Ergod. Th. & Dynam. Sys. (2002), **22**, 1343–1364
DOI: 10.1017/S0143385702001062

The development of dynamics in the 20th century and the contribution of Jürgen Moser

BORIS HASSELBLATT† and ANATOLE KATOK‡

† *Department of Mathematics, Tufts University, Medford, MA 02155, USA*
(e-mail: Boris.Hasselblatt@tufts.edu)
‡ *Department of Mathematics, The Pennsylvania State University, University Park, PA 16802-6401, USA*
(e-mail: katok_a@math.psu.edu)

Jürgen Moser (1928–1999) was one of the most accomplished mathematicians of the second half of the 20th century and his work had a major impact in broad areas of analysis, especially partial differential equations and dynamical systems, and geometry. In this article we discuss only his contributions to dynamics and closely related areas. We feel that the best way to make the reader understand and appreciate the impact of Moser's work is to discuss it within the framework of some major trends in the development of dynamical systems during the last century. We apologize in advance for omitting many important references, including some key original work. These can be found in secondary sources to which we refer for detailed accounts of various areas. Needless to say, we included in our references all papers by Moser relevant to our discussion. For brief surveys of Moser's work in all areas see [**45, 91**].

The leading theme of virtually all of Moser's work in dynamics is the search for elements of stable behavior in dynamical systems with respect to either initial conditions or perturbations of the system.

1. *Dynamics from Newton to Poincaré*

The major dynamical issues to whose study Moser made contributions can be traced back to the inception of the discipline of dynamical systems in the 19th century. Celestial mechanics had long been a driving force for scientific and mathematical development. It was given a physical basis by Newton when he established universal gravitation as its basic mechanism and as the system of the world [**78**]. The development of calculus and classical mechanics supported a flourishing of celestial mechanics into a universal model of unprecedented predictive power, epitomized by Laplace's monumental *Mécanique Celeste* [**38**]. In particular, the 'great inequality' in the Jupiter–Saturn interaction was described in a way that only a sophisticated theory could: an apparent systematic drift in observational data over time was shown to be a periodic phenomenon with a period of some

900 years (due to resonant periods). The sense that this established supreme order in the universe is reflected in Laplace's famous statement about physical determinism [**39**, p. 2]: 'An intellect which at any given moment knew all the forces that animate nature and the mutual positions of the beings that comprise it ... could condense into a single formula the movement of the greatest bodies of the universe and that of the lightest atom: for such an intellect nothing could be uncertain; and the future just like the past would be present before its eyes.'

Laplace claimed to have established the stability of the solar system, but his argument was found to fall apart when the perturbation theory he used was carried beyond first order. This was noticed by le Verrier, who also found that higher-order perturbation theory led to series solutions where small denominators might cause convergence problems. From then on it was understood that stability of the solar system should be established by obtaining series solutions that converge uniformly in time. Throughout the 18th and most of the 19th century the leading theme of research in classical mechanics in general and celestial mechanics in particular was the search for more or less explicit expressions of solutions either via elementary functions and quadratures or as convergent power series in properly chosen coordinates. The hope was that accomplishing this would clarify the qualitative long-term behavior. One situation in which this approach works is the completely integrable case. Using modern language, if there are sufficiently many commuting constants of motion (first integrals) then most of the phase space of a mechanical system decomposes into invariant manifolds, and the compact ones among these are tori on which the motion is quasiperiodic with respect to properly chosen coordinates. While the classics never hoped that the most important systems studied in classical mechanics, such as the n-body problem, were completely integrable in full generality, complete integrability was an attractive paradigm of regular and divinely ordained behavior, which motivated the search for explicit solutions in a more general context.

At the end of the 19th century Poincaré made great contributions to classical celestial mechanics and in the process discovered that even in the three-body problem the long-term behavior could be vastly more complicated than anyone would have anticipated [**9, 79**]. Specifically, he noticed the presence of homoclinic tangles and inferred that the series solutions of the equations of motion diverge. ('These intersections form a type of trellis, tissue, or grid with infinitely fine mesh ... The complexity of this figure is striking, and I shall not even try to draw it. Nothing is more suitable for providing us with an idea of the complex nature of the three-body problem' [**80**, Ch. 33, §396].)

This watershed work marks the beginning of the theory of dynamical systems in the modern sense. In it Poincaré began to develop qualitative (topological), probabilistic, and global methods of analysis of dynamical systems. Poincaré recognized the study of small perturbations of completely integrable Hamiltonian systems as the central problem in dynamics: the solar system can be viewed as a (relatively) small perturbation of the integrable Kepler problem. Since homoclinic tangles can appear for arbitrarily small perturbations of integrable systems and are persistent, it came to be expected that away from integrable systems highly unstable behavior is typical. From a different direction the Boltzmann ergodic hypothesis also brought about the necessity of studying complicated behavior in systems with many degrees of freedom.

2. *Dynamics in the first half of the 20th century*

Thus, the dominant theme of research in qualitative dynamics in the first half of the 20th century, or more precisely, in the 65 years from the seminal insights of Poincaré to the Kolmogorov discovery in 1954, was a somewhat cautious exploration of the complex unstable behavior which much later became known as 'chaotic'. In the process of this development the modern language of discourse in dynamics was developed. In particular, certain notions were recognized as characteristic for complex behavior: coexistence of periodic orbits of arbitrarily large periods, transitivity on significant parts of the phase space, local exponential divergence of nearby orbits, existence of non-trivial invariant measures, and ergodicity with respect to such measures.

It is interesting that the most important concept in this circle of ideas, entropy, was discovered only in 1957 and again by Kolmogorov, who brought into dynamics insights from Shannon's information theory.

With a little simplification one can say that the progress revolved around two themes: first, the search for tractable models with complex behavior to augment the old model of completely integrable quasiperiodic behavior, and second, the search for methods of establishing the above-mentioned qualitative features in various classes of concrete systems. In the process the emphasis shifted from traditional techniques based on classical analysis and differential equations towards new geometric, variational and probabilistic methods. Without any pretense to a comprehensive account let us mention some of the highlights.

The search for tractable models led to the study of the free motion of a particle (the geodesic flow) on surfaces of constant negative curvature. By 1898 Hadamard had considered (non-compact) surfaces in $\mathbb{R}^3$ of negative curvature, showing that each homotopy class (except for the 'waists' of cusps) contains a unique geodesic [**21**, III.31,37]. A popular classic by Duhem [**17**] used this to illustrate that determinism in classical mechanics does not imply any practical long-term predictability. To today's reader his description is a shrewd translation of symbolic dynamics into everyday language. Indeed, several authors trace back symbolic dynamics to this paper of Hadamard [**21**, II.20], including Birkhoff ('the symbols effectively introduced by Hadamard' [**12**, p. 184]). Between 1920 and 1940 the geodesic flows of surfaces of constant negative curvature with finite volume and finitely generated fundamental group were extensively studied. Morse introduced symbolic dynamics to produce non-periodic recurrent geodesics [**47**] and density of closed geodesics [**48**]; symbolic dynamics was also used by Artin [**7**] and Koebe [**34**]. Topological transitivity and density of periodic orbits were established by Koebe and Löbell [**34, 43**] (after an earlier work for special cases by Artin [**7**] and Nielsen) and topological mixing by Hedlund [**24**]. This combination of properties is among the hallmarks of 'chaotic' behavior in a contemporary sense.

Ergodic theory can be said to have begun with Poincaré's memoir [**79**], where the recurrence theorem first appeared. Preservation of phase volume in conservative mechanical systems made ergodic theory a natural tool. The 1930s witnessed the birth and the first flourishing of ergodic theory in the modern sense. The ergodic theorems of von Neumann [**76**] and Birkhoff [**10**] established the existence of time averages, which adds significant physical motivation, especially in the ergodic case when these must be constant. The founding text of modern ergodic theory is von Neumann's paper [**77**]. Its title clearly indicates that classical mechanics was von Neumann's principal motivation while in its content Hamiltonian and other classical systems do not figure prominently. Von Neumann found an abstract version of quasi-periodicity (pure point spectrum) and classified such systems. He also found that special flows over a circle rotation with piecewise differentiable roof and non-zero sum of jumps provide simple geometric examples of weakly mixing systems, i.e. systems without factors with pure point spectrum. E. Hopf wrote a textbook on ergodic theory as early as 1937 [**29**]. A further motivation for ergodic theory as a powerful and general tool in dynamics came from the work of Krylov and Bogolubov who showed that any topological dynamical system has an invariant measure [**37**]. Hopf and Hedlund looked at ergodic properties of classical systems which by then had been shown to possess complicated topological orbit structure, proving ergodicity [**28**], and mixing [**25**] for geodesic flows on compact surfaces of constant negative curvature. At the time mixing looked like the most natural hallmark of complicated behavior from the statistical point of view, contrasting with a pure point spectrum, which is characteristic for stable behavior. It took another 20 years until more fitting concepts, entropy and derivative notions of strong mixing, such as completely positive entropy (the K-property), were discovered.

The crowning achievement in the work on ergodic theory of classical systems in the 1930s was the work of Hopf [**30**] in which he proved that for compact surfaces of non-constant (predominantly) negative curvature the Liouville measure (phase volume) is ergodic and indeed mixing. This work provides the bridge to the next phase in the development of the theory which began in the 1960s primarily with the work of Anosov and Sinai.

Parallel to these developments Birkhoff advanced the original program of Poincaré by looking at complicated phenomena in sufficiently general Hamiltonian and other classical systems. Among other things Birkhoff proposed billiards inside convex plane curves as a testing ground for the qualitative study in mechanical systems ('In this problem the formal side, usually so formidable in dynamics, almost completely disappears, and only the interesting qualitative questions need to be considered.' [**11**, §VI.6, p. 170]). Convex billiards give rise to twist maps, which have been especially useful, often as the simplest models in which to test ideas and methods intended for application to Hamiltonian or Lagrangian systems. Birkhoff paid great attention to topological methods. He was proud of his proof of Poincaré's Last Geometric Theorem, which in particular implies existence of periodic orbits with arbitrarily high periods for a general class of Hamiltonian systems with two degrees of freedom.

In 1927 B. van der Pol and J. van der Mark reported on experiments with a 'relaxation oscillator' circuit built from a capacitor and a neon lamp (this is the nonlinear element) and a periodic driving voltage [**81**]. They were interested in the fact that in contrast to a linear oscillator, which exhibits multiples of a base frequency, these oscillations were at 'submultiples' of the basic frequency, i.e. half that frequency, a third, and so on down to 1/40th, as the driving voltage increased. They obtained these frequencies by listening 'with a telephone coupled loosely in some way to the system' and reported that 'Often an irregular noise is heard in the telephone receivers before the frequency jumps to the next lower value. However, this is a subsidiary phenomenon, the main effect being the regular frequency demultiplication.' This irregular noise was one of the first experimental encounters with complex dynamics. Although it was then overlooked, in retrospect it is a natural precursor to the work of Cartwright and Littlewood on relaxation oscillations.

Much of the search for methods to establish orbit complexity in concrete systems can retrospectively be seen as related to the phenomena that arise from homoclinic (or heteroclinic) tangles. Other than in Poincaré's initial discovery, this is fully explicit in the work of Birkhoff, who established that also in this context one obtains periodic points of arbitrarily high periods. This was much later sharpened by Smale to the existence of a horseshoe (the Birkhoff–Smale Theorem), which allows us now to see connections to situations where tangles as such are not the most apparent feature. This is the case with the work of Cartwright and Littlewood [**13, 14**], which was to lead Smale to the discovery of horseshoes. Here the concrete system was a radar circuit that exhibited erratic relaxation oscillations. (Cartwright and Littlewood were aware of and guided by Poincaré's work.)

The variational approach, introduced by Hadamard, Birkhoff and Morse, served as one of the tools for finding complicated behavior. Until the 1980s variational methods were only applied to finding motions of the simplest possible type, i.e. closed orbits for differential equations and periodic points for discrete-time systems. However, by setting many variational problems for the same system these authors were able to find periodic orbits of arbitrarily large periods, thus establishing one of the features of complicated behavior. In these early applications of variational methods the objectives were always critical points of the action functionals of finite Morse index, such as extrema or 'mountain passes'. Such problems usually pose only moderate analytical difficulties and classical homology theory provides the proper topological tools. The earliest work dealing with the infinite index case was a 1948 paper by Seifert [**84**] that involved a degenerate Jacobi metric.

3. *Second half of the century: stable and random motions*

A crucial breakthrough in the development of the general theory of dynamical systems appeared with Kolmogorov's discovery of large sets of stable motions within Hamiltonian systems of a fairly general kind. While the motions themselves fit into the classical quasiperiodic paradigm, the set of tori filled with such motions is complicated, typically (transversely) a Cantor set of positive Lebesgue measure. We discuss this topic in more detail in the next section. At this point it is enough to point out that Kolmogorov's work led to a fundamental change in the general outlook of the theory.

First, analytical methods were back, albeit at a new, more sophisticated level, to detect and describe various stable elements in the orbit structure.

Second, the quest for complexity now focused on two directions: (i) understanding how complex behavior coexists with large sets of stable motions, and (ii) finding proper models for genuinely stochastic or chaotic behavior unencumbered by any admixture of stability; the latter usually goes under the name of hyperbolic dynamics, often with various adjectives, such as 'uniform', 'partial', 'non-uniform' (for an introduction and partial overview see [**33**]). Direct analytical, geometric, and variational methods were developed to achieve this end.

Third, a strong revival of the interest in the classical subject of completely integrable systems started in the 1960s, prompted on the one hand by the discovery of new mechanisms of complete integrability, which originally appeared in some important infinite-dimensional systems coming from partial differential equations, and, on the other hand, by KAM bringing completely integrable systems back into the focus of attention.

Another important development was a blossoming of the global variational approach, which often involves a blend of hard analysis and geometric considerations. Variational methods provide an opening for the coexistence problem (i) above, which otherwise often looks intractable. In addition to finding periodic solutions for Hamiltonian and Lagrangian systems, variational methods were now used to find new elements of the orbit structure that either reflect some traces of stable behavior (Aubry–Mather sets) or complicated motions (homoclinic, heteroclinic, oscillating) that accompany stable structures.

Moser was a major participant in all of these developments. Always keenly interested in the work of others, he was able to discern the fundamental trends and invariably made essential, often fundamental, contributions. On several occasions lecture series of his at the early stages of a development brought great clarity, opened new perspectives, and provided major insights on his own part [**53, 55, 57, 64, 65**]. Moser's own contributions to different major developments in dynamics are of unequal volume and importance. Still, we cannot think of another mathematician in the period after 1960 who had such a broad view and comprehensive understanding of virtually all major trends in dynamics and influenced their development to a similar degree. In fact, the only major development that mostly escaped Moser's creative attention seems to be the circle of ideas from probability theory and statistical mechanics, which entered dynamics through symbolic dynamics and ergodic theory.

Moser was usually well aware of the work of his predecessors and was gracious and generous in giving them credit. This knowledge also facilitated his search for deep understanding of mechanisms as well as connections between different parts of mathematics. In his work he usually searched for wisdom rather than simply knowledge, and thus he strongly emphasized developments of methods and insights over pushing a specific result to the limit. Accordingly, he sometimes described the outcome of his own work as methods rather than theorems. This also influenced his exposition. In order to convey his understanding he wrote beautiful papers with lucid prose, and often chose to add inessential simplifying hypotheses to make the central idea stand out more clearly. A prime example is the excessive number of derivatives (333 was chosen not very seriously) assumed in his original KAM paper [**51**]. Several of his papers are easier to read than their review in *Mathematical Reviews*. This and his finely crafted lecture notes greatly magnified his influence. In addition he was diligent in bringing to the physics community those insights that would be useful to them.

As it turned out, the search for paradigms for complex behavior, at least in the earlier part of the period, also centered on the search for stability, but in a different sense than stability of quasiperiodic motions: preservation of the global (often complex) orbit structure under perturbations of the system (structural stability and related notions). The leading paradigm in this area is that of exponential instability of orbits, or hyperbolic behavior. The functional-analytic methods based on implicit-function and fixed-point type theorems played an important role and while Moser's contribution here was not large in volume, it has been quite influential [**56, 58**].

A major role in the development of global variational methods and symplectic geometry was played by two students of Moser, C. Conley and P. Rabinowitz, and by E. Zehnder who was greatly influenced by Moser [**15, 16, 82**]. In the rest of this note we discuss the above-mentioned major areas and outline Moser's contributions. The order we have chosen is based on the importance of Moser's contributions (as we perceive it), and thus differs both from the historical order of appearance of major ideas, and from the order of Moser's own work.

4. *KAM*

In 1954 Kolmogorov announced [**35**] that under general non-degeneracy conditions many invariant tori in a completely integrable analytic Hamiltonian system persist under any sufficiently small analytic perturbation. Non-degeneracy in a system with n degrees of freedom means that the frequencies of the quasiperiodic motions on the n-parametric family of n-dimensional invariant tori change locally in all possible ways. The persistent tori are those whose frequency vector $\omega = (\omega_1, \dots, \omega_n)$ satisfies a Diophantine condition for constants c and r depending on the size of the perturbation, i.e.

$$|\langle \omega, k\rangle - m| > c|k|^{-r}, \quad \text{for } k_1, \dots, k_n,\ m \in \mathbb{Z} \tag{1}$$

where $k = (k_1, \dots, k_n)$, $\langle \omega, k\rangle = \sum_{i=1}^{n} \omega_i k_i$, and $|k| = \sum_{i=1}^{n} |k_i|$.

Thus, stable Diophantine tori are not isolated objects, like hyperbolic equilibria or periodic orbits whose stability had been known before. On the contrary, they cover a substantial part of the phase space: their union has positive Lebesgue measure, although it is usually nowhere dense. Kolmogorov's discovery dramatically altered the perceived picture of typical long-term behavior in non-integrable systems. On the one hand, stable behavior (quasiperiodic motions fully within the classical paradigm) is abundant; on the other, complexity and disorder exist in the rest of the phase space.

Kolmogorov's revolutionary work had two deficiencies. First, it did not directly apply to the solar system for which the non-degeneracy assumption does not hold, and second, results were not supplied with complete proofs. The paper [**35**] did, however, contain a description of the method, which is based on a sophisticated modification of an infinite-dimensional version of the classical Newton method for solving equations. We postpone a description of this method in a form given to it by Moser in order not to interrupt the remarkable story of how Kolmogorov's prophetic insight developed into a major tool in several areas of contemporary mathematics. Jürgen Moser played the central role in this transformation. The early history of what became known as 'KAM theory' or the 'KAM method' is well recorded, including Moser's personal account [**74**], and we recall some parts of it here.

At the time of Kolmogorov's announcement Moser was a young mathematician trained mostly in differential equations and classical analysis, and deeply interested in the classical problems of celestial mechanics. The primary mathematical influence in his early development was Carl Ludwig Siegel. In his earlier work, Siegel managed to overcome the difficulties of 'small denominators' (which were also central for carrying out Kolmogorov's approach) in a somewhat simpler 'center problem' [**85, 86**]. Siegel used a high-powered version of the classical Cauchy majorization method, but, as Moser showed later, the Siegel problem was also an excellent stepping stone for the application of Kolmogorov's method (see [**53**] for details and references). Moser became interested in Kolmogorov's work as soon as he saw the announcement [**36**], which he was asked by *Mathematical Reviews* to review. Not persuaded by the outline of the convergence arguments in Kolmogorov's paper [**35**], he wrote a letter to Kolmogorov, asking for the details. It is not known whether Kolmogorov received the letter (it was after all only the very beginning of the post-Stalin 'thaw'), but in any case Moser never received an answer. Paradoxically, this may have been fortunate for the development of mathematics, because in due time two powerful insights into the problem were developed independently, by V. I. Arnold, a brilliant student of Kolmogorov, and by Moser. This arguably made the resulting method more versatile and powerful than it might otherwise have been.

Having only Kolmogorov's announcement to guide him, Moser developed a powerful functional-analytic machinery which in its various versions was applicable to a number of problems in partial differential equations, geometry, and dynamics. One part of it became known as the 'Nash–Moser Implicit Function Theorem' [**50**] (being hesitant to turn methods into theorems, Moser preferred the term 'Nash–Moser Implicit Function Method'), and another led to a realization of the Kolmogorov iteration scheme. Moser announced his result in the latter area at a conference in 1961 and published it in 1962 [**51**]. The Moser result, usually referred to as the 'Invariant Curve Theorem', is different from the theorem announced by Kolmogorov in a number of respects. First, it deals only with the case of Hamiltonian systems with two degrees of freedom. More precisely, Moser considers maps that appear as a section (first return map) of such systems on a hypersurface of constant energy. For a completely integrable system, in a tubular neighborhood of an invariant torus, the section map has the form

$$f(r, \theta) = (r, \theta + \alpha(r)), \tag{2}$$

where θ is the angular coordinate on the circle, which appears as the section of an invariant torus, and r is a first integral of the system. The non-degeneracy condition implies the *twist condition* $\alpha' \neq 0$ for the section map. Naturally, for a small perturbation of an integrable system the section map is a small perturbation of the map (2). Existence of an invariant torus for the perturbed system near a torus for the integrable system is equivalent to the existence of an invariant curve for the perturbed section map near a circle $r = r_0$. The Diophantine condition (1) becomes a Diophantine condition for the rotation number $\alpha(r_0)$. The section maps for Hamiltonian systems are area preserving (if the transverse coordinate r is chosen properly) and exact, i.e. the area under the graph of any function $r = \varphi(\theta)$ does not change under the map. This, in particular, implies that any such graph intersects its image. Moser used only this weaker property to prove the preservation of invariant circles with Diophantine rotation number. Thus his result holds for a more general class of systems than Hamiltonian systems with two degrees of freedom. Another difference was that Moser only assumed the existence of a finite, albeit (for expository reasons) large, number of derivatives of the integrable map, as well as closeness of these derivatives for the perturbation, instead of analyticity in Kolmogorov's scheme. As Moser points out in [**74**], this came to be considered as his main achievement. Naturally, the invariant curves in Moser's result are also only finitely differentiable. Moser mentions that at the time he was, in fact, unable to verify Kolmogorov's claim precisely, i.e. to prove existence of analytic invariant curves for small analytic perturbations of integrable twist maps (he succeeded in that a bit later; see, for example, [**54**]). We would like to stress that Moser's is the first proof of a new kind of stability result applicable to Hamiltonian systems. Since Moser was dealing with the finite smoothness situation where a loss of regularity inevitably appears, he made a very fruitful contribution by introducing smoothing operators, an idea influenced by the work of Nash and J. Schwartz. About a year and a half later V. I. Arnold published his proof of Kolmogorov's Theorem in the original setting of analytic Hamiltonian systems for any number of degrees of freedom [**5**]. This is how the acronym KAM came about.

As Arnold himself acknowledges, his proof differed considerably from Kolmogorov's scheme. It is adapted to a variety of problems in Hamiltonian mechanics, in particular the problem of stability of the solar system which Arnold attacked in [**6**]. We refer the reader to [**45**] for a brief account that explains some of the technical points, and to an excellent account by de la Llave [**42**] for a detailed description and comparison of various versions and flavors of the KAM method. According to de la Llave, Arnold's method, while technically more complicated and more difficult to carry out, provides more information in the analytic Hamiltonian case. On the other hand, as de la Llave points out, this method has never been fully carried out in the differentiable case.

Here is a general outline of the KAM method following [**53**] for a situation where one seeks a smooth or analytic conjugacy between a certain dynamical system g (the 'model') and its perturbation f on a certain part of the phase space, such as an invariant torus.

The idea is to cast the conjugacy equation as an implicit-function problem. For that one writes the conjugacy equation as

$$g = \mathcal{F}(f, h) := h^{-1} \circ f \circ h.$$

The main feature of the operator $\mathcal{F}$ is the 'group property':

$$\mathcal{F}(f, \varphi \circ \psi) = \mathcal{F}(\mathcal{F}(f, \varphi), \psi), \quad \mathcal{F}(f, \mathrm{Id}) = f. \tag{3}$$

As in the elementary Newton method, we want to linearize the operator and hence we need to assume that there is a linear structure on a neighborhood of (g, Id) in the appropriate functional space and that f is close to g. Then one can linearize $\mathcal{F}$ on this neighborhood. Write $D_1\mathcal{F}$ and $D_2\mathcal{F}$ for the partial differentials with respect to f and h, respectively. To look for an 'approximate solution' $h = \mathrm{Id} + w$ of the conjugacy equation linearized at (g, Id), write

$$\mathcal{F}(f, h) = \mathcal{F}(g, \mathrm{Id}) + D_1\mathcal{F}(g, \mathrm{Id})(f - g) + D_2\mathcal{F}(g, \mathrm{Id})(h - \mathrm{Id}) + \mathcal{R}(f, h),$$

where $\mathcal{R}(f, h)$ is of second order in $(f - g, h - \mathrm{Id})$. In other words, if h solves the linearized equation (obtained by dropping $\mathcal{R}$), then $w = h - \mathrm{Id}$ is a solution of

$$\mathcal{F}(g, \mathrm{Id}) + D_1\mathcal{F}(g, \mathrm{Id})(f - g) + D_2\mathcal{F}(g, \mathrm{Id})w = g.$$

Using that $D_1\mathcal{F}(g, \mathrm{Id}) = \mathrm{Id}$ (since $\mathcal{F}(\cdot, \mathrm{Id}) = \mathrm{Id}(\cdot)$ by (3)), this simplifies to

$$(f - g) + D_2\mathcal{F}(g, \mathrm{Id})w = 0.$$

If $D_2\mathcal{F}(g,\mathrm{Id})$ is invertible, then $w = -(D_2\mathcal{F}(g,\mathrm{Id}))^{-1}u$, where $u = f - g$. In this case, w is of the same order as u, and substituting $h = \mathrm{Id} + w$ into $\mathcal{F}(f,h)$ we obtain a function $f_1 = h^{-1} \circ f \circ h = \mathcal{F}(f,h) = g + \mathcal{R}(f,h)$, so the size of $u_1 = f_1 - g = \mathcal{R}(f,h)$ should be of second order in the size of $u = f - g$. To justify this, one needs to estimate the difference between $\mathcal{F}$ and its linearization near (g,Id).

Thus, consider an iterative process as follows. Assuming that $f_1,\dots,f_n$ have been constructed, we solve the equation

$$f_n - g + D_2\mathcal{F}(g,\mathrm{Id})w_{n+1} = 0$$

and set

$$h_{n+1} = h_n \circ (\mathrm{Id} + w_{n+1}) \quad \text{and} \quad f_{n+1} = (\mathrm{Id} + w_{n+1})^{-1} \circ f_n \circ (\mathrm{Id} + w_n).$$

The last step of the construction is the proof of convergence of the sequence h_n in an appropriate topology. It follows from the same estimates that provide the fast decrease of the size of the $f_n - g$.

Notice that at every step the linear part is inverted at (g,Id), rather than at the intermediate points as in the elementary Newton method. This is the main feature that makes analytic difficulties manageable (however often still quite formidable).

Moser's emphasis was on exploring broad implications of the KAM method (as he points out in [**74**], he liked this name better than 'KAM theory'), both in the framework of stability of Hamiltonian and other dynamical systems, and beyond. Accordingly, he worked both on transforming the iteration scheme into a precise general theorem of the implicit function type, and on a variety of applications of the scheme. The latter included finding invariant tori in a variety of classes of dynamical systems and different problems in dynamics, geometry and partial differential equations. He achieved this during the 1960s and early 1970s by the cumulative impact of his work [**53–55, 57**] Moser's parting view of the subject's development over 40 years is in [**73**].

We do not feel that in an overview note like the present one it would do full justice to Moser's work to list and briefly discuss his specific achievements in this area, which is most important for his whole *oeuvre*. Many of the most important points are in details, in the sophisticated ways of setting functional equations, choosing classes of spaces, finding a proper combination of elegance and broad applicability, and carrying out hard estimates. Aside from the above-mentioned works by Moser himself, we again recommend the book-length exposition by de la Llave [**42**], which will appear in an expanded form as a separate book, as well as some references therein, and the expository articles in the KAM section of the collection [**1**]. In particular, [**40**] provides an elegant exposition based on a 1985 lecture by Moser for the simplest case of the Lagrangian version of the KAM method which unlike the description above does not use iterated coordinate changes. This method is based on [**90**] and was later developed in [**83**].

Finally, let us mention later pioneering work of Moser's in the area of KAM theory [**68**]. It deals with simultaneous linearization of two commuting diffeomorphisms of a circle. For a single map the theory was developed to fairly definitive results well beyond the perturbation case considered by Arnold [**4**], mostly due to efforts of M. Herman and J.-C. Yoccoz [**26, 88**]. Diophantine-type conditions for the rotation number are not only sufficient but also necessary for a map with this rotation number to be smoothly conjugate to a rotation. For commuting maps a new phenomenon appears: the rotation numbers of individual maps in a $\mathbb{Z}^k$ action may be non-Diophantine, but still the action is linearizable due to the absence of too strong resonances. Moser's work still stands as the principal result in this direction.

5. *Aubry–Mather theory and beyond: weak traces of stable behavior*

Cantor sets and similar structures (e.g. solenoids) are recognizable hallmarks, almost icons, of complex behavior in dynamics. They appear in horseshoes, 'strange attractors', at various global bifurcations and suchlike. This makes it remarkable that such structures also appear as the last vestiges of stable behavior, as 'ghosts' of vanished invariant tori.

We begin with a brief general account. Most of the details as well as references can be found in [**33**]. These can be supplemented by [**8**] and [**46**].

In one of his early works Poincaré provided the first ever exhaustive analysis for a general class of dynamical systems. What allowed him to do so was the low dimension of the phase space he considered. Starting from differential equations on the two-dimensional torus and taking a section, Poincaré obtained invertible maps on the circle. For such maps Poincaré gave a complete description of possible behavior by means of a crucial invariant, the rotation number. Interesting qualitative behavior appears if the rotation number is irrational, and in this case Poincaré proved the following dichotomy: either all orbits are dense and the map reduces to an irrational rotation by a continuous coordinate change, or the limit points form a Cantor set. In the latter case the orbits within the Cantor set are dense in it, and the whole system can be described as an irrational rotation with each point from a finite or countable set of orbits blown up into an interval. Thus the order of points on each orbit is the same as for the corresponding rotation. Poincaré, of course, was interested not so much in an arbitrary homeomorphism, as in maps that appear as section maps for differential equations, i.e. diffeomorphisms. Poincaré left open the question of whether the curious second case in the dichotomy may actually appear in a smooth situation. In reading Poincaré's work in new areas of mathematics that he discovered or anticipated (as opposed to his work in well-established areas) one is often left with the impression that he did not have strong enough motivation to elaborate on many questions that posed even moderate technical difficulty. Probably, Poincaré had such a tremendous wealth of new ideas and concepts that he preferred to move on, leaving loose ends to his followers. Unfortunately, there were not that many direct followers of Poincaré working in dynamical systems, and almost 50 years passed until Denjoy resolved the problem by showing that for twice differentiable maps the second possibility in the Poincaré dichotomy never occurs. Denjoy also showed that there are slightly less regular maps (C^1 diffeomorphisms whose derivative is α-Hölder continuous, where α can be arbitrarily close to 1), for which this possibility does occur. Such maps are usually referred to as Denjoy counterexamples. Thus, limit Cantor sets were shown to be an oddity related to low regularity.

Nevertheless, behavior that exactly corresponds to Denjoy counterexamples inevitably appears in broad and natural classes of *two-dimensional* dynamical systems. We already encountered such maps as section maps of small perturbations of Hamiltonian systems with two degrees of freedom, and more general maps for which Moser proved existence of invariant curves. The proper class is that of twist maps of an annulus $A = \mathbb{R} \times S^1$:

$$f(x, \theta) = (X(x, \theta), \Theta(x, \theta)), \quad \text{where } \frac{\partial \Theta}{\partial x} > 0. \tag{4}$$

Recall that even for small perturbations of integrable twist maps (2) only a nowhere dense set of invariant curves close to the corresponding circles $r =$ constant persists. What happens with the remaining circles? The answer, in retrospect, is remarkably simple and holds not only for small perturbations of integrable twists, but for all twist maps with the graph-intersection property, including exact area-preserving maps: there are traces or ghosts of all of them. These traces are invariant sets that are partial graphs of Lipschitz functions $x = f(\theta)$, and they consist of orbits whose θ coordinate changes in exactly the same order as for the corresponding rotations. For rational rotation numbers those are periodic orbits, at least two for each rational number p/q with p and q relatively prime. They have period q and make p rounds around the circle. These orbits were known to Birkhoff (they are often called Birkhoff periodic orbits), who extensively studied twist maps as one of the crucial models for mechanical problems. It is quite remarkable that Birkhoff did not take the simple step of looking at limits of those orbits corresponding to sequences of rational numbers approximating an irrational number α. If a twist map has an invariant curve with rotation number α and dense orbits then the corresponding Birkhoff orbits converge to that curve (it is necessarily unique due to the twist property (4)). However, even if an invariant curve does not exist, there is still a limit, and it is exactly a Cantor set with the dynamics of a Denjoy counterexample. More precisely, the limit set for the Birkhoff orbits always contains such a Cantor set, often together with other ordered orbits traveling inside its gaps. These sets were discovered independently by Aubry and Mather around 1980 for volume-preserving maps. For general twist maps with the graph intersection property their existence was proven several years later independently by D. Bernstein and R. Hall.

Existence of Aubry–Mather sets brings a new twist into the story of stable behavior in general systems. The motions on those sets still can be viewed as fairly regular and ordered, albeit in a weaker sense than those on invariant tori. It should also be noted that, unlike KAM, Aubry–Mather theory only partially extends to higher dimensions [**44**].

Now we come to Moser's highly non-trivial and original part in this story. He did not take part in the early development of Aubry–Mather theory, which was mostly completed by 1985. Around that time Moser became interested in the theory (see his lectures [**65**]) and came up with an amazing observation and a highly non-trivial extension of the theory.

First, Moser realized that in a geometric guise an essential part of the Aubry–Mather work was contained in early work of Hedlund [**23**] which slightly predates even the Denjoy paper. Aubry–Mather theory has a close counterpart in the context of geodesics for an arbitrary metric on the two-dimensional torus. The geodesics in questions are *globally minimal*, i.e. their lifts to the universal cover are the shortest curves between any two points on a geodesic. Of course, these geodesics are of great geometric significance. Following Moser's suggestion V. Bangert developed and presented Hedlund's work in modern language [**8**], showing that for any irrational direction on the plane there is either a unique foliation of the torus $\mathbb{T}^2$ into minimal geodesics with this asymptotic direction or a *lamination* of a part of the torus into such geodesics. A lamination is a closed set consisting of geodesics which transversally has a Cantor-like structure.

As we pointed out, Aubry–Mather theory does not fully extend to higher dimensions, and, in fact, Hedlund addressed this question by displaying metrics on the three-dimensional torus that have only few minimal geodesics. Moser, however, found a highly non-trivial multidimensional generalization of minimal geodesics on the two-torus by considering *minimal hypersurfaces* on tori of higher dimension. The problem of the existence of such hypersurfaces is not a problem of dynamics or differential equations any more but the general structure of the variational approach can be carried out because the desired objects have co-dimension one. In goes without saying that the analytic aspects of the problem are far more formidable than in the geodesics problem. Moser carried out his brilliant insight and showed that the Aubry–Mather theory indeed has a counterpart: there are always either foliations or laminations of $\mathbb{T}^n$ with any Riemannian metric by $(n-1)$-dimensional minimal hypersurfaces [**67, 69**].

The relationship between KAM theory and Aubry–Mather theory is quite mysterious; it is one of the big riddles of modern dynamics. Aubry–Mather sets, Hedlund–Bangert minimal geodesics, and Moser's laminations are all constructed by global variational methods and there is a striking dichotomy: for a small perturbation of an integrable twist map or a flat metric on the torus sometimes (as for Diophantine values of the appropriate parameter) the solutions of the variational problem are regular smooth objects such as smooth invariant curves, geodesic foliations, and foliations by minimal surfaces [**66**]; other times they are quite singular. While there are effective sufficient criteria for guaranteeing that the solutions of a variational problem are not smooth there are no such criteria guaranteeing smoothness. In other words, even basic KAM-type results have not been obtained by variational methods. Moser provided insights into this extremely difficult problem in his last works related to Aubry–Mather theory [**71, 72**] by showing how Aubry–Mather sets can be approximated by smooth solutions of certain variational problems.

6. *Revival of completely integrable systems*

Solving the equations of motion of a mechanical system completely in terms of elementary functions is the ultimate success one can strive for, but experience quickly showed that one can rarely expect this much. Hamilton was the first to make an observation that was to lead to anything like a method. He found that integrating the Lagrangian of a mechanical system (along solutions) gives a function S that satisfies the partial differential equation

$$H\left(\frac{\partial S}{\partial q}, q\right) = \text{constant},$$

which is called the Hamilton–Jacobi equation. Jacobi's name is attached to it because he realized that a solution of the differential equation gives rise to a coordinate change (as a generating function) that makes the mechanical problem trivial, by rendering the Hamiltonian independent of p. This then yields explicit solutions in terms of elementary functions (and their integrals, and inverse functions), i.e. by quadrature.

Solving a partial differential equation is not a simple matter, so the approach taken by Jacobi was largely to look for applications in which a given coordinate transformation could be useful. Among his famous successes is the geodesic flow on an ellipsoid, which he found to be integrable. Others picked up on his technique (e.g. [**75**]) and a small collection of integrable systems resulted.

The theorem of Liouville (or the Liouville–Arnold–Jost Theorem) established that this is the way all integrable systems arise: if a Hamiltonian system with n degrees of freedom has n integrals in involution then there is a canonical coordinate change (to action-angle variables) in which compact joint level sets are standard tori on which the flow is linear. This is a precursor of an important modern result by Noether: every constant of motion corresponds to a one-parameter group of symmetries of the equations of motion. Thus, partial or complete integrability of a mechanical system is directly connected with symmetries of the system, many of which may not otherwise be readily apparent. Physicists refer to these as 'hidden symmetries'.

Theorems such as those by Liouville and Noether are milestones, but in terms of the discovery of integrable systems there was only intermittent progress (by Jacobi, Neumann and Kovalevskaya), which served to underscore the rarity of complete integrability, and even this ended at the time of Poincaré's discovery of tangles and orbit complexity, which laid to rest all hopes for an abundance of integrable systems. And thus this field went dormant for some seven decades. New motivation and new techniques needed to emerge. As they did, Moser made substantial contributions. Being well versed in classical results and methods, and quickly picking up the new techniques and the connection to partial differential equations, he identified new integrable systems and noted their relation to partial differential equations, but also emphasized the rarity of integrable systems by relating the new ones to classically known integrable problems.

In 1955, at about the time of Kolmogorov's insight, Fermi, Pasta, and Ulam conducted the first ever numerical simulation of a dynamical system [**18**], and their oscillator lattice appeared to exhibit quasiperiodicity. In retrospect this is a hint at later developments that revived interest in integrable systems. The continuum limit of their model is the Korteweg–de Vries (KdV) equation which was later recognized to be an integrable infinite-dimensional Hamiltonian system, i.e. Fermi, Pasta, and Ulam were studying a perturbation of an integrable system, which according to KAM theory should exhibit quasiperiodicity. What makes the KdV equation interesting is that while it is nonlinear it has special solutions (solitons) whose superpositions are also solutions (up to small error and a phase shift), that is, the nonlinearity notwithstanding, some analogies to linear phenomena are present.

It turned out that quasiperiodicity and superposition are due to integrability of the partial differential equation written as a Hamiltonian system. Indeed, Kruskal and others had found that the KdV equation has infinitely many constants of motion [**19**]. Peter Lax, with whom Moser overlapped at the Courant Institute, studied integrability of partial differential equations such as the KdV equation in pivotal work [**41**] that introduced *Lax pairs* as a means to find integrals. The general scheme of these is as follows. Consider a partial differential equation for an unknown function q, and for square matrices A, B depending on q and q_x consider the linear system $E_x = EA$, $E_t = EB$. This is called a Lax pair for the original partial differential equation if the compatibility condition $A_t - B_x = [A, B]$, which arises from $E_{xt} = E_{tx}$ and is hence necessary for the existence of a solution E, translates into the original partial differential equation. This, in turn, can be re-expressed in terms of the operator $L(t) := (\partial/\partial x) + A$ as $\partial L/\partial t = [L, B]$, which implies that $L(t)$ is isospectral because $L(t)$ is conjugate to $L(0)$ for all t. The interest in this association lies in the fact that the eigenvalues of the linear operator are integrals, i.e. constants of motion. Thus, Lax pairs provide for many integrals. This subject is also interesting because of the interest in isospectral families, specifically for the Schrödinger operator $L := -d^2/dx^2 + q$. The spectral problem is to determine the spectrum from the potential q. If q has compact support or decays sufficiently rapidly at infinity (Bargmann potential) the spectrum consists of points on the negative real axis and continuous spectrum on the positive real axis. This corresponds to bound states of an electron in a well and the free (positive energy) case, and this spectrum corresponds directly to experimental observations of light spectra due to excitation of atoms. For periodic potentials (corresponding to a crystal) one gets a *band spectrum* of finitely many compact intervals plus a half-infinite interval at $+\infty$ (conduction bands). These were well known, and Moser contributed to the study of quasiperiodic potentials, on which we elaborate below. The inverse spectral problem is to decide which potentials give rise to a given spectrum (what inference can an experimenter draw about the potential from spectroscopic observations). Spatial translations of a potential clearly give rise to the same spectrum, but the family of potentials that do so is rather larger.

The Lax approach bore fruit in 1974 when the KdV equation was integrated in terms of action-angle variables [**20**] (in fact, the point spectrum corresponds to solitons for the KdV equation). In the same year Moser made the first of several significant contributions to this subject. While it is clear how to use Lax pairs, it requires insight and ingenuity to find them. Moser, who possessed both, established integrability of numerous important systems and found some deep connections between various different such systems, both classical and modern. He first proved integrability of three finite-dimensional Hamiltonian systems: finitely many particles on the line or circle with inverse-cube interaction forces, and a system related to a discretization of the KdV equation. [**60**] describes these ideas most clearly by considering the Toda lattice, where the potential is exponential. In this case integrability is due to Hénon and Flaschka. [**59**] carries out the same approach for the Calogero lattice (inverse square interaction potential) and the Sutherland lattice (the $\sin^{-2}$ or $\sinh^{-2}$ interaction potential). Moser viewed integrable systems as rare beauties deserving of careful study, and in [**62**] (as well as the Fermi lectures [**64**] and [**63**]) Moser treats these three systems as special examples, but at the same time unifies them by showing that they are all related to the geodesic flow on an ellipsoid and to the Neumann system (a harmonic oscillator constrained to a sphere [**75**]), both classical integrable systems (known to be equivalent), as well as the KdV equation. Without diminishing the remarkable success of finding several new integrable systems, this established that upon closer inspection these systems are closely related to the classical ones. In that sense these accomplishments again reflect Moser's unique blend of taste, prowess, and erudition.

For quasiperiodic potentials Moser made inroads by defining a rotation number of a potential as follows [**31**]. If φ is an eigenfunction of the Schrödinger operator, i.e. $L\varphi = \lambda\varphi$, set $\rho(\lambda) := \lim_{x\to\infty}(1/x)\arg(\varphi'(x) + i\varphi(x))$ (this is π times the density of states) [**31**]. Extending to complex λ this turns out to be the imaginary part of a function w

(of λ and q), an analogue of the Floquet exponent. The functionals $w(\lambda, \cdot)$ commute with respect to a Poisson bracket commonly used in the study of the KdV equation. Indeed, if one expands $w(\lambda, q) = -\sqrt{-\lambda}\left(1+\sum w_i(q)/\lambda^i\right)$ then the w_i are in involution. By explicit calculation $16w_3$ is KdV, so here is a new way in which one obtains countably many integrals for the KdV equation.

The rotation number was put to use in [**64**], where Moser also shows that the problem of finding all almost periodic potentials for which the one-dimensional Schrödinger equation has a specified spectrum consisting of finitely many intervals is equivalent to finding the integrals for the geodesic flow on an ellipsoid. The connection is rather straightforward to describe when the spectrum is a finite band spectrum, i.e. it consists of finitely many intervals $I_1, \dots, I_d$. The set of corresponding potentials is a d-dimensional torus $\mathbb{T}^d$, and since translations of the potential leave the spectrum invariant we obtain a flow on T^d, which can be mapped to an integral surface of the Neumann problem. Since the latter is well understood, Moser was able to obtain formulas for finite-band potentials.

7. *Moser's contributions to hyperbolic dynamics*

Hyperbolic dynamics has a long and distinguished prehistory, including works of Hadamard, Perron, Hedlund, Hopf, and many others. The field in its modern form was born in the 1960s in two parallel developments centered at Berkeley ('the Smale school') and at Moscow (Anosov, Sinai, Alexeev, Arnold). From the beginning there was a considerable exchange of ideas between the two groups and mutual influence. Still, due to different backgrounds and motivation, the seminal work done in these two places exhibits striking differences both in form and substance. This is not a good place to tell the story in any detail so we restrict ourselves to one, albeit crucial, aspect, the question of structural stability. A dynamical system is structurally stable if its *topological* orbit structure does not change when the system is perturbed in a *differentiable* way (for a precise definition see, for example, [**33**, §2.3]). The concept was introduced by Andronov and Pontryagin in the 1930s who applied it to flows on surfaces, which have little recurrence. One of the main impulses for the development of hyperbolic dynamics was the discovery by Smale that systems with complicated behavior may exhibit structural stability on sufficiently large invariant sets ('Smale horseshoe') [**87**].

A compact invariant set Λ of a diffeomorphism $f : M \to M$ is said to be hyperbolic if the tangent bundle splits invariantly into a direct (Whitney) sum of two subbundles that are expanded and contracted by the map: $T_\Lambda M = E^u \oplus E^s$ such that $\|Df^{-n}\restriction_{E^u}\| \le C\lambda^n$ and $\|Df^n\restriction_{E^s}\| \le C\lambda^n$ for some $\lambda < 1$ and all $n \in \mathbb{N}$. If the whole manifold M is a hyperbolic set then f is said to be an Anosov diffeomorphism. In [**2**] Anosov developed the principal technical tools of hyperbolic dynamics and proved the structural stability of Anosov systems. The central element of Anosov's proofs is the existence of invariant families of stable and unstable manifolds, i.e. unique integrability of the distributions E^s and E^u. (Smale's proof for the horseshoe uses symbolic dynamics.)

In [**56**] Moser gives an alternative proof of the Anosov result. As is often the case with Moser's work, the method is most significant. He realized that the conjugacy could be obtained directly as the fixed point of a contraction defined on a space on candidate conjugacies by an operator cleverly constructed to reflect the conjugacy equation and to include some linear approximation. In other words, he obtains the conjugacy by Picard iteration. The resulting proof is remarkably short and transparent. Moser's inspiration is to recast the stability of the dynamical system into the stability of the fixed point of a contraction on a functional space with respect to perturbations of the contraction.

Moser's proof modified by Mather [**87**, Appendix] made a great impact on the further development of hyperbolic dynamics by the Smale school. Reducing complicated global problems in finite dimensions, such as conjugacy or integrability of distributions, to local problems near a fixed point in infinite-dimensional spaces is a crucial element in the work of Robbin on structural stability, the Hirsch, Pugh, and Shub work on partially hyperbolic systems, and a number of other major developments. It also influenced the Russian school; see, for example, an account of the general Anosov Shadowing Theorem in [**33**, §18.1]. [**89**] shows how effective this approach is at obtaining the basic structural information about hyperbolic sets.

Hyperbolic dynamics is one of two central themes of [**57**]. This little book is remarkable in two respects. First, in it Moser paints a very broad picture demonstrating the coexistence of stable (invariant tori) and 'random' motions (special types of horseshoes) within the context of classical problems of celestial mechanics, in particular the three-body problem. Second, it has been quite influential in turning the attention of the mainstream mathematical community to the modern theory of dynamical systems (including ergodic theory), in particular to the achievements of the Russian school.

8. *Two early gems*

The short 1965 paper [**52**] is one of Moser's most influential and most often quoted works. Although it does not deal with dynamics directly, its impact on the development of the theory of dynamical systems has been considerable. Locally any two smooth volume forms ω_0, ω_1 on a manifold M are the same up to a smooth coordinate change. If the manifold is compact then the total volume is finite and an obvious necessary condition for existence of a globally defined diffeomorphism $f : M \to M$ such that $f_*\omega_1 = \omega_0$ is equality of total volumes: $\int_M \omega_0 = \int_M \omega_1$. It is natural to ask whether this condition is sufficient. Richard Palais put this question to Moser, and the solution (or rather two different solutions) did not only have the usual elegance, but involved new ideas which turned out to be applicable to a number of other problems. Using an approach first suggested by R. Thom in the context of singularity theory, Moser observed that if two forms are connected by a smooth path ω_t $(0 \le t \le 1)$ along which the total volume is fixed, then one can write and solve a time-dependent differential equation for a family of diffeomorphisms f_t, such that $(f_t)_*\omega_t = \omega_0$. In the case in question one can simply take

$$\omega_t = (1-t)\omega_0 + t\omega_1. \tag{5}$$

Since $\int_M \omega_0 = \int_M \omega_1$, one has $\omega_1 - \omega_0 = d\Theta$ for some $(n-1)$-form Θ by the Poincaré Lemma. Since ω_t is non-degenerate, there is a unique (smooth) vector field X_t such that $\omega_t \lrcorner X_t = \omega_t(X_t, \cdot, \dots, \cdot) = -\Theta$. This is the desired time-dependent system of differential equations. Since M is compact one can integrate X_t to get a one-parameter family of diffeomorphisms $\{f_t\}_{t\in[0,1]}$ such that $\dot{f}_t = X_t$ and $f_0 = \mathrm{Id}$. Then

$$\begin{aligned}\frac{d}{dt} {f_t}^* \omega_t &= {f_t}^*(\mathcal{L}_{X_t}\omega_t) + {f_t}^* \frac{d}{dt}\omega_t \\ &= {f_t}^* d(\omega_t \lrcorner X_t) + {f_t}^*(\omega_1 - \omega_0) = {f_t}^*(-d\Theta + \omega_1 - \omega_0) = 0,\end{aligned}$$

so $f_{1*}\omega_1 = f_{0*}\omega_0 = \Omega_0$, i.e. f_1 is the desired coordinate change.

In another version of the proof, Moser provided a localized version of deformation, where at each step the form changes only in a small neighborhood, and demonstrated that the resulting diffeomorphisms can also be taken locally supported. Moser also applied his method to closed non-degenerate 2-forms (symplectic structures) where the place of total volume is taken by the homology class. In this case the straight-line deformation (5) may destroy non-degeneracy. The question of when a deformation preserving both the homology class and non-degeneracy is possible is highly non-trivial and is related to subsequent major developments in symplectic topology (see, for example, [**27**]). Notice that the local version of this argument gives a very nice proof of the classical Darboux Theorem.

As an example of the fruitfulness of the Moser method, let us point out that both authors of the present article have had occasions to use this idea in their work. Hasselblatt used it to obtain local symplectic coordinates with rather special additional properties [**22**], and in the work of Anosov and Katok [**3**] a modified version of Moser's local result played an important role in the construction of an iterative procedure, which can also be viewed as an 'ugly cousin' of the KAM method. It is also used in the classification of area-preserving flows on surfaces of higher genus [**33**, §14.7].

Finally, we turn to the earliest work of Moser we are going to discuss here. It is his seventh published paper [**49**]. In this paper Moser treats a classical problem, analytic normal forms for an area-preserving transformation near a hyperbolic fixed point, with a version of the classical method of Cauchy majorization. The formal part was known to Birkhoff. In the smooth category local normal forms with a given formal power series can usually be achieved by a smooth coordinate change, as was shown in great generality for hyperbolic maps by Sternberg and Chen several years after Moser's work. The smooth situation however is quite 'soft'; in particular, the centralizer of the normal form is huge. Moser, on the other hand, found essentially the only case with infinitely many resonances (due to preservation of area) with a fairly rigid normal form (in the analytic category). The map becomes a 'hyperbolic twist' and the local partition into invariant hyperbolas is canonically defined, giving an infinite-dimensional but tractable centralizer.

Acknowledgement. We would like to thank P. Rabinowitz and E. Zehnder for reading the draft of this paper and making a number of valuable suggestions.

References

[1] *Smooth Ergodic Theory and its Applications (Summer Research Institute Seattle, WA, 1999) (Proc. Symp. in Pure Mathematics, 69).* Eds. A. Katok, R. de la Lave, Y. Pesin and H. Weiss. American Mathematical Society, Providence, RI, 2001.

[2] D. V. Anosov. *Geodesic Flows on Riemannian Manifolds with Negative Curvature (Proc. Steklov Institute of Mathematics, 90)*. American Mathematical Society, Providence, RI, 1967.

[3] D. V. Anosov and A. B. Katok. New examples in smooth ergodic theory. Ergodic diffeomorphisms. *Trans. Moscow Math. Soc.* **23** (1970), 1–35.

[4] V. I. Arnold. Small denominators. I. Mapping the circle onto itself. *Izv. Akad. Nauk SSSR Ser. Mat.* **25** (1961), 21–86; Correction to V. Arnold's paper: 'Small denominators. I.'. *Izv. Akad. Nauk SSSR Ser. Mat.* **28** (1964), 479–480.

[5] V. I. Arnold. Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. *Usp. Mat. Nauk* **18**(5) (1963), 13–40.

[6] V. I. Arnold. Small denominators and problems of stability of motion in classical and celestial mechanics. *Usp. Mat. Nauk* **18**(6) (1963), 91–192.

[7] E. Artin. Ein mechanisches System mit quasiergodischen Bahnen. *Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität* **3** (1924), 170–175.

[8] V. Bangert. *Mather Sets for Twist Maps and Geodesics on Tori (Dynamics Reported, 1)*. Wiley, Chichester, 1988, pp. 1–56.

[9] J. Barrow-Green. *Poincaré and the Three Body Problem (American Mathematical Society/London Mathematical Society History of Mathematics, 11)*. American Mathematical Society, Providence, RI, 1997.

[10] G. D. Birkhoff. Proof of the ergodic theorem. *Proc. Natl. Acad. Sci. USA* **17** (1931), 656–660.

[11] G. D. Birkhoff. *Dynamical Systems (Colloquium Publications, 9)*. American Mathematical Society, Providence, RI, 1927

[12] G. D. Birkhoff. *Nouvelles recherches sur les systèmes dynamique*, 'Memoriae Pont. Acad. Sci. Novi Lyncaei, s. 3, Vol. 1, 1935, pp. 85–216' according to the collected works.

[13] M. L. Cartwright. Forced oscillations in nonlinear systems. *Contributions to the Theory of Nonlinear Oscillations (Annals of Mathematics Studies, 20)*. Princeton University Press, Princeton, NJ, 1950, pp. 149–241.

[14] M. L. Cartwright and J. E. Littlewood. On non-linear differential equations of the second order. I. The equation $\ddot{y} - k(1 - y^2)\dot{y} + y = b\lambda k \cos(\lambda t + a)$, k large. *J. London Math. Soc.* **20** (1945), 180–189.

[15] C. Conley. *Isolated Invariant Sets and The Morse Index (CBMS Regional Series in Mathematics, 38)*. American Mathematical Society, Providence, RI, 1978.

[16] C. Conley and E. Zehnder. Morse-type index theory for flows and periodic solutions for Hamiltonian equations. *Commun. Pure Appl. Math.* **37**(2) (1984), 207–253.

[17] P. M. M. Duhem. *La Théorie Physique: Son Objet, Sa Structure*. Marcel Rivière & Cie., Paris, 1906, 1914; Engl. Transl. *The Aim and Structure of Physical Theory*. Princeton Science Library, Princeton University Press, 1991.

[18] E. Fermi, J. Pasta and S. Ulam. Studies of nonlinear problems. I. *Los Alamos Report, LA-1940*, 1955. Reproduced in *Nonlinear Wave Motion (Proc. of the Summer Seminar, sponsored by the American Mathematical Society and the Society for Industrial and Applied Mathematics, Clarkson College of Technology, Potsdam, NY, 1972) (Lectures in Applied Mathematics, 15)*. Ed. A. C. Newell. American Mathematical Society, Providence, RI, 1974, pp. 143–156.

[19] C. S. Gardner, M. D. Kruskal and R. M. Miura. Korteweg–de Vries equation and generalizations. II. Existence of conservation laws and constants of motion. *J. Math. Phys.* **9** (1968), 1204–1209.

[20] C. S. Gardner, J. M. Greene, M. D. Kruskal and R. M. Miura. Korteweg–de Vries equation and generalization. VI. Methods for exact solution. *Commun. Pure Appl. Math.* **27** (1974), 97–133.

[21] J. S. Hadamard. Les surfaces à courbures opposées et leurs lignes géodesiques. *J. Math. Pures Appliquées* **4**(5) (1898), 27–73.

[22] B. Hasselblatt. Regularity of the Anosov splitting and of horospheric foliations. *Ergod. Th. & Dynam. Sys.* **14** (1994), 645–666.

[23] G. A. Hedlund. Geodesics on a two-dimensional Riemannian manifold with periodic coefficients. *Ann. Math.* **33** (1932), 719–739.

[24] G. A. Hedlund. Fuchsian groups and transitive horocycles. *Duke Math. J.* **2** (1936), 530–542.

[25] G. A. Hedlund. Fuchsian groups and mixtures. *Ann. Math.* (2) **40**(2) (1939), 370–383.

[26] M. R. Herman. Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. *Inst. Hautes Études Sci. Publ. Math.* **49** (1979), 5–233.

[27] H. Hofer and E. Zehnder. *Symplectic Invariants and Hamiltonian Dynamics*. Birkhäuser, 1994.

[28] E. Hopf. *Fuchsian Groups and Ergodic Theory (Transactions of the American Mathematical Society, 39)*. American Mathematical Society, Providence, RI, 1936, pp. 299–314.

[29] E. Hopf. *Ergodentheorie (Ergebnisse der Mathematik, 5)*. Springer, 1937.

[30] E. Hopf. Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung. *Berichte über die Verhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig, Mathematisch-Physikalische Klasse* **91** (1939), 261–304.

[31] R. Johnson and J. K. Moser. The rotation number for almost periodic potentials. *Comm. Math. Phys.* **84**(3) (1982), 403–438; Erratum: 'The rotation number for almost periodic potentials'. *Comm. Math. Phys.* **90**(2) (1983), 317–318.

[32] A. B. Katok and K. Burns. Infinitesimal Lyapunov functions, invariant cone families and stochastic properties of smooth dynamical systems. *Ergod. Th. & Dynam. Sys.* **14** (1994), 757–785.

[33] A. B. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, 1995.

[34] P. Koebe. Riemannsche Mannigfaltigkeiten und nicht euklidische Raumformen. *Sitzungsberichte der Preußischen Akademie der Wissenschaften*, *I*, (1927) 164–196; *II, III*, (1928) 345–442; *IV*, (1929) 414–557; *V, VI*, (1930) 304–364, 504–541; *VII*, (1931) 506–534.

[35] 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; Engl. Transl. *Stochastic Behavior in Classical and Quantum Hamiltonian Systems (Volta Memorial Conference, Como, 1977) (Lecture Notes in Physics, 93)*. Springer, 1979.

[36] A. N. Kolmogorov. Théorie générale des systèmes dynamiques et mécanique classique. *Proceedings of the International Congress of Mathematicians, Amsterdam, 1954, Vol. 1*. Noordhoff, Groningen; North-Holland, Amsterdam, 1957, pp. 315–333; Engl. Transl. R. Abraham. *Foundations of Mechanics*. Benjamin, 1967, Appendix D.

[37] N. Kryloff and N. Bogoliouboff. La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire. *Ann. Math.* **38**(1) (1937), 65–113.

[38] P. Simon marquis de Laplace. *Traité de mécanique céleste*, 1799–1825; Engl. Transl. N. Bowditch. Laplace's Mécanique Céleste. Reprint Services Corp.

[39] P. Simon marquis de Laplace. *Philosophical Essay on Probabilities*. Translated from the 5th French edition of 1925 by A. I. Dale. Springer, New York, 1995.

[40] M. Levi and J. K. Moser. A Lagrangian proof of the invariant curve theorem for twist mappings. *Smooth Ergodic Theory and its Applications (Summer Research Institute Seattle, WA, 1999) (Proc. Symp. in Pure Mathematics, 69)*. Eds. A. Katok, R. de la Lave, Y. Pesin and H. Weiss. American Mathematical Society, Providence, RI, 2001, pp. 733–746.

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

[42] R. de la Llave. A tutorial on KAM theory. *Smooth Ergodic Theory and its Applications (Summer Research Institute Seattle, WA, 1999) (Proc. Symp. in Pure Mathematics, 69)*. Eds. A. Katok, R. de la Lave, Y. Pesin and H. Weiss. American Mathematical Society, Providence, RI, 2001, pp. 175–292.

[43] F. Löbell. Über die geodätischen Linien der Clifford–Kleinschen Flächen. *Math. Z.* **30** (1929), 572–607.

[44] J. N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.* **207** (1991), 169–207.

[45] J. N. Mather, H. P. McKean, L. Nirenberg and P. H. Rabinowitz. Jürgen K. Moser (1928–1999). *Notices Amer. Math. Soc.* **47**(11) (2000), 1392–1405.

[46] J. N. Mather and G. Forni. Action minimizing orbits in Hamiltonian systems. *Transition to Chaos in Classical and Quantum Mechanics (Montecatini Terme, 1991) (Lecture Notes in Mathematics, 1589)*. Springer, Berlin, 1994, pp. 92–186.

[47] M. Morse. Recurrent geodesics on a surface of negative curvature. *Trans. Amer. Math. Soc.* **22** (1921), 84–100.

[48] M. Morse. A fundamental class of geodesics on any closed surface of genus greater than one. *Trans. Amer. Math. Soc.* **26** (1924), 25–61.

[49] J. K. Moser. The analytic invariants of an area-preserving mapping near a hyperbolic fixed point. *Commun. Pure Appl. Math.* **9** (1956), 673–692.

[50] J. K. Moser. A new technique for the construction of solutions of nonlinear differential equations. *Proc. Natl. Acad. Sci. USA* **47** (1961), 1824–1831.

[51] J. K. Moser. On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen Math.-Phys.* Kl. II **1962** (1962), 1–20; Remark on the paper: 'On invariant curves of area-preserving mappings of an annulus'. *Regul. Chaotic Dyn.* **6**(3) (2001), 337–338.

[52] J. K. Moser. On the volume elements on a manifold. *Trans. Amer. Math. Soc.* **120** (1965), 286–294.

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

[54] J. K. Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.* **169** (1967), 136–176.

[55] J. K. Moser. *Lectures on Hamiltonian Systems (Mem. Amer. Math. Soc., 81)*. American Mathematical Society, Providence, RI, 1968.

[56] J. K. Moser. On a theorem of Anosov. *J. Diff. Eqns* **5** (1969), 411–440.

[57] J. K. Moser. *Stable and Random Motions in Dynamical Systems (Ann. Math. Stud., 77)*. Princeton University Press, Princeton, NJ, 1973; Reprint, Princeton University Press, Princeton, NJ, 2001.

[58] J. K. Moser. A lemma on hyperbolic fixed points of diffeomorphisms. *Usp. Mat. Nauk* **29**(2) (1974), 228–232; Engl. Transl. *Russian Mathematical Surveys* **29**(2) (1974), 236–241.

[59] J. K. Moser. Three integrable Hamiltonian systems connected with isospectral deformations. *Surveys in Applied Mathematics (Proc. First Los Alamos Symp. Math. in Natural Sci., Los Alamos, NM, 1974)*. Academic Press, New York, 1976, pp. 235–258.

[60] J. K. Moser. 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, 1974) (Lecture Notes in Physics, 38)*. Springer, Berlin, 1975, pp. 467–497.

[61] J. K. Moser. Stable and unstable motion in dynamical systems. *Nonlinear Dynamics and the Beam-beam Interaction (Symp., Brookhaven Nat. Lab., New York, 1979)*. American Institute of Physics, New York, 1980, pp. 222–235.

[62] J. K. Moser. *Geometry of Quadrics and Spectral Theory (The Chern Symposium, 1979)*. Springer, 1980, pp. 147–188.

[63] J. K. Moser. Various aspects of integrable Hamiltonian systems. *Dynamical Systems (Bressanone, 1978)*. Liguori, Naples, 1980, pp. 137–195; also published in *Dynamical Systems (C.I.M.E. Summer School, Bressanone, 1978) (Progress in Mathematics, 8)*. Birkhäuser, Boston, 1980, pp. 233–289.

[64] J. K. Moser. *Integrable Hamiltonian Systems and Spectral Theory*. Lezioni Fermiane, Scuola Normale Superiore, Pisa, 1983.

[65] J. K. Moser. Break-down of stability. *Nonlinear Dynamics Aspects of Particle Accelerators (Santa Margherita di Pula, 1985) (Lecture Notes in Physics, 247)*. Springer, Berlin, 1986, pp. 492–518.

[66] J. K. Moser. A stability theorem for minimal foliations on a torus. *Ergod. Th. & Dynam. Sys.* **8*** (1988), Charles Conley Memorial Issue, 251–281.

[67] J. K. Moser. Minimal solutions of a variational problem on a torus. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **3**(3) (1986), 229–272.

[68] J. K. Moser. On commuting circle mappings and simultaneous Diophantine approximations. *Math. Z.* **205**(1) (1990), 105–121.

[69] J. K. Moser. Minimal foliations on a torus. *Lecture Notes in Mathematics, vol. 1365*. Springer, 1989, pp. 62–69.

[70] J. K. Moser and A. P. Veselov. Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Commun. Math. Phys.* **139**(2) (1991), 217–243.

[71] J. K. Moser. An unusual variational problem connected with Mather's theory for monotone twist mappings. *Seminar on Dynamical Systems (St. Petersburg, 1991)*. Birkhäuser, Basel, 1994, pp. 81–89.

[72] J. K. Moser. Smooth approximation of Mather sets of monotone twist mappings. *Commun. Pure Appl. Math.* **47**(5) (1994), 625–652.

[73] J. K. Moser. Old and new application of KAM theory. *Hamiltonian Systems with Three or More Degrees of Freedom (NATO Adv, Sci. Inst. Ser. Math Phys. Sci., 533)*. Kluwer, 1999, pp. 184–192.

[74] J. K. Moser. Recollections. Concerning the early development of KAM theory. *The Arnoldfest (Toronto, ON, 1997)*. American Mathematical Society, Providence, RI, 1999, pp. 19–21.

[75] C. G. Neumann. De problemate quodam mechanico, quod ad primam integralium ultraellipticorum classem revocatur. *J. Reine Angew. Math.* **56** (1859), 46–63.

[76] J. von Neumann. Proof of the quasi-ergodic hypothesis. *Proc. Natl. Acad. Sci. USA* **18** (1932), 70–82.

[77] J. von Neumann. Zur Operatorenmethode in der klassischen Mechanik. *Ann. Math.* **33** (1932), 587–642.

[78] I. Newton. *Philosophiæ Naturalis Principia Mathematica*, 3rd edn. Will & John Innys, London, 1729; Translation and Guide: *The Principia, Mathematical Principles of Natural Philosophy*. University of California Press, 1999.

[79] J. H. Poincaré. Sur le problème des trois corps et les equations de la dynamique. *Acta Mathematica* **13** (1890), 1–270.

[80] J. H. Poincaré. *Les méthodes nouvelles de la mécanique céleste*. Paris, 1892–1899; Engl. Transl. *New Methods of Celestial Mechanics (History of Modern Physics and Astronomy, 13)*. Ed. D. Goroff. American Institute of Physics, New York, 1993.

[81] B. van der Pol and J. van der Mark. Frequency demultiplication. *Nature* **120** (1927), 363–364.

[82] P. H. Rabinowitz. Periodic solutions of Hamiltonian systems. *Commun. Pure Appl. Math.* **31** (1978), 157–184.

[83] D. Salamon and E. Zehnder. KAM theory in configuration space. *Comment. Math. Helv.* **64** (1989), 84–132.

[84] H. Seifert. Periodische BEWEGUNGEN mechanischer Systeme. *Math. Z.* **51** (1948), 197–216,

[85] C. L. Siegel. Iteration of analytic functions. *Ann Math.* (2) **43** (1942), 607–612.

[86] C. L. Siegel and J. K. Moser. *Lectures on Celestial Mechanics*. Springer, New York, 1971; Reprint of the 1971 translation. Springer, Berlin, 1995.

[87] S. Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.* **73** (1967), 747–817.

[88] J.-C. Yoccoz. Conjugaison différentiable des difféomorphismes du cercle dont le nombre de rotation vérifie une condition diophantienne. *Ann. Sci. École Norm. Sup.* (4) **17**(3) (1984), 333–359.

[89] J.-C. Yoccoz. Introduction to hyperbolic dynamics. *Real and Complex Dynamical Systems (Proceedings of the NATO Advanced Study Institute held in Hillerød, June 20–July 2, 1993) (NATO Advanced Science Institutes Series C: Mathematical and Physical Sciences, 464)*. Eds. B. Branner and P. Hjorth. Kluwer, Dordrecht, 1995, pp. 265–291.

[90] E. Zehnder. Generalized implicit function theorems with applications to small divisors problems. I. *Commun. Pure Appl. Math.* **28** (1975), 91–140.

[91] E. Zehnder. Cantor-Medaille für Jürgen Moser. *Jahresber. Deutsch. Math.-Verein.* **95**(2) (1993), 85–94.