Vladimir I. Arnold

From superpositions to KAM theory

Foreword. V. I. Arnold (12 June 1937 – 3 June 2010) published several papers where he described, in the form of recollections, his two earliest research problems (superpositions of continuous functions and quasi-periodic motions in dynamical systems), the main results and their interrelations: [1], [2] (reprinted as [4]), and [3] (translated into English by the author as [5]). The first exposition [1] has never been translated into English; however, it contains many details absent in the subsequent articles. It seems therefore that publishing the English translation of the paper [1] would not be superfluous. What follows is this translation prepared by M. B. Sevryuk.

A related material is contained also in Arnold's recollections "On A. N. Kolmogorov". Slightly different versions of these reminiscences were published several times in Russian and English [6–10].

Kew words and phrases: Hilbert's 13th problem · Superpositions of continuous functions · Invariant tori carrying quasi-periodic motions · KAM theory and its applications · Kolmogorov as a supervisor

MSC (*2010*): 26B40 37J40 70H08 01A60 01A70 37N05

A. N. Kolmogorov proved his theorem on the persistence of invariant tori under a small analytic perturbation of a completely integrable Hamiltonian system in 1954.[1] He later related that he had been thinking about this problem for decades, starting from his childhood when he had read Flammarion's *Astronomy*,[2] but the success had come only after Stalin's death in 1953 when a new epoch had begun in the Russian life. The hopes this death raised had a deep impact on Kolmogorov, and the years 1953–1963 were one of the most productive periods in his life.

The initial point of the 1954 work on invariant tori was the mathematical practicum for sophomores of the Faculty for Mechanics and Mathematics of the Moscow State University. This practicum was introduced by Kolmogorov into the compulsory program at the time when computers were virtually unavailable in Russia. As one of the problems for the practicum, Kolmogorov proposed a study of integrable dynamical systems (geodesics on surfaces of revolution, the motion of a heavy particle on a horizontal torus, and so on). To his surprise, in all these integrable systems, one observed conditionally periodic motions along invariant tori in the phase space.

[1] *Translator's note:* A. N. Kolmogorov, On the persistence of conditionally periodic motions under a small change in the Hamilton function, *Doklady Akad. Nauk SSSR*, 1954, **98**, no. 4, 527–530 (in Russian). The English translation in: Stochastic Behavior in Classical and Quantum Hamiltonian Systems. Proceedings of the Volta Memorial Conference (Como, 1977). Edited by G. Casati and J. Ford. Lecture Notes in Physics, **93**. Springer, Berlin, 1979, 51–56 and in: Selected Works of A. N. Kolmogorov. Vol. 1. Mathematics and Mechanics. Edited by V. M. Tikhomirov. Kluwer, Dordrecht, 1991, 349–354.

[2] *Translator's note:* C. Flammarion's famous book *Astronomie populaire* (1880) was published in Russian several times under various titles. The edition of 1897 was probably the first one.

In trying to understand this phenomenon, Kolmogorov examined its abstract version: a dynamical system on the torus afforded by a field of divergence zero with respect to some volume element. In his note published in 1953 in *Doklady*,[3] Kolmogorov proved that such a "generic" vector field (i.e., a field satisfying certain Diophantine conditions on the mean frequencies that hold almost always) is equivalent to a standard (translationally invariant) field on the torus.[4] Such a field determines quasi-periodic motions ("conditionally periodic motions", as Kolmogorov used to say following old-fashioned terminology[5]). The system is ergodic (i.e., it does not possess nontrivial measurable invariant sets), but it does not mix up the particles of the phase space (the torus). The field flow carries over these particles preserving their shapes.

However, for some "exceptional" mean frequencies of revolution along the torus, Kolmogorov observed intermixing motions rather than quasi-periodic ones, even in the case of analytic fields. This intermixing is explained by non-uniformity of the motions along the orbits filling the torus in a quasi-periodic way, as parallel straight lines.

Immediately, the question arose whether such exceptional intermixing systems on tori were of real importance for studies of Hamiltonian dynamical systems.

In integrable systems, the motions along the tori are quasi-periodic. To discover real applications of the abstract theory of vector fields on the torus constructed by Kolmogorov, one had therefore to find invariant tori in non-integrable systems.

The simplest way to find such tori is to try to employ some variant of the perturbation theory for integrable systems. This was the course took by Kolmogorov and it was in this way that he arrived at his 1954 theorem on the persistence of invariant tori.

However, the initial goal here was by no means achieved. The motions along the perturbed invariant tori found by Kolmogorov are quasi-periodic. It is still unknown whether in the phase space of a typical nearly integrable Hamiltonian system, there are invariant tori carrying flows intermixing along those tori. Kolmogorov supposed that such tori do exist, so that the effect revealed in his 1953 note is observed in generic nearly integrable Hamiltonian systems.

It is interesting to remark that the "partial" success of Kolmogorov's 1954 work (which result is known nowadays as the KAM theorem[6]) is of much greater importance than the technical question of intermixing which Kolmogorov did not manage to answer. Kolmogorov's achievement was similar to that of Columbus whose attempt to find a Western route to India had

[3] *Translator's note:* A. N. Kolmogorov, On dynamical systems with an integral invariant on the torus, *Doklady Akad. Nauk SSSR*, 1953, **93**, no. 5, 763–766 (in Russian). The English translation in: Selected Works of A. N. Kolmogorov. Vol. 1. Mathematics and Mechanics. Edited by V. M. Tikhomirov. Kluwer, Dordrecht, 1991, 344–348.

[4] This theorem of Kolmogorov admits a natural multidimensional generalization in the theory of polyintegrable systems (*Algebra i Analiz*, 1992, **4**, no. 6, 54–62). [*Translator's note:* The English translation: V. I. Arnold, Polyintegrable flows, *St. Petersburg Math. J.*, 1993, **4**, no. 6, 1103–1110.]

[5] *Translator's note:* In fact, the term "conditionally periodic motions" is still in use nowadays. One often defines a quasi-periodic motion to be a conditionally periodic motion with incommensurable frequencies.

[6] In the American literature of the sixties, one can find papers with proofs of the "analytic counterpart of Moser's theorem" (which is of course Kolmogorov's original theorem). J. Moser never supported these attempts to attribute Kolmogorov's theorem to him.

failed.

Several years later, Kolmogorov delivered a course of lectures on the theory of dynamical systems at the Faculty for Mechanics and Mathematics of the Moscow State University. I was then a student and I heard his exposition of the 1953 and 1954 papers. There were no formal proofs in the course (that was a characteristic feature of Kolmogorov's lectures in general), but all the ideas were presented quite clearly, and I had no doubts that the proofs Kolmogorov had in view were correct. I still have no such doubts (although Kolmogorov never wrote down his proofs). In my opinion, the text of the 1954 note in *Doklady* enables one to restore the missing details of the proofs. Kolmogorov held that life is short and that it was better for him to devote the time remained to discoveries of new results.

I entered the Faculty for Mechanics and Mathematics of the Moscow State University in 1954 (before Stalin's death in 1953 or after the invasion to Czechoslovakia in 1968, this would probably have been impossible for me because my mother was a Jew while my grandfather was shot dead in 1938 on the flagrantly false charge of espionage for England, Germany, Greece, and Japan[7]).

In 1955–1956, I was incited by Kolmogorov to work on Hilbert's 13th problem concerning the possibility of representing any continuous function in three real variables in the form of a finite superposition of continuous functions in two variables. Kolmogorov had just reduced this problem to the following question:

Is it possible to embed the universal tree in the Euclidean space in such a way that any continuous function on this tree can be represented as the sum of continuous functions, each depending on one coordinate only?

I managed to construct such an embedding of the universal tree in the three-dimensional Euclidean space. This result implies the affirmative answer to Hilbert's question: the desired superposition does exist. Kolmogorov told me that I had solved Hilbert's 13th problem,[8] and that I had to choose the next problem myself.

The work on Hilbert's problem was my first serious mathematical research. I was still a junior student entirely ignorant in almost all the remaining mathematics. I therefore decided to keep on representing functions in the form of the sums of summands depending on one of the coordinates only. It was natural to try to figure out what happens when the set where the functions to be represented in the form of the sum of functions in the coordinates are defined is not a tree.

[7]The family was first informed that he had been condemned to ten years of forced labor camps without the right of correspondence. [*Translator's note:* This was the standard euphemism used by the Soviet secret police in Stalin's era for a death sentence.] Later on, the KGB provided us twice with various false data on the cause, location, and moment of his death. [*Translator's note:* Such lies were also quite usual.] Finally, in the nineties—no earlier—we were given a rehabilitation certificate and acquainted with the corresponding documents which included a troika's sentence to be shot and information on its immediate execution. [*Translator's note:* Under Stalin, the so-called "troikas" ("troĭka" means "triple" or "triad" in Russian) were commissions of three persons who convicted people without trial. They were a very widespread instrument of extrajudicial repressions.]

[8]Nowadays, I hold that the question of a representation of algebraic functions in three variables by superpositions of algebraic functions in two variables would be more reasonable than Hilbert's own formulation.

The simplest non-tree curve is a circle. So I attempted to represent a function on a closed planar curve in the form of the sum of functions in the coordinates.

This problem immediately led me to a dynamical system defined on the curve.

I started with simplest (convex) smooth closed curves like an ellipse. Such a curve is intersected by the straight lines parallel to the coordinate axes at two points (which rarely coincide). The involution A interchanging the two intersection points with each of the lines $x = \text{const}$ and the involution B interchanging the two intersection points with each of the lines $y = \text{const}$ determine the orientation preserving diffeomorphism $T = AB : S^1 \to S^1$ of the curve to itself.

The solvability of the problem of representing a function $f : S^1 \to \mathbb{R}$ in the form $a(x) + b(y)$ depends on the properties of the dynamical system $T : S^1 \to S^1$. For instance, if P is a periodic point ($T^n P = P$), then the function f to be represented in the form indicated has to satisfy the relation

$$f(P) + f(TP) + \ldots + f(T^{n-1}P) = f(Q) + f(TQ) + \ldots + f(T^{n-1}Q),$$

where $Q = AP$. On the other hand, if the dynamical system T has no periodic points, then it is easy to construct a formal representation $f = a + b$ by choosing arbitrarily the value of a in one of the points. (If we know the sum f and one of the summands a or b in each successive point of the polygonal line P, Q, TP, TQ, T^2P, T^2Q, $\ldots$, we are able to compute the other summand.) However, the question of the continuity and smoothness of the functions a and b constructed this way is far from being so straightforward.

The dynamical system T given on the circle is characterized by the *rotation number* defined by Poincaré. This number can be either rational or irrational. In the rational ("resonant") case, the mapping T admits, as a rule, attracting and repelling periodic points. In the case of an irrational rotation number, the mapping T can be reduced to the rotation at an angle incommensurable with 2π by a homeomorphic coordinate change on the circle (according to the theorem of Poincaré and Denjoy).

I tried to prove that this homeomorphic change of the variable is a diffeomorphism (for typical rotation numbers satisfying the usual Diophantine conditions—the same conditions as those in Kolmogorov's theorem on invariant tori).

I considered the analytic category and supposed that the mapping is analytically close to a (rigid) rotation. This is the case if, for example, the original planar curve is close to an ellipse. Kolmogorov's method can be applied to this problem in the perturbation theory and enables one to prove the smoothness of the Denjoy homeomorphism.

It was really amazing that my attempt to become independent of my teacher and to solve my own problem avoiding problems raised by others had led me to a problem so close to another field of Kolmogorov's research. The mysterious interrelations between different branches of mathematics with seemingly no connections are still an enigma for me. A discovery of such interrelations is one of the greatest enjoyments provided by mathematics, and I have been lucky enough to feel such a delight several times in various mathematical contexts. By the way, the problem of representing functions on a curve by the sum of functions in the coordinates appeared again in the singularity theory where its solution (S. M. Voronin, J.-P. Dufour, 1981) turned out to be different in the infinitely smooth category and in the analytic one.

Having proven the smoothness of the Denjoy homeomorphism for typical small perturbations of a rotation, I started inquiring whether somebody had examined all these questions between A. Denjoy's work of 1932 and my work of 1958. Among others, I found C. L. Siegel's papers on the linearization of holomorphic mappings near fixed points. To be more precise, I first invented this problem myself (as a simplified model of the problem of circle mappings) and solved it by Kolmogorov's method. Only later on, I discovered Siegel's work who had obtained the same result by another method[9] in about 1940.

"We are in a good company," Kolmogorov told me when I let him know of my bibliographic findings. As far as I understand, he was aware of neither Siegel's works nor J. E. Littlewood's works on the exponential slowness of an increase in perturbations.

I tried to prove some more conjectures raised by myself and concerning analytic diffeomorphisms of a circle. The first conjecture was that the Denjoy homeomorphism (which conjugates an analytic mapping with a rotation) is smooth (analytic) under the usual Diophantine conditions on the rotation number.

The smoothness of the Denjoy homeomorphism means the existence of a smooth invariant measure for the original diffeomorphism. I constructed examples of analytic diffeomorphisms for which the invariant measure was singular (not absolutely continuous with respect to the Lebesgue measure). The rotation numbers of these diffeomorphisms are pathologically well approximable by rationals. Such unusual numbers form a set of measure zero.

It was M.-R. Herman who managed to prove (more than twenty years later) the conjecture on the smoothness of the invariant measure without the assumption that the mapping is close to a rotation.[10]

The second conjecture concerned the natural boundary of the complex neighborhood (of the real circle) where the Denjoy homeomorphism may be extended holomorphically. Such a neighborhood cannot contain periodic points of the original holomorphic mapping. The conjecture was that (say, for mappings of the form $x \mapsto x + a + b \sin x$, $|b| < 1$) there exist periodic points in a vicinity of any point of the boundary of the domain of holomorphy of the Denjoy mapping (and on the boundary itself).

Analogous conjectures were recently proven by J.-C. Yoccoz and R. Pérez-Marco for the particular case of rational mappings. Counterexamples to the original conjecture were also constructed but only for exceptional rotation numbers (which form a set of measure zero). As far as I know, the question whether the initial conjecture holds for non-rational mappings (similar to the one pointed out above) remains open.

[9] After several recent works by L. H. Eliasson, E. Trubowitz, and others, it has become clear that Siegel's method is close to diagrammatic methods of the quantum field perturbation theory.

[10] *Translator's note:* In fact, Herman proved this conjecture in 1976, slightly less than two decades after Arnold's work, see M.-R. Herman, Conjugaison C^∞ des difféomorphismes du cercle pour presque tout nombre de rotation, *C. R. Acad. Sci. Paris, Sér. A–B*, 1976, **283**, no. 8, Aii, A579–A582.

In 1958–1959, I discussed these works on mappings of a circle onto itself with many mathematicians. I. M. Gelfand explained to me the relations between mappings of a circle and his investigations (joint with M. L. Tsetlin) of cardiac arrhythmias. S. L. Sobolev and two members of his school, N. N. Vakhaniya and R. A. Aleksandryan, let me know of the fact that these problems were close to some questions in the spectral theory of differential operators. Those questions had arisen in connection with Sobolev's works on liquid vibrations in rotating rockets. These works by Sobolev (carried out in 1943) had just been declassified (they were published in 1960). Professor A. L. Goldenveĭzer[11] explained to me the relation between the Dirichlet problem for the wave equation (I had started examining the mappings of a circle onto itself just for this problem) and studies of the stability of thin shells of hyperbolic curvature. Within the latter problem, the resonances correspond to a special instability of the shells distending along the polygons formed by the asymptotic lines of the shell surface (these lines are pairwise connected at the vertices located at the boundary of this surface).

All these applications were discussed in my diploma work defended at the Department for Function Theory and Functional Analysis of the Faculty for Mechanics and Mathematics of the Moscow State University in April 1959.

The same year, this work was submitted to *Izvestiya Akad. Nauk SSSR, Ser. Matem.* However, Kolmogorov advised me to exclude two subsections from the work. One of them was about heartbeats and the other one, about the effect of a small noise on the invariant measure.

"The heartbeat theory is an interesting application," he said, "but this is not one of the classic problems a mathematician should work on." "You would do better," he added, "to start developing applications of these ideas to celestial mechanics, to the motion of a rigid body, and to such classic problems as the stability problem for the Solar System."

The subsection about the effect of a small noise on the invariant measure dealt with the asymptotics of the solutions of the Fokker–Planck equation (and of its discrete analogues) as the diffusion coefficient tends to zero and the time tends to infinity. Nowadays, these asymptotics are included in the general scheme of the Morse–Witten theory (where the case of discrete time seems to have not been considered yet, however). Naturally, Kolmogorov did not approve my amateurish intrusion into the classic branch of the theory of stochastic processes that belonged to him by right.

This was the only occasion where Kolmogorov interfered in my studies. Although he was my scientific advisor, after my work on Hilbert's problem Andreĭ Nikolaevich gave me complete freedom to study whatever I wanted. He even told me that he would regard all his advices as harmful for me, especially those concerning the choice of a further research direction. However, later on A. N. advised me to include J. W. Milnor's works on differentiable structures on the spheres in my post-graduate curriculum which was extremely helpful for me (and led me to learning some mathematics from S. P. Novikov, D. B. Fuks,[12] and V. A. Rokhlin).

[11] *Translator's note:* In the Western literature, this name is very often spelt "Goldenweiser".
[12] *Translator's note:* In the Western literature, this name is usually spelt "Fuchs".

So, I shortened my diploma work, and its abridged version was published two years later (*Izvestiya Akad. Nauk SSSR, Ser. Matem.*, 1961, **25**, no. 1, 21–86). Since the English translation appeared in 1965,[13] this work is usually referred to as "Arnold 1965". Later on, the subsection on heartbeats I had deleted was partially published in I. M. Gelfand's Collected Works (vol. 3).[14] By that moment, most of its content had been rediscovered by the Canadian physiologist L. Glass (who had employed mathematical theorems in "Arnold 1965").[15]

One day, Kolmogorov invited two famous physicists, his old friend and collaborator M. A. Leontovich[16] and L. A. Artsimovich from the Atomic Energy Institute of the USSR Academy of Sciences,[17] to present a talk at his 1958–1959 seminar on the dynamical systems theory at the Faculty for Mechanics and Mathematics of the Moscow State University. They reported on mathematical problems in the theory of particle confinement in reactors for controlled thermonuclear fusion.

One of the problems was to prove the existence of so-called magnetic surfaces in systems of toroidal geometry of the Tokamak type. Magnetic surfaces are surfaces filled with magnetic field lines. An example of such a surface had been pointed out by I. E. Tamm[18] in his manual on the electricity theory in the twenties.

Kolmogorov's theorem implies the persistence of the most of magnetic surfaces (enclosed in one another) in systems with "shear" (where the rotation number changes as one passes from a torus to another torus).

The persistence of magnetic surfaces was a more or less direct consequence of Kolmogorov's theorem but for the physicists at those times, all the contemporary theory of dynamical systems was new and unexpected, and it seemed improbable to them. It was written in the manual by L. D. Landau and E. M. Lifshits[19] that any Hamiltonian system is either completely integrable or ergodic on each level manifold of the first integrals known.[20] However, invariant tori (in integrable systems) were mentioned in Born's *Introduction to Atomic Mechanics*. In the Russian translation of 1934, they were called "trëkhizmeritel′nye raznovidnosti".[21]

---

[13] *Translator's note:* V. I. Arnold, Small denominators. I. On mappings of a circle onto itself, *Amer. Math. Soc. Transl., Ser.* 2, 1965, **46**, 213–284.

[14] *Translator's note:* V. I. Arnold, Cardiac arrhythmias and circle mappings. In: I. M. Gelfand. Collected Papers. Vol. III. Edited by S. G. Gindikin, V. W. Guillemin, A. A. Kirillov, B. Kostant, and S. Sternberg. Springer, Berlin, 1989, 1019–1024. Reprinted (with some corrections) in: *Chaos*, 1991, **1**, no. 1, 20–24.

[15] *Translator's note:* See, e.g., L. Glass, Cardiac arrhythmias and circle maps—A classical problem, *Chaos*, 1991, **1**, no. 1, 13–19 and references therein.

[16] It is interesting to note that in their joint work on the trajectories of Brownian motion [*Translator's note:* A. Kolmogoroff and M. Leontowitsch, Zur Berechnung der mittleren Brownschen Fläche, *Phys. Z. Sowjetunion*, 1933, **4**, H. 1, 1–13], the physical part was due to Kolmogorov and the mathematical one, to Leontovich.

[17] *Translator's note:* Now the National Research Center "Kurchatov Institute".

[18] Artsimovich, Leontovich, and Tamm were among the most active theoretical physicists of the Soviet nuclear project. Leontovich and Tamm were close friends of my father and helped my family to survive hard years after his death (I was then eleven years old).

[19] *Translator's note:* In the Western literature, this name is usually spelt "Lifshitz".

[20] *Translator's note:* In the later editions of *Mechanics* by Landau and Lifshits, this was corrected, of course.

[21] *Translator's note:* Here the author has in view M. Born's treatise *Vorlesungen über Atommechanik* ("Lectures on Atomic Mechanics"). In 1934, it was translated into Russian: M. Born, Lektsii po Atomnoĭ Mekhanike,

The second problem formulated by Leontovich and Artsimovich consisted in studying the behavior of adiabatic invariants, for instance, in magnetic mirror traps. In this theory, a charged particle moves fast along a spiral around a magnetic field line (under the Lorentz force). The Larmor radius (the distance from the magnetic field line) is small, and the adiabatic approximation predicts that it will remain small even after many reflections from the magnetic mirrors.

As a typical example here, one may consider the motion of a charged particle in the magnetic field of the Earth between the two magnetic poles. The vibrations as the Larmor center moves along a line of force with reflections near the poles are observed in the form of the polar aurorae (the northern and southern lights).

In the problems of particle confinement in a trap, one has to know what happens after many millions of reflections. The adiabatic approximation cannot answer this question, and new methods are required.

I started working on this question while not forgetting applications to celestial mechanics. While reading Poincaré's *New Methods of Celestial Mechanics* and discussing it with V. M. Alekseev, I realized that to apply Kolmogorov's method to the stability problem of planetary systems, one had to overcome several difficulties. Some of them are present already in the problem of adiabatic invariants.

Born called the main difficulty "proper degeneracy": some of the frequencies of the perturbed quasi-periodic motion vanish as the perturbation magnitude tends to zero.[22] This difficulty is encountered already in the problem of the perpetual adiabatic invariance for a system with a parameter undergoing a slow periodic variation. Here the ratio of the slow frequency of the parameter variation to the fast frequency of the basic motion is a small parameter of the problem.

In celestial mechanics, the main small parameter is the ratio of the masses of the planets to the mass of the Sun. This ratio is of the order of 10^{-3}. The Keplerian ellipses are fixed in the space provided that one does not take into account the mutual attraction of the planets. As one begins taking into account the mutual attraction of the planets, the ellipses start moving slowly. The angular frequencies of their slow rotation are small as soon as the perturbation is small. Consequently, in the next-to-Keplerian quasi-periodic approximation to the planetary motions, there are fast (Keplerian) frequencies as well as slow ("secular") frequencies that vanish as the perturbation magnitude tends to zero.

ONTI – Gos. Nauchn.-Tekhn. Izd. Ukrainy, Khar′kov–Kiev, 1934. "Trëkhizmeritel′nye raznovidnosti" is a pointless term meaning something like "three-measuring races". I have failed to find this term in the Russian translation of Born's book. Nevertheless, on p. 38, there are the words "dvukhizmeritel′nye raznovidnosti" (≈ "two-measuring races") for "zweidimensionale Mannigfaltigkeiten" ("two-dimensional manifolds"). Born discussed there any two-dimensional surfaces in the phase space, not necessarily invariant 2-tori. The correct Russian for "two-dimensional manifolds" is "dvumernye mnogoobraziya". The translation has many other flaws.

[22] *Translator's note:* In *Vorlesungen über Atommechanik*, Born contrasted "proper degeneracy" ("eigentliche Entartung") and "chance degeneracy" ("zufällige Entartung") of completely integrable systems. The latter referred to the situation where the unperturbed frequencies are commensurable for some particular values of the action variables. The term "chance degeneracy" is almost out of use nowadays.

Another peculiarity of the planetary problem is the so-called "limit degeneracy".[23] The eccentricities and mutual inclinations of the planetary orbits are small. The limit case of zero eccentricities and inclinations corresponds to a degeneracy of the first approximation motions described above. This kind of degeneracy is analogous to that of a family of toroidal magnetic surfaces that are enclosed in one another and shrink down to the central circle lying inside all the surfaces.

The simplest mathematical model with this sort of degeneracy is the problem (formulated by G. D. Birkhoff) on the stability of a fixed point of an area preserving mapping of the plane onto itself. I decided to start with this problem.

A solution of this problem of Birkhoff was completed in 1960 and published in *Doklady* in 1961 (*Doklady Akad. Nauk SSSR*, 1961, **137**, no. 2, 255–257).[24] This result is not so different from Kolmogorov's original theorem of 1954, but it provides a solution of an old classic problem and, at the same time, it is a necessary step towards exploring the problem of planetary motions.

After that, I turned to proper degeneracy. First, as a model problem, I considered the case of a non-Hamiltonian system where the frequency ratio was proportional to the perturbation parameter (*Doklady Akad. Nauk SSSR*, 1961, **138**, no. 1, 13–15).[25]

This case is no longer within the standard framework of the perturbation theory and Kolmogorov's method since the solution cannot be expanded in the Taylor series in the perturbation parameter.

Now it became possible to apply the techniques developed in the problem of adiabatic invariants. As soon as I accomplished that, Kolmogorov suggested that I should submit the paper on perpetual adiabatic invariance to *ZhÈTF*,[26] the main physical journal in the USSR.

A few weeks later, M. A. Leontovich (who was, as far as I remember, a deputy to the editor-in-chief of *ZhÈTF*[27]) invited me to his home (near the Atomic Energy Institute of the USSR Academy of Sciences) to discuss the manuscript. Having fed me, as usual, by boiled buckwheat and calling me, as usual, "Dimka"[28] (M. A. called me in such a way until his death), Mikhail Aleksandrovich explained to me that the paper could not be published in *ZhÈTF* due to the following reasons.

1. The manuscript contained the words "theorem" and "proof" forbidden in *ZhÈTF*.

2. The manuscript claimed that "A implies B" while every physicist knew examples showing that B does not imply A.

[23] *Translator's note:* "Grenzentartung" in Born's *Vorlesungen über Atommechanik*.

[24] *Translator's note:* The English translation: V. I. Arnold, On the stability of an equilibrium point of a Hamiltonian system of ordinary differential equations in the general elliptic case, *Soviet Math. Dokl.*, 1961, **2**, no. 2, 247–249.

[25] *Translator's note:* The English translation: V. I. Arnold, On the birth of a quasi-periodic motion out of a family of periodic motions, *Soviet Math. Dokl.*, 1961, **2**, no. 3, 501–503.

[26] *Translator's note: Zhurnal Èksperimental'noĭ i Teoreticheskoĭ Fiziki* (translated into English as *Journal of Experimental and Theoretical Physics*).

[27] *Translator's note:* Indeed, Leontovich was a deputy editor-in-chief of *ZhÈTF* in 1957–1981, from vol. **32**, no. 3 through vol. **80**, no. 6 (until his death).

[28] *Translator's note:* This is one of the (many) diminutive forms of "Vladimir".

3. The manuscript used the unintelligible terms "Lebesgue measure", "invariant tori", "Diophantine conditions".

Mikhail Aleksandrovich therefore proposed that I should rewrite the paper.

Now I realize how right he was in defending a physical journal from the Bourbaki-like mathematical jargon.

For instance, indeed, while claiming that "A implies B" the author must point out explicitly whether the converse holds, otherwise any reader not spoiled by the mathematical slang would understand the claim as "A is equivalent to B".

Nowadays, every physicist who studies Hamiltonian chaos or employs KAM theory in the problems of plasma confinement or of the particle accelerator theory freely uses the Lebesgue measure, invariant tori, and Diophantine conditions. But in 1961 one of the first papers on the theory now called KAM was, as we see, rejected by a leading physical journal for the use of these terms (and also of the words "theorem" and "proof").

As a result, I withdrew the article from *ZhÈTF*, and it appeared no earlier than a year later (*Doklady Akad. Nauk SSSR*, 1962, **142**, no. 4, 758–761).[29] The first long paper on KAM theory ("Small denominators. I" where in 1959, Kolmogorov's method had been described in detail for the first time) was also initially rejected by the editorial board of *Izvestiya Akad. Nauk SSSR, Ser. Matem.* but finally appeared in January 1961.

Mathematical connections of Russia with the West started restoring after Stalin's death due to the Khrushchëv "Thaw". Kolmogorov had been isolated from the Western colleagues for about 15 years (1938–1953) but at the beginning of the sixties, foreign mathematicians started visiting Moscow.

My first serious mathematical contact with a foreign mathematician was a meeting with S. Smale who came to Moscow in fall 1961. Among various interesting things Smale told me about *on the roof* of the skyscraper of the Moscow State University (in his reminiscences, he writes *on the steps*[30]), there was the information that J. Moser had started working on the problem of invariant tori. A few months later, Moser's remarkable paper in *Proc. Nat. Acad. Sci. USA* appeared,[31] where he combined Kolmogorov's method with J. Nash's smoothing and applied these ideas to the isometric imbedding problem of a Riemannian manifold into a Euclidean space.[32]

Later on, this powerful method enabled him to prove a Kolmogorov-type theorem for per-

[29] *Translator's note:* The English translation: V. I. Arnold, On the behavior of an adiabatic invariant under a slow periodic variation of the Hamilton function, *Soviet Math. Dokl.*, 1962, **3**, no. 1, 136–140.

[30] *Translator's note:* Here the author probably has in view the recollections: S. Smale, On the steps of Moscow University, *Math. Intelligencer*, 1984, **6**, no. 2, 21–27; in: From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, August 5–9, 1990). Edited by M. W. Hirsch, J. E. Marsden, and M. Shub. Springer, New York, 1993, 41–52. However, Smale recounts there an entirely different event, namely, his anti-war press conference on the steps of the Moscow State University on August 26, 1966 (the last day of the Moscow International Congress of Mathematicians where Smale was awarded a Fields Medal).

[31] *Translator's note:* J. Moser, A new technique for the construction of solutions of nonlinear differential equations, *Proc. Nat. Acad. Sci. U.S.A.*, 1961, **47**, no. 11, 1824–1831.

[32] *Translator's note:* To be more precise, of a 2-torus into $\mathbb{R}^5$.

turbing functions of finite smoothness. To my shame, I should confess that I have not read Moser's proofs and confined myself with applying his ideas on my own. Because of this, a strange technical inaccuracy in all the proofs by Moser (which was absent in my version of his theory) remained unnoticed for about 30 years until it was revealed by M. B. Sevryuk.[33]

Moser's results were amazing. Kolmogorov had supposed that even infinite smoothness of the perturbation is insufficient for the persistence of invariant tori. "Moser's achievement," he told me, "changes all our philosophy."

At the Stockholm International Congress of Mathematicians (in August 1962), I met J. Moser for the first time. He was an invited speaker and gave a talk about his solution of the Birkhoff problem on the stability of fixed points of area preserving mappings of a plane onto itself.

Instead of the analyticity requirement in my 1961 paper, Moser assumed the existence of 333 derivatives only (for subsequent 30 years, the order of the derivatives needed in the proofs reduced to three[34]). What was even more astonishing to me was that Moser replaced the Birkhoff nonresonance condition by the assumption of the absence of a finite number of the "strong" resonances only (of orders less than five). In this problem, the rational numbers with the denominators greater than four behave like irrational ones. In my 1961 work, I missed this circumstance although my proof (in the analytic category) could be carried over verbatim from the non-resonant case to the case of weak resonances (of orders greater than four).

In that article, instead of thinking of the nature of things, I tried to solve the "sports" problem raised by Birkhoff. In Birkhoff's formulation, all the resonances are forbidden. Being hypnotized by this fact, I did not observe that my proof provided more than a solution of Birkhoff's problem. This was a good lesson: one should never yield to the hypnotic influence of famous experts.

The trip to the Stockholm Congress of 1962 was my first visit abroad. My report was devoted to the stability problem for planetary systems, but the "panel" (the committee selecting the speakers) and the Program Committee did not regard planetary systems as deserving an invited lecture at a Mathematical Congress.

Fortunately, at those times, brief communications of uninvited speakers were heard at special sessions of 15-minute talks, so I gave my talk all the same.

The Organizing Committees of the International Congresses of Mathematicians have made great efforts to wipe out uninvited reports. The question whether this is good provokes disputes. The best (and the most important for me) talk I have ever heard at the Mathematical Congresses

[33] *Translator's note:* In fact, for less than 25 years. I detected this (minor) inaccuracy in 1985 (I was then Arnold's post-graduate student) in Moser's famous paper on invariant curves of planar mappings: J. Moser, On invariant curves of area-preserving mappings of an annulus, *Nachr. Akad. Wiss. Göttingen, Math.-Phys. Kl. II*, 1962, **1**, 1–20. Then I communicated my observation to Arnold and Moser. Subsequently, Moser published an erratum: J. Moser, Remark on the paper "On invariant curves of area-preserving mappings of an annulus", *Regul. Chaotic Dyn.*, 2001, **6**, no. 3, 337–338.

[34] *Translator's note:* In fact, for 25 years. The smoothness class C^3 (for annular mappings) was achieved in 1986: M.-R. Herman, Sur les courbes invariantes par les difféomorphismes de l'anneau, Vol. 2, *Astérisque*, **144**, Soc. Math. France, Paris, 1986.

was F. Hirzebruch's lecture at the Moscow Congress of 1966 on E. Brieskorn's works concerning the connection between the singularity theory and the Milnor spheres. Hirzebruch was not an invited speaker. I hold that the harm from uninvited reports of little interest is smaller than the loss due to rejections of important and interesting reports. Galois would hardly be invited to a Congress. In 1992, V. A. Vasil′ev was not given an official permission to attend the sessions of the First European Congress of Mathematics in Paris although four invited speakers discussed his works in their talks.[35]

My paper on the stability problem of planetary systems appeared in 1962 (*Doklady Akad. Nauk SSSR*, 1962, **145**, no. 3, 487–490).[36] From the viewpoint of the techniques, this work is very complicated. It depends on a lemma on Diophantine approximations in an intricate situation with several small parameters (this is due to the presence of both proper and limit degeneracies in the problem at hand). A detailed proof was published in *Uspekhi Matem. Nauk*, 1963, **18**, no. 6, 91–192 ("Small denominators. III").[37] It is a combination of the ideas of the works on adiabatic invariants and on the Birkhoff stability problem.

Kolmogorov's sixtieth anniversary was celebrated in the Assembly Hall of the Moscow State University in April 1963. I gave a talk on his works on invariant tori. In this talk, I showed how these works could be used to prove the stability of fast rotation of a heavy non-symmetrical rigid body. This work "Small denominators. II" (*Uspekhi Matem. Nauk*, 1963, **18**, no. 5, 13–40)[38] contained the first detailed proof of Kolmogorov's theorem of 1954 on the persistence of invariant tori under a small analytic perturbation of a completely integrable Hamiltonian system (Kolmogorov's *method* had been published in detail already in 1961 in my 1959 work "Small denominators. I" on mappings of a circle).

Of my own achievements in the perturbation theory for non-integrable Hamiltonian systems, the main one was published in *Doklady* in 1964 (*Doklady Akad. Nauk SSSR*, 1964, **156**, no. 1, 9–12).[39]. This paper describes the universal mechanism of instability in Hamiltonian systems with many degrees of freedom. Later on, the physicists called this mechanism "Arnold diffusion".[40] This diffusion contradicted Kolmogorov's intuition; he thought that stability can take place in generic multidimensional systems as well despite the fact that in these cases, stability is not ensured by the existence of invariant tori.

[35] *Translator's note:* Nevertheless, he managed to attend.

[36] *Translator's note:* The English translation: V. I. Arnold, On the classical perturbation theory and the stability problem for planetary systems, *Soviet Math. Dokl.*, 1962, **3**, no. 4, 1008–1012.

[37] *Translator's note:* The English translation: V. I. Arnold, Small denominators and problems of stability of motion in classical and celestial mechanics, *Russian Math. Surveys*, 1963, **18**, no. 6, 85–191.

[38] *Translator's note:* The English translation: V. I. Arnold, Proof of a theorem by A. N. Kolmogorov on the persistence of conditionally periodic motions under a small change in the Hamilton function, *Russian Math. Surveys*, 1963, **18**, no. 5, 9–36.

[39] *Translator's note:* The English translation: V. I. Arnold, On the instability of dynamical systems with many degrees of freedom, *Soviet Math. Dokl.*, 1964, **5**, no. 3, 581–585.

[40] *Translator's note:* The term "Arnold diffusion" was proposed by B. V. Chirikov in 1969: B. V. Chirikov, Research in the theory of nonlinear resonance and stochasticity, Preprint of the Novosibirsk Institute for Nuclear Physics of the USSR Academy of Sciences, 1969, no. 267 (in Russian). The English translation: *CERN Transl.*, 1971, no. 71-40.

In systems with the phase space of small dimension, the invariant tori entrap the regions between them and thus ensure stability (for instance, in the Birkhoff problem). In the paper of 1964, I constructed an example of instability in a situation where the Kolmogorov tori are preserved. I supposed then (and I still do) that the "diffusion" mechanism described in that article works in generic systems. So, what is typical is the existence of trajectories connecting a vicinity of any invariant n-torus of a nearly integrable Hamiltonian system with a vicinity of any other n-torus on the same energy level hypersurface (if the dimension $2n-1$ of this hypersurface is no less than 5, that is, if $n > 2$). However, this has not been proven yet.[41]

I do not hold that one may speak of a KAM theorem. Kolmogorov proved correctly his theorem in 1954, and this is clearly seen from his note in *Doklady*. Most probably, it would be more rightful to speak of KAM *theory* constituted by many theorems.[42] In the present note, I have tried to describe the contribution of each of the three authors.

The bibliography below pertains to the foreword and has been added by the translator.

References

[1] Arnold VI (1997) From superpositions to KAM theory. In: Vladimir Igorevich Arnold. Selected–60. PHASIS, Moscow, pp. 727–740 (in Russian)

[2] Arnold VI (1999) From Hilbert's superposition problem to dynamical systems. In: The Arnoldfest. Proceedings of a Conference in Honour of V. I. Arnold for his Sixtieth Birthday (Toronto, ON, June 15–21, 1997). Edited by E. Bierstone, B. Khesin, A. Khovanskii, and J. E. Marsden. Fields Institute Communications, vol. 24. American Mathematical Society, Providence, RI, pp. 1–18

[3] Arnold VI (2003) From Hilbert's superposition problem to dynamical systems. In: Mathematical Events of the XXth Century. PHASIS, Moscow, pp. 19–51 (in Russian)

[4] Arnold VI (2004) From Hilbert's superposition problem to dynamical systems. Amer. Math. Monthly 111(7): 608–624

[5] Arnold VI (2006) From Hilbert's superposition problem to dynamical systems. In: Mathematical Events of the Twentieth Century. Edited by A. A. Bolibruch, Yu. S. Osipov, and Ya. G. Sinai. Springer, Berlin + PHASIS, Moscow, pp. 19–47

[6] Arnold VI (1993) On A. N. Kolmogorov. In: Kolmogorov in Recollections. Edited by A. N. Shiryaev. Nauka, Moscow, pp. 144–172 (in Russian)

[41] *Translator's note:* This problem is still open although many important partial results have been obtained by various authors.

[42] *Translator's note:* The term "KAM theory" was coined in 1968: F. M. Izraĭlev, B. V. Chirikov, Stochasticity of the simplest dynamical model with divided phase space, Preprint of the Novosibirsk Institute for Nuclear Physics of the USSR Academy of Sciences, 1968, no. 191 (in Russian).

[7] Arnold VI (1993) On A. N. Kolmogorov. In: Golden Years of Moscow Mathematics. Edited by S. Zdravkovska and P. L. Duren. History of Mathematics, vol. 6. American Mathematical Society, Providence, RI + London Mathematical Society, London, pp. 129–153

[8] Arnold VI (2000) On A. N. Kolmogorov. In: Kolmogorov in Perspective. History of Mathematics, vol. 20. American Mathematical Society, Providence, RI + London Mathematical Society, London, pp. 89–108

[9] Arnold VI (2006) On A. N. Kolmogorov. In: Kolmogorov in Recollections of his Students. Second edition. Edited by A. N. Shiryaev. MTsNMO, Moscow, pp. 34–53 (in Russian)

[10] Arnold VI (2007) On A. N. Kolmogorov. In: Golden Years of Moscow Mathematics. Second edition. Edited by S. Zdravkovska and P. L. Duren. History of Mathematics, vol. 6. American Mathematical Society, Providence, RI + London Mathematical Society, London, pp. 129–153