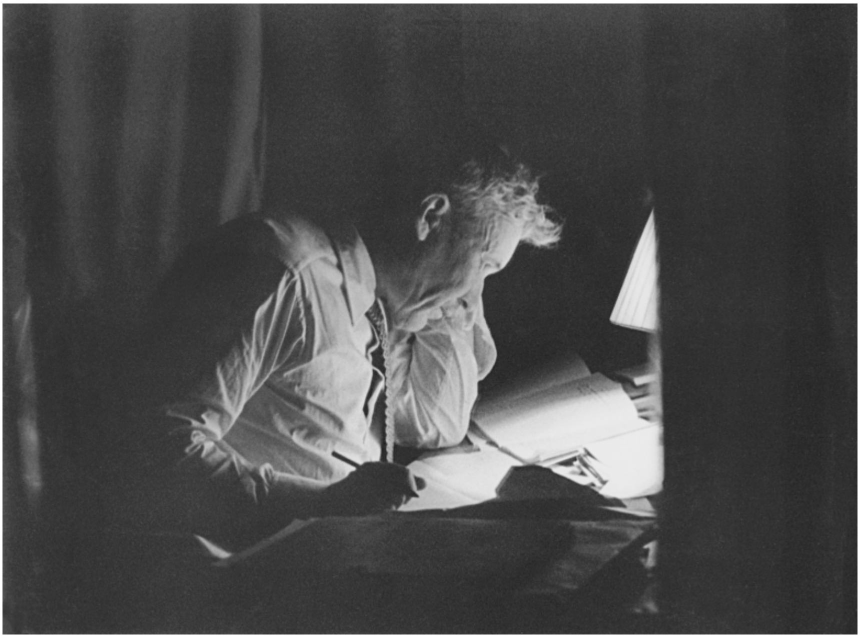
Kolmogorov's Heritage in Mathematics



Picture taken from the Kolmogorov Archive with kind permission of A.N. Shiryaev

Éric Charpentier · Annick Lesne
Nikolaï K. Nikolski (Eds.)

Kolmogorov's Heritage in Mathematics

With 22 Figures

Springer

Éric Charpentier
Institut de Mathématiques
Université Bordeaux 1
351 cours de la Libération
33405 Talence Cedex
France
eric.charpentier@math.u-bordeaux1.fr

Nikolaï K. Nikolski
Institut de Mathématiques
Université Bordeaux 1
351 cours de la Libération
33405 Talence Cedex
France
Nikolai.Nikolski@math.u-bordeaux1.fr

Annick Lesne
Laboratoire de Physique Théorique
de la Matière Condensée
Université Pierre et Marie Curie
Case 121, Tour 24-25 - 4ème étage, pièce 407
4 Place Jussieu
75252 Paris
France

And

Institut des Hautes Études Scientifiques
Le Bois-Marie, 35 route de Chartres
91440 Bures-sur-Yvette
France
lesne@lptmc.jussieu.fr

Édition originale: *L'héritage de Kolmogorov en mathématiques* © Éditions Belin – 2004
Ouvrage publié avec le concours du Ministère français chargé de la culture – Centre national du livre
This edition was supported by the French Ministry of Culture – Centre national du livre

Cover picture taken from the Kolmogorov Archive with kind permission of A. N. Shiryaev

Library of Congress Control Number: 2007923591

ISBN 978-3-540-36349-1 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media
springer.com

Typesetting: by the authors and Integra, India using a Springer LaTeX macro package
Cover design: WMX-Design, Heidelberg

Printed on acid-free paper SPIN: 11775072 5 4 3 2 1 0

Contents

List of Contributors

Vasco Brattka
Laboratory of Foundational Aspects of Computer Science
Department of Mathematics and Applied Mathematics
University of Cape Town
South Africa
`http://cca-net.de/vasco`
`BrattkaV@maths.uct.ac.za`

Victor M. Buchstaber
Steklov Mathematical Institute of Russian Academy of Sciences, and Moscow State University, Russia
`http://www.mi.ras.ru/~buchstab`
`buchstab@mi.ras.ru`
and School of Mathematics, The University of Manchester, UK,
`Victor.Buchstaber@manchester.ac.uk`

Loïc Chaumont
Laboratoire Angevin de Recherche en Mathématiques (LAREMA)
University of Angers, France
`http://math.univ-angers.fr/~chaumont`
`loic.chaumont@univ- angers.fr`

Thierry Coquand
Department of Computer Science and Engineering
Chalmers University of Technology and Göteborg University
Sweden
`http://www.cs.chalmers.se/~coquand`
`coquand@cs.chalmers.se`

Giuseppe Da Prato
Scuola Normale Superiore
Pisa, Italy
`http://www.sns.it/scienze/menunews/docentiscienze/daprato`
`daprato@sns.it`

Bruno Durand
Laboratoire d'informatique fondamentale de Marseille
University of Provence (Aix-Marseille I), France
`http://www.lif.univ-mrs.fr/~bdurand`
`Bruno.Durand@lif.univ-mrs.fr`

Kevin Ford
Department of Mathematics
University of Illinois at Urbana-Champaign, USA
`http://www.math.uiuc.edu/~ford`
`ford@math.uiuc.edu`

Étienne Ghys
Unité de mathématiques pures et appliquées (UMPA)

CNRS and École normale supérieure de Lyon, France
http://www.umpa.ens-lyon.fr/~ghys
etienne.ghys@umpa.ens-lyon.fr

John H. Hubbard
Department of Mathematics
Cornell University, Ithaca, NY, USA
http://www.math.cornell.edu/People/Faculty/hubbard.html
jhh8@cornell.edu

J.-P. Kahane
Département de Mathématiques d'Orsay
University Paris-sud, Orsay, France
http://www.academie-sciences.fr/Membres/K/Kahane_JP.htm
Jean-Pierre.Kahane@math.u-psud.fr

Denis V. Kosygin
Courant Institute of Mathematical Sciences
New York University, USA
kosygin@cims.nyu.edu

Laurent Mazliak
Laboratoire de Probabilités et Modèles Aléatoires
Institut de mathématiques de Jussieu
University Paris VI, France
http://www.proba.jussieu.fr/users/lma/mazliak.html
mazliak@ccr.jussieu.fr

Mikhaïl Nikouline (M. Nikulin)
Statistique Mathématique et ses Applications
University Victor Segalen (Bordeaux II), France
http://www.sm.u-bordeaux2.fr/stat
nikouline@sm.u-bordeaux2.fr

Karl Sigmund
Faculty for Mathematics
University of Vienna
and International Institute for Applied Systems Analysis, Austria
http://homepage.univie.ac.at/Karl.Sigmund
karl.sigmund@univie.ac.at

Yakov G. Sinai
Department of Mathematics
Princeton University, USA
sinai@Math.Princeton.EDU

Valentin Solev
Laboratory of Statistical Methods,
Steklov Institute of Mathematics,
RAS, Saint Petersburg, Russia
http://www.pdmi.ras.ru/staff/solev.html
solev@pdmi.ras.ru

Vladimir M. Tikhomirov
Moscow State University and Independent University of Moscow, Russia

Paul Vitanyi
Centrum voor Wiskunde en Informatica (CWI), Amsterdam
The Netherlands
http://www.cwi.nl/~paulv
Paul.Vitanyi@cwi.nl

Marc Yor
Laboratoire de Probabilités et Modèles Aléatoires
Institut de mathématiques de Jussieu
University Paris VI, France
http://www.academie-sciences.fr/Membres/Y/Yor_Marc.htm

Alexander Zvonkin
Laboratoire Bordelais de Recherche en Informatique (LaBRI)
University Bordeaux I, France
http://www.labri.fr/perso/zvonkin
zvonkin@labri.fr

Introduction

Éric Charpentier[1], Annick Lesne[2], and Nikolaï Nikolski[3]

[1] Institut de Mathématiques de Bordeaux, University Bordeaux I, France
`http://www.math.u-bordeaux1.fr/imb/`
`eric.charpentier@math.u-bordeaux1.fr`

[2] Laboratoire de physique théorique de la matière condensée, University Paris VI, and Institut des Hautes Études Scientifiques, Bures-sur-Yvette, France
`http://www.lptmc.jussieu.fr/users/lesne/`
`lesne@lptmc.jussieu.fr`

[3] Institut de Mathématiques de Bordeaux, University Bordeaux I, France
`http://www.math.u-bordeaux1.fr/imb/`
`Nikolai.Nikolski@math.u-bordeaux1.fr`

Translated from the French by Elizabeth Strouse

Andrei Nikolaevich Kolmogorov (Tambov 1903, Moscow 1987) was one of the most brilliant mathematicians that the world has ever known. Incredibly deep and creative, he was able to approach each subject with a completely new point of view: in a few magnificent pages, which are models of shrewdness and imagination, and which astounded his contemporaries, he changed drastically the landscape of the subject. Most mathematicians prove what they can, Kolmogorov was of those who prove what they want.

In this book we have asked several world experts to present (one part of[4]) the mathematical heritage left to us by Kolmogorov[5]. Each chapter treats one of Kolmogorov's research themes, or a subject that was invented as a consequence of his discoveries. We present here his contributions, his methods, the perspectives he opened to us, the way in which this research has evolved up to now, along with examples of recent applications and a presentation of the modern prospects.

We hope that this book can be read by anyone with a master's (or even a bachelor's) degree in mathematics, computer science or physics, or more

[4] In the book *Kolmogorov in perspective* (History of Mathematics, Vol. 20, American Mathematical Society, 2000) one can find a (more or less) complete bibliography of Kolmogorov's work, which consists of about 500 publications. A lot of these are at the heart of very active research going on today

[5] A book entitled *The Kolmogorov Legacy in Physics* has already appeared (Springer, 2004). It contains contributions to dynamical systems, complexity, turbulence and probability. We strongly recommend it to mathematicians

generally by anyone who likes mathematical ideas. Rather than presenting detailed proofs, we give the main ideas, and a bibliography for those who wish to understand the technical details. One can see that sometimes very simple reasoning (with the right interpretation and tools) can lead in a few lines to very substantial results.
Here is a quick summary of the themes that are treated here, with, for each, some significant examples of the "master's touch".

Fourier Series (Chap. 1). In 1922, at the age of 19, Kolmogorov managed to construct a Lebesgue integrable function on $[0, 2\pi]$ whose Fourier series diverges almost everywhere, that is, except on a set of Lebesgue measure 0. (Up to this point the Fourier series of all the functions that had been considered converged almost everywhere.) Three years later, he found an example of a function whose Fourier series diverged *everywhere*! In the same time, Kolmogorov obtained important results on lacunary Fourier series, harmonic conjugates...

Logic (Chap. 2). In 1925, Kolmogorov became interested in intuitionistic logic. For Brouwer (the father of Intuitionism), Intuitionism and Formalism were antinomic approaches. Moreover, intuitionistic logic was generally considered as a weakening of classical logic. But, baffling all expectations, *Kolmogorov managed to formalize intuitionistic logic* and to present it as an *extension* of classical logic. He then deduced that any "finitary" proposition with a classical proof could be proved using intuitionistic logic. All this in only two pages!

If intuitionistic logic is perceived as an extension of classical logic, obtained by adding connectors with no analogues in classical logic, one needs to find an interpretation of these connectors. This is what Kolmogorov did in 1932 when he interpreted intuitionistic logic as a "calculus of problems". This interpretation later turned out to be relevant to computer science.

Probabilities (Chaps. 3 and 4). It was also around 1925 that Kolmogorov began working on probability theory. He began with a clever generalization of the Bienaymé-Chebyshev inequality: this "Kolmogorov's inequality" quickly revealed its usefulness. He used it, with Khinchin, to obtain a famous convergence rule for series of random variables.

In 1930 he used the inequality to deduce a version of the strong law of large numbers ("almost sure" convergence – i.e. convergence with probability equal to one – of the empirical mean towards the mathematical expectation) which contained all prior versions of this law (Borel, Cantelli, Khinchin) like a set of Russian dolls: this is the "L^2 version" given in Chap. 3. A little later he obtained the "L^1 version" which is in a certain sense optimal since it holds whenever the variables (independant and with the same law) have a finite expectation. Kolmogorov announced this result in his book *Foundations of probability theory*, in 1933, without giving a proof; he explained the proof to Maurice Fréchet and let him have the honor of being the first to publish it in the first book of his *Recherches théoriques modernes sur le calcul des probabilités* (Gauthier-Villars, 1937).

In 1931, Kolmogorov began thinking about continuous time Markov processes. Instead of studying the trajectories of a process, he studied the transition probability densities and determined the differential equations they satisfied. These are, of course, the fundamental equations from which the modern theory of diffusion has been constructed. Because of the applications to modern physics, they have even been generalized to Hilbert spaces of infinite dimension: Chapter 4 describes this active area of research which has had several recent applications.

Lomnicki, Steinhaus, Cantelli had sketched out axiomatizations of probability theory, based on Borel's idea to formulate it in the langage of measure theory; but it is Kolmogorov, in 1933, who solved the problem: his axiomatization, both natural and powerful, is the one that is still used today.

Statistics (Chaps. 5 and 6). In 1933, Glivenko et Cantelli proved what is sometimes called the *fundamental theorem of statistics*: almost surely, the empirical distribution function of a real random variable converges uniformly towards the true distribution function when the size of the sample tends towards infinity. Soon afterwards, Kolmogorov gave the precise law for the convergence, in terms of a universal function (independant of the – sought after – law of the random variable) which leads to a very efficient goodness-of-fit test (a test for the unknown law).

The Chap. 5 retraces this discovery of Kolmogorov and presents certain improvements with a surprising application to number theory: it gives the asymptotic behavior (when n tends to infinity) of the probability that an integer chosen at random has at least one divisor between n and $2n$, and of the probability that it has exactly r such divisors.

Epsilon-entropy, a quasi-universal tool introduced by Kolmogorov in the 1950's (see Chap. 8) has become a fundamental tool in statistics for measuring the quality of estimators: the Chap. 6 gives a survey of the current state of knowledge, intended for master's degree students as well as for researchers.

Topology (Chap. 7). During the years 1934–1937, Kolmogorov was very excited about topology. From 1934 to 1935, at the same time as Alexander, he constructed the cohomology of topological spaces and discovered its ring structure. In 1935, Samuel Eilenberg asked if an open mapping (one that takes open sets to open sets) can increase the dimension of its domain; Kolmogorov answered this question with a three page article, which appeared in 1937, in which he gave a very clever construction of an open mapping which transforms a topological space of dimension 1 into a topological space of dimension 2. This result was later used to analyze groups' actions on topological spaces.

Geometry (Chap. 8). One of Kolmogorov's great gifts was his extraordinary geometrical intuition. This played an essential role in almost all his work, as shown in the fascinating Chap. 8. This chapter describes the contributions of Kolmogorov to geometry, approximation theory, the invention of ε-entropy, etc.

Mathematical Ecology (Chap. 9). In 1936, Kolmogorov published a short note on the predator-prey model of Lotka and Volterra: they had (independently) expressed the rates of growth of populations by explicit functions, depending on a small number of parameters, and they solved the equations explicitely. Clearly such a method can only apply to highly simplified models. Kolmogorov, on the other hand, did not choose specific functions, but instead concentrated on monotonicity properties (which give robust conditions); then he applied the qualitative methods of Poincaré to study the long-time behavior of populations (equilibria, limit cycles). Of course, this qualitative approach is the only one that can be used in realistic models.

Dynamical Systems. KAM Theorem (Chaps. 10, 11) and entropy (Chap. 12). Around 1953, Kolmogorov became interested in dynamical systems. He proved a fundamental theorem which explains why, when a system stays close enough to a system in which numerous laws of conservation hold, most movements of the former remain close to the regular movements of the latter, instead of capitalizing on the lack of laws to wander. Kolmogorov presented his proof in talks, but did not publish it. Arnold and Moser later published the first proofs of Kolmogorov's theorem (with different hypotheses) and this is where the name of the KAM (Kolmogorov, Arnold, Moser) theorem comes from. For more than 30 years, all known proofs of this theorem were extremely complicated, but in 1984 a wonderfully simple proof was published. This proof was improved upon in 2002. When one of the authors of this improved proof explained it in Moscow in 2002, members of the audience who had heard Kolmogorov in 1957 said that the new proof was actually the same as Kolmogorov's original one! (Chap. 11, p. 215.) Chapter 10 gives a historical survey of the problem of stability of movements in Celestial mechanics, explains the role of resonance and small divisors and presents an analogue of KAM theorem in a toy model where mathematical difficulties are weakened. Chapter 11 illustrates the KAM theorem in the cases of the solar system and of the forced pendulum; then, it gives a rigorous statement of the theorem and sketches the main ideas of the new proof.

In 1958, Kolmogorov introduced the idea of entropy of dynamical systems, a quantity which is invariant by metric isomorphisms. This turned out to be a very powerful tool which permits one to show that two dynamical systems are not metrically isomorphic. At first, Kolmogorov thought that deterministic systems determined by differential equations would necessarily have zero entropy (roughly: "no disorder") unlike probabilistic systems. But Kolmogorov and Sinai noticed that there exist deterministic systems with nonzero entropy. This discovery is one of the starting points of the modern theory of *deterministic chaos*.

The modern (and fundamental) notion of *hyperbolicity* of a dynamical system was developed as a result of the investigation of these deterministic systems with nonzero entropy. Chapter 12 describes this discovery and the latest results on the subject.

The superposition theorem (chaps. 8 and 13). The 13th Hilbert problem (Paris, 1900) was to determine if any continuous function of three variables could be constructed in a finite number of steps, each one an application of a continuous function of *one* or *two* variables. (This was an idealization of a method of graphical resolution of equations.) Hilbert expected a negative answer. In 1956 Kolmogorov showed that any continuous function of n variables on $[0,1]^n$ is constructible if one permits the use of continuous functions of *three* variables as auxiliaries, and in 1957 his student Arnold proved that any continuous function of three variables is constructible using functions of *two* variables, thus resolving Hilbert's 13th problem (with a different result than that expected by Hilbert). Soon after, Kolmogorov showed that any continuous function of n variables on $[0,1]^n$ is constructible using only continuous functions of one variable and additions: this is the *Kolmogorov's Superposition Theorem.*

In 1987, Hecht-Nielsen deduced that any continuous function could be implemented by a certain type of neural network with continuous activation functions and real weights. Constructive versions of the superposition theorem have recently made possible the transposition of this result to *computable* functions, activation functions and weights, and may give a way to construct networks corresponding to given functions. This application of the superposition theorem to neural networks would surely have pleased Kolmogorov: in fact, according to Arnold, one of Kolmogorov's last mathematical works was motivated by his curiosity about the structure of the brain (cf. chapitre 9, p. 177).

More recently other applications of the superposition theorem have appeared: Kolmogorov would surely have been happy to know that it applies to subjects such as the Radon transform and topological groups.

Complexity of description (Chaps. 14 and 15). In 1965, Kolmogorov defined the *complexity of description* of an object; this is, more or less, the length of the shortest algorithm which can describe this object (the exact definition is given in Chap. 14). This is a wonderful tool. One of its applications was a surprisingly simple and short proof by Gregory Chaitin of Gödel's incompleteness theorem (Chap. 14), which can be understood by everybody! Chaitin was very pleased to find this proof which, he explained, gives one the impression that the incompleteness phenomenon discovered by Gödel is natural – unlike the traditional proof, based on the "liar paradox" which makes the phenomenon seem rather pathological and uncommon.

Kolmogorov's complexity theory also makes sense of propositions such as "the sequence is random": for Kolmogorov this means more or less that the sequence has no regularity with which one can "summarize" it, it is "incompressible".

The existence of such incompressible objects gives a simple way to prove things in all branches of mathematics: this is the "incompressibility method". Two examples of this method are given in Chap. 15, one in number theory (how to show in a few lines, and almost without calculation, that the nth prime number has an order of magnitude less than $n \log^2 n$, for instance), and the other in graph theory (finding the maximum size of complete graphs contained in a random graph).

But the principal object of Chap. 15 is to propose a new approach to deterministic chaos: instead of studying the unpredictibility in terms of the probabilities of ensembles of trajectories, one uses the complexity theory of Kolmogorov to express the fact that an *individual* trajectory is "unpredictable".

Acknowledgements

We of course thank all our authors for the wonderful texts they contributed.

Our gratitude also extends to the translators and to the colleagues whose comments and suggestions made this book possible: Michel Balazard, Pierre Castéran, Robert Deville, Étienne Matheron, Cyril Mauvillain, Michel Mendès France, Vesselin Petkov, Elizabeth Strouse, Philippe Thieullen, John Tromp and Alexander Zvonkin.

1

The youth of Andrei Nikolaevich and Fourier series

Jean-Pierre Kahane

Département de Mathématiques d'Orsay, University Paris-sud, Orsay, France
http://www.academie-sciences.fr/Membres/K/Kahane_JP.htm
Jean-Pierre.Kahane@math.u-psud.fr

Translated from the French by Elizabeth Strouse

Andrei Nikolaevich Kolmogorov was 14 years old in October 1917. In Moscow this was neither a time for peace nor for reflection. According to his biography in [Kol91], it seems that he was a railroad worker during the difficult years of 1919 and 1920. In 1920–21 he began his studies at the University of Moscow where he would become a student of N.N. Lusin. At the same time he was conducting a study, using land registers, of the evolution of land ownership during the 15th and 16th centuries in the region of Novgorod. During the spring of 1922, at the age of 19, he constructed a Lebesgue integrable function whose Fourier series diverged almost everywhere. In spite of his continuing love for history, this Fourier series would become an irresistible force bringing him into mathematics.

1.1 Convergence and divergence of Fourier series

In truth, the battle between convergence and divergence of Fourier series is also a part of history. For Daniel Bernoulli, in the 18th century, it was clear that a sound was a superposition of harmonics, and thus that a periodic function could be represented as the sum of a trigonometric series:

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} \left(a_n \cos \frac{2\pi nx}{T} + b_n \sin \frac{2\pi nx}{T} \right). \qquad (1.1)$$

Fourier at the beginning of the nineteenth century [DR98], believed that, using his formulas

$$a_n = \frac{2}{T} \int_0^T f(x) \cos \frac{2\pi nx}{T} dx, \quad b_n = \frac{2}{T} \int_0^T f(x) \sin \frac{2\pi nx}{T} dx, \qquad (1.2)$$

he had totally justified Bernoulli's belief. He had verified the convergence of the series in certain cases – in particular for the characteristic function of an interval – and declared that this convergence would hold for an arbitrary function as long as the Fourier coefficients were calculated using his formulas. It was Dirichlet, in 1829, who produced the first general convergence theorem, concerning functions which were piecewise continuous and piecewise monotone[1]; thinking that he could come back and eliminate the assumption of monotonicity. It turns out that this assumption can be weakened, as was done by Jordan ([Jor81], 1881) when he introduced functions of bounded variation, but continuity alone does not guarantee convergence as shown by the amazing (at the time) example furnished by Paul du Bois-Reymond of a continuous function whose Fourier series diverges at a point ([DuB73], 1873). For the next quarter century Fourier series had a bad reputation and were thought to be useful only for producing monsters like the continuous functions without derivatives which made Hermite turn away in "dread and horror".

The scene changed in 1900 when the very young Fejér Lipót[2] [Fej00] showed that the arithmetic means of the partial sums converged to the (bounded) function's value at any point where the function was continuous. The extensions and applications of this result revived interest in Fourier series. Then Lebesgue gave a new framework with his integral. After the publication of his *Leçons sur les séries trigonométriques* in 1906, when one spoke of a "Fourier series" one meant a Fourier-Lebesgue series, that is, a Fourier series whose coefficients are obtained using the Lebesgue integral. This framework

---

[1] Let us mention an exception. Camille Deflers, in 1819, in a *Note sur quelques intégrales définies, et application à la transformation des fonctions en séries de quantités périodiques* (Bulletin de la Société Philomatique de Paris, november 1819, pp. 161–166), proved what we call today the "Riemann-Lebesgue Lemma" using integration by parts (his proof works for $\mathcal{C}^1$ functions) and introduced the "Dirichlet kernel" ten years before Dirichlet; then he deduced that a sufficiently regular function was the sum of its Fourier series (his argument works in particular for $\mathcal{C}^2$ functions, but Deflers was not precise about the conditions under which his results were true). By using a version of "Riemann-Lebesgue Lemma" for continuous functions (which follows from the $\mathcal{C}^1$ case by uniform approximation: a tool which is now elementary, but was discovered after Deflers), the argument of Deflers proves that a periodic *differentiable* function is the sum of its Fourier series: it is this result which is taught today and wrongly attributed to Dirichlet. In fact the Dirichlet theorem works for functions f which are *piecewise monotone* (and continuous, if one wants the sum to be equal to the function). There is certainly a tie between the two theorems: a periodic $\mathcal{C}^1$ function is continuous and of bounded variation, and so, on any interval, it can be expressed as the difference of two continuous increasing functions (Jordan, 1881), to which the Dirichlet theorem can be applied: the Deflers theorem, in the case where f' is continuous, can thus be deduced from that of Dirichlet; but Deflers' proof (essentially the one which is taught today) is much more simple. (Editor's note.)

[2] Or Léopold Fejér: it is, in fact, with this gallicized form of his name that he signed his note in *Comptes rendus*. (Editor's note.)

would soon obtain a name, it would be the space L^1. The Riesz-Fischer theorem in 1907 established the isomorphism of the spaces L^2 and l^2 by using Fourier's formulas. Once again it seemed worthwhile to study Fourier series.

Divergence phenomena interested and inspired Fejér and Lebesgue in very different ways. They both worked with continuous functions and with Fourier series which diverged at a point. Kolmogoroff (as he signed his articles at the time) published an example in 1923 which made quite a sensation. It was complemented by a note in *Comptes rendus* in 1926, with an even more elaborate example: that of a (Lebesgue-)integrable function whose Fourier series diverged everywhere.

The example of Kolmogorov is even more interesting today. Indeed, at the time one could wonder if a stronger result could not be obtained, that is, if there existed a continuous function whose Fourier series diverged everywhere. But we know since Carleson ([Car66], 1966) and Hunt ([Hun67], 1967), that this is not possible: all functions belonging to a space L^p with $p > 1$ are limits almost everywhere of the partial sums of their Fourier series. The best result in this direction is that of Antonov ([Ant96], 1996): if $|f| \log^+ |f| \log^+ \log^+ \log^+ |f|$ is integrable on the circle (where $\log^+$ is equal to the positive part of log: $\log^+ x = \max\{\log x, 0\}$), the Fourier series of f converges to f almost everywhere.

If one assumes only that f is integrable on the circle the best that one can hope to obtain for the partial sums is thus an estimation of type $S_n(f) = o(l(n))$ almost everywhere. Hardy, in 1913, established such a formula with $l(n) = \log n$, and conjectured that it was the best possible result of this type. His conjecture is still open.

What is left to do today is thus, either to improve Hardy's formula (and to contradict Hardy's conjecture), or to improve the construction of Kolmogorov by giving a result with Ω (the opposite of o) for the S_n almost everywhere, if not everywhere. We have known for a long time that a result with Ω almost everywhere holds for any sequence $l(n)$ growing more slowly than $\log \log n$, and the result has also been established everywhere (Chen 1962). A big step was completed in 1999 by Konyagin when he constructed an integrable function whose partial Fourier sums are everywhere $\Omega(l(n))$ where the squares of the $l(n)$ are $o(\log n / \log \log n)$. One can try to improve Konyagin's estimation, as did Bochkarev in the case where the circle was replaced by the Cantor group (2003). But it would be difficult to attain $l(n) = o((\log n)^p)$ with $p > 1/2$.

1.2 Harmonic conjugates and Fourier series

In February of 1923 (when he had not yet turned twenty), Andrei Nikolaevich made his second major contribution to the theory of Fourier series, the paper *"Sur les fonctions harmoniques conjuguées et les séries de Fourier"* published in 1925 in the review *Fundamenta Mathematica*. The operation which takes a harmonic function to its conjugate can be interpreted in many ways: changing the real part of an analytic function into its imaginary part, or applying the Hilbert transform; or, in terms of Fourier series, changing the sign of the coefficients which correspond to negative frequencies and setting the constant term to zero. Today Kolmogorov's theorem can be stated in the following fashion: the operation of harmonic conjugation takes the space L^1 into weak L^1. (The functions discussed here are always real-valued functions on the circle.)

What is weak L^1 ? It is the set of functions f verifying the following condition: the inverse image of the ray $]y, +\infty[$ has a measure $m(y)$ bounded by C/y where C is a constant which depends only on the function f. In the case of L^1, $m(y)$ is integrable, and, since it is also decreasing, it verifies the given condition.

What was known about harmonic conjugation in 1923? First of all, by the Riesz-Fisher theorem (1907) it was known that it maps L^2 into L^2. Then, as noted in Lusin's work[3] which appeared in 1915, that it does not map L^1 into L^1 (this is rather clear to us today when we think of the conjugate of the Dirac measure). It was later, in 1927, that Marcel Riesz, the young brother of Frederic Riesz, showed that, for all $p > 1$, conjugation maps L^p into L^p.

In 1925, when it was published, the theorem of Kolmogorov seemed to be only an interesting curiosity. The weak L^1 space was not yet defined (in fact, it was first defined because of this theorem). Kolmogorov gives an application to the partial sums of Fourier series as a consequence of his result: they converge to the function in L^p for $p < 1$, and so they converge in measure. It was, according to Zygmund, the first appearance of the idea that the partial sums could be expressed in terms of harmonic conjugates ([Zyg59], Vol. 1, p. 381, §6). This curiosity, and especially its consequences for the partial sums, immediately attracted attention, as witnessed by Littlewood's article in 1926 *"On a theorem of Kolmogoroff "* [Lit26], which brings complex variable methods to bear on the problem.

Today this theorem is an important one because the Hilbert transform is the prototype of a whole class of singular integrals which have the same property. And this property of mapping L^1 into weak L^1 together with that of mapping L^2 into L^2, leads to the proof of Marcel Riesz's theorem by using a powerful tool, the interpolation theorem of Marcinkiewicz ([Mar39], 1939). In the second edition of Zygmund's book *Trigonometric series* [Zyg59] (1959) Kolmogorov's theorem is treated as a major result and two proofs of it are given, one using real analysis and the other using complex analytic functions.

1.3 Fourier series, integration and probability

Kolmogorov's first publications, which I will discuss a bit later, are about Fourier series. From 1925 on, he worked on a host of other subjects, in particular integration and probability. These two subjects are related to Fourier series.

[3] "The integral and the trigonometric functions" (in Russian), Moscow, 1915

1.3.1 The integral of the harmonic conjugate of an integrable function

Historically, every definition of the integral generates a new class of Fourier series; those whose Fourier coefficients are obtained with the Fourier formula when the integral is evaluated using the given definition. In fact, Fourier series have often served as a testing ground for different definitions of the integral. We thus can define Fourier-Riemann series, Fourier-Lebesgue series, Fourier-Stieltjes series, Fourier-Denjoy series, Fourier-Schwartz series, and on and on. Kolmogorov himself was essentially interested in Fourier-Lebesgue series. At the same time, he asked the question: while the harmonic conjugate of a Lebesgue-integrable function is not necessarily Lebesgue-integrable, could it be integrable in a more general sense? This question is answered by an article which appeared in 1928. This article is in French and entitled *"Sur un procédé d'intégration de M. Denjoy"*. This title is translated in the *Selected Works* by *"On the Denjoy integration process"*, which is a mistake (it should be *"On a Denjoy integration process"*). The English title makes one think that the article concerns Denjoy's totalization, either the first [Den12], which showed how to integrate any differentiable function, or the second [Den33, Den41], which gave a method for calculating the coefficients of an everywhere convergent trigonometric series given its sum. But Kolmogorov's article had nothing to do with all this. In 1919, Denjoy had published in *Comptes rendus* a note, *"Sur l'intégration riemannienne"* [Den19], in which he gave three generalizations of the Riemann integral, all three of them containing the Lebesgue integral (which he called in this note *"besguienne"*; this would not have pleased Lebesgue, who, they say, had threated Denjoy that he would call the totalization *"l'intégrale joyeuse"*, that is "joyous integral"; in any case the name *"besguien"* did not survive). These three generalization were called A), B) and C). Kolmogorov is quite clear: he is talking about the integral B). It has no relationship, as far as I know, with either the first or the second totalization. As far as I know, this is the first and the only interesting usage of the integral B). There is an explanation of all this and Kolmogorov's result on pp. 262–263 of the first volume of [Zyg59].

1.3.2 Series of independent random variables and lacunary Fourier series

The knowledge and the ideas of Kolmogorov concerning measure and integration were formed at this time, and later had a major impact on his interpretation of probabilities. But we must say that his first work in probability theory came essentially from Rademacher series, considered as series of independent random variables. In 1922, Rademacher [Rad22] published his theorem about the Rademacher orthogonal systems: L^2 convergence implies convergence almost everywhere. The article [KK25] of Khinchin (or Khintchine) and Kolmogorov in 1925 refines and extends this result: L^2 convergence and almost everywhere convergence are equivalent for a series of independent random variables, as are L^2 divergence and almost everywhere divergence. This article is mentioned in the *List of Works* of the *Selected Works* where it says that it appears in Vol. 1 along with all of the articles about integration and orthogonal series; in fact, it appears in Vol. 2, along with the other articles which are dedicated to probability. This article motivated many later studies of lacunary Fourier series, while, on the other hand, the study of lacunary Fourier series and of random Fourier series furnished new tools to be used in probability. The classical reference for this subject, as for the others, is Zygmund's book.

It is wonderful that Kolmogorov himself had established the theorem about almost everywhere convergence of lacunary Fourier series in 1922, before any mention had been made of independent random variables and at the same time as Rademacher's article appeared. This work appeared in 1924, before the article with Khinchin. He deals with more than the L^2 case; he shows in a few lines that almost everywhere convergence holds for any Fourier series which is lacunary in the sense of Hadamard, that is, in the sense that the ratio of a frequency to the preceding frequency is always greater than some number $q > 1$.

1.4 The descendants of the articles of young Kolmogorov

In 1923 and 1924 young Andrei Nikolaevich had only published articles about Fourier series. We have already discussed two aspects of these articles: divergence and lacunary Fourier series. It seems appropriate to complete this visit before finishing our promenade.

All of these articles are written in French and announce both the date when they were written and the date when they were published. The first, written on June 2nd, 1922, and published in 1923 in *Fundamenta Mathematica*, is the one concerning almost everywhere divergence which I said above made a great sensation at the time. Although the article was short, five pages, the technique seemed so elaborate that Zygmund decided not to put it into his treatise. *A fortiori* the proof of the existence of an everywhere divergent Fourier-Lebesgue series, of which a short indication had been given in a note to *Comptes rendus* in 1926, seemed to him to be unpublishable in his book. Today these proofs have been simplified, and, in particular, the reduction of the second result to the first has become part of a rather general framework. In Katznelson's book, a new edition of which has appeared recently, he gives the scope of this result with all the details and the description of the divergence sets for other classes of functions, such as the continuous functions, in a few pages at the end of his chapter about the convergence of Fourier series.

The second article, in the order in which they were written, is that of October 7, 1922, published in *Fundamenta Mathematica* in 1924. This is a very short article, two pages, which contains two theorems. The second is the one about lacunary Fourier series which I have already discussed. It has generated a remarkable number of articles, because the subject of lacunary Fourier series is related to all of functional analysis and, in particular, with the geometry of Banach spaces simultaneously with probability. One gets an idea of the work in this area (in particular that of Banach, Sidon, Zygmund, Kaczmarz and Steinhaus, Rudin, Marcus and Pisier, Bourgain) by looking at the chapter *"Lacunes et randon"* of my book with Pierre-Gilles Lemarié-Rieusset.

Today, that is, after the work of Carleson in 1966, the first theorem is totally outdated. But it seems in any case worth discussing as the beginning of a beautiful story. Kolmogorov shows that, if f is an L^2 function and if $m(n)$ is a lacunary sequence in the Hadamard sense, then the partial sums of order $m(n)$ converge almost everywhere to f. This date by which this result should be applied has passed, but we describe the simple and interesting method of its proof anyway; first to decompose the series into blocs $B(n)$ corresponding to frequencies situated between $m(n)$ and $m(n+1)$, then to regroup the $B(n)$ into two series according to whether n is even or odd. For each of these series the convergence of the Fejér sums implies the convergence of the partial sums with large gaps at the end, thus, the convergence of either the partial sums of order $m(2n)$ or those of order $m(2n+1)$ (depending on the series). Thus, because of the gaps, this gives for each of the two series convergence almost everywhere of the partial sums of order $m(n)$. Until 1966 this was the best known result on convergence for the L^2 class, and many specialists thought that it could not be improved (as far as I remember Zygmund was one of them). It could then have been tempting to extend it to L^p, $p > 1$, by using an analogous decomposition of the Fourier series. This was the monumental work of Littlewood and Paley in the 1930's; this decomposition is now called the Littlewood-Paley decomposition, and we just saw that it is exactly that of Kolmogorov. But it was necessary to develop an arsenal of weapons to do it, which is justifiably called the Littlewood-Paley theory. This theory was not yet finished when the first edition of Zygmund's book appeared in 1935, but there is an excellent explanation of it in the second edition which appeared in 1959.

The justification at the time was the extension to L^p of Kolmogorov's L^2 theorem. In 1966 and 1967, the theorems of Carleson and Hunt changed things completely: a much better result was obtained without using the decomposition and the theory of Littlewood-Paley. If Zygmund had wished to keep everything up to date, he should have included these new results in a new edition and renounced the old methods. Instead, he renounced the preparation of a 3rd edition and settled for several *"reprints"* of the second edition, with only slight additions. A "third edition" was produced recently by *Cambridge University Press*. Luckily it is nothing more than the second edition with the addition of an introduction by Robert Fefferman. The *Bulletin of the American Mathematical Society* asked me to review this book for them, a review I was quite pleased to write, as the book is a masterpiece, and because the Littlewood-Paley theory, although it no longer serves its initial purpose, has become an essential tool in other areas of analysis particularly in the study of partial differential equations. The exposition given by Zygmund is one of the best that exist. It is sometimes good, in mathematics, to think twice before throwing old things away.

The third article, concerning the order of magnitude of Fourier coefficients, was written on December 3, 1922, and published in 1923 in the *Bulletin of the Polish Academy of Sciences.* It establishes that, if the coefficients of a cosine series form a convex sequence converging to zero, then the series is a Fourier-Lebesgue series and, therefore, there exists Lebesgue-integrable functions whose Fourier coefficients converge to zero as slowly as anyone could want them to. This fact was discovered over and over, independently, by different mathematicians. The priority for the discovery should surely be given to W. H. Young, in his article *"On the Fourier series of bounded functions"* from 1912. For Fourier integrals, the corresponding result, that a continuous even function which converges to zero at infinity and is convex to the right of zero is equal to the Fourier transform of a positive integrable function, is attributed to Pólya 1949 (the functions in question are called *"Pólya functions"*, and it would be easy to justify calling the corresponding sequences *"Young sequences"*). But Kolmogorov adds a necessary and sufficient condition for the cosine series that he treats to converge towards the function in the L^1 metric, that is, the condition that the coefficients be $o(1/\log n)$: this is totally new.

The fourth article, written in common with Seliverstov (or Seliverstoff), is a note to *Comptes rendus* from January 14th, 1924; it was a time when the *Comptes rendus* attracted some very good works because they were published so quickly. It concerns almost everywhere convergence and an improvement of the conditions given by Hardy for trigonometric series and by Mensov (or Menchoff) for orthogonal series. It is the first of a series of articles of Kolmogorov and Seliverstov (1926), Plessner (1926), Hardy and Littlewood (1944), whose content is found in Zygmund's book from 1959 and which are only interesting from a historical point of view because of Carleson's theorem. The methods are in any case very clever and might be useful some day for other purposes.

I now stop the promenade in 1924, when Andrei Nikolaevich was not yet twenty one years old. He had been a precocious genius, in tune with the latest developments of a theory which had been copiously "plowed", and he left a very deep mark on it. It is true that in the beginning of the twentieth century some very young men made huge contributions to the rise of mathematics. The beauty of mathematics can attract young people, even in difficult and troubled times, and it seems to me that other Kolmogorovs are in the process of appearing, men or women, individuals or groups of individuals, on one continent or on another. Perhaps in Europe, perhaps in France, and there is no better way to leave Andrei Nikolaevich than to wish them welcome.

Appendix: Two other aspects of Kolmogorov's results concerning harmonic conjugates

by Nikolaï Nikolski

I The A-integral of Kolmogorov

The A-integral was introduced by Kolmogorov in his classic book [Kol33] (p. 65 of the American edition) under the name of generalized expected value of a random variable (the name "A-integral" was proposed later by Kolmogorov himself). A function f is A-integrable if

1. f belongs to weak L^1 (that is, if $\lim_{t\to\infty} t|\{|f|>t\}| = 0$), and:
2. the limit $\lim_{t\longrightarrow\infty} \int_{\{|f|<t\}} f(x)dx$ exists (we call this limit $(A)\int fdx$).

Here dx is an arbitrary finite measure. One can view this notion as a lebesguian analogue of the "principal value" of Cauchy $\lim_{t\longrightarrow\infty} \int_{-t}^{t} f(x)dx$ on the real axis. In the case of the interval $[0,2\pi]$ the definition had been proposed before Kolmogorov, by E. Titchmarch [Tit29], at first without the condition (1). But, as was proved in [Tit29], in this case the "integral" is not additive. On the other hand, the combination of conditions (1)-(2) has become useful in probability, in harmonic analysis and in complex analysis. Dozens of articles have been published on the subject. In particular:

1. E. Titchmarch [Tit29] showed (1929) that if $f \in L^1([0,2\pi])$, then its harmonic conjugate $\tilde{f}$ is A-integrable and the conjugate Fourier series is the Fourier series of $\tilde{f}$ in the sense of the A-integral.
2. P. Ulyanov [Uly57] showed that a Cauchy transform $f(z) = \frac{1}{2\pi i}\int_\gamma \frac{h(t)dt}{t-z}$ of a function $h \in L^1(\gamma)$ on a smooth contour γ is a Cauchy A-integral of its boundary values (this result is not true if, e.g. γ has angular points).
3. A. B. Aleksandrov [Ale81] showed that Ulyanov's result extends to functions f of the Smirnov class Nev^+ whose boundary values satisfy condition (1) of the definition. He also proved analogues of this result for harmonic functions on $\mathbb{R}^n$, and he used the A-integral to implement the duality $\mathrm{CI}(\mathbb{D})^* = H^\infty(\mathbb{D})$ between the space of Cauchy integrals $\mathrm{CI}(\mathbb{D}) = L^1/H^1_-$ and the algebra $H^\infty(\mathbb{D})$.
4. Other important publications on this subject are [ABGHV95], [Luk82], [Sal88], [Mad84]; in particular, [ABGHV95] contains an application of the A-integral to the description of those subsets of the unit circle for which one can arbitrarily choose the Cauchy data for the Laplace equation.

II On the weak (1,1) type of the harmonic conjugate

S. A. Vinogradov [Vin81] strengthened Kolmogorov's theorem in the following way: the map $\mu \mapsto \tilde{\mu}$ is continuous as a map $U^* \to$ weak L^1 where U^* is the dual of the space of uniformly convergent Fourier series (replacing $C(\mathbb{T})^* \to$ weak L^1 in Kolmogorov). This theorem has many applications, see [Kis87] for the general panorama.

References

[ABGHV95] Aleksandrov, A.B., Bourgain, J., Giesecke, M., Havin V., Vymenets, Yu.: Uniqueness and free interpolation for logarithmic potentials, and the Cauchy problem for the Laplace equation in $\mathbb{R}^2$. GAFA, **5**:3, 530–571 (1995)

[Ale81] Aleksandrov, A.B.: A-integrability of boundary values of harmonic functions (in Russian). Mat. Zametki, **30**, 59–72 (1981)

[Ant96] Antonov, N.Yu.: Convergence of Fourier series. East J. Approximation, **2**, 187–196 (1996)

[Boc03] Bochkarev, S.V.: Everywhere divergent Fourier-Walsh series. Dokl. Russian Akad Nauk, **390**:1 (2003), 11–14. Doklady Math., **67**:3, 307–310 (2003)

[Car66] Carleson, L.: Convergence and growth of partial sums of Fourier series. Acta Mathematica, **116**, 135–157 (1966)

[Che62] Chen, Y.M.: A remarkable divergent Fourier series. Proc. Japan Acad., **38**, 239–244 (1962)

[Den12] Denjoy, A.: Une extension de l'intégrale de M. Lebesgue. Comptes Rendus Acad. Sciences Paris, **154**, 859–862 (1912)

[Den19] Denjoy, A.: Sur l'intégration riemannienne. Comptes Rendus Acad. Sciences Paris, **169**, 219–221 (1919).

[Den33] Denjoy, A.: Sur le calcul des coefficients des séries trigonométriques. Comptes Rendus Acad. Sciences Paris, **196**, 237 (1933)

[Den41] Denjoy, A.: Leçons sur le calcul des coefficients d'une série trigonométrique. Gauthier-Villars, Paris (1941)

[DR98] Dhombres, J., Robert, J.-B.: Fourier, créateur de la physique mathématique. Coll. "Un savant, une époque", Belin, Paris (1998)

[DuB73] Du Bois-Reymond, P.: Über die Fourierschen Reihen. Nachr. Akad. Wiss. Göttingen, 571–584 (1873)

[Fej00] Fejér, L.: Sur les fonctions bornées et intégrables. Comptes Rendus Acad. Sciences Paris **131**, 984–987 (1900)

[Fis07] Fischer, E.: Sur la convergence en moyenne. Comptes Rendus Acad. Sciences Paris, **144**, 1022–1024 (1907)

[Har69] Hardy, G.H.: Collected Papers, Vol. 3. Oxford (1969)

[Hun67] Hunt, R.A.: On the convergence of Fourier series. Proc. Conference Edwardsville, IIlinois, 235–255 (1967)

[Jor81] Jordan, C.: Sur la série de Fourier. Comptes Rendus Acad. Sciences Paris, **92**, 228–230 (1881)

[Kat04] Katznelson, Y.: An introduction to harmonic analysis. Wiley (1968) Cambridge University Press (2004)

[Kis87] Kislyakov, S.V.: Classical themes of Fourier analysis. In Khavin, V., Nikolski, N. (ed) "Itogi Nauki i Techniki, Contemp. Problems Math." (Russian Math. Encyclopaedia), Vol. 15, 135–196 (1987), VINITI, Moscow (in Russian). Engl. transl.: *Encyclopaedia of Math. Sci.*, Vol. 15, 113–166, Springer, Berlin Heidelberg New York (1991)

[KK25] Khinchin, A.Y., Kolmogorov, A.N.: Über Konvergenz von Reihen, deren Glieder durch den Zufall bestimmt werden. Mat. Sb., **32**, 668–677 (1925)

[KL95] Kahane, J.-P., Lemarié-Rieusset, P.G.: Fourier series and wavelets. Gordon and Breach, London (1995) French transl.: Séries de Fourier et ondelettes. Cassini, Paris (1998)

[Kol23] Kolmogorov, A.N.: Une série de Fourier-Lebesgue divergente presque partout. Fund. Math., **4**, 324–328 (1923)

[Kol25] Kolmogorov, A.N.: Sur les fonctions harmoniques conjuguées et les séries de Fourier. Fund. Math., **7**, 23–28 (1925)

[Kol26] Kolmogorov, A.N.: Une série de Fourier-Lebesgue divergente partout. Comptes Rendus Acad. Sciences Paris, **183**, 1327–1329 (1926)

[Kol33] Kolmogorov, A.N.: Grundbegriffe der Wahrscheinlichkeitrechnung. Ergebnisse der Mathematik (1933) Russian transl.: ONTI, Moscow (1936) Engl. transl.: Chelsea, NY (1950)

[Kol91] Kolmogorov, A.N., Selected Works, Volume 1. Kluwer (1991)

[Kon00] Konyagin, S.V.: On divergence of trigonometric Fourier series everywhere. Comptes Rendus Acad. Sciences Paris, **239**, série 1, 693–697 (1999). Sb. Math., **191**, 97–120 (2000)

[Lit26] Littlewood, J.E.: On a theorem of Kolmogoroff. J. London Math. Soc., **1**, 229–231 (1926)

[Luk82] Lukashenko, T.P.: On the A-integral representation of the Hilbert transform and conjugate function. Anal. Math., **8**:4, 263–275 (1982)

[Mad84] Madan, S.: On the A-integrability of singular integral transforms. Ann. Inst. Fourier, **34**:2, 53–62 (1984)

[Mar39] Marcinkiewicz, J.: Sur l'interpolation des opérations. Comptes Rendus Acad. Sciences Paris, **208**, 1272–1273 (1939)

[Rad22] Rademacher, H.: Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen. Math. Ann., **87**, 112–138 (1922)

[Rie60] Riesz, F.: Œuvres, Vol. 1. Gauthier-Villars, Paris, Akadémiai kiadó Budapest (1960)

[Rie07a] Riesz, F.: Sur les systèmes orthogonaux de fonctions. Comptes Rendus Acad. Sciences Paris, **144**, 615–619 (1907)

[Rie07b] Riesz, F.: Sur une espèce de géométrie analytique des systèmes de fonctions sommables. Comptes Rendus Acad. Sciences Paris, **144**, 1409–1411 (1907)

[Rie24] Riesz, M.: Sur les fonctions conjuguées et les séries de Fourier. Comptes Rendus Acad. Sciences Paris **178**, 1464–1467 (1924)

[Rie28] Riesz, M.: Sur les fonctions conjuguées. Math. Zeit., **27**, 218–244 (1928)

[Sal88] Salimov, T.S.: The A-integral and boundary values of analytic functions (in Russian), Mat. Sb. (N.S.), **136 (178)**:1, 24–40 (1988). Engl. transl.: USSR-Sbornik, **64**:1, 23–39 (1989)

[Tit29] Titchmarch, E.: On conjugate functions. Proc. London Math. Soc., **29**, 49–80 (1929)

[Uly57] Ulyanov, P.L.: On A-integrals for contours. Doklady Acad. Sci. USSR, **112**:3, 383–385 (1957)

[Vin81] Vinogradov, S.A.: A refinement of Kolmogorov's theorem on the conjugate function and interpolation properties of uniformly convergent power series. Trudy Mat. Inst. Steklova, **155**, 7–40 (1981). Engl. transl.: Proc. Steklov Inst. Math., **155**, 3–37 (1983)

[You00] Young, W.H.: Selected Papers. Lausanne (2000)

[Zyg59] Zygmund, A.: Trigonometric series, Volumes 1 et 2. Cambridge Univ. Press (1959)

2

Kolmogorov's contribution to intuitionistic logic

Thierry Coquand

Department of Computer Science and Engineering, Chalmers University of Technology and Göteborg University, Sweden
http://www.cs.chalmers.se/~coquand
coquand@cs.chalmers.se

Translated from the French by Emmanuel Kowalski

Kolmogorov published only two papers related to mathematical logic. They are concerned with aspects of intuitionism and contain simple and fundamental contributions. Although they were written at the beginning of his mathematical career, it seems that Kolmogorov's interest for mathematical logic was long-lasting, as shown for instance by his work on the closely related topic of complexity[1]. We will start by explaining the content of these works, presenting their historical context, and discussing their current relevance. We conclude with a presentation of a recent development which involves the results from both papers. To write this survey, I have been helped by the references [Hei67] and [KolI], to which I direct the reader for further information[2]. I wish to thank Hugo Herbelin and Per Martin-Löf for their many comments on a preliminary version of this text.

2.1 The first paper (1925). Formalization of intuitionistic logic

2.1.1 Historical context

In this section, we will present Kolmogorov's first paper [Kol25] concerning mathematical logic, which was published in Russian in 1925 and only translated in English in 1967 [Hei67]. The context in which it appears needs to be explained in order to fully understand the subject of this paper. At that

[1] See the Chaps. 14 (by B. Durand and A. Zvonkin) and 15 (by P. Vitanyi) in this volume. (Editor's note.)

[2] The references [Hes03] and [Man98], that I discovered after writing this survey, are also directly relevant

time, the debate between Brouwer and Hilbert concerning the foundations of mathematics was raging. Weyl, who was Hilbert's most brilliant disciple, had just written a paper [Wey21] where, in dramatic language[3], he allied himself to Brouwer's cause, and strongly criticized Hilbert's Axiomatic approach. The paper of Hilbert [Hil23] quoted by Kolmogorov, although less virulent than the previous one [Hil22], was partly an answer to this article of Weyl and to Brouwer's criticisms. Many eminent mathematicians, such as Borel and Lebesgue, insisted on the doubtful status of transfinite reasonings [Bor50]. It may be guessed that such questions were subjects of animated discussions among Lusin's students[4], among whom were Khintchine, P. S. Novikov and Kolmogorov. Echoes of such discussions can be found in the examples quoted by Kolmogorov at the end of his paper[5].

Hilbert's article [Hil23] raises the problem of justifying the rules of quantification (both existential and universal) over an infinite domain, in particular the following principle

$$(\neg\forall x.A) \rightarrow \exists x.\neg A,$$

which follows from the Principle of Excluded Middle, and may be used to deduce the existence of an element $\exists x.\neg A$ from a proof of the impossibility of its non-existence $\neg\forall x.A$[6]. This is a typical instance of what Hilbert calls a *transfinite argument*, a terminology which is also used in Kolmogorov's paper (these terms may be somewhat surprising, since the adjective "transfinite" is associated nowadays with the use of the class of countable ordinals, or more generally of uncountable classes). Hilbert remarks that this rule is justified in the case of quantification applied to a finite domain. But, accepting the critique of Brouwer and Weyl, he admits that the intuitive meaning of $\forall x.A$ and $\exists x.A$ is far from clear in general. As, in analysis, it is not legitimate to "extrapolate to infinite sums and products the theorems which are valid for finite sums and products", it is not possible here to treat quantifiers applied to an infinite domain with the same semantics used in the finite case. Hence, the general idea will be to treat *formally* such quantifications as $\forall x.A$ and $\exists x.A$, stating precise rules of inference that apply to them, and show that those

[3] As shown by this sample quote: *"In this coming dissolution of the empire of analysis, even if few are forewarned, I was looking for a firm ground...* Because this state is not tenable*, as I convinced myself, and Brouwer, here is the revolution!"*. This polemic culminated in 1928 when Hilbert fired Brouwer, and the editors opposing this measure, from the editorial board of the prestigious journal *Mathematische Annalen* [vanD90]

[4] Lusin, following Borel and Lebesgue, was also rather critical of purely formal uses of transfinite reasonings [Lus30]

[5] Indeed, Khintchine [Khi26] and Kolmogorov [Kol29] both published general articles presenting the debate on intuitionism

[6] In this text, we use the following notation: $\neg A$ for the negation of A, $A \rightarrow B$ for A implies B, and we denote $A_1 \rightarrow (A_2 \rightarrow B)$ by $A_1 \rightarrow A_2 \rightarrow B$. The notation $A(x)$ simply means that x may be a free variable in A and we denote by $A(t/x)$, or simply $A(t)$, the result of substituting t for x in A

rules can only lead to correct results. (This is quite similar to the treatment of possibly divergent series in analysis.) As Hilbert stated: "By staying within finitary territory, the goal is to succeed in handling freely the transfinite and in dominating it entirely!". The method that Hilbert suggests is extremely original. It consists first in defining the quantifiers from a symbol τ subject to a unique axiom, namely

$$A(\tau_x A) \to A(x),$$

which expresses a strong form of the Axiom of Choice; then, one must show that this symbol may be eliminated from any proof of a finitary result[7].

But Hilbert's paper is only a programme, and contains no definitive result[8]. In this context, the main result of Kolmogorov's first paper is very remarkable, since it purports to prove conclusively that transfinite methods can only yield correct finitary results. Moreover, this proof is different from the one suggested by Hilbert, and is very natural. It confronts directly the problem of interpretation of quantifiers, and avoids considering the τ symbol (in fact, we will see below that Kolmogorov's method does not apply directly to an interpretation of the strong form of the Axiom of Choice that is derived from this symbol).

2.1.2 Kolmogorov's formalization of intuitionistic logic

Kolmogorov's first contribution in this paper is a complete formalization of *minimal* propositional calculus (a strict subset of intuitionistic logic which is usually attributed to Johansson [Joh36]) and minimal predicate calculus. As indicated by Wang, Kolmogorov's formalization is no less remarkable than Heyting's [Hey30]. The very possibility of such a formalization is already quite surprising, if we reflect that the motivations behind intuitionism were opposed to the process of formalization[9]. Kolmogorov's work is final concerning propositional calculus, but less precise with respect to predicate calculus.

---

[7] Using Hilbert's example, if A is the predicate "to be corruptible", then $\tau_x A$ would be a man "of such absolute integrity that if he turned out to be corruptible, then all men would be corruptible" [Hil23]. It is then possible to define $\forall x.A$ as equivalent to $A(\tau_x A)$. Afterwards, the symbol τ will be replaced by a dual symbol ϵ, designating a choice function. Bourbaki uses the symbol τ in its formulation of set theory, but with the dual meaning of the choice symbol ϵ. Hilbert's method is explained very suggestively in the reference [Wey44]

[8] A general result of elimination of the symbol τ in analysis was published by Ackermann [Ack24]; von Neumann [vonN27] found a problem in this proof and showed that its range of applicability is much more restricted than stated, in fact similar to Kolmogorov's result. It seems that the problem of elimination of the τ symbol in analysis remains open [Kre65]

[9] According to Wang [Wan87], Brouwer considered this result to be more remarkable and surprising than Gödel's celebrated incompleteness theorem [Göd31]

The formalization is directly inspired from Hilbert [Hil23], who had suggested the following axioms for implication and negation:

1. $A \to B \to A$
2. $(A \to A \to B) \to A \to B$
3. $(A \to B \to C) \to B \to A \to C$
4. $(B \to C) \to (A \to B) \to A \to C$
5. $A \to \neg A \to B$
6. $(A \to B) \to (\neg A \to B) \to B$

The only inference rule is the *modus ponens*: one may deduce B from $A \to B$ and A. For example, taking $B = A$ in the first two axioms (i.e. *by instantiation of B by A*), we obtain the following two special cases (two *instances*):

$$A \to A \to A$$

$$(A \to A \to A) \to A \to A$$

and by means of *modus ponens*, it follows that we also have $A \to A$.

The notions of semantic were not yet widely developed and it is interesting to notice how Kolmogorov states the property of *completeness* of this axiomatic system: it is not possible to add a new axiomatic schema without contradiction. More precisely, a new axiom, such as

$$(\neg\neg(A \to B)) \to \neg\neg A \to \neg\neg B,$$

is either provable from the axioms above, or leads to a contradiction, i.e. one may then deduce an arbitrary proposition[10].

The first four axioms deal only with implication, and Kolmogorov raises the pertinent question of their completeness for implication. As explained by Wang (p. 416 of [Hei67]), this system is not complete, but it becomes so if *Peirce's law*

$$((A \to B) \to A) \to A$$

is added to it.

The last two axioms concerning negation are both rejected by Kolmogorov, and replaced by a unique axiom called *principle of contradiction*

$$(A \to B) \to (A \to \neg B) \to \neg A$$

The last of Hilbert's axioms must be rejected because it may be seen as a formulation of the principle of excluded middle $A \vee \neg A$, as explained by Kolmogorov in his Note 9[11]. As noticed by Kolmogorov in his Note 3, the principle

$$(\neg \forall x.A) \to \exists x.\neg A$$

[10] In this example, one may deduce this new axiom from the ones given previously

[11] The numbering of notes refers to that used in the English translation in [Hei67]