μ (definiert durch $\mu(A) = \mathrm{card}\,\{i : z_i \in A\}$, $A \subset E$), deren Gesamtheit wir mit $N(E)$ bezeichnen. Für den

Aufbau der Theorie ist es von Vorteil anzunehmen, daß E ein vollständiger, separabler metrischer Raum ist. $N(E)$ läßt sich dann einbetten in die Menge $M(E)$ aller Borelmaße auf E, deren schwache Topologie derart metrisiert werden kann, daß $M(E)$ ebenfalls zu einem vollständigen, separablen metrischen Raum wird. Dies gibt das Grundgerüst für eine einigermaßen reichhaltige Theorie. Ein Punktprozeß ist in dieser Sichtweise nichts anderes als eine Zufallsvariable mit Werten in $N(E)$. Der überwiegende Teil des Buches von Daley und Vere-Jones, insbesondere Kapitel 5–12 und 14 ist der Behandlung von derartigen Zufallsvariablen gewidmet. Betrachtet werden zufällige Maße und Integrale, Momentenmaße, einfache und markierte Punktprozesse, Cluster-Prozesse, unendlich teilbare Prozesse, Cox-Prozesse, Konvergenzbegriffe, Überlagerung und Verdünnen von Punktprozessen, Stationarität, Spektraltheorie, Campbellmaße und Palmverteilungen und manches mehr. Das Bestreben der Autoren war es offenbar, dieses Material sorgfältig und in allen Einzelheiten darzulegen, Argumente vollständig zu entwickeln und damit den Einstieg in die Spezialliteratur zu ermöglichen. Standardreferenzen für dieses Gebiet waren bisher das konzentrierte Büchlein von O. Kallenberg: *Random Measures,* sowie die umfangreichere Monographie von J. Kerstan, K. Matthes, J. Mecke: *Unbegrenzt teilbare Punktprozesse.* Beide gelten als schwierige Lektüre. Als ergänzende Darstellung ist hier das von Daley und Vere-Jones vorgelegte Lehrbuch gerade für Nichtspezialisten sicherlich willkommen. (Für Spezialisten gibt es wohl nicht sehr viel neues zu entdecken.) In den ersten vier Kapiteln wird versucht, den Leser durch Darstellung historischer Entwicklungen sowie einfacher Modelle und Resultate auf den Themenkomplex vorzubereiten. In drei Anhängen werden technische Voraussetzungen erläutert, so daß die Darstellung für einen Leser mit Grundkenntnissen in der maßtheoretischen Stochastik zugänglich sein dürfte.

Knapp fällt der Teil der Darstellung aus, der sich mit Martingalmethoden befaßt, von den 14 Kapiteln des Buches ist ihm gerade ein Kapitel gewidmet. Wenn man bedenkt, daß dieser Aspekt in den letzten Jahren die wissenschaftliche Diskussion beherrscht hat, könnte man hier größere Ausführlichkeit für wünschenswert halten. Immerhin werden die Grundbegriffe erläutert, die Aufgabe des Filtrierens wird mit der Methode der „Referenzwahrscheinlichkeiten" („Girsanovscher Satz") gelöst (der Innovationsansatz samt Darstellung von Martingalen durch stochastische Integrale kommt allerdings nicht zur Sprache) und ein paar Grenzwertsätze, die auf Konvergenzbetrachtungen von Kompensatoren beruhen, werden abgeleitet.

Einige kritische Bemerkungen lassen sich jedoch nicht umgehen. Die Autoren unternehmen einige Anstrengungen, um Resultate zu erläutern und Begriffe zu motivieren (wobei sie sich nicht scheuen, auch Trivialitäten auszubreiten). Man kann jedoch nicht verschweigen, daß dies an mancher Stelle mißlingt. Man trifft auf Aussagen, die, nimmt man sie wörtlich, nicht stimmen. Im Einzelfall mag das harmlos sein, es irritiert allerdings, daß sich kleinere oder auch größere Ungenauigkeiten doch immer wieder finden lassen. (Symptomatisch ist, daß eine auf S. 174 angekündigte „more general question" das zuvor behandelte Problem nicht umfaßt.)

Frankfurt G. Kersting

Number Theory

Proceedings of the First Conference of the Canadian Number Theory Association held at the Banff Center, Banff, Alberta, April 17-27, 1988

Richard A. Mollin (Editor)

1990. XIV, 659 pages. 17 x 24 cm. Cloth DM 198,- ISBN 3 11 011723 1

Contents:

W DE G **de Gruyter · Berlin · New York**