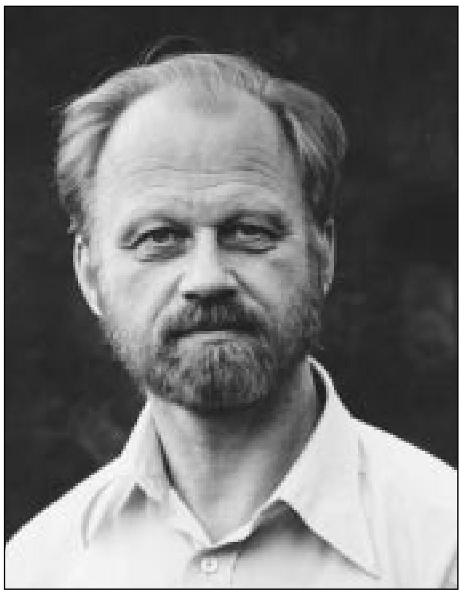
Jürgen K. Moser (1928–1999)

John N. Mather, Henry P. McKean, Louis Nirenberg, and Paul H. Rabinowitz



Jürgen K. Moser

Jürgen K. Moser died of prostate cancer on December 17, 1999, in Zürich, Switzerland. He was seventy-one years young. One of those rare people with a gift for seeing mathematics as a whole, he was very much aware of its connections to other branches of science. His research had a profound influence on mathematics as well as on astronomy and physics.

Moser was born on July 4, 1928, in Königsberg, in East Prussia. He spent 1947–53 at the university in Göttingen, receiving his doctorate in 1952 under the direction of Franz Rellich. It was Rellich who kindled his interest in physics. A more important influence on him, however, was Carl Ludwig Siegel, who returned to Göttingen in 1950 after ten years in Princeton. Moser learned number theory and celestial mechanics from him and wrote the notes for Siegel's course on the latter topic. These notes became the first draft of Siegel's 1956 book. Moser revised the book in 1971, and it was reissued under their joint authorship.

A Fulbright Fellowship allowed Moser to spend a stimulating year (1953–54) at New York University (NYU). He went back to Göttingen as Siegel's assistant (1954–55), and this interval was followed by a two-year assistant professorship at NYU and three years as associate professor at the Massachusetts Institute of Technology (MIT). He returned to NYU as a full professor in 1960 and remained there until 1980. Moser then joined the faculty of the Eidgenössische Technische Hochschule (ETH) in Zürich, Switzerland. He retired in 1995 at the mandatory retirement age of sixty-seven.

Moser made deep and important contributions to an extremely broad range of questions in dynamical systems and celestial mechanics, partial differential equations, nonlinear functional analysis, complex geometry, and the calculus of variations. More specifically, he did major work on KAM theory (named in honor of Kolmogorov, Arnold, and Moser), regularity questions and Harnack inequalities for elliptic and parabolic partial differential equations (PDEs), Nash-Moser theory, biholomorphic equivalence, and completely integrable Hamiltonian systems.

For his outstanding research achievements, Moser received many honors and awards. He was elected to membership in the National Academy of Science of the USA (NAS) in 1973, and he was a foreign member of several national academies. In addition, he received the Craig Watson Medal of the NAS for his contributions to dynamic astronomy (1967), the first AMS-SIAM George David Birkhoff Prize in Applied Mathematics (1968), the L. E. J. Brouwer Medal of the Dutch Scientific Society for his work in analysis and classical mechanics (1984), the Cantor Medal of the German Mathematical Society (1992), and the Wolf Prize (1995). He delivered the AMS Gibbs Lecture (Dallas, 1973), the Pauli Lectures (ETH, 1975), the AMS Colloquium Lectures (Toronto, 1976), the Hardy Lectures (Cambridge, 1977), the Fermi Lectures (Pisa, 1981), the von Neumann Lecture (SIAM, Seattle, 1984), and three invited addresses at International Congresses of Mathematicians.

To those who knew him Moser exemplified a creative scientist and, perhaps even more important, a human being. His standards were

---

Paul H. Rabinowitz is professor of mathematics at the University of Wisconsin, Madison. His e-mail address is rabinowi@math.wisc.edu.

All photographs are courtesy of Klaus Moser.

Göttingen, 1947.

high and his taste impeccable. His papers were elegantly written. Not merely focused on his own research, he worked successfully for the well-being of mathematics in many ways. He stimulated several generations of younger people by his penetrating insights into their problems, scientific and otherwise, and by his warm and wise counsel, concern, and encouragement. My experience as his student was typical: then and afterwards I was made to feel like a member of his family. Moser had an important influence in his role as director of the Courant Institute of Mathematical Sciences at NYU (1967–70), as director of the Mathematics Research Institute of the ETH (1984–95), as president of the International Mathematical Union (1983–86), and through his active membership on numerous advisory bodies. He was a lifelong music lover—he played the cello—and amateur astronomer. He enjoyed biking and swimming. At the age of sixty, he took up paragliding.

With his untimely death at seventy-one, many of us in the mathematics community lost a hero, friend, and mentor.

Some of Moser's mathematics will be discussed in the three segments below. John Mather treats KAM theory, Louis Nirenberg describes Moser's work in partial differential equations, and Henry McKean writes about completely integrable Hamiltonian systems. A three-volume selecta of his work, as chosen by Moser himself, is planned by the publisher Izhevsk. The first volume, *Integrable Systems and Spectral Theory*, 296 pages (translated into Russian), appeared in 1999. The two further translated volumes, as well as an edition of the selecta in English, are expected to appear soon.

—Paul H. Rabinowitz

John N. Mather

The theme of Moser's work in dynamical systems was "stable and random motions", a phrase that appears in the title of a very influential book [M9] of his. He explained "the larger picture of the stability problem for Hamiltonian systems" as follows:[1]

> Speaking quite intuitively, we are familiar with systems with very unstable, even ergodic, behavior, as …demanded by statistical mechanics. …On the other hand, there are other systems exhibiting clearly stable behavior, as, for example, the planetary motion of the solar system. The problem is to decide which systems have stable and which unstable behavior.
>
> At this stage of the knowledge one can establish stability of a Hamiltonian system only if a system is sufficiently close to so-called *integrable systems*—these are systems for which so many integrals are known that stability is evident. Moreover small perturbations of such systems do not affect the stability behavior too much. This is the meaning of KAM theory …. But as the system recedes further from the exceptional integrable system stability will deteriorate and ultimately get lost. This is the expected phenomenon which one would like to understand. For example, the planetary system has the special feature that the masses of the planets are small compared to that of the sun, so that the forces between the planets are much smaller than those between the planets [and] the sun. If one neglects the former, one has the Kepler approximation, in which all planets move on ellipses. This is the integrable approximation which is evidently stable, and this system will retain its stability for small masses of the planets. If we enlarge the masses of the planets sufficiently, we would expect the system to become unstable.
>
> More generally one studies small perturbations $H = H_0 + \varepsilon P$ of an integrable Hamiltonian H_0. For such an integrable system of n degrees of freedom, one has n independent integrals in involution whose level sets are n-dimensional tori. In this case stability is obvious. After small perturbation of such a system, i.e., for small ε, the system corresponding to the Hamiltonian $H = H_0 + \varepsilon P$ still possesses a large set of invariant tori. This is the content of the so-called KAM theory. In other words, for most initial points in phase space the orbits lie on such tori and exhibit stable behavior. For $n \geqslant 3$ the exceptional orbits not lying on such tori may leak out and escape slowly. For

John N. Mather is professor of mathematics at Princeton University. His e-mail address is `jnm@math.princeton.edu`.

[1] *Recent developments in the theory of Hamiltonian systems,* SIAM Review ***28*** *(1986), pp. 459–485.*

Doctoral Students of Jürgen Moser

Daniel Slotnick, NYU (1957)
Charles C. Conley, MIT (1961)
Seymour Sherman, NYU (1962)
Wolfe Snow, NYU (1964)
Robert Sacker, NYU (1964)
Marcelle Friedman, NYU (1966)
Paul Rabinowitz, NYU (1966)
Burton Lieberman, NYU (1967)
Martin Braun, NYU (1968)
Neil Fenichel, NYU (1970)
Howard Jacobowitz, NYU (1970)
Samuel Graff, NYU (1971)
Peter Kammeyer, NYU (1974)
Mark Adler, NYU (1976)
Matthew Bottkol, NYU (1977)
Mark Levi, NYU (1979)
Edward Belbruno, NYU (1979)
Robert Sachs, NYU (1980)
Jürgen Pöschel, ETH (1982)
Håkan Eliasson, U. Stockholm (1984)
Alessandra Celletti, ETH (1989)
Christophe Genecand, ETH (1990)
Martin Flucher, ETH (1991)
Jochen Denzler, ETH (1992)

> $n = 2$ such slow escape is not possible since the exceptional orbits on a three-dimensional energy surface are trapped between the two-dimensional tori. Nevertheless, the set of these tori generally forms a rather complicated Cantor set, even for simple and smooth Hamiltonians.

The stability problem in celestial mechanics goes back to Newton. Nevertheless, the modern formulation is due to Poincaré (1890), who first recognized the importance of random motions (or "chaos", in the currently popular terminology). This aspect of Poincaré's work was greatly extended by G. D. Birkhoff. However, the fundamental problem remained unsolved, as Moser explained in his historical comments on pp. 8–9 of [M9]. This is the problem of constructing quasi-periodic solutions (i.e., invariant tori) in the N-body problem. Weierstrass knew formal power series expansions for such solutions, and he suggested the problem of their convergence to Mittag-Leffler as a prize question to be sponsored by the Swedish king. Poincaré won the prize with his famous 1890 paper. The prize was awarded for the wealth of ideas; Poincaré did not solve the problem. From the remarks of Weierstrass and Poincaré quoted in [M1, pp. 8–9], it would appear that Weierstrass favored the view that the series converge and Poincaré favored the opposite view. Both agreed, however, that the question had not been settled rigorously.

Moser remarked [M9, p. 9] that the work of Kolmogorov [K1] showed that the series do converge at least if a certain Hessian determinant does not vanish. Here, H is analytic. In action-angle coordinates $I = (I_1, \dots, I_n)$ and $\theta = (\theta_1, \dots, \theta_n)$, the unperturbed H_0 is a function of I alone: $H_0 = H_0(I)$. Here I is in an open ball B^n in $\mathbb{R}^n$, and θ is in $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. The Kolmogorov theorem assumes the nonvanishing of the Hessian determinant: $\det(\partial^2 H_0/\partial I^2) \neq 0$.

The equations of motion have Hamiltonian form in action-angle coordinates, namely,

$$\dot{\theta} = \partial H/\partial I, \quad \dot{I} = -\partial H/\partial \theta. \tag{1}$$

For the unperturbed Hamiltonian H_0, the second equation reduces to $\dot{I} = 0$. So each torus $I^\circ \times \mathbb{T}^n$ is invariant, I° denoting a particular point I. On this torus, the equations of motion take the form $\dot{\theta} = \omega^\circ := \partial H/\partial I(I^\circ)$. Every trajectory has the form $t \mapsto \theta^\circ + t\omega^\circ$; such a flow on the torus is called a *Kronecker flow* with rotation vector ω°. Kolmogorov's assertion is that the tori whose rotation vectors satisfy a Diophantine condition, namely, there exist $C > 0$ and $\nu > 0$ such that

$$\Big| \sum m_i \omega_i^\circ \Big| \geqslant C \Big/ \Big(\sum |m_i| \Big)^\nu \quad \text{for all } m \in \mathbb{Z}^n \setminus 0, \tag{2}$$

persist. This means that for ε small, the Hamiltonian system $H_0 + \varepsilon P$ has an invariant torus with a Kronecker flow of rotation vector ω°, and this torus is close to the invariant torus of the unperturbed system with the same rotation vector.

If the Diophantine *exponent* ν is greater than n, then almost every vector $\omega^\circ \in \mathbb{R}^n$ satisfies the Diophantine condition (2) for some Diophantine *coefficient* $C > 0$. It follows that $\omega^\circ = \partial H/\partial I(I^\circ)$ satisfies a Diophantine condition with exponent ν for almost every $I^\circ \in B$, in view of Kolmogorov's nondegeneracy condition $\det(\partial^2 H_0/\partial I^2) \neq 0$.

Moser tells[2] how *Mathematical Reviews* asked him to review [K2]. He was disappointed that neither [K1] nor [K2] contained a proof of the result announced in [K1]. He wrote to Kolmogorov asking for an argument, but never received a reply, so he indicated in his review that there was only a sketch of a proof. Moser was very excited about Kolmogorov's result because of its relevance to the stability problem of elliptic fixed points of area-preserving mappings, a problem C. L. Siegel had urged him to pursue.

[2] *Recollections,* The Arnoldfest: Proceedings of a Conference in Honour of V. I. Arnold for His Sixtieth Birthday *(E. Bierstone, B. Khesin, A. Khovanskíi, J. Marsden, eds.), Fields Institute Communications, vol. 24, Amer. Math. Soc., 1999, pp. 19–21.*

Since Kolmogorov's proof was not available to Moser, he continued to try to prove the invariant curve theorem for area-preserving twist maps. In 1961 he succeeded, but in the smooth category rather than the analytic category. Moser saw it as a shortcoming that he was unable to give a proof in the analytic category. The Russians, on the other hand, considered that Moser had extended Kolmogorov's theorem from the analytic category to the smooth category and that this was its main virtue. This has come to be the accepted view.

Moser's proof [M4] of his invariant curve theorem combined ideas from Kolmogorov's announcement [K1] with a fast iteration method Moser developed to prove a general implicit function theorem [M3] similar to that formulated by J. Schwartz in a 1960 paper generalizing Nash's embedding theorem [N1]. Moser's fast iteration method was based on a generalization of Newton's method, familiar to students of freshman calculus as a numerical method of finding roots of equations. Moser got the idea of using a method like Newton's from Kolmogorov's papers [K1] and [K2]. He combined this with a smoothing technique due to Nash [N1], which permitted him to carry out the proof in the smooth category.

Moser discussed a number of applications of fast iteration methods in his Pisa lectures [M7]. In particular, he proved the Kolmogorov theorem with estimates. Then he proved his differentiable version of Kolmogorov's theorem by using his estimates for Kolmogorov's theorem together with approximation techniques of Bernstein and Jackson.

There is a loss of derivatives in the conclusion of Moser's theorem: if H is C^r, then Moser's theorem asserts the existence of invariant tori of class C^{r-d}, where $d > 0$ depends on the number n of degrees of freedom and the Diophantine exponent ν, but not the differentiability class r. The method of [M7] gives a better upper bound for d than the method of [M3] and [M4].

Let K_δ be the subset of B^n consisting of vectors whose distance to the boundary is $\geq \delta$. Moser's proof shows that there exists $r_0 > 0$ and $d > 0$ such that for any $\delta > 0, C > 0, \nu \geq n, r \geq r_0$, there exists $\varepsilon > 0$ such that if the conditions

- P has C^r norm $< \varepsilon$,
- $I^\circ \in K_\delta$, and
- $\omega^\circ := \partial H/\partial I(I^\circ)$ satisfies a Diophantine condition with coefficient C and exponent ν

are satisfied, then the invariant torus $I^\circ \times \mathbb{T}^n$ for H_0 "persists" for the perturbed system $H = H_0 + P$, in the sense that this system has an invariant torus with rotation vector ω°. Moreover, this torus is of differentiability class C^{r-d}, it is C^{r-d} close to the original torus, and the coordinate transformation that exhibits the flow on it as a Kronecker flow is of class C^{r-d} and is close to the identity transformation in the C^{r-d} topology.

Kolmogorov (left) and Moser in Stockholm, 1962.

Moser's student Pöschel showed in 1982 that such tori nearly fill phase space, in the following sense: if $\eta > 0$ is given, they fill out a set of measure $\geqslant \text{meas}(K_\delta) - \eta$ provided ε is small enough. (Here, the measure is ordinary Lebesque measure on $\mathbb{R}^n$ or $\mathbb{R}^{2n}$.) The earlier analytic version of this theorem is due to Arnold [A2].

Both Kolmogorov's and Moser's proofs rely on a rapid iteration method. Here is a very approximate description of the proofs: Both authors introduced a nonlinear functional equation whose solution would imply the persistence of the invariant tori. They linearized this equation and solved the linearized equation. Kolmogorov iterated and stated that the resulting sequence converged. In [M4] Moser, using the method of [M3] that he created to simplify Nash's proof [N1], smoothed and iterated and proved convergence.

This very general description also describes the Picard method of proving the ordinary implicit function theorem (except for smoothing in the case of Moser's proof). However, the Picard convergence proof does not apply in either the Kolmogorov or the Moser setting, and entirely new ideas were needed. The difficulty concerns the linearized equation, which, of course, may be expressed in the form "solve $Ax = y$ for x". In the context of Kolmogorov's and Moser's proofs, A is a linear operator on an infinite-dimensional function space. It is invertible, but A^{-1} is unbounded with respect to any reasonable norm. The unboundedness of A^{-1} makes it impossible to apply the Picard convergence proof.

The linear operator that appears in Kolmogorov's and Moser's proofs is closely related to the differential operator $\omega^\circ \cdot \nabla := \sum \omega_i^\circ(\partial/\partial\theta_i)$ acting on functions of vanishing mean value on the n-torus $\mathbb{T}^n$. In the case of Kolmogorov's proof, one would consider real analytic functions; in the case of Moser's, highly differentiable functions. The properties of the linear operators that appear in

Moser at New York University in 1963.

Kolmogorov's and Moser's proofs are similar to the properties of this operator, and we will use it for illustration.

Consider the Fourier expansion

$$x(\theta) = \sum \hat{x}(m) e^{2\pi i m\cdot\theta}. \tag{3}$$

The equation $\omega^\circ \cdot \nabla x = y$ is equivalent to the infinite family of equations

$$2\pi i \omega^\circ \cdot m \hat{x}(m) = \hat{y}(m), \quad m \neq 0 \in \mathbb{Z}^n, \tag{4}$$

which is formally solvable as long as $\omega^\circ \cdot m := \sum \omega_i^\circ m_i$ never vanishes for $m \neq 0$, i.e., as long as $\omega_1^\circ, \dots, \omega_n^\circ$ are linearly independent over $\mathbb{Q}$.

For the formal solution to represent a function, some further condition is needed. For example, if ω° satisfies the Diophantine condition (2), then x is analytic (resp. C^∞) if y is. Thus, $\omega^\circ \cdot \nabla$ is invertible on the space $C_0^\omega(\mathbb{T}^n)$ (resp. $C_0^\infty(\mathbb{T}^n)$) of real analytic (resp. infinitely differentiable) real-valued functions of vanishing mean value on $\mathbb{T}^n$. In fact, in the infinitely differentiable case, the Diophantine condition is the necessary and sufficient condition for invertibility of $\omega^\circ \cdot \nabla$. In the real analytic case, a weaker condition suffices.

On the Banach space $C^r(\mathbb{T}^n)$, the operator $(\omega^\circ \cdot \nabla)^{-1}$ is unbounded. This unboundedness is due to "small divisors". In fact, if $\varepsilon > 0$, there are infinitely many $m \neq 0 \in \mathbb{Z}^n$ such that $|\omega^\circ \cdot m| \leqslant |m|^{-n+\varepsilon}$. From this, it follows that $(\omega^\circ \cdot \nabla)^{-1} : C^r \to C^{r-n+\varepsilon}$ is unbounded. On the other hand, $(\omega^\circ \cdot \nabla)^{-1} : C^r \to C^{r-d}$ is bounded if $d > \nu + n$, in the case that ω° satisfies a Diophantine condition of exponent ν. This is the phenomenon of "finite loss of derivatives". These facts are elementary exercises in Fourier series. The loss of derivatives makes Moser's proof difficult; the fact that d is independent of r makes it possible.

Both Kolmogorov's and Moser's proofs are based on fast iteration methods. These contrast with Picard's method, where the $(k+1)^{\text{st}}$ iterate u_{k+1} is related to the earlier iterates by the inequality $\|u_{k+1} - u_k\| \leqslant \lambda \|u_k - u_{k-1}\|$, where $0 < \lambda < 1$. According to Sinai, Kolmogorov's convergence proof depends on an estimate like the following to relate certain constructed norms:

$$\|u_{k+1} - u_k\|_{k+1} \leqslant C\Lambda^k \|u_k - u_{k-1}\|_k^\nu, \tag{5}$$

where C and Λ are (large) positive constants and $\nu > 1$. If $\|u_1 - u_0\|_1$ is small enough, (5) implies that $\|u_k - u_{k-1}\|_k$ converges very rapidly to zero.

The inequality (5) is obtained only at the cost of introducing different norms on the two sides of (5). Roughly speaking,

$$\|u\|_k = \sup \Big\{ |u(I,\theta)| : |I - I^\circ| \leqslant \delta_k, \\ |\operatorname{Im}\theta| \leqslant \rho(1/2 + 1/2^k) \Big\}.$$

Here I° is the action of the unperturbed torus that we wish to prove persists. The supremum is taken over the set of *complex* n-vectors $I = (I_1, \dots, I_n)$ and $\theta = (\theta_1, \dots, \theta_n)$ satisfying the given inequalities. This means that we extend the real analytic function u of $2n$ real variables $(I_1, \dots, I_n,\ \theta_1, \dots, \theta_n)$ to a holomorphic function of $2n$ complex variables. In the case that no such extension to the given domain is possible, $\|u\|_k$ is defined to be $+\infty$. The positive numbers δ_k must be chosen to tend very rapidly, but not too rapidly, to zero. A sequence satisfying a recursion relation $\delta_{k+1} = \delta_k^\alpha$ with $1 < \alpha < \nu < 2$ will do, provided that δ_1 is small enough. Finally, ρ is a positive number, related to H_0 and P. Roughly speaking, H_0 and P must extend holomorphically to the domain $|\operatorname{Im}\theta| \leqslant \rho$, and $|P(I,\theta)|$ must be small in that domain.

As Kolmogorov noted in [K1], his iteration method is a form of Newton's method. This is what permits one to prove (5) with any exponent $\nu < 2$. For details of this method in proofs of related theorems, see Arnold [A1], Moser [M7, §6], and [M9, Chap. V]. Arnold [A2] gave a proof of Kolmogorov's theorem by a different method.

Thus the proof of Kolmogorov shows that an invariant torus whose rotation vector ω^0 satisfies (2) persists for the perturbed Hamiltonian $H = H_0 + P$. How small $|P(I,\theta)|$ must be in the domain $|\operatorname{Im}\theta| < \rho$ is related to the Diophantine exponent and coefficient.

The proof of the estimate that we have represented by (5) is very long, although it is a series of straightforward applications of Cauchy's integral formula and elementary manipulations.

In the differentiable case, the difficulties caused by the unboundedness of $(\omega^0 \cdot \nabla)^{-1}$ are of a very different nature. When one solves the linearized problem, there is a loss of derivatives. Imitating the method of Kolmogorov would mean a loss of derivatives. After a finite number of iterations, all derivatives would be lost, and it would be impossible to continue the iteration. Moser's solution (following Nash) was to solve the linearized problem and to smooth, e.g., by convolution with a bump function $\rho_{\lambda_k}(x) = \rho(\lambda_k x)$. One has to choose λ_k going to infinity very fast, but not too fast. For example, a sequence satisfying a recursion relation $\lambda_{k+1} = \lambda_k^\alpha$ with $1 < \alpha < 2$ will do, if λ_1 is large enough. However, the estimates required to prove convergence are difficult. The knowledge that Nash had carried out similar estimates for the isometric embedding problem clearly was a great help, but Moser's insight that these methods could also apply to the very different stability problem in Hamiltonian mechanics was most remarkable.

In his Pisa lectures [M7] Moser showed that fast iteration methods constitute a very versatile tool, with many applications to analysis. The version of the fast iteration method used in [M3] and [M4] is now known as the Nash-Moser method.

The first result Moser proved [M4] as a result of his theory solved a fundamental problem proposed by G. D. Birkhoff [Bi, pp. 662–663] after deep study of Poincaré's ideas. Let f be an area- and orientation-preserving diffeomorphism of a region of the plane having the origin as a fixed point. Birkhoff describes the origin as *stable* if it has arbitrarily small invariant neighborhoods and *unstable* otherwise.

Moser [M4] obtained sufficient conditions for stability when f is highly differentiable. Let λ and ω be the eigenvalues of the derivative df_0 of f at the origin. Since f is area- and orientation-preserving, $\lambda\omega = 1$. The only case that is interesting for the stability question is when λ is nonreal. In this case, $\lambda = \overline{\omega}$ and so $|\lambda| = 1$. Birkhoff showed [Bi, pp. 111–229] that if λ is imaginary and is not a third or fourth root of unity, then there exists an analytic coordinate system x, y, centered at the origin, such that $dx \wedge dy$ is the standard area form on $\mathbb{R}^2$ and

$$f(z) = N(z) + R(z)$$
$$N(z) = ze^{2\pi i(\beta_0+\beta_1 r^2)}, \quad R(z) = O(r^5), \tag{6}$$

where $z = x + iy$ and $r = \sqrt{x^2 + y^2}$. This is the special case of the *Birkhoff normal form* that is most relevant to Moser's theorem. Here β_0 and β_1 are real numbers, and $e^{2\pi i\beta_0} = \lambda$. The number β_1 is called the *first Birkhoff invariant*.

Moser [M4] showed that if $\beta_1 \neq 0$ (in addition to the conditions stated above), then the origin is stable. In particular, stability is a consequence of a condition on the 3-jet of f at the origin, and this condition is satisfied for an open dense set of 3-jets. This is contrary to what Birkhoff expected, even in the analytic case. According to Arnold,[3] Kolmogorov expected such a result in the analytic case, but not in the differentiable case. Indeed, in 1954 he even wrote that the differentiable case of his theorem would "obviously" be false [K2]. It would seem that Moser's differentiable version of the Kolmogorov theorem was totally unexpected.

We have discussed KAM theory above in the context of differential equations, but it applies also to diffeomorphisms. For an f as above, the area- and orientation-preserving condition

[3] *Jürgen Moser (1928–1999): Déclin des Mathématiques (après la mort de Jürgen Moser),* Gaz. Math., *n^{o} 84 (Avril 2000), 92–95.*

Selected Publications of Jürgen Moser

References

[A-M] H. Airault, H. P. McKean, and J. Moser, Rational and elliptic solutions of the Korteweg-de Vries equation and a related many-body problem, *Comm. Pure Appl. Math.* **30** (1977), 95–148.

[C-M] S. S. Chern and J. K. Moser, Real hypersurfaces in complex manifolds, *Acta Math.* **133** (1974), 219–271; Erratum, ibid. **150** (1983), 297.

[D-M] B. Dacorogna and J. Moser, On a partial differential equation involving the Jacobian determinant, *Annales Inst. H. Poincaré Anal. Non Linéaire* **7** (1990), 1–26.

[M1] J. K. Moser, A new proof of de Giorgi's theorem concerning the regularity problem for elliptic differential equations, *Comm. Pure Appl. Math.* **13** (1960), 457–468.

[M2] ____, On Harnack's theorem for elliptic differential equations, *Comm. Pure Appl. Math.* **14** (1961), 577–591.

[M3] ____, A new technique for the construction of solutions of nonlinear differential equations, *Proc. Nat. Acad. Sci. USA* **47** (1961), 1824–1831.

[M4] ____, On invariant curves of area-preserving mappings of an annulus, *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* **1962** (1962), 1–20.

[M5] ____, A Harnack inequality for parabolic differential equations, *Comm. Pure Appl. Math.* **17** (1964), 101–134.

[M6] ____, On the volume elements on a manifold, *Trans. Amer. Math. Soc.* **120** (1965), 286–294.

[M7] ____, A rapidly convergent iteration method and non-linear partial differential equations I, *Annali Scuola Norm. Sup. Pisa* (3) **20** (1966), 265–315; II, 499–535.

[M8] ____, A sharp form of an inequality by N. Trudinger, *Indiana Univ. Math. J.* **20** (1971), 1077–1092.

[M9] ____, *Stable and Random Motions in Dynamical Systems: With Special Emphasis on Celestial Mechanics*, Ann. of Math. Studies, vol. 77, Princeton Univ. Press, Princeton, 1973.

[M10] ____, On a nonlinear problem in differential geometry, *Dynamical Systems (Proc. Sympos. Univ. Bahia, Salvador, 1971)*, Academic Press, New York, 1973, pp. 273–280.

[M11] ____, Finitely many mass points on the line under the influence of an exponential potential—An integrable system, *Dynamical Systems, Theory and Applications (Rencontres, Battelle Res. Inst., Seattle, WA, 1974)*, Lecture Notes in Phys., vol. 38, Springer, Berlin, 1975, pp. 467–497.

[M12] ____, Three integrable Hamiltonian systems connected with isospectral deformations, *Advances in Math.* **16** (1975), 197–220.

[M13] ____, The scattering problem for some particle systems on the line, *Geometry and Topology*, Lecture Notes in Math., vol. 597, Springer, Berlin, 1977, pp. 441–463.

[M14] ____, Geometry of quadrics and spectral theory, *The Chern Symposium 1979 (Proc. Internat. Sympos., Berkeley, CA, 1979)*, Springer, New York and Berlin, 1980, pp. 147–188.

[M15] ____, Minimal foliations on a torus, *Topics in Calculus of Variations*, Springer Lecture Notes in Math., vol. 1365, 1988, pp. 62–99.

corresponds to Hamilton's equations. The nonvanishing of the first Birkhoff invariant corresponds to the nondegeneracy condition $\det(\partial^2 H_0/\partial I^2) \neq 0$. The normal form N in the decomposition (6) plays the role of the integrable system. Since β_0 and β_1 are real, all small circles $\{r = \epsilon\}$ are invariant under N. Since $\beta_1 \neq 0$, the rotation number varies with ϵ. The remainder term R plays the role of the perturbation term. In this case, the persistence result means that f has many invariant curves that surround the origin and so the origin is stable, as Moser showed in [M4].

Even in the case f is analytic, it does not appear to be possible to prove this result by the Kolmogorov method; the problem is that the remainder term is too large. In this sense the Nash-Moser method gives a better result than the Kolmogorov method, even in the analytic case: it shows that tori persist for C^r small perturbations, not just for C^ω small perturbations (although it shows the existence of C^r tori, not real analytic ones). This extra robustness of the Nash-Moser method seems to play an essential role in applications Moser gave of his theorem in three papers in 1966 and 1968, one of them joint with W. H. Jeffreys, to the 3-body problem and to containment in a magnetic bottle, even though these systems are analytic. Arnold [A2] improved Kolmogorov's technique in another way, and this led to a very difficult theorem of Arnold's concerning the 3-body problem.

Acknowledgments: I wish to thank M. R. Herman for many enlightening discussions over many years, V. I. Arnold for a very stimulating discussion, and Ya. Sinai for telling me that Kolmogorov explained his proof to Moscow mathematicians and for suggesting improvements in a draft of this article.

Louis Nirenberg

Jürgen Moser was one of the most profound analysts of the last half century. His work ranged over different fields of mathematics, both pure and applied. Much of his deepest work was concerned with dynamical systems, especially small divisor problems and relations with celestial mechanics. He also did fundamental work in functional analysis—the Nash-Moser theory—and partial differential equations, and he made deep contributions in completely integrable systems, geometry, and complex analysis. Moser was also a master of exposition. His papers and published lectures are beautifully written.

This segment of the article is devoted primarily to some of Moser's contributions in partial differential equations. In the 1960s Moser published a series of wonderful papers in the field.

Louis Nirenberg is professor emeritus of mathematics at the Courant Institute of Mathematical Sciences at New York University. His e-mail address is `nirenl@cims.nyu.edu`.

Jürgen Moser and wife Gertrude, 1975.

Nash-Moser Theory

I start with the fundamental papers [M3] and [M7], which developed what is now called the Nash-Moser theory, or technique, in nonlinear functional analysis.

In a most remarkable paper [N1], John Nash proved that any Riemannian manifold may be embedded isometrically in a Euclidean space of sufficiently high dimension. To do so he devised a deep extension of the classical implicit function theorem. In its simplest form it requires solving an equation of the form

$$F(u) = f, \tag{1}$$

for a function u (of some variables). Here f is a function and F is a nonlinear operator (possibly a partial differential operator). The function f is supposed to be "close" to a fixed function f_0, for which we have a solution u_0, i.e., $F(u_0) = f_0$. One assumes the Frechet derivative $F'(u)$ is invertible for u and f "close" to u_0 and f_0. However, and here lies the difficulty, in applying the inverse F_u^{-1}, there is some loss of derivatives. Thus the usual Picard iteration scheme:

$$u_{i+1} = u_i + F'(u_0)^{-1}(f - F(u_i)),$$
$$\text{with } u_i = u_0 \text{ for } i = 0,$$

fails: one runs out of derivatives. Nash's scheme involved using smoothing of the functions. J. Schwartz in 1960 presented an abstract version of Nash's procedure. In [M3], in just a few pages, Moser presented a different, very applicable, way of attacking such problems, also involving smoothing operators, but using Newton's iteration scheme:

$$u_{i+1} = u_i + F'(u_i)^{-1}(f - F(u_i)). \tag{2}$$

With this scheme, even though the error $e_i = F(u_i) - f$ satisfies, formally,

$$(3) \qquad e_{i+1} = O(e_i^2),$$

i.e., there is accelerated convergence, one still runs out of derivatives.

Moser modified Newton's method by using a family of smoothing operators T_N, $N = 1, 2, \ldots$. Let $|u|_s$ be the C^s norm of u for $s = 0, 1, \ldots$. The operators T_N send the spaces C^s into C^∞ and are such that T_N tends to the identity operator as $N \to \infty$. Also, for some $C > 0$ and $\delta > 0$,

$$(4) \quad |T_N v|_{r+s} \leq CN^{s+\delta}|v|_r \qquad \text{for all } r, s \geq 0,$$

$$(5) \quad |(I - T_N)v|_r \leq CN^{-s+\delta}|v|_{r+s} \quad \text{for all } r, s > 0.$$

Moser's scheme is: For a suitable sequence of increasing integers N_i, use the iteration:

$$u_{i+1} = u_i + T_{N_i} F'(u_i)^{-1}(f - F(u_i)).$$

There is now no loss of derivatives, and estimates (3), (4), and (5) enable one to obtain convergence of this u_i in some C^s topology.

Moser used a variant of the technique of [M3] in his famous paper [M4] on invariant curves of area-preserving maps of an annulus—the beginning of his contribution to KAM theory—which led to Moser's solution of Birkhoff's problem, as described by J. Mather earlier in this article.

In his lectures [M7] Moser presented useful modifications of the Nash-Moser techniques of [M3] and [M4] and applied them to a number of problems, including (i) invariant manifolds of vector fields and (ii) a simplified form of results of Kolmogorov and Arnold concerning vector fields on a torus. In terms of (1), (ii) involves a situation where $F'(0)$ is invertible while for $u \neq 0$, $F'(u)$ may not be. But , as in [M4], the problem treated is a conjugacy problem, and this additional information is used to help make the method work. These self-contained papers are extremely beautiful—full of fascinating mathematics. I recommend that all graduate students read them.

Since Moser's work, many others have used and modified the Nash-Moser technique for a variety of problems. See, for example, the work of R. S. Hamilton (1982), L. Hörmander (1985), and Alinhac-Gérard (1991).

Another word about the Nash isometric embedding: Mathias Günther found an ingenious proof of it in 1989—it even lowers the dimension of the space in which the embedding occurs—which uses standard elliptic theory and completely avoids the Nash-Moser technique. In recent years a number of results first proved via Nash-Moser have been re-proved without it. Nevertheless, I regard the technique as one of the most important developments in nonlinear analysis in the last half century.

Moser's Derivation and Extension of the DeGiorgi-Nash Estimates for Elliptic and Parabolic Equations

Hilbert's Nineteenth Problem asked about the regularity of stationary points of the functional

$$(6) \qquad \int_\Omega f(x, u, \nabla u)\, dx \cdots dx_n.$$

Here Ω is a domain (open connected set) in $\mathbb{R}^n$ in which we consider functions u. The function f is assumed to be smooth, even analytic, in all its variables, and the corresponding Euler equation is assumed to be uniformly elliptic, i.e., for some $c_0 > 0$,

$$c_0|\xi|^2 \leq \sum_{i,j=1}^{n} f_{u_{x_i} u_{x_j}}(x, u, \nabla u)\xi_i \xi_j \leq \frac{1}{c_0}|\xi|^2$$

for all values of the arguments of f and all $\xi \in \mathbb{R}^n$. Hilbert posed the

Question. *If f is smooth (or real analytic), is every solution u smooth (or real analytic)?*

In dealing with the variational problem, one normally already knows that u and ∇u are in L^2, i.e., u is in the Sobolev space $W^{1,2}$. For simplicity I will consider f in (6) depending only on ∇u, say $f = f(u_{x_1}, \ldots, u_{x_n})$.

In 1938 C. B. Morrey proved the desired regularity in case $n = 2$. But for $n > 2$ the problem remained open for about twenty years. By well-known elliptic theory at the time, it was known that to prove the desired regularity it sufficed to show that u is in $C^{1,\alpha}$ locally, i.e., u is in C^1 and ∇u is Hölder continuous of some order α. This was proved independently by E. De Giorgi [DeG] and J. Nash [N2]. An extremal u of (6), if it were smooth, would satisfy the Euler equation

$$\sum_i \partial_i f_{u_i} = 0;$$

here $\partial_i = \frac{\partial}{\partial x_i}$.

Any first derivative of u, $v = u_{x_k} = u_k$, would satisfy

$$(7) \qquad \sum_{i,j} \partial_i (f_{u_i u_j} \partial_j v) = 0.$$

For an extremal u in $W^{1,2}$, it is easily seen that $v = \partial_k u$ is also in $W^{1,2}_{\text{loc}}$. Furthermore, v is a weak solution of (7), i.e., for every function $\phi \in C_0^\infty(\Omega)$ (in $C^\infty(\Omega)$ with compact support)

$$(8) \qquad \sum_{i,j} \int f_{u_i u_j} v_j \phi_i = 0.$$

Equation (8) then automatically holds for every ϕ in $W^{1,2}$ having compact support. One wants to prove that v is in C^α_{loc} for some $0 < \alpha < 1$, i.e., v is Hölder continuous locally. That is what De Giorgi and Nash proved. They considered a general linear elliptic equation in divergence form,

$$Lv = \sum_{i,j} \partial_i (a_{ij}(x) v_j) = 0, \tag{9}$$

with coefficients a_{ij} merely bounded and measurable, such that, for some $c_0 > 0$,

$$c_0 |\xi|^2 \le \sum_{i,j} a_{ij} \xi_i \xi_j \le \frac{1}{c_0} |\xi|^2. \tag{10}$$

Theorem 1. (De Giorgi, Nash). *A weak solution v in $W^{1,2}_{\text{loc}}$ of* (9) *belongs to some C^{α}_{loc}.*

Nash proved the analogous result for linear parabolic equations. Their proofs are quite different and quite elaborate. In a ball $B_r = \{|x| < r\}$ they established

$$|v(0)|^2 \le \frac{C}{r^n} \int_{B_r} v^2 \tag{11}$$

and, for some α in $(0,1)$,

$$|v(x) - v(y)| \le \frac{C}{r^{\alpha + n/2}} \left[\int_{B_r} v^2 \right]^{1/2} |x - y|^{\alpha} \tag{12}$$

for $x, y \in B_{r/2}$. Here C and α depend only on n and c_0.

A general principle is operating here: progress in partial differential equations is intimately connected with discovery of new a priori estimates for solutions, and new inequalities.

In [M1] Moser discovered new, very elegant, proofs of their theorem using Sobolev's inequality in an iterative way. Many people have since used his ideas. His ingenious proof of (11) begins with the weak form of (9) in B_r: for every ϕ in $W^{1,2}$ with compact support in B_r,

$$\sum_{i,j} \int \phi_i a_{ij} v_j = 0. \tag{13}$$

Moser's proof of (11) proceeds by making various choices for ϕ, depending on v. Roughly, he takes ϕ's of the form a power of a fixed cut off function times a power of v. Recursively using the Sobolev inequality leads to estimates of the form

$$\left[\int_{B_{r/2}} |v|^{2p^{j+1}} \right]^{\frac{1}{p^{j+1}}} \le C_j \int_{B_r} v^2, \tag{14}$$

where $p = \frac{n}{n-1}$ and C_j is bounded independently of j. As $j \to \infty$, the left-hand side of (14) tends to $(\text{meas}\, B_{r/2}) \times \max_{B_{1/2}} |v|^2$, from which (11) follows.

Moser then gives an elegant proof of (12) using a kind of Harnack inequality. But Moser's next paper in the subject, [M2], derives a full analogue of the classical inequality of Harnack for positive harmonic functions. Namely, he proves

Theorem 2. *Let v be a positive weak solution of* (9) *in a domain D, with $v \in W^{1,2}$. For any compact subset D' of D*

$$\max_{D'} v \le c \min_{D'} v,$$

where c depends only on D, D', and c_0.

The Hölder continuity of general solutions, namely (12), then follows easily from this theorem.

In his proof of Theorem 2, Moser works again with powers of the solution. He also makes use of a result of Fritz John and me on functions of bounded mean oscillation (BMO). One day he was on a train going home from New York with Fritz, who told him about our work. That night Moser realized that he could use it to complete his proof of the Harnack inequality. In 1970 E. Bombieri [Bo] found a proof of it that did not need BMO.

In [M2] Moser also gives some beautiful applications of Theorem 2. One states that any bounded solution of (9) in $|x| > R$ has a limit as $|x| \to \infty$. Using this, he proves an extension of "Bernstein's theorem" to higher dimensions—but with an additional hypothesis:

Theorem 3. *Let u be a solution of the minimal surface equation in $\mathbb{R}^n$,*

$$\sum_i \partial_i \left(\frac{u_{x_i}}{\sqrt{1 + |\nabla u|^2}} \right) = 0,$$

and assume that

$$|\nabla u| \le C. \tag{15}$$

Then u is an affine function.

Bernstein's original theorem was for $n = 2$ and did not assume (15). In fact, the result, without requiring (15), has been proved to hold for $n \le 7$. In a 1969 paper Bombieri, De Giorgi, and E. Giusti gave examples showing that the strong statement, i.e., without assuming (15), is false for $n \ge 8$. In a 1988 paper by L. Caffarelli, J. Spruck, and me, it was proved that Theorem 3 holds even if (15) is replaced by the weaker condition

$$|\nabla u(x)| = o(|x|^{1/2}) \quad \text{near infinity.}$$

The result also holds for a class of fully nonlinear equations.

Returning to the more general equation (6), one has to extend the results of Theorem 1 to more general linear equations of the form

$$\sum_{i,j} \partial_i (a_{ij}(x) \partial_j v) + \sum_i b_i(v) \partial_i v + cv = 0.$$

This was done by G. Stampacchia (1958) using De Giorgi's kind of arguments. He also established similar estimates up to the boundary for boundary value problems.

Paper [M5] is the first of two extending the results of Theorem 1 to parabolic equations of the form

$$(16) \qquad v_t = Lv = \sum_{i,j} \partial_i(a_{ij}(t,x)v_j)$$

with a_{ij} bounded measurable and satisfying (10). Besides giving another proof of Nash's result for these equations, Moser in [M5] extends the Harnack inequality (Theorem 2) to positive solutions of (16). The extension takes the following form.

***Theorem* 2′.** *Let v be a positive weak solution of* (16) *in a cylinder $D = \Omega \times (0,T)$, Ω being a domain in $\mathbb{R}^n$, and suppose*

$$\sup_{t\in(0,T)} \int_\Omega v^2\,dx + \int_0^T \int_\Omega |\nabla_x v|^2 < \infty.$$

Let K be a subdomain of Ω, with $\overline{K} \subset \Omega$, and consider two subintervals of $(0,T)$:

$$I^- : t_1 < t < t_2 \text{ and } I^+ : t_3 < t < t_4, \quad \text{with} \quad 0 < t_1 < t_2 < t_3 < t_4 < T.$$

Set

$$D^- = I^- \times K \text{ and } D^+ = I^+ \times K.$$

Then

$$\sup_{D^-} v \le \exp[C(c_0 + c_0^{-1})] \inf_{D^+} v,$$

where $C > 0$ depends only on D, D^+, and D^-.

In his proof of Theorem 2′ in [M5] Moser extended the result for BMO by John and me to the parabolic situation—no simple matter. However, in a subsequent paper Moser carried over Bombieri's idea of [Bo] to give a simpler proof of Theorem 2′ without need of BMO. Both are lovely papers.

Excellent sources for all the material presented here and further developments are the books [G-T] and [Kr]. Nash's methods in proving Theorem 2 were little exploited until a very interesting 1986 paper by E. Fabes and D. Stroock, which took up and simplified his arguments.

More PDE, Geometry, and Complex Analysis

There is a problem in global Riemannian geometry that is still not completely settled, though a number of people have worked on it.

Problem. *On the sphere S^2 with standard metric ds_0^2, which functions K may be the Gauss curvature of a metric ds^2 conformal to ds_0^2?*

(The same has been asked on S^n or other manifolds.) By the Gauss-Bonnet Theorem,

$$\int K\,dA = 4\pi;$$

here dA is the element of area for the metric ds^2. Thus a necessary condition on K is that it is positive somewhere. The problem may be formulated as a variational problem. In [M10] Moser proves that if K is even on the sphere (i.e., symmetric) and $K > 0$ somewhere, then the desired conformal metric ds^2 exists. The proof is based on an inequality, a sharp form of one by N. Trudinger in [M8]. As I have mentioned, estimates, or inequalities, play a crucial role in differential equations—indeed, in much of analysis. Paper [M8] is beautiful. Here is an inequality of [M8] used in the geometric problem above:

Theorem 4. *On S^2 in the standard metric let u be a smooth function satisfying*

$$\int_{S^2} |\nabla u|^2 \le 1, \quad \int u = 0.$$

Then there is an absolute constant c such that

$$\int e^{4\pi u^2} \le c.$$

The constant 4π is optimal.

Corollary.

$$\log \int e^u \le \frac{1}{16\pi} \int |\nabla u|^2 + \frac{1}{4\pi} \int u + c' ,$$

with c' a different constant, for any u for which the right-hand side is finite.

This inequality is used in the proof of the geometric result above. In [M8] the sharp form of Trudinger's inequality for a function u in a domain D in $\mathbb{R}^n$ is

Theorem 5. *For u with compact support in D and*

$$\int |\nabla u|^n \le 1,$$

there is a constant c depending only on n such that

$$\int e^{au^p} \le c\,\text{meas}(D),$$

for $p = n/(n-1)$ whenever $a \le a_n = n\,\omega^{\frac{1}{n-1}}$, where ω is the surface area of S^{n-1}. For $a > a_n$, however, the integral on the left can be arbitrarily large.

Trudinger treated $p < n/(n-1)$.

Papers [M6] and [D-M], the latter with B. Dacorogna, treat mappings preserving volume. In [M6] Moser proved that if there is a diffeomorphism between two compact Riemannian manifolds and if the manifolds have equal total volume, then there is a diffeomorphism preserving the volume element. [D-M] treats more general problems. It studies existence and regularity of a diffeomorphism u of the closure $\overline{\Omega}$ of a bounded open set Ω such that

$$\begin{aligned} \det \nabla u(x) &= f(x) \quad \text{in } \Omega \\ u(x) &= x \qquad \text{on } \partial\Omega. \end{aligned}$$

Here f is a given positive function. As an application they prove the existence of a volume-preserving map u with given boundary data, i.e., $u = \psi$ on $\partial\Omega$ with $\psi \in \text{Diff}(\overline{\Omega})$. This is carried out

At the telescope, 1979.

in the space $C^{k,\alpha}$, with $k \geq 1$ and $0 < \alpha < 1$, using elliptic theory and Schauder estimates. An interesting case is that of C^k, k an integer > 0. Here they use the implicit function theorem. In a bounded domain Ω in $\mathbb{R}^n$, with $\partial\Omega$ in C^k, suppose f and g are functions > 0 in $C^k(\overline{\Omega})$ such that

$$\int_\Omega f = \int_\Omega g.$$

They prove that there exists $\phi \in \mathrm{Diff}^k(\Omega)$ with $\phi(x) = x$ on $\partial\Omega$ such that for every open subset E of Ω,

$$\int_E f = \int_{\phi(E)} g.$$

In the 1980s Moser derived very interesting results on foliations of codimension one on a torus $\mathbb{T}^m$ whose leaves are minimals of a nonlinear variational problem (for example, minimizing area). A foliation is called *minimal* if every leaf minimizes the variational problem when viewed on the covering space $\mathbb{R}^m$. He studied existence and stability of leaves that are graphs. With any minimal foliation he associated an asymptotic normal vector $\overline{\alpha}$, which can be written as

$$\overline{\alpha} = (\alpha_1, \ldots, \alpha_n, -1) = (\alpha, -1);$$

α is called the *slope*. He proved that for any leaf of the form $x_{n+1} = u(x')$ with $x' = (x_1, \ldots, x_n)$, there is a unique slope $\alpha \in \mathbb{R}^n$ such that

$$\sup_x |u(x) - (\alpha, x)| < \infty.$$

Furthermore, α is the same for all leaves. To every α in $\mathbb{R}^n$ one cannot always find a corresponding foliation, but there exists a minimal "lamination", a foliation of a certain subset of $\mathbb{T}^m$, which is a Cantor set if not all of $\mathbb{T}^m$. The minimal foliation is described by a nonlinear partial differential equation, and a "lamination" corresponds to a weak solution with discontinuities. The work is related to Aubry-Mather Theory. Much of this work is presented in the beautiful lectures [M15], where other references may be found. Very recently, L. Caffarelli and R. de la Llave devised a very different approach to these foliation problems.

In a 1988 paper Moser extended KAM theory to nonlinear partial differential equations to construct quasi-periodic solutions. Three interesting examples are presented in a 1989 paper.

One should add that Moser did work in symplectic geometry (in particular, on finding infinitely many periodic points near a fixed point of a symplectic map, under suitable conditions), on pseudo-holomorphic curves on an almost complex torus, and on electrical network theory. Much of his early work was concerned with spectra of operators and perturbation theory.

We turn finally to complex analysis. In their famous paper [C-M], S. S. Chern and Moser determined local invariants of real hypersurfaces of real codimension one in $\mathbb{C}^n$ under local biholomorphic maps of $\mathbb{C}^n$. If p and p' are points on such hypersurfaces M and M', one says that M and M' are *equivalent* if there is a biholomorphic map ϕ near p in $\mathbb{C}^n$ taking p and M into p' and M'. They determine necessary and sufficient conditions for equivalence. For $n = 1$ the problem is trivial if M and M' are real analytic curves: they are equivalent. For $n \geq 2$, the signature of the Levi form at p is invariant. But they find many more invariants. For $n = 2$ these were already found by É. Cartan in the 1930s.

[C-M] presents two approaches: (i) extrinsic and (ii) intrinsic. I will describe just (i): They approximate M near p by the image of a hyperquadric. In terms of suitable local coordinates in $\mathbb{C}^n$, $(z^1, \ldots, z^{n-1}, w) = (z, w)$ and $w = u + iv$ (p is the origin), M may be written as

$$v = F(z, \overline{z}, u), \quad \nabla F(0) = 0,$$

with F real analytic in case M is real analytic. The Levi form

$$\langle z, \overline{z} \rangle = \sum_{\alpha,\beta=1}^{n} F_{z^\alpha \overline{z^\beta}}\, z^\alpha \overline{z^\beta}$$

is assumed to be nondegenerate. They choose a biholomorphic map ϕ near the origin, preserving 0, such that the image of M osculates the hyperquadric

$$v = \langle z, \overline{z} \rangle$$

to a high degree. The osculation takes place along a curve, and thus they are led to a holomorphically invariant family of distinguished curves on M, called "chains"; these satisfy second-order differential equations. Quite remarkable! Furthermore, Chern and Moser find a parallel transport of certain frames on M; these correspond to a connection found by the approach (ii). There is a big difference when n goes from 2 to $n > 2$. One result is the following: In case M is real analytic and the Levi form $\langle z, \overline{z} \rangle$ is nondegenerate, there is a unique biholomorphic map ϕ, satisfying some normalizing conditions, such that the transformed hypersurface has the form

$$v = \langle z, \overline{z} \rangle + N(z, \overline{z}, u)$$

with N of higher order in $z, \overline{z}$. In the expansion of N in $(z, \overline{z})$, certain specific low-order terms do not appear. (The coefficients in the expansion of N are not holomorphic invariants, however.) These and other results are then extended to sufficiently smooth M, not necessarily analytic.

In 1976 C. Feffermen gave a totally different description of the "chains" and showed that they are governed by Hamiltonian systems of differential equations.

For the case that $M \subset \mathbb{C}^n$ has codimension > 1, not much is known. In 1983 Moser and S. M. Webster treated the case $M^2 \subset C^2$ under a nondegeneracy condition on the complex tangent.

A few personal remarks. Moser and I first met in 1952 in Göttingen, where he was a graduate student, and we became friends when he came to New York University in 1953 for a year. Later he spent many years at the Courant Institute, and we all greatly appreciated his unfailing good sense and mature judgment. It was a privilege to have him as a colleague and a friend.

Henry P. McKean

Complete integrability is an odd subject. One begins with a Hamiltonian flow of $2d$ degrees of freedom, regulated by $dQ/dt = \partial H/\partial P$ and $dP/dt = -\partial H/\partial Q$, in which $Q \in \mathbb{R}^d$ represents the "positions" and $P \in \mathbb{R}^d$ the "momenta" of d classical particles. The Hamiltonian $H = H(Q, P)$ is the total energy; it is a constant of motion: $dH/dt = 0$. Additional (independent) constants of motion $H_2, H_3, \ldots, H_m$ may exist over and above $H_1 = H$. They *commute* if the corresponding flows do so. The original flow and all these new ones too are said to be *(completely) integrable* if the number of commuting constants of motion is as big as it could be ($m = d$). Then, in "action-angle variables", referred to by Mather, the typical invariant manifold $\mathfrak{M}$ obtained by fixing the values of $H_1, \ldots, H_d$ (actions) appears as a product $\mathbb{T}^p \times \mathbb{R}^q$ with $p + q = d$, and the flow reduces there to straight-line motion at constant speed in the natural coordinates (angles).

And there for a long time the story stopped, apart from a series of nice examples: the simple pendulum for starters; then Kepler's (2-body) problem of planetary motion; Jacobi's integration of the geodesic flow on the surface of a 3-dimensional ellipsoid; Kovalevskya's top (1889); C. Neumann's harmonic oscillators (1859), constrained by an external force to move on the surface of a $(d-1)$-dimensional sphere; and, of course, the much simpler example of (free or coupled) harmonic oscillators and their infinite-dimensional analogue, the string with tied ends, regulated by the wave equation in the form $\dot{Q} = \partial H/\partial P$ and $\dot{P} = -\partial H/\partial Q$ with Hamiltonian $H = \int_0^1 (P^2 + Q'^2)\, dx$. Indeed, the subject came to a sad end, lapsing into obscurity for some seventy-five years with the discovery of Poincaré (1890) that the 3-body problem (and, indeed, most any Hamiltonian system) is not integrable.

Playing the cello, 1968.

The most curious aspect of the whole subject—and this despite the last twenty-five years' spectacular discovery of a whole series of infinite-dimensional instances of integrability, summed up in the catchword "KdV and all that"—has been (and still is) a lack of theorems. Folklore aside, one has no means of testing a particular Hamiltonian H for its integrability, nor, should integrability obtain, of finding the right action-angle variables. One has simply to integrate the flow by hand, with tears.

The rebirth of integrability began with the (numerical) discovery by Kruskal and Zabusky in 1965 of "solitons"[4] for the Korteweg-de Vries equation (KdV)

$$\partial Q/\partial t + 6Q\, \partial Q/\partial x + \partial^3 Q/\partial x^3 = 0,$$

this being a leading-edge approximation to long waves in shallow water. This was followed in 1974 by the integration of KdV by Gardiner, Greene, Kruskal, and Miura via action-angle variables provided by the scattering theory of the 1-dimensional Schrödinger operator $L = -d^2/dx^2 + Q$, following ideas of P. Lax [La]. Later this idea served to integrate sine-Gordon, cubic Schrödinger, and the other infinite-dimensional examples cited above, but now I leave these and come directly to Moser's contributions, which clarified and extended the most classical aspects of the subject in a number of ways.

Moser [M11] presented the new developments in their most transparent form in connection with Toda's lattice,[5] a linear assembly of particles coupled by, not Hooke's, but exponential restoring forces:

$$\ddot{x}_k = e^{x_{k-1} - x_k} - e^{x_k - x_{k+1}} \quad \text{for } k = 1, \ldots, n,$$

subject to $x_0 = -\infty$ and $x_{n+1} = +\infty$. Flaschka in 1974 had already proved the integrability of this system in the periodic case, $x_0 = x_n$, reducing it to an isospectral motion of a tri-diagonal matrix L via a clever substitution. The (simple) eigenvalues of L are commuting constants of motion, and it is the same for Moser's (free) case, only now things are,

Henry P. McKean is professor of mathematics at the Courant Institute of Mathematical Sciences at New York University. His e-mail address is mckean@cims.nyu.edu.

[4] *Solitons are the subject of an article by C.-L. Terng and K. Uhlenbeck in the January 2000* Notices.

[5] *Toda introduced his lattice about 1967. See [To] for an exposition.*

so to say, rational rather than algebraic-geometrical and much simpler. In particular, the "scattering" (from time $-\infty$ to time $+\infty$) can be described by explicit initial and final velocities and phase shifts, suggestive of free motion at times $\pm\infty$, plus the effects of simple pairwise interactions, as for hard balls. The actual integration employs the "Green's function" $R(\lambda) = (\lambda I - L)^{-1}$. The entry $R_{nn}(\lambda)$ may be represented as a sum $\sum_{k=1}^{n} r_k^2(\lambda - \lambda_k)^{-1}$ and also as a continued fraction from which L can be read off directly, as Stieltjes already knew. The numbers r^2 represent norming constants for the associated eigenvectors; the continued fraction is the "inverse spectral map", from spectrum and norming, back to the mechanical variables x and $\dot{x}$ via L. What is so pretty is that the "angle" variables hidden in the norming constants straighten out the flow; in fact, the spectrum does not move, while $(r_k^2)^\bullet = -2\lambda_k r_k^2$ for $k = 1, \dots n$ "projectively", meaning that one solves for $r_k^2(t) = r_k^2(0)\exp(-2\lambda_k t)$ and then restores the necessary $\sum_{k=1}^{n} r_k^2 = 1$.

This is, in miniature, the method by which KdV, sine-Gordon, cubic Schrödinger, and all the other nonclassical examples have been solved. The analogous development for the periodic Toda lattice (and for KdV, etc., too) had, at this date, not yet been made. This development involves 2-sheeted (hyperelliptic) projective curves, a linearizing (Abel) map to a naturally associated torus (Jacobi variety) where the angle variables live, and an inversion from the latter back to the original mechanical variables via Riemann's theta function, but it is unnecessary to go into these matters here.[6]

The method extends, phase shift and all, to Calogero's lattice,

$$\ddot{x} = -\text{grad} \sum_{i<j} U(|x_i - x_j|) \quad \text{with } U(x) = x^{-2},$$

and to Sutherland's lattice with $U(x) = \sin^{-2} x$ (or $\sinh^{-2} x$). This is the content of [M12]. Paper [M13] deals with phase shifts for more general (nonintegrable) repulsive forces. Here the pretty question of recovering U from the phase shifts is raised and solved; in particular, it is proved that the phase shifts vanish only if $U(x) = \text{constant} \times x^{-2}$.

The paper [M14] relates the Toda-Calogero-Sutherland story to Jacobi's geodesic flow on the ellipsoid, to C. Neumann's oscillators, and also to KdV and its accompanying Hill's equation in the "finite-gap" case. These systems are all related and yield to the same method, supplemented by the apparatus of projective curves, Jacobi varieties, and theta functions alluded to before. What [M14] brings out—and this is fascinating—is a further relation of isospectral classes $\mathfrak{M}$ of $n \times n$ matrices to the classical (and not so classical) geometry of confocal quadrics Q: The common eigenvalues pick out $n-1$ distinguished quadrics, $\mathfrak{M}$ is identified with the family of lines simultaneously tangent to each of these, and the eigenvectors of a member of $\mathfrak{M}$ are interpreted as the directions normal to the distinguished quadrics at the points of contact of the associated line. Moreover, the identification of lines to points (plus some reflections) puts $\mathfrak{M}$ into one-to-one correspondence with the Jacobi variety of a hyperelliptic curve, making contact thereby with Staude's 1883 geometrical interpretation of the addition theorem for hyperelliptic integrals. Fascinating indeed—and this is just a sample of the ramifications of "KdV and all that", ranging as they do from solitary waves in shallow water to the purest of pure mathematics, exemplified by projective curves[7] and even, as now seems not implausible, the Riemann hypothesis![8]

Another aspect of KdV is explained in [A-M]: to wit, its surprising connection to Calogero's lattice with $U(x) = x^{-2}$. The latter is completely integrable: if L is the $n \times n$ matrix with $\dot{x}_k$ on the diagonal and $\sqrt{-1}\,(x_i - x_j)^{-1}$ off the diagonal, then the eigenvalues of L are commuting constants of motion. More precisely, the traces $F_m = \text{Tr}\, L^m$ are commuting constants of motion. Now it is a standard fact of mechanics that, in such a case, the flow with Hamiltonian F_k, say, restricts to the locus $\text{grad}\, F_\ell = 0$. Here $\frac{1}{2}F_2 = \frac{1}{2}\dot{x}^2 + \sum_{i<j}(x_i - x_j)^{-2}$ is Calogero's Hamiltonian, and the F_3 flow reduces, on locus $\text{grad}\, F_2 = 0$, to $\dot{x}_i = 12\sum_{j\neq i}(x_i - x_j)^{-2}$ for $i = 1, \dots, n$.

What is remarkable is the connection to KdV: If a solution of the latter is of rational character in x, then it must be of the form $Q(t,x) = 2\sum_{i=1}^{n}(x - x_i)^{-2}$ with time-dependent poles $x_1, \dots x_n$, and these move according to the reduction of Calogero's lattice. The locus $\text{grad}\, F_2 = 0$ is nontrivial only if the poles $x_1, \dots, x_n$ are permitted to be complex, and in this case $(\text{grad}\, F_2 = 0) \simeq \mathbb{C}^d$ provided only that n is triangular number $n = \frac{1}{2}d(d+1)$; moreover, the whole locus is produced by flowing out from the "origin" $Q_0(x) = d(d+1)\,x^{-2}$ by means of the first d flows of the KdV hierarchy.

It may be remarked that such connections are not in any way exceptional; in fact, they are the rule. After all, any completely integrable flow reduces,

[6]*For an exposition see H. P. McKean, Integrable systems and algebraic curves,* Global Analysis, *Lecture Notes in Math., vol. 755, Springer, Berlin, 1979, pp. 83–200.*

[7]*For references see E. Arbarello, Periods of abelian integrals, theta functions, and differential equations of KdV type,* Proc. Internat. Congr. Math. (Berkeley, 1986), *Amer. Math. Soc., Providence, RI, 1987, pp. 623–627.*

[8]*See A. M. Odlyzko, On the distribution of spacings between zeros of the zeta function,* Math. Comp. ***48*** *(1987), 273–308.*

in action-angle variables, to straight-line motion at constant speed, and all such motions look alike.

Adler and Moser presented a remarkably simple, much more detailed description of rational KdV. The general function Q produced by flowing out from Q_0 is of the form $Q(x) = -2D^2 \ell n\, \Theta(x+t_1, t_2, \ldots, t_d)$, in which the time parameters refer to the first d of the KdV flows. Here $\Theta = \prod_{i=1}^{n}(x - x_i)$ is a highly degenerate "theta function" coming from a projective line with just one singular point of degree d, and it is the chief point of the Adler-Moser work that it can be built up recursively, from $\Theta_0 = 1$ and $\Theta_1 = x + t_1$, by solving

$$\Theta'_{n+1}\Theta_{n-1} - \Theta_{n+1}\Theta'_{n-1} = (2n+1)\,\Theta_n^2$$
$$\text{from } n = 2 \text{ up to } n = d-1,$$

each step producing a new constant of integration: $t_2, t_3, \ldots, t_d$.

Finally, I must say something about the work of Kolmogorov, Arnold, and Moser (KAM), which overturned the perception of complete integrability as it had been received for seventy-five years. Moser's work on KAM has been discussed by Mather in the present article.

First, a bit of history. Poincaré had proven the nonintegrability of the 3-body problem, as mentioned before; indeed, he discovered that a very general Hamiltonian system can have no constant of motion besides the total energy itself. This striking fact came to be enshrined in the conventional wisdom as the belief that most small perturbations of such a system must produce "metric transitivity". Fermi even "proved" this in 1923. The advent of fast computation changed the picture entirely. I refer to the discovery of Fermi-Pasta-Ulam that a certain lattice of oscillators with cubic coupling could exhibit almost periodic behavior. By hindsight, this is not so strange: The cubic lattice is a caricature of Boussinesq's completely integrable approximation to long waves in shallow water, and it is the content of KAM that a "nearly integrable" system preserves, perhaps not all, but most of the character of its integrable parent. The parent is assumed to have compact invariant manifolds (tori). Then, under suitable technical conditions, a small change of the parent may and often will cause the breakup of some small proportion of these, but the vast majority survive and almost periodic motion is seen.

This fact has revolutionized the common perception of integrability. After Poincaré it was thought to be just too rare to be of any practical importance, but it is not so: Integrable systems are simple enough to be capable of (more or less) explicit solution, and as they are (more or less) stable in view of KAM and embody genuinely nonlinear behavior in a way that linear approximations could never do, they may help us to a deeper understanding. G. Lamb's beautiful 1971 work in optics provides a striking example.

References

References to selected publications by Jürgen Moser may be found in a sidebar.

[A1] V. I. Arnold, Small denominators I. Mapping the circle onto itself, *Izv. Akad. Nauk SSSR Ser. Mat.* **25** (1961), 21–86 (Russian); English transl., *Transl. Amer. Math. Soc.* **46** (1965), 213–284.

[A2] ____, Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian, *Uspehi Mat. Nauk* **18** (1963), no. 5 (113), 13–40 (Russian); English transl., *Russian Math. Surveys* **18** (1963), 9–36.

[Bi] G. D. Birkhoff, *Collected Mathematical Papers*, vol. II, Amer. Math. Soc., New York, 1950.

[Bo] E. Bombieri, *Theory of Minimal Surfaces and a Counterexample to the Bernstein Conjecture in High Dimensions*, mimeographed notes of lectures held at Courant Institute, 1970.

[DeG] E. De Giorgi, Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari, *Mem. Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur.* (3) **3** (1957), 25–43.

[G-T] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, second edition, Springer-Verlag, 1983.

[K1] A. N. Kolmogorov, On conservation of conditionally periodic motions for a small change in Hamilton's function, *Dokl. Akad. Nauk SSSR (N.S.)* **98** (1954), 527–530 (Russian); English transl., Casati and Ford (eds.), *Stochastic Behavior in Classical and Quantum Hamiltonian Systems* (Volta Memorial Conf., Como, 1977), Lecture Notes in Phys., vol. 93, Springer, Berlin and New York, 1979, pp. 51–56.

[K2] ____, Théorie générale des systèmes dynamiques et mécanique classique, *Proc. Internat. Congr. Math., Amsterdam, 1954*, vol. 1, Noordhoff, Groningen, 1957, pp. 315–333; English transl. in Appendix D of R. Abraham, *Foundations of Mechanics*, Benjamin, 1967.

[Kr] N. V. Krylov, *Nonlinear Elliptic and Parabolic Equations of the Second Order*, Mathematics and Its Applications (Soviet Series), D. Riedel, Dordrecht, 1987.

[La] P. D. Lax, Integrals of nonlinear equations of evolution and solitary waves, *Comm. Pure. Appl. Math.* **21** (1968), 467–490.

[N1] J. Nash, The imbedding problem for Riemannian manifolds, *Ann. of Math.* **63** (1956), 20–63.

[N2] ____, Continuity of solutions of parabolic and elliptic equations, *Amer. J. Math.* **80** (1958), 931–954.

[St] G. Stampacchia, Contributi alla regolarizzazione delle soluzioni dei problemi al contorno per equazioni del secondo ordine ellittiche, *Annali Scuola Norm. Sup. Pisa* (3) **12** (1958), 223–245.

[To] M. Toda, *Theory of Nonlinear Lattices*, Springer Series in Solid-State Sciences, vol. 20, Springer-Verlag, Berlin and New York, 1981; second edition, 1989.