A. G. Vitushkin, On Hilbert's thirteenth problem and related questions, *Russian Mathematical Surveys*, 2004, Volume 59, Issue 1, 11–25

DOI: 10.1070/RM2004v059n01ABEH000698

Russian Math. Surveys **59**:1 11–25
Uspekhi Mat. Nauk **59**:1 11–24 DOI 10.1070/RM2004v059n01ABEH000698

On Hilbert's thirteenth problem and related questions

A. G. Vitushkin

Abstract. Hilbert's thirteenth problem involves the study of solutions of algebraic equations. The object is to obtain a complexity estimate for an algebraic function. As of now, the problem remains open. There are only a few partial algebraic results in this connection, but at the same time the problem has stimulated a series of studies in the theory of functions with their subsequent applications. The most brilliant result in this cycle is Kolmogorov's theorem on superpositions of continuous functions.

Contents

1. Superpositions of functions. The problems and the results will be formulated in terms of superpositions of functions, and therefore we recall the corresponding definition. A function $f(x_1,\dots,x_n)$ is called a *superposition* of the functions $f_{\alpha_i}(y_{\alpha_i,1},\dots,y_{\alpha_i,k})$ $(i=0,1,\dots,s,\ \alpha_i=(\beta_0,\dots,\beta_i)$ is a multi-index, $\beta_0=0$, $\beta_j=1,\dots,k$ for $j>0$, and every argument $y_{\alpha_s,j}$ is equal to one of the coordinates $x_1,\dots,x_n)$ if we obtain $f(x_1,\dots,x_n)$ when we replace the arguments y_{α_i} successively by the functions $f_{\alpha_i}(y_{\alpha_i,1},\dots,y_{\alpha_i,k})$.

Let us take an example of using superpositions in numerical mathematics. Suppose that an algorithm computing a function of n variables is given and it is necessary to produce another algorithm computing this function together with

AMS 2000 *Mathematics Subject Classification*. Primary 28B40; Secondary 68Q30, 26B45, 41A46, 68P30.

its gradient. This problem arises, for instance, when computing an extremum of a function. The question is to what extent the number of arithmetic operations increases. The usual answer is that the number of operations becomes at least n times greater. This is wrong.

K. V. Kim observed that this problem can be solved in such a way that the number of operations becomes only 3 or 4 times greater (asymptotically for large n). The idea is as follows. Every formula for executing a sequence of arithmetic operations is a superposition of functions of two variables. Therefore, the formula computing each partial derivative of the initial function is a superposition of the same representing functions of two variables and their partial derivatives. Approximately, the complexity of computing a function of two variables together with its gradient is only at most three times greater than the complexity of computing this function without the gradient. The computation of the values of all the representing functions with their gradients is the largest part of the work. Hence, the new total number of operations is obtained by multiplying the former number by a factor which is substantially less than n. Kim designed a computer program based on this idea (1984, [1]).

2. Hilbert's conjecture. The solution of an algebraic equation of degree $n \leq 4$ is given by a formula containing only radicals and arithmetic operations, and therefore it is a superposition of functions of two variables. For $n > 4$ an algebraic substitution, the so-called Tschirnhausen transformation given by a formula containing only radicals and arithmetic operations, reduces the general nth-degree algebraic equation $x^n + a_1 x^{n-1} + \cdots + a_n = 0$ to the form $y^n + b_{n-4} y^{n-4} + \cdots + b_1 y + 1 = 0$. Here a_i and b_i are complex numbers. It follows that the solution of an algebraic equation of degree n can be represented as a superposition of functions of two variables if $n < 7$ and as a superposition of functions of $n-4$ variables if $n \geq 7$.

In particular, the solution of a general equation of degree 7 is a superposition of arithmetic operations, radicals, and a single function of 3 variables which is a solution of the equation $f^7 + xf^3 + yf^2 + zf + 1 = 0$. A further simplification of this equation by means of algebraic transformations seems to be impossible.

In the famous list of Hilbert's problems (1900, [2]) this conjecture is formulated as problem number 13 as follows: *A solution of the general equation of degree* 7 *cannot be represented as a superposition of continuous functions of two variables.* Here and below, when speaking about a solution of an algebraic equation, we mean a single-valued branch of the general solution. However, judging from Hilbert's comments on this problem, one cannot be sure that this is precisely what he meant.

Hilbert's conjecture does not become less attractive if in its statement one means the complete (seven-valued) solution and multivalued continuous functions of two complex variables. This understanding of the conjecture stresses the topological aspect of the problem. The 'complexity' of the discriminant curve of the complete solution is possibly an obstacle to representing the solution as a superposition of functions of two variables.

On the whole, different authors have understood this problem in different ways. Hilbert related this problem in his comments [1] to needs in nomography (a nomogram is a collection of plane curves designed for the graphic computation of values of some elementary function of several variables). He later studied superpositions of algebraic functions [3]. Wiman [4] and Chebotarev [5] reckoned on a possible development of Galois theory. In his youth, Kolmogorov related the problem to the theory of analytic functions and advised P. S. Aleksandrov to study it by concentrating on its topological aspect (1932, [6], p. 444).

Kolmogorov proved that *every continuous function of several variables can be represented as a superposition of continuous functions of one variable and the operation of addition* (1957, [7]). Thus, it is as if there are no functions of several variables at all. There are only simple combinations of functions of one variable. Unfortunately, the representing functions are pretty bad, they are at best only continuous, and they cannot be chosen to be differentiable even if one represents an analytic function (1964, [8]).

Seriously speaking, *Kolmogorov's theorem is a brilliant example of his mastery.* In particular, the theorem shows that *Hilbert's conjecture is wrong.*

However, it should be noted that the algebraic core of the problem remains untouched. An affirmative answer to the problem is still possible, that is, it is still possible that *the solution of the equation of degree* 7 *cannot be represented as a superposition of functions of* 2 *variables* (provided, of course, that these functions are assumed to be smooth in some sense). It is not known whether or not there is an algebraic function of n variables ($n > 2$) which cannot be represented as a superposition of algebraic functions of fewer variables. On the other hand, it is not ruled out that every algebraic function is a superposition of algebraic functions of one variable and arithmetic operations. Thus, the *problem remains open* and, *by the highest standards, the range of issues is in fact as broad as it was at the beginning of the* 20*th century.*

3. Superpositions of algebraic functions. Hilbert himself studied the algebraic aspect of the problem. He proved (1927, [3]) that the solution of the general equation of degree 9 can be represented as a superposition of algebraic functions of 4 variables. (We recall that the Tschirnhausen transformation gives a superposition of functions of five variables in this case.)

There are a few results concerning equations of higher degrees ([4], [5], [9], [10]) similar to this theorem of Hilbert.

Let us present an example of another kind which is an attempt to find an approach to solving the problem in the affirmative. Khovanskii [Hovanski] showed (1971, [11], [12]) that the solution of the equation $f^5 + xf^2 + yf + 1 = 0$ cannot be represented by a superposition of entire algebraic functions of one variable and polynomials in arbitrarily many variables. We recall that the Tschirnhausen transformation reduces the general equation of degree 5 to an equation with a single parameter, that is, Khovanskii's function can be represented as a superposition of algebraic functions of one variable and arithmetic operations. This example shows that the prohibition against using the division operation is substantial.

4. Superpositions of analytic functions. No one before Kolmogorov dared to doubt the validity of Hilbert's conjecture. Therefore, no one studied superpositions of continuous functions. The only exception was an erroneous paper of Bieberbach [13]. Counting on a positive solution of the problem, it is reasonable to first restrict the class of representing functions by considering, for instance, superpositions of sufficiently smooth functions.

One can readily show that there is an analytic function of three variables which cannot be represented as a superposition of analytic functions of two variables. The idea of the proof is that the number of partial derivatives of order p for a function of three variables is proportional to p^2, whereas the number of partial derivatives of order p for a function of two variables is proportional to p, and hence there are 'more' functions of three variables than superpositions of functions of two variables.

There are some specific examples. Ostrowski showed (1920, [14]) that the analytic function $\xi(x, y) = \sum_{n=1}^{\infty} \frac{x^n}{n^y}$ of two variables cannot be represented as a superposition of infinitely differentiable functions of one variable and algebraic functions of arbitrarily many variables. The proof of this result is based on the fact that the function $\xi(x, y)$ satisfies no algebraic partial differential equation (that is, an equation of the form $\Phi\bigl(\xi, \frac{\partial\xi}{\partial x}, \frac{\partial\xi}{\partial y}, \dots, \frac{\partial^{\mu+\lambda}\xi(x,y)}{\partial x^{\mu}\partial y^{\lambda}}\bigr) = 0$, where Φ is a polynomial with constant coefficients in the function ξ and its partial derivatives up to some finite order). At the same time, it is rather easy to prove that any function of two variables which is a finite superposition of the above form satisfies an algebraic partial differential equation of this kind.

5. Superpositions of smooth functions. A large cycle of papers on the theory of functions with diverse subsequent applications appeared in the 1950s and 60s in connection with the Hilbert problem under consideration. The first result of this cycle was obtained in Kronrod's seminar (for this seminar, see [15]). It was proved that *there is a continuously differentiable function of three variables which cannot be represented as a superposition of continuously differentiable functions of two variables* (1954, [16]).

I remember the talk about this result particularly well. At that time (in 1953), the Department of Mechanics and Mathematics was located in the old university building at the corner of Gertsen [Herzen] and Mokhovaya streets. It was a warm spring day, and Kronrod suggested holding the seminar outside, in the university yard. In such cases the seminars were held in the front garden near the Lomonosov statue, and the participants were sitting on the parapet fence around the monument. We did not need a blackboard, there were only a few of us, and we readily understood one another. Our study was rather noisy as a rule, which sometimes attracted onlookers.

The unusual form of the statue also attracted attention. Lomonosov was represented in all his magnitude, with a scroll in his left hand resting against the upper part of his thigh. On seeing the statue in profile from the right, one could only guess what the protruding scroll represented. The first-year university girls, the more observant of them, blushed and looked away when passing by. In time the scroll was removed. Now the monument to the founder of Moscow University looks decent.

The result on superpositions gave me an occasion to make Kolmogorov's acquaintance and then to become one of his postgraduate students. I asked him to recommend my two notes [16] and [17] to *Doklady Akad. Nauk SSSR*. After listening to the formulation of the result, Kolmogorov uttered his usual drawn out "eh ... " and added: "Yes, it is correct, and I see how it can be done." I was a bit discouraged. Pavel Sergeevich Aleksandrov, who took part in the conversation, comforted me: "Don't get upset, Tolya. Andrei Nikolaevich understands everything."

A week after this conversation, when I met with Kolmogorov to learn his decision, I had a surprise of another kind. He had completely rewritten both communications and even typed them up, explaining in this unusual way how the material should be presented and how an article should be prepared for publication.

Later, when it came to writing surveys and it was desirable to improve the results if possible, the theorem about superpositions of smooth functions was reformulated in the following form: if the complexity of a function is measured by the ratio of the number of variables to the order of smoothness, then *almost every function of a given complexity* (*except for a first-category set of functions*) *cannot be represented as a superposition of functions of lower complexity* [18].

6. Superpositions of continuous functions. Kolmogorov used to say that Hilbert's thirteenth problem was good material for students. In time it became clear that this problem was good not only for students but also for Kolmogorov himself. *The definitive result was Kolmogorov's paper claiming that every continuous function of several variables can be represented as a superposition of continuous functions of three variables* (1956, [19]).

Kolmogorov said that the idea of this construction arose when, as was his custom, he was looking through out-dated journals and came across an article by Kronrod [20] which treated functional trees among other things. The tree of a function is the space of components of its levels. A tree is one-dimensional and acyclic, and hence can be embedded in the plane homeomorphically. The values of the function can be naturally transferred to its tree, and thus a function of several variables becomes (in a sense) a function of only two variables. Constructing superpositions required an additional variable, which resulted in superpositions of functions of three variables.

This result of Kolmogorov was a sensation: Hilbert's idea of characterizing the complexity of solutions of algebraic equations in terms of superpositions of *continuous* functions turned out to be inadequate. In fact, it became clear that Hilbert's conjecture about the algebraic equation of degree 7 was also incorrect. However, additional efforts were required to disprove this conjecture.

Kolmogorov was an excellent research supervisor. When working with students, he always created opportunities for a gifted student to play a solo part, so to say. This was the case again: Dima Arnol'd, a third-year student, improved Kolmogorov's construction and showed that any continuous function of three variables (in particular, the solution of the general equation of degree 7) can be represented as a superposition of continuous functions of two variables, thus proving that *Hilbert's conjecture was incorrect* (1957, [21], 1958, [22], 1959, [23]).

It was in the character of Kolmogorov to carry any work to completion. Shortly thereafter, following the rule of improving every result to its sharpest form, he found a new construction, avoiding functional trees, and proved that every function f defined and continuous on the n-dimensional cube can be represented by a superposition of the form

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{2n+1} f_i\left(\sum_{j=1}^{n} \phi_{i,j}(x_j)\right),$$

where all the functions $\{f_i\}$ and $\{\phi_{i,j}\}$ are defined and continuous on $\mathbb{R}^1$ and all the functions $\{\phi_{i,j}\}$ are independent of the choice of f (1957, [7]).

There have been many comments, refinements of this theorem, and illustrations of its proof ([8], [15], [18], [24]–[40]). There has been only one improvement of Kolmogorov's theorem. B. Fridman (1967, [34]) showed that the functions $\{\phi_{i,j}\}$ can be chosen to be Lipschitzian with exponent 1. However, it should be noted that, if one replaces the functions $\{\phi_{i,j}\}$ by some continuously differentiable functions in Kolmogorov's formula, then it turns out that there are even analytic functions which cannot be represented by this formula (1964, [8]).

Kolmogorov's theorem concludes the cycle of papers on superpositions of continuous functions. Of course, new problems related to this topic still arise in the theory of functions. For instance, the problem of *whether or not every analytic function of two variables can be represented as a superposition of continuously differentiable functions of one variable and the operation of addition.*

A paper of A. A. Milyutin (1951) could have become a precursor of the results discussed here (on superpositions of smooth and continuous functions). He proved that the linear space of continuous functions of n variables is isomorphic to the linear space of continuous functions of one variable for every positive integer n. He also proved that such an isomorphism is impossible if we assume in addition that the smooth functions in one space are taken into smooth functions in the other space (the second part of the theorem was obtained by comparing the cardinalities of ε-nets for the corresponding compact sets of smooth functions). Milyutin's paper could indeed have become such a precursor, but this possibility was not realized. A. Pełczyński claimed that his work was incorrect, and the paper remained unpublished and was forgotten. Fifteen years later, G. M. Khenkin [Henkin] obtained the same result. He learned about the work of Milyutin (possibly from Pełczyński), and curiosity compelled him to find Milyutin's manuscript in the library of the Department of Mechanics and Mathematics. It turned out that there were no errors at all, and Khenkin arranged for the publication of Milyutin's research (1966, [41]).

7. Variations of sets. Let us now discuss some ideas that arose in connection with the problem of superpositions of smooth functions. We recall [42]–[44] that to a subset of an n-dimensional space one can assign variations of all orders from 0 to n. The kth variation is the integral, over the space of $(n-k)$-dimensional planes, of the function defined as the number of connected components in the intersection of the given subset with each of the planes.

Consider a set in $\mathbb{R}^3$ looking like a rope, for instance, like a clothesline. Four variations are defined in this case. The zeroth variation of the rope is the number of its pieces. The first variation is equal to its length in the usual sense (the Hausdorff length of a rope is infinite). The second variation is equal to the surface area of the rope. The last (the third) variation is the volume of the given set.

Kolmogorov pointed out that the highest non-zero variation of a set is equal to its Favard measure of the same order. (The Favard measure is defined in the same way as the variation, with the single difference that the integrand is given by the number of points rather than by the number of components.) There is nothing to prove here; this is only an observation. However, it is still nice: one of variations of any set is numerically equal to a measure of the corresponding order.

The variations are convenient because they characterize the 'width' of the set in all dimensions simultaneously. This fact can be used to estimate the complexity of diverse mathematical objects. In particular, this property of variations was successfully applied to the study of largeness of diverse compact sets in spaces of smooth functions. We note that the variations of a set of functions are defined as variations of an appropriately chosen projection of this set of functions to some space $\mathbb{R}^n$, for instance, of a projection obtained by restricting the functions to a sufficiently dense net of points.

The variations of a set can be estimated by approximating it by an algebraic surface and using known bounds for the number of components of algebraic surfaces. There are investigations by Petrovskii and Oleinik containing sharp estimates for the Betti numbers (in particular, for the number of components) of a real algebraic hypersurface (1952, [45], [46]). These sharp estimates and the technique obtained in working with variations of sets were used in the proof of the above theorem on superpositions of smooth functions (see § 5).

8. Variations of functions. Variations of sets arose as a generalization of the notion of Kronrod variations of functions of two variables [20]. In terms of variations of sets one can obtain the following good general definition of variations for functions of arbitrarily many variables. For a function of n variables let us define the variations of all orders from 1 to n. The kth variation of a function f is the integral (over all real values t) of the $(k-1)$st variation of the level set $f = t$. Almost all of Kronrod's results concerning the variations of functions of two variables [20] can readily be generalized to functions of arbitrarily many variables. The only exception was the proof of the fact that the linear variation (variation of order one) of an n times continuously differentiable function of n variables is finite [42]. By induction, this theorem on the linear variation easily implies that if a function of n variables is $(n-k+1)$ times continuously differentiable, then it has finite variation of order k $(1 \leq k \leq n)$.

Unfortunately, there are no formulae for computing variations of functions, and thus the proof of this theorem turned out to be quite complicated. The complication is not related to the fact that the minimal smoothness guaranteeing finiteness of the variation is assumed (the bound indicated above is indeed sharp). The proof cannot be simplified even if one assumes that the function in question is analytic.

This theorem took two of my student years: a year for $n = 3$ and a year for $n = 4$, and only for $n = 5$ did the proof became suitable for arbitrary n. The proof remained technically complicated, but its idea was substantially simplified. The point is that the linear variation of a function over a small neighborhood of a point is comparable to the linear variation of the Taylor expansion of the function at this point. It is easy to estimate the variation of the Taylor polynomial, because the number of components of its level set is bounded by some constant depending only on the degree of the polynomial and the number of variables. Any estimate (however rough) for the number of components of the level sets of a polynomial was sufficient for the proof. This was not hard. However, there was no need to worry about that, because sharp estimates for the number of components of real algebraic hypersurfaces had been obtained in the meantime by Petrovskii and Oleinik in the papers mentioned above.

9. ε-Entropy estimates for compact sets of smooth functions. Let us fix two positive integers n and s and denote by F the compact set (with respect to the uniform norm) of functions defined on the n-dimensional cube and bounded by 1 in modulus together with all their partial derivatives up to order s inclusive. Let H_ε be the ε-entropy of the compact set F (the logarithm with base 2 of the number of elements of a minimal ε-net for F).

Kolmogorov proved (1955, [47]) that $A(1/\varepsilon)^{n/s} < H_\varepsilon(F) < B(1/\varepsilon)^{n/s}$, where A and B are positive constants independent of ε.

The theorem about superpositions of smooth functions (see § 5) can readily be proved by using this inequality. However, the theorem appeared earlier than the inequality, and moreover, it led to the discovery of the inequality.

Kolmogorov proved the inequality by using only the local Taylor formula, leaving aside the estimates for the variations of compacta and the corresponding bounds for the Betti numbers. The new proof of the theorem on superpositions of smooth functions obtained in this way became more understandable and, which is of special importance, the estimates of the ε-entropy turned out to be useful in many similar cases.

The computation of entropy became fashionable, and this stimulated diverse investigations in approximation theory, information theory, estimates for the complexity of algorithms, and so on (see [48]–[79]). As far as estimates of the variations of sets are concerned, they turned out to be very useful in estimating the complexity of algorithms.

10. An example of estimating the complexity of algorithms. The papers about superpositions led to a series of investigations on estimates of complexity of algorithms. For the main classes of smooth and analytic functions, some estimates of the complexity of algorithms computing functions of these classes were obtained by the method used in the case of superpositions, by computing the variations of the corresponding compact sets of functions (1957, [80]–[83]). Namely, under certain conditions on a compact set F of functions it was proved that if an algorithm gives an approximation of every function in F with accuracy ε, the number of parameters determining the function is p, and the parameters enter the formula for computing the function with total degree (with respect to the total set of parameters) not exceeding k, then $p\log_2(k+1) \geq cH_\varepsilon$. Here c is a constant depending on the compact set F, and H_ε stands for the ε-entropy of F.

The known approximation schemes for smooth and analytic functions are close to being optimal from the point of view of these estimates for the complexity of algorithms. This fact can be understood in the following way: *there are no significantly better approximation methods for these classes of functions.* The above inequalities also show that if one tries to reduce the number of parameters (for instance, by representing a function as a superposition of functions of fewer variables) when forming a table of some function, then the degree of dependence of the function on the parameters increases, and thus the number of necessary arithmetic operations increases drastically.

After reading the manuscript of this paper, Kolmogorov sent me a letter congratulating me on the interesting result and advising me how to correct its formulation. His secretary delivered the letter (at the time I lived in a student hostel in the main building of the university). This form of communication surprised me because Kolmogorov was at his office at that time (in the central part of the same building), and he could have invited me to come by. Nevertheless, he sent a letter, and I regarded this fact as his particular evaluation of the result. However, my talk at the Moscow Mathematical Society was not equally successful. Kolmogorov scheduled my lecture too soon; the material was in rough form, and I could not prepare it properly. The audience was noisy and often interrupted me. At the end I became one of the audience and Kolmogorov explained the essence of the result at the blackboard.

Kolmogorov always prepared his lectures carefully and expected the same from others. Having got a deserved scolding after my report, I tried to escape with a quip: "At least the talk was amusing." As a penalty for this "at least", Kolmogorov did not speak to me for half a year and forgave me only when I came up with a new result.

Kolmogorov's scientific talks always attracted a large audience. Attendance of his lectures by students was quite good, too. However, it was hard to understand his lectures because he obviously overestimated the abilities of the audience. By the way, he was a 'comfortable' examiner. He did not like to spend time waiting for an answer to his question and sometimes answered the question himself. If the student could successfully discuss the details of this answer, he got a high grade.

11. Encoding of signals with finite spectrum. It was planned to hold the First All-Union Congress on Communications Reconstruction and Low-Current Industry Development in 1933. Vladimir Aleksandrovich Kotel'nikov prepared a talk "On the carrying capacity of the ether and wire in telecommunications." The congress was not held, but the proceedings were published [84].

The main thesis of Kotel'nikov's report was that *the amount of information received through a communications channel is proportional to the frequency bandwidth of the channel.* A more rigorous formulation of this assertion is now called Kotel'nikov's theorem in textbooks, namely, *an entire function f of type σ which is square-integrable on the real axis can be represented in the form*

$$f(t) = \sum_{k=-\infty}^{\infty} f(t_k) q_k(t), \quad \text{where} \quad t_k = \frac{k\pi}{\sigma}, \quad q_k(t) = \frac{\sin(\sigma(t - t_k))}{\sigma(t - t_k)}\,.$$

Indeed, it follows from this formula that the 'amount of information', that is, the number of independent numbers carried by a signal of spectrum σ per unit time, is equal to $\frac{\sigma}{\pi}$.

Formulae of this kind were known before Kotel'nikov. His discovery was a reasonable choice of the function class and the understanding of potential applications.

As time went by, Kotel'nikov's thesis was refined several times. Shannon characterized the information content of a signal by calculating the asymptotics of the density of a signal code (the length of binary code per unit time) for a random process with finite spectrum (1948, [85]). Kolmogorov obtained a similar asymptotic formula for the Bernstein class of functions (the finite-type entire functions which are real and bounded on the real axis) (1957, [86]). There are also estimates of the code density for signals with finite spectrum in a special metric adapted to describe high-quality sound recording devices (1974, [87]).

Kolmogorov and Tikhomirov showed (1959, [58]) that the ε-entropy (in the uniform norm) of the compact set of functions obtained as the restrictions to the interval $[-T, T]$ of the entire functions of type σ which are real and bounded in modulus by 1 on the real axis is asymptotically equal (for small values of ε and large values of T) to $\frac{2T\sigma}{\pi}\log_2\frac{1}{\varepsilon}$. Hence, if every function from the above class, when transmitted through a communications channel, can be recovered with accuracy ε at the output of the channel, then the amount of information (that is, the number of binary digits) which can be transmitted through this channel per unit time is asymptotically equal to $\frac{\sigma}{\pi}\log_2\frac{1}{\varepsilon}$.

Shannon entirely stopped his research at an early stage and still kept his position of professor at the Massachusetts Institute of Technology. When he visited Moscow State University, he asked Petrovskii (who was the rector at that time) to arrange a meeting with Kolmogorov. They met in Petrovskii's office. When Kolmogorov arrived there, he was far from being in his best spirits. The meeting was very short. Apparently, this was due to the language barrier as well. Kolmogorov was fluent in French and German and read English well, but his spoken English was not very good. Shannon, with some sympathy, expressed his regret that they could not understand each other well. Kolmogorov replied that there were five international languages, he could speak three of them, and, if his interlocutor were also able to speak three languages, then they would have no problems.

12. Digital sound recording. The paper on encoding of signals with finite spectrum was completely re-appraised only at the end of the 1970s, when it became possible to replace analog recording systems by digital ones. In particular, in sound recording the expense of necessary digital devices became acceptable, and the quality of the playback signal on a CD turned out to be almost perfect. However, minimizing the complexity of algorithms is always a problem. For instance, the flow of sound and video recordings nowadays on the Internet is so broad that it becomes necessary to restrict ourselves to a low quality of playback in order to save memory and time. For this reason, new encoding schemes appear from time to time for reducing file sizes.

It is wonderful to listen to music not only in concert halls but also at home. However, bad acoustics in a room can change the frequency characteristics of the entire reproducing system, thus significantly constraining the quality of the playback. If a powerful computer is available, then Kotel'nikov's formula provides the possibility of smoothing out the characteristics. This can be achieved by replacing the function $f(t)$ received from the record player by another function $f^*(t) = \sum_{n=-N}^{N} \varepsilon_n f(t - t_n)$. One can readily see that the spectrum of the function $f^*(t)$ (the Fourier transform of this function) is the product of two factors, namely, the spectrum of the function $f(t)$ and a correcting factor which has the form of a segment of the Fourier series with the coefficients ε_n. Therefore, for a sufficiently large value of the number N one can take the numbers ε_n in such a way that the frequency characteristic of the whole system (including the acoustics of the room) becomes sufficiently close to a constant. It is important that no preliminary processing of the function $f(t)$ is needed. For a sufficiently fast processor the chain of numbers $f(t_k)$ can be transformed into the chain $f^*(t_k)$ in real time.

In sound recording the quality of the equipment is characterized by three parameters: the frequency response range, the non-linear signal distortion (measured as the percentage of the mean-square signal norm), and the device's dynamic range (defined as the ratio of the norms of the maximum and minimum signals with guaranteed small non-linear distortions). Sufficiently 'good' values of the last two parameters can be obtained at the expense of high accuracy of approximation of the function. However, if one uses Kotel'nikov's formula, taking the numbers $f(t_k)$ as the encoding sequence, then the length of the code increases. *It is more efficient to use the Weierstrass formula* reproducing a function from its zeros, and thus take the coordinates of the zeros as the corresponding code.

It turned out, and this was quite unexpected, that *there is an encoding system providing an arbitrarily wide dynamic range of a communications channel or device without increasing the length of the code.* This fact explains the mechanism of the action of the existing noise-suppression systems used in both analog and digital sound recording. This work was completed by V. I. Buslaev and me in 1974 [78].

For video signals we have not yet been able to get similar estimates for the length of the code in dependence on the dynamic range of a communications channel. Such a result would give an estimate of how close the encoding methods used in video devices are to being optimal.

Bibliography

[1] K. V. Kim, Yu. E. Nesterov, and B. V. Cherkasskii, "Estimation of the complexity of computing the gradient", *Dokl. Akad. Nauk SSSR* **275** (1984), 1306–1309; English transl., *Soviet Math. Dokl.* **29** (1984), 384–387.

[2] D. Hilbert, "Mathematische Probleme", *Nachr. Akad. Wiss. Göttingen* (1900), 253–297; Gesammelte Abhandlungen, Bd. 3, Springer, Berlin, 1935, pp. 290–329.

[3] D. Hilbert, "Über die Gleichung neunten Grades", *Math. Ann.* **97** (1927), 243–250; Gesammelte Abhandlungen, Bd. 2, Springer, Berlin, 1933, pp. 393–400.

[4] A. Wiman, "Über die Anwendung der Tschirnhausen-Transformation auf die Reduktion algebraischer Gleichungen", *Nova Acta Soc. Sci. Upsal.* (1928), 3–8.

[5] N. G. Chebotarev, "On certain questions of the problem of resolvents", Collected Works, vol. I, 1949, pp. 255–340. (Russian)

[6] A. N. Shiryaev (ed.), *A. N. Kolmogorov, Anniversary Book*, vol. 2 Selected correspondence of A. N. Kolmogorov with P. S. Aleksandrov, Fizmatlit, Moscow 2003. (Russian)

[7] A. N. Kolmogorov, "On the representation of continuous functions of several variables by superpositions of continuous functions of one variable and addition", *Dokl. Akad. Nauk SSSR* **114** (1957), 953–956; English transl., *Amer. Math. Soc. Transl.* (2) **28** (1963), 55–59.

[8] A. G. Vitushkin, "Proof of the existence of analytic functions of several variables not representable by linear superpositions of continuously differentiable functions of fewer variables", *Dokl. Akad. Nauk SSSR* **156** (1964), 1258–1261; English transl., *Soviet Math. Dokl.* **5** (1964), 793–796.

[9] N. G. Chebotarev, "On the problem of resolvents", *Uchen. Zap. Kazan. Gos. Univ.* **114** (1954), 189–193. (Russian)

[10] V. V. Morozov, "On certain questions connected with the problem of resolvents", *Uchen. Zap. Kazan. Gos. Univ.* **114** (1954), 173–187. (Russian)

[11] A. G. Khovanskii, "Superpositions of holomorphic functions with radicals", *Uspekhi Mat. Nauk* **26**:3 (1971), 213–214. (Russian)

[12] A. G. Khovanskii, "The representability of functions by quadratures", *Uspekhi Mat. Nauk* **26**:4 (1971), 251–252. (Russian)

[13] L. Bieberbach, "Bemerkung zum dreizehnten Hilbertschen Problem", *J. Reine Angew. Math.* **165** (1931), 89–92.

[14] A. Ostrowski, "Über Dirichletsche Reihen und algebraische Differentialgleichungen", *Math. Z.* **8** (1920), 241–298.

[15] A. G. Vitushkin, "Half a century as one day", *Uspekhi Mat. Nauk* **57**:1 (2002), 191–206; English transl., *Russian Math. Surveys* **57** (2002), 199–220.

[16] A. G. Vitushkin, "On Hilbert's thirteenth problem", *Dokl. Akad. Nauk SSSR* **96** (1954), 701–704. (Russian)

[17] A. G. Vitushkin, "On certain estimates of variations of sets", *Dokl. Akad. Nauk SSSR* **95** (1954), 433–434. (Russian)

[18] A. G. Vitushkin, "On representation of functions by means of superpositions and related topics", *Enseign. Math.* (2) **23** (1977), 255–320.

[19] A. N. Kolmogorov, "On the representation of continuous functions of several variables by superpositions of continuous functions of fewer variables", *Dokl. Akad. Nauk SSSR* **108** (1956), 179–182; English transl., *Amer. Math. Soc. Transl.* (2) **17** (1961), 369–373.

[20] A. S. Kronrod, "On functions of two variables", *Uspekhi Mat. Nauk* **5**:1 (1950), 24–134. (Russian)

[21] V. I. Arnol'd, "On functions of three variables", *Dokl. Akad. Nauk SSSR* **114** (1957), 679–681; English transl., *Amer. Math. Soc. Transl.* (2) **28** (1963), 51–54.

[22] V. I. Arnol'd, "On the representation of functions of several variables by superpositions of functions of fewer variables", *Mat. Prosveshchenie* **1958**, no. 3, 41–61. (Russian)

[23] V. I. Arnol'd, "On the representation of continuous functions of three variables by superpositions of continuous functions of two variables", *Mat. Sb.* **48** (1959), 3–74; English transl., *Amer. Math. Soc. Transl.* (2) **28** (1963), 61–147.

[24] Yu. P. Ofman, "On best approximation of functions of two variables by functions of the form $\varphi(x)+\psi(y)$", *Izv. Akad. Nauk SSSR Ser. Mat.* **25** (1961), 239–252; English transl., *Amer. Math. Soc. Transl.* (2) **44** (1965), 12–28.

[25] V. P. Motornyi, "On best approximation of functions of two variables by functions of the form $\varphi(x)+\psi(y)$", *Izv. Akad. Nauk SSSR Ser. Mat.* **27** (1963), 1211–1214. (Russian)

[26] G. M. Khenkin, "Imbedding the space of s-smooth functions of n variables into a space of sufficiently smooth functions of fewer variables", *Dokl. Akad. Nauk SSSR* **153** (1963), 57–60; English transl., *Soviet Math. Dokl.* **4** (1963), 1633–1636.

[27] R. Doss, "On the representation of the continuous functions of two variables by means of addition and continuous functions of one variable", *Colloq. Math.* **10** (1963), 249–259.

[28] A. G. Vitushkin, "Some properties of linear superpositions of smooth functions", *Dokl. Akad. Nauk SSSR* **156** (1964), 1003–1006; English transl., *Soviet. Math. Dokl.* **5** (1964), 741–744.

[29] G. M. Khenkin, "On linear superpositions of continuously differentiable functions", *Dokl. Akad. Nauk SSSR* **157** (1964), 288–290; English transl., *Soviet. Math. Dokl.* **5** (1964), 948–950.

[30] D. A. Sprecher, "On the structure of continuous functions of several variables", *Trans. Amer. Math. Soc.* **115** (1965), 340–355.

[31] D. A. Sprecher, "On the structure of representation of continuous functions of several variables as finite sums of continuous functions of one variable", *Proc. Amer. Math. Soc.* **17** (1966), 98–105.

[32] A. G. Vitushkin, "On the possibility of representation of functions by superpositions of functions of fewer variables", Proc. International Congress of Mathematicians (Moscow, 1966), Nauka, Moscow 1968, pp. 322–329; English transl., *Amer. Math. Soc. Transl.* (2) **86** (1970), 101–108.

[33] L. A. Bassalygo, "On the representation of continuous functions of two variables by means of continuous functions of one variable", *Vestnik Moskov. Univ. Ser.* I *Mat. Mekh.* **1966**, no. 21, 58–63. (Russian)

[34] B. L. Fridman, "An improvement in the smoothness of the functions in Kolmogorov's theorem on superpositions", *Dokl. Akad. Nauk SSSR* **177** (1967), 1019–1022; English transl., *Soviet Math. Dokl.* **8** (1967), 1550–1553.

[35] A. G. Vitushkin and G. M. Khenkin, "Linear superpositions of functions", *Uspekhi Mat. Nauk* **22**:1 (1967), 77–124; English transl., *Russian Math. Surveys* **22**:1 (1967), 77–126.

[36] A. G. Vitushkin, "On the 13th problem of Hilbert", *Hilbert's problems*, Nauka, Moscow 1969, pp. 163–170; German transl. in *Ostwalds Klassiker Exakt. Wiss*, vol. 252, Akad. Verlag, Leipzig 1971.

[37] B. L. Fridman, "Nowhere denseness of the space of linear superpositions of functions of several variables", *Izv. Akad. Nauk SSSR Ser. Mat.* **36** (1972), 814–846; English transl., *Math. USSR-Izv.* **6** (1972), 807–837.

[38] J.-P. Kahane, "Sur le théorème de superposition de Kolmogorov", *J. Approx. Theory* **13** (1975), 229–234.

[39] G. G. Lorentz, "The 13th problem of Hilbert", *Proc. Sympos. Pure Math.* **28** (1976), 419–430.

[40] V. Ya. Lin, "Superpositions of algebraic functions", *Funktsional. Anal. i Prilozhen.* **10**:1 (1976), 37–45; English transl., *Funct. Anal. Appl.* **10** (1976), 32–38.

[41] A. A. Milyutin, "Isomorphism of spaces of continuous functions on compacta with the cardinality of the continuum", *Teor. Funktsii Funktsional. Anal. i Prilozen.* **1966**, no. 2, 150–156. (Russian)

[42] A. G. Vitushkin, "Variations of functions of several variables and sufficient conditions for the variations to be finite", *Dokl. Akad. Nauk SSSR* **96** (1954), 1089–1091. (Russian)

[43] A. G. Vitushkin, *On multidimensional variations*, Gostehizdat, Moscow 1955. (Russian)

[44] L. D. Ivanov, *Variations of sets and functions*, Nauka, Moscow 1975. (Russian)

[45] I. G. Petrovskii and O. A. Oleinik, "On the topology of real algebraic surfaces", *Izv. Akad. Nauk SSSR Ser. Mat.* **13** (1949), 389–402; English transl., *Amer. Math. Soc. Transl.* (1) **7** (1962), 399–417.

[46] O. A. Oleinik, "Estimates of Betti numbers of real algebraic hypersurfaces", *Math. Sb.* **28** (1951), 635–640. (Russian)

[47] A. N. Kolmogorov, "Estimates of the minimal number of elements of ε-nets in various classes of functions and their applications to the problem of representing functions of several variables by superpositions of functions of fewer variables", *Uspekhi Mat. Nauk* **10**:1 (1955), 192–194. (Russian)

[48] A. N. Kolmogorov, "On certain asymptotic characteristics of totally bounded metric spaces", *Dokl. Akad. Nauk SSSR* **108** (1956), 385–388. (Russian)

[49] A. N. Kolmogorov, A. M. Yaglom, and I. M. Gel'fand, "Amount of information and entropy for continuous distributions", Proc. 3rd All-Union Mathematical Congress (Moscow, 1956), vol. III, Surveys, Izdat. Akad. Nauk SSSR, Moscow 1958, pp. 300–320; English transl., *Selected works of A. N. Kolmogorov*, vol. III: *Information theory and the theory of algorithms*, Kluwer, Dordrecht 1993, paper 4.

[50] A. G. Vitushkin, "Absolute ε-entropy of metric spaces", *Dokl. Akad. Nauk SSSR* **117** (1957), 745–747; English transl., *Amer. Math. Soc. Transl.* (2) **17** (1961), 365–367.

[51] V. M. Tikhomirov, "On ε-entropy of some classes of analytic functions", *Dokl. Akad. Nauk SSSR* **117** (1957), 191–194. (Russian)

[52] A. Pełczyński, "On the approximation of S-spaces by finite dimensional spaces", *Bull. Acad. Polon. Sci. Cl.* III **5** (1957), 879–881.

[53] A. N. Kolmogorov, "On the linear dimension of topological vector spaces", *Dokl. Akad. Nauk SSSR* **120** (1958), 239–242; English transl., *Selected works of A. N. Kolmogorov*, vol. I: *Mathematics and mechanics*, Kluwer, Dordrecht 1991, 388–392.

[54] K. I. Babenko, "On the entropy of a class of analytic functions", *Nauchn. Dokl. Vysshei Shkoly Ser. Fiz. Mat. Nauk* **1958**, no. 2, 9–16. (Russian)

[55] V. D. Erokhin, "On conformal transformations of annuli and the fundamental basis of the space of functions analytic in an elementary neighborhood of an arbitrary continuum", *Dokl. Akad. Nauk SSSR* **120** (1958), 689–692. (Russian)

[56] V. D. Erokhin, "On asymptotics of ε-entropy for analytic functions", *Dokl. Akad. Nauk SSSR* **120** (1958), 949–952. (Russian)

[57] V. D. Erokhin, "On best approximation of analytic functions by rational functions with free poles", *Dokl. Akad. Nauk SSSR* **128** (1959), 29–32. (Russian)

[58] A. N. Kolmogorov and V. M. Tikhomirov, "ε-entropy and ε-capacity of sets in a function space", *Uspekhi Mat. Nauk* **14**:2 (1959), 3–86; English transl., *Amer. Math. Soc. Transl.* (2) **17** (1961), 277–364.

[59] Yu. A. Brudnyi and A. F. Timan, "Constructive characteristics of compact sets in Banach spaces and ε-entropy", *Dokl. Akad. Nauk SSSR* **126** (1959), 927–930. (Russian)

[60] N. S. Bakhvalov, "On the approximate calculation of multiple integrals", *Vestnik Moskov. Univ. Ser.* I *Mat. Mekh.* **1959**, no. 4, 3–18. (Russian)

[61] V. M. Tihomirov, "Widths of sets in function spaces and the theory of best approximations", *Uspekhi Mat. Nauk* **15**:3 (1960), 81–120; English transl., *Russian Math. Surveys* **15**:3 (1960), 75–111.

[62] S. A. Smolyak, "ε-entropy of the classes $E_s^{\alpha,k}(B)$ and $W_s^{\alpha}(B)$ in the metric of L_2", *Dokl. Akad. Nauk SSSR* **131** (1960), 30–33; English transl., *Soviet Math. Dokl.* **1** (1960), 192–196.

[63] B. S. Mityagin, "Relation between the ε-entropy, approximation rate, and nuclearity of a compactum in a linear space", *Dokl. Akad. Nauk SSSR* **134** (1960), 765–768; English transl., *Soviet Math. Dokl.* **1** (1960), 1140–1143.

[64] V. Ya. Pan, "Approximation of analytic functions by rational functions", *Uspekhi Mat. Nauk* **16**:5 (1961), 195–197. (Russian)

[65] B. S. Mityagin, "Approximate dimension and bases in nuclear spaces", *Uspekhi Mat. Nauk* **16**:4 (1961), 63–132; English transl., *Russian Math. Surveys* **16**:4 (1961), 59–128.

[66] C. Bessaga, A. Pełczyński, and S. Rolewicz, "Approximative dimension of linear topological spaces and some of its applications" (Summary of a report), *Studia Math.* (*Ser. Specjalna*) *Zeszyt.* **1** (1963), 27–29.

[67] N. S. Bakhvalov, "Estimating the amount of calculation work required to approximate solutions of problems", S. K. Godunov and V. S. Ryaben'kii, *Introduction to the theory of difference schemes*, Fizmatgiz, Moscow 1962, pp. 316–329; English transl., North-Holland, Amsterdam 1964.

[68] N. S. Bakhvalov, "On the rate of convergence of indeterministic integration processes in the function classes $W_p^{(l)}$", *Teor. Veroyatnost. i Primenen.* **7** (1961), 238; English transl., *Theory Probab. Appl.* **7** (1961), 227, 463.

[69] G. G. Lorentz, "Lower bounds for the degree of approximation", *Trans. Amer. Math. Soc.* **97** (1960), 25–34.

[70] G. G. Lorentz, "Metric entropy, widths, and superpositions of functions", *Amer. Math. Monthly* **69** (1962), 469–485.

[71] V. M. Tikhomirov, "Kolmogorov's work on the ε-entropy of function classes and superpositions of functions", *Uspekhi Mat. Nauk* **18**:5 (1963), 55–92; English transl., *Russian Math. Surveys* **18**:5 (1963), 51–87.

[72] N. M. Korobov, *Number-theoretic methods in numerical analysis*, Fizmatgiz, Moscow 1963. (Russian)

[73] S. A. Smolyak, "Quadrature and interpolation formulas for tensor products of certain classes of functions", *Dokl. Akad. Nauk SSSR* **148** (1963), 1042–1045; English transl., *Soviet Math. Dokl.* **4** (1963), 240–243.

[74] Yu. A. Brudnyi and B. D. Kotlyar, "On the order of growth of the ε-entropy for certain compact classes of functions", *Dokl. Akad. Nauk SSSR* **148** (1963), 1001–1004; English transl., *Soviet Math. Dokl.* **4** (1963), 196–199.

[75] Yu. P. Ofman, "On the approximate realization of continuous functions on automata", *Dokl. Akad. Nauk SSSR* **152** (1963), 823–826; English transl., *Soviet Math. Dokl.* **4** (1963), 1439–1443.

[76] H. S. Shapiro, "Some negative theorems of approximation theory", *Michigan Math. J.* **11** (1964), 211–217.

[77] P. A. Ostrand, "Dimension of metric spaces and Hilbert's problem 13", *Bull. Amer. Math. Soc.* **71** (1965), 619–622.

[78] A. G. Vitushkin and V. I. Buslaev, "An estimate of the length of a signal code with a finite spectrum in connection with sound recording problems", *Izv. Akad. Nauk SSSR Ser. Mat.* **38** (1974), 867–895; English transl., *Math. USSR-Izv.* **8** (1975), 867–894.

[79] V. V. Zmushko, "The entropy of a class of entire functions in a frequency-dependent metric", *Izv. Akad. Nauk SSSR Ser. Mat.* **40** (1976), 1173–1186; English transl., *Math. USSR-Izv.* **10** (1976), 1119-1132.

[80] A. G. Vitushkin, "Some estimates from tabulation theory", *Uspekhi Mat. Nauk* **12**:2 (1957), 227–228. (Russian)

[81] A. G. Vitushkin, "Some estimates from tabulation theory", *Dokl. Akad. Nauk SSSR* **114** (1957), 923–926. (Russian)
[82] A. G. Vitushkin, "On best approximations of smooth and analytic functions", *Dokl. Akad. Nauk SSSR* **119** (1958), 418–420. (Russian)
[83] A. G. Vitushkin, *Estimation of the complexity of the tabulation problem*, Fizmatgiz, Moscow 1959; English transl., *Theory of the transmission and processing of information*, Pergamon Press, Oxford 1961.
[84] V. A. Kotel'nikov, "On the carrying capacity of the 'ether' and wire in telecommunications", Proc. 1st All-Union Congress on Questions of Communications, Izdat. Red. Upr. Svyazi RKKA, Moscow 1933. (Russian)
[85] C. E. Shannon, "A mathematical theory of communication", *Bell. System Tech. J.* **27** (1948), 379–423, 623–656.
[86] A. N. Kolmogorov, "Theory of transmission of information", Session on scientific problems of automatization in industry, vol. 1, Plenary talks, Izdat. Akad. Nauk SSSR, Moscow 1957, pp. 66–99; English transl., *Selected works of A. N. Kolmogorov*, vol. III: *Information theory and the theory of algorithms*, Kluwer, Dordrecht 1993, paper 3.
[87] A. G. Vitushkin, "Coding of signals with finite spectrum and sound recording problems", Proc. International Congress of Mathematicians (Vancouver, 1974), vol. 1, Canad. Math. Congress, Montréal 1975, pp. 221–226.

Steklov Institute of Mathematics

Received 17/JUN/03
Translated by IPS(DoM)

Typeset by $\mathcal{AMS}$-TEX