ISSN 1560-3547, Regular and Chaotic Dynamics, 2024.

Nineteen Fifty-Four: Kolmogorov's New "Metrical Approach" to Hamiltonian Dynamics

Luigi Chierchia[1*] and **Isabella Fascitiello**[2**]

[1]*Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, Largo San Leonardo Murialdo 1, 00146 Roma, Italy*

[2]*Dipartimento of Education, Università Roma Tre, 00185 Roma, Italy*

Received January 31, 2024; revised May 22, 2024; accepted May 27, 2024

Abstract—We review Kolmogorov's 1954 fundamental paper *On the persistence of conditionally periodic motions under a small change in the Hamilton function* (Dokl. akad. nauk SSSR, 1954, vol. **98**, pp. 527–530), both from the historical and the mathematical point of view. In particular, we discuss Theorem 2 (which deals with the measure in phase space of persistent tori), the proof of which is not discussed at all by Kolmogorov, notwithstanding its centrality in his program in classical mechanics.
In Appendix, an interview (May 28, 2021) to Ya. Sinai on Kolmogorov's legacy in classical mechanics is reported.

MSC2010 numbers: 01A60, 37J40, 37J05, 37J25, 70H08

DOI: 10.1134/S1560354724550021

Keywords: Kolmogorov's theorem on invariant tori, KAM theory, history of dynamical systems, small divisors, Hamiltonian systems, perturbation theory, symplectic transformations, nearly-integrable systems, measure of invariant tori

1. INTRODUCTION

Kolmogorov's 1954 paper *On the persistence of conditionally periodic motions under a small change in the Hamilton function* [5] is probably one of — if not, "the" — most influential contribution to the modern development of classical mechanics and dynamical systems: in four pages, it started the celebrated KAM theory, with precise statements and a clear outline of the main result. However, a complete discussion of this brief paper (four pages with a bibliography containing three items), both from a historical and a mathematical point of view, is still missing[1)].

The plan of the present paper is the following.

In Section 2 (*Historical remarks on Kolmogorov's influence on classical mechanics in the 20th century*), following [34], and [33], we discuss the circumstances of publication and diffusion of Kolmogorov's revolutionary research program in classical mechanics, as it emerges from his 1950s papers and his lecture at the 1954 International Congress of Mathematicians in Amsterdam. Such a program implied a deep conceptual change intimately related to a mathematical reformulation of classical mechanics, with impact also on celestial mechanics.

In Section 3 (*The theorems in Kolmogorov's 1954 paper*), we continue the mathematical discussion started in [26], considering, in particular, Theorem 2, where Kolmogorov states that

[*]E-mail: luigi.chierchia@uniroma3.it

[**]E-mail: isabella.fascitiello@uniroma3.it

[1)]For a "friendly introduction" to the history and mathematics of KAM theory, see [30]; for a short review of KAM theory and Kolmogorov's legacy, see also [24].

the Lebesgue measure of persistent invariant tori of analytic nearly-integrable Hamiltonian systems (with a nondegenerate integrable Hamiltonian and bounded phase space) tends to full measure, as the size of the perturbation goes to zero. Theorem 2 (unlike[2)] Theorem 1) is not followed by any mathematical details or technical comment[3)], and in Section 3 we propose a proof of it based on the scheme of proof of Theorem 1 given in [26].

2. HISTORICAL REMARKS ON KOLMOGOROV'S INFLUENCE ON CLASSICAL MECHANICS IN THE 20TH CENTURY

The International Congress of Mathematicians in Amsterdam, Netherlands, in 1954, ended with a lecture by Andrej Nikolaevich Kolmogorov (1903–1987). It was the second International Congress following the hiatus caused by political tensions and the Second World War, and the first to feature a Soviet delegation[4)]. The printed text of his lecture in the Proceedings of the Congress (published in 1957 [7]) starts with the following words:

> My aim is to elucidate ways of applying basic concepts and results in the modern general metrical and spectral theory of dynamical systems to the study of conservative dynamical systems in classical mechanics [7, 1991, p. 355].

Such a text (20 pages ca., depending on translation) is a literary piece aiming at the reader's engagement, with an echo of Kolmogorov's original oral style. The bibliography consists of 24 references to papers and monographs written by ca. 20 authors, published between 1917 and 1954 mainly in the Soviet Union but also in France, the USA and Germany.

It has to be noted that Kolmogorov's 1954 research program in dynamical systems and classical mechanics was elaborated starting from the early 1930s[5)]. In the new age marked by Stalin's death (March 5, 1953), Kolmogorov encouraged the implementation of his program as a collective endeavor:

> My papers on classical mechanics appeared under the influence of von Neumann's papers on the spectral theory of dynamical systems and, particularly under the influence of the Bogolyubov – Krylov paper of 1937.
>
> [...] To accumulate specific information we organized a seminar on the study of individual examples. My ideas concerning this topic and closely related problems aroused wide response among young mathematicians in Moscow [16, pp. 503–504].

Indeed, in the following years, in Moscow, Kolmogorov directed some of his students toward the field of dynamical systems, with special attention to classical mechanics and celestial mechanics.

The initial building blocks of the new mathematical landscape presented in the Amsterdam conference had been published by Kolmogorov in two short papers (4 pages each, [4] and [5]) in Russian in the proceedings of the Soviet Academy of Sciences (Doklady Akademii Nauk SSSR):

[2)] Theorem 1 is followed by a precise outline of its proof, without, however, including estimates and, in particular, without discussing the convergence of the Newton scheme; the missing analytical details have been discussed in [26], following closely Kolmogorov's outline.

[3)] Apart from a brief final remark; compare note 6 below.

[4)] Suspended in 1936, the International Congress was reinstated only in 1950 in Cambridge, Massachusetts (USA). However, in that occasion, the Russian academic community did not participate. In Volume I of the 1950 ICM proceedings, in the section entitled "Report of the Secretary", the following is stated [3, p. 122]:

> Shortly before the opening of the Congress, the following cable was received from the President of the Soviet Academy of Sciences: "The USSR Academy of Sciences appreciates having received a kind invitation for Soviet scientists to participate in the International Congress of Mathematicians to be held in Cambridge. Soviet mathematicians are very busy with their regular work, unable to attend the congress. I hope that the upcoming congress will be a significant event in mathematical science. Desire for success in congress activities. S. Vavilov, President, USSR Academy of Sciences."

For a historical account of International Congresses of Mathematicians, see [27] and [18].

Notice that Kolmogorov's last travel abroad dated back to the early 1930s: "en 1934 [...] quoique la fondation Rockefeller lui eût accordé une bourse, Kolmogorov ne fut pas autorisé à se rendre à Paris pour travailler prés d'Hadamard" in [28, p. 133].

[5)] For historiographical issues concerning the origins of Kolmogorov's new paradigm in dynamical systems, see [33].

the first dated November 13, 1953, and the second one August 31, 1954, nine days before the Amsterdam conference.

The two papers, with bare bibliographical references (respectively two and three items), consist of the statements and some hints at the proofs of three theorems, without conceptual framing.

Thus, we are faced with two distinct literary genres: on one side, the lecture/speech text [7], more literary and refined in its expressiveness, and, on the other, two concise typical twentieth-century research papers [4, 5].

From a historiographical perspective, Kolmogorov's theorems can be better understood in their cultural significance within the framework established by the research program presented at the Amsterdam lecture text. Conversely, the classical mechanics program is validated by the initial mathematical steps in the research papers, especially in the fundamental 1954 paper [5], which we now proceed to analyze in more detail.

The paper [5] contains a short introduction, including notation, and two theorems: Theorem 1 — followed by a precise outline of proof — which deals with the persistence, under perturbation, of a *single* invariant Diophantine Lagrangian torus, and Theorem 2 — stated at the end of the paper without any mathematical discussion, apart from a brief concluding remark[6)] — where Kolmogorov claims that, under a suitable nondegeneracy condition, the Lebesgue measure of persistent invariant tori in nearly-integrable analytic Hamiltonian systems tends to full measure as the size of the perturbation goes to zero.

Both theorems are considered in the Amsterdam lecture text [7], where Kolmogorov discusses explicitly the deep meaning he attaches to their significance:

> Theorems 1 and 2 in my paper[7)] [22] assert that in the above-described situation the only change in the entire pattern for small[8)] θ is that some of the tori corresponding to systems of frequencies for which the expression[9)] (n, λ) decreases too rapidly with increasing
>
> $$|n| = \sqrt{\sum n_\alpha^2}$$
>
> may disappear while the majority of the tori[10)] T_p^s, retaining the character of motions occurring on them, are somewhat displaced in[11)] Ω^{2s}, and still fill for small θ the region[12)] G to within a set of small measure. Thus, under small variations of H the dynamical system remains nontransitive and the region G continues to be decomposable, to within a residual set of small measure, into ergodic sets with discrete spectra (of the indicated specific nature) [7, 1991, p. 366].

The centrality of Theorem 2 for Kolmogorov appears particularly evident in the introduction of the conference text:

> In conservative systems, asymptotically stable motions are impossible.
> Therefore, for instance, the determination of individual periodic motions, however interesting it may be from the viewpoint of mathematics, has only a rather restricted real physical significance in the case of conservative systems. For conservative systems, the metrical

6) In the following citation M_θ denotes the set of all quasi-periodic trajectories, θ being the size of the perturbation, [3] refers to [4] in our bibliography:

> It seems that, in a sense, the "general case" is the case when the set M_θ has an everywhere dense complement for all positive θ. Complications of this kind appearing in the theory of analytic dynamical systems were indicated in my paper [3] in connection with a more specific situation (in [5, 1991, p. 354]).

7) He refers to [5].

8) In Kolmogorov's nomenclature, θ is the size of the perturbation.

9) (n, λ) denotes inner product and λ is the frequency.

10) T_p^s are invariant unperturbed s-dimensional tori, s being the number of degrees of freedom.

11) Ω^{2s} is the phase space.

12) G is a bounded region in Ω^{2s}, where H is defined.

> approach[13] is of basic importance making it possible to study properties of a major part of motions [7, 1991, p. 356].

These theorems gave rise to contrasting interpretations and doubts.

Jürgen Moser (1928–1999) questioned the validity of the proof of Theorem 1 already in his 1959 review published in Mathematical Reviews [9]:

> In the center of this talk is the author's new statement on the conservation of conditionally periodic solutions [...] (The proof of this theorem was published in Dokl. Akad. Nauk SSSR 98 (1954), 527–530 [MR0068687], but the convergence discussion does not seem convincing to the reviewer.)

Yakov Grigor'evich Sinai (b. 1935), within the volume *Kolmogorov in Perspective* [21], speaks about the proof of Theorem 1 in more detail:

> In the fall of 1957 I became a graduate student under Andrei Nikolaevich. At the same time he began a famous course of lectures on the theory of dynamical systems, which later was continued as a seminar. Much has already been written about this seminar. Among those present, besides us, were V. M. Alekseev, V. I. Arnol'd, L. D. Meshalkin, M. S. Pinsker, M. M. Postnikov, K. A. Sitnikov, and many others. The first part of the course definitely had a probabilistic bias [...]. Later in the course he presented the theorem[14] that was to become the basis for the famous KAM theory, together with a complete proof. In early 1958 Andrei Nikolaevich departed to spend half a year in France and left Meshalkin and me a program for preparation for the examination in classical mechanics, which included this proof [21, pp. 117–118].

Vladimir Igorevich Arnol'd (1937–2010), retrospectively, on the occasion of the 1997 Arnoldfest (Toronto, CA), states:

> Moser complained that a proof of the theorem in the case of analytic Hamiltonians was never published by Kolmogorov. I think that Kolmogorov was reluctant to write down the proof because he had other things to do in his remaining years of active work — which is a challenge when you are sixty. According to Moser, the first proof was published by Arnol'd. My opinion, however, is that Kolmogorov's theorem was proved by Kolmogorov [19, 2004, pp. 622–623].

On the other hand, Arnol'd might have considered the lack of discussion of the proof of Theorem 2 the main motivation for his celebrated paper entitled *"Proof of a theorem of A. N. Kolmogorov on the persistence of conditionally periodic motions under a small perturbation of the Hamiltonian"* [12]. In support of this comment, Sinai says[15]:

> There were some gaps in the estimates of the measure of invariant sets. That was the main point where Arnol'd complained about the proof by Kolmogorov. In Kolmogorov's paper, complete estimates of such a measure were not given.

In [12] Arnol'd provides a detailed proof of Kolmogorov's Theorem 2, using a quadratic convergent scheme different from that of Kolmogorov[16].

Paradoxically, Moser seemed to believe in[17] 1957–59 that Theorem 2 was a straightforward consequence of Theorem 1. Indeed, in the above-mentioned 1959 Mathematical Review [9], even though no direct reference to Theorem 2 is ever made, Moser claims:

> This very interesting theorem[18] would imply that for an analytic canonical system which is close to an integrable one, all solutions but a set of small measure lie on invariant tori.

---

[13] The title of our present paper derives from these words.

[14] Arguably, Sinai refers here to Theorem 1 in [5]; see, also, the interview in the Appendix.

[15] See the interview in Appendix.

[16] In Section 2 we shall provide, instead, a proof of Theorem 2 based on Kolmogorov's original quadratic scheme.

[17] Notice that the first contribution on small divisors by Moser is the famous 1962 paper [11] on area-preserving maps; (in 2001 an addendum [22] to this paper appeared). Note, however, that Moser's first article on the quadratic convergence method is [10] (1961).

[18] He refers to Theorem 1 in [5].

Regarding the relevance in a *longue durée* context of Kolmogorov's new impulse in dynamical systems — also in connection with George D. Birkhoff (1884–1944)'s work — Stephen Smale (b. 1930) wrote in his essay *The mathematics of time* [15]:

> It may be stated in conclusion that the outstanding unsolved problem in the ergodic theory is the question of the truth or falsity of metrical transitivity for general Hamiltonian systems. In other words, the Quasi-Ergodic Hypothesis has been replaced by its modern version: the Hypothesis of Metrical Transitivity.
> This hypothesis played an important role in Birkhoff's later work. He not only believed it but part of his work is written assuming that it is true.
> [...] These beliefs held sway in mathematical physics until Kolmogoroff's famous Amsterdam Congress paper in 1954 and subsequent work of Arnol'd and Moser in 1961–1962. The work of Kolmogoroff, Arnol'd, and Moser, KAM, showed that near "elliptic" closed orbits of a general Hamiltonian system on an energy surface, ergodicity failed. In that case there exist families of invariant tori of positive measure.
> Furthermore, these elliptic orbits occur frequently in Hamiltonian systems. Thus, the hypothesis of metrical transitivity is false in a definite way [15, 1980, pp. 138–139].

The bulk of techniques and results stemmed out, in subsequent years, from the seminal works of Kolmogorov, Moser and Arno'ld in the decade 1953–1963, is usually referred to as "KAM theory". Incidentally, regarding the usage of the KAM acronym, Arnol'd is quite resolute:

> This theory is called KAM, or Kolmogorov – Arnold – Moser, and people say that there is even a KAM theorem. I was never able to understand what theorem it is [19, 1999, p. 16].

Doubts concerning the completeness of Kolmogorov's theorems and issues related to a "KAM theorem" or "KAM theory" might also have been derived from the late availability of translations of the original 1954 paper [5]. From a historiographical standpoint[19)], it is therefore crucial to analyze the dissemination and transmissions of Kolmogorov's contributions after 1957. Volume I of Proceedings of the 1954 ICM in Amsterdam [6] was published only in 1957: here, the original Russian text of Kolmogorov's lecture is found, titled in Russian and French; see, also, [7]. In March 1958, Kolmogorov delivered a presentation, on the same topic, at the Seminar on Analytical Mechanics and Celestial Mechanics hosted by Maurice Janet (1888–1983) at the Faculty of Sciences of the Sorbonne in Paris; for a historical account of this trip by Kolmogorov, see [32]. The French translation of the 1957 Russian Kolmogorov's text [7] appears in the proceedings of the seminar of the young mathematician Jean – Paul Benzécri (1932–2019) [8]. A note stated[20)]: "L'auteur prévoit la publication prochaine de dètails complèmentaires, dans un autre recueil", however, this publication never materialized.

In France, there was a certain interest in classical mechanics, yet it is noteworthy that all other seminars published in that issue of the Janet seminar, except for Kolmogorov's one, dealt with general relativity[21)]. The original Russian version of Kolmogorov's lecture was reprinted in the first volume of the selected works edited by Vladimir Mikhailovich Tikhomirov (b. 1934) and published in 1985 by the publisher Nauka, two years before Kolmogorov's death. Until 1967, there were only two circulating versions of the conference text: in Russian (1957) and in French (1958). In 1967, an English translation appeared as Appendix D in R. H. Abraham's (b. 1936) *Foundations of Mechanics*[22)] published by Benjamin (New York); compare [7] (1967). A second English translation

---

19) Compare also [25].

20) p. 1 of [8], in a footnote.

21) In this regard S. Dumas says: "But KAM theory $[\cdots]$ also had the misfortune of playing out over roughly the same interval during which the revolutions of modern physics took place" in [30, Preface, p. viii].

22) The book was published with the assistance of Jerrold E. Marsden (1942–2010); (a second edition was published in 1978). In the introduction of the 1967 edition Abraham writes:

> In the spring of 1966, I gave a series of lectures in the Princeton University Department of Physics, aimed at recent mathematical results in mechanics, especially the work of Kolmogorov, Arnold, and Moser and its application to Laplace's question of the stability of the solar system. Mr. Marsden's notes of the lectures, with some revision and expansion by both of us, became this book. Although the lectures were attended equally by mathematicians and physicists, our goal was to make the subject available to the nonspecialists (p. xvii).

can be found in the National Aeronautics and Space Administration (NASA), [7] (1972). As for the 1954 paper [5], for decades only the Russian original version was available. An English translation (Los Alamos Scientific Laboratory translation LA-TR-71-67 by Helen Dahlby) was published in 1979 in the proceedings of the Volta Memorial conference at (Como, 1977) *Stochastic behavior in classical and quantum Hamiltonian systems*. Finally, the Amsterdam conference text and both the 1953 and 1954 papers were included in the 1991 American edition by Kluwer in vol. 1 of Kolmogorov's Selected Works[23)].

In addition to the difficulties concerning the availability of translations of the three papers in the 1960s, the circulation of Kolmogorov's research program beyond the Iron Curtain was conditioned also by the process of restoring international contacts face-to-face among mathematicians, contacts, which were formally resumed only after a period of stagnation. Moreover, the shift from French and German to Russian and English as the dominant languages within the international mathematical community after World War II created a new communication challenge[24)].

One more issue, in this context, is related to the relevance of classical mechanics in the 20th century. Classical mechanics seemed to be neglected, both by theoretical physicists and mathematicians, in many countries west of the Iron Curtain (except France) in the mid-century as compared to the prominent position it had held in the 19th and early 20th century[25)]. For example, Clifford Truesdell (1919–2000) in his *History of Classical Mechanics*, writes:

> The word “classical” has two senses in scientific writing; (1) acknowledged as being of the first rank or authority, and (2) known, elementary, and exhausted (“trivial” in the root meaning of that word). In the twentieth century mechanics based upon the principles and concepts used up to 1900 acquired the adjective “classical” in its second and pejorative sense, largely because of the rise of quantum mechanics and relativity. [...] Engineers still had to be taught classical mechanics, because in terms of it they could understand the machines with which they worked and could devise new machines for new purposes. Research in mechanics came to be slanted toward the needs of engineers and to be carried out largely by university teachers who regarded mathematics as a scullery-maid, not a goddess or even a mistress [14, pp. 127–128].

Let us conclude this section with a few remarks.

From the above analysis, discussions on Kolmogorov's theorems in [5] appear to be linked to two profound conceptual circumstances:

1. The level of detail required for a proof to be truly convincing and the limits of the general conception of absolute deductive proof.

2. The connection between an overall research program and the building of a mathematical theory. Indeed, the statements of both theorems implicitly contain the core of Kolmogorov's research program, and the revolutionary features in the 1954 paper [5] could have been better understood by considering Kolmogorov's detailed presentation accompanied by the bibliographic references in [7].

Finally, the complete absence of mathematical details concerning Theorem 2 in [5], together with Moser's and Arnol'd's extremely different reactions to this fact, still needs clarification, particularly in view of the importance given by Kolmogorov to the metrical approach.

In Section 3 below we shall show how Theorem 2 can be derived from Theorem 1, adding a few standard technical details[26)].

[23)] See [4, 5] and [7]. Our quotations in Section 1 of the Amsterdam lecture and of the 1954 paper are taken from this translation. The Russian mathematical physicist Vladimir Markovich Volosov was in charge of the translation of the whole volume.

[24)] Starting in 1945, for instance, the British Mathematical Society had initiated the systematic English translation of the Russian journal *Uspekhi Matematicheskikh Nauk*, titled *Russian Mathematical Surveys*.

[25)] A revival of classical mechanics and dynamical systems took place in the 1960–80s, also thanks to the contribution by Smale; compare [17, 23].

[26)] After all, the already cited words of Arnol'd (“Kolmogorov's theorem was proved by Kolmogorov”) might be particularly appropriate.

3. THE THEOREMS IN KOLMOGOROV'S 1954 PAPER

Here, we analyze in detail the two theorems appearing in Kolmogorov's paper [5]: in Section 3.1 we recall the discussion made in [26] of Theorem 1, adding a few remarks; in Section 3.2 we propose a proof of Theorem 2 based on the scheme of proof of Theorem 1 given by Kolmogorov and implemented in [26].

3.1. Theorem 1

The following is an extended statement of Theorem 1 following [26]. Such a theorem deals with small analytic perturbations of a real analytic Hamiltonian in "Kolmogorov normal form", namely, a Hamiltonian K of the form[27)]

$$K = K(y, x) = E + \omega \cdot y + Q(y, x), \qquad Q = O(|y|^2), \tag{3.1}$$

where $E \in \mathbb{R}$,

$$\omega \in \mathbb{R}^d_{\gamma,\tau} := \Big\{\omega \in \mathbb{R}^d : \ |\omega \cdot n| \geqslant \frac{\gamma}{|n|^\tau}, \quad \forall\, n \in \mathbb{Z}^d \backslash \{0\}\Big\}, \tag{3.2}$$

is a Diophantine frequency, for some $\tau \geqslant n-1$, $\gamma > 0$, and Q is nondegenerate in the sense that

$$\det\langle \partial_y^2 Q(0, \cdot)\rangle := \int_{\mathbb{T}^d} \partial_y^2 Q(0, x) \frac{dx}{(2\pi)^d} \neq 0. \tag{3.3}$$

The phase space is[28)] $\mathcal{M} = B_\xi(0) \times \mathbb{T}^d$, endowed with the standard symplectic form $dy \wedge dx = \sum_j dy_j \wedge dx_j$, and the Hamiltonian flow generated by the Hamiltonian K, $t \to \Phi^t_K(y, x)$ is the solution of the system of equations

$$\begin{cases} \dot y(t) = -\partial_x K(y(t), x(t); \varepsilon), \\ \dot x(t) = \partial_y K(y(t), x(t); \varepsilon), \end{cases} \quad \begin{cases} y(0) = y, \\ x(0) = x. \end{cases}$$

The special feature of a Hamiltonian K in Kolmogorov's normal form is that *the torus $\mathcal{T}_0 := \{0\} \times \mathbb{T}^d \subseteq \mathcal{M}$ is a Lagrangian transitive invariant torus for K*, since $\Phi^t_K(0, x) = (0, x + \omega t)$. The Diophantine vector ω is called *the frequency vector of the invariant torus $\mathcal{T}_0$*.

Given $\xi, \varepsilon_0 > 0$, define the complex domain

$$W_{\xi,\varepsilon_0} := D^d_\xi(0) \times \mathbb{T}^d_\xi \times D^1_{\varepsilon_0}(0) \subseteq \mathbb{C}^{2d+1}, \tag{3.4}$$

where $D^m_r(z)$ denotes the complex d-ball of radius r centered at $z \in \mathbb{C}^m$, and $\mathbb{T}^d_\xi$ is the complex neighborhood of the torus $\mathbb{T}^d$ given by $\{x \in \mathbb{C}^d : |\operatorname{Im} x_j| < \xi, \forall j\}/(2\pi\mathbb{R}^d)$. For a real analytic function $f : W_{\xi,\varepsilon_0} \to \mathbb{C}$ we denote its sup-norm on W_{ξ,ε_0} by $\|f\|_{\xi,\varepsilon_0}$, and its sup-norm (at fixed ε) by $\|f\|_\xi$. Then, Theorem 1 in [5] can be formulated as follows[29)].

Theorem 1. (i) *Let $\omega \in \mathbb{R}^d_{\gamma,\tau}$ and K be a Hamiltonian in Kolmogorov's normal form as in* (3.1)–(3.3) *with K real analytic and bounded on W_{ξ,ε_0} for some $\xi, \varepsilon_0 > 0$; let $P = P(y, x; \varepsilon)$ be a real analytic function on W_{ξ,ε_0}. Then, for any $0 < \xi_* < \xi$, there exists $0 < \varepsilon_* \leqslant \varepsilon_0$ and, for any $0 \leqslant \varepsilon < \varepsilon_*$, a near-identity symplectic transformation $\phi_* : D^d_{\xi_*}(0) \times \mathbb{T}^d_{\xi_*} \to D^d_\xi(0) \times \mathbb{T}^d_\xi$, real analytic on W_{ξ_*,ε_*}, such that the Hamiltonian $H \circ \phi_*$, where $H := (K + \varepsilon P)$, is in Kolmogorov normal form:*

$$H \circ \phi_* = K_* = E_* + \omega \cdot y + Q_*, \qquad Q_* = O(|y|^2). \tag{3.5}$$

---

27) $\omega \cdot y = \sum_j \omega_j y_j$ is the standard inner product and $Q = O(|y|^2)$ means that Q vanishes together with its y-derivatives at $y = 0$; $|\cdot|$ denotes the standard Euclidean norm, and $\partial_y^2 Q$ denotes the Hessian matrix $\partial^2_{y_i y_j} Q$.

28) $B_\xi(y)$ denotes the Euclidean d-ball with radius ξ, centred at y_0 and $\mathbb{T}^d := \mathbb{R}^d/(2\pi\mathbb{Z}^d)$ is the standard flat d-dimensional torus.

29) Part (i) is essentially Kolmogorov's original statement, part (ii) contains the associated estimates.

(ii) In the above statement one can take $\varepsilon_* = \min\{\varepsilon_0, \mathsf{c}_*^{-1}\}$ *where*

$$\mathsf{c}_* = \mathsf{c}\gamma^{-4}(\xi-\xi_*)^{-\nu}C^{\nu}\,\|P\|_{\xi,\varepsilon_0}, \quad \text{and} \quad \begin{cases} C := \max\left\{|E|, |\omega|, \|Q\|_{\xi,\varepsilon_0}, \|T\|, 1\right\}, \\ T := \|\langle \partial_y^2 Q(0,\cdot)\rangle^{-1}\|, \end{cases} \tag{3.6}$$

and $\mathsf{c}, \nu > 1$ *are suitable constants depending only on* d *and* τ*. Furthermore, for any complex* ε *with* $|\varepsilon| < \varepsilon_*$*, one has*

$$\|\phi_* - \mathrm{id}\|_{\xi_*}, |E - E_*|, \|Q_* - Q\|_{\xi_*}, \|\langle \partial_y^2 Q(0,\cdot)\rangle^{-1} - \langle \partial_y^2 Q_*(0,\cdot)\rangle^{-1}\| \leqslant \mathsf{c}_*\,|\varepsilon|.$$

Let us make a few remarks.

(1.1) (*On the dependence of the smallness condition upon the Diophantine constant* γ)
A detailed proof, *apart* from the explicit dependence upon the Diophantine constant γ (which plays an important role in the analysis of the measure of persistent tori), based on Kolmogorov's original outline, has been given in [26]; compare, in particular, Lemma 5, Eq. (27) (the factor 2 in the definition of C in Eq. (26) has, here, been absorbed in the constant c), and Eq. (31). The way the constant c_* depends upon γ needs a short discussion.

The dependence upon γ comes in through the constant $\bar{c}$ in Eq. (18) of [26] (beware that the Diophantine constant γ is denoted by κ in [26]). Now, in the first line of Eq. (18) one can actually take $\bar{c} = \gamma^{-2}\bar{c}_0$ with $\bar{c}_0 = \bar{c}_0(d,\tau)$ depending only on d and τ: indeed, the factor γ^{-1} appears every time the small-divisor operator D_ω^{-1} is applied[30], and the formulae defining the functions on the left-hand side of (18) involve D_ω^{-1} at most *twice*; compare the formulae at the beginning of p. 135 of [26]. Then, it is easy to check that the constant $\bar{c}$ in the estimate on the norm of the "new" perturbing function P' in the second line of Eq. (18) can be taken to be[31] $\bar{c} = \gamma^{-4}\bar{c}_1$, with $\bar{c}_1 = \bar{c}_1(d,\tau)$. Therefore also, c in Eq. (22) and c_* in Eq. (27) in [26] are proportional to a constant $\bar{c}_*(d,\tau)\gamma^{-4}$, which leads to (3.6) above.

Incidentally, we observe that the argument sketched here shows that *the relation* $c\varepsilon_*\gamma^{-4}\|P\|_\xi < 1$, *with a constant* c *independent of* γ*, cannot be improved following Kolmogorov's scheme*: indeed, the norm of $\|P'\|$ cannot be estimated better than by $\gamma^{-4}\|P\|^2$ times a constant independent of γ, and iterating these relations (i. e., replacing $\|P\|$ with $\|P_{j-1}\|$, $\|P'\|$ with $\|P_j\|$; $P_0 := P$), one finds that $|\varepsilon^{2^j}|\|P_j\| \sim \gamma^4(|\varepsilon|\gamma^{-4}\|P\|)^{2^j}$, so that, in order for the Newton scheme to converge, it is *necessary* that $c\,\varepsilon_*\gamma^{-4}\|P\| < 1$.

On the other hand, following Arnol'd's approach [12] — which is a Newton scheme based on approximate solutions of Hamilton – Jacobi equations, where the new perturbing function is of order $\varepsilon^2\gamma^{-2}\|P\|^2$ — allows for a final condition of the form $c\,\varepsilon_*\gamma^{-2} < 1$, which turns out to be optimal (as far as *primary tori* are concerned[32]); compare, e. g., [31].

(1.2) (*On the structure of Kolmogorov's transformation*)
Kolmogorov's transformation ϕ_* has a particularly simple form. Indeed, Kolmogorov describes in detail the transformation ϕ_1, which is the first transformation of the iteration, conjugating the starting Hamiltonian $H = K + \varepsilon P$ to a new Hamiltonian $H_1 := K_1 + \varepsilon^2 P_1 := H \circ \phi_1$, with K_1 in Kolmogorov's normal form with the same[33] ω.

---

[30] D_ω^{-1} is the inverse of the directional derivative $D_\omega = \sum_j \omega_j \partial_{x_j}$ acting on zero-average, real analytic functions on $\mathbb{T}^d$.

[31] The factor $(\gamma^{-2})^2$ comes from the term $P^{(1)}$; compare Eq. (11).

[32] Primary tori are invariant tori which are a deformation of integrable tori and which, in particular, are graphs over $\mathbb{T}^d$; for a recent discussion of a KAM theory for *primary and secondary* tori, see [35].

[33] After the description of ϕ_1 Kolmogorov adds:

> The construction of further approximations is not associated with new difficulties. Only the use of condition (3) for proving the convergence of the recursions, $\mathrm{K}_\theta^{(k)}$, to the analytical limit for the recursion K_θ is somewhat more subtle [5, 1979, p. 55].

In our notation, $\mathrm{K}_\theta^{(k)}$ and K_θ correspond to, respectively, ϕ_j and ϕ_*; Eq. (3) in the citation corresponds to the Diophantine condition (3.2) above.

Now, the transformation ϕ_1 belongs to the (formal) group of near-identity symplectic transformations $\mathcal{G}$ of the form

$$\phi : (y', x') \mapsto \begin{cases} y = y' + \varepsilon\big(u(x') + U(x')y'\big) \\ x = x' + \varepsilon\alpha(x') \end{cases}$$

with U a $(d \times d)$ matrix (depending periodically on x'): such transformations are defined, for small ε, in a neighborhood of the origin $y' = 0$ times $\mathbb{T}^d$; compare Remark 2, and in particular, Eq. (9), in [26]. In the recursion, ϕ_j will have the same form but with ε replaced by $\varepsilon^{2^{j-1}}$, and[34] $\phi_* = \lim_j \phi_1 \circ \cdots \phi_j$ will be given by

$$(I, \theta) \mapsto \phi_*(I, \theta) = (I, \theta) + \varepsilon\big(u_*(\theta) + U_*(\theta)I, \theta + \varepsilon\alpha_*(\theta)\big) \in \mathcal{G}.$$

Thus, defining

$$\zeta_*(\theta) := \phi_*(0, \theta) = \big(\varepsilon u_*(\theta), \theta + \varepsilon\alpha_*(\theta)\big), \tag{3.7}$$

the final invariant torus for the original Hamiltonian H is given by

$$\mathcal{T}_* := \big\{(y, x) = \zeta_*(\theta) : \theta \in \mathbb{T}^d\big\}, \qquad \text{and} \qquad \Phi^t_H\big(\zeta_*(\theta)\big) = \zeta_*(\theta + \omega t).$$

Observe that, since the map $\theta \mapsto \theta + \varepsilon\alpha_*(\theta)$ is a diffeomorphism of $\mathbb{T}^d$ with inverse of the form $x \mapsto x + \varepsilon a_*(x)$, the *invariant torus $\mathcal{T}_*$ is a graph over $\mathbb{T}^d$ given by*

$$\mathcal{T}_* = \big\{(y, x) = \big(\varepsilon \bar{y}_*(x), x\big) : x \in \mathbb{T}^d\big\}, \qquad \bar{y}_*(x) := u_*\big(x + \varepsilon a_*(x)\big).$$

(1.3) (*On ε-analyticity and the convergence of Lindstedt series*)

Let us make the obvious remark — which, however, seems to have been completely overlooked! — that from Theorem 1, it follows immediately that *the invariant torus $\mathcal{T}_*$ depends analytically on ε*: indeed, ϕ_* is real analytic on W_{ξ_*, ε_*}, so that the above function ζ_* is analytic in $\{\varepsilon \in \mathbb{C} : |\varepsilon| < \varepsilon_*\}$.

This observation implies at once that *the Lindstedt series proposed for the first time in* [1] — i. e., the formal ε-expansion of quasi-periodic trajectories for nearly-integrable Hamiltonian systems (which in the present setting is given by ζ_*) — *are actually convergent ε-power series*, a fact that was formally settled, after eighty years from Lindstedt's memoirs and thirteen years after Kolmogorov's paper, by J. Moser in 1967 [13] using his version of KAM theory (which, also, is rather different from Kolmogorov's approach).

Incidentally, it is worthwhile to mention that H. Poincaré, apparently, thought that Lindstedt series were divergent, as it appears from his comments in [2] (vol. II, § IX, n. 123):

> M. Lindstedt ne démontrait pas la convergence des développements qu'il avait ainsi formès, et, en effet, ils sont divergents

and later, [2] (vol. II, § XIII entitled "Divergence des series de M. Lindstedt", n. 149):

> Il semble donc permis de conclure que le séries (2) ne convergent pas. Toutfois le raisonnement qui précède ne suffit pas pour établir ce point avec une rigueur complète[...] Tout ce qu'il m'est permis de dire, c'est qu'il est fort invraisemblable.

3.2. Theorem 2

In Theorem 2, Kolmogorov considers real analytic nearly-integrable Hamiltonian systems, namely, one-parameter families of Hamiltonian systems governed by a real analytic Hamiltonian

$$(\mathtt{y}, \mathtt{x}, \varepsilon) \in W := V \times \mathbb{T}^d \times (-\varepsilon_0, \varepsilon_0) \mapsto \mathtt{H}(\mathtt{y}, \mathtt{x}; \varepsilon) := \mathtt{H}_0(\mathtt{y}) + \varepsilon \mathtt{P}(\mathtt{y}, \mathtt{x}; \varepsilon), \tag{3.8}$$

[34] Of course, all the symbols indexed by $*$ depend on ε (and on the fixed ω).

where $V \subseteq \mathbb{R}^d$ is a bounded regular open connected set, and $\varepsilon_0 > 0$; "regular", here, means that[35)]

$$\lim_{\delta\to 0} \mathrm{meas}(V\backslash V^{(\delta)}) = 0, \qquad \text{where} \qquad V^{(\delta)} := \{\mathtt{y} \in V : B_\delta(\mathtt{y}) \subseteq V\}.$$

The phase space is the set $\mathcal{M} := V \times \mathbb{T}^d$, endowed with the standard symplectic form $d\mathtt{y} \wedge d\mathtt{x} = \sum_j d\mathtt{y}_j \wedge d\mathtt{x}_j$, and ε is a small parameter. Denote by $\phi^t_{\mathtt{H}}(\mathtt{y},\mathtt{x})$ the Hamiltonian flow starting at $(\mathtt{y},\mathtt{x}) \in \mathcal{M}$.

In considering such systems, Kolmogorov says[36)]:

> There arises the natural hypothesis that at small θ the "displaced tori" obtained in accordance with Theorem 1 fill a larger part of region G. This is also confirmed by Theorem 2, pointed out later [5, 1979, p. 56].

Then, Kolmogorov defines the set $\mathcal{Q}_\varepsilon$ of Hamiltonian trajectories in $\mathcal{M}$, which are quasi-periodic with frequencies $\omega \in \mathbb{R}^d$, i. e., trajectories of the form $\phi^t_{\mathtt{H}}(\mathtt{y},\mathtt{x}) = \big(Y(\omega t), X(\omega t)\big)$ for suitable analytic functions $\theta \in \mathbb{T}^d \mapsto \big(Y(\theta), X(\theta)\big) \in \mathcal{M}$, and, at the end of [5], states the following[37)]

Theorem 2. *Let* $\mathtt{H}$ *be as in* (3.8) *and assume* $\det \partial^2_{\mathtt{y}} \mathtt{H}_0 \neq 0$ *on* V. *Then,* $\lim_{\varepsilon\to 0} \mathrm{meas}(\mathcal{M}\backslash\mathcal{Q}_\varepsilon) = 0$.

In the rest of this section, we will show how one can deduce Theorem 2 from Theorem 1 and its proof.

Proof (*of Theorem* 2).

(2.1) *Local reduction*

The claim of Theorem 2 is actually of local nature. Indeed, since V is a regular set, it is enough to show that, for each $\delta > 0$, $\lim_{\varepsilon\to 0} \mathrm{meas}((V^{(\delta)} \times \mathbb{T}^d)\backslash\mathcal{Q}_\varepsilon) = 0$. Furthermore, since $\mathtt{H}$ is real analytic on $V \times \mathbb{T}^d$ and $V^{(\delta)}$ is compact, $\mathtt{H}$ is real analytic and bounded on $\cup_{\mathtt{y}\in V^{(\delta)}} D^d_{\xi_0}(\mathtt{y}) \times \mathbb{T}^d_{\xi_0}$ for a suitable $0 < \xi_0 < \delta$ (for all $|\varepsilon| < \varepsilon_0$). Also, since $\det \mathtt{H}_0'' \neq 0$ on V, by the implicit function theorem, there exists $0 < r < \xi_0/2$ such that the unperturbed frequency map

$$\mathtt{y} \in V \mapsto \omega_0(\mathtt{y}) := \partial_{\mathtt{y}} \mathtt{H}_0(\mathtt{y}),$$

is an analytic diffeomorphism from B onto $\Omega := \omega_0(B)$, for any closed ball $B = B_r(\mathtt{y})$ with $\mathtt{y} \in V^{(\delta)}$. Therefore, since $V^{(\delta)}$ can be covered by a finite number of such balls B, *it is enough to prove that* $\lim_{\varepsilon\to 0} \mathrm{meas}((B \times \mathbb{T}^d)\backslash\mathcal{Q}_\varepsilon) = 0$, *for any set* $B = B_r(\mathtt{y}_0)$ *with* $\mathtt{y}_0 \in V^{(\delta)}$.

(2.2) *Application of Theorem* 1

Before the statement of Theorem 2, Kolmogorov observes:

> The condition of absence of "small divisors" (3) [i. e., the Diophantine inequalities in (3.2)] should be presumed to be fulfilled "in general" since for any $\eta > s - 1$ there exists $c(\lambda)$ such that[38)]
>
> $$|(n,\lambda)| \geqslant c(\lambda)/|n|^\eta$$
>
> at all points of the s-dimensional space $\lambda = (\lambda_1, \ldots, \lambda_s)$, except at a set of Lebesgue measure zero, for any integers n_1, n_2,…, n_s [5, 1991, pp. 352–353].

Indeed, it is an elementary observation that, if $\tau > d - 1$ and we define $\Omega_\gamma := \{\omega \in \Omega : \omega \in \mathbb{R}^d_{\gamma,\tau}\}$, then

$$\mathrm{meas}(\Omega\backslash\Omega_\gamma) \leqslant c\gamma, \tag{3.9}$$

for a suitable constant c depending on d, τ and on the diameter of[39)] Ω.

35) "meas" denotes Lebesgue measure. This regularity assumption on the boundary of the set V (which is satisfied, e. g., by sets with piece-wise regular boundary) is not present in Kolmogorov's paper. Kolmogorov speaks simply of "bounded regions".

36) In this citation, θ and G correspond, in our notation, to, respectively, ε and $\mathcal{M}$.

37) As already mentioned, this statement is not accompanied by any remark, nor references.

38) η, s, c, λ correspond in our notation, respectively, to τ, d, γ, ω.

To proceed in the discussion, fix $\delta > 0$ and pick a ball $B = B_r(\mathtt{y}_0)$ as in **(2.1)** above; fix (once and for all) $\tau > d-1$ and let $\gamma > 0$ (eventually, γ will be chosen as a suitable power of ε). Denote

$$\mathtt{B}_\gamma := \big\{\mathtt{y} \in B : |\omega_0(\mathtt{y}) \cdot n| \geqslant \frac{\gamma}{|n|^\tau},\ \forall\, n \in \mathbb{Z}^d \backslash\{0\}\big\},\ \Omega_\gamma := \omega_0(\mathtt{B}_\gamma) = \{\omega \in \omega_0(B) : \omega \in \mathbb{R}^d_{\gamma,\tau}\}.$$

Observe that $\mathtt{B}_\gamma$ and Ω_γ are nowhere dense sets and that ω_0 is a lipeomorphism (bi-Lipschitz homeomorphism) between them, being an analytic diffeomorphism of B onto $\Omega = \omega_0(B)$. Given $\mathtt{y} \in \mathtt{B}_\gamma$, consider the trivial symplectic map

$$\phi_0 : (y, x) \in B_\xi(0) \times \mathbb{T}^d \mapsto \phi_0(y, x; \mathtt{y}) := (\mathtt{y} + y, x),$$

where $\xi := \xi_0/2$. Then, define

$$H(y, x) := \mathtt{H}_0 \circ \phi_0 + \varepsilon \mathtt{P} \circ \phi_0 =: K + \varepsilon P, \quad P = P(y, x; \varepsilon) := \mathtt{P}(\mathtt{y} + y, x; \varepsilon),$$

and observe that H is real analytic and bounded on[40] W_{ξ,ε_0}. By Taylor's formula, one has

$$K := E + \omega \cdot y + Q, \quad \text{with} \quad \begin{cases} E := \mathtt{H}_0(\mathtt{y}), \quad \omega := \omega_0(\mathtt{y}) \\ Q = \Big(\displaystyle\int_0^1 (1-t)\partial_y^2 \mathtt{H}_0(\mathtt{y} + ty)dt\Big) y \cdot y. \end{cases}$$

For any $\mathtt{y} \in \mathtt{B}_\gamma$, one has that $\omega \in \Omega_\gamma$ so that *we can apply Theorem 1* to H and get a near-identity symplectic transformation ϕ_* so that (3.5) holds for any $\varepsilon < \varepsilon_*$. Notice that *everything here* (H, ϕ_*, etc.) *depend on* $\mathtt{y} \in \mathtt{B}_\gamma$. Thus,

$$\mathcal{T}_* = \mathcal{T}_*(\mathtt{y}) := \psi(\{0\} \times \mathbb{T}^d), \qquad \text{where} \quad \psi := \phi_0 \circ \phi_*\,,$$

is a real analytic Lagrangian torus invariant for the flow of $\mathtt{H}$ and spanned by Diophantine quasi-periodic trajectories. In fact, defining "Kolmogorov's transformation"

$$\psi_{_K} : (\mathtt{y}, \theta) \in \mathtt{B}_\gamma \times \mathbb{T} \mapsto \psi_{_K}(\mathtt{y}, \theta) := \psi(0, \theta; \mathtt{y}) \overset{(3.7)}{=} \big(\mathtt{y} + \varepsilon u_*(\theta; \mathtt{y}), \theta + \varepsilon \alpha_*(\theta; \mathtt{y})\big), \tag{3.10}$$

we find

$$t \mapsto \Phi^t_{\mathtt{H}} \psi_{_K}(\mathtt{y}, \theta) = \psi_{_K}(\mathtt{y}, \theta + \omega t). \tag{3.11}$$

(2.3) *Kolmogorov's set*

In view of (3.9), in order to get a full measure set as ε goes to zero, it is natural to choose γ as a suitable power of ε so that the smallness condition of Theorem 1 holds *uniformly* in phase space. For example, if we take $\gamma = \varepsilon^{1/5}$, we see that ε_* in point (ii) of Theorem 1, for ε small enough, is given by $\varepsilon_* \sim \varepsilon^{4/5}$, so that the condition $\varepsilon < \varepsilon_*$ is fulfilled for any $\varepsilon > 0$ small enough and any $\mathtt{y} \in \mathtt{B}_\gamma = \mathtt{B}_{\varepsilon^{1/5}}$. With these choices, the set

$$\mathcal{K}_\varepsilon := \psi_{_K}(\mathtt{B}_\gamma \times \mathbb{T}^d), \qquad \gamma = \varepsilon^{1/5}, \tag{3.12}$$

defines a set of invariant tori for $\mathtt{H}$, which, by (3.11), is made up of quasi-periodic trajectories, so that $\mathcal{K}_\varepsilon \subseteq \mathcal{Q}_\varepsilon$. Also, from (3.9), since $\mathtt{B}_\gamma = \omega_0^{-1}(\Omega_\gamma)$ and ω_0 is a diffeomorphism, it follows that

$$\operatorname{meas}\big((\mathtt{B} \times \mathbb{T}^d) \backslash (\mathtt{B}_\gamma \times \mathbb{T}^d)\big) \leqslant c'\ \varepsilon^{1/5}. \tag{3.13}$$

All this, in our opinion, must have been rather obvious to Kolmogorov. Furthermore, the Kolmogorov's map $\psi_{_K}$ in (3.10) is a near-identity map (analytic in ε) and it is very tempting, at this point, to conclude that also $\operatorname{meas}\big((\mathtt{B} \times \mathbb{T}^d) \backslash \mathcal{K}_\varepsilon\big) \to 0$ as $\varepsilon \to 0$ concluding the proof of Theorem 2.

---

[39] If δ_Ω denotes the diameter of Ω, one has $\Omega \backslash \Omega_\gamma \subseteq \Big\{\omega \in \Omega : \exists n \neq 0\, \text{s.t.}\, \big|\omega \cdot \frac{n}{|n|}\big| < \frac{\gamma}{|n|^{\tau+1}}\Big\}$, which implies $\operatorname{meas}(\Omega \backslash \Omega_\gamma) \leqslant \displaystyle\sum_{n \neq 0} \frac{\gamma}{|n|^{\tau+1}} \delta_\Omega^{d-1} =: c\gamma$.

[40] Recall the definition in (3.4), and that, by **(2.1)**, $\mathtt{H}$ is real analytic and bounded on $\cup_{\mathtt{y} \in V^{(\delta)}} D^d_{\xi_0}(\mathtt{y}) \times \mathbb{T}^d_{\xi_0}$ for all $|\varepsilon| < \varepsilon_0$.

Clearly, to complete the argument, one needs to have more information on the regularity of ψ_K also in $(\mathtt{y}, \theta)$ in order to control how Lebesgue measure changes under its action. For example, the theorem would follow easily if one proved that ψ_K is a lipeomorphism with Lipschitz constant arbitrarily close to 1, and this, in turn, (because of (3.10)) follows if one proves that the functions u_* and α_* are Lipschitz functions with uniformly bounded Lipschitz constants on $\mathbb{T}^d \times \mathtt{B}_\gamma$.

(2.4) *Lipschitz properties*

As mentioned above, it is enough to show that the functions u_* and α_* are Lipschitz functions with uniformly bounded Lipschitz constants on $\mathbb{T}^d \times \mathtt{B}_\gamma$. The θ dependence is analytic and we may focus on the dependence on $\mathtt{y}$ in $\mathtt{B}_\gamma$ (*joint* lipschitzianity will follow easily).
The way the variable $\mathtt{y}$ enters in the construction of ϕ_* is only through $\omega = \omega_0(\mathtt{y})$, and, since ω_0 is an analytic function, it is enough to check that ϕ_* (*and hence* u_* *and* α_*) *is a Lipschitz function of* ω *with uniformly bounded Lipschitz constants on* Ω_γ.

The starting simple observation is that, if $u = \sum_{n\neq 0} u_n e^{in\cdot x}$ is an analytic map on $\mathbb{T}^d$ with zero average, then

$$(D_\omega^{-1} u)(x) := \sum_{n\neq 0} \frac{u_n}{i\omega \cdot n} e^{in\cdot x} \tag{3.14}$$

depends in a Lipschitz way on $\omega \in \Omega_\gamma$, as we will shortly see. We collect in the following two elementary lemmata what is needed in evaluating Lipschitz constants in Kolmogorov's scheme. Let f be real analytic on W_{ξ,ε_0} and depending also on $\omega \in \Omega \subseteq \mathbb{R}^d_{\gamma,\xi}$ and assume it is uniformly Lipschitz in ω, i. e.:

$$\mathrm{Lip}_{\xi,\varepsilon_0}(f) := \sup \frac{|f(y,x,\omega) - f(y,x,\omega')|}{|\omega - \omega'|} < \infty,$$

where the supremum is taken over all $\omega \neq \omega' \in \Omega$ and over all $(y, x, \varepsilon) \in W_{\xi,\varepsilon_0}$.

Lemma 1. *Let f be as above, let $\lambda = \mathrm{Lip}_{\xi,\varepsilon_0}(f)$, and let $0 < \delta < \xi$. Then, the following holds.*
(i) Let[41] *$\alpha, \beta \in \mathbb{N}_0^d$ be multiindices. Then,*

$$\mathrm{Lip}_{\xi-\delta,\varepsilon_0}(\partial_y^\alpha \partial_x^\beta f) \leqslant c\,\delta^{-(|\alpha|+|\beta|)}\lambda,$$

for a suitable constant c depending only on d and $|\alpha| + |\beta|$.

(ii) $\forall n \in \mathbb{Z}^d$, the Fourier coefficients $f_n(y,\omega)$ of $x \mapsto f(y,x,\omega)$ satisfy[42]

$$|f_n(y,\omega) - f_n(y,\omega')| \leqslant \lambda e^{-|n|\xi}\, |\omega - \omega'|, \quad \forall\, \omega, \omega' \in \Omega. \tag{3.15}$$

(iii) Assume $f_0(y,\omega) = \langle f(y,\cdot,\omega)\rangle = 0$, for all $(y,x) \in D^d_\xi \times \Omega$. Then, $F(y,x,\omega) := D_\omega^{-1} f(y,x,\omega)$ is Lipschitz in ω and

$$|F(y,x,\omega) - F(y,x,\omega')| \leqslant \lambda' |\omega - \omega'|, \quad \forall\; y \in D^d_\xi,\; x \in \mathbb{T}^d_{\xi-\delta},\; \omega, \omega' \in \Omega,$$

where, for suitable constants[43] *c, k depending only on d and τ,*

$$\lambda' := c\,\delta^{-k}\gamma^{-2}(m + \lambda\gamma), \quad m := \sup_{W_{\xi,\varepsilon_0}\times\Omega} |f|. \tag{3.16}$$

Proof.

(i) follows immediately by standard Cauchy estimates;

[41] $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$
[42] As usual, in Fourier analysis, the norm in the exponents are 1-norms.
[43] We take $k \geqslant k_1$ where k_p is as in the "small-divisor" estimate Eq. (6) of [26].

(*ii*) follows immediately by the standard (n-dependent) shift-of-contour argument based on the Cauchy theorem of complex analysis, observing that

$$f_n(y,\omega) - f_n(y,\omega') = \int_{\mathbb{T}^d} \Big(f(y,x,\omega) - f(y,x,\omega') \Big) e^{-in\cdot x} \frac{dx}{(2\pi)^d},$$

and that $|f(y,x,\omega) - f(y,x,\omega')| \leqslant \lambda|\omega - \omega'|$ on $W_{\xi,\varepsilon_0} \times \Omega$.

(*iii*) By (3.14), (3.15), one has, $\forall\ y \in D^d_\xi$, $x \in \mathbb{T}^d_{\xi-\delta}$, and $\omega, \omega' \in \Omega$:

$$\begin{aligned} |F(y,x,\omega) - F(y,x,\omega')| &= \Big| \sum_{n\neq 0} \Big(f_n(y,\omega) \frac{(\omega'-\omega)\cdot n}{(\omega\cdot n)(\omega'\cdot n)} + \frac{f_n(y,\omega) - f_n(y,\omega')}{\omega'\cdot n} \Big) e^{in\cdot x} \Big| \\ &\leqslant |\omega - \omega'| \sum_{n\neq 0} \Big(m e^{-|n|\xi} \frac{|n|^{2\tau+1}}{\gamma^2} + \lambda \frac{|n|^\tau}{\gamma} e^{-|n|\xi} \Big) e^{|n|(\xi-\delta)}, \end{aligned}$$

which, since $\omega \in \mathbb{R}^d_{\gamma,\tau}$, yields the claim by standard estimates[44]. □

Lemma 2. (*i*) *Let $\Omega \subseteq \mathbb{R}^d$ and let $A = A(\omega)$ be an invertible matrix such that $\|A(\omega) - A(\omega')\| \leqslant \lambda|\omega - \omega'|$ and $\|A^{-1}(\omega)\| \leqslant m$, $\forall\ \omega, \omega' \in \Omega$. Then,*

$$\|A^{-1}(\omega) - A^{-1}(\omega')\| \leqslant \lambda'|\omega - \omega'|, \quad \forall\ \omega, \omega' \in \Omega,$$

with $\lambda' = \lambda m^2$.

(*ii*) *For any $\omega \in \Omega \subseteq \mathbb{R}^d$ and any $|\varepsilon| < \varepsilon_*$, let $x \in \mathbb{T}^d \mapsto \varphi(x) = \varphi(x,\omega) = x + \varepsilon a(x,\omega) \in \mathbb{T}^d$ be a near-identity C^1 diffeomorphism*[45]*, with inverse given by $\psi(x') = \psi(x',\omega) = x' + \varepsilon\alpha(x',\omega)$, and satisfying $\varepsilon_*\|a_x\|_\infty < 1$. Assume that $|a(x,\omega) - a(x,\omega')| \leqslant \lambda|\omega - \omega'|$ for any x, ω, ω'. Then,*

$$|\alpha(x',\omega) - \alpha(x',\omega')| \leqslant \lambda'|\omega - \omega'|\,, \qquad \forall\ \omega, \omega' \in \Omega\,,$$

with $\lambda' = \lambda/(1 - |\varepsilon|\|a_x\|_\infty)$. An analogous statement holds in complex neighborhoods of $\mathbb{T}^d$.

Proof. (*i*) Let $v \neq 0$ and let $v' = A^{-1}(\omega)v$. Then, for any $\omega, \omega' \in \Omega$,

$$|A^{-1}(\omega)v - A^{-1}(\omega')v| = \big|A^{-1}(\omega') \big(A(\omega') - A(\omega)\big) v'\big| \leqslant m^2\lambda|\omega - \omega'|\,|v|.$$

(*ii*) Let $x_1 = \psi(x_1', \omega_1)$ and $x_2 = \psi(x_1', \omega_2)$. Then,

$$\begin{aligned} |\alpha(x_1',\omega_1) - \alpha(x_1',\omega_2)| &= |a(x_1,\omega_1) - a(x_2,\omega_2)| \\ &= |a(x_1,\omega_1) - a(x_1,\omega_2) + a(x_1,\omega_2) - a(x_2,\omega_2)| \\ &\leqslant \lambda|\omega_1 - \omega_2| + \|a_x\|_\infty |x_1 - x_2| \\ &= \lambda|\omega_1 - \omega_2| + |\varepsilon|\|a_x\|_\infty |\alpha(x_1',\omega_1) - \alpha(x_1',\omega_2)|, \end{aligned}$$

which implies the claim. The complex case is treated in the same way. □

To describe the iterative step needed to control Lipschitz constant in Kolmogorov's scheme we refer to [26] and, in particular, to Lemma 4 and[46] its proof in [26].

Proposition 1. *Let E, Q, T and P be as in* **(2.2)** *above, and assume that they depend in a Lipschitz way on $\omega \in \Omega_\gamma$ with uniform (on their complex domain of definition) Lipschitz constant Λ. Let*

$$C := \max\big\{|E|, |\omega|, \|Q\|_{\xi,\varepsilon_0}, \|T\|, \Lambda, 1\big\},$$

---

[44] See, e. g., footnote 10 in [26].

[45] The dependence upon ε of the functions is not explicitly indicated.

[46] There is a small correction to be done in the statement of Lemma 4 in [26], namely, the bound on $\|P'\|_{\bar{\xi}}$ in Eq. (18) should be given *after* hypothesis (19).

assume that[47] $\gamma \leqslant 1/2 \min\{1, \Lambda\}$, *and let* $0 < \delta < \xi < 1$. *Finally, let* L, $\phi_1 = \mathrm{id} + \varepsilon\tilde{\phi}$, $E_1 = E + \varepsilon\tilde{E}$, $Q_1 = Q + \varepsilon\tilde{Q}$, $T_1 = T + \varepsilon\tilde{T}$, *and* P_1 *be as in step* (i) *and Lemma 4 of* [26]. *Then,* $\tilde{E}$, $\tilde{Q}$, $\tilde{T}$ *and* $\tilde{\phi}$ *are Lipschitz in* $\omega \in \Omega_\gamma$ *uniformly on* $W_{\bar{\xi},\varepsilon_0}$, $\bar{\xi} := \xi - \frac{2}{3}\delta$, *with Lipschitz constant given by*

$$\tilde{\Lambda} = c'\gamma^{-a'}C^{\mu'}\delta^{-\nu'}M \geqslant L, \qquad M := \sup_{W_{\xi,\varepsilon_0}\times\Omega_\gamma} \|P\|,$$

where $c', a', \mu'\,\nu'$ *are suitable positive constants depending on* τ, d. *Furthermore, if* $\varepsilon_* \leqslant \varepsilon_0$ *is such that* $\varepsilon_*\tilde{\Lambda} \leqslant \delta/3$, *then* P_1 *is Lipschitz in* $\omega \in \Omega_\gamma$ *uniformly on* W_{ξ',ε_*} *with* $\xi' := \xi - \delta$ *with Lipschitz constant* $\tilde{\Lambda}M$. *Finally, for any* $|\varepsilon| < \varepsilon_*$, E_1, Q_1 *and* T_1 *are uniformly Lipschitz in* $\omega \in \Omega_\gamma$ *with Lipschitz constant* $\Lambda_1 := \Lambda + |\varepsilon|\tilde{\Lambda}$.

Proof. Let us give the details for the estimate on the Lipschitz constant of $\tilde{E}$.

$\tilde{E}$ is defined as[48] $\omega \cdot b + P_0(0;\omega)$ where

$$b = -T(\omega)\big(\langle Q_{yy}(0,\cdot;\omega)s\rangle + \langle P_y(0,\cdot;\omega)\rangle\big), \qquad s := -D_\omega^{-1}\big(P_y(0,\cdot;\omega) - P_0(0;\omega)\big).$$

Then, by Lemma 1(iii) with $f = P_y(0,\cdot;\omega) - P_0(0;\omega)$, $\lambda = \Lambda$, and using that $\Lambda\gamma < 1 \leqslant M$, we get[49]

$$\mathrm{Lip}_{\xi-\frac{\delta}{3},\varepsilon_0}(s) \leqslant c\,\delta^{-k}\gamma^{-2}M,$$

and by Lemma 1(i),

$$\mathrm{Lip}_{\xi-\frac{2\delta}{3},\varepsilon_0}(s) \leqslant c\,\delta^{-(k+1)}\gamma^{-2}M.$$

Now, by Lemma 2(i) we get $\mathrm{Lip}(T) \leqslant C^3$, and therefore[50]

$$\mathrm{Lip}_{\xi-\frac{2\delta}{3},\varepsilon_0}(b) \leqslant c\,C^4\delta^{k+1}\gamma^{-2}M, \qquad \mathrm{Lip}_{\xi-\frac{2\delta}{3},\varepsilon_0}(E_1) \leqslant \Lambda + |\varepsilon|\cdot(c\,C^5\delta^{k+1}\gamma^{-2}M).$$

It is not difficult to check that also the Lipschitz constants of[51] $\tilde{Q}$, β_0, β (defined in Remark 2 (a) of [26]), $\tilde{T}$ and $\tilde{\phi}$ satisfy similar estimates; also the estimate on $\mathrm{Lip}(P_1)$ is of the same type, but with an extra factor M, since in the definition of P_1 there appears a term ($P^{(1)}$ in Eq. (11) in [26]), which is quadratic in β. □

Now, the inductive argument follows easily as in the proof of Lemma 5 of [26]. We give a sketch of it.

Let, as in Lemma 5 of [26], $\xi_{j+1} = \xi_j - \delta_j$, $\delta_j = \delta_0/2^j$, $\delta_0 = (\xi - \xi_*)/2$,

$$\phi_j : W_{\xi_j,\varepsilon_*} \times \Omega_\gamma \to D^d_{\xi_{j-1}} \times \mathbb{T}^d_{\xi_{j-1}}, \quad \Phi_j = \Phi_{j-1}\circ\phi_j, \qquad (j \geqslant 1,\ \Phi_0 = \mathrm{id}),$$

so that ϕ_* in **(2.2)** above is given by $\phi_* = \lim \Phi_j$. From the proof of Lemma 5 in [26] and from the Cauchy estimate it follows that

$$\sup_{B\times\mathbb{T}\times\Omega_\gamma} \|\partial_z\Phi_j\| \leqslant 2, \qquad z = (y, x). \tag{3.17}$$

Let

$$\tilde{\Lambda}_i = c'\gamma^{-a'}C^{\mu'}\delta_i^{-\nu'}M_i \geqslant \mathrm{Lip}_{\xi_i,\varepsilon_*}(\phi_i), \tag{3.18}$$

---

47) This assumption, which is armless (since, eventually, γ will be chosen small with ε), is made to simplify the estimate (3.16).

48) Compare step (i) at p. 135 of [26] and recall that $T(\omega) := \langle Q_{yy}(0,\cdot;\omega)\rangle^{-1}$. Usually, we do not indicate the dependence upon ε of the various functions involved.

49) We denote possibly different constants depending on d and τ by c.

50) For products, use $\mathrm{Lip}(fg) \leqslant \mathrm{Lip}(f)\sup|g| + \mathrm{Lip}(g)\sup|f|$, and observe that $\|s\|_{\xi-2/3\delta} \leqslant c\,\delta^{-b}\gamma^{-1}M$; compare, e. g., Eq. (6) in [26].

51) Notice that, since $\tilde{\Lambda} \geqslant L$, the hypotheses of Lemma 4 (compare Eq. (19)) are met.

be as in the i^{th} iteration of[52] Proposition 1, and let $\lambda_j := \mathrm{Lip}(\Phi_j)$ be the Lipschitz constant of Φ_j over $B \times \Omega_\gamma$.

Let $\omega, \omega' \in \Omega_\gamma$, $z_j = \phi_j(y, \theta; \omega)$ and $z'_j = \phi_j(y, \theta; \omega')$, for $y \in B$ and $\theta \in \mathbb{T}^d$. Then, by (3.17) and (3.18)

$$\begin{aligned}
|\Phi_j(y,\theta;\omega) - \Phi_j(y,\theta;\omega')| &= |\Phi_{j-1}(z_j;\omega) - \Phi_{j-1}(z'_j;\omega')| \\
&\leqslant |\Phi_{j-1}(z_j;\omega) - \Phi_{j-1}(z'_j;\omega)| + |\Phi_{j-1}(z'_j;\omega) - \Phi_{j-1}(z'_j;\omega')| \\
&\leqslant 2|\phi_j(z,\theta;\omega) - \phi_j(z,\theta;\omega')| + \lambda_{j-1}|\omega - \omega'| \\
&\leqslant 2|\varepsilon|^{2^j}\tilde{\Lambda}_j|\omega - \omega'| + \lambda_{j-1}|\omega - \omega'|,
\end{aligned}$$

which (dividing by $|\omega - \omega'|$ and taking the supremum over $y \in B$, $\theta \in \mathbb{T}^d$ and $\omega \neq \omega'$) yields the relation

$$\lambda_j \leqslant \lambda_{j-1} + 2\varepsilon^{2^j}\tilde{\Lambda}_j,$$

which, iterated, implies[53], for $|\varepsilon|$ small enough,

$$\lambda_j \leqslant 1 + 2\sum_{i=0}^{\infty} \varepsilon^{2^i}\tilde{\Lambda}_i < 2, \qquad \forall j.$$

Taking the limit as $j \to \infty$, we get $\mathrm{Lip}(\phi_*) \leqslant 2$, which, as discussed above, is all what is needed to conclude the proof of Theorem 2. □

APPENDIX. AN INTERVIEW TO YA. SINAI

One of the authors (I.F.), during her doctoral thesis [34], supervised by Luca Biasco and Ana Millán Gasca, had the opportunity to interview Yakov Sinai on May 28, 2021, in his quality of student and witness to Kolmogorov's legacy. Here, we provide the transcription of this interview[54].

F: The first question concerns Siegel's work on Diophantine estimates. These techniques are also used by Kolmogorov in his proof of the theorem in 1954, but he did not mention Siegel in the bibliography. Do you know if Kolmogorov was aware of Siegel's work on such a matter?

S: In my opinion, he didn't know Siegel's work. Siegel's work was discussed later in Arnol'd's seminar, and I assume that Arnol'd explained Siegel's work to Kolmogorov. As you know, they both used small denominators.

F: Do you know what inspired Kolmogorov for Diophantine estimates?

S: I'm not so sure about this.

F: Okay. So, I move on to the next question: in the published text of the Amsterdam conference, Kolmogorov cited in bibliography "Mathematische Grundlagen der Quantenmechanik" (1932) by von Neumann. Did Kolmogorov ever work on problems in quantum mechanics?

S: Kolmogorov never worked on problems of quantum mechanics because he used to say that he didn't find interesting problems for himself in that field.

F: Okay, but I have a puzzle to solve. I read a sentence written by Kolmogorov that I quote here: "My papers on classical mechanics appeared under the influence of von Neumann's papers on

---

[52] For $i \geqslant 0$, E, $\tilde{E}$, Q, $\tilde{Q}$, Λ, $\tilde{\Lambda}$,…,ξ, δ, ε correspond to E_i, $\tilde{E}_i$, Q_i, $\tilde{Q}_i$, Λ_i, $\tilde{\Lambda}_i$,…,ξ_i, δ_i, ε^{2^i}, while E_1, Q_1, Λ_1, … correspond to E_{i+1}, Q_{i+1}, Λ_{i+1}, etc.

[53] The superexponential series is treated as in [26, p. 138].

[54] **F** = Isabella Fascitiello; **S** = Yakov Sinai; **B** = Luca Biasco.

the spectral theory of dynamical systems..."[55]. In this sentence the reference is to "Mathematical Foundations of Quantum Mechanics".

S: No, I remember it was another von Neumann's paper; there was a paper written by von Neumann about the ergodic theory.

F: Okay. Actually, in another note, written by Shiryaev[56] on Kolmogorov, the author wrote that there is another reference, "Operator Methods in Classical Mechanics"[57].

S: It's possible. That was the main contribution in the operator method.

B: Professor Sinai, do you think that Kolmogorov, for his theorem on the persistence of invariant tori, was also motivated by the foundations of statistical mechanics?

S: He never mentioned this. He just mentioned the work of Chazy[58]. Chazy was a friend, a mathematician, or maybe a physicist; he was the first person who wrote about statistical and central limit theory and other papers on probability theory, which can be used in classical mechanics.

F: Another question concerns your article in the book *Kolmogorov in Perspective*. You wrote that "in the fall of 1957 Kolmogorov began a famous course of lectures on the theory of dynamical systems" and that, I quote, "Kolmogorov presented the theorem[59] that was to become the basis for the famous KAM theory, together with a complete proof"[60]. What did you mean? The history of science says that the first complete proof is due to Arnol'd in 1963.

S: There is a very good proof of Kolmogorov's theorem given by a student of Gallavotti[61]. I forgot his name.

B: Maybe it's Luigi Chierchia.

S: Maybe it was him. Yes.

F: But in 1957, Kolmogorov did present a complete proof in this seminar? Is this assertion true?

S: You see, there is a controversy about this. For example, Arnol'd thought that Kolmogorov did not give a complete proof, that his proof had some gaps. And this was a reason why Arnol'd wrote his paper.

F: Okay. What about you? Do you think Kolmogorov give a complete proof of his theorem?

S: It is a controversial question. I believed that Kolmogorov gave a complete proof, but Arnol'd convinced me that Kolmogorov's proof was not complete.

B: According to Arnol'd, was the proof incomplete because Kolmogorov omitted certain steps, or were there indeed certain gaps that Kolmogorov did not address?

---

55) [16, p. 503]

56) In [20, p. 53].

57) *Zur Operatorenmethode in der klassischen Mechanik*. Princeton, Annals of Mathematics, Second Series 33(3), (1932) pp. 587–642.

58) Jean – François Chazy (1882–1955). Two Chazy's papers, 1929 and 1932, are included in the references of [7], both titled *Sur l'allure finale du mouvement dans le problème des trois corps*.

59) He refers to the theorem on the persistence of invariant tori for quasi-integrable Hamiltonian systems, Theorem 1 in [5].

60) See the complete excerpt cited in Section 2 of this article, taken from [21].

61) He refers to Giovanni Gallavotti (b. 1941).

S: This is a complicated matter. There were some gaps in the estimates of the measure of invariant sets. That was the main point where Arnol'd complained about the proof by Kolmogorov. In Kolmogorov's paper, complete estimates of such a measure were not given.

B: So, only regarding this specific point?

S: Yes.

F: Another question concerning the connection between Kolmogorov and Arnol'd. Did Arnol'd ever make a comparison between his form of the theorem on the persistence of invariant tori and Kolmogorov's original one? What were his motivations for giving a different proof of this theorem?

S: Arnol'd wrote a complete proof of Kolmogorov's theorem which was published in a Russian journal, and it was exactly motivated by the fact that the proof in Kolmogorov's paper was not complete.

F: Was Arnol'd thinking about celestial mechanics?

S: Arnol'd was used to think about celestial mechanics, but there was another student of Kolmogorov diligently working on the subject. I am specifically referring to Sitnikov[62)], who, in one of his articles, provides a comprehensive example of the solution to oscillations.

F: So, Kolmogorov also was thinking about celestial mechanics in 1954?

S: He was very much interested in problems in celestial mechanics. Alekseev's papers[63)] on celestial mechanics were certainly influenced by discussions with Kolmogorov.

ACKNOWLEDGMENTS

We are grateful to Ana Millán Gasca and Luca Biasco for their support. In particular, A. M. Gasca has given precious hints concerning Section 2. We are also indebted to the Referee for a very accurate revision, particularly, of the bibliography.

FUNDING

Partially supported by the grant *NRR-M4C2-I1.1-PRIN 2022-PE1-Stability in Hamiltonian dynamics and beyond-F53D23002730006-Financed by E.U.–NextGenerationEU.*

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

[62)] He refers to Kirill Aleksandrovich Sitnikov (b. 1926). Sitnikov is cited by Kolmogorov in the conference text [7], with the article "On the possibility of capture in the three-body problem", published in the Russian Journal *Matematicheskii Sbornik* in 1953. Incidentally, Sitnikov was mainly a student of P. S. Aleksandrov.

[63)] He refers to the following papers of Vladimir Mikhailovich Alekseev (1932–1980), a student of Kolmogorov specializing in celestial mechanics: "Quasirandom vibrations and the problem of capture in the bounded three-body problem" (1967) in *Doklady Akademii Nauk SSSR*, "On the possibility of capture in the three-body problem with a negative value for the total energy constant" (1969) in *Uspekhi Matematicheskikh Nauk*, and "Final motions in the three-body problem and symbolic dynamics" (1981) in *Uspekhi Matematicheskikh Nauk* (translated into English in *Russian Mathematical Surveys*).

REFERENCES

1. Lindstedt, A., Beitrag zur Integration der Differentialgleichungen der Störungstheorie, *Mém. Acad. Sci. St.-Pet.,* 1883, vol. 31, no. 4, 20 pp.
2. Poincaré, H., *Les méthodes nouvelles de la mécanique céleste: In 3 Vols.,* Paris: Gauthier-Villars, 1892, 1893, 1899.
3. *Proceedings of the International Congress of Mathematicians (Cambridge, Mass., 1950): In 2 Vols.,* Providence, R.I.: AMS, 1952.
4. Kolmogorov, A. N., On Dynamical Systems with an Integral Invariant on the Torus, in *Selected Works of A. N. Kolmogorov: Vol. 1. Mathematics and Mechanics,* V. M. Tikhomirov (Ed.), Dordrecht: Kluwer, 1991, pp. 344–348; see also: *Dokl. Akad. Nauk. SSSR,* 1953, vol. 93, no. 5, pp. 763–766.
5. Kolmogorov, A. N., Preservation of Conditionally Periodic Movements with Small Change in the Hamilton Function, in *Stochastic Behaviour in Classical and Quantum Hamiltonian Systems (Volta Memorial Conference, Como, 1977),* G. Casati, J. Ford (Eds.), Lect. Notes Phys. Monogr., vol. 93, Berlin: Springer, 1979, pp. 51–56; see also: *Selected Works of A. N. Kolmogorov: Vol. 1. Mathematics and Mechanics,* V. M. Tikhomirov (Ed.), Dordrecht: Kluwer, 1991, pp. 349–354; see also: *Dokl. Akad. Nauk SSSR (N. S.),* 1954, vol. 98, pp. 527–530.
6. *Proceedings of the International Congress of Mathematicians (Amsterdam, 1954): Vol. 1,* Groningen/Amsterdam: Noordhoff/North-Holland, 1957.
7. Kolmogorov, A. N., Théorie générale des systèmes dynamiques et mécanique classique, in *Proc. of the Internat. Congr. of Mathematicians (Amsterdam, 1954): Vol. 1,* Groningen/Amsterdam: Noordhoff/North-Holland, 1957, pp. 315–333; see also: *Selected Works of A. N. Kolmogorov: Vol. 1. Mathematics and Mechanics,* V. M. Tikhomirov (Ed.), Dordrecht: Kluwer, 1991, pp. 355–374.
8. Kolmogorov, A. N., Théorie générale des systèmes dynamiques et mécanique classique, in *Séminaire Janet. Mécanique analytique et mécanique céleste,* 1957/1958, vol. 1, talk no. 6, 20 pp.
9. Moser, J. K., Review of "Théorie générale des systèmes dynamiques et mécanique classique" (1957), *Math. Rev.,* Providence, R.I.: AMS, 1959.
10. Moser, J., A New Technique for the Construction of Solutions of Nonlinear Differential Equations, *Proc. Natl. Acad. Sci. USA,* 1961, vol. 47, no. 11, pp. 1824–1831.
11. Moser, J., On Invariant Curves of Area-Preserving Mappings of an Annulus, *Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl. IIa,* 1962, vol. 1962, no. 1, pp. 1–20.
12. Arnol'd, V. I., Proof of a Theorem of A. N. Kolmogorov on the Invariance of Quasi-Periodic Motions under Small Perturbations of the Hamiltonian, *Russian Math. Surveys,* 1963, vol. 18, no. 5, pp. 9–36; see also: *Uspekhi Mat. Nauk,* 1963, vol. 18, no. 5, pp. 13–40.
13. Moser, J., Convergent Series Expansions for Quasi-Periodic Motions, *Math. Ann.,* 1976, vol. 169, no. 1, pp. 136–176.
14. Truesdell, C., History of Classical Mechanics: Part 2. The 19th and 20th Centuries, *Naturwissenschaften,* 1976, vol. 63, pp. 119–130.
15. Smale, S., *The Mathematics of Time: Essays on Dynamical Systems, Economic Processes, and Related Topics,* New York: Springer, 1980.
16. Kolmogorov, A. N., *Selected Works: Vol. 1. Mathematics and Mechanics,* Math. Appl. (Soviet Ser.), vol. 25, Dordrecht: Kluwer, 1991.
17. *Chaos et déterminisme,* A. Dahan-Dalmedico, J.-L. Chabert, K. Chemla (Eds.), Paris: Seuil, 2016.
18. Lehto, O., *Mathematics without Borders: A History of the International Mathematical Union,* New York: Springer, 1998.
19. Arnold, V. I., 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),* E. Bierstone, B. Khesin, A. Khovanskii, J. E. Marsden (Eds.), Fields Institute Communications, vol. 24, Providence, R.I.: AMS, 1999, pp. 1–18. See also (reprinted with some changes): *Amer. Math. Monthly,* 2004, vol. 111, no. 7, pp. 608–624; *Mathematical Events of the Twentieth Century,* A. A. Bolibruch, Yu. S. Osipov, Ya. G. Sinai (Eds.), Berlin: Springer, 2006, pp. 19–47; *Arnold: Swimming against the Tide,* B. A. Khesin, S. L. Tabachnikov (Eds.), Providence, R.I.: AMS, 2014, pp. 11–29.
20. Shiryaev, A. N., Andrei Nikolaevich Kolmogorov (April 25, 1903 to October 20, 1987): A Biographical Sketch of His Life and Creative Paths, in *Kolmogorov in Perspective,* Hist. Math., vol. 20, Providence, R.I.: AMS, 2000, pp. 1–87.
21. Sinai, Ya. G., Remembrances of A. N. Kolmogorov, in *Kolmogorov in Perspective,* Hist. Math., vol. 20, Providence, R.I.: AMS, 2000, pp. 117–120.
22. Moser, J., Remark on the Paper "On Invariant Curves of Area-Preserving Mappings of an Annulus", *Regul. Chaotic Dyn.,* 2001, vol. 6, no. 3, pp. 337–338.

23. Aubin, D. and Dahan Dalmedico, A., Writing the History of Dynamical Systems and Chaos: *Longue durée* and Revolution, Disciplines and Cultures, *Historia Math.,* 2002, vol. 29, no. 3, pp. 273–339.
24. Broer, H. W., KAM Theory: The Legacy of A. N. Kolmogorov's 1954 Paper, *Bull. Amer. Math. Soc. (N. S.),* 2004, vol. 41, no. 4, pp. 507–521.
25. *Landmark Writings in Western Mathematics 1640 – 1940,* I. Grattan-Guinness (Ed.), Amsterdam: Elsevier, 2005.
26. Chierchia, L., A. N. Kolmogorov's 1954 Paper on Nearly-Integrable Hamiltonian Systems, *Regul. Chaotic Dyn.,* 2008, vol. 13, no. 2, pp. 130–139.
27. Curbera, G. P., *Mathematicians of the World, Unite!: The International Congress of Mathematicians — A Human Endeavor,* Wellesley, Mass.: Peters/CRC, 2009.
28. Demidov, S. S., Les relations mathématiques franco-russes entre les deux guerres mondiales, *Rev. Hist. Sci.,* 2009, vol. 62, no. 1, pp. 119–142.
29. Sevryuk, M. B., Translation of the V. I. Arnold Paper "From Superpositions to KAM Theory", *Regul. Chaotic Dyn.,* 2014, vol. 19, no. 6, pp. 734–744.
30. Dumas, H. S., *The KAM Story: A Friendly Introduction to the Content, History, and Significance of Classical Kolmogorov – Arnol'd – Moser Theory,* Singapore: World Sci., 2014.
31. Chierchia, L. and Koudjinan, C., V. I. Arnold's "Pointwise" KAM Theorem, *Regul. Chaotic Dyn.,* 2019, vol. 24, no. 6, pp. 583–606.
32. Mazliak, L., Andrei Nikolaevitch Kolmogorov's Visits to France. Resuming the Scientific Relationship between France and Soviet Union after Stalin's Death, in *26th Internat. Congr. of History of Science and Technology (Prague, Czech Republic, Jul 2021),* 13 pp.
33. Fascitiello, I., Thirty Years After: Insights on the Cultural Origins of Kolmogorov's 1954 Invariant Tori Theorem from a Short Conversation with Arnold, arXiv:2212.06030 (2022).
34. Fascitiello, I., Andrej N. Kolmogorov's 1954 Theorem on the Persistence of Invariant Tori: A Historical Perspective on Its Cultural Roots and Its Meaning in the History of Classical Mechanics, *PhD Dissertation,* Roma, Università Roma Tre, 2023, 216 pp.
35. Biasco, L. and Chierchia, L., Singular KAM Theory, arXiv:2309.17041 (2023).

Publisher's note. Pleiades Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.