The Annals of Probability
1989, Vol. 17, No. 3, 866–944

KOLMOGOROV: LIFE AND CREATIVE ACTIVITIES

By A. N. Shiryaev

Steklov Mathematical Institute

> "Kolmogorov occupies a unique place in modern mathematics and in the scientific world. By the diversity and breadth of his scientific interest he recalls the classical natural scientists of past centuries."
>
> Bogolyubov, Gnedenko and Sobolev ([21], page 24)

In 1985, 1986 and 1987 Nauka (Science) publishers issued three volumes of Andrei Nikolaevich's selected works (in Russian) with commentaries by him as well as by his pupils and followers:

Mathematics and Mechanics
Probability Theory and Mathematical Statistics
Information Theory and the Theory of Algorithms

Though these three volumes [MM, PS, IA] contain as many as 60, 53 and 13 papers, respectively (author's selection), they cover far from all that he accomplished in these areas of science (see the list of Andrei Nikolaevich Kolmogorov's works on pages 945–964). However, even a brief review of the lists of contents astonishes the reader with the breadth and profoundness of the material therein.

Topics in the theory of trigonometric series, theory of measure and sets, studies in the theory of integration, approximation theory, constructive logic, topology, theory of superposition of functions and Hilbert's 13th problem, topics in classical mechanics, ergodic theory, theory of turbulence, diffusion and patterns (models) in the dynamics of populations, papers on the foundations of probability theory, limit theorems, theory of stochastic (Markov, stationary, branching, ...) processes, mathematical statistics, theory of algorithms, information theory, ... —even this is hardly a complete list of all the branches of science in which Andrei Nikolaevich obtained fundamentally important results, which determined the state of many fields of 20th century mathematics and possible directions for their development.

Kolmogorov's papers on the applications of mathematical methods in the social sciences (including articles on the theory of poetry and the statistics of text and literature), the history and methodology of mathematics and the teaching of mathematics in schools, together with popular works for schoolchildren and schoolteachers of mathematics, will supposedly be included in the forthcoming volumes of his selected works, already scheduled for publication.

The exceptional breadth of Andrei Nikolaevich Kolmogorov's scientific interests, and his extraordinary scientific productivity and generosity, are clearly indicated by the titles of his lectures, delivered at meetings of the Moscow Mathematical Society over many years.

Received August 1988.

The list of references at the end of this article consists of papers and books by other authors that are cited in the present article. These are given as [1], [2],.... Citations of items from the general list of principal publications by A. N. Kolmogorov on pages 945–964 are given as [K1], [K2],....

Whenever possible the quotations from Kolmogorov's works are referenced in two ways: Quotations from original sources and those of [MM, PS, IA] ([K467, K471, K473], where the articles are given in Russian).

D. Reidel has committed itself to publication of the volumes [MM, PS, IA] in English translation.

Childhood and school years (1903–1920). Andrei Nikolaevich Kolmogorov was born on April 25, 1903 in Tambov where his mother Maria Yakovlevna Kolmogorova had been delayed on the way back from the Crimea. Maria Yakovlevna died in childbirth and her son of 10 days was taken first to Yaroslavl and then to his maternal grandfather's house in the village of Tunoshna (17 miles from Yaroslavl down the Volga), where he was adopted by Maria Yakovlevna's sister, Vera Yakovlevna. Andrei Nikolaevich Kolmogorov's father, Nikolai Matveevich Kataev, was a qualified agronomist and statistician (what they called at the time a "learned agronomist") who was exiled to Yaroslavl. After the Great October Socialist Revolution he became director of the educational department in Narkomzem (an agricultural ministry), and later perished on the southern front during the offensive by Denikin in 1919.

The three Kolmogorov sisters—Maria Yakovlevna, Vera Yakovlevna and Nadejda Yakovlevna—were independent women with high social ideals. They aided the revolution underground: A clandestine printing press was located in their house, and Kolmogorov's postal address was used for communications from abroad. Referring to the family chronicles, Andrei Nikolaevich wrote in one of his letters that, the apartment having been searched on one occasion, forbidden literature was saved by being hidden under his cradle. Both Vera Yakovlevna and Maria Yakovlevna were arrested and kept for several months in the house of preliminary imprisonment in St. Petersburg.

Andrei Nikolaevich spent the first years of his life until 1910 in this house near Yaroslavl. He wrote as follows in [K315], "How I became a mathematician" (see also [K476], page 7):

"Very early I experienced the joy of mathematical discovery, having noticed the following law at the age of 5 or 6:

$$1 = 1^2,$$

$$1 + 3 = 2^2,$$

$$1 + 3 + 5 = 3^2,$$

$$1 + 3 + 5 + 7 = 4^2, \text{ etc.}$$

"My aunts set up a small school in our house near Yaroslavl where they taught a dozen children of various ages by the most advanced pedagogical methods of the time. The magazine *Spring Swallows* was edited at this school.

It published my discoveries, and also some arithmetical problems that I posed." (Kolmogorov recalled the following such problem: "There is a button with four holes in it. Thread should go through at least two of them to fix the button. In how many ways can this be done?")

In 1910 Vera Yakovlevna and her adopted son moved to Moscow where the boy joined the private E. A. Repman Gymnasium (renamed the Section Grade School No. 23 after the Great October Socialist Revolution).

Andrei Nikolaevich repeatedly emphasized the auspicious climate of the gymnasium, founded by a circle of democratically minded intellectuals. It was also one of the least expensive in terms of tuition.

He recalled that "the gymnasium rooms were small, each holding 15–20 pupils. The teachers were enthusiasts of science, some of them being university lecturers. Our teacher of geography was involved in interesting expeditions. Many schoolchildren competed against each other in their private studies, sometimes even with the intention of shaming the less experienced teachers.

"The school experimented with a new practice of defending final compositions (similar to diploma papers in institutes of higher education).

"I ranked among the first in mathematics in my class, but at the time my major scientific interests were biology first and Russian history second." In this regard Andrei Nikolaevich remarked in answer to "Rovesnik's" questionnaire: "I owe my general goal of a quest for serious and useful activity to the family tradition, and primarily to my aunt Vera Yakovlevna Kolmogorova, who brought me up, and also to the very atmosphere of the remarkable Repman Gymnasium where I studied. My scientific ambitions were inspired by the gymnasium teachers and were ardently cherished in a circle of friends. (I shall name the Seliverstov brothers: Gleb was a mathematician and Nikolai a historian.)"

Recollecting his last year at school Andrei Nikolaevich wrote ([K315]; [K476], page 8):

"Life was not easy in the Moscow of 1918–1920. Only the most persistent studied seriously. Together with other senior students I left Moscow for the construction of the railway from Kazan to Ekaterinburg (now Sverdlovsk). In addition to this work I continued with my independent studies preparing for the secondary school examination, and for taking the secondary school degree externally. Back in Moscow I felt a certain disappointment: They issued a certificate (the secondary school degree) without taking the trouble of testing my knowledge."

Student years and postgraduate schooling (1920–1925, 1925–1929). Having obtained his secondary school degree in 1920, and hesitating in his choice of subject, Andrei Nikolaevich enrolled in the physics and mathematics faculty at Moscow University. Anyone could be admitted then without examination. He wrote in [K470]:

"I arrived at Moscow University with a fair knowledge of mathematics. I knew in particular the beginnings of set theory from the book *New Ideas in Mathematics*. I studied many questions in the encyclopedia of Brokhaus and Efron, filling the gaps resulting from the over-brief coverage of the articles."

Concurrently Andrei Nikolaevich enrolled in the Metallurgy Department of the D. I. Mendeleev Chemical and Technological Institute, having passed the required examinations in mathematics, and studied there for some time. ("I never gave up the idea of a career in technology; somehow, I do not know why I was interested in metallurgy." "Technology was seen then as something more essential and necessary than a pure science," Kolmogorov recalled.) At the same time he still continued his studies of history, attending Professor S. V. Bakhrushin's seminar on ancient Russian history at the Historical Faculty of the University. In this seminar he delivered his first report on land relationships in Novgorod on the basis of his analysis of documents from the 15th and 16th centuries, where he used certain techniques from mathematics ([K315]; [K476], page 8), in particular, Bayes' formula. Very soon Andrei Nikolaevich's interest in mathematics prevailed over his reservations about the relevance of the mathematical profession.

"Having passed my first year exams I was then entitled as a second year student to 16 kg of bread and 1 kg of butter monthly—sufficient for comparatively good physical health. Clothing I had already and wooden-soled shoes I made myself.

"However, in 1922–1925 the need to add other earnings to my rather insubstantial scholarship led me back to the secondary school. It is with great pleasure that I recall now my work at the Potylikhin Experimental School, RSFSR Narcompros (Ministry of Education of the Russian Soviet Federated Socialist Republic). I taught mathematics and physics (they were not afraid then to entrust two subjects to a 19-year-old teacher) and took a very active part in the school public life (I was the secretary of the school council and the school tutor)" —Andrei Nikolaevich describing his student years ([K315]; [K476], page 9).

As a university student Kolmogorov used to attend only specialized courses and seminars. In his first year (1920–1921) he attended lectures on the theory of analytic functions by N. N. Luzin and on projective geometry by A. K. Vlasov. In one such lecture devoted to the proof of Cauchy's theorem Luzin made the following statement: "Let the square be divided into a finite number of squares. Then for any constant C there exists C' such that, for any curve of length not bigger than C, the sum of perimeters of squares touching the curve will not exceed C'." Luzin assigned the proof as an exercise. "I managed to show," Andrei Nikolaevich recalls ([K476], page 11) that "actually the statement was wrong. Nikolai Nikolaevich Luzin immediately understood the idea of my counterexample. It was decided that I would report it at the student mathematical seminar."

Pavel Samuilovich Uryson undertook the verification of all the constructions and proofs. Thus resulted a manuscript dated January 4, 1921, "A report on squaring to the student mathematical seminar"; this was supposed lost but has been recently rediscovered and included in the third volume [IA] of Andrei Nikolaevich's works, *Information Theory and the Theory of Algorithms*, as enclosure No. 1 ([K473], pages 290–294).

In the autumn of 1921 Kolmogorov continued attending lectures by Luzin and Vlasov, and began to attend lectures by Aleksandrov and Uryson. "By the way," Aleksandrov recalled, "Andrei Nikolaevich noticed a mistake in the complicated construction in the lecture by Uryson, in his proof of the theorem on the dimension of three-dimensional space. Uryson corrected the mistake the very next day, but he was greatly impressed by the sharpness of mathematical vision displayed by the 18-year-old student Kolmogorov."

At the invitation of Uryson, Andrei Nikolaevich began calling on him for mathematical supervision. Later Andrei Nikolaevich recollected: "Moscow mathematics of that time was rich in vivid and talented personalities. But even against that background Pavel Samuilovich stood out for the diversity of his interests combined with the purposefulness of his research studies, and his intelligibility in posing problems (in particular those he set for me as part of his responsibility for directing my work), clear assessment of his and others' achievements combined with benevolence towards even the smallest successes" (from an article by Kolmogorov in a book by L. Neiman, *Happiness of Discovery*, Moscow, 1972, dedicated to P. S. Uryson (1848–1923); see [K388]).

While under the influence of P. S. Aleksandrov's lectures, Andrei Nikolaevich began his studies in the very general area of the descriptive theory of sets and had the idea of a rather general "theory of operations on sets", following up and generalizing the studies by Borel, Baire, Lebesgue, P. S. Aleksandrov and M. Ya. Suslin. His work on this topic was finished by January 3, 1922, but its first publication was delayed until 1928 (through no fault of the author) [K15], [MM-13]. (Kolmogorov himself remarked that "my descriptive works gathered dust in Luzin's desk until 1926 as they were believed to be methodologically incorrect".) The second part of this manuscript was accessible to a number of researchers in the descriptive theory of sets, but it first appeared only in 1987 as an enclosure (No. 2, pages 294–303) in the third volume [IA] of Kolmogorov's works.

This work [K15], [MM-13] was the first to introduce the notion of the δS-operation X on sets defined in the following way. Let us take the closed subsets of the interval $(0,1)$, including the empty set, as a basic class of "elementary" sets. The X-operation is determined by the following two objects: a certain collection $\{U^X\}$ of numerical sequences $U^X = \{n_1, n_2, \dots\}$ which are subsets of the sequence of positive integers $\{1, 2, \dots\}$ and a certain sequence of elementary sets $E_1, E_2, \dots$. Every U^X determines the corresponding sequence of sets $E_{n_1}, E_{n_2}, \dots$ and the nucleus $\bigcap_k E_{n_k}$. The union of all such nuclei of sequences corresponding to the given collection $\{U^X\}$ and to the sequence of sets $E_1, E_2, \dots$ constitutes the result of applying the X-operation to this collection of sets.

Later Kolmogorov defined the notion of a complementary operation $\overline{X}$ for the given X-operation and proved the following remarkable result: There exists an X-set [on $(0,1)$], whose complement is not an X-set. Experts in the descriptive theory of sets will appreciate the significance of this result, generalizing in particular Suslin's theorem on the existence of A-sets (the analytical sets, introduced by P. S. Aleksandrov) that are not B-sets (Borel sets).

In the autumn of 1921, as a second-year student Andrei Nikolaevich started to work in V. V. Stepanov's seminar on trigonometric series as well. There he solved the problem of the construction of a Fourier series whose coefficients tend to zero as slowly as desired, a problem with which Luzin was particularly concerned. In the resulting paper, which Kolmogorov described as his "first independent work" (submitted for printing only in late 1922), he formulated his major result on the order of values of the Fourier coefficients [K2], [MM-2]:

"It is known that the Fourier coefficients of a summable function tend to zero. In this article we prove the following proposition concerning the cosine series:

"For any sequence $\{a_n\}_{n=1}^{\infty}$ tending to zero there is a sequence $\{a'_n\}_{n=1}^{\infty}$ such that

1. $|a_n| < |a'_n|$,
2. $\sum_{n=1}^{\infty} a'_n \cos nx$ is the Fourier series of the summable function."

Kolmogorov recalled ([K469]; [K476], page 20) that "as soon as Luzin was told about this he approached me (I remember it happened on the University staircase) and suggested that I should regularly take classes from him."

Thus Kolmogorov became the pupil of N. N. Luzin, whose method of teaching students consisted of weekly scientific discussions on fixed days. Such "intensive work with students was one of those innovations that were introduced by N. N. Luzin" ([K469]; [K476], page 21).

In 1922 Andrei Nikolaevich obtained his most celebrated result on trigonometric series—he constructed an example of a Fourier–Lebesgue series that diverges almost everywhere. In his paper dated June 2, 1922 he says of this result:

"The purpose of this article is to provide an example of a summable (that is, 'integrable' with respect to Lebesgue measure) function whose Fourier series diverges almost everywhere (that is, at every point outside some set of measure zero). The function constructed in this paper is not square-summable and I know nothing about the order of its Fourier coefficients."

He also mentioned here that the methods employed do not allow the construction of a Fourier series diverging everywhere. Later, in 1926, Kolmogorov slightly changed his original method and constructed an example of a summable function with an everywhere divergent Fourier series [K12], [MM-11].

These two examples really shocked and greatly impressed the mathematical community. For the rest of his life Andrei Nikolaevich Kolmogorov retained his interest in the theory of trigonometric functions and orthogonal series, returning now and then to these problems and posing a number of questions to younger mathematicians. He published about a dozen papers in this area, each initiating directions of research which continue to this day. (P. L. Ul'yanov has reviewed this aspect of Kolmogorov's work together with more recent progress in [195]; see also Zygmund [216].)

Along with his interest in the theory of trigonometric series and the descriptive theory of sets, Andrei Nikolaevich concurrently did research in classical analysis, i.e., differentiation, integration, theory of measure and also mathematical logic.

In numerous works of the 1920s attempts were made to generalize the notion of "differentiation." It was hoped to obtain a general definition of the derivative under which any measurable (or at least continuous) function could be differentiated in its natural sense. As a rule, however, for any of the definitions proposed one could construct as a counterexample a continuous function that was not differentiable in the proposed sense. Kolmogorov investigated the problem in its most general form [K7], [MM-7]. He formulated a series of requirements that the "generalized derivative" $f'(x)$ of the function $f(x)$ should satisfy, for example, that it should coincide with the ordinary derivative whenever the latter exists and that if $\phi(x) := af(x)$, then $\phi(x)$ has a generalized derivative at the same points as $f(x)$ and $\phi'(x) = af'(x)$. He showed then that if the function

$$f(x) = \sum_{n=1}^{\infty} \frac{\cos 3^n x}{3^n}$$

has a finite or infinite "generalized derivative" on a set of positive measure, then it is a nonmeasurable function. This example shows that the quest for an effective definition of derivative for the whole class of continuous functions will inevitably lead to the same difficulties as arise in attempts to construct nonmeasurable sets. Similar results concerning the summation of divergent series and the general definition of integral were also formulated in the same article. In a later work ([K26], [MM-16]) of 1930, Kolmogorov analyzed in depth certain established and new constructions of the integral, bringing harmony and clarity to the whole theory of integration, where previous results had generally been uncoordinated and disjointed.

In his introduction to this work [K26] Kolmogorov wrote of its goal as follows:

"...to clarify the logical nature of integration. In combining the various approaches to the idea of the integral by the notion of a generalized integral, the real point is that the generalization of a notion may often be useful in appreciating the essence of its nature. In addition, such generalizations can shed light on the application of the theory. Besides, I see merit in the more general approach through the simplicity and clarity introduced by the new ideas."

The same series of works includes "La définition axiomatique de l'intégrale" [K5], [MM-5], 1925; "Sur le bornes de la généralisation de l'intégrale" [K6], [MM-6] (initially published in the first volume of the selected papers, *Mathematics and Mechanics*, together with proofs of the results obtained in [K5], [MM-5]), 1925; and also a work of 1928, "Sur un procédé d'intégration de M. Denjoy" [K17], [MM-14]. For further details see [93] and [164].

In 1925 there appeared Andrei Nikolaevich's first work in intuitionist logic, entitled "On the tertium non datur principle" [K9], [MM-9]. (The second work is [K36], [MM-19].) Kolmogorov described the main idea of these works as follows (see his commentary in *Mathematics and Mechanics*, page 393):

"It was intended as an introduction to a wider concept. The construction of various branches of classical mathematics within the framework of intuitionist mathematics should have served as evidence of their consistency. The consistency of intuitionist mathematics was believed to be a consequence of its intuitive nature. Such a way is surely unwarranted as a justification of the consistency of the classical logic of statements, but this method may perhaps be applied to establish the consistency of classical arithmetic (cf. Gödel's work of

1933 [71]). I wrote the work [K36] hopeful that the logic of the solution of the problem would subsequently become a permanent part of the logic curriculum. It was supposed to establish an integrated logical system for dealing with objects of two types—statements and problems."

In their extensive commentary to the papers ([K9] and [K36], in the volume [MM], pages 394–404), V. A. Uspenskii and V. E. Plisko continued the development of Kolmogorov's ideas.

In his fourth year at the University in 1924 Andrei Nikolaevich developed the beginnings of his interest in that branch of science where his name was to become greatest—the theory of probability.

His first article in this branch new to him, "Über Konvergenz von Reihen, deren Glieder durch den Zufall bestimmt werden" [K10], [PS-1], dated December 3, 1925, was written in collaboration with A. Ya. Khinchin (also Luzin's pupil). "All my joint work in probability with Khinchin," Andrei Nikolaevich recalled ([K469]; [K476], pages 19–22), "as well as the whole initial period of my work in this area, were marked by the application of the methods developed in the metrical theory of functions. Such topics as conditions for the validity of the law of the large numbers and conditions for convergence of series of independent random variables were actually tackled by methods developed by N. N. Luzin and his pupils in the general theory of trigonometric series."

The work [K10] is made up of four parts. The first was written by A. Ya. Khinchin, the other three by Kolmogorov. In modern notation its results can be written as follows.

Let $\xi_1, \xi_2, \ldots$ be a sequence of independent random variables. Then:

(i) The convergence of the two series $\sum_k E\xi_k$ and $\sum_k D\xi_k$ is sufficient for the almost sure convergence of the series $\sum_k \xi_k$.

(ii) If $\xi_1, \xi_2, \ldots$ are uniformly bounded $[P(|\xi_k| \le C) = 1,\ k \ge 1,\ C < \infty]$, then the convergence of the two series $\sum_k E\xi_k$ and $\sum_k D\xi_k$ is not only sufficient but also necessary for the almost sure convergence of the series $\sum_k \xi_k$.

(iii) If $\xi^c = \xi I(|\xi| \le c)$, then for the almost sure convergence of the series $\sum_k \xi_k$ it suffices that for some $c > 0$ the three series

$$\sum_k E\xi_k^c, \qquad \sum_k D\xi_k^c, \qquad \sum_k P(|\xi_k| \ge c)$$

converge, and if the series $\sum_k \xi_k$ converges also surely, then the three given series necessarily converge for any $c > 0$.

In [K10] Khinchin (part 1) and Kolmogorov (part 2) proved the result by different methods. Khinchin applied a generalization of Rademacher's method (1922), and the latter treated the case of random variables ξ_k taking the two values of C_k and $-C_k$ with probabilities $1/2$ each. Kolmogorov based his proof on the same ideas as those applied in the proof of the now classical "Kolmogorov's inequality":

If $\eta_1, \eta_2, \ldots$ are independent random variables, $E\eta_i = 0$, $S_k = \eta_1 + \cdots + \eta_k$, then

$$P\Big(\max_{1\le k\le n} |S_k| \ge \varepsilon\Big) \le \frac{ES_n^2}{\varepsilon^2}. \tag{1}$$

This inequality is stated exactly thus in Kolmogorov's subsequent work, "Über die Summen durch den Zufall bestimmter unabhängiger Grössen" [K18], [PS-4], written in late 1927, where a second proof of (i) is also given.

The proof of statement (ii), given in the second part of [K10], [PS-1], is largely based on the estimation of the probability $P(\max_{1\le k\le n}|S_k| \ge \varepsilon)$ from below for the case of bounded random variables η_k, $k = 1,\dots, n$. The present form of this inequality is contained in [K18], [PS-4]:

If $\eta_1, \eta_2, \dots$ are independent random variables,

$$E\eta_k = 0,\ P(|\eta_k| \le C) = 1, \qquad k = 1,\dots, n,$$

then

$$P\Big(\max_{1\le k\le n} |S_k| \ge \varepsilon\Big) \ge 1 - \frac{(C+\varepsilon)^2}{ES_n^2}.$$

Thus this rather brief collaboration between Kolmogorov and Khinchin in probability theory includes the "Kolmogorov–Khinchin two-series theorem", and the "Kolmogorov three-series theorem", as well as the simply formulated "Kolmogorov–Khinchin criterion" for the almost sure convergence of a series $\Sigma_n \xi_n$ of mutually independent random variables with zero means, nowadays part of all probability textbooks:

If $\Sigma_n E\xi_n^2 < \infty$, then the series $\Sigma_n \xi_n$ converges almost surely.

The subsequent development of probability theory has shown that the impact of this work ([K10]) goes far beyond providing complete solutions to these problems, important though they are. It gave birth to new methods which have been widely employed since and were later applied to study random processes of more general structure, such as martingale difference sequences.

In 1925 Kolmogorov graduated from Moscow University as a student and enrolled in the University postgraduate school, where Luzin continued to be his scientific supervisor. On the subject of postgraduate training, Andrei Nikolaevich recalled ([K469]; [K476], page 21) that "it did not result then in a thesis paper as happens nowadays: the present scientific degrees were introduced only in 1934." (The scientific degree of Doctor of Physics and Mathematics was awarded to Andrei Nikolaevich Kolmogorov in 1935 for his collection of published work without his submitting a thesis.)

Kolmogorov's fundamental works on the conditions for the validity of the law of large numbers and strong law of large numbers date back to 1927–1929. By the end of 1927 he had completed his research on the sufficient, and the necessary and sufficient, conditions for the validity of the weak law of large numbers, initiated by J. Bernoulli and continued by P. L. Chebyshev and A. A. Markov.

In his introduction to [K472] Kolmogorov wrote: "The cognitive value of probability theory lies in the establishment of strict regularities resulting from the combined effects of mass random phenomena. The very notion of mathematical probability would have been fruitless if it were not realized as the frequency of a certain result under repeated experimentation. That is why the works by Pascal and Fermat can be viewed as only the prehistory of probability, while its true history begins with J. Bernoulli's law of large numbers," formulated as follows:

If $\xi_1, \xi_2, \ldots$ are independent identically distributed (Bernoulli) random variables taking the two values 1 and 0,

$$P(\xi_n = 1) = p, \qquad P(\xi_n = 0) = 1 - p,$$

then for every $\varepsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right) \to 0, \qquad n \to \infty, \tag{2}$$

where

$$S_n = \xi_1 + \cdots + \xi_n.$$

Poisson generalized Bernoulli's law of large numbers to the case of nonidentically distributed Bernoulli variables, as follows:

If $\xi_1, \xi_2, \ldots$ are independent (Bernoulli) random variables, assuming the two values 1 and 0,

$$P(\xi_n = 1) = p_n, \qquad P(\xi_n = 0) = 1 - p_n,$$

and $\sum_{n=1}^{\infty} p_n(1 - p_n) = \infty$, then for any $\varepsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - \frac{ES_n}{n}\right| > \varepsilon\right) \to 0, \qquad n \to \infty, \tag{3}$$

where

$$ES_n = p_1 + \cdots + p_n.$$

In 1867 P. L. Chebyshev investigated the validity of the law of large numbers in the form (3) for an arbitrary sequence of independent random variables (not necessarily Bernoulli distributed). Chebyshev's method is applicable to random variables with finite expectations and variances, and it implies (3) under the weak condition

$$\frac{1}{n^2} \sum_{k=1}^{n} D\xi_k \to 0, \qquad n \to \infty. \tag{4}$$

(This condition is commonly attributed to Markov, who was the first to clearly emphasize its sufficiency; Chebyshev assumed the $E\xi_k^2$ uniformly bounded.)

In [K18], [PS-4], submitted for printing on December 24, 1927, Kolmogorov gets necessary and sufficient conditions for the validity of the "generalized" law of large numbers for the "scheme of series." Suppose one is given a sequence of independent random variables $\xi^n = (\xi_{n1}, \ldots, \xi_{nn})$ for every $n \geq 1$. It is said that the averages

$$\Sigma_n = \frac{\xi_{n1} + \cdots + \xi_{nn}}{n}$$

are *stable* if there exists a sequence of numbers $A_1, A_2, \ldots$ such that for any $\varepsilon > 0$,

$$P(|\Sigma_n - A_n| > \varepsilon) \to 0, \qquad n \to \infty. \tag{5}$$

One also says that the two systems of random variables $\xi^n = (\xi_{n1}, \dots, \xi_{nn})$ and $\bar{\xi}^n = (\bar{\xi}_{n1}, \dots, \bar{\xi}_{nn})$ are *equivalent* if

$$P(\Sigma_n \neq \bar{\Sigma}_n) \to 0, \qquad n \to \infty,$$

where

$$\bar{\Sigma}_n = \frac{\bar{\xi}_{n1} + \cdots + \bar{\xi}_{nn}}{n}.$$

Kolmogorov formulated his generalized law of large numbers as follows. It is necessary and sufficient for the stability of the averages Σ_n, $n \geq 1$, that there exist random vectors $\bar{\xi}^n = (\bar{\xi}_{n1}, \dots, \bar{\xi}_{nn})$ of independent random variables, equivalent to the system $\xi^n = (\xi_{n1}, \dots, \xi_{nn})$, which satisfy

$$\frac{1}{n^2} \sum_{k=1}^{n} D\bar{\xi}_{nk} \to 0, \qquad n \to \infty.$$

He deduced the following result which finally settled the problem of finding natural conditions for the validity of the law of the large numbers [K18], [PS-4]:

Let $\xi_1, \xi_2, \dots$ be a sequence of independent random variables and $S_n = \xi_1 + \cdots + \xi_n$. The necessary and sufficient condition for

$$P\left(\left|\frac{S_n}{n} - \frac{ES_n}{n}\right| > \varepsilon\right) \to 0, \qquad n \to \infty,\ \varepsilon > 0,$$

is the existence for every $n \geq 1$ of a sequence of independent random variables $\bar{\xi}_{n1}, \dots, \bar{\xi}_{nn}$ such that

$$\sum_{k=1}^{n} P(\xi_k \neq \bar{\xi}_{nk}) \to 0,$$

$$\frac{1}{n} \sum_{k=1}^{n} [E\xi_k - E\bar{\xi}_{nk}] \to 0$$

and

$$\frac{1}{n^2} \sum_{k=1}^{n} D\bar{\xi}_{nk} \to 0$$

as $n \to \infty$.

For independent, identically distributed random variables Kolmogorov obtained a similar result:

For the stability of the averages $\Sigma_n = (\xi_1 + \cdots + \xi_n)/n$ the condition

$$nP(|\xi_1| > n) \to 0, \qquad n \to \infty, \tag{6}$$

is necessary and sufficient.

If $E|\xi_1| < \infty$, then the condition (6) holds and the well-known (weak) law of large numbers, earlier obtained by Khinchin, follows from the previous statement:

Let $\xi_1, \xi_2, \dots$ be a sequence of independent identically distributed random variables with $E|\xi_1| < \infty$. Then the law of the large numbers holds:

$$P\left(\left|\frac{S_n}{n} - E\xi_1\right| > \varepsilon\right) \to 0, \qquad n \to \infty.$$

Later (in [K24], [PS-8], 1930) Kolmogorov obtained his other famous result now included in any textbook on probability—the strong law of large numbers:

If $\xi_1, \xi_2, \ldots$ is a sequence of independent random variables with finite second moments and $\Sigma_{n=1}^{\infty} D\xi_n/n^2 < \infty$, then

$$\frac{S_n - ES_n}{n} \to 0, \qquad p\text{-a.s.}$$

In 1933 Kolmogorov, by the way, remarked in his classic book [K40] for the case of identically distributed terms that the following final result is derived from the previous one:

If $\xi_1, \xi_2, \ldots$ is a sequence of independent identically distributed random variables and $E|\xi_1| < \infty$, then

$$\frac{S_n}{n} \to E\xi_1, \qquad p\text{-a.s.};$$

if $E|\xi_1| = \infty$, then S_n/n diverges p-a.s., or, equivalently, $E|\xi_1| < \infty$, $E\xi_1 = \mu \Leftrightarrow (1/n)\Sigma_{k=1}^{n} \xi_k \to \mu$ (p-a.s.).

By contrast, recall that in 1909 Borel [23] was the first to formulate the strong law of large numbers for the Bernoulli case (using the language of number theory):

Let $\xi_1, \xi_2, \ldots$ be a sequence of independent Bernoulli random variables, $P(\xi_n = 1) = P(\xi_n = 0) = \frac{1}{2}$. Then as $n \to \infty$,

$$\frac{S_n}{n} \to \frac{1}{2}, \qquad p\text{-a.s.}$$

Later F. Cantelli proved (in 1917) the following:

If $\xi_1, \xi_2, \ldots$ is a sequence of independent random variables with finite fourth moment and $E|\xi_n - E\xi_n|^4 \leq C < \infty$, $n \geq 1$, then

$$\frac{S_n - ES_n}{n} \to 0, \qquad p\text{-a.s.},$$

as $n \to \infty$.

The very term "strong law of large numbers" was introduced by Khinchin (1927–1928), who provided certain sufficient conditions for its validity, also applicable to the dependent case.

Kolmogorov's results above on the strong law of large numbers for independent random variables are distinguished for both completeness of formulation and clarity of proofs. Many modern probability textbooks highlight these results, both for their intrinsic importance and as evidence of the power of Kolmogorov's inequality (1). For background on the historical emergence of Kolmogorov's strong law of large numbers, see Krengel [101].

With regard to the Kolmogorov strong law of large numbers for independent identically distributed random variables $\xi_1, \xi_2, \ldots$ it is appropriate to mention here its connection with the Birkhoff–Khinchin ergodic theorem. This theorem states that if (ξ_n) is an ergodic stationary sequence, then

$$\frac{1}{n} \sum_{k=1}^{n} \xi_k \to E\xi_1, \qquad p\text{-a.s. and } L^1.$$

For the case of continuous time and ergodic stationary processes $(\xi_t)_{t\geq 0}$, Khinchin formulated the corresponding analogue

$$\frac{1}{t}\int_0^t \xi_s\,ds \to E\xi_0, \qquad p\text{-a.s. and } L^1$$

in 1938 in [96]. In the same volume of *Uspekhi Matematicheskikh Nauk* there appeared a paper by Kolmogorov, "A simplified proof of the Birkhoff–Khinchin ergodic theorem" [K99], [MM-39], in which he shows how the case of continuous time is reduced to discrete time, and explains the form in which Birkhoff formulated his result in 1931. Thus Kolmogorov writes in [K99] that Birkhoff formulated his result "as a theorem of mechanics, or, if you like, a theorem concerning the evolution of an arbitrary system whose state is completely determined by a finite number of parameters and rate of change by differential equations admitting an integral invariant." In the same article Kolmogorov shows that Birkhoff's theorem, proved by him under rather more restrictive conditions, can be reformulated as

$$\frac{1}{t}\int_0^t \xi(T^s\omega)\,ds \to E\xi(\omega), \qquad p\text{-a.s.},$$

where (T^s) is a semigroup of ergodic measure-preserving transformations and $E|\xi(\omega)| < \infty$, and explains how this result is linked with the formulation of Khinchin's corresponding ergodic theorem for stationary processes.

We should note also that there now exist various proofs of Kolmogorov's strong law of large numbers. For example, J. L. Doob's proof is based on his observation that $(S_n/n)_{n\geq 1}$ gives a reverse martingale. Etemadi [49] (see also Grimmett and Stirzaker [72]) gave a proof of Kolmogorov's strong law of large numbers [for identically distributed and pairwise independent random variables (ξ_n)] using only the Borel–Cantelli lemma and the method of *truncation* [passage from ξ_n to $\tilde{\xi}_n := \xi_n I(|\xi_n| < n)$], frequently employed by Kolmogorov, beginning with his first probabilistic work [K10].

In late 1927 Andrei Nikolaevich completed the work [K21] (published in 1929) on the law of the iterated algorithm—one of the remarkable probability theorems that sharpens the strong law of the large numbers.

Khinchin discovered the law of the iterated logarithm (1924) for the Bernoulli scheme (and afterwards for the Poisson scheme):

If $\xi_1, \xi_2, \dots$ are independent identically distributed Bernoulli random variables, $P(\xi_n = 1) = P(\xi_n = -1) = \frac{1}{2}$, then

$$\limsup_n \frac{S_n}{\sqrt{2n\ln\ln n}} = 1, \qquad p\text{-a.s.}$$

This sharpens the strong law of large numbers, since it implies that $S_n/n \to 0$, p-a.s.

Kolmogorov's law of the iterated logarithm is ([K21], [PS-5]):

Let $\xi_1, \xi_2, \dots$ be a sequence of independent random variables with zero means, $\sigma_n^2 = E\xi_n^2$, $B_n = \sum_{k=1}^n \sigma_k^2$, $n \geq 1$. Let there exist a sequence of constants M_n,

$n \geq 1$, such that

$$M_n = o\left(\left(\frac{B_n}{\ln\ln B_n}\right)^{1/2}\right),$$

$$|\xi_n| \leq M_n, \qquad p\text{-a.s.}$$

Then

$$\limsup_n \frac{S_n}{\sqrt{2B_n \ln\ln B_n}} = 1, \qquad p\text{-a.s.}, \tag{7}$$

where $S_n = \xi_1 + \cdots + \xi_n$.

Just as Kolmogorov's inequality, and its proof, is a key technique in obtaining a.s. convergence of series of random terms, so Kolmogorov's proof of the law of the iterated logarithm has become a key part of the arsenal of fundamental probability tools.

Here is Kolmogorov's proof (in its general form and with minor modifications): Statement (7) is equivalent to the validity of the following two statements:

(A) For every $\varepsilon > 0$, the function

$$\phi^\varepsilon(n) = (1+\varepsilon)\sqrt{2B_n \ln\ln B_n}$$

is an upper function for S_n, $n \geq 1$, that is,

$$P\{S_n > \phi^\varepsilon(n) \text{ i.o.}\} = 0. \tag{8}$$

(B) For any $\varepsilon > 0$, the function

$$\phi_\varepsilon(n) = (1-\varepsilon)\sqrt{2B_n \ln\ln B_n}$$

is a lower function for S_n, $n \geq 1$, that is,

$$P\{S_n > \phi_\varepsilon(n) \text{ i.o.}\} = 1. \tag{9}$$

Let $\{n_k\}$ be a nondecreasing sequence of integers such that $n_k \to \infty$ as $k \to \infty$; later we shall choose the sequence explicitly. Then

$$\{S_n > \phi^\varepsilon(n) \text{ i.o.}\} \subseteq \left\{\max_{n_{k-1} < n \leq n_k} S_n > \phi^\varepsilon(n_{k-1}) \text{ i.o.}\right\}$$
$$\subseteq \left\{\max_{n \leq n_k} S_n > \phi^\varepsilon(n_{k-1}) \text{ i.o.}\right\}.$$

In establishing (8) by the Borel–Cantelli lemma, it suffices to show for a special choice of the subsequence $\{n_k\}$, that

$$\sum_{k=1}^{\infty} P\left(\max_{n \leq n_k} S_n > \phi^\varepsilon(n_{k-1})\right) < \infty. \tag{10}$$

A direct application of the Kolmogorov inequality (1) fails to give (10) and Kolmogorov took another route:

(i) He obtained the inequality

$$P\Big(\max_{1\le k\le n} S_k \ge x\Big) \le 2P\left(S_n \ge x - \sqrt{2\sum_{k=1}^{n} E\xi_k^2}\right). \tag{11}$$

(ii) His second major step was an exponential bound for the probabilities of large deviations:

$$P(S_n \ge n) \le \begin{cases} \exp\left(-\dfrac{x^2}{2B_n}\left(1-\dfrac{xM_n}{2B_n}\right)\right), & 0 \le xM_n \le B_n, \\ \exp\left(-\dfrac{x^2}{4M_n}\right), & xM_n \ge B_n. \end{cases} \tag{12}$$

The sequence $\{n_k\}$ is chosen so that for a given $\tau > 0$ we have

$$B_{n_{k-1}} \le (1+\tau)^k \le B_{n_k};$$

under this condition,

$$P\big(S_{n_k} \ge \phi^{\varepsilon}(n_k)\big) \le [k\ln(1+\tau)]^{-(1+\varepsilon)^2(1-\mu)}$$

follows from (12), for any $\mu > 0$. For μ small enough, $(1+\varepsilon)^2(1-\mu) > 1$, whence

$$\sum_{k=1}^{\infty} P\big(S_{n_k} \ge \phi^{\varepsilon}(n_k)\big) < \infty.$$

Together with (11), this implies the required inequality (10).

In order to prove that the functions $\phi_\varepsilon(n)$ are lower for S_n, $n \ge 1$, one uses the second Borel–Cantelli lemma: If the events $A_1, A_2, \ldots$ are independent and $\sum_{n=1}^{\infty} P(A_n) = \infty$, then $P(A_n \text{ i.o.}) = 1$. Kolmogorov showed first that for any $\varepsilon > 0$ one may find such a subsequence $\{n_k\}$ and $\gamma > 0$ such that

$$P\big(S_{n_k} > \phi_\varepsilon(n_k)\text{ i.o.}\big) \ge P\big(S_{n_k} - S_{n_{k-1}} > (1-\gamma)\phi_0(n_k)\text{ i.o.}\big). \tag{13}$$

He next proved that for some $\tau > 0$ one has

$$\begin{aligned} &P\big(S_{n_k} - S_{n_{k-1}} > (1-\gamma)\phi_0(n_k)\big) \\ &\quad \ge P\big(S_{n_k} > (1-\gamma/2)\phi_0(n_k)\big) - P\big(S_{n_k} > \tfrac{1}{3}\gamma\sqrt{\tau}\,\phi_0(n_{k-1})\big). \end{aligned} \tag{14}$$

The last term may be estimated above by (13), and this leaves the task of finding a good lower bound for $P(S_{n_k} > (1-\gamma/2)\phi_0(n_k))$. The required bound is obtained as follows: If $x > 0$, $xM_n/B_n \to 0$ and $x^2/B_n \to \infty$, then

$$P(S_n \ge x) \le \exp\left(-\frac{x^2}{2B_n}(1+\mu)\right)$$

for any fixed $\mu > 0$ and for all sufficiently large n. Thus

$$P\left(S_{n_k} > (1-\gamma/2)\phi_0(n_k)\right) \geq \left(\ln B_{n_k}\right)^{-(1+\mu)(1-\gamma/2)^2}.$$

We now choose μ sufficiently small and τ sufficiently large to obtain the following bound (for sufficiently large k):

$$P\left(S_{n_k} - S_{n_{k-1}} > (1-\gamma)\phi_0(n_k)\right) \geq Ck^{-\alpha},$$

where $C > 0$, $\alpha < 1$. This bound, together with (13), (14) and the second Borel–Cantelli lemma, implies that

$$P\left(S_{n_k} > \phi_\varepsilon(n_k) \text{ i.o.}\right) = 1$$

for any $\varepsilon > 0$. Thus the functions $\phi_\varepsilon(n)$ are the lower functions for S_n, $n \geq 1$, as required.

Kolmogorov's law of the iterated logarithm and the methods of its proof provided resources for many subsequent studies. We note some examples.

In 1937 Marcinkiewicz and Zygmund [121] showed that "o" could not be replaced by "O" in the condition

$$M_n = o\left(\left(\frac{B_n}{\ln\ln B_n}\right)^{1/2}\right)$$

of Kolmogorov's formulation of the law of the iterated logarithm.

In 1941 Hartman and Wintner [77] determined the validity of the law of the iterated logarithm for the case of independent identically distributed variables $\xi_1, \xi_2, \ldots,$ under the assumptions $E\xi_1 = 0$, $E\xi_1^2 < \infty$ only.

In 1966 Strassen [189] obtained the "converse" of the law of the iterated logarithm for independent identically distributed random variables $\xi_1, \xi_2, \ldots,$ by showing the necessity of the assumption of finite second moments. More specifically, if $E\xi_1 = 0$ and

$$\limsup_n \frac{S_n}{\sqrt{2n\ln\ln n}} < \infty, \qquad p\text{-a.s.},$$

then $E\xi_1^2 < \infty$. Strassen [188] also obtained a functional version of the law of the iterated logarithm.

Among Feller's results on the law of the iterated logarithm [53, 54] are statements for the case of variables with infinite second moment. The function $\phi^\varepsilon(t) = (1+\varepsilon)\sqrt{2t\ln|\ln t|}$ is an upper function in both limits $t \to \infty$ (global form) and $t \to 0$ (local form) for the standard Wiener process (or Brownian motion) $W = (W_t)_{t\geq 0}$, as Khinchin showed in his monograph of 1933 [94]. Similarly, the function $\phi_\varepsilon(t) = (1-\varepsilon)\sqrt{2t\ln|\ln t|}$ is a lower function.

The concept of upper and lower functions was introduced by Khinchin, who worked on the problem of how to characterize them. The paper of 1935 by Petrovskii [147] was a breakthrough in this direction, as the author applied methods of the theory of differential equations instead of the probabilistic techniques employed by Khinchin and Kolmogorov.

Petrovskii shows in his paper that the nondecreasing function $\phi = \phi(t)$ with $\phi(0) = 0$ and $\phi(t)/\sqrt{t} \uparrow \infty$, $t \downarrow 0$, is an upper function (in a neighborhood of 0) for the Wiener process $W = (W_t)_{t \geq 0}$ if and only if

$$\int_{0+}^{1} \phi(t) t^{-3/2} \exp\left(- \frac{\phi^2(t)}{2t} \right) dt < \infty.$$

Using the time-inversion property of the Wiener process (if W_t, $t \geq 0$, is a Wiener process, then

$$W_t' = \begin{cases} tW_{1/t}, & t > 0, \\ 0, & t = 0, \end{cases}$$

is also a Wiener process), one derives from this result that the nondecreasing function $\psi = \psi(t)$ is an upper function for W_t, $t \geq 0$, for large t, if and only if

$$\int_{1}^{\infty} \psi(t) t^{-1} \exp\left(- \frac{\psi^2(t)}{2} \right) dt < \infty.$$

(Kolmogorov formulated this result in the early 1930s but omitted the proof in the published paper.)

In modern terms Petrovskii's method may be summarized as follows. Consider the process $X_t = (t, W_t)$ on the domain

$$D = \{(t, w)\colon |w| \leq \phi(t), 0 \leq t \leq 1\}.$$

The point $x_0 = (0,0)$ is called *regular* for the process $X = (X_t)$, $0 \leq t \leq 1$, if $P_{x_0}(\tau_D > 0) = 0$, where $\tau_D = \inf\{s > 0\colon X_s \in \overline{D}\}$. Clearly if x_0 is regular, then $\phi = \phi(t)$ is a lower function, whereas if x_0 is irregular, then ϕ is an upper function. The regularity criterion for the point $x_0 = (0,0)$ stems from the construction of barriers ("superharmonic functions for the process X") for the operator

$$\mathfrak{A} = \frac{\partial}{\partial t} + \frac{1}{2} \frac{\partial^2}{\partial w^2}.$$

(In an extensive and detailed paper by Bingham [19], one finds a "survey of the numerous results of recent years related to the classical law of the iterated logarithm, with particular reference to developments in the decade or so since the survey in Chapter 5 of Stout [187].")

In May 1929 Andrei Nikolaevich finished his four years as a postgraduate student, with 18 mathematical papers to his credit, dated 1923–1928. There followed the question of his future employment. Andrei Nikolaevich described the solution as follows ([K470], page 227):

"There was one vacancy for a senior scientific researcher in the Institute of Mathematics and Mechanics, Moscow University. Along with me, one of the older generation of mathematicians had a claim to the post, and Dmitrii Fedorovich Egorov, the director, though fully aware of my scientific achievements, had always been loyal to the criterion of seniority in employment. I was also attracted by another possibility. The Ukranian Mathematical Institute had been founded in 1926 in Kharkov and was headed by Sergei Natanovich Bernshtein, who was then at the peak of his international fame and domestic authority. The building was already there, but the staff still had to be appointed. Sergei Natanovich offered me the post of research fellow, and also suggested a one-year apprenticeship abroad prior to joining. He proceeded to apply for a Rockefeller scholarship on my behalf. But P. S. Aleksandrov vigorously rebelled against the idea, and finally prevailed upon Egorov to give me priority in employment."

So in June 1929 Kolmogorov joined the Institute of Mathematics and Mechanics at Moscow University which was to be associated with his subsequent work. (Anticipating the chronological development we note here that in March 1931 Kolmogorov became Professor at Moscow University and on December 1, 1933 he was appointed Director of the Scientific and Research Institute of Mathematics of Moscow University.)

The summer of 1929 was marked by the beginning of a close friendship between Andrei Nikolaevich Kolmogorov and Pavel Sergeevich Aleksandrov, a friendship described thus by Aleksandrov in March 1981 [3]:

"My friendship with Andrei Nikolaevich Kolmogorov has been of unparalleled and unprecedented value in my life; this friendship has lasted 50 years, and in its entire half century has neither suffered a single breach, nor been marred by a quarrel, nor did we ever experience any mutual misunderstanding in matters of any significance for our lives and our philosophy; even when our views did not coincide, each treated the other with fullest understanding and sympathy."

The beginning of this friendship happened in 1929 when Kolmogorov decided upon a boat trip down the Volga, as he had once done before.

Recalling the invitations to his companions (including Pavel Sergeevich among them) Kolmogorov wrote (1986) in [K470]: "My personal contacts with Pavel Sergeevich were then pretty limited, although we met frequently, for example at the concerts in the small hall of the conservatory. We would greet each other but never enter into conversation. I may have been put off by his stiff collars and a certain overall impression of primness.... It was never clear to me how I dared to invite Pavel Sergeevich as a third companion.

"On June 16, 1929 we started from Yaroslavl down the Volga. Pavel Sergeevich was new to boating, but he immediately appointed himself our quartermaster and purchased lots of tasty delicacies in Moscow. It is since June 16, our departure day, that I calculate my friendship with Pavel Sergeevich."

Kolmogorov wrote [K470] that the years of their friendship "were the reason why my entire life was on the whole full of happiness, and the basis of that happiness was the unceasing thoughtfulness on the part of Pavel Sergeevich." In the same year (1986) Kolmogorov said at the meeting of the Moscow Mathematical Society (May 27) dedicated to the memory of Aleksandrov (who died on November 16, 1982, at the age of 87): "Probably I could have become a mathematician independently, but my merits as a human being were greatly shaped under Pavel Sergeevich's influence. By the wealth and breadth of his views Aleksandrov was a really extraordinary man.... His knowledge of music and art and his warm and sympathetic attitude to people were remarkable."

In [K470] Kolmogorov vividly and humorously describes this 21-day trip to Samara and on to the Caucasuses (Baku, Lake Sevan, Erivan, Tiflis, ...). On the Sevan Pavel Sergeevich was working on some chapters of his monograph *Topologie* (written in collaboration with Hopf) [5]. Andrei Nikolaevich was writing an article on the theory of integration and was busy pondering the analytical description of Markov processes in continuous time that later on resulted in a memoir, "Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung" [K28], [PS-9].

The thirties (1930–1939). The late 1920s and early 1930s marked a great expansion in Andrei Nikolaevich Kolmogorov's creative activity in a number of branches of mathematics simultaneously.

In 1929 Andrei Nikolaevich published "The general theory of measure and the calculus of probability" [K19], [PS-7] (generally the work is little known to the mathematical public), where he gave the first version of his axiomatic construction of the foundations of probability theory, which subsequently became the well-known "Kolmogorov's axioms" described in *Grundbegriffe der Wahrscheinlichkeitsrechnung* ([K40], 1933).

Kolmogorov speaks in [K19], [PS-7] of the necessity of constructing probability theory as a "rather general and purely mathematical theory," emphasizes the urgent need "to distinguish those elements of probability (theory) that will determine its internal logical structure," says also that "the axiomatization of probability should be constructed on the basis of the general theory of measure and metrical theory of functions—the theory dealing with studies of those properties of functions which depend exclusively on the measure of the sets where these functions assume this or that collection of values" (for example, orthogonality of two functions, or the completeness property of systems of orthogonal functions). He speaks of "the space of the elementary events of the given problem and the probability of the various sets of these events"; notes that "the strength of probability methods in their application to pure mathematics is largely based on the employment of the notion of independent random variables"; focuses attention on the lack of "clear and pure mathematical formulations of the notion of independence of random variables, though there is hardly any difficulty to provide such a formulation."

It was in 1909 that Borel [23] considered the significance of the general theory of measure for the construction of the foundations of probability theory; some aspects of this general idea were highlighted by Łomnicki in 1923 ([117]).

At the beginning of this century Bohlmann [22] attempted the axiomatization of probability theory. The article by Bernshtein [14] on the construction of the foundations of probability theory was published in 1917. (In the axiomatization by Bernshtein the collection of events was considered as a Boolean algebra and was based on the qualitative comparison of random events by the size of their probabilities.) von Mises had a different approach to the foundations of probability theory [200–203]; his notion of the probability of a random event was associated with the result of a certain idealized experiment and with the assumption of the existence of the frequency limit.

In 1933, four years after the appearance of the article "The general theory of measure and the calculus of probability" [K19], [PS-7], Kolmogorov published (Springer-Verlag) the subsequently classic monograph *Grundbegriffe der Wahrscheinlichkeitsrechnung*, where Borel's initial idea took its final shape. This monograph became the background for all subsequent developments of probability theory, a model for exposition and an introductory probability text for many mathematicians.

Itô [86] wrote: "Having read Kolmogorov's *The Foundations of Probability Theory*, I became convinced that probability theory could be developed in terms of measure theory as rigorously as other fields of mathematics." Kac [89, pages 48–49], describing the beginning of his mathematical life and his collaboration with Hugo Steinhaus, wrote about 1935–1938:

"Our work began at a time when probability theory was emerging from a century of neglect and was slowly gaining acceptance as a respectable branch of pure mathematics. The turnabout came as a result of a book by the great Soviet mathematician A. N. Kolmogorov on foundations of probability theory, published in 1933."

Lévy [111, pages 67–68] wrote:

"Dès 1924, je m'étais peu à peu habitué à l'idée qu'il ne fallait pas se borner à ce que j'appelais les vraies lois de probabilité. J'avais cherché à prolonger une vraie loi. Si arbitraire que ce fût, j'étais arrivé à l'idée d'une loi définie dans une certaine famille borelienne. Je ne songeais pas à me dire que c'était là la vraie base du calcul des probabilités; je n'avais pas l'idée de publier cette idée si simple. Puis, un jour, je reçus le mémoire d'A. Kolmogorov sur les fondements du calcul des probabilités. Je compris quelle occasion j'avais perdue. Mais c'était trop tard. Quand saurai-je distinguer ce qui, dans mes idées, merite d'être publié?"

In the introduction to his *The Foundations of Probability Theory* (the Russian edition [K63] appeared in 1936, the English edition in 1950 and the second Russian edition [K403] in 1974) Kolmogorov notes that he would like to indicate those aspects "which go beyond the limits of the above-mentioned range of ideas which are rather familiar to experts in their general terms." These aspects included the following:

1. The distribution of probability in infinite-dimensional spaces;
2. differentiation and integration of expectation with respect to a parameter; and especially
3. the theory of conditional expectation.

He also notes here that "all these new notions and problems necessarily arise in the consideration of entirely concrete physical problems," referring to his joint work with M. A. Leontovich [K42], [PS-14] and Leontovich's work [105].

The significance of all these new results, and of the successful axiomatization of probability, is very clear today, more than 55 years after *Grundbegriffe der Wahrscheinlichkeitsrechnung* was published.

Thus the basic theorem in [K40], paragraph 4, Chapter 3, on the possibility of the construction of a probability measure in an infinite-dimensional space emanating from a consistent collection of finite-dimensional distributions was fundamental to the theory of random processes, which subsequently became a large independent branch of probability with an immense number of applications.

Supported by the Radon–Nikodym theorem (by the way, its modern form dates back to Nikodym's work [139] of 1930), Kolmogorov defines the notion of the conditional probability $P(A\,|\,\mathscr{G})$ of an event A with respect to a σ-subalgebra $\mathscr{G}$, the conditional probability $P(A\,|\,\eta)$ of the event A w.r.t. a random element η, the conditional expectation $E(\xi\,|\,\mathscr{G})$ of the random variable ξ w.r.t. a σ-subalgebra $\mathscr{G}$, and the conditional expectation $E(\xi\,|\,\eta)$ of the random variable ξ w.r.t. the random element η—these are the concepts which are now in the main arsenal of contemporary probability.

In the summer of 1930 Andrei Nikolaevich completed one of his most remarkable probability works, "Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung" [K28], [PS-9], where he laid the foundation of the general theory of Markov random processes and revealed the deep relationship between both this theory and the whole theory of probability, and the theory of ordinary and partial differential equations and mathematical physics.

Kolmogorov says in [K28] that the subject of his studies is processes with continuous time; he specifically emphasizes this aspect and the essential novelty of the suggested methods deriving from it.

The paper [6] by Aleksandrov and Khinchin dedicated to Kolmogorov's 50th birthday says about his work [K28]:

"In the whole of 20th century probability one can hardly find another study that could be similarly as fundamental to the further development of science and its applications, as this work by A. N. Kolmogorov. An extensive branch of probability theory has now developed from it: that is the theory of random processes, which competes with the classical parts of probability in its size and number of applications. Kolmogorov's "differential equations" describing the Markov processes included as special cases, with full mathematical rigor, all those equations (by von Smoluchowski, Chapman, Fokker–Planck, etc.) that had been so far derived and used by physicists on various occasions without rigorous proofs; he provided sufficient justification and clear statement of the conditions needed. An immense number of studies throughout the world have been and still are based on these Kolmogorov equations; they proved fundamental to the further development of theory and to the mathematical development of the most diverse applied problems."

The main subject of "Analytical methods" (as the work [K28] is often referred to) is the transition probability $P(s, x; t, A)$ that we are in the set A at time t under the condition that at time s we are in the state x, $t > s$. A fundamental equation, nowadays called the Kolmogorov–Chapman equation (which expresses the Markov property),

$$P(s, x; t, A) = \int P(s, x; u, dy)P(u, y; t, A), \qquad 0 \le s < u < t, \tag{15}$$

is satisfied by this transition probability, apart from the relevant boundary conditions. (As far as we know, Chapman pointed at this equation in [29].)

It is known now that under suitable conditions (15) allows one to construct a Markov process $X = (X_t)_{t \geq 0}$ whose conditional probability $P(X_t \in A | X_s = x)$ coincides with $P(s, x; t, A)$. In [K28], [PS-9] Kolmogorov does not deal directly with the realization of $X = (X_t)_{t \geq 0}$ but derives the differential equations for the transition probabilities proceeding from (15); thus he creates a new analytical method in its fullest generality and breadth—a method based on differential equations, on the studies of the probability properties of random processes in continuous time, whose evolution obeys the Markov property (15).

In his "Analytical methods" Kolmogorov introduces a notion of "differential characteristics," considering first processes with discrete state-space $E = \{\ldots, i, j, \ldots\}$ and second, continuous diffusion processes with values on the real line $E = R$.

In the first case the limits

$$(16) \qquad \begin{aligned} A_{ii}(t) &= \lim_{\Delta \downarrow 0} \frac{p_{ii}(t, t+\Delta) - 1}{\Delta}, \\ A_{ij}(t) &= \lim_{\Delta \downarrow 0} \frac{p_{ij}(t, t+\Delta)}{\Delta}, \qquad i \neq j \end{aligned}$$

(assumed to exist), act as these differential characteristics for the $p_{ij}(s, t) = P(i, s; t, \{j\})$.

In the second case these characteristics for $F(s, x; t, y) = P(s, x; t, (-\infty, y])$, assumed to have a density

$$f(s, x; t, y) = \frac{\partial F(s, x; t, y)}{\partial y},$$

are the limits

$$(17) \qquad \begin{aligned} A(s, x) &= \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \int_{-\infty}^{\infty} (y - x) f(s, x; s+\Delta, y)\, dy, \\ B^2(s, x) &= \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \int_{-\infty}^{\infty} (y - x)^2 f(s, x; s+\Delta, y)\, dy. \end{aligned}$$

Kolmogorov determines the conditions for the existence of these limits and reveals their real essence: $A(s, x)$ is the "instantaneous mean" and $B^2(s, x)$ the "instantaneous variance."

Then Kolmogorov derives his famous differential equations in each case, that is:

The first, or backward, differential equation on s, x,

$$(18) \qquad -\frac{\partial p_{ij}(s, t)}{\partial s} = \sum_{k \in E} A_{ik}(s) p_{kj}(s, t),$$

$$(19) \qquad -\frac{\partial}{\partial s} f(s, x; t, y) = A(s, x) \frac{\partial}{\partial x} f(s, x; t, y) + \frac{B^2(s, x)}{2} \frac{\partial^2}{\partial x^2} f(s, x; t, y)$$

and the second, or direct differential equation on t, y,

$$\frac{\partial p_{ij}(s,t)}{\partial t} = \sum_{k\in E} p_{ik}(s,t)A_{kj}(t), \tag{20}$$

$$\begin{aligned}\frac{\partial f(s,x;t,y)}{\partial t} &= -\frac{\partial}{\partial y}[A(t,y)f(s,x;t,y)] \\ &\quad + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[B^2(t,y)f(s,x;t,y)\right].\end{aligned} \tag{21}$$

[Equation (21) was derived by Planck and Fokker in connection with their work in diffusion theory.]

In [K28] Kolmogorov poses the problem of existence and uniqueness of the solution of these equations and the differentiability of the transition probabilities.

Papers by Kolmogorov (e.g., [K186]), Feller [52] and many others were subsequently devoted to these problems.

The differential characteristics introduced by Kolmogorov subsequently found their realization in the general situation within the framework of the semigroup approach to the theory of Markov processes.

This theory introduces the infinitesimal operator of the corresponding semigroup of the Markov process as its differential characteristics. Necessary and sufficient conditions were obtained for the infinitesimal operator for the unique determination of the transition function (Feller [55], Dynkin [46]).

As was mentioned above, the Markov process was not considered in "Analytical methods" from the point of its trajectory properties. Kolmogorov deals only with transition probabilities and their differential characteristics. A powerful method for the construction of Markov processes is the method of stochastic differential equations of Itô [81–85, 87] (see also [86] and Gikhman [62–64]), developed in the 1940s and 1950s. (The finite-difference stochastic differential equations were considered by Bernshtein in 1934 [15].)

The gist of Itô's method is to proceed from the simplest process—that is from the Wiener process $W = (W_t)_{t\geq 0}$ with $E\Delta W_t = 0$, $E(\Delta W_t)^2 = \Delta t$ and to construct the process $X = (X_t)_{t\geq 0}$ as the solution of the stochastic differential equation

$$dX_t = A(t, X_t)\,dt + B(t, X_t)\,dW_t. \tag{22}$$

Itô intuitively believed that starting from the point x at time t such a process will behave locally like the Wiener process with "drift" $A(t,x)$ and "diffusion" $B^2(t,x)$; he also proceeds from the essence of the "Kolmogorov differential characteristics" $A(t,x)$ and $B(t,x)$ and gives a new and original method of constructing a Markov diffusion process $X = (X_t)_{t\geq 0}$, whose transition probability $P(s,x;t,\Gamma) = P(X_t \in \Gamma | X_s = x)$ satisfies the Kolmogorov equations.

To this end Itô first assigned a precise meaning to the notion of "stochastic differential equation." He was thus compelled to develop the now commonly used "stochastic integral of a nonanticipating function w.r.t. the Wiener process."

Second, by using successive approximations he showed that Lipschitz conditions and linear growth on x of the coefficients $A(t,x)$ and $B(t,x)$ assure the existence and uniqueness of the solution of (22); third, he established that this solution is a Markov process; and fourth, he applied his famous "Itô formula" (change of variables): If $f = f(t,x)$ belongs to the class C^2,

$$df(t,X_t) = \left[\frac{\partial f}{\partial t}(t,X_t) + A(t,X_t)\frac{\partial f}{\partial x}(t,X_t) + \frac{1}{2}B^2(t,X_t)\frac{\partial^2 f}{\partial x^2}(t,X_t)\right] + B(t,X_t)\frac{\partial f}{\partial x}(t,X_t)\,dW_t,$$

to show that the backward Kolmogorov equations are satisfied for the transition density of the Markov process.

Later, in the 1960s and 1970s the so-called martingale approach was launched, especially after the work by Stroock and Varadhan [190, 191]. It allowed one to prove existence and uniqueness of the so-called weak solution of the stochastic differential equation (22) under rather weak assumptions on $A(t,x)$ and $B(t,x)$, which provided a considerable advance in the solution of the problem (posed in "Analytical methods") on the existence of the diffusion whose "differential characteristics" $A(t,x)$ and $B(t,x)$ are subject to almost no substantial restrictions (see also [65, 88, 114]).

"Analytical methods," together with another work of 1931, "Eine Verallgemeinerung des Laplace–Liapounoffschen Satzes" [K31], [PS-12], showed how to study the transition probability by using differential equations, but it also provided a radically new proof of the Laplace–Lyapunov–Lindeberg theorem, based on the idea that sum S_n, $n \geq 1$, of independent random variables $\xi_1, \xi_2, \ldots$ (with zero means) constitutes a Markov process which converges to a diffusion process under suitable normalization.

By this means Kolmogorov provided a method of constructing an asymptotic expansion for the probabilities

$$P_n(x) = P\left(S_n\big/\sqrt{DS_n} \leq x\right).$$

He also formulated a problem (in [K31]): "What is the probability that all the inequalities $a(t_k) < S_k < b(t_k)$, $k = 1, \ldots, n$, hold?"

Actually this is (in modern terms) a typical boundary problem for the "invariance principle":

Let $\xi_{n1}, \ldots, \xi_{nn}$ be a sequence (for every $n \geq 1$) of independent random variables, $S_{nk} = \sum_{i \leq k}\xi_{ni}$, $E\xi_{ni} = 0$, $\sum_{i \leq n} D\xi_{ni} = 1$, $\sum_{i \leq n} E|\xi_{ni}|^3 = L_n \to 0$, $n \to \infty$. The problem is when and how fast [for sufficiently smooth boundaries $a(t)$ and $b(t)$, $0 \leq t \leq 1$]

$$R_n \equiv |P(a(DS_{nk}) \leq S_{nk} \leq b(DS_{nk}),\ k = 1,\ldots,n) - P(a(t) \leq W_t \leq b(t),\ 0 \leq t \leq 1)| \to 0, \qquad n \to \infty,$$

where $W = (W_t)_{t \geq 0}$ is a standard Wiener process. Kolmogorov showed that the boundary problem of determining the probabilities $P = P(a(t) \leq W_t \leq b(t),$

$0 \le t \le 1$) reduces to and gives an asymptotic expansion for the probabilities

$$P_n = P(a(DS_{nk}) \le S_{nk} \le b(DS_{nk}),\ k = 1, \dots, n),$$

the first term of which is P.

Afterwards Prokhorov [157] obtained an estimate

$$R_n \equiv |P_n - P| = o\left(L_n^{1/4}(\ln L_n)^2\right)$$

for sufficiently smooth functions $a(t)$, $b(t)$, $a(0) < 0 < b(0)$. Skorokhod [182] obtained an estimate $O(\ln n/\sqrt{n})$ for R_n in the case of identically distributed bounded random variables by the "method of the single probability space" (now known as "strong approximation"; see, e.g., Csörgő and Révész [36]), and Nagaev [137] and Sakhanenko [163] refined the estimate to $O(1/\sqrt{n})$, omitting $\ln n$ and removing the assumption of boundedness. (Concerning developments in recent years of methods of approximation of the probabilities P_n, see, e.g., Skorokhod [182], Borovkov [24, 25], Komlós, Major and Tusnády [99], Stout [187], Csörgő and Révész [36] and Bingham [19].)

From June 1930 to March 1931 Kolmogorov was on a nine-month business trip to Germany and France. Together with P. S. Aleksandrov he spent three days in Berlin and moved to Göttingen. In "Reminiscences about P. S. Aleksandrov" [K470], Kolmogorov wrote:

"In those years Göttingen was regarded as the leading mathematical center in Germany and a competitor to Paris in France and Princeton in the USA. Such a status was attained despite a very limited permanent staff. There were only four full professors of mathematics: Hilbert, Courant, Landau and apparently Bernshtein (Hilbert was already 66 years old and due to retire; Hermann Weyl had already been invited to fill the vacancy). Many of Courant's junior staff had the status of assistants. Even Emmy Noether, already regarded as the leading figure of modern general algebra, did not have her full professor's title. Her pupils van der Waerden and Deuring were also assistants.

"The major part of the Göttingen mathematical community was clustered around Hilbert, Courant, Landau and Emmy Noether. It was a very friendly group and Pavel Sergeevich was never viewed as a stranger.... I had extensive scientific contacts in Göttingen. With Courant and his pupils I spoke on limit theorems, where the diffusion processes are limits of discrete random processes; with H. Weyl on intuitionistic logic; and finally with Landau on the theory of functions."

After Göttingen Andrei Nikolaevich went to Munich to visit Carathéodory, who, Kolmogorov recalled [K470], "happened to like my work on measure theory and insisted on its earliest publication," though he was rather cool about the work on the generalization of the notion of integral.

Having been invited by Fréchet, Kolmogorov and Aleksandrov called on him at the Mediterranean in Sanary-sur-Mer (not far from Toulon) and they worked together (on probability in Kolmogorov's case and on set-theoretic topology in Aleksandrov's case); after a short trip (the Bavarian Alps, Ulm, Freiburg in Germany and Lake Annecy and Marseille in France) they finally came to Sanary-sur-Mer.

"Fréchet was busy then with Markov chains in discrete time and with the various types and sets of states. We discussed with him all the Markov problems in their broadest setting. This rather monotonous life—occasionally disrupted by small excursions—went on for a month," Andrei Nikolaevich recalled [K470]. Having turned up in Paris, he proceeds, "... it was natural to enquire about the evaluation of my work and to be somehow advised by the senior mathematical savants, Borel and Lebesgue, about the continuation of my work. But unfortunately my contacts with them were reduced to short formal visits. However, Borel's intervention proved essential for the extension of my French visa. The clearance was issued immediately after the submission of a letter signed *Émil Borel, Ancien Ministre de la Marine.*

"In mathematical matters I gained much from my contacts with P. Lévy. I was repeatedly invited to his home, where we had long and substantial scientific discussions." (See also Lévy [111], pages 87–88, about this visit by Kolmogorov.)

In March 1931 Andrei Nikolaevich became Professor at Moscow University and on December 1, 1933 he was appointed Director of the Scientific and Research Institute of Mathematics of Moscow University, where he remained until April 15, 1939 (returning for a short period, in 1951–1953).

In 1930 and 1932 Kolmogorov published two works in geometry: "Zur topologisch-gruppentheoretischen Begründung der Geometrie" [K25], [MM-15] and "Zur Begründung der projektiven Geometrie" [K37], [MM-20].

The first develops the classical geometries of constant curvature for n-dimensional space, based on topology and the theory of groups. The second gives a new construction of projective geometry on the basis of L. S. Pontryagin's theorem, stating that the only connected locally compact topological (skew) fields with countable base are the following: the field of real numbers, the field of complex numbers and the skew field of quaternions. This theorem enables a direct construction of both real and complex projective geometries.

The later of Kolmogorov's classical works in topology date back to the extremely prolific 1930s. Kolmogorov's main contribution lies in the introduction (in [K67], [MM-29]; this was simultaneous with and independent of Alexander [7, 8]) to algebraic topology of the notion of *operator* and the construction with its help of the *cohomology groups*, also referred to as "upper Betti groups" or "∇-groups," which supplied a powerful and convenient tool for studies of various topological problems, in particular, those related to continuous mappings. Secondly, Kolmogorov [K69], [MM-30] and Alexander [7, 8] determined the operation of multiplication in the cohomology group, thus turning a cohomology into a ring (cohomology ring) which was vital for subsequent studies. Kolmogorov's third outstanding contribution to topology is the "duality law," which refers to the closed sets in any locally compact completely regular topological space, satisfying the condition of acyclicity. (See also the comment by G. S. Chogoshvili, "On A. N. Kolmogorov's works in homology theory" [MM], pages 405–411.)

To this list of remarkable topological works by Kolmogorov should be added "Über offene Abbildungen" [K79], [MM-36], published in 1937, where he constructed a masterly example of a continuous open mapping (that is, mapping open sets to open sets) of the one-dimensional continuum onto the two-dimensional. Kolmogorov wrote in his commentary to the work ([MM], page 412): "P. S. Aleksandrov was very keen about the possibility of increasing the dimension under an open mapping. We were sweating together for a while over the impossibility of increasing the dimension. Gradually this search revealed the reason for our failure. It was the analysis of our failure that finally resulted in the counterexample"; let us add that it also stimulated Soviet topologists to go on with their research on open mappings (L. V. Keldysh, B. A. Pasynkov, ...).

Kolmogorov's topological works of the 1930s also include the one, "Zur Normierbarkeit eines allgemeinen topologischen linearen Raumes" (1934) [K49], [MM-23], where in particular he defines a *topological linear space*, boundedness and convexity of sets in such spaces, and necessary and sufficient conditions for the normability of a topological linear space.

In 1935 and 1936 two more Kolmogorov papers on approximation theory appeared ("Zur Grössenordnung des restgliedes Fourierschen Reihen differenzierbarer Funktionen" [K61], [MM-27] and "Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse" [K62], [MM-28]), bringing about the emergence—so typical of Kolmogorov's work—of new directions in approximation theory.

In the first of these works Kolmogorov considered the class $\mathbb{F}^{(p)}$ of all periodic functions $f = f(x)$, continuous together with their derivatives of order $p - 1$, where the $(p - 1)$th derivative satisfies the Lipschitz condition $|f^{(p-1)}(x) - f^{(p-1)}(y)| \leq |x - y|$ and $\sup_x |f^{(p)}(x)| \leq 1$. Writing

$$R_n(f, x) = f(x) - \left[\tfrac{1}{2}a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)\right]$$

for the remainder term of the Fourier series of the function $f = f(x)$, Kolmogorov considered the problem of finding the value

$$C_n^{(p)} = \sup_{f \in \mathbb{F}^{(p)}} |R_n(f, x)|.$$

In 1910 Lebesgue showed that $C_n^{(1)}$ has order $\log n/n$. In the article under consideration Kolmogorov showed that in general

$$C_n^{(p)} = \frac{4}{\pi^2}\frac{\log n}{n^p} + O\left(\frac{1}{n^p}\right)$$

and proved the explicit formula for the case of p odd,

$$C_n^{(p)} = \frac{1}{\pi}\int_0^{2\pi}\left|\sum_{k=n+1}^{\infty}\frac{\sin kx}{k^p}\right| dx.$$

This work inspired generalizations in various directions: partial Fourier sums were replaced by other approximating expressions; the class $\mathbb{F}^{(p)}$ was replaced by other functional classes, etc. Considerable results in approximation theory were obtained in the 1940s by Kolmogorov's pupil, academician Sergei Mikhailovich Nikol'skii, and by their pupils and followers (see [MM], pages 382–386). In his article [140] Nikol'skii writes of the regular visits by P. S. Aleksandrov and A. N. Kolmogorov to Dnepropetrovsk, where they delivered lectures and conducted scientific seminars, which particularly promoted research in the theory of approximation of functions in this city.

The second Kolmogorov paper, "Über die beste Annäherung von Funktionen einer gegebenen Funktionklasse," was of great and distinctive value; it introduced a new characteristic of approximating properties of classes of functions, later called "Kolmogorov's diameter," which has been attracting wide attention especially since the 1960s.

Kolmogorov formulates his problem as follows:

Let us suppose that we introduce a certain distance for the functions $f, g, \ldots$ under consideration and examine the problem of approximation of the function f by linear combinations $\phi_c = c_1\phi_1 + \cdots + c_n\phi_n$ with fixed functions $\phi_1, \ldots, \phi_n$. P. L. Chebyshev considered the problem of the choice of such coefficients $c = (c_1, \ldots, c_n)$ that the distance $\rho(f, \phi_c)$ can be made arbitrarily small. Kolmogorov now poses a new problem:

Let $F = \{f\}$ be a certain class of functions and

$$D_n(F) = \inf_{(\phi_1, \ldots, \phi_n)} \sup_{f \in F} \inf_{c=(c_1, \ldots, c_n)} \rho(f, \phi_c).$$

It is required, for given n and F, to find $D_n(F)$, and to clarify the existence of the optimal functions $\phi_1, \ldots, \phi_n$ and their uniqueness (to within linear transformations).

For the case of $\rho(f, g) = [\int_0^1 (f - g)^2\, dx]^{1/2}$ and the class F_1 of differentiable functions $f = f(x)$ with $\int_0^1 (f'(x))^2\, dx \leq 1$, Kolmogorov showed that

$$D_n(F_1) = \frac{1}{\pi n}, \qquad n = 1, 2, \ldots,$$

and that the optimal functions $\phi_1, \ldots, \phi_n$ are

$$1, \sqrt{2}\cos \pi k x, \qquad k = 1, \ldots, n-1.$$

For the class F_p^*, $p \geq 1$, consisting of all p times differentiable functions $f = f(x)$ satisfying the conditions

$$\int_0^1 \left(f^{(p)}(x)\right)^2 dx \leq 1,\ f(0) = f(1),\ f'(0) = f'(1), \ldots, f^{(p-1)}(0) = f^{(p-1)}(1)$$

Kolmogorov found that

$$D_{2m-1}\left(F_p^*\right) = D_{2m}\left(F_p^*\right) = \frac{1}{(2\pi m)^p}, \qquad m = 1, 2, \ldots,$$

and showed that for $n = 2m + 1$ the functions 1, $\sqrt{2}\sin 2\pi k x$, $\sqrt{2}\cos 2\pi k x$, $k = 1, \ldots, m$, are optimal.

In paragraph 5 of [K62], [MM-28], Kolmogorov gives in his proof a geometric interpretation, on the basis of which the value $D_n(F)$ was called the nth *diameter* of the set F.

The 1930s were very fruitful in Kolmogorov's scientific activity. The spectrum of topics was very diverse; the number of articles published in those years can be seen from the data: 1931—5 papers, 1932—6, 1933—9, 1934—10, 1935—4, 1936—17, 1937—9, 1938—16 and 1939—5.

Doing research in various branches of mathematics, A. N. Kolmogorov (in the 1930s) obtained a series of results in probability theory of fundamental value in addition to the ones described above.

In the late 1920s and early 1930s there came a number of papers by de Finetti (e.g., [38]), devoted (in modern terms) to the probability properties of random processes $X = (X_t)_{t \geq 0}$ with homogeneous independent increments. In other terms the point lies in the structure of distributions of the so-called infinitely divisible laws.

The random variables $\xi = \xi(\omega)$ with infinitely divisible laws are characterized by the coincidence of their distributions with that of the sum $\xi_{n1} + \cdots + \xi_{nn}$ of independent identically distributed random variables $\xi_{n1}, \ldots, \xi_{nn}$ for every $n \geq 1$. (The importance of this class of infinitely divisible distributions lies in their acting as limit laws, under general conditions, for normalized sums of independent random variables.)

de Finetti [38] suggested a certain rather general formula for the characteristic function $f(t) = Ee^{it\xi}$ of an infinitely divisible random variable ξ. It was he who suggested the following formula for $f(t)$:

$$f(t) = \exp\left\{iat - \frac{\sigma^2}{2}t^2 + c\int(e^{iut} - 1)\,dF(u)\right\}, \tag{23}$$

where $F(u)$ is the distribution function of sizes of jumps, by combining normal and compound Poisson types.

This formula (23) failed to include the general case, but described a certain subclass of the infinitely divisible distributions.

In 1932 Kolmogorov gave an exhaustive reply to de Finetti's problem for the case of a random variable ξ with *finite* second moment, $E\xi^2 < \infty$:

The function $f = f(t)$ is the characteristic function of the infinitely divisible law of a random variable ξ, $E\xi^2 < \infty$, if and only if it may be written as

$$f(t) = \exp\left\{iat + \int_{-\infty}^{\infty} \frac{e^{itx} - 1 - itx}{x^2}\,dK(x)\right\}, \tag{24}$$

where $a \in R$, $K = K(x)$ is a nondecreasing bounded function, and the integrand is equal to $-t^2/2$ at the point $x = 0$.

The general case, which included the possibility of infinite variance, was investigated by Lévy [108] in 1934. In 1937 Khinchin [95] showed that Lévy's result can be obtained by Kolmogorov's method as well.

"The Lévy–Khinchin formula" (to use its now commonly accepted title) for the characteristic function $f = f(x)$ of an infinitely divisible distribution has the form

$$f(t) = \exp\left\{i\alpha t - \frac{t^2\sigma^2}{2} + \int_{-\infty}^{\infty}\left(e^{itx} - 1 - \frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}\Lambda(dx)\right\}, \tag{25}$$

where $\alpha \in R$, $\sigma^2 \geq 0$ and Λ is a certain finite measure on $(R, \mathscr{B}(R))$ with $\Lambda(\{0\}) = 0$. Other formulations are also used, for example,

$$f(t) = \exp\left\{i\alpha t - \frac{t^2\alpha^2}{2} + \int_{-\infty}^{\infty} (e^{itx} - 1 - ith(x))F(dx)\right\},$$

where F is a nonnegative measure such that

$$\int_{-\infty}^{\infty} \min(1, x^2)F(dx) < \infty$$

and $h = h(x)$ is a bounded Borel function with compact support, behaving similarly to x in the neighborhood of zero.

Kolmogorov's remarkable work "Sulla determinazione empirica di una legge di distribuzione" [K43], [PS-15] also dates back to 1933; it has become classical and one of the key points in a whole area of statistics, that of nonparametric goodness-of-fit tests.

The formulation of its main result is simple and beautiful:

Let $\xi = (\xi_1, \xi_2, \dots)$ be a sequence of independent identically distributed random variables with continuous distribution function $F(x) = P(\xi_1 \leq x)$ and let

$$F_n(x; \xi) = \frac{1}{n} \sum_{k=1}^{n} I(\xi_k \leq x)$$

be the empirical distribution function. Then

$$\lim_n P\left\{\sqrt{n} \sup_x |F_n(x; \xi) - F(x)| \leq \lambda\right\} = \mathscr{K}(\lambda), \tag{26}$$

where

$$\mathscr{K}(\lambda) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2\lambda^2 k^2}.$$

It should be recalled for a proper appreciation of this result that Cramér in 1928 [32] and von Mises in 1931 [202] considered the "omega-square" statistic

$$\omega_n^2 = \int_{-\infty}^{\infty} [F_n(x; \xi) - F(x)]^2 \, dx$$

for testing the hypothesis $F = F(x)$ on the basis of the observations $\xi_1, \dots, \xi_n$. However, no exact statement on the asymptotic behavior was obtained. (See the commentary by E. V. Khmaladze in [PS], pages 514–520, on this work of Kolmogorov.)

It evidently follows from Kolmogorov's result (26) that

$$D_n = \sup_x |F_n(x; \xi) - F(x)| \to 0$$

in probability. It may be enlightening that this same fourth volume of *Giorn.*

Istit. Ital. Attuari of 1933 contained the two famous papers by Glivenko [67] and Cantelli [27] bearing the same title, "Sulla determinazione empirica delle legge di probabilitá," which proved the convergence of $D_n \to 0$ with probability 1. [Glivenko did it for continuous $F(x)$ and Cantelli considered the general case.] Later on Smirnov [183], in 1944, and Dvoretzky, Kiefer and Wolfowitz [44], in 1956, showed that the inequality

$$P(D_n > d) \le Ce^{-2nd^2}, \qquad d > 0,\ n \ge 1,$$

holds for Kolmogorov's statistic D_n. By the Borel–Cantelli lemma one may deduce from this that $D_n \to 0$ with probability 1.

In 1936 and 1937 Kolmogorov launched extensive research [K68, K81] on the asymptotic behavior of the transition probabilities for Markov chains with countable state-space. The classification of Markov chains by arithmetic properties of the transition probabilities $p_{ij}^{(n)}$ from state i to state j in n steps (essential and inessential states, indecomposable classes, cyclic subclasses, ...) and the classification by the asymptotic properties of the probabilities $p_{ii}^{(n)}$ (recurrent states, nonrecurrent states, positive and null states, ...) still present a brilliant model of how to resolve the knottiest issues of the possible behavior of such a complicated stochastic object—even in discrete time—as a Markov chain with a countable set of states.

The distinctive breadth of A. N. Kolmogorov's scientific interests is shown in his "more applied" work where the probabilistic approach is directed to problems of biology, genetics, physics, geology, ... Thus in his paper, "On the solution of a biological problem" [K101], [PS-25] dealing with a simple model of the branching random process, Kolmogorov found the asymptotic behavior of the extinction probability as the number of generations increases.

In discussion on genetics in the autumn of 1939 much attention was given to the validity of Mendel's laws (its simplest case means a splitting in the ratio 3 : 1). In this connection Kolmogorov wrote "On a new confirmation of Mendel's law" [K115], [PS-26], where he analyzed the statistical data of N. I. Ermolayeva, a pupil of T. D. Lysenko [*Yarovizatsiya* **2**(23) (1939) 79–86], and her results, saying: "These data, controversial in Ermolayeva's own judgment, provide a brilliant new confirmation of Mendel's laws."

Within certain schematic, but rather general, assumptions Kolmogorov managed to provide, in his work "On the statistical theory of crystallization in metals" [K83], a strict solution of a problem on crystallization rates, and pointed out the "essential significance for metallurgy of the study of the process of crystal growth under random formation of crystallization centers" and marked "certain difficulties in the registration of clashes between the seeds of the crystallizing substance, which are being developed around the separate crystallization centers."

Kolmogorov's formulas on the probability of inclusion of a given point in the already crystallized mass (see (3) in [K83]) and on the number of crystallization centers ((6) and (6a) in [K83]), are still fundamental to the general theory of crystallization in metals.

In 1933 A. N. Kolmogorov and M. A. Leontovich published an article [K42], [PS-14] in a physics journal, "Zur Berechnung der mittleren Brownschen Fläche," where they solved a problem, suggested by S. I. Vavilov, on the expectation ES_t of the area S_t, covered during the length of time t by a circle of radius ρ, whose center moves on the plane as a Brownian particle.

If we reduce ourselves to the main term of the formula for ES_t, then

$$ES_t \sim \frac{4\pi Dt}{\ln(1.26Dt\rho^{-2})}, \qquad Dt\rho^{-2} \gg 1,$$

where D is the diffusion coefficient.

It is worth following Kolmogorov's methods for the solution of this problem, as they are very tightly linked to those developed in "Analytical methods." (It may be interesting to remark that the purely physical part in [K42], [PS-14] was written by Kolmogorov and its purely mathematical part by Leontovich.)

Namely, let $P_L(x, y; t)$ denote the probability of the Brownian particle located at the point (x, y) at time $t = 0$ crossing the boundary Γ of the domain G, containing the point (x, y), at least once during the t-time, with the first crossing falling on a given part L of the boundary Γ. Then $P_L(x, y; t)$ satisfies the "first Kolmogorov equation" (19) with the conditions: $P_L(x, y; 0) = 0$, if $(x, y) \in G$, $P_L(x, y; t) \to 1$, for every $t > 0$, when (x, y) tends to a point of the part L of the boundary Γ, and $P_L(x, y; t) \to 0$, if $t > 0$ and (x, y) tends to a point from $\Gamma \setminus L$. These conditions uniquely determine the function $P_L(x, y; t)$ and enable its determination. Simultaneously the same method was suggested by Pontryagin, Andronov and Vitt [151]. (The Kolmogorov–Leontovich problem on S_t [K42], [PS-14] is now called the "Wiener sausage" problem; see, e.g., Donsker and Varadhan [41], Varadhan [197] and Le Gall [103].)

A brief Kolmogorov note, "Zufällige Bewegungen" (1934) [K57], [PS-19], subtitled "Zur Theorie der Brownschen Bewegung," is devoted to a general description of Brownian motion with inertia, in which physicists were so vigorously interested. In the theory by Einstein [47] and von Smolukhovskii [206], the inertia of the Brownian particle was neglected and thus the particle did not have finite velocity. In 1930 Uhlenbeck and Ornstein [194] were busy developing the theory of Brownian motion with inertia. In this amended theory the trajectories of the particle motion were differentiable (but with infinite acceleration).

Kolmogorov considered this package of problems in general, supposing that the state of the system considered is described by $2n$ coordinates $q = (q_1, \dots, q_n)$ and $\dot{q} = (\dot{q}_1, \dots, \dot{q}_n)$, whose probability density $G(t, q, \dot{q}; t', q', \dot{q}')$, $t < t'$, is determined. Then he shows (in the spirit and on the basis of "Analytical methods") that this density satisfies the corresponding forward equation (of Fokker–Planck type) ((9) in [K57]), a degenerate equation of parabolic type, which owes the beginning of its theoretical development to Piskunov, Kolmogorov's pupil, and others (see Nelson [138]).

Kolmogorov's article [K85], [PS-24] on the general issue of the reversibility of the statistical laws of nature (completed in 1936) may be also included in the series of his works on the theory of Brownian motion.

The essence of the problem under consideration is as follows.

In the thermodynamic interpretation, Brownian motion is nonreversible in the sense that if the number of particles is large and time grows, there results a "levelling" of the probabilities describing the location of the particles. If the time goes down, then the "heterogeneity" of this distribution conversely increases. Schrödinger [169] was perhaps the first to notice the fact that the diffusion process will nevertheless possess certain reversibility, in case the probabilities are fixed both for initial time t_0 and final time t_1, the behavior of the process being considered in the time interval $[t_0, t_1]$.

In his paper [K85], [PS-24] Kolmogorov provided necessary and sufficient conditions for the statistical reversibility (in the sense of "usual" densities of the transition probability coinciding with the "reverse" ones) for the very general situation of an n-dimensional Markov diffusion process.

Kolmogorov's paper of 1937 (together with I. G. Petrovskii and N. S. Piskunov), "Studies of the diffusion equation, with the increasing quantity of the substance and its application to a biological problem" [K82], [MM-38], belongs to his studies of the 1930s on the theory of Brownian motion and diffusion processes and is the first work to establish the existence of wave solutions of parabolic equations and the convergence to them of the solution of the Cauchy problem, when $t \to \infty$.

The diffusion equations considered in [K82], [MM-38] have the form ($K > 0$)

$$\frac{\partial u}{\partial t} - K\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = F(u) \tag{27a}$$

and (if u does not depend on y)

$$\frac{\partial u}{\partial t} - K\frac{\partial^2 u}{\partial x^2} = F(u), \tag{27b}$$

where $F = F(u)$ is a sufficiently smooth function which is determined at $[0,1]$ and satisfies the conditions $F(0) = F(1) = 0$, $F(u) > 0$ for $0 < u < 1$, $F'(0) = \alpha > 0$, $F'(u) < \alpha$, $0 < u \leq 1$.

One searches for the solution $u = u(x, t)$ of the equation satisfying the initial condition $u(x, 0) = f(x)$, where

$$f(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases} \tag{28}$$

It turns out that (27) has a solution of travelling-wave type, that is,

$$u(x, t) = w(x - ct), \tag{29}$$

for all $c \geq \sqrt{2k\alpha}$. The problem given by (27b) and (28) then has solutions which converge as $t \to \infty$, in form and in velocity, to solutions of travelling-wave type.

The commentary by G. I. Barenblatt ([MM], pages 416–420) gives a rather detailed description of contemporary studies in mathematical physics (the theory of combustion in particular), initiated by this work of Kolmogorov, Petrovskii and Piskunov.

Quite apart from the purely mathematical interest of the paper [K82], [MM-38], Kolmogorov in his lectures and talks repeatedly saw the need to emphasize that in its origin and in its formulation of problems this work owed much to biology. In paragraph 1 of [K82], [MM-38], citing the book by Fisher, *The Genetical Theory of Natural Selection* [56], he points to the formulation of the problem of the study of the evolution of the "concentration" $p := p(t, x, y)$ of the biological species under investigation. Under natural assumptions on $p = p(t, x, y)$ one obtains

$$\frac{\partial p}{\partial t} = \frac{1}{4}\rho^2\left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2}\right) + \alpha p(1 - p^2),$$

giving an equation of the type (27), and following that Kolmogorov formulates mathematical problems of purely "biological" interest, one of which is the determination of the rate of advance of the boundary of the area populated by the species under study. It is interesting to note that Fisher published his work [57] on the very same subject and in the very same year. (Of recent "probabilistic" studies linked with [K82] one may cite, for example, the articles by Bramson [26] and Gärtner [60].)

A very small Kolmogorov paper, "La transformation de Laplace dans les espaces linéaires" (1935) [K60], [MM-26], was the first to define the characteristic functional of a probability measure in a Banach space, the generalization of the notion of characteristic function to the infinite-dimensional case. In the same work he defines the normal distributions and moment forms of the n-order, mentions the possibility of generalizing the central limit theorem to linear spaces and emphasizes the significance of these notions for the construction of nonlinear quantum theory, saying:

"If we were working towards a nonlinear quantum theory, it would have been necessary to consider the distributions themselves, or their characteristic functions, or finally the entire set of moments."

Later L. Le Cam (1947) and E. Mourier (1950, 1953) arrived at the notion of the characteristic functional. The modern theory of probability distributions in Banach spaces and detailed references are presented in monographs by Vakhaniya, Tarieladze and Chobanyan [196], Araujo and Giné [10] and Linde [113].

In 1938 there appeared Kolmogorov's work "A simplified proof of the Birkhoff–Khinchin ergodic theorem" [K99], [MM-39]. In the same year Kolmogorov completed his work (together with I. M. Gel'fand) "On rings of continuous functions on topological spaces" [K107], [MM-41], which showed that in quite general situations the algebraic structure alone of the ring of continuous functions, given on the topological space with a sufficiently "good" topology, might determine this topological space (to within homeomorphism). In 1939 came his paper "On the inequalities between upper bounds of the successive derivatives of an arbitrary function on an infinite interval" [K106], [MM-40], dating back to E. Landau and J. Hadamard. Kolmogorov formulates the problem and result as follows:

A function $f = f(x)$ is considered on the real line. Let

$$M_k(f) = \sup_x |f^{(k)}(x)|, \qquad k = 0, 1, \ldots, n;$$

then:

"In order that three positive numbers M_0, M_k, M_n, $0 < k < n$, should correspond to the function $f(x)$ via

$$M_0 = M_0(f), \qquad M_k = M_k(f), \qquad M_n = M_n(f),$$

it is necessary and sufficient that

$$M_k \leq C_{nk} M_0^{(n-k)/n} M_n^{k/n},$$

where C_{nk} are the indicated constants" ([MM-40], page 253).

For the history of the matter, discussion of the results and subsequent studies see the commentaries by V. M. Tikhomirov and G. G. Magaril-Il'yaev ([MM], pages 387–390).

We will return to certain aspects of Kolmogorov's extremely prolific activity in the 1930s (in particular, his works on the teaching of mathematics), but for the present let us remark the following.

In 1930 the State Scientific Council of Narcompros, RSFSR (Educational Ministry of the Russian Soviet Socialist Republic) approved Kolmogorov's academic rank of Professor of Mathematics (PR No. 014075) and in 1935 the Qualification Committee of Narcompros, RSFSR awarded him the scientific degree (dispensing with a dissertation defense) of Doctor of Physics and Mathematics (DT No. 000038).

The Second All-Union (National) Mathematical Congress (Leningrad, 1934) decided to launch a new mathematical journal, *Uspekhi Matematicheskikh Nauk*. (The English translation, published by the London Mathematical Society and begun in 1960, is *Russian Mathematical Surveys*.) From its founding in 1936 to 1944 publication was of separate volumes; in 1946 it started as a periodical. From 1936 to his death Kolmogorov served on the editorial board, and he was Editor-in-Chief from 1946 to 1955 and from 1982 to 1987.

From December 1, 1933 to April 15, 1939 Kolmogorov was Director of the Scientific and Research Institute of Mathematics (Moscow University).

On January 29, 1939 Kolmogorov was elected Full Member (Academician) of the USSR Academy of Sciences. From 1939 to 1942 he was the Academician-Secretary of the Physics and Mathematics Department of the USSR Academy of Sciences and a member of the Presidium of the USSR Academy of Sciences. From 1938 to 1958 he was Head of the Probability Department in the Steklov Mathematical Institute, USSR Academy of Sciences.

The forties (1940–1949). The late 1930s and early 1940s are marked by Kolmogorov's works in the theory of random processes with stationary increments and the related theory of isotropic turbulence, works extraordinary for their conceptual value, profoundness of ideas and diverse possibilities of application.

Kolmogorov noted in [PS], page 473, that his "interests in the spectral theory of stationary random processes developed in connection with works by Khinchin and Slutskii," who had dealt with them in the early 1930s. Kolmogorov referred to his report, "Statistical theory of oscillation with a continuous spectrum," delivered at the General Meeting of the USSR Academy of Sciences in 1947 [K141], [PS-34], where he emphasized the great value of Stieltjes integrals for the general presentation of stationary oscillating processes, comprising almost periodic oscillations as well as those with a continuous spectrum.

The works of 1940, "Kurven im Hilbertschen Raum, die gegenüber einer einparametrigen Gruppe von Bewegungen invariant sind" [K110], [MM-42] and "Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum" [K111], [MM-43], deal with the L^2-theory of random processes $\xi = (\xi_t)$, $-\infty < t < \infty$, with stationary increments and their various subclasses (including processes stationary in the wide sense, Wiener processes and others) from the point of view of the structure of the covariance function

$$B_\xi(\tau_1, \tau_2) = E\left[\xi_{t+\tau_1} - \xi_t\right]\left[\overline{\xi_{t+\tau_2} - \xi_t}\right]$$

(Theorem 2 in [K110], [MM-42]) and of the possibility of the spectral representation of the process $\xi = (\xi_t)$, $-\infty < t < \infty$, which Kolmogorov (see Theorem 3 in [K110], [MM-42]) gives as

$$\xi_t = \int_{-\infty}^{\infty} (e^{i\lambda t} - 1)\Phi(d\lambda) + x_0 + x_1 t. \tag{30}$$

If the process $\xi = (\xi_t)$, $-\infty < t < \infty$, is itself stationary, as well as having stationary increments, then $x_0 = \int_{-\infty}^{\infty}\Phi(d\lambda)$, $x_1 = 0$ and the spectral representation of the stationary random process $\xi = (\xi_t)$, $-\infty < t < \infty$, can be derived from (30) as a Stieltjes integral,

$$\xi_t = \int_{-\infty}^{\infty} e^{i\lambda t}\Phi(d\lambda), \tag{31}$$

with respect to a random measure $\Phi(\Delta)$ with orthogonal increments ($E\Phi(\Delta_1)\overline{\Phi}(\Delta_2) = 0$, $\Delta_1 \cap \Delta_2 = \varnothing$). This was also discovered independently by Cramér (1942, [33]) and Maruyama (1949, [126]). See also Loève [116].

The paper [K111], [MM-43] directly adjoins [K110], [MM-42] and deals with some special cases of processes $\xi = (\xi_t)$, $-\infty < t < \infty$, with stationary independent increments. Actually Kolmogorov considers processes ξ with self-similarity, which means that for any $k \neq 0$ there exists a similarity transformation A_k such that for all t,

$$\xi_{kt} = A_k \xi_t.$$

As it turns out, the "structural" function $B_\xi(\tau_1, \tau_2)$ of these processes can be presented as

$$B_\xi(\tau_1, \tau_2) = c\left[|\tau_1|^\gamma + |\tau_2|^\gamma - |\tau_1 - \tau_2|^\gamma\right], \tag{32}$$

where c and γ are real constants, satisfying the inequalities $c \geq 0$, $0 \leq \gamma \leq 2$. [The zero mean Gaussian process with covariance function given by (32) is now called fractional Brownian motion of index $\frac{1}{2}\gamma$.]

It is to be noted that in recent years many papers applicable for example to statistical physics (see Sinai [181] and Taqqu and Levy [192]) have been devoted to random processes with the property of self-similarity (for background on self-similarity see, e.g., Vervaat [198]).

These works on processes with stationary increments were followed by Kolmogorov's classical works on stationary (in the wide sense) random processes where (as in [K110, K111]) he extensively employed Hilbert space techniques, as reflected in the title of his 1941 work, "Stationary sequences in Hilbert space" [K116], [PS-27]. Here he introduced new concepts (subordination of one stationary sequence to another, regularity, singularity and minimality) which gave rise to many subsequent studies of vector processes with continuous time (see [161, 162]).

The notion of the stationary sequence $\eta = (\eta_n)$, $n = 0, \pm 1, \ldots$, being subordinate to another stationary sequence $\xi = (\xi_n)$, $n = 0, \pm 1, \ldots$, means that they are stationarily related and the closed linear subspace $H(\xi)$ generated by the elements ξ_n, $n = 0, \pm 1, \ldots$, comprises all the elements η_n, $n = 0, \pm 1, \ldots$. The surprise of Kolmogorov's result is in the possibility for a subordination to be expressed in purely spectral terms. Namely, the sequence η is subordinate to ξ if and only if there exists a function $\phi(\lambda) \in L^2(F_{\xi\xi})$ such that the spectral functions $F_{\eta\eta}(\lambda)$ and $F_{\xi\eta}(\lambda)$ in the representations of the covariance functions

$$B_{\eta\eta}(n) = \int_{-\pi}^{\pi} e^{i\lambda n}\, dF_{\eta\eta}(\lambda), \qquad B_{\xi\eta}(n) = \int_{-\pi}^{\pi} e^{i\lambda n}\, dF_{\xi\eta}(\lambda)$$

satisfy

$$F_{\eta\eta}(\lambda) = \int_{-\infty}^{\lambda} |\phi(u)|^2\, dF_{\xi\xi}(\lambda), \qquad F_{\xi\eta}(\lambda) = \int_{-\infty}^{\lambda} \phi(u)\, dF_{\xi\xi}(\lambda),$$

where

$$B_{\xi\xi}(n) = \int_{-\pi}^{\pi} e^{i\lambda n}\, dF_{\xi\xi}(\lambda), \qquad B_{\xi\eta}(n) = E\xi_{n+k}\bar{\eta}_k.$$

The property of singularity (in other terms—determinacy) of the sequence $\xi = (\xi_n)$, $E\xi_n = 0$, $n = 0, \pm 1, \ldots$, means that the space $H(\xi)$ coincides with $H_{-\infty}(\xi) = \bigcap_n H_n(\xi)$, where $H_n(\xi)$ is the closed linear space generated by the random variables ξ_k, $k \leq n$.

The property of regularity (in other terms—pure nondeterminacy) means that the space $H_{-\infty}(\xi)$ is trivial.

Kolmogorov employs the results on boundary properties of functions analytic in the disc and obtains his well-known result:

For the nondegenerate stationary sequence ξ to be regular, it is necessary and sufficient that the spectral function $F_{\xi\xi}(\lambda)$ possess a density $f_{\xi\xi}(\lambda)$ such that

$$\int_{-\pi}^{\pi} \ln f_{\xi\xi}(\lambda) > -\infty. \tag{33}$$

Kolmogorov defines the minimality of the sequence ξ to mean that the space $\hat{H}(\xi)$, which is the minimal closed subspace generated by all ξ_n, $n \neq 0$, does not coincide with the minimal linear closed subspace $H(\xi)$ generated by all ξ_n, $n = 0, \pm 1, \ldots$.

A. N. Kolmogorov states:

The stationary sequence ξ is minimal if and only if there exists a spectral density $f_{\xi\xi}(\lambda)$ such that $f_{\xi\xi}(\lambda) > 0$ almost everywhere w.r.t. Lebesgue measure and

$$\int_{-\pi}^{\pi} \frac{d\lambda}{f_{\xi\xi}(\lambda)} < \infty. \tag{34}$$

Besides, if these conditions are fulfilled, then

$$d_\xi \equiv \inf_{\hat{\xi}_0 \in \hat{H}(\xi)} E\left[\xi_0 - \hat{\xi}_0\right]^2 = (2\pi)^2 \left[\int_{-\pi}^{\pi} \frac{d\lambda}{f_{\xi\xi}(\lambda)}\right]^{-1}. \tag{35}$$

This work is closely adjoined and followed by Kolmogorov's "Interpolation and extrapolation of stationary random sequences" [K117], [PS-28], where he says in the Introduction that the paper "sets the spectral conditions for the possibility of extrapolating, to within any given accuracy, a stationary random sequence, with a sufficiently large number of observations."

In this work Kolmogorov gives the first results on the value of the error in problems of extrapolation and interpolation, as well as providing a precise formulation of such problems.

These results by A. N. Kolmogorov and the results by Wiener [210] (see also Doob [43], Chapter XII) created an entirely new branch in the theory of random processes, with wide application in science and technology.

Kolmogorov writes

$$\sigma_E^2(n, m) = \inf_{(a_1, \ldots, a_n)} E\left[\xi_{t+m} - \tilde{\xi}_{[t-1, \ldots, t-n]}\right]^2$$

for the minimal possible error of prediction of ξ_{t+m}, $m \geq 0$, by values $\tilde{\xi}_{[t-1,\ldots,t-n]}$ of the type $a_1\xi_{t-1} + \cdots + a_n\xi_{t-n}$, $n > 0$, and thus via $\sigma_E^2(m) = \lim_{n\to\infty}\sigma_E^2(n, m)$ he obtains the explicit formula for $\sigma_E^2(m)$, which is expressed in spectral terms. He also shows here that, if the integral $\int_0^\pi \log f_{\xi\xi}(\lambda)\, d\lambda = -\infty$, then the error of prediction $\sigma_E^2(m)$ vanishes for all $m \geq 0$. In the case of regular sequences, when the integral $\int_0^\pi \log f_{\xi\xi}(\lambda)\, d\lambda$ is finite, Kolmogorov gives an explicit formula for the error $\sigma_E^2(m)$ (Theorem 2 in [K117], [PS-28]).

In the interpolation problems Kolmogorov introduces the value

$$\sigma_I^2(n) = \inf E\left[\xi_t - \hat{\xi}_{[t-n, t+n]}\right]^2$$

of the minimal possible error of the interpolation of ξ_t by values

$$\hat{\xi}_{[t-n, t+n]} = \sum_{k=-n}^{n} a_k \xi_{t+k}$$

and finds for $\sigma_I^2 = \lim_{n\to\infty} \sigma_I^2(n)$ that if the integral

$$R = \frac{1}{\pi}\int_0^{\pi} \frac{d\lambda}{f_{\xi\xi}(\lambda)}$$

is equal to $+\infty$, then $\sigma_I^2 = 0$, and if $R < \infty$, then $\sigma_I^2 = R^{-1}$ (Theorem 3, [K117], [PS-28]).

Kolmogorov's report [K141], [PS-34] to the General Meeting of the USSR Academy of Sciences, 1947, mentioned above, was described by A. M. Yaglom (see his commentary in [PS], pages 491–496) as "the first popular review of the spectral theory of stationary random processes, which is one of the most important branches of the mathematical theory of random functions, and has been developed only recently (with active participation by Kolmogorov himself) and was hardly known to anyone outside a narrow circle of experts." (In this commentary by Yaglom the reader will find detailed historical and bibliographical data relating to the spectral theory of stationary processes; see also [161, 162, 212, 43, 35].)

One can hardly overestimate the importance of Kolmogorov's work of the early 1940s in the *theory of turbulence*, which promoted further development of the concept and the theory and applications of the local structure of turbulent motions.

Kolmogorov said, commenting on his works on turbulence ([MM], page 421):

"My interest in turbulent processes in liquids and gases developed in the late 1930s. Immediately I became aware that the newly emerging theory of random functions of many variables (random fields) was to become the major mathematical tool of research in turbulence. Besides, it soon became clear that one could hardly rely upon the creation of a pure theory closed in itself. The lack of such theory will mean reliance on hypotheses derived from the processing of experimental data. It was also essential to employ talented staff, who could manage work in such a mixed field combining theory with experimentation.

"I was lucky with the latter: A. M. Obukhov, who had been posted in Moscow University from Saratov University, wrote his diploma (1939) and postgraduate thesis papers under my scientific supervision. Almost simultaneously M. D. Millionshchikov became my postgraduate student in the Moscow Aviation Institute. Later A. S. Monin and A. M. Yaglom also became my postgraduate students.

"In 1946 O. Yu. Shmidt suggested that I should head the Turbulence Laboratory in the Institute of Theoretical Geophysics, USSR Academy of Sciences. In 1949 this post was passed to Obukhov. I was not engaged in experimentation myself, but I worked extensively with other researchers on computation and graphical processing of the data."

The presence of chaotic pulsations of velocity $U(x, t)$ and pressure $P(x, t)$ and of other hydrodynamical characteristics in the flow of liquids and gases (called turbulence) makes the study of individuals fields of turbulent motion hardly feasible at all. This very aspect gave value and interest to the statistical description of flows by O. Reynolds, the founder of the theory, who had been aware of it even at the end of the 19th century. However, his suggested averaging over a given interval of space or time proved rather inconvenient, because of the difficulties in obtaining sufficiently simple and reliable equations for the average (mean) field.

Kolmogorov took the averaging in a probabilistic sense, that is, averaging over ensembles (package averaging). Thus he suggested viewing the fields of hydrodynamic characteristics as random functions of spatial and time coordinates, as is now commonly accepted.

His in-depth intuition in physics helped Kolmogorov to distinguish the general qualitative and quantitative laws which determine the stochastic nature of the small-scale pulsation for developed turbulence with sufficiently large Reynolds number on the basis of two hypotheses of similarity, formulated in Kolmogorov's famous paper, "The local structure of turbulence in incompressible viscous liquid for very large Reynolds' numbers," written in 1940 ([K119], [MM-45]).

These hypotheses enabled one to state the fundamental quantitative relations, including first of all Kolmogorov's famous "law of two thirds," saying that the average (mean) square of the difference of the velocities at two points, located at distance r (neither too big nor too small), is proportional to $r^{2/3}$.

Kolmogorov's so-called longitudinal and transverse structural functions $B_{dd}(r)$ and $B_{nn}(r)$ of the fields of velocities have been extensively tested experimentally, and the "law of two thirds" ($B_{dd}(r) \sim r^{2/3}$) as well as the formula $B_{nn}(r) \sim \frac{4}{3}B_{dd}(r)$ were verified for considerable ranges of values of r. (For more details see [K119], [MM-45]; [MM], pages 421–433.)

Kolmogorov's works [K119], [MM-45] and [K121], [MM-46] on turbulence were further developed in his report to the International Colloquium on the Mechanics of Turbulence, Marseille, 1961 [K306], [MM-58], and in [K307], adjacent to the paper by A. M. Obukhov [141]. Kolmogorov suggests replacement of his two hypotheses of similarity (from [K119]) by two more detailed ones, referring to the normalized velocity difference, and also supplements them by a third one postulating the logarithmic normality of the probability distributions of the dissipation of energy ε_r (obtained as a result of averaging on the sphere of radius r) and the linearity of the variance $\log \varepsilon_r$ from $\log(L/r)$, where L is the characteristic scale of length for the flow considered.

These three hypotheses led to the refinement of the "law of two thirds," that is, to the new formula $B_{dd}(r) \sim r^{2/3}(L/r)^{-k}$, which already takes into account an early remark of L. D. Landau on Kolmogorov's paper [K119], on the impossibility of neglecting the variation of energy dissipation with the growth of L/r. (For more details see [MM], pages 349 and 428.)

Summarizing Kolmogorov's contribution to the theory of turbulence, one may quote the final lines from "Kolmogorov flow and laboratory simulation of it" by A. M. Obukhov [142]:

"The personal contribution of Kolmogorov to the study of turbulence and his ideas relating to the general theory of dynamical systems are fundamental reference points in the development of investigations of the most complex phenomenon in nature, namely, turbulence, in connection with diverse areas of knowledge."

Kolmogorov was very responsive both as a universal mathematician and as an applied researcher; he had an extraordinary gift of penetrating into the very essence of a given problem, of selecting its basic and vital aspects and bringing clarity to contentious issues.

This is vividly illustrated by Kolmogorov's works in ballistic theory, which date back to the Great Patriotic War (1941–1945). Kolmogorov's paper, "The determination of the center of dispersion and the measure of accuracy resulting from a limited number of observations" (handed to the printer on September 15, 1941) [K126], says, in particular, that the author "was requested to settle an artillery dispute on how to estimate the measure of accuracy from experimental data." Kolmogorov modestly remarks that the article does not claim anything more than a certain methodological value and then proceeds to a critical comparison of the different approaches.

In collaboration with the Steklov Mathematical Institute of USSR Academy of Sciences, the Mathematics and Mechanics Faculty of Moscow University, the Artillery, Scientific and Research Marine Institute and others Kolmogorov undertook a profound theoretical and computational investigation of the efficiency of firing systems. One may appreciate the nature of these extensive studies from Kolmogorov's two papers, "Number of hits in several shots and the general principles of estimating the efficiency of firing systems" [K129] and "The artificial dispersion of single-shot hitting and one-dimensional dispersion" [K130].

The work [K129] considers the number μ of hits in a group of n shots, $\mu = 0, 1, \ldots, n$. Kolmogorov writes $P_m = P(\mu = m)$, $R_m = P(\mu \geq m)$ and $E\mu$ for the expectation of the hitting numbers and suggests a definition of the "efficiency characteristics of the firing system." He says that the usual reasoning about the comparative advantages and disadvantages of "estimation by the expectation" and "estimation by the probability" often lacks sufficient sharpness. He also poses the problem of whether a set of probabilities $P_0, P_1, \ldots, P_n$, characterizing the firing system through the probability distribution of the number of hits, can be "replaced by a reliable single value $W = f(P_0, P_1, \ldots, P_n)$, which could be called an efficiency characteristic."

After his analysis of the subject (paragraph 1, [K129]), Kolmogorov obtains a series of explicit formulae for the probabilities P_k and supplies practically convenient approximations and tables of their accuracy.

The other group of problems in the paper is to select the best firing system by classifying the factors which affect the fire results and to resolve the problem of "artificial dispersion."

Let us denote by $p_i = p_i(\alpha_i, \beta_i)$ the probability of a hit in the ith shot depending on the azimuth α_i and the aim β_i. Let $(\bar{\alpha}_i, \bar{\beta}_i)$ be a combination (usually unique) of values α_i and β_i which maximizes the probability of a hit,

$$\max p_i(\alpha_i, \beta_i) = p_i(\bar{\alpha}_i, \bar{\beta}_i),$$

and let

$$\bar{\alpha} = (\bar{\alpha}_1, \ldots, \bar{\alpha}_n), \qquad \bar{\beta} = (\bar{\beta}_1, \ldots, \bar{\beta}_n).$$

The question is whether

$$\max W = W(\bar{\alpha}, \bar{\beta}),$$

i.e., the maximum fire efficiency is sufficiently ensured by maximizing the probability of a hit in every shot.

The article says that in two special cases, when

$$W = E\mu$$

and when W is presented as

$$W = c_1 P_1 + \cdots + c_n P_n, \qquad c_i \geq 0,$$

and the events B_i, $i = 1, \ldots, n$ (hitting the target in the ith shot), are independent, then the property $\max W = W(\bar{\alpha}, \bar{\beta})$ will hold and, consequently, the optimal firing system in these two cases is to maximize the probability of a hit in every separate (single) shot.

However, generally speaking this is not correct for other fire efficiency criteria W. Therefore "in order to achieve the maximum overall fire efficiency one should deliberately deviate now and then from the maximum probability of every shot hitting." This is firing with *artificial dispersion* and it will prove useful in such a typical situation as:

"It is most essential to achieve at least a small number of hits, considerably less than the total number of shots n."

In his second work [K130], Kolmogorov considers the case of "artificial dispersion, where one hit and one-dimensional dispersion suffice for the given purpose, for example, fire at a narrow long strip (e.g., a bridge) which is perpendicular to the firing surface.

In 1949 Kolmogorov wrote a paper, "The solution of a probability problem related to the question of the formation of strata" [K154], [PS-37]. A. B. Vistelius says in his commentary ([PS], pages 527–531):

"At that time the geological sciences had hardly possessed such notions as random variable, probability distribution function and sequence of values of a random variable. It was a first step in the construction of such a scientific foundation of geology to introduce stochastic methods. This basic reconstruction, which subsequently gave birth to mathematical geology, was greatly stimulated by this article and Kolmogorov's personal advice and comments during the period from 1945 to 1950."

In 1946 there appeared the book by H. Cramér, *Mathematical Methods of Statistics* [34]. A champion of mathematical and statistical education and an enthusiast of the development of statistical research in the USSR, Kolmogorov wrote a detailed introduction to the Russian edition and edited it [K149]. He says in this Introduction:

"The existing courses in mathematical statistics had been erected on a theoretical basis, which definitely fails to meet modern requirements" and "studies of the specific issues of mathematical statistics have outgrown the old level of presentation of the mathematical and probabilistic prerequisites" and therefore Cramér's book is "an attempt to present systematically the fundamental issues of mathematical statistics from a rather up-to-date position."

With a view to stimulating statistical research in the USSR, Kolmogorov delivered the following reports at the Second All-Union Conference on Mathematical Statistics, Tashkent, 1948: "Basic problems of theoretical statistics" [K156] and "The real meaning of the analysis of variance" [K157]. In March 1950 Kolmogorov completed a fundamental work, "Unbiased estimates" [K164], [PS-38], where he systematically analyzed the properties of unbiased estimates and the different methods of their construction by means of sufficient statistics, and also described significant application of unbiased estimates in problems of statistical control and mass industrial quality control.

Kolmogorov's "Unbiased estimates" and the subsequent "Statistical quality control with the allowed number of defective items equal to zero" [K189] initiated vast theoretical and practical probability research in sampling. (See commentaries by Yu. K. Belyaev and Ya. P. Lumelskii in [PS], pages 522–523.)

Kolmogorov is rightfully mentioned among the founding fathers of the modern theory of *branching* random processes. (The very notion was introduced at Kolmogorov's university seminar, 1946–1947.)

Though some specific problems related to simple models of branching processes had already been considered by Fisher [56], Steffenson [186], Leontovich [105] and Kolmogorov himself [K101], [PS-25], the vigorous development of the theory of branching random processes, an independent new area of probability, was launched by Kolmogorov's "Branching random processes" (in collaboration with N. A. Dmitriev) [K139], [PS-32] and "The computation of the final probabilities for branching random processes" (in collaboration with B. A. Sevast'yanov) [K140], [PS-33].

The first post-war publication on branching processes in the West seems to be Harris (1948) [75]. For historical background see Kendall [90, 91] (note the 1944 paper of Hawkins and Ulam—unpublished because of its Los Alamos security classification—and Harris' Princeton dissertation of 1947).

The works [K139] and [K140] consider the patterns of Markov branching processes with several types of particles for both discrete and continuous time. Later on research switched to more complicated patterns, which studied the dependence of reproduction on the age of particles, their location, energy, etc. Apart from giving the state of the art in branching processes at the time of publication, the commentary by B. A. Sevast'yanov in [PS], pages 485–486, and also [12, 76, 170] contain rich material on the various applications of the theory of branching random processes to biology, chemistry, physics, technology, etc.

Probability theory owes a great debt to the book by B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables* (1949) [K151] (English edition, 1954), devoted to the theory of limit theorems, whose centerpiece is the notion of infinitely divisible and stable laws. In their introduction to this book Gnedenko and Kolmogorov describe the problems of limit theorems which brought them to these laws as follows.

If $\xi_1, \xi_2, \ldots$ is a sequence of independent identically distributed random variables, $S_n = \xi_1 + \cdots + \xi_n$, then it is natural to ask a general question about the conditions under which the limit property

$$P\left(\frac{S_n - A_n}{B_n} \leq x\right) \to V(x), \quad \text{as } n \to \infty,$$

holds for some choice of constants A_n and B_n, and which limit distributions $V(x)$ may arise here.

This problem was fully solved by Khinchin, who established that the possible laws $V(x)$ are the so-called "stable" distributions and, as [K151] says: "the range of real applied problems where they might play an essential part could in time prove to be rather extensive," as subsequent developments have borne out.

Taking up the issue of infinitely divisible distributions, the authors especially emphasized the importance of the scheme of series of random variables

$$\xi^n = (\xi_{n1}, \ldots, \xi_{nn})$$

that are independent within each row, as this scheme may "contain in itself all the meaningful and practically valuable limit theorems, relating to sums of independent terms and leading to limit laws very much unlike the normal ones."

In asking under which conditions $S_n = \xi_{n1} + \cdots + \xi_{nn}$ may have the limit property

$$P\left(\frac{S_n - A_n}{B_n} \leq x\right) \to V(x), \quad \text{as } n \to \infty,$$

and which limit laws $V(x)$ may arise, Gnedenko and Kolmogorov restricted themselves to arrays satisfying the negligibility condition

$$\sup_{1 \leq k \leq n} P(|\xi_{nk} - a_{nk}| \geq \varepsilon B_n) \to 0, \quad \text{as } n \to \infty,$$

for some constants a_{nk}. Under this assumption they gave a complete solution of the limit problems considered, which has been taught to generations of students of probability.

In Chapter 8 Gnedenko and Kolmogorov consider results on the speed of convergence of

$$V_n(x) = P\left(\frac{S_n - A_n}{B_n} \leq x\right)$$

to $V(x)$. They note that P. L. Chebyshev emphasized, surprisingly early, the significance of an asymptotic expansion for $V_n(x)$ and provided a central limit theorem (without however a precise proof) with such an asymptotic expansion with terms of order of $n^{-1/2}$.

The book *Limit Distribution for Sums of Independent Random Variables* by B. V. Gnedenko and A. N. Kolmogorov was awarded the P. L. Chebyshev Prize (on December 14, 1951) of the USSR Academy of Sciences.

In the late 1940s Kolmogorov began his significant work in *The Large Soviet Encyclopedia* (2nd edition), as head of its mathematical section. He made up a glossary, selected the contributors, edited articles and also made contributions himself in the most diverse mathematical branches. From 1949 to 1958 he wrote 93 articles (1949—6, 1950—20, 1951—7, 1952—27, 1953—14, 1954—9, 1955—5, 1956—2 and 1958—3). Among Kolmogorov's contributions to the encyclopedia one must especially distinguish his famous article "Mathematics" (*The Large Soviet Encyclopedia*, 2nd edition, volume 26, pages 464–483, 1954 [K247]), where "he briefly outlined the history of mathematics, marked the milestones of its development and suggested an ingenious scheme of its periodization" [21].

In 1941 Kolmogorov and Khinchin were awarded the Stalin Prize.

In 1945 Kolmogorov was awarded a medal "For Valiant Labor in the Great Patriotic War," and in 1944 and 1945 the Order of Lenin.

In the fall of 1942 Kolmogorov married Anna Dmitrievna Egorova, a school friend of his.

The fifties (1950–1959).

The general theory of Hamiltonian systems.
Information theory.
The ergodic theory of dynamical systems.
ε-entropy.
The superposition of functions and Hilbert's 13th problem... .

These are the mathematical branches in which Kolmogorov was doing research, those which are definitively associated with his name, which gave rise to entire scientific fields and schools.

Kolmogorov's works in the theory of dynamical systems consist of two cycles. The first one was directly initiated by and is related to problems of classical mechanics ([K227, K242, K243], or 51, 52 and 53 in [MM]) and the second one deals with information theory.

Kolmogorov says, commenting on his works in classical mechanics ([MM], page 433):

"My works in classical mechanics were influenced by J. von Neumann's papers (e.g., [204]) on the spectral theory of dynamical systems and most of all by the classical work of N. M. Krylov and N. N. Bogolyubov (1937) [102].

"I was then deeply intrigued by the question: Which are the ergodic sets (in the sense of Krylov and Bogolyubov) in the dynamical systems of classical mechanics, and which types of them may fill sets of positive measure (the question has not been yet solved)? A special seminar on the study of some specific examples was set up to accumulate concrete information on the subject. My ideas in these and adjacent areas found a wide response from the young Moscow mathematicians."

At the closing meeting of the International Mathematical Congress, Amsterdam, 1954, Kolmogorov delivered a report, "The general theory of dynamical systems and classical mechanics" [K243], [MM-53], which refers to "the basic problem of dynamics" (in H. Poincaré's terms), that is, to research on the quasiperiodical motions of Hamiltonian systems under small perturbations of the Hamiltonian. [The "smallness" of the variation of the Hamiltonian $W(p)$ is understood as a transition to the Hamiltonian $W(p) + \theta S(q, p, \theta)$ with small parameter θ.]

The remarkable theoretical result by A. N. Kolmogorov says that quasiperiodical motions will be preserved for the case of general position, i.e., $\det(\partial^2 W/\partial p^2) \neq 0$, and for the majority of the initial conditions.

Kolmogorov's theory and its subsequent development provided the long-awaited solution to many problems. For instance, it implies the stability of rapid rotation of an asymmetrical solid body about a fixed point, the stability of the motion of an asteroid of negligible mass in the three-body problem; it implies the stability of most magnetic surfaces under small variation of the magnetic field in toroidal systems.

Speaking of his methods, Kolmogorov explained (see his report [K243], [MM-53]) that the proof was based on developing the "idea of the possibility of avoiding abnormally 'small denominators' in the computation of orbit perturbations, which was widely disputed in celestial mechanics." (See, e.g., Siegel and Moser [178].)

(We know well the following example of the "small denominators": $2\omega_1 - 5\omega_2 = 0.007$, where $\omega_1 = 299''.1$ and $\omega_2 = 120''.5$ are the frequencies of movement for Jupiter and Saturn. These "small denominators" lead to big mutual perturbations in the movements of the planets as the expressions $m\omega_1 + n\omega_2$ appear as denominators in the series

$$\sum_{m,n\neq 0} a_{mn} \frac{\exp[i(m\omega_1 + n\omega_2)]}{m\omega_1 + n\omega_2}$$

in the theory of perturbation.)

Kolmogorov's method circumventing these "small denominators" was subsequently improved by V. I. Arnol'd, Kolmogorov's pupil, and J. K. Moser, and it is presently known as the KAM theory, for Kolmogorov–Arnol'd–Moser. (See details and references in Arnol'd [MM, pages 433–444] and the book by Abraham and Marsden [1].)

In the second cycle of his works on the theory of dynamical systems Kolmogorov applied certain ideas of information theory to research on the ergodic properties of these systems.

In the early 1950s the work of Shannon stimulated Kolmogorov's direct approach to the problems of information theory. Kolmogorov said about it ([K316], Introduction):

"Shannon's contribution to pure mathematics was denied immediate recognition. I can recall now that even at the International Mathematical Congress, Amsterdam, 1954, my American colleagues in probability seemed rather doubtful about my allegedly exaggerated interest in Shannon's work, as they believed it consisted more of techniques than of mathematics itself. Nowadays such views hardly need a denouncement.

"However Shannon did not provide rigorous mathematical justification of the complicated cases and left it all to his followers. Still his mathematical intuition is amazingly correct."

Thus it became evident that information theory was in need of a mathematical basis. The first steps in this direction were taken by Khinchin [97, 98], who proved the basic theorems of information theory for the discrete case, and by I. M. Gel'fand, A. N. Kolmogorov and A. M. Yaglom (see "On a general definition of the amount of information" [K267], [IA-2] and "The amount of information and entropy for continuous distributions" [K276], [IA-4]), who took the general case, established the general properties of the quantity of information in the Gaussian case and formulated the coding theorem for the transmission of messages with given precision.

In 1956 Kolmogorov reported his "The theory of transmission of information" [K272], [IA-3] to the session of the USSR Academy of Sciences devoted to the scientific problems of industrial automation. The report contained the basic concepts of information theory and clarified the limits of its applicability.

All these works "set a tradition of presenting the results of information theory in compliance with high standards of mathematical rigor, which was carefully observed by both categories of researchers: mathematicians and engineers."

His understanding of Shannon's ideas in information theory brought Kolmogorov to the unexpected combination of this theory with his theory of approximation and theory of algorithms, which dated back to the 1930s. Let us discuss this further.

Shannon described the uncertainty measure of discrete messages ξ, which assume discrete values $x_1, x_2, \ldots$ with probabilities $p_1, p_2, \ldots$, by the application of "entropy" $H(\xi)$, as

$$H(\xi) = -\sum_i p_i \log p_i.$$

He also defined the notion of the information $I(\xi, \eta)$ contained in the object ξ with respect to η by

$$I(\xi, \eta) = \sum_{i,j} p_{ij} \log \frac{p_{ij}}{p_i q_j}$$

in the case of discrete random variables ξ and η with

$$p_{ij} = P(\xi = x_i, \eta = y_j), \qquad p_i = P(\xi = x_i), \qquad q_j = P(\eta = y_j)$$

and

$$I(\xi, \eta) = \iint \log \frac{p(x, y)}{p(x)q(y)} p(x, y)\, dx\, dy$$

when the variables ξ and η have joint density $p(x, y)$ and one-dimensional densities $p(x)$ and $q(y)$, respectively.

In the case of continuous messages all the natural analogues to the Shannon entropy lead to infinite values. In this regard Kolmogorov repeatedly emphasized that for arbitrary messages the basic notion is the quantity of information $I(\xi, \eta)$ of one object ξ w.r.t. η, rather than the entropy.

Starting from this idea Kolmogorov defines the ε-entropy $H_\varepsilon(\xi)$ of a random object ξ as

$$H_\varepsilon(\xi) = \inf I(\xi, \eta),$$

where the infimum is taken (under the fixed distribution P_ξ of the object ξ) over all pairs of random variables (ξ, η) which satisfy the restriction that their joint distribution $P_{\xi\eta}$ belongs to a given class W_ε, depending on a parameter ε (for example,

$$W_\varepsilon = \{(\xi, \eta): E\rho(\xi, \eta) \le \varepsilon\},$$

where ρ is a certain metric in the space of values of the objects considered).

The values $H_\varepsilon(\xi)$ had already been considered by Shannon as "rate of message generation." Kolmogorov writes in [K273], [IA-3]:

"Though the selection of a new name will not change the nature of the problem, I will nevertheless venture to do so, as it emphasizes the wider interest of the notion and its deeper similarity to the usual notion of entropy... I would especially note the interest of research on the asymptotic behavior of the ε-entropy as $\varepsilon \to 0$. The cases previously studied are nothing more than very special ones of the possible regularities. My paper [K266], though using different terminology, may cast light on the emerging prospects."

The paper [K266] is the work "On certain asymptotic characteristics of totally bounded metric spaces," published in 1956. In this work Kolmogorov introduces the notion of the ε-entropy $\mathscr{H}_\varepsilon(C)$ of a nonrandom object C, a set in a metric space (X, ρ), defined as the binary logarithm of $N_\varepsilon(C)$, the minimal number of sets of diameter not bigger than 2ε which can cover C.

Along with the ε-entropy $\mathscr{H}_\varepsilon(C)$, subsequently called absolute entropy, Kolmogorov also introduces the relative ε-entropy $\mathscr{H}_\varepsilon(C, X)$ and determines it as the binary logarithm $N_\varepsilon(C, X)$ of the minimal number of elements in an ε-net from X for the set C.

Values $\mathscr{H}_\varepsilon(C)$ and $\mathscr{H}_\varepsilon(C, X)$ are generally constructed in the same way as the so-called "Kolmogorov diameters," introduced by him in 1936 [K62]. For example, the binary logarithm of the inverse function of the diameters

$$\mathscr{E}_N(C, X) = \inf_{A \in \Sigma_N} \sup_{x \in C} \inf_{y \in \Lambda} \|x - y\|,$$

where Σ_N is the collection of N-point approximating sets, coincides with the ε-entropy $\mathscr{H}_\varepsilon(C, X)$.

As had happened to almost all of Kolmogorov's notions, the ε-entropy techniques of estimating the "metric massiveness" of functional classes and spaces laid the foundation for an entirely new area of research in approximation theory. (See the commentary by V. M. Tikhomirov [IA], pages 262–269.)

In 1958 a new Kolmogorov paper, "A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces" [K280], [IA-5] appeared. (A slightly revised version was published later in [K468].) In this work the ideas of information theory led Kolmogorov to the introduction of entropy-characteristics to the theory of dynamical systems (this is "the second cycle" mentioned above on page 911).

A dynamical system is understood in [K280], [IA-5] as a monoparametric group $\{S^t\}$ of transformations of a probability space $(X, \mathscr{X}, \mu)$ which preserve measure. As in the theory of stationary random processes, Kolmogorov introduces the notion of quasiregular dynamical systems, now known as K-systems. The significance of this notion for ergodic theory was revealed a few years later, when Ya. G. Sinai showed that many classical dynamical systems which have nothing in common with the theory of probability are K-systems.

For quasiregular dynamical systems the notion of entropy was introduced by Kolmogorov in [K280], [IA-5]. (A slightly modified and sharpened version of the corresponding definition is given in [K468].) Shortly after that Sinai provided a definition of entropy applicable to an arbitrary dynamical system. (For more details see the commentary by Sinai in [IA], pages 275–279.)

In the case of discrete time ($t = 1, 2, \dots; S^1 = S$) the commonly accepted definition of the "Kolmogorov–Sinai entropy" is as follows.

Let $A = \{A_1, \dots, A_N\}$ be a finite partition of X, i.e., $\bigcup A_i = X$, $A_i \cap A_j = \varnothing$, $i \neq j$. This partition defines the entropy

$$H(A) = -\sum_{i=1}^{N} \mu(A_i) \log \mu(A_i).$$

Write

$$\mu_{i_1, \dots, i_r} = \mu\left(A_{i_1} \cap SA_{i_2} \cap \cdots \cap S^{r-1}A_{i_r}\right)$$

and

$$H_r(A) = -\sum_{i_1, \dots, i_r} \mu_{i_1, \dots, i_r} \log \mu_{i_1, \dots, i_r}.$$

The Kolmogorov–Sinai entropy of the dynamical system $(X, \mathscr{X}, \mu; S)$ is defined by

$$H(S) = \sup_A \lim_{r \to \infty} \frac{H_r(A)}{r}.$$

The entropy of a dynamical system plays a central role in ergodic theory, primarily in the solution of the metric classification of the dynamical system, that is, the problem of describing a complete set of invariants, which imply the metrical isomorphism of dynamical systems.

It should be recalled that the first example of a metric invariant is the spectrum of a dynamical system.

In case of ergodic dynamical systems with pure point spectrum the complete set of metric invariants is given by that spectrum (von Neumann [204] and Halmos and von Neumann [74]). But if we take dynamical systems with continuous spectra, in particular their most important subclass, the systems with countably multiple Lebesgue spectrum (for example, for the Bernoulli automorphisms) we shall see that there have been no approaches to their metric classification before Kolmogorov's.

The entropy of a dynamical system proved to be a radically new invariant under metrical isomorphism of dynamical systems, which is independent of their spectrum, as the entropy may assume any of the admissible values in the class of systems with countable Lebesgue spectrum. Thus the new invariant allowed the decomposition of dynamical systems with countable Lebesgue spectrum into a continuum of invariant subclasses, with different values of entropy and therefore metrically nonisomorphic.

All K-systems (from the viewpoint of the theory of random processes these correspond to processes with a very weak condition of dependence between values at mutually remote time intervals) have countable Lebesgue spectrum and positive entropy. In the absence of other metric invariants for distinguishing K-systems, besides the entropy, it was natural to ask whether it is true that K-systems with equal values of entropy are metrically isomorphic.

The first examples of a nontrivial isomorphism of a Bernoulli automorphism were provided by L. D. Meshalkin, a pupil of Kolmogorov. Sinai showed that the Bernoulli automorphisms with equal entropies are weakly isomorphic, i.e., each can be realized as a factor of the other. The complete solution of the problem of isomorphism was obtained by the American mathematician Ornstein [143], who showed that Bernoulli automorphisms with equal entropy are metrically isomorphic. However, it has recently been discovered that in the class of all K-systems the entropy does not give a full system of metric invariants. (D. Ornstein and P. Shields showed that the number of nonisomorphic types of K-systems with equal entropy is uncountable.) For more details see the commentaries on Kolmogorov's work in ergodic theory by Sinai ([IA], pages 275–279) and [100, 101, 124, 145]. It is shown here that nowadays the entropy theory of dynamical systems, which was initiated by Kolmogorov's works, has become an important branch of ergodic theory.

Kolmogorov's entropy characteristics of "metric massiveness" [$\mathscr{H}_\varepsilon(C)$, $\mathscr{H}_\varepsilon(C, X)$, $\ldots$] enabled him to assign a clear interpretation to the results by A. G. Vitushkin on the nonrepresentability of a function of n variables of smoothness r as the superposition of a function of m variables of smoothness l, if $n/r > m/l$.

This research brought Kolmogorov right to the 13th Hilbert problem, that is, to the existence of a continuous function of three variables which is not expressible as the superposition of continuous functions of two variables.

In 1955 Kolmogorov started a seminar for students on the theory of approximate representation of functions of several variables, including also problems in approximate nomography.

Kolmogorov recalled about this seminar ([MM], page 444):

"Even in my introductory lecture I have formulated the 13th Hilbert problem as a very remote and hardly realistic target."

In Hilbert's formulation his "13th problem" lies in showing that the solution $f = f(x, y, z)$ of the equation of the seventh degree

$$f^7 + xf^3 + yf^2 + zf + 1 = 0$$

(which an arbitrary algebraic equation of the seventh degree can be transformed into) cannot be represented as a superposition of continuous functions of two variables. (See [2, 118].)

Kolmogorov begins his work of 1956, "On the representation of continuous functions of several variables as superpositions of continuous functions of fewer variables" [K265], [MM-55], by saying:

"Theorem 4, stated below, has the following unexpected consequence: Any continuous function of however many variables is representable as a finite superposition of continuous functions of at most three variables. Here is the representation of an arbitrary function of four variables:

$$f(x_1, x_2, x_3, x_4) = \sum_{r=1}^{4} h^r[x_4, g_1^r(x_1, x_2, x_3), g_2^r(x_1, x_2, x_3)].\text{"}$$

In 1957 V. I. Arnol'd showed [11] that every continuous function of *three* variables can be represented as a superposition of continuous functions of *two* variables (thus disproving Hilbert's conjecture). Finally in the same year, 1957, Kolmogorov [K273], [MM-56] made the last step by showing that *every continuous function $f(x_1, x_2, \dots, x_n)$ of n variables is representable as a superposition of continuous functions of one variable and the operation of addition*,

$$f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n+1} \chi_q \left[\sum_{p=1}^{n} \phi^{pq}(x_p) \right],$$

where the "inner" functions ϕ^{pq} are universal and only the "outer" function χ_q depends on the given function $f(x_1, x_2, \dots, x_n)$.

Arnol'd [MM, page 445] says that Kolmogorov described this result as his "most technically difficult achievement."

In 1953 Kolmogorov published his paper "Some recent work in the field of limit theorems in probability theory" [K226], [PS-41]. He says in the introduction:

"In the mid-forties it was believed that the problems of classical limit theorems (that is, problems on the limit behavior of the distribution of sums of many terms, either independent or connected in a Markov chain) are almost exhausted. Actually, however, in the late 1940s one would see considerable activity in these classical branches." It can be attributed to "the insufficient accuracy of the bounds of the remainder terms in the limit theorems" and to the fact that a number of traditional problems, which used to be soluble only under complicated and restrictive conditions, achieved "rather simple and complete solution."

This work of Kolmogorov is remarkable as it analysed and proposed various measures of closeness between probability distributions and different types of their convergence. In the same work Kolmogorov suggested a new formulation of a problem on the approximation of the sums $S_n = \xi_{n1} + \cdots + \xi_{nn}$ of independent random variables satisfying the condition of asymptotic negligibility. (Under such an assumption the limit distribution is infinitely divisible, if it exists.)

The essence of this new approach is the following.

Studies of convergence to concrete infinitely divisible laws had not fully revealed how the distributions of the sums S_n of independent random variables behave. Kolmogorov radically changed the formulation of the problem by suggesting that, instead of studying the approximation of the distribution of sums by individual distributions (together with the corresponding rate of convergence estimates), one should employ the whole family of infinitely divisible distributions. This paper also raised the possibility of uniform theorems for whole classes of random variables as well as for fixed sequences $(\xi_{n1}, \dots, \xi_{nn})$.

In 1955 Yu. V. Prokhorov proved in [156] that for an arbitrary distribution F one may construct a sequence of infinitely divisible distributions D_n, such that

$$\rho(F^{*n}, D_n) \to 0, \quad \text{as } n \to \infty,$$

where $F^{*n} = F * \cdots * F$ is the n-fold convolution of the distribution F and $\rho(F,G) = \sup_x |F(x) - G(x)|$.

In other words,

$$\rho(F^{*n}, \mathscr{D}) = \inf_{D \in \mathscr{D}} \rho(F^{*n}, D) \to 0, \quad \text{as } n \to \infty, \tag{36}$$

where $\mathscr{D}$ is the class of all infinitely divisible distributions.

Kolmogorov's "Two uniform limit theorems for the sums of independent random variables" [K261], [PS-43], dated November 12, 1956, is a real breakthrough: He manages to show that in (36) the convergence is uniform over the class $\mathscr{F}$ of all distributions F, that is,

$$\psi(n) \equiv \sup_{F \in \mathscr{F}} \inf_{D \in \mathscr{D}} \rho(F^{*n}, D) \to 0, \quad \text{as } n \to \infty, \tag{37}$$

and

$$\psi(n) \le Cn^{-1/5}, \tag{38}$$

where C is a certain constant.

The same work supplied the corresponding theorem for nonidentically distributed terms as well.

These results of Kolmogorov gave a powerful impetus to the study of the right order of decrease of the function $\psi(n)$ when $n \to \infty$.

In 1960 Prokhorov [158] showed that

$$\psi(n) \le cn^{-1/3}(1 + \ln n)^2.$$

In his work "On the approximation of distributions of sums of independent terms by infinitely divisible distributions" ([K308], [PS-51]), dated 1963, Kolmogorov obtained the bound $\psi(n) \le cn^{-1/3}$.

L. D. Meshalkin ([130, 131], in 1960) obtained a lower bound

$$\psi(n) \ge cn^{-2/3}(1 + \ln n)^{-7/2}.$$

In 1980–1983 T. V. Arak and A. Yu. Zaitsev obtained the final result,

$$c_1 n^{-2/3} \le \psi(n) \le c_2 n^{-2/3}, \tag{39}$$

where c_1 and c_2 are certain constants. For this and other results, as well as a detailed historical and bibliographical description of the problems, see the book by Arak and Zaitsev [9].

The value of (39) lies in particular in the fact that it leads to the existence for F^{*n} of infinitely divisible approximations whose order in the uniform metric is much better than the $n^{-1/2}$ provided by the classical Berry–Esseen bound (in the case of the central limit theorem).

In his work [K261], [PS-43] Kolmogorov gives a number of inequalities for the concentration function $Q(l, \xi) \equiv \sup_x P(x \le \xi \le x + l)$ of the random variable ξ, introduced by P. Lévy. This function is useful in characterizing the "dispersion" of random variables, especially in qualitatively describing the growth of the "dispersion" when independent random variables are summed. Kolmogorov develops Lévy's results on the properties of the concentration function and obtains the bound $Q(l, S_n)$ for the sum $S_n = \xi_1 + \cdots + \xi_n$ via the concentration function of each independent term. In 1961 Rogozin [160] reinforced Kolmogorov's inequality and presented it as

$$Q(l, S_n) \le Cl\left\{\sum_{i=1}^{n} l_i^2[1 - Q(l_i, \xi_i)]\right\}^{-1/2}. \tag{40}$$

Later, Miroshnikov and Rogozin [132, 133] obtained the inequality

$$Q(l, S_n) \le Cl\left\{\sum_{i=1}^{n} E\left[\min\left(|\tilde{\xi}_i|, \frac{l_i}{2}\right)\right]^2 Q_i^{-2}(l_i)\right\}^{-1/2}, \tag{41}$$

where $\tilde{\xi}_i$ is the symmetrization of the random variable ξ_i, i.e., $\tilde{\xi}_i = \xi_i - \xi_i'$, where ξ_i and ξ_i' are iid random variables [the bound (40) follows from (41) when $2l \ge l_i$, $i = 1, \ldots, n$].

The 1950s gave probability a new branch in the theory of random processes, that is, the theory of functional limit theorems (in particular those of the "invariance principle"), whose foundation benefitted much from the works by A. N. Kolmogorov, Yu. V. Prokhorov and A. V. Skorokhod.

Even in 1931 in Kolmogorov's "Eine Verallgemeinerung des Laplace–Liapounoffschen Satzes" [K31], [PS-12], he considered a problem which in modern terms could be referred to as a typical boundary problem for the "invariance principle." (For more details see page 889.) A series of separate results appeared in the 1940s and 1950s, by Erdös and Kac in 1947 [48], by Doob in 1949 [42], by Donsker in 1951–1952 [39, 40], by Fortet and Mourier in 1953 [58] and by Maruyama in 1955 [127], which all refer to the "invariance principle."

On November 30, 1948, Kolmogorov reported to the Moscow Mathematical Society his "Measures and distributions of probabilities in functional spaces." Here he suggested considering the distribution of a random process as a measure on the Borel σ-algebra of some functional space, and pointed out that it is then natural to take convergence of the distributions of a random process as weak convergence of their corresponding measures in function space. In 1953 Kolmogorov's pupil Yu. V. Prokhorov formulated a significant result [155], saying that tightness of a family of probability measures implies relative compactness in an arbitrary metric space; then he constructed the general theory of weak convergence of distributions of random processes (necessity holds for complete separable metric spaces) [157].

In his consideration of random processes with values in the space D (functions continuous from the right and with limits on the left) Skorokhod (in 1955–1956) introduced a metric which turns D into a Hausdorff topological space. In the work "On Skorokhod convergence" [K260], [PS-42] Kolmogorov suggests a more convenient metric (equivalent to Skorokhod's) where the space D proves to be separable. In the same work Kolmogorov shows that general topological considerations enable the space D to be given a metric, under which it becomes a complete separable metric space, and sets the task of finding a simple construction of such a metric. (This was soon done by Yu. V. Prokhorov.)

The basic ideas of the general theory of weak convergence of probability measures on metric spaces and its results were presented in the joint report by A. N. Kolmogorov and Yu. V. Prokhorov [K262], [PS-44] to the Conference on Probability and Mathematical Statistics in Berlin in 1956.

Kolmogorov's ideas on algorithms date back to the 1950s. On the one hand he intended to provide a most general mathematical definition of the notion of algorithm and, on the other hand, he wished to clarify whether such a notion would lead to the expansion of the already established notion of a computable function (see [193, 152, 123, 37]). These ideas were presented in Kolmogorov's report "On the concept of algorithm" [K225], [IA-1], delivered at a meeting of the Moscow Mathematical Society (March 17, 1953), and in the diploma paper by his pupil V. A. Uspenskii, "The general definition of algorithmic computability and algorithmic reducibility" (written in the first half of 1952). The detailed analysis of the development of these ideas in the theory of algorithms and its state at that time are presented in an article by Kolmogorov and Uspenskii, "On the definition of algorithm" ([K283], [IA-6], 1958), and in the commentaries by V. A. Uspenskii and A. L. Semenov ([IA], pages 279–289). In particular, in these commentaries one might find a detailed explanation of the connections between "Algorithms, or Kolmogorov machines" and "Computable functions, or Turing machines" (for interesting material about the life and work of A. M. Turing see [78, 209]).

In the late 1940s and early 1950s Kolmogorov initiated a new course, "Analysis-III," for students of the Mathematics and Mechanics Faculty of Moscow State University and read lectures himself. This replaced separate courses in the theory of functions of a real variable, theory of measure, integral equations and the calculus of variations.

These lectures by Kolmogorov and S. F. Fomin appeared in 1954 and 1960 in two volumes (*Elements of Function Theory and Functional Analysis* [K253, K297]). A second one-volume edition came out in 1968 [K351], survived several subsequent editions (6th edition, 1989) and was translated into many languages.

Kolmogorov devoted much to his research, organizational and pedagogical activities. During 1951–1953 he was Director of the Scientific and Research Institute of Mathematics and Mechanics of Moscow State University. From December 25, 1954 until February 1, 1958 he was Dean of the Mathematics and Mechanics Faculty of Moscow State University; he headed the Department of Probability of this Faculty (during 1939–1966) and headed the Probability

Division in the Steklov Mathematical Institute, USSR Academy of Sciences (from 1939 to 1958). In the spring of 1958 he was at the University of Paris as Visiting Professor.

The sixties (1960–1969). This period is noted for Kolmogorov's work on the logical foundations of information and probability theories and their reconstruction on an algorithmic basis.

It might be said that unlike the 1950s, when Kolmogorov's efforts were concentrated on the application of information theory to the various branches of mathematics, this time he created new mathematical fields:

Algorithmic information theory.
Algorithmic probability theory.

Analyzing the various approaches to information theory, Kolmogorov distinguishes three of them ([IA], pages 251–253):

1. The purely combinatorial approach.
2. The purely probabilistic approach.
3. The algorithmic approach.

In the commentaries on his works in information theory and its various applications ([IA], pages 251–253) Kolmogorov describes the essence of the first approach:

"In the combinatorial approach the quantity of information, transmitted by indication of a certain element in a set of N objects, is taken as the binary logarithm of N (R. Hartley, 1928). For example, there are

$$C(m_1,\dots,m_S) = \frac{n!}{m_1!\cdots m_S!}$$

different words in the alphabet of S elements, including the ith letter of our alphabet m_i times ($m_1+\cdots+m_S=n$). Therefore the required quantity of information is

$$H = \log_2 C(m_1,\dots,m_S).$$

If $n, m_1,\dots,m_S$ tend to infinity, one obtains the asymptotic formula

$$\text{(i)} \qquad H \sim n\sum_i \frac{m_i}{n}\log_2\frac{m_i}{n}.$$

The reader must have noticed its similarity to that in information theory,

$$\text{(ii)} \qquad H = n\sum_i p_i \log_2 p_i.$$

"If our work is constructed by a well-known pattern by means of independent trials, then the asymptotic formula (i) is an evident consequence of (ii) and the law of large numbers, but the range of applicability of (i) is much wider. (See, for example, works on the transmission of information through nonstationary channels.) In general I believe it very useful to get rid of surplus probability assumptions whenever possible. I have repeatedly pointed out in my lectures the proper value of the purely combinatorial approach to problems of information theory.

"It is the purely combinatorial approach to entropy that supported my works and those of my colleagues on ε-entropy and the ε-capacity of compact classes of functions. Here the ε-entropy $H_\varepsilon(K)$ is the quantity of information, required for the selection of an individual function from a class of functions, and the ε-capacity $C_\varepsilon(K)$ is the quantity of information which might be encoded by the elements from K if the elements from K, located at a distance not smaller than ε from one another, are reliably distinguishable."

The probability approach to the basic notions of information theory has been described above (the 1950s). As to the algorithmic approach, Kolmogorov's idea is to define the entropy (alternatively, complexity) and the quantity of information on the basis of algorithms and computable functions.

Kolmogorov described the essence and background of the algorithmic approach in his report to the Probability Section of the Moscow Mathematical Society on April 24, 1963:

"One often has to deal with very long sequences of symbols. Some of them, for example, the sequences of symbols in the 5-digit logarithm table, permit a simple logical definition and therefore might be obtained by the computations (though clumsy at times) of a simple pattern.

"Others seem not to admit any sufficiently simple 'legitimate' way to construct them. It is supposed that such is the case for a rather long segment in a table of random numbers.

"There arises the question of constructing a rigorous mathematical theory to account for these differences of behavior.

"Let us follow the tradition of information theory and reduce ourselves to binary sequences, that is, those of the type

$$x = (x_1, x_2, \dots, x_n),$$

where $x_i = 0$ or 1. Let us denote by D^n a set of these sequences of length n and let $E = D^1 \cup D^2 \cup \cdots$ be the set of all binary sequences."

Kolmogorov emphasized that different methods are possible in the introduction of measures $K(x)$ of "complexity" of sequences x, corresponding to this idea, which is well-formulated, though some arbitrariness can hardly be avoided here. "The basic discovery," Kolmogorov wrote in [IA, page 253], "which I have accomplished independently from and simultaneously with R. Solomonoff lies in the fact that the theory of algorithms enables us to limit this arbitrariness by the determination of a 'complexity' which is almost invariant (the replacement of one method by another leads only to the supplement of the bounded term)."

Let us identify each sequence $x = (x_1, x_2, \dots, x_n)$ with a natural number whose binary expansion is uniquely determined by the ordered set $\langle 1, x_1, \dots, x_n \rangle$. This identification enables us to speak about the partially recursive (computable) functions f, determined on E, with values in E. By definition Kolmogorov supposes that for each function f,

$$K_f(x) = \begin{cases} \min\{n\colon x \in f(D^n)\}, & x \in f(E), \\ +\infty, & x \notin f(E), \end{cases}$$

and calls it a "complexity of the object x" at the "mode f of the assignment."

Then Kolmogorov introduces the class $\mathscr{F}_0$ of such functions f_0, called optimal, which have the property that for any other function f there exists a constant C, depending on f_0 and f, such that for all $x \in E$

$$K_{f_0}(x) \le K_f(x) + C.$$

A fundamental result, discovered by Kolmogorov and Solomonoff [185] independently, states that this class $\mathscr{F}_0$ of computable functions f_0 is not empty. This basic result suggests calling a function

$$K(x) = K_{f_0}(x)$$

(where $f_0 \in \mathscr{F}_0$) a "complexity" ("measure of complexity") of the sequence x. For any two functions f_1 and f_2 from $\mathscr{F}_0$ we have

$$|K_{f_1}(x) - K_{f_2}(x)| \le C(f_1, f_2)$$

at any $x \in E$. In this sense all the optimal functions from the class $\mathscr{F}_0$ are equivalent and thus ([K461], [IA-13], page 243): "from the asymptotic point of view the complexity $K(x)$ of the element x does not depend on the random peculiarities of the chosen optimal method."

Along with the "complexity" $K(x)$, also referred to as the "simple Kolmogorov entropy" of the individual object x, he introduces the conditional entropy $K(y|x)$ of the object y at the known x and the information $\mathscr{I}(y|x)$ about y, contained in x. [Note that Kolmogorov's entropies $K(x)$, $K(y|x), \dots$ are closely related to the notion of the "complexity of recursive functions," introduced by Schnorr [166–168] as the logarithm of the number (of the function) with respect to an "optimal enumeration."]

All these concepts introduced and researched by Kolmogorov in his papers, "Three approaches to the definition of the quantity of information" [K320], [IA-10], "On the logical foundations of information theory and probability theory" [K354], [IA-12] and "Combinatorial foundations of information theory and the calculus of probability" [K461], [IA-13], and also in the works by his pupils and other researchers (for more details see the commentary by A. H. Sheny in [IA], pages 257–261) gave birth to the development of a new subject, algorithmic information theory. See also the book by Chaitin [28].

Kolmogorov's ideas related to the introduction of the "complexity" led him to consider afresh which class of concrete sequences $x = (x_1, \dots, x_n)$ consisting of 0's and 1's, for instance, might be naturally viewed as "random," and which might not.

If, for example, one fairly tosses an unbiased coin $2n$ times and describes the results by 1 and 0, then for sufficiently large $2n$, outcomes such as $(0, 0, 0, 0, \dots, 0)$

or $(0, 1, 0, 1, \ldots, 0, 1)$ can hardly be considered random, though from the probability point of view every such sequence (like any other) has one and the same probability 2^{-2n}. Thus classical probability is unable to answer the question of how to distinguish "random" and "nonrandom" sequences and what is the real meaning of "randomness" of an individual sequence.

The very idea of distinguishing a subclass of the class of infinite (for instance, binary) sequences $x = (x_1, x_2, \ldots)$, whose elements could be called random, dates back to R. von Mises, who employed here the German word "Kollectiv." Under his scheme it is first of all necessary, in order that the sequence $x = (x_1, x_2, \ldots)$ be "random," that there exist a limit $\lim_n(S_n/n)$, where $S_n = x_1 + \cdots + x_n$ is the number of ones in $(x_1, \ldots, x_n)$. The example of the sequence $(0, 1, 0, 1, 0, 1, \ldots)$, for which this limit exists, shows that this condition is necessary, but can by no means be viewed as sufficient, for "randomness."

This is why von Mises adds a further requirement, saying that the average frequency of 1's should be maintained, if the sequence is replaced by one of its infinite subsequences, obtained via any acceptable rule of selection. But he did not provide an exact definition. In 1940 Church [31] gave a possible definition of an "acceptable rule of selection," thus formally defining the random sequence $x = (x_1, x_2, \ldots)$ and making the concept of von Mises precise.

The importance of Kolmogorov's work "On tables of random numbers" [K311], [IA-9], representing, as he said in the introduction to [IA-9], a certain stage in his "drive for the comprehension of the frequency interpretation of probability by von Mises," lies primarily in the introduction of a more general pattern of selection, which enlarges the class of acceptable selection rules and states that "the order of terms in the subsequence is not bound to coincide with their order in the initial sequence." (Loveland reached this result independently in 1966 [119, 120].) The class of sequences so defined is called the class of random sequences in the sense of von Mises–Kolmogorov–Loveland, and all these sequences are random in the sense of von Mises–Church. However, the converse is not true (as Loveland showed).

In [K311] (see also "The author's remarks" in the Russian translation of the work [IA], page 204), Kolmogorov discusses his broader (compared to von Mises) definition of the algorithm of random selection and emphasizes that "the main difference as compared to von Mises is the strictly finite nature of the whole concept, and in the introduction of a quantitative bound on the frequency stability." By this Kolmogorov meant that the measure of the randomness can be defined for sufficiently long finite sequences $x = (x_1, x_2, \ldots, x_n)$ as well as for infinite ones $x = (x_1, x_2, \ldots)$.

The approach to the problem, developed by Kolmogorov in [K311] and based on the introduction of a finite system of selection rules, can naturally be called the "frequency" one.

Later A. N. Kolmogorov, Martin-Löf [125] and Levin [107] developed an approach, based on the notion of complexity, where those sequences $x = (x_1, \ldots, x_n)$ are called random whose complexity is maximum. [As the complexity $K(x_1, \ldots, x_n) \leq n + C$, it is natural to call a sequence $x = (x_1, \ldots, x_n)$ random, when $K(x_1, \ldots, x_n) \geq n - C$.] For more details see [IA], page 272.

Afterwards Kolmogorov repeatedly returned to the problems of "randomness and complexity." He delivered the report, "Combinatorial foundations of information theory and the calculus of probability" at the International Mathematical Congress, Nice, 1970, unfortunately published only in 1983 [K461], [IA-13]. Kolmogorov reported his "On logical foundations of probability theory" to the Fourth Soviet–Japanese Symposium on Probability and Mathematical Statistics, Tbilisi, 1982 [K462], [PS-53], where he took the view that "randomness means absence of regularities," and explained how the complexity of a finite object enables this idea to be made precise.

A Forum report, "Algorithms and randomness," by Kolmogorov and V. A. Uspenskii, was presented to the First World Congress of the Bernoulli Society, 1986; the full text of it appeared in [K475]. It contains the most exhaustive picture of the concepts and results of Kolmogorov and his pupils and followers on the algorithmic approach to the definition of "randomness." See Zvonkin and Levin [215] and Vovk [207], and also an exhaustive survey by Li and Vitányi [112].

Among Kolmogorov's studies of the 1960s a special role belongs to his works in linguistics and philology, devoted to the statistics of speech and to poetry studies. His concepts in these branches are, on the one hand, closely connected to information theory, implementing both the probabilistic and algorithmic approaches, and, on the other hand, they reflect his long-standing interest in the analysis of the regularities inherent in the form of literary works and their style. (Our description of these studies is based on materials prepared by A. V. Prokhorov at the author's request.)

Kolmogorov was reportedly interested in poetry studies even in the 1940s. The basic idea that guided Kolmogorov's studies is that the entropy of speech (measure of the quantity of information conveyed) can be decomposed into two components: the nonspeech information (essential, semantic) and the proper speech information (linguistic). The first one describes diversity, enabling the communication of semantically varied information; the second one, called "residual entropy," describes the diversity of means of expressing the same (or equivalent) semantic information—in other words, it characterizes the *flexibility of expression*. The presence of the "residual" entropy ensures the possibility of assigning special artistic (e.g., auditory) expressiveness to the communication of the given semantic information. In view of this general idea, concrete problems on the computation of the full and residual entropies were set and solved. Together with Kolmogorov other contributions were also made by his pupils A. V. Prokhorov, N. G. Rychkova-Khimchenko, N. D. Svetlova and A. P. Savchuk.

In 1960–1961 Kolmogorov developed a new method for the consistent estimation of the entropy of speech. Originating as the result of a considerable refinement of the method of Shannon [171] on the determination of the entropy by "guessing-the-sequel" tests, Kolmogorov's method avoids the uncertainty of the previous one, namely it leads (under the correct strategy of guessing) to the consistent estimation of entropy instead of the upper and lower bounds of Shannon. Guessing-the-sequel procedures provided only an upper bound for the entropy. However the bound proved more precise than Shannon's. (Deviating from the optimal strategy may result in overstating all the bounds and unreliability of the lower bounds [165, 149].)

In 1962 Kolmogorov suggested a purely combinatorial approach to the entropy of speech: The combinatorial entropy of speech was defined as the limit

$$H_k = \lim_{n \to \infty} \frac{N_n}{n},$$

where N_n is the number of texts of length n composed from the words of a given vocabulary and obeying given rules of grammar. As had been conjectured, the entropy bounds so obtained proved higher on average than the upper bounds obtained by guessing-the-sequel methods when applied to concrete prose texts ([K320], [IA-10]).

From the point of defining the "flexibility" of speech and estimating the "residual" entropy Kolmogorov focused his efforts on poetry, as here one dealt with regularities that were independent of the meaning.

The possibility of assigning the proper auditory expressiveness to poetic speech is based on the multiplicity of ways of communicating the same (or equivalent) contents. Modern poetry studies contain many works which apply statistical methods to research on metrics and rhythms of poetry. On Kolmogorov's initiative substantial research was carried out to revise and make precise the results by such famous poetry scholars as A. Belyi, B. Tomashevskii, G. Shengeli, K. Taranovsky, R. Jakobson, etc. The major results obtained by Kolmogorov and his pupils and followers can be divided into the following groups.

1. *Revelations of metric laws*. General and partial definition of meter, idea of the image of meter and acoustic meter, strict formal logical definition of classic meters [K313, K352, K353, K465]; description and classification of the nonclassic Russian meters [K303, K312, K313, K317, K318, K325].
2. *The classification and statistics of the meter's rhythmic varieties*. These works formulated and verified the principal argument saying that the acoustic structure of speech obeys simple statistical regularities, which can be computed by probability theory (these regularities are realized under pressure of the demand to convey semantic information, if such pressure is not balanced by a systematically introduced artistic trend); marked the general method of construction of the theoretical patterns of various meters; and formulated the hypothesis of "randomness imitation" [K314, 153, 154, K466].
3. *The analysis of the "residual" entropy and its bound*. A bound was obtained for the "residual" entropy and the "consumption of the entropy" was calculated for the separate means of poetic auditory expressiveness.

The investigations headed by Kolmogorov initiated a large flow of studies on mathematical methods for research on the language of artistic works. In the 1960s there were two seminars at Moscow State University, involving such prominent figures in linguistics and philology as A. Zaliznyak, V. Ivanov, M. Gasparov, and V. Rozentsveig. Many works in poetry studies were directly influenced by Kolmogorov's ideas, for example, those of V. Ivanov, M. Gasparov, M. Krasnoperova and others. Academician V. Zhirmunskii, Professor S. Bondi and Professors K. Taranovsky and R. Jakobson of Harvard University displayed continuous attention to and interest in this work.

In the 1960s Kolmogorov's scientific and organizational activity was marked by two major events: the establishment of the Laboratory of Statistical Methods, Department of Probability, Mathematics and Mechanics Faculty, Moscow State University and the establishment of the Physics and Mathematics School (No. 18), sponsored by Moscow State University.

The Laboratory of Statistical Methods was set up to unify efforts and intensify research in the application of probability and statistical methods.

Kolmogorov outlined its major goals as: the theory of optimal control and statistical decisions (I. V. Girsanov); the theory of reliability (B. V. Gnedenko and Yu. K. Belyaev); design of experiments (V. V. Nalimov and V. V. Fedorov); statistics and linguistics (A. V. Prokhorov); statistics in medicine (L. D. Meshalkin); statistics in geology (A. M. Shurygin); and nonlinear spectral analysis of random processes (A. N. Kolmogorov, A. N. Shiryaev and I. G. Zhurbenko).

Some scholars of the academic institutes, such as L. K. Bolshev and A. N. Shiryaev, worked in the Laboratory as consultants.

The widely known All-Union Seminar on Turbulence in Liquids and Gases was organized within the framework of the Laboratory and with the direct participation of A. N. Kolmogorov, M. D. Millionshchikov, A. M. Obukhov, S. A. Khristianovich, L. I. Sedov, A. S. Monin, A. M. Yaglom, etc. This seminar directly prompted Kolmogorov to commit himself to join two voyages of the scientific research ship "Dmitrii Mendeleev" in 1970 and in 1971–1972, and thus to study the turbulence of the ocean.

As scientific supervisor of these voyages Kolmogorov (together with A. S. Monin, V. Pak, I. G. Zhurbenko, M. V. Kozlov, etc.) dealt with the research and application of highly robust methods (as compared to those which lead to distortion from the neighboring frequencies) of spectral analysis. The special significance of these methods is their potential application to the spectral analysis of nonstationary processes. Admitting that the computation of the statistical parameters of random space–time turbulent fields is rather lengthy and complicated, Kolmogorov viewed his role in these voyages as "the prompt clarification of the techniques directly on board; determination of the necessary duration of the realization, and the interval of the discretization, etc.," so that the analysis of the result will facilitate "the further planning of measurements, evaluating their representativeness; proper judgment of the applicability and quality of the tested samples of the measuring equipment."

Under the influence of the work of J. Neyman and E. Scott, Kolmogorov launched research in the Statistical Laboratory on the statistical analysis of the active effects on atmospheric phenomena (in particular, artificial stimulation of rainfall).

The studies of Neyman's methods showed Kolmogorov that the parametric methods are too "sophisticated" (nonrobust) for the relatively "rough" atmospheric data. In this connection he created his already tested nonparametric methods (close to Fisher's concepts). Together with Neyman he argues strongly for the randomization of the experiment and for the necessity of obtaining "pure" data. Kolmogorov vigorously supports Neyman, who said that "statistics is helpless and powerless unless the initial data is not spoiled."

Since the foundation of the Laboratory of Statistical Methods (1960) Kolmogorov was its Director, giving it much of his time, energy and strength. Kolmogorov granted much of his personal currency reserves (from the International Balzan award of 1963) for purchasing foreign books and journals for the Library in Probability and Statistics of the Laboratory.

From February 1, 1976 Kolmogorov headed the new Division of Mathematical Statistics, Mathematics and Mechanics Faculty, Moscow State University, which was established on his initiative.

In 1963 the USSR Council of Ministers opened four schools in mathematics and physics in the universities of Moscow, Leningrad, Kiev and Novosibirsk.

The School No. 18 of the new type sponsored by Moscow University is inseparably linked to Kolmogorov, and therefore it is often referred to as "Kolmogorov's school."

Answering a question on how he viewed the first steps in familiarizing the future scholar with science, Kolmogorov said: "Following the biographies of celebrated scientists, we see that they started with encouraging school teachers who supported the capable pupils; then came the first scientific supervisor, who outlined the first topic of independent research, which was very often specially adjusted to the capabilities of this pupil. We also see one or two close friends, encouraging one another. I suppose that these fragile human relationships which shape scientists will maintain their value in the future as well.

"Now, when our country is in need of many capable and well-educated researchers in the most diverse branches of science and technology it becomes imperative to establish a wide system of institutional measures with extracurricular lessons with the senior school children: specialized schools, various types of nonschool activities, wide familiarization of the young with the specific nature of work in the universities and technical colleges of the new technology (such as the Moscow Physical-Technical College), proper organization of entrance examinations, and wide involvement in research of students in colleges, where the teaching of future researchers is subsidiary only. Certainly, all these organizational measures will not provide an exhaustive result, if they are not followed by the individual call for every boy—a potential scholar—that I was talking about in the beginning."

Following these principles Kolmogorov committed himself entirely to this School No. 18 right from its first day; he viewed his work with the schoolchildren, and the other broader work in the improvement of mathematical teaching in secondary school, as necessary and valuable for his country, as his civic duty, as his personal responsibility for mathematical education.

Kolmogorov's personal efforts meant much for the school both in its founding period and in its first days and years. He did not restrict himself (and it went on for *15 years*!) to lectures and exercises, but also wrote summaries for the pupils, told them about music and literature and went camping with them.

As already noted it was in 1922–1925 that Kolmogorov took up school teaching of mathematics and physics in Potylikhin Experimental School. (His employment list has as entry No. 1: "Total employment period of 3 years prior to joining Moscow University.") In this school Kolmogorov was Secretary of the School Council and a tutor, a matter of pride for him. Forty years later, as the Chairman of the School Trustee Council, he immediately looked into the needs and problems of the pupils, many of whom came from small cities and villages. (There is no admission to this school for the residents of university cities!) He also supervised the work in general.

For more details about this school see [K413, K428]. A documentary film, "Ask your questions, boys," released in 1970, was devoted to this school.

In 1964 Kolmogorov headed a Mathematics Section of the Commission of the USSR Academy of Sciences and the Academy of Pedagogical Sciences on the contents of the secondary education. In 1968 this section issued a new school curriculum in mathematics for grades 6–8 and 9–10, which contributed to the further improvement of mathematical education and provided a background for new textbooks. Kolmogorov participated directly in preparation of the textbooks *Algebra and the Beginnings of Analysis, 9–10* [K458] and *Geometry, 6–8* [446].

Let us complete this portion of the school-related summary of Kolmogorov's activities by one of his annual reports.

"The report of the member of the USSR Academy of Pedagogical Sciences for 1969:

1. Three first contributions to *Mathematics at School* on the scientific foundation of school mathematics (two of them published in 1969, the third to appear in 1970).
2. Editing an experimental textbook on geometry (by Cherkasov, Nagibin and Semenovich) for the sixth grade.
3. In the first six months I headed the Mathematics Section in the Commission on the contents of mathematics education. Number of critical reviews made on the textbooks.
4. Supervised teaching of mathematics in the school at Moscow State University. Lectures for students of the ninth grade (the first six months) and of the tenth grade (second six months).
5. Collection of information for *Quantum*, a magazine for the senior school children (publication scheduled for 1970)."

In the late 1950s and early 1960s Kolmogorov suggested to his pupils V. P. Leonov and A. N. Shiryaev a series of problems related to the issues of nonlinear analysis of random processes (in particular, in radio technology) which brought about the techniques of calculating cumulants under nonlinear transformations, and the development of the theory of spectral analysis of the high-order moments of stationary random processes [104; 173; 176, pages 287–291].

Kolmogorov participated in and supervised research on the conditions for the validity of various properties of ergodic and mixing type, and conditions for the validity of certain basic limit theorems for random processes (Volkonskii and Rozanov [199], Leonov [104], Sinai [179] and Shiryaev [174]).

In his work [K304], [PS-50], Kolmogorov (in collaboration with M. Arato and Ya. G. Sinai) constructed a mathematical theory for estimation of the parameters describing the displacement of the Earth's rotation axis, based on the idea that the fine structure of the movement is governed by a complex stationary Gauss–Markov process.

In 1964 the Laboratory of Statistical Methods under Kolmogorov's supervision launched research on the nature of the eleven-year periodicity of sunspot activity, which was to confirm a hypothesis by E. E. Slutskii ("a precise periodicity with constant phase over hundreds of periods, obscured by noise, instead of an oscillating process with varying phase").

A report by A. N. Kolmogorov and A. N. Shiryaev, "The application of Markov processes to the detection of disruption in industrial processes," was presented to the Sixth All-Union Conference on Probability and Mathematical Statistics, Vilnius, 1960. It initiated broad development and research in statistical sequential analysis [K477, 175], optimal nonlinear filtering [114] and martingale theory [115, 88].

On April 5, 1961 Kolmogorov spoke at Moscow State University on "Automata and life," a lecture which impressed by its strength, depth of ideas and remarkable vision.

In 1967 there came a work by A. N. Kolmogorov and Ya. M. Barzdin', "On the realization of networks in three-dimensional space" [K335], [IA-11], which was inspired by an attempt to explain the following construction of the human brain: Nervous fibres (axons) absorb the bulk of space and the nerve cells along with their appendages (neurons) are located on the surface only. The presented construction confirmed the optimal character of such structure for nerve networks.

On April 25, 1963, the day of Kolmogorov's sixtieth birthday, the Presidium of the Supreme Soviet of the USSR announced a Decree on "Awarding Academician A. N. Kolmogorov with the title of Hero of Socialist Labor":

"For his outstanding merits in mathematics and on the occasion of his sixtieth birthday Academician A. N. Kolmogorov is awarded the title of Hero of Socialist Labor with the presentation of the Order of Lenin and the Gold Medal 'Hammer and Sickle'."

The true appraisal of Kolmogorov's outstanding merit was the presentation to him in 1963 of the International Prize in Mathematics by the Fondation Internationale Balzan (in other fields the award was given to Pope John XXIII, historian S. Morison, biologist K. Frisch and composer P. Hindemith).

From December 1, 1964 to December 13, 1966 Kolmogorov was President of the Moscow Mathematical Society.

In 1965 A. N. Kolmogorov and V. I. Arnol'd were awarded the Lenin Prize for their works in the theory of perturbations of Hamiltonian systems.

The seventies and eighties (1970–October 20, 1987). The improvement of school mathematics remained a centerpiece of Kolmogorov's activities in both the 1970s and 1980s. He worked at school, created (together with Academician I. K. Kikoin) a popular scientific magazine in physics and mathematics for school children, *Quantum*, committed himself to work on the school curriculum and worked on commissions on school education.

The uneasy quest for optimal systems in teaching of mathematics at school, which is still under way, shows how complicated the matter is that Kolmogorov was so concerned with. He was happy with the successes and grieved much about misunderstandings of his concepts and positions in this area, but never sought sympathy and never complained.

One can boldly admit that school mathematics was the subject of his permanent interest and concern since 1927, when he had started teaching in Potylikhin Experimental School, and throughout his whole life. On November 22 and 28, 1937, two meetings of the Moscow Mathematical Society discussed the draft of a new textbook on elementary algebra (*Algebra, Part I* [K104]), submitted by P. S. Aleksandrov and A. N. Kolmogorov. The authors stated here (the book came out in 1939) certain principles, which constituted background for the textbook:

"We always tried to combine simplicity of presentation with sufficient depth and logical flawlessness. We started from the presumption that our book will be a student's reliable guide in both his first exposure to the subject and in his further studies of mathematics. The authors always sought the complete and exhaustive comprehension by the student of the meaning of all the operations. In particular much had been done to avoid the separation of operations with symbols from the arithmetic operations with numbers."

In 1941 *Mathematics at School* published articles by Kolmogorov and Aleksandrov, "Irrational numbers" [K124] and "Properties of inequalities and the concept of approximation" [K123]. In 1961 these articles were reprinted in the book *Issues of Teaching Mathematics in the Secondary School* [K301].

Let us again draw attention to some of Kolmogorov's reports so as to give the idea of the *concreteness* of his work in the 1970s and of his personal contribution to mathematics education.

"A report on the work by A. N. Kolmogorov, member of the USSR Academy of Pedagogical Sciences, 1970:

1. I head the methodological association of the mathematicians in School No. 18, Moscow State University, give lectures and do some general work as Chairman of the Trustee Council. On the basis of the school experiences a new textbook, *The Mathematics Course for the Physics and Mathematics Schools*, was prepared together with V. A. Gusev, A. A. Shershevskii and A. V. Sosinskii. (See [K381], published in 1971.) I have contributed certain chapters to this book.

   During 24 summer days I was completely busy in the summer school, which is responsible for the final selection of students to School No. 18.
2. Work on *Quantum*, as head of its mathematics section. Contributed a number of notes and the big article on the modern understanding of the concept of function. (See *Quantum*, 1970, Nos. 1–2 [K365, K367].)
3. A program has been worked out for a new course on "Scientific foundation of the school course of mathematics" for the students of pedagogical colleges; the program has been approved by the USSR Ministry of Education. One

more article contributed to *Mathematics at School* (next in turn in a series of articles on this subject).

4. "Teaching materials for fifth grade geometry" [K364] has been written in collaboration with R. S. Cherkasov and A. F. Semenovich. A textbook of geometry for grades 6–8 worked out (in collaboration with the same coauthors and F. F. Nagibin). The trial edition of the textbook for the sixth grade has been released already, the one for the seventh grade has been handed for printing. Participated in the Conference on Experimental Teaching in Vladimir.
5. Presided in the Mathematics Section of the Commission for the Contents of Secondary Education, USSR Academy of Pedagogical Sciences, where detailed critical reviews have been made on the textbooks (edited by A. I. Markushevich) for fourth, fifth and sixth grades.
6. The Commission being abolished, I started my work with the Mathematics Commission of the Scientific Methodological Council, USSR Ministry of Education. A report contributed, "A system of basic concepts and notation for a school mathematics course"; its broad presentation will come in *Mathematics at School* ([K374]).
7. Supervision of the postgraduate schooling of A. Abramov.

January 5, 1971."

"Report of A. N. Kolmogorov, acting member of the Academy of Pedagogical Sciences for the year 1974.

1. Work on textbooks for mass schools.
   1.1. The revision of the textbook on algebra and the elements of analysis for the ninth grade (in collaboration with O. S. Ivashev-Musatov and S. I. Shvartsburd), originated by myself, B. E. Veits and I. G. Demidov (general editing by me). The textbook will start in the mass school during the fall of 1975.
   1.2. Revision of the geometry textbook for grades 6–8 (in collaboration with R. S. Cherkasov and A. F. Semenovich). The work to be completed in 1975.
2. Work in School No. 18.

   Lectures for students in the ninth grade. Leadership of the summer school at Pushchino and of the admission to the school itself.

   Collection of material for a textbook for the physics and mathematics schools and for after-class studies in the ordinary schools (based on the summer school experiences).
3. Leadership of the mathematics section of the Scientific Council, USSR Ministry of Education.
4. Editorial work on *Quantum*. Collaboration with the editorial board of *Mathematics at School.*

December 12, 1974."

Though Kolmogorov's school-related legacy has hardly been studied or documented in full it presents an inexhaustible source of clear-cut and uncommonly stated ideas and concepts, which amazingly accurately and exactly grasp the essence of the problem. One vital task is to preserve this legacy, organize it and bring it to public attention.

In this respect the reader may be interested in Kolmogorov's small book, *Mathematics: Science and Profession*, *Quantum* library, No. 64, 1988 [K476], composed by G. A. Gal'perin.

Kolmogorov's scientific and organizational work of the 1970s and 1980s is associated with Moscow State University and the Steklov Mathematical Institute.

From February 1, 1976 to January 1, 1980 Kolmogorov headed the Division of Mathematical Statistics of the Mathematics and Mechanics Faculty of Moscow State University and from January 1, 1980 its Division of Mathematical Logic. On October 3, 1983 Kolmogorov joined the Steklov Mathematical Institute on a permanent basis. He headed the Department of Mathematical Statistics and Information Theory, while still maintaining his appointment in the Division of Mathematical Logic at Moscow State University.

From 1973 to October 15, 1985 Kolmogorov was President of the Moscow Mathematical Society and from 1982 to 1987 he was Editor-in-Chief of *Uspekhi Matematicheskikh Nauk.*

From August 22 to August 29, 1982, Tbilisi hosted the Fourth Soviet–Japanese Symposium on Probability and Mathematical Statistics. Although not well, Kolmogorov participated and delivered a lecture, "On the logical, semantic and algorithmic foundations of probability theory." (It appeared in [K462], [PS-53] under the title, "On logical foundations of probability theory.") This symposium was attended by 45 Japanese and 270 Soviet scientists, the Japanese delegation being headed by K. Itô. The participation by Kolmogorov and Itô was a very meaningful event.

In the early 1980s the Presidium of the USSR Academy of Sciences decided to publish Kolmogorov's selected works. Kolmogorov made up a list of his papers to be included, sorted them by subject, and wrote and dictated his commentaries, and he thoroughly studied the commentaries on groups of his papers, mainly written by his pupils. (The first volume appeared in 1985, [K467] = [MM], the second in 1986, [K471] = [PS] and the third in 1987, [K473] = [IA].)

A. N. Kolmogorov himself, the main editors (S. M. Nikol'skii and Yu. V. Prokhorov), the editors (V. M. Tikhomirov and A. N. Shiryaev) and the acting editor (V. I. Bityutskov) were greatly assisted by Kolmogorov's pupils and followers, who were very responsive to the call for Russian translations, commentaries, editing and checking the proofs.

On April 25, 1985, A. N. Kolmogorov's eighty-second birthday, he dictated the following "Epilogue", reproduced here in full [IA, page 303]:

"The suggested three volumes of *Selected Works* actually comprise all my works in mathematics, classical mechanics, the theory of turbulence, probability theory, mathematical logic and information theory. They do not include the papers on teaching and history of mathematics, on poetry and my general articles.

"In certain areas the results achieved seem sufficiently unified and complete and so in my 82 years I happily leave these things to my successors.

"In other areas things do not stand like that and the published materials seem to present only the fragments of future work, which I can only hope will be achieved by others. Progress to date is mainly covered in the commentaries by the group of my pupils, to whom I am most grateful."

A. N. Kolmogorov created a number of scientific schools, many of them headed by his pupils. The scientific atmosphere of high demands and high moral standards, his ability to encourage creativity and to spot a fitting problem or task for everyone, his extraordinarily generous attitude to ideas—this atmosphere is really unforgettable for all his pupils.

Among his pupils are:

Academicians: I. M. Gel'fand, A. I. Maltsev, M. D. Millionshchikov, V. S. Mikhalevich, S. M. Nikol'skii, A. M. Obukhov and Yu. V. Prokhorov.

Academician of the Ukranian Academy of Sciences: B. V. Gnedenko.

Academician of the Uzbek Academy of Sciences: S. H. Sirazhdinov.

Corresponding Members of the USSR Academy of Sciences: V. I. Arnol'd, L. N. Bolshev, A. A. Borovkov, A. S. Monin and B. A. Sevast'yanov.

Doctors and Candidates of Sciences, researchers: A. M. Abramov, V. M. Alekseev, M. Arato, D. A. Asarin, G. M. Bavli, G. I. Barenblatt, L. A. Bassalygo, Yu. K. Belyaev, E. P. Bezhich, V. I. Bityutskov, A. V. Bulinskii, I. Ya. Verchenko, V. G. Vinokurov, V. G. Vovk, G. A. Gal'perin, A. N. Dvoichenkov, N. A. Dmitriev, R. L. Dobrushin, E. B. Dynkin, V. D. Erokhin, I. G. Zhurbenko, V. N. Zasukhin, V. M. Zolotarev, O. S. Ivashev-Musatov, M. V. Kozlov, V. V. Kozlov, A. T. Kondurar', L. A. Levin, V. P. Leonov, R. F. Matveev, P. Martin-Löf, Yu. T. Medvedev, L. D. Meshalkin, R. A. Minlos, Yu. P. Ofman, Yu. S. Ochan, A. A. Petrov, B. Penkov, M. S. Pinsker, A. V. Prokhorov, Yu. A. Rozanov, M. Rosenblatt-Roth, Ya. G. Sinai, V. M. Tikhomirov, L. N. Tulaikov, V. A. Uspenskii, S. V. Fomin, M. K. Fage, A. N. Shilov, A. N. Shiryaev, F. I. Shmidov, B. M. Yunovich and A. M. Yaglom.

In April 1986, celebrating his birthday, A. N. Kolmogorov invited his pupils to his country cottage in the famous Komarovka. Talking about the Teacher, everyone remarked on his invariable youthfulness of spirit. The severe illness that Kolmogorov suffered from in his last years was accompanied by speech failures and he could not convey what he wanted to say. The next day he dictated the following text to one of his pupils:

Reply to Pupils

"There was talk about my allegedly inexhaustible youth. I am grateful for such an appraisal, but I'd better introduce certain limits to it. Age is anyway objective and one cannot escape it. Happy age... How can it be realized? Either by the refusal to produce any new results, or by the tolerance of an actually shallow existence. Leaving that aside, the old man can view this period as bright and happy, but it will be inevitably combined with sad feelings about whether I can do this or that. It refers to something more than cold baths and sport successes.

"The physicians objectively view my state as comparatively favorable. But anyhow the quantitative output has become considerably less, and that brings out those sad restrictions.

"In my case I see my scientific career as finished in the sense of obtaining new results. Sorry about that, but I have to bow to the inevitable.

"In recent years my activity has been developing in other directions, I mean my contribution to the school reform so vital for the nation. Here, firstly I think that if age does not interfere I would introduce many useful and even indispensable things, working on school textbooks and manuals for young people fond of science. Both directions are highly absorbing and I'd like to do things most vigorously and with youthful ardor. But the time goes by, months are gone, and this or that work has been scheduled, but is postponed.

"That is why the choice of that branch, where one is the most indispensable, acquires now special priority. If I concentrate on textbooks for the most advanced, then I'll not manage those for ordinary schools. And now you caught me at such a crossroad. If I agree to work actively and sweepingly in one direction, then I'll fail with same in the other. Such feelings intensify in old age. That is why I greatly appreciate those young assistants, many of whom were invited today."

The First World Congress of the Bernoulli Society (September 8–14, 1986) did much for probability and mathematical statistics. For health reasons, Kolmogorov unfortunately could not attend.

The inauguration was immediately followed by the forum report by A. N. Kolmogorov and V. A. Uspenskii, "Algorithms and randomness" (delivered by Uspenskii) [K475], which considered the general issues of the applicability of probability to real phenomena of a random nature and showed that the theory of algorithms and recursive functions could give a precise mathematical meaning to the concepts of "complexity" and "randomness," and also outlined the programs of further research in this area.

The report by Uspenskii was preceded by Kolmogorov's greeting to the participants of the First World Congress of the Bernoulli Society (full text in [K474], page 200).

Kolmogorov's merits are highly appreciated by the Soviet State. He was awarded the title of Hero of Socialist Labor (1963), seven Orders of Lenin (1944, 1945, 1953, 1961, 1963, 1973, 1975), a "Gold Star" medal (1963), Order of the Labor Red Banner (1940), Order of the October Revolution (1983) and many medals.

In 1941 he was awarded the Stalin Prize and in 1965 the Lenin Prize.

In 1939 Kolmogorov was elected a full member of the USSR Academy of Sciences and in 1966 a full member of the USSR Academy of Pedagogical Sciences.

In 1949 Kolmogorov was honored by the P. L. Chebyshev Prize of the USSR Academy of Sciences and in 1987 by the N. I. Lobachevskii Prize of the Academy.

The high ranking position of A. N. Kolmogorov in world science is reflected in the fact that he has been elected a member of many academies, universities and societies:

1955 honorary degree of Doctor of Science, University of Paris.
1956 corresponding member of the Romanian Academy of Sciences.
foreign member of the Polish Academy of Sciences.
honorary member of the Royal Statistical Society, Great Britain.
1957 honorary member of the International Statistical Institute.
1959 honorary member of the American Academy of Arts and Sciences.
member of "Leopoldina," the German Academy of Natural Sciences, GDR.
1960 honorary degree of Doctor of Sciences, Stockholm University.
1961 foreign member of the American Philosophical Society, Philadelphia.
1962 honorary degree of Doctor of Science, Indian Statistical Institute, Calcutta.
honorary member of the American Meteorological Society.
honorary member of the Indian Mathematical Society.
honorary member of the London Mathematical Society.
1963 foreign member of the Netherlands Royal Academy of Sciences.
1964 Fellow of the Royal Society of London.
1965 honorary member of the Romanian Academy.
honorary member of the Hungarian Academy.
1967 member of the National Academy of Sciences, USA.
1968 foreign member of the French Academy of Sciences.
1973 doctor of science honoris causa, Hungary.
1977 honorary member of the International Academy of History of Science.
foreign member of the GDR Academy of Sciences.
member of the Society of the Order of "Pour le Mérite," FRG.
1983 foreign member of the Finnish Academy of Sciences.

In 1963 A. N. Kolmogorov was awarded the International Prize in Mathematics from the Balzan Fund (Fondation Internationale Balzan).

In 1980 A. N. Kolmogorov was awarded the International Mathematical Prize of the Wolf Foundation for his "deep and original discoveries in Fourier analysis, in probability theory and ergodic theory and in dynamical systems."

In the summer of 1987 Andrei Nikolaevich's health (in recent years he had been suffering from Parkinson's disease) deteriorated to the point that he agreed to preventive treatment in special clinics and to make full medical analyses, including computer tomography, and to clarify the possibility of fighting his almost complete blindness. The illness was rapidly progressing, and serious lung failures were revealed at the beginning of October. At that time he was moved to the lung department, where we had a small chat:

"Where am I actually now?" he asked.

"Lung department."

"But why?"

"They revealed some lung trouble."

"What is the consequence?"

"You'll have to stay here for a while so as to get home with healthy lungs."

"That is all right."

It was actually the very last conversation with Andrei Nikolaevich. During the next few days his temperature and blood pressure fluctuated wildly and his breathing became complicated—the destructive forces of the disease were taking their toll. On October 20th at 14 hours, 9 minutes the control oscillograph displayed a straight line instead of the usual curve for the heart rhythm—Andrei Nikolaevich's heart had stopped. Death immediately revealed his characteristic curved nose, which was hardly visible in life, especially when his warm gaze was lit by a big smile.

This is the end of life, the life of Academician A. N. Kolmogorov.

An obituary (*Pravda* and *Izvestia*, October 23, 1987), signed by the leaders of the Communist Party and the Soviet State read:

"The whole life of A. N. Kolmogorov is an unparalleled feat for the cause of science. He has been a symbol of nobility, selflessness and the highest morality in the service of the Socialist Motherland. A. N. Kolmogorov has entered the Pleiad of the great Russian and World scientists."

Epilogue. One article can hardly embrace fully and exhaustively the versatile life and creative activities of such a unique individual as A. N. Kolmogorov. Despite its rather large scale much has been omitted. Kolmogorov's work with pupils and colleagues, the trips and walks with him, which were really extraordinary in their scientific content and emotional impact, musical parties in his well-known country house, Komarovka, excursions to the ancient Russian cities, where Kolmogorov was a wonderful guide,... .

This article lacks (one hopes that the gap will be filled by his pupils) Kolmogorov's portrait—his broad smile, his eyes, characteristic voice uncomparable with anything else.... Unlike anyone else Andrei Nikolaevich could penetrate to the very essence of the problem being discussed and grasp its gist, making others see it anew. It might be recalled that at one of the preparatory meetings of the Bernoulli Society Congress (in 1986) Andrei Nikolaevich seemingly shrank into himself, but suddenly started and spoke out. Silence came immediately and everyone heard his question: "What is the actual distribution by age of the invited speakers?" (The computer plot revealed that they were mainly in their forties.)

Primarily following chronological order, the author has tried to supplement the broad description of the basic scientific results and inventions of A. N. Kolmogorov (with emphasis on the probabilistic and statistical aspects) by remarks bringing out the impact of his ideas and works on the origin and development of many branches of science.

This article drew on various sources, in particular a number of articles about Kolmogorov in *Uspekhi Matematicheskikh Nauk* and *Theory of Probability and Its Applications*, and the commentaries to his works in the three volumes [MM, PS, IA].

Acknowledgments. An unrepayable debt is owed to Anna Dmitrievna, widow of A. N. Kolmogorov, for her advice and access to the family archives.

The advice and remarks by Yu. V. Prokhorov, V. A. Uspenskii, Ya. G. Sinai, V. M. Tikhomirov, A. V. Prokhorov, A. M. Abramov, N. H. Bingham, and G. R. Grimmett greatly improved the completeness and accuracy of the text.

The draft version of this English translation was prepared by M. M. Romanovskaya and the author with participation and assistance by B. S. Stechkin, S. S. Barsov, A. A. Gushchin and Yu. O. Usenko.

The final version of the English translation was carried out by N. H. Bingham and G. R. Grimmett.

STEKLOV MATHEMATICAL INSTITUTE
VAVILOVA STR. 42
GSP-I, 117966 MOSCOW
UNION OF SOVIET SOCIALIST REPUBLICS

REFERENCES

[1] ABRAHAM, R. and MARSDEN, J. E. (1978). *Foundations of Mechanics*, 2nd ed. Benjamin/Cummings, Reading, Mass.

[2] ALEKSANDROV, P. S., ED. (1969). *The Problems of Hilbert.* Nauka, Moscow. (In Russian.)

[3] ALEKSANDROV, P. S. (1983). A few words on A. N. Kolmogorov. *Russian Math. Surveys* **38**(4) 5–7.

[4] ALEKSANDROV, P. S. and GNEDENKO, B. V. (1963). A. N. Kolmogorov as a teacher. *Russian Math. Surveys* **18**(5) 111–115.

[5] ALEKSANDROV, P. and HOPF, H. (1935). *Topologie*. Springer, Berlin.

[6] ALEKSANDROV, P. S. and KHINCHIN, A. YA. (1953). Andrei Nikolaevich Kolmogorov (on his fiftieth birthday). *Uspekhi Mat. Nauk* **8**(3) 177–200. (In Russian.)

[7] ALEXANDER, J. W. (1935). On the chains of a complex and their duals. On the ring of a compact metric space. *Proc. Nat. Acad. Sci. U.S.A.* **21** 509–512.

[8] ALEXANDER, J. W. (1936). On the connectivity ring of an abstract space. *Ann. of Math.* (*2*) **37** 698–708.

[9] ARAK, T. V. and ZAITSEV, A. YU. (1986). Uniform limit theorems for sums of independent random variables. *Trudy Mat. Inst. Steklov.* **174**. (In Russian.)

[10] ARAUJO, A. and GINÉ, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*. Wiley, New York.

[11] ARNOL'D, V. I. (1957). On functions of three variables. *Dokl. Akad. Nauk SSSR* **114** 679–681. (In Russian.)

[12] ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. Springer, Berlin.

[13] BACHELIER, L. (1900). Théorie de la speculation. *Ann. Sci. École Norm. Sup.* (*3*) **17** 21–86.

[14] BERNSHTEIN, S. (1917). An essay on the axiomatic foundation of the theory of probability. *Proc. Khar'kov Math. Soc.* (*2*) **15** 209–274. (In Russian.)

[15] BERNSHTEIN, S. (1934). Principes de la théorie des équations différentielles stochastiques. *Trudy Fiz.-Mat. Inst. Steklov.* **5** 95–124.

[16] BERNSHTEIN, S. (1938). Équations différentielles stochastiques. *Act. Sci. Ind.* **738** 5–31. Conf. Internat. Sci. Math. Univ. Genève, Hermann, Paris.

[17] BILLINGSLEY, P. (1965). *Ergodic Theory and Information*. Wiley, New York.

[18] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[19] BINGHAM, N. H. (1986). Variants on the law of the iterated logarithm. *Bull. London Math. Soc.* **18** 433–467.

[20] BIRKHOFF, G. D. (1931). Proof of the ergodic theorem. *Proc. Nat. Acad. Sci. U.S.A.* **17** 656–660.

[21] BOGOLYUBOV, N. N., GNEDENKO, B. V. and SOBOLEV, S. L. (1983). Andrei Nikolaevich Kolmogorov (on his eightieth birthday). *Russian Math. Surveys* **38**(4) 9–27.

[22] BOHLMANN, G. (1909). Die Grundbegriffe der Wahrscheinlichkeitsrechnung in ihrer Anwendung auf die Lebensversicherung. *Atti del IV Congresso Internazionale dei Matematici* (*Roma, 6–11 Aprile 1908*) **3** Sec. 11b.

[23] BOREL, É. (1909). Les probabilités dénombrables et leurs applications arithmétiques. *Rend. Circ. Mat. Palermo* **27** 247–271.

[24] BOROVKOV, A. A. (1976). Convergence of measures and random processes. *Russian Math. Surveys* **31**(2) 1–69.

[25] BOROVKOV, A. A. (1983). Boundary-value problems, the invariance principle, and large deviations. *Russian Math. Surveys* **38**(4) 259–290.

[26] BRAMSON, M. D. (1978). Maximal displacement of branching Brownian motion. *Comm. Pure Appl. Math.* **31** 531–581.

[27] CANTELLI, F. P. (1933). Sulla determinazione empirica delle leggi di probabilitá. *Giorn. Istit. Ital. Attuari* **4** 421–424.

[28] CHAITIN, G. J. (1987). *Algorithmic Information Theory*. Cambridge Univ. Press, Cambridge.

[29] CHAPMAN, S. (1928). On the Brownian displacements and thermal diffusion of grains suspended in a non-uniform fluid. *Proc. Roy. Soc. London Ser. A* **119** 34–54.

[30] CHUNG, K. L. (1974). *A Course in Probability Theory*, 2nd ed. Academic, New York.

[31] CHURCH, A. (1940). On the concept of a random sequence. *Bull. Amer. Math. Soc.* **46** 130–135.

[32] CRAMÉR, H. (1928). On the composition of elementary errors. Statistical applications. *Skand. Aktuarietidskr.* **11** 141–180.

[33] CRAMÉR, H. (1942). On harmonic analysis in certain functional spaces. *Ark. Mat. Astr. Fys.* **28B**(12) 1–7.

[34] CRAMÉR, H. (1946). *Mathematical Methods of Statistics*. Princeton Univ. Press, Princeton, N.J.

[35] CRAMÉR, H. and LEADBETTER, M. R. (1967). *Stationary and Related Stochastic Processes*. Wiley, New York.

[36] CSÖRGŐ, M. and RÉVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics*. Academic, New York.

[37] DAVIS, M. (1958). *Computability and Unsolvability*. McGraw-Hill, New York.

[38] DE FINETTI, B. (1930). Le funzioni caratteristiche di legge istantanea. *Atti Accad. Naz. Lincei Rend.* **12** 278–282.

[39] DONSKER, M. (1951). An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.* **6**.

[40] DONSKER, M. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov–Smirnov theorems. *Ann. Math. Statist.* **23** 277–281.

[41] DONSKER, M. and VARADHAN, S. R. S. (1975). Asymptotics for the Wiener sausages. *Comm. Pure Appl. Math.* **28** 525–565.

[42] DOOB, J. L. (1949). Heuristic approach to the Kolmogorov–Smirnov theorems. *Ann. Math. Statist.* **20** 393–403.

[43] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[44] DVORETZKY, A., KIEFER, J. and WOLFOWITZ, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* **27** 642–669.

[45] DYNKIN, E. B. (1961). *Theory of Markov Processes*. Prentice-Hall, Englewood Cliffs, N.J.

[46] DYNKIN, E. B. (1965). *Markov Processes* **1, 2**. Springer, Berlin.

[47] EINSTEIN, A. (1905). Über die von der molekularkinetischen Theorie der Wärme gefordete Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Physik (4)* **17** 549–560.

[48] ERDÖS, P. and KAC, M. (1947). On the number of positive sums of independent random variables. *Bull. Amer. Math. Soc.* **53** 1011–1021.

[49] ETEMADI, N. (1981). An elementary proof of the strong law of large numbers. *Z. Wahrsch. verw. Gebiete* **55** 119–122.

[50] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York.

[51] EWENS, W. J. (1979). *Mathematical Population Genetics*. Springer, Berlin.

[52] FELLER, W. (1936). Zur Theorie der stochastischen Prozesse (Existenz- und Eindeutigkeitssätze). *Math. Ann.* **113** 113–160.

[53] FELLER, W. (1943). The general form of the so-called law of the iterated logarithm. *Trans. Amer. Math. Soc.* **54** 373–402.

[54] FELLER, W. (1946). The law of the iterated logarithm for identically distributed random variables. *Ann. of Math.* (2) **47** 631–638.

[55] FELLER, W. (1952). The parabolic differential equations and the associated semi-groups of transformations. *Ann. of Math.* (2) **55** 468–519.

[56] FISHER, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford Univ. Press, Oxford.

[57] FISHER, R. A. (1937). The wave of advance of advantageous genes. *Ann. Eugenics* **7** 355–369.

[58] FORTET, R. and MOURIER, E. (1953). Convergence de la répartition empirique vers la répartition théorique. *Ann. Sci. École Norm. Sup.* (3) **70** 267–285.

[59] FRIEDMAN, N. A. (1970). *Introduction to Ergodic Theory*. Van Nostrand Reinhold, New York.

[60] GÄRTNER, J. (1982). Location of wave fronts for the multidimensional KPP equation and Brownian first exit densities. *Math. Nachr.* **105** 317–351.

[61] GEORGII, H.-O. (1988). *Gibbs Measures and Phase Transitions*. de Gruyter, Berlin.

[62] GIKHMAN, I. I. (1947). A method of constructing random processes. *Dokl. Akad. Nauk SSSR* **58** 961–964. (In Russian.)

[63] GIKHMAN, I. I. (1950). Certain differential equations with random functions. *Ukrain. Mat. Zh.* **2**(3) 45–69. (In Russian.)

[64] GIKHMAN, I. I. (1950). On the theory of differential equations of random processes. *Ukrain. Mat. Zh.* **2**(4) 37–63. (In Russian.)

[65] GIKHMAN, I. I. and SKOROKHOD, A. V. (1972). *Stochastic Differential Equations*. Springer, Berlin.

[66] GIKHMAN, I. I. and SKOROKHOD, A. V. (1979). *The Theory of Stochastic Processes* **3**. Springer, Berlin.

[67] GLIVENKO, V. (1933). Sulla determinazione empirica delle leggi di probabilitá. *Giorn. Istit. Ital. Attuari* **4** 92–99.

[68] GNEDENKO, B. V. (1963). The work of A. N. Kolmogorov in the theory of probability. *Russian Math. Surveys* **18**(5) 1–7.

[69] GNEDENKO, B. V. (1973). Andrei Nikolaevich Kolmogorov (on the occasion of his seventieth birthday). *Russian Math. Surveys* **28**(5) 5–16.

[70] GÖDEL, K. (1931). Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme. I. *Monatsh. Math. Phys.* **38** 173–198.

[71] GÖDEL, K. (1933). Zur intuitionistischen Arithmetik und Zahlentheorie. *Ergeb. Math. Kolloq.* **4** 34–38.

[72] GRIMMETT, G. and STIRZAKER, D. (1982). *Probability and Random Processes*. Clarendon Press, Oxford.

[73] HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application*. Academic, New York.

[74] HALMOS, P. R. and VON NEUMANN, J. (1942). Operator methods in classical mechanics. II. *Ann. of Math.* (2) **43** 332–350.

[75] HARRIS, T. E. (1948). Branching processes. *Ann. Math. Statist.* **19** 474–494.

[76] HARRIS, T. E. (1963). *The Theory of Branching Processes*. Springer, Berlin.

[77] HARTMAN, P. and WINTNER, A. (1941). On the law of the iterated logarithm. *Amer. J. Math.* **63** 169–176.

[78] HODGES, A. (1983). *Alan Turing. The Enigma*. Simon and Schuster, New York.

[79] HOPF, E. (1937). *Ergodentheorie*. Springer, Berlin.

[80] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. Kodansha, Tokyo/North-Holland, Amsterdam.

[81] ITÔ, K. (1942). Differential equations determining a Markoff processes. *J. Pan-Japan Math. Colloq.* No. 1077 1352-1400. (In Japanese.)

[82] ITÔ, K. (1942). On stochastic processes (infinitely divisible laws of probability). *Japan. J. Math.* **18** 261–301.

[83] ITÔ, K. (1944). Stochastic integral. *Proc. Imperial Acad. Tokyo* **20** 519–524.

[84] ITÔ, K. (1946). On a stochastic integral equation. *Proc. Imperial Acad. Tokyo* **22** 32–35.

[85] ITÔ, K. (1951). On stochastic differential equations. *Mem. Amer. Math. Soc.* **4**.

[86] ITÔ, K. (1987). *Selected Papers* (D. W. Stroock and S. R. S. Varadhan, eds.). Springer, New York.

[87] ITÔ, K. and MCKEAN, H. P., JR. (1974). *Diffusion Processes and Their Sample Paths*. Springer, Berlin.

[88] JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, Berlin.

[89] KAC, M. (1985). *Enigmas of Chance. An Autobiography*. Univ. California Press, Berkeley.

[90] KENDALL, D. G. (1966). Branching processes since 1873. *J. London Math. Soc.* **41** 385–406.

[91] KENDALL, D. G. (1975). The genealogy of genealogy: Branching processes before (and after) 1873. *Bull. London Math. Soc.* **7** 225–253.

[92] KENDALL, D. G., BATCHELER, G. K. and MOFFATT, H. K. (1987). Obituary. Dr. Andrei Kolmogorov. Giant of mathematics. *The Times* Oct. 26.

[93] KESTERMAN, H. (1937). *Modern Theory of Integration*. Oxford Univ. Press, Oxford.

[94] KHINCHIN, A. YA. (1933). *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*. Springer, Berlin. (Reprinted by Chelsea, New York, 1948.)

[95] KHINCHIN, A. YA. (1937). Déduction nouvelle d'une formule de M. Paul Lévy. *Bull. Univ. État Moscou. Ser. Int., Sect. A, Math. Mecan.* **1** 1–5.

[96] KHINCHIN, A. YA. (1938). The correlation theory of stationary stochastic processes. *Uspekhi Mat. Nauk* **5** 42–51. (In Russian.)

[97] KHINCHIN, A. YA. (1953). On the notion of entropy in the theory of probability. *Uspekhi Mat. Nauk* **8**(3) 3–20. (In Russian.)

[98] KHINCHIN, A. YA. (1956). On the basic theorems of information theory. *Uspekhi Mat. Nauk* **11**(1) 17–75. (In Russian.)

[99] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975 / 76). An approximation of partial sums of independent rv's and the sample df. I, II. *Z. Wahrsch. verw. Gebiete* **32** 111–131; **34** 33–58.

[100] KORNFEL'D, I. P., SINAI, YA. G. and FOMIN, S. V. (1980). *Ergodic Theory*. Nauka, Moscow. (In Russian.)

[101] KRENGEL, U. (1985). *Ergodic Theorems*. de Gruyter, Berlin.

[102] KRYLOV, N. M. and BOGOLYUBOV, N. N. (1937). La théorie génerale de la mesure dans son application à l'étude des systémes dynamiques de la mécanique nonlinéaire. *Ann. of Math.* (*2*) **38** 65–113.

[103] LE GALL, J.-F. (1986). Sur la saucisse de Wiener et les points multiples du mouvement brownien. *Ann. Probab.* **14** 1219–1244.

[104] LEONOV, V. P. (1964). *Some Applications of High Order Semiinvariants to the Theory of Stationary Random Processes*. Nauka, Moscow. (In Russian.)

[105] LEONTOVICH, M. A. (1933). Zur Statistik der kontinuierlichen Systeme und des zeitlichen Verlaufes der physikalischen Vorgänge. *Physik. Z. Sowjetunion* **3** 35–63.

[106] LEONTOVICH, M. A. (1935). Basic equations of kinetic gas theory from the point of view of the theory of random processes. *Zh. Èksper. Teoret. Fiz.* **5** 211–231. (In Russian.)

[107] LEVIN, L. A. (1973). On the notion of a random sequence. *Soviet Math. Dokl.* **14** 1413–1416.

[108] LÉVY, P. (1934). Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. *Ann. Scuola Norm. Sup. Pisa* (*2*) **3** 337–366.

[109] LÉVY, P. (1937). *Théorie de l'addition des variables aléatoires*. Gauthier-Villars, Paris.

[110] LÉVY, P. (1948). *Processus stochastiques et mouvement brownien*. Gauthier-Villars, Paris.

[111] LÉVY, P. (1970). *Quelques aspects de la pensée d'un mathématicien*. Blanchard, Paris.

[112] LI, M. and VITÁNYI, P. M. B. (1988). Two decades of applied Kolmogorov complexity. *Proc. IEEE Third Structure in Complexity Theory Conference, Washington, June, 1988* 1–23. (Also Report TR-02-88, Center for Research in Computing Technology, Harvard Univ.)

[113] LINDE, W. (1986). *Probability in Banach Spaces—Stable and Infinitely Divisible Distributions*, 2nd ed. Wiley, Chichester.

[114] LIPTSER, R. S. and SHIRYAEV, A. N. (1977 / 78). *Statistics of Random Processes* **1**, **2**. Springer, New York.

[115] LIPTSER, R. S. and SHIRYAEV, A. N. (1986). *Theory of Martingales*. Nauka, Moscow. (In Russian.)

[116] LOÈVE, M. (1948). Fonctions aléatoires du second ordre. Appendix to P. Lévy, *Processus stochastiques et mouvement brownien*. Gauthier-Villars, Paris.

[117] ŁOMNICKI, A. (1923). Nouveaux fondements du calcul des probabilités. *Fund. Math* **4** 34–71.

[118] LORENTZ, G. G. (1976). The 13th problem of Hilbert. In *Mathematical Developments Arising from Hilbert's Problems* (F. E. Browder, ed.). *Proc. Symp. Pure Math.* **28** 419–430. Amer. Math. Soc., Providence, R.I.

[119] LOVELAND, D. W. (1966). A new interpretation of the von Mises concept of random sequence. *Z. Math. Logik Grundlag. Math.* **12** 279–294.

[120] LOVELAND, D. W. (1966). The Kleene hierarchy classification of recursively random sequences. *Trans. Amer. Math. Soc.* **125** 497–510.

[121] MARCINKIEWICZ, J. and ZYGMUND, A. (1937). Remarque sur la loi du logarithme itéré. *Fund. Math.* **29** 215–222.

[122] MARKOV, A. A. (1906). Extension of the law of large numbers to dependent events. *Bull. Soc. Phys. Math. Kazan (2)* **15** 135–156. (In Russian.)

[123] MARKOV, A. A. (1951 / 54). The theory of algorithms. *Trudy Mat. Inst. Steklov.* **38** 176–189; **42** 7–375. (In Russian.)

[124] MARTIN, N. F. G. and ENGLAND, J. W. (1981). *Mathematical Theory of Entropy*. Addison-Wesley, Reading, Mass.

[125] MARTIN-LÖF, P. (1966). The definition of random sequences. *Inform. and Control* **9** 602–619.

[126] MARUYAMA, G. (1949). The harmonic analysis of stationary stochastic processes. *Mem. Fac. Sci. Kyushu Univ. Ser. A* **4** 45–106.

[127] MARUYAMA, G. (1955). Continuous Markov processes and stochastic equations. *Rend. Circ. Mat. Palermo (2)* **4** 48–90.

[128] MCKEAN, H. P., JR. (1969). *Stochastic Integrals*. Academic, New York.

[129] MCKEAN, H. P., JR. (1975). Application of Brownian motion to the equation of Kolmogorov–Petrovskii–Piskunov. *Comm. Pure Appl. Math.* **28** 323–331.

[130] MESHALKIN, L. D. (1960). A lower estimate of the rate of approach of the distributions of sums to a set of infinitely divisible laws. *Soviet Math. Dokl.* **1** 648–651.

[131] MESHALKIN, L. D. (1960). On the approximation of polynomial distributions by infinitely divisible laws. *Theory Probab. Appl.* **5** 106–114.

[132] MIROSHNIKOV, A. L. and ROGOZIN, B. A. (1980). Inequalities for the concentration function. *Theory Probab. Appl.* **25** 176–180.

[133] MIROSHNIKOV, A. L. and ROGOZIN, B. A. (1982). Remarks on an inequality for the concentration function of sums of independent variables. *Theory Probab. Appl.* **27** 848–850.

[134] MOFFATT, H. K. (1981). Some developments in the theory of turbulence. *J. Fluid Mech.* **106** 27–47.

[135] MONIN, A. S. and YAGLOM, A. M. (1971 / 75). *Statistical Fluid Mechanics: Mechanics of Turbulence* **1**, **2**. MIT Press, Cambridge, Mass.

[136] MOSTELLER, F. and WALLACE, D. L. (1964). *Inference and Disputed Authorship: The Federalist*. Addison-Wesley, Reading, Mass. [2nd ed. (1984), *Applied Bayesian and Classical Inference. The Case of the Federalist Papers*. Springer, New York.]

[137] NAGAEV, S. V. (1970). On the speed of convergence in a boundary value problem. I, II. *Theory Probab. Appl.* **15** 163–186, 403–429.

[138] NELSON, E. (1967). *Dynamical Theories of Brownian Motion*. Princeton Univ. Press, Princeton, N.J.

[139] NIKODYM, O. (1930). Sur une généralisation des intégrales de M. J. Radon. *Fund. Math.* **15** 131–179.

[140] NIKOL'SKII, S. M. (1983). Aleksandrov and Kolmogorov in Dnepropetrovsk. *Russian Math. Surveys* **38**(4) 41–55.

[141] OBUKHOV, A. M. (1962). Some specific features of atmospheric turbulence. *J. Fluid Mech.* **13** 77–81.

[142] OBUKHOV, A. M. (1983). Kolmogorov flow and laboratory simulation of it. *Russian Math. Surveys* **38**(4) 113–126.

[143] ORNSTEIN, D. (1974). *Ergodic Theory, Randomness and Dynamical Systems*. Yale Univ. Press, New Haven, Conn.

[144] PAPANICOLAOU, G. C., STROOCK, D. W. and VARADHAN, S. R. S. (1977). Martingale approach to some limit theorems. In *Papers from the Duke Turbulence Conference. Duke Univ. Math. Ser.* **3**, No. 6.

[145] PARRY, W. (1981). *Topics in Ergodic Theory. Cambridge Tracts in Mathematics* **75**. Cambridge Univ. Press, Cambridge.

[146] PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces*. Academic, New York.

[147] PETROVSKII, I. G. (1935). Zur ersten Randwertaufgabe der Wärmeleitungsgleichung. *Compositio Math.* **1** 383–419.

[148] PETROVSKII, I. G. (1987). *Selected Papers. Differential Equations. Theory of Probability.* Nauka, Moscow. (In Russian.)

[149] PIOTROVSKII, R. G. (1968). *Information Measurement of Speech (Language)*. Nauka, Leningrad. (In Russian.)

[150] POINCARÉ, H. (1892 / 93 / 99). *Les méthodes nouvelles de la mécanique céleste* **1**, **2**, **3**. Gauthier-Villars, Paris. (Reprinted by Dover, New York, 1957.)

[151] PONTRYAGIN, L. S., ANDRONOV, A. A. and VITT, A. A. (1933). On statistical consideration of dynamical systems. *Zh. Èksper. Teoret. Fiz.* **3** 165–180. (In Russian.)

[152] POST, E. L. (1936). Finite combinatory processes—formulation. I. *J. Symbolic Logic* **1**(3) 103–105.

[153] PROKHOROV, A. V. (1970). Teoria prawdopodobienstwa w badaniach rytmu wiersza. *Pamietnik Literacki* **61**(3) 113–127. (In Polish.)

[154] PROKHOROV, A. V. (1984). On random versification (to the problem of theoretical and speech models of poetry). In *Problems of the Theory of Verse* 89–98. Nauka, Leningrad. (In Russian.)

[155] PROKHOROV, YU. V. (1953). Distributions of probabilities in functional spaces. *Uspekhi Mat. Nauk* **8**(3) 165–167. (In Russian.)

[156] PROKHOROV, YU. V. (1955). On sums of identically distributed random variables. *Dokl. Akad. Nauk SSSR* **105** 645–647. (In Russian.)

[157] PROKHOROV, YU. V. (1956). Convergence of stochastic processes and limit theorems in probability theory. *Theory Probab. Appl.* **1** 157–214.

[158] PROKHOROV, YU. V. (1960). On a uniform limit theorem of A. N. Kolmogorov. *Theory Probab. Appl.* **5** 98–106.

[159] ROGERS, L. C. G. and WILLIAMS, D. (1987). *Diffusions, Markov Processes and Martingales* **2**: *Itô Calculus*. Wiley, New York.

[160] ROGOZIN, B. A. (1961). An estimate for concentration functions. *Theory Probab. Appl.* **6** 94–97.

[161] ROZANOV, YU. A. (1967). *Stationary Random Processes*. Holden-Day, San Francisco.

[162] ROZANOV, YU. A. (1974). *Theory of Innovation Processes*. Nauka, Moscow. (In Russian.)

[163] SAKHANENKO, A. I. (1974). On the speed of convergence in a boundary problem. *Theory Probab. Appl.* **19** 399–403.

[164] SAKS, S. (1937). *Theory of the Integral*. Hafner, New York.

[165] SAVCHUK, A. P. (1964). On estimates for the entropy of a language according to Shannon. *Theory Probab. Appl.* **9** 138–140.

[166] SCHNORR, C.-P. (1971). *Zufälligkeit und Wahrscheinlichkeit. Eine algorithmische Behandlung der Wahrscheinlichkeithstheorie. Lecture Notes in Math.* **218**. Springer, Berlin.

[167] SCHNORR, C.-P. (1972). Optimal Gödel numberings. In *Information Processing 71: Proc. IFIP Congress 1971 Ljubljana* **1** 56–58. North-Holland, Amsterdam.

[168] SCHNORR, C.-P. (1975). Optimal enumerations and optimal Gödel numberings. *Math. Systems Theory* **8**(2) 182–191.

[169] SCHRÖDINGER, E. (1931). Über die Umkehrung der Naturgesetze. *Sitzungsber. Preuss. Akad. Wiss. Phys. Math. Kl.* 144–153.

[170] SEVAST'YANOV, B. A. (1971). *Branching Processes*. Nauka, Moscow. (In Russian.)

[171] SHANNON, C. E. (1951). Prediction and entropy of printed English. *Bell System Tech. J.* **30** 50–64.

[172] SHIELDS, P. (1973). *The Theory of Bernoulli Shifts*. Univ. Chicago Press, Chicago.

[173] SHIRYAEV, A. N. (1960). Some problems in the spectral theory of higher-order moments. I. *Theory Probab. Appl.* **5** 265–284.

[174] SHIRYAEV, A. N. (1963). On conditions for ergodicity of stationary processes in terms of higher order moments. *Theory Probab. Appl.* **8** 436–439.

[175] SHIRYAEV, A. N. (1976). *Optimal Stopping Rules*. Springer, Berlin.

[176] SHIRYAEV, A. N. (1984). *Probability*. Springer, Berlin.

[177] SIEGEL, C. L. (1954). Über die Existenz einer Normalform analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtslösung. *Math. Ann.* **128** 144–170.

[178] SIEGEL, C. L. and MOSER, J. K. (1971). *Lectures on Celestial Mechanics*. Springer. Berlin.

[179] SINAI, YA. G. (1963). On higher order spectral measures of ergodic stationary processes. *Theory Probab. Appl.* **8** 429–436.

[180] SINAI, YA. G. (1976). *Introduction to Ergodic Theory*. Princeton Univ. Press, Princeton, N.J.

[181] SINAI, YA. G. (1982). *Theory of Phase Transitions: Rigorous Results*. Pergamon, New York.

[182] SKOROKHOD, A. V. (1965). *Studies in the Theory of Random Processes*. Addison-Wesley, Reading, Mass.

[183] SMIRNOV, N. V. (1944). On approximation of distribution laws of random variables by empirical data. *Uspekhi Mat. Nauk* **10** 179–206. (In Russian.)

[184] SMOLLER, J. (1983). *Shock Waves and Reaction-Diffusion Equations*. Springer, Berlin.

[185] SOLOMONOFF, R. J. (1964). A formal theory of inductive inference. I, II. *Inform. and Control* **7** 1–22, 224–254.

[186] STEFFENSON, J. F. (1933). Deux problèmes du calcul des probabilités. *Ann. Inst. H. Poincaré* **3** 319–344.

[187] STOUT, W. F. (1974). *Almost Sure Convergence*. Academic, New York.

[188] STRASSEN, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. verw. Gebiete* **3** 211–226.

[189] STRASSEN, V. (1966). A converse to the law of the iterated logarithm. *Z. Wahrsch. verw. Gebiete* **4** 265–268.

[190] STROOCK, D. W. and VARADHAN, S. R. S. (1969). Diffusion processes with continuous coefficients. I, II. *Comm. Pure Appl. Math.* **22** 345–400, 479–530.

[191] STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes*. Springer, Berlin.

[192] TAQQU, M. S. and LEVY, J. B. (1986). Using renewal processes to generate long-range dependence and high variability. In *Dependence in Probability and Statistics* (E. Eberlein and M. S. Taqqu, eds.) 73–89. Birkhäuser, Boston.

[193] TURING, A. M. (1937). On computable numbers, with an application to the Entscheidungsproblem. *Proc. London Math. Soc.* (2) **42** 230–265.

[194] UHLENBECK, G. E. and ORNSTEIN, L. S. (1930). On the theory of Brownian motion. *Phys. Rev.* **36** 823–841.

[195] UL'YANOV, P. L. (1983). Kolmogorov and divergent Fourier series. *Russian Math. Surveys* **38**(4) 57–100.

[196] VAKHANIYA, N. N., TARIELADZE, V. I. and CHOBANYAN, S. A. (1985). *Probability Distributions in Banach Spaces*. Nauka, Moscow. (In Russian.)

[197] VARADHAN, S. R. S. (1984). *Large Deviations and Applications*. SIAM, Philadelphia.

[198] VERWAAT, W. (1985). Sample path properties of self-similar processes with stationary increments. *Ann. Probab.* **13** 1–27.

[199] VOLKONSKII, V. A. and ROZANOV, YU. A. (1959 / 61). Some limit theorems for random functions. I, II. *Theory Probab. Appl.* **4** 178–197, **6** 186–198.

[200] VON MISES, R. (1919). Grundlagen der Wahrscheinlichkeitsrechnung. *Math. Z.* **5** 52–99.

[201] VON MISES, R. (1928). *Wahrscheinlichkeit, Statistik und Wahrheit*. Springer, Wien. (*Probability, Statistics and Truth*, 2nd rev. English ed. Dover, New York, 1981.)

[202] VON MISES, R. (1931). *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und Theoretischen Physik*. Franz Deuticke, Leipzig.

[203] VON MISES, R. (1964). *Mathematical Theory of Probability and Statistics*. Academic, New York.

[204] VON NEUMANN, J. (1932). Zur Operatorenmethode in der klassischen Mechanik. *Ann. of Math.* (2) **33** 587–642.
[205] VON NEUMANN, J. (1932). *Mathematische Grundlagen der Quantenmechanik*. Springer, Berlin.
[206] VON SMOLUCHOWSKI, M. (1916). Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen. *Physik. Z.* **17** 557–571.
[207] VOVK, V. G. (1987). The law of the iterated logarithm for random Kolmogorov, or chaotic, sequences. *Theory Probab. Appl.* **32** 413–425.
[208] WALTERS, P. (1982). *An Introduction to Ergodic Theory*. Springer, Berlin.
[209] WELCHMAN, G. (1982). *The Hut Six Story. Breaking the Enigma Codes*. Allen Lane, London.
[210] WIENER, N. (1949). *The Extrapolation, Interpolation and Smoothing of Stationary Time Series*. MIT Press, Cambridge, Mass.
[211] WILLIAMS, D. (1979). *Diffusions, Markov Processes and Martingales* **1**: *Foundations*. Wiley, New York.
[212] YAGLOM, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions* **1**, **2**. Springer, New York.
[213] YAGLOM, A. M. and YAGLOM, I. M. (1983). *Probability and Information*. Reidel, Dordrecht, Holland.
[214] ZHURBENKO, I. G. (1982). *Spectral Analysis of Time Series*. Izd. Moskov. Gos. Univ., Moscow. (In Russian.)
[215] ZVONKIN, A. K. and LEVIN, L. A. (1970). The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russian Math. Surveys* **25**(6) 83–124.
[216] ZYGMUND, A. (1959). *Trigonometric Series* **1**, **2**, 2nd ed. Cambridge Univ. Press, Cambridge.