


## Review

# Review of Methods, Applications and Publications on the Approximation of Piecewise Linear and Generalized Functions

Sergei Aliukov <sup>1,\*</sup>, Anatoliy Alabugin <sup>2</sup> and Konstantin Osintsev <sup>3</sup> 

<sup>1</sup> Department of Automotive Engineering, Institute of Engineering and Technology, South Ural State University, 76 Prospekt Lenina, 454080 Chelyabinsk, Russia

<sup>2</sup> Department of Digital Economy and Information Technology, School of Economics and Management, South Ural State University, 76 Prospekt Lenina, 454080 Chelyabinsk, Russia

<sup>3</sup> Department of Energy and Power Engineering, Institute of Engineering and Technology, South Ural State University, 76 Prospekt Lenina, 454080 Chelyabinsk, Russia

\* Correspondence: alysergey@gmail.com

**Abstract:** Approximation of piecewise linear and generalized functions is an important and difficult problem. These functions are widely used in mathematical modeling of various processes and systems, such as: automatic control theory, electrical engineering, radio engineering, information theory and transmission of signals and images, equations of mathematical physics, oscillation theory, differential equations and many others. The widespread use of such functions is explained by their positive properties. For example, piecewise linear functions are characterized by a simple structure over segments. However, these features also have disadvantages. For example, in the case of using piecewise linear functions, solutions have to be built in segments. In this case, the problem of matching the obtained solutions at the boundaries of the segments arises, which leads to the complication of the research results. The use of generic functions has similar disadvantages. To eliminate shortcomings in practice, one resorts to the approximation of these functions. There are a large number of well-known methods for approximating piecewise linear and generalized functions. Recently, new methods for their approximation have been developed. In this study, an attempt was made to generalize and discuss the existing methods for approximating the considered functions. Particular emphasis is placed on the description of new approximation methods and their applications in various fields of science and technology. The publication-based review discusses the strengths and weaknesses of each method, compares them, and considers suitable application examples. The review will undoubtedly be interesting not only for mathematicians, but also for specialists and scientists working in various applied fields of research.

**Keywords:** piecewise linear functions; generalized functions; methods of approximation; advantages and disadvantages; application examples

**MSC:** 37L65



**Citation:** Aliukov, S.; Alabugin, A.; Osintsev, K. Review of Methods, Applications and Publications on the Approximation of Piecewise Linear and Generalized Functions. *Mathematics* **2022**, *10*, 3023. <https://doi.org/10.3390/math10163023>

Academic Editor: Jay Jahangiri

Received: 18 July 2022

Accepted: 19 August 2022

Published: 22 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In many areas of mathematics and its applications, a problem often arises related to replacing functions that are complex or inconvenient in one sense or other with simpler ones, but close to the original ones. This problem is called the approximation of functions. In many areas of mathematics and its applications, the problem often arises of replacing functions that are complex or inconvenient in one sense or other with simpler ones, but close to the original ones. A number of studies have been devoted to solving this problem, and a large amount of scientific and educational literature has been published. As an example, one can point to the fundamental books [1–4] and many others.

The review considers methods of approximation of piecewise linear and generalized functions [5–17], which are universal and find wide application for solving a variety

of problems related to mathematical modeling of systems, processes and phenomena described using such functions.

Piecewise linear and generalized functions are widely used in various fields of scientific research. Their traditional fields of application are technical and mathematical disciplines, for example, automatic control theory, electrical engineering, radio engineering, information theory and transmission of signals and images, equations of mathematical physics, oscillation theory, differential equations and many others [18–24]. Using such functions, for example, they describe the dynamics of mechanical systems with nonlinear elasticities, elastic-dissipative characteristics of vehicle suspensions, systems with loads of the “dry friction” type, impulse transformations during transmission and reception of signals, distributed and concentrated loads, and many other processes.

The widespread use of piecewise linear functions is explained by the simplicity of their structure, especially by areas. At each site, these functions are straight lines and their segments, which in many cases allows one to obtain solutions using the methods of the theory of linear systems. At the same time, problems often arise when constructing solutions over the entire domain of definition of piecewise-linear functions, linking solutions for sections with the need to use special mathematical methods. Systems with piecewise linear characteristics and functions are referred to as essentially nonlinear structures, emphasizing the complexity of obtaining solutions for such structures. Despite the simplicity of piecewise linear functions over sections, the construction of solutions in problems with piecewise linear functions over the entire domain of definition requires the use of special mathematical methods, for example, the matching method [25], with the need to match solutions over sections and switching surfaces. So, when constructing periodic solutions, it is necessary to monitor the fulfillment of the conditions for the transition of the system from section to section, fixing the values of the variables at the end of the previous section and taking these values as the initial conditions for the next section. The need to match the values of the variables over the sections, as well as at the beginning and end of the cycle, leads to a cumbersome system of transcendental equations. Therefore, the application of the fitting method often requires overcoming significant mathematical difficulties, and even if the periodic solution can be found in analytical form, it is usually obtained in the form of complex expressions. In addition, the derivatives of piecewise linear functions are not continuous, which in some cases also complicates research. To simplify calculations, one often resorts to approximating piecewise linear functions. The existing approximation methods have their drawbacks.

Generalized functions were introduced in connection with the problems of physics and mathematics that appeared in the twentieth century and required a new understanding of the concept of a function. Such problems include the problems of determining the density of a point mass, the intensity of a point charge and a point dipole, problems of quantum theory and many others. Generalized functions are now widely used in a wide variety of research areas. For example, generalized functions and their derivatives represent a very convenient mathematical tool for developing control systems. The use of impulse controls significantly increases the control capabilities of various systems. An example of a generalized function is the  $\delta$ -function or the Dirac function.

Generalized singular functions are very different from regular functions. In many practical applications, these unusual functions are used as purely abstract mathematical constructions, completely divorced from their physical understanding. This approach does not seem to be correct. For a better understanding of generalized functions and a conscious decision on their basis of practical problems, you can use the approximations of these functions. Some of these possible approximations are suggested in this review.

The content of the new methods is explained by a number of practical examples. The examples given are taken from a wide variety of fields: structural mechanics, medicine, quantum theory, signal theory, semiconductor theory, mechanical engineering, heat engineering, and others. The variety of the considered examples emphasizes the rather broad possibilities of practical application of the conceded approximation methods. Therefore, the

review will undoubtedly be interesting not only for mathematicians, but also for specialists and scientists working in various applied fields of research.

## 2. Basic Provisions and Methods of Approximation Theory

This section of the review introduces the basic concepts of approximation theory, discusses the canonical methods of approximation of continuous and discontinuous functions [26–29]. The considered methods are illustrated with examples, the positive and negative sides of these methods are noted.

### 2.1. The Main Idea of the Approximation of the Original Function: Basic Concepts

Let  $M$  be some point set in  $n$ -dimensional space and let  $f(A)$ ,  $A \in M$  be some function defined on this set. This function  $f(A)$ , due to some considerations that we will talk about below, we want to replace with other, the so-called approximation function  $\varphi(A, B_1, B_2, \dots, B_k)$ , also defined on the set  $M$ , where  $B_1, B_2, \dots, B_k$  are parameters. It is necessary to determine the parameters so that the deviation of the function  $\varphi(A, B_1, B_2, \dots, B_k)$  from the original function  $f(A)$  would be the least according to some criteria.

It is clear that the meaning of such a replacement will be only when the original function does not suit us in some way, and therefore we want to go to a function as close to the original, as possible, but devoid of shortcomings of the original function. For example, dissimilar to the original function  $f(A)$ , the approximation function  $\varphi(A, B_1, B_2, \dots, B_k)$  can have a simpler structure, be continuous, differentiable, analytical, allow the use of any mathematical methods and so on. The approximation function must belong to a certain type of function, which has these advantages over the original function and the properties of functions of this type are well studied in mathematics. For example, an algebraic polynomial, a fractional rational function, a trigonometric sum, a spline function, and so on can act as an approximation function.

When constructing an approximating function, the question arises: what is to be understood by the deviation, or, in other words, by the proximity of functions, to determine the approximation error. To solve this issue, the concepts of metric and metric space are introduced in functional analysis.

**Definition 1.** A set  $Y$  is called a metric space if each pair of its elements  $f_1$  and  $f_2$  is associated with a real number  $\rho(f_1, f_2) \geq 0$ , for which the following axioms hold:

1. Identity axiom:  $\rho(f_1, f_2) = 0$  if and only if  $f_1 = f_2$ ;
2. Symmetry axiom:  $\rho(f_1, f_2) = \rho(f_2, f_1)$ ;
3. Triangle axiom:  $\rho(f_1, f_2) + \rho(f_2, f_3) \geq \rho(f_1, f_3)$ , for  $\forall f_3 \in Y$ .

This number  $\rho(f_1, f_2)$  is called the metric of the set  $Y$  or the distance between the elements  $f_1$  and  $f_2$ .

Within the framework of this monograph, we mainly consider the issues of approximation of such mathematical objects as functions. In this case, the elements (points) of the sets under consideration are functions. Therefore, we will often denote points of sets and spaces by the letter symbol  $f$ , as is conducted in the above definition of a metric space.

Consider examples of metrics and metric spaces that are important from the point of view of approximation of functions.

1. Let  $Y$  be the set of all continuous functions on the segment  $[a, b]$ . As a metric, we can take the maximum modulus of the difference  $\rho(f_1, f_2) = \max_{x \in [a, b]} |f_1(x) - f_2(x)|$ .

Without losing the generality of reasoning, the segment  $[a, b]$  with the introduction of a new variable can always be reduced to the segment  $[0, 1]$ . Therefore, the set of all continuous functions defined on a closed interval with the metric introduced in this way is called the space of all continuous functions and is denoted by  $C[0, 1]$ ;

2. Let  $Y$  be a set of all real bounded functions on the interval  $[0,1]$ . In this case, as a metric, we can take the supremum of the modulus of the difference  $\rho(f_1, f_2) = \sup_{x \in [0,1]} |f_1(x) - f_2(x)|$ .

The set of all real bounded functions with the metric introduced in this way is called the space  $M[0,1]$ . It is clear that  $C[0,1] \subset M[0,1]$ ;

3. Let  $Y$  be the set of all measurable functions defined on the interval  $[0,1]$ . Two functions that differ only on a set of measure zero (coinciding almost everywhere) will be considered identical. As a metric, we can take the equality  $\rho(f_1, f_2) = \int_0^1 \frac{|f_1(x) - f_2(x)|}{1 + |f_1(x) - f_2(x)|} dx$ . Such a space is called the space  $S[0,1]$ . Convergence in this space is convergence in measure, that is, a sequence of elements  $f_n \rightarrow f$ , if  $\rho(f_n, f) \xrightarrow{n \rightarrow \infty} 0$ ;
4. Let  $Y$  be the set of all functions with integrable  $p$ -th power on the interval  $[0,1]$ . Recall that a function  $f(x)$  is called a function with integrable  $p$ -th power on the interval  $[a,b]$ , if  $\int_a^b |f(x)|^p dx < \infty$ . The integral is considered in the sense of Lebesgue. As in the previous case, we will consider two functions that differ only on a set of measure zero to be identical. In this case, as a metric, we can take the integral  $\rho(f_1, f_2) = \left( \int_a^b |f_1(x) - f_2(x)|^p dx \right)^{1/p}$ . Such a space is called the space  $L_p[0,1]$ . For  $p = 2$  we obtain the so-called functional Hilbert space.

In the further description, we need concepts such as norm and linear normed space.

**Definition 2.** A set  $E$  is called a linear space if the operations of adding elements and multiplying an element by a number are defined in this set, and for any elements  $f_1, f_2, f_3 \in E$  and for any numbers  $\alpha, \beta$  the conditions are met:

1.  $f_1 + f_2 = f_2 + f_1$ ;
2.  $(f_1 + f_2) + f_3 = f_1 + (f_2 + f_3)$ ;
3.  $f_1 + 0 = f_1$ ;
4.  $f_1 + (-f_1) = 0$ ;
5.  $\alpha(f_1 + f_2) = \alpha f_1 + \alpha f_2$ ;
6.  $(\alpha + \beta)f_1 = \alpha f_1 + \beta f_1$ ;
7.  $\alpha(\beta f_1) = (\alpha\beta)f_1$ ;
8.  $1 \cdot f_1 = f_1$ .

**Definition 3.** The norm of an arbitrary element  $f$  of the set  $E$  is a nonnegative number, which is denoted by  $\|f\|$ , for which the conditions are met:

1.  $\|f\| = 0 \Leftrightarrow f = 0$ ;
2.  $\|\alpha \cdot f\| = |\alpha| \cdot \|f\|$ ;
3.  $\|f_1 + f_2\| \leq \|f_1\| + \|f_2\|$ .

**Definition 4.** The linear space  $E$  with the introduced norm is called a normed linear space.

A metric in a normed linear space can be introduced using the equality  $\rho(f_1, f_2) = \|f_1 - f_2\|$ . Convergence in a normed linear space is convergence in the norm, that is, a sequence of elements  $f_n \rightarrow f$ , if  $\|f_n - f\| \xrightarrow{n \rightarrow \infty} 0$ .

Let us give examples of some normed linear spaces that are of great importance from the point of view of approximation theory.

1. The space of all continuous functions  $C[0,1]$  with operations of addition and multiplication by a number, defined in the usual way. The norm is introduced by the relation  $\|f\| = \max_x |f(x)|$ ;



2. The space  $L_p[0,1]$  of all functions with integrable  $p$ -th power on the interval  $[0,1]$  with operations of addition and multiplication by a number, defined in the usual way. We introduce a norm in the space of such functions using the equality  $\|f\| = \left(\int_a^b |f(x)|^p dx\right)^{1/p}$ ;
3. The space of all continuous functions on the segment  $[0,1]$  with derivatives continuous on this segment up to the  $k$ -th order inclusive. The notation for such a space is  $C^k[0,1]$ . As a norm in this space, one usually takes the relation  $\|f\| = \max\left\{\max_x |f(x)|, \max_x |f'(x)|, \dots, \max_x |f^{(k)}(x)|\right\}$ .

The basic idea of approximation in a normed linear space can be expressed by the following theorem.

**Theorem 1.** Let  $E$  be some normed linear space in which the elements  $f_1, f_2, \dots, f_n$  are linearly independent. Let some element  $f \in E$  be given. For the element  $f$  one can choose such numbers  $\lambda_1, \lambda_2, \dots, \lambda_n$ , that the value  $\Delta(\lambda_1, \lambda_2, \dots, \lambda_n) = \|f - (\lambda_1 f_1 + \lambda_2 f_2 + \dots + \lambda_n f_n)\|$  obtains the smallest value.

Let us present the proof of the theorem [22].

Let  $\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*$  be an arbitrary set of numbers

Using the reverse triangle inequality, we estimate the modulus of the difference

$$\begin{aligned} |\Delta(\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*) - \Delta(\lambda_1, \lambda_2, \dots, \lambda_n)| &= \left| \|f - \sum_{k=1}^n \lambda_k^* f_k\| - \|f - \sum_{k=1}^n \lambda_k f_k\| \right| \leq \left\| \sum_{k=1}^n (\lambda_k^* - \lambda_k) f_k \right\| \\ &\leq \sum_{k=1}^n |\lambda_k^* - \lambda_k| \|f_k\| \leq \max_k |\lambda_k^* - \lambda_k| \cdot \sum_{k=1}^n \|f_k\|. \end{aligned}$$

Therefore, for  $(\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*) \rightarrow (\lambda_1, \lambda_2, \dots, \lambda_n)$  we obtain that  $(\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*) \rightarrow \Delta(\lambda_1, \lambda_2, \dots, \lambda_n)$ . Therefore, the function  $\Delta(\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*)$  is continuous.

Consider a continuous function  $\Omega(\lambda_1, \lambda_2, \dots, \lambda_n) = \|\lambda_1 f_1 + \lambda_2 f_2 + \dots + \lambda_n f_n\|$ .

The continuity of the function  $\Omega(\lambda_1, \lambda_2, \dots, \lambda_n)$  can be proved similarly to the proof of the continuity of the function  $\Delta(\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*)$ .

The sphere  $\sum_{k=1}^n |\lambda_k|^2$  is a bounded closed set in a finite-dimensional Euclidean space, therefore, based on the well-known Weierstrass theorem, the function  $\Omega(\lambda_1, \lambda_2, \dots, \lambda_n)$  attains its minimum on this set, which we denote by  $m$ . Note that  $m > 0$ , since the function  $\Omega(\lambda_1, \lambda_2, \dots, \lambda_n)$  is a norm, and the elements  $f_1, f_2, \dots, f_n$  are linearly independent.

Let  $\overline{\Omega} \geq 0$  be some lower bound for the set of values of the function  $\Omega(\lambda_1, \lambda_2, \dots, \lambda_n)$ .

If  $\sqrt{\sum_{k=1}^n |\lambda_k|^2} > \frac{1}{m}(\overline{\Omega} + 1 + \|f\|) = R$ , then, it is easy to obtain

$$\Delta(\lambda_1, \lambda_2, \dots, \lambda_n) \geq \|\lambda_1 f_1 + \lambda_2 f_2 + \dots + \lambda_n f_n\| - \|f\| \geq \sqrt{\sum_{k=1}^n |\lambda_k|^2} m - \|f\| > \overline{\Omega} + 1,$$

therefore, to find the minimum, we can consider the function  $(\lambda_1, \lambda_2, \dots, \lambda_n)$  only in a closed bounded domain  $\sum_{k=1}^n |\lambda_k|^2 < R^2$ . According to the well-known Weierstrass theorem, a continuous function reaches its minimum in such a domain. Therefore, there are numbers  $\lambda_1, \lambda_2, \dots, \lambda_n$ , that provide the best approximation of the element  $f$  by the linear combination  $\lambda_1 f_1 + \lambda_2 f_2 + \dots + \lambda_n f_n$ .

## 2.2. Contraction Mapping Principle

The contraction mapping principle is one of the most important mathematical achievements and is widely used to prove existence and uniqueness theorems, find solutions to equations of various types by the method of successive approximations, and prove the convergence of iterative procedures used, for example, in approximation theory. The contraction mapping principle was formulated by the Polish mathematician S. Banach [30–34].

First, we give the following definition.

Let  $Y$  be some metric space and  $\{f_n\}$ —a sequence of elements of this space.

**Definition 5.** A sequence  $\{f_n\}$  is called Cauchy sequence, if for any positive number  $\varepsilon$  there is a number  $N_0$ , depending on  $\varepsilon$ , such that for any numbers  $n$  and  $m$  it follows from the condition  $n, m \geq N_0$  than  $\rho(f_n, f_m) < \varepsilon$ .

In quantifiers, this definition will be written as follows

$$\forall \varepsilon > 0 \exists N_0 = N_0(\varepsilon) \in \mathbb{N} \forall n, m \in \mathbb{N} : n, m > N_0 \Rightarrow \rho(f_n, f_m) < \varepsilon.$$

A space  $Y$  is called complete, if any Cauchy sequence converges to some limit that is an element of the same space.

**Theorem 2. (Principle of contraction mappings).** Let  $Y$  be a complete metric space, in which some operator  $A$ , is given, transforming elements of a given space into elements of the same space, i.e.,  $A: Y \rightarrow Y$ . Let, for any  $f, \varphi \in Y$ , there exist a number  $\alpha < 1$ , independent of  $f$  and  $\varphi$ , such that the condition  $\rho(Af, A\varphi) \leq \alpha \rho(f, \varphi)$  is satisfied. Then, there is a single element (point)  $f_0: Af_0 = f_0$ . This point is called the fixed point of the operator  $A$ .

**Proof.** Let  $f$  be an arbitrary fixed element of the set  $Y$ .

Let us create a sequence  $f_1 = Af, f_2 = Af_1, \dots, f_n = Af_{n-1}, \dots$ , which is fundamental.

Really,

$$\rho(f_1, f_2) = \rho(Af, Af_1) \leq \alpha \rho(f, Af),$$

$$\rho(f_2, f_3) = \rho(Af_1, Af_2) \leq \alpha \rho(f_1, f_2) \leq \alpha^2 \rho(f, Af) \text{ and so on.}$$

We obtain,  $\rho(f_n, f_{n+1}) \leq \alpha^n \rho(f, Af)$ .

By the triangle axiom for the metric, we write

$$\begin{aligned} \rho(f_n, f_{n+p}) &\leq \rho(f_n, f_{n+1}) + \rho(f_{n+1}, f_{n+2}) + \dots + \rho(f_{n+p-1}, f_{n+p}) \\ &\leq (\alpha^n + \alpha^{n+1} + \dots + \alpha^{n+p-1}) \rho(f, Af) = \frac{\alpha^n - \alpha^{n+p}}{1 - \alpha} \rho(f, Af), \quad p \in \mathbb{N}. \end{aligned}$$

Taking into account that  $\alpha < 1$ , we obtain  $\rho(f_n, f_{n+p}) \leq \frac{\alpha^n}{1 - \alpha} \rho(f, Af)$ .

Therefore,  $\rho(f_n, f_{n+p}) \xrightarrow{n \rightarrow \infty} 0$ . Therefore, the sequence  $\{f_n\}$  is fundamental, and since, by the hypothesis of the theorem, the space  $Y$  is complete, there is an element  $f_0 \in Y: \lim_{n \rightarrow \infty} f_n = f_0$ .

Let us show that this element  $f_0$  is a fixed point.

Really,

$$\rho(f_0, Af_0) \leq \rho(f_0, f_n) + \rho(f_n, Af_0) = \rho(f_0, f_n) + \rho(Af_{n-1}, Af_0) \leq \rho(f_0, f_n) + \alpha \rho(f_{n-1}, f_0).$$

Take an arbitrary number  $\varepsilon > 0$ . Since the sequence  $\{f_n\}$  converges to the element  $f_0$ , then, there is a number  $N_1$ , for which for any  $n \geq N_1$  the condition  $\rho(f_0, f_n) < \frac{\varepsilon}{2}$ , will be satisfied, and there is a number  $N_2$ , for which for any  $n \geq N_2$  the condition  $\alpha \rho(f_{n-1}, f_0) < \frac{\varepsilon}{2}$  is satisfied.

Then, we obtain

$$\forall \varepsilon > 0 \exists N(\varepsilon) = \max\{N_1, N_2\} \forall n : n > N \Rightarrow \rho(f_0, Af_0) \leq \rho(f_0, f_n) + \alpha \rho(f_{n-1}, f_0) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Since the number  $\varepsilon$  is arbitrary, then  $\rho(f_0, Af_0) = 0$ , whence we obtain  $Af_0 = f_0$ , that is, the element  $f_0$  is a fixed point.

Suppose that along with the fixed point  $f_0$  there is one more, belonging to  $Y$ , which we denote by  $f_0^*$ . By the definition of a fixed point, we write  $Af_0 = f_0$  and  $Af_0^* = f_0^*$ . In this case  $\rho(f_0, f_0^*) = \rho(Af_0, Af_0^*) < \alpha \rho(f_0, f_0^*)$  and  $\alpha > 1$ , if the fixed points are different. The obtained value of  $\alpha$  contradicts the hypothesis of the theorem; therefore, the fixed point is the only one.  $\square$

### 2.3. Weierstrass Theorems on the Convergence of a Sequence of Approximating Functions

Let us define the concept of uniform convergence of a sequence  $\{f_n(x)\}$  of elements of a metric space.

Let  $Y$  be a metric space.

**Definition 6.** A sequence  $\{f_n(x)\}$  of elements of this space  $Y$  converges uniformly on the segment  $[0,1]$  to an element  $f(x) \in Y$ , if

$$\forall \varepsilon > 0 \exists N_0(\varepsilon) \in \mathbb{N} \forall n \in \mathbb{N} : n \geq N_0 \Rightarrow \rho(f_n(x), f(x)) < \varepsilon \forall x.$$

**Theorem 3. Weierstrass' First Theorem.** If the function  $f(x)$  is continuous on the interval  $[0,1]$ , then  $\forall \varepsilon > 0 \exists n(\varepsilon) \in \mathbb{N}$ , that there is a sequence of polynomials  $P_n(x)$  of degree  $n$ , for which the relation holds:

$$|f(x) - P_n(x)| < \varepsilon, \forall x \in [0, 1].$$

In other words, one can construct a sequence of polynomials  $\{P_n(x)\}$ , uniformly converging to the original function  $f(x)$  on the segment  $[0,1]$ . The polynomials of this sequence can be used to approximate the function  $f(x)$ .

**Proof of Theorem.** Consider a sequence of polynomials  $A_n(x) = \sum_{k=0}^n C_n^k \cdot x^k (1-x)^{n-k} f\left(\frac{k}{n}\right)$ , where  $C_n^k$  are binomial coefficients.

Differentiating the binomial relation

$$\sum_{k=0}^n C_n^k \cdot p^k q^{n-k} = (p+q)^n$$

in  $p$  twice and carrying out simple transformations, we write down the relations

$$\begin{aligned} \sum_{k=0}^n \frac{k}{n} C_n^k \cdot x^k (1-x)^{n-k} &= x; \\ \sum_{k=0}^n \frac{k^2}{n^2} C_n^k \cdot x^k (1-x)^{n-k} &= \left(1 - \frac{1}{n}\right)x^2 + \frac{1}{n}x. \end{aligned}$$

Taking into account the identity  $\sum_{k=0}^n C_n^k \cdot x^k (1-x)^{n-k} = 1$ , the obtained relations allow us to find  $\sum_{k=0}^n \left(\frac{k}{n} - x\right)^2 C_n^k \cdot x^k (1-x)^{n-k} = \frac{x(1-x)}{n}$ .

Let's write the expression

$$f(x) - A_n(x) = \sum_{k=0}^n \left(f(x) - f\left(\frac{k}{n}\right)\right) C_n^k \cdot x^k (1-x)^{n-k} = B_1 + B_2,$$

where  $B_1$  corresponds to terms for which  $\left|\frac{k}{n} - x\right| \leq n^{-1/4}$ .  $B_2$  matches all other terms.

We set  $\lambda_n = \max_{\left|\frac{k}{n} - x\right| \leq n^{-1/4}} \left|f(x) - f\left(\frac{k}{n}\right)\right|$ . Then, we obtain  $|B_1| \leq \lambda_n \sum_{k=0}^n C_n^k \cdot x^k (1-x)^{n-k} = \lambda_n$ . Note that  $\lambda_n \xrightarrow{n \rightarrow \infty} 0$ .

By the conditions of the theorem, the function  $f(x)$  is continuous on a closed segment and, therefore, is bounded. Therefore,  $f(x) \leq M$  on this segment. Where we can obtain

$$\begin{aligned} |B_2| &\leq 2MB_2 C_n^k x^k (1-x)^{n-k} = 2MB_2 \frac{(k-nx)^2}{(k-nx)^2} C_n^k x^k (1-x)^{n-k} \leq \\ &\leq 2MB_2 \frac{(k-nx)^2}{n^{3/2}} C_n^k x^k (1-x)^{n-k} \leq 2M \frac{1}{n^{3/2}} \sum_{k=0}^n (k-nx)^2 C_n^k x^k (1-x)^{n-k} = \\ &= 2M \frac{1}{n^{3/2}} nx(1-x) \leq \frac{M}{2\sqrt{n}}. \end{aligned}$$

We will finally write down

$$|f(x) - A_n(x)| = |B_1 + B_2| \leq |B_1| + |B_2| \leq \lambda_n + \frac{M}{2\sqrt{n}}.$$

It is easy to see that  $\lambda_n + \frac{M}{2\sqrt{n}} \rightarrow 0$ .

Whence it follows that the sequence of polynomials  $A_n(x)$  tends uniformly to the function  $f(x)$  on a given interval.  $\square$

**Theorem 4. Weierstrass' Second Theorem.** *If a periodic function  $f(x)$  with period  $2\pi$  is continuous, then  $\forall \varepsilon > 0 \exists n(\varepsilon) \in \mathbb{N}$ , there is a trigonometric sum  $S_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \sin kx + b_k \cos kx)$ , for which the relation  $|f(x) - S_n(x)| < \varepsilon \forall x$  holds.*

**Proof.** We introduce two even continuous functions with period  $2\pi$  using the relations

$$\mu(x) = \frac{f(x) + f(-x)}{2} \text{ and } \nu(x) = \frac{f(x) - f(-x)}{2} \sin x.$$

Let  $t = \cos x$ . Then, if  $x \in [0, \pi]$ , then  $t \in [-1, 1]$ . Note that due to the parity and periodicity of the functions under consideration, all conclusions that are valid for  $x \in [0, \pi]$ , will be valid for any  $x$ .

Let us introduce the functions  $\varphi(t) = \mu(x)$  and  $\psi(t) = \nu(x)$ , which are continuous for  $t \in [-1, 1]$ . Based on the first Weierstrass theorem, for any number  $\varepsilon > 0$  there are polynomials  $P(t)$  and  $Q(t)$ , for which the conditions  $|\mu(x) - P(\cos x)| < \frac{\varepsilon}{4}$  and  $|\nu(x) - Q(\cos x)| < \frac{\varepsilon}{4}$ .

Since the relation  $f(x) \sin x = \mu(x) \sin x + \nu(x)$  holds, there is a trigonometric sum  $A(x) = Q(\cos x) + P(\cos x) \sin x$ , for which, for all  $x$  the relation holds

$$\begin{aligned} |f(x) \sin x - A(x)| &= |\mu(x) - P(\cos x) \sin x + \nu(x) - Q(\cos x)| \\ &\leq |\mu(x) - P(\cos x) \sin x| + |\nu(x) - Q(\cos x)| < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

Similarly, for the function  $f(\frac{\pi}{2} - x)$  there is a trigonometric function  $B(x)$ , for which  $|f(\frac{\pi}{2} - x) \cos x - B(x)| < \frac{\varepsilon}{2}$ .

Making the substitution  $\frac{\pi}{2} - x = x$ , we rewrite the last inequality in the form  $|f(x) \cos x - B(\frac{\pi}{2} - x)| < \frac{\varepsilon}{2}$ .

From the inequalities obtained, we find

$$\left| f(x) \sin^2(x) - A(x) \sin x \right| < \frac{\varepsilon}{2}; \left| f(x) \cos^2(x) - B\left(\frac{\pi}{2} - x\right) \cos x \right| < \frac{\varepsilon}{2}.$$

where do we obtain

$$\begin{aligned} \left| f(x) - A(x) \sin x - B\left(\frac{\pi}{2} - x\right) \cos x \right| &= \left| f(x) \sin^2(x) - A(x) \sin x + f(x) \cos^2(x) - B\left(\frac{\pi}{2} - x\right) \cos x \right| \\ &\leq \left| f(x) \sin^2(x) - A(x) \sin x \right| + \left| f(x) \cos^2(x) - B\left(\frac{\pi}{2} - x\right) \cos x \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Since  $\varepsilon$  can be any positive number, including arbitrarily small, it can be concluded that the trigonometric sum is uniformly convergent  $S(x) = A(x) \sin x + B(\frac{\pi}{2} - x) \cos x$  to the function  $f(x)$ .  $\square$

Note that the considered Weierstrass theorems can be generalized to the case of the space  $L_p[0, 1]$ .

#### 2.4. Approximation by Algebraic Polynomials

The approximation by algebraic polynomials of functions [35–39], in particular, piecewise-linear ones, is often performed using algebraic polynomials in the system of polynomials  $1, x, x^2, \dots$ . These polynomials are simple and well-studied mathematical constructions, with the possibility of simple differentiation, and the derivative is again a polynomial.

Let the function  $f: R \rightarrow R$  be subject to approximation. Let  $x \in [a, b]$  and the values of the function  $f_k = f(x_k)$  are known at the points  $x_k, k = 0, 1, \dots, n$ . We approximate our function by the algebraic polynomial  $P_n(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n, a_0 \neq 0$ , for which the condition  $P_n(x_k) = f(x_k)$ . Such a polynomial exists and is unique.

The approximating polynomial can be found by the Lagrange formula

$$L_n(x) = \sum_{k=0}^n \frac{\prod_{i=0, i \neq k}^n (x - x_i) f(x_k)}{(x - x_k) \prod_{i=0, i \neq k}^n (x_k - x_i)}.$$

Example. Approximate the piecewise linear function on the segment  $[0, 7]$

$$f(x) = \begin{cases} 2x, & \forall x \in [0, 2]; \\ 6 - x, & \forall x \in (2, 7]. \end{cases}$$

Let's calculate the values of the function at several points

$$f(0) = 0, f(1) = 2, f(2) = 4, f(3) = 3, f(5) = 1, f(7) = -1.$$

Using the calculated values, we construct an approximating polynomial in the Lagrange form

$$\begin{aligned} L_5(x) = & \frac{x(x-1)(x-2)(x-3)(x-5)(x-7) \cdot 0}{-210x} + \frac{x(x-1)(x-2)(x-3)(x-5)(x-7)2}{48(x-1)} \\ & + \frac{x(x-1)(x-2)(x-3)(x-5)(x-7)4}{-30(x-2)} + \frac{x(x-1)(x-2)(x-3)(x-5)(x-7)3}{48(x-3)} \\ & + \frac{x(x-1)(x-2)(x-3)(x-5)(x-7)1}{-240(x-5)} + \frac{x(x-1)(x-2)(x-3)(x-5)(x-7)(-1)}{1680(x-7)}. \end{aligned}$$

After transformations, we find

$$L_5(x) = -0.034x^5 + 0.548x^4 - 2.941x^3 + 5.495x^2 - 1.068x.$$

Figure 1 shows the graphs of the original function (thickened line) and the approximating function (thin line). As you can see, despite the relatively high degree of the approximating polynomial, the approximation error is large.

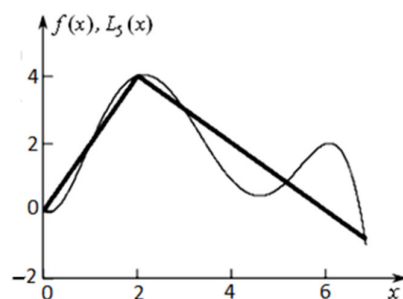


Figure 1. Graphs of the original and approximating functions.

The approximation error can be found by the relation [35]

$$|f(x) - L_n(x)| \leq \frac{M_{n+1} |\prod_{k=0}^n (x - x_k)|}{(n+1)!}, \text{ where } M_{n+1} = \sup_{x \in [a, b]} |f^{(n+1)}(x)|.$$

This assessment should be carefully considered. It may not be valid for all continuous functions due to the appearance of a derivative of order  $n+1$ . Convergence issues are considered in the works [1–4].

Since the expression for the error includes the product  $\prod_{k=0}^n (x - x_k)$ , the error depends on the choice of points  $x_k$ . The approximation error reaches the smallest value on the

interval  $[-1,1]$  if the points  $x_k$  are the roots of the Chebyshev polynomial  $G_{n+1}(x) = \frac{\cos((n+1)\arccos x)}{2^n}$ , which are calculated by the formula  $x_k = \cos \frac{2k+1}{2(n+1)}$ ,  $k = 0, 1, \dots, n$ .

In the case  $x \in [a, b]$ , there are optimal points as follows

$$x_k = \frac{a+b}{2} + \frac{b-a}{2} \cos\left(\frac{2k+1}{2(n+1)}\right).$$

Moreover,  $\max_{x \in [a,b]} |\prod_{k=0}^n (x - x_k)| = \frac{(b-a)^{n+1}}{2^{2n+1}}$ , and the approximation error estimate takes the form  $|f(x) - L_n(x)| \leq \frac{M_{n+1}(b-a)^{n+1}}{2^{2n+1}(n+1)!}$ .

The approximation accuracy can be improved by increasing the points (nodes) of the approximation.

To approximate the function, you can apply a polynomial in the Newtonian form

$$N_n(x) = f(x_0) + (x - x_0)f(x_0, x_1) + (x - x_0)(x - x_1)f(x_0, x_1, x_2) + \dots + (x - x_0)(x - x_1) \dots (x - x_{n-1})f(x_0, x_1, \dots, x_n),$$

where  $f(x_i, x_j) = \frac{f(x_j) - f(x_i)}{x_j - x_i}$ ,  $i, j = 0, 1, \dots, n$ , ( $i \neq j$ ) are separated first-order differences,  $f(x_i, x_j, x_k) = \frac{f(x_j, x_k) - f(x_i, x_j)}{x_k - x_i}$ ,  $i, j, k = 0, 1, \dots, n$ , ( $i \neq j \neq k$ ) are separated second-order differences,  $f(x_i, x_{i+1}, \dots, x_{i+k}) = \frac{f(x_{i+1}, x_{i+2}, \dots, x_{i+k}) - f(x_i, x_{i+1}, \dots, x_{i+k-1})}{x_{i+k} - x_i}$  are separated  $k$ -th-order differences.

Newton's formula is applied at equidistant points at which the values of the original function are calculated. The advantage of Newton's formula over Lagrange's formula is that when adding new approximation points, all the coefficients in the Lagrange's formula have to be recalculated, whereas only new terms are added in Newton's formula, while the old ones remain unchanged.

Newton's formula is a difference analogue of Taylor's formula, which is used in the case of approximation by algebraic polynomials of an analytic function in a neighborhood of some point  $x_0$ , and which has the form

$$f(x) = f(x_0) + \sum_{k=1}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(n+1)}(x_0 + \theta(x - x_0))}{(n+1)!} (x - x_0)^{n+1}, \theta \in [0, 1],$$

where the remainder  $\frac{f^{(n+1)}(x_0 + \theta(x - x_0))}{(n+1)!} (x - x_0)^{n+1}$  is written in Lagrange form.

In some cases, the original function cannot be approximated with the required accuracy by algebraic polynomials. Sometimes such an approximation is possible, but the sequence of polynomials converges very slowly. In these cases, rational fractions or fractional rational functions representing the ratio of polynomials are used to approximate the function [1–3].

## 2.5. Approximation of Piecewise Linear Functions by Fourier Series

This paragraph is based on the publications [40–44].

Let  $Y$  be a linear space. Let us introduce the scalar product operation on this space, which is defined as follows.

**Definition 7.** To each pair of elements  $f_1, f_2 \in Y$  we associate some (generally complex) number  $(f_1, f_2)$ , satisfying the conditions:

1.  $(f_1, f_1) \geq 0$ , and  $(f_1, f_1) = 0 \Leftrightarrow f_1 = 0$ ;
2.  $(f_1, f_2) = \overline{(f_2, f_1)}$ ;
3.  $(f_1 + f_2, f_3) = (f_1, f_3) + (f_2, f_3)$ ,  $\forall f_3 \in Y$ ;
4.  $(\lambda f_1, f_2) = \lambda (f_1, f_2)$  for any complex number  $\lambda$ .



So, the entered number is called the dot product of the elements  $f_1$  and  $f_2$ .

The norm of an element  $f \in Y$  can be introduced using the relation  $\|f\| = \sqrt{(f, f)}$ , and the metric in this space can be defined by the relation  $\rho(f_1, f_2) = \|f_1 - f_2\|$ . It is easy to verify that all the axioms of the norm and metrics are satisfied.

Elements  $f_1, f_2$  are called orthogonal if  $(f_1, f_2) = 0$ .

A system of elements  $f_1, f_2, \dots, f_n \in Y$  is called orthonormal if for any elements  $f_i, f_j$  of this system the condition  $(f_i, f_j) = \delta_{ij}$  is satisfied, where  $\delta_{ij}$  is the Kronecker symbol equal to one for  $i = j$  and zero for  $i \neq j$ .

Let the system of functions  $\{f_1(x), f_2(x), \dots, f_k(x), \dots\}$  be orthogonal on the segment  $[a, b]$ . A series of the form  $\sum_{k=1}^{\infty} c_k f_k(x)$ , in which the coefficients  $c_k$  are found by the formulas  $c_k = \frac{1}{\|f_k(x)\|^2} \int_a^b f(x) f_k(x) dx$ , is called the Fourier series in the orthogonal system  $\{f_1(x), f_2(x), \dots, f_k(x), \dots\}$  for the function  $f(x)$ . For the orthonormal system of functions  $\{f_1(x), f_2(x), \dots, f_k(x), \dots\}$  the coefficients of the Fourier series can be found by the formulas  $c_k = \int_a^b f(x) f_k(x) dx$ .

There are various orthogonal systems of functions. Often the trigonometric system is  $\{1, \sin x, \cos x, \sin 2x, \cos 2x, \dots, \sin kx, \cos kx, \dots\}$ , for which the scalar product is defined by the relation  $(f_i, f_j) = \int_a^b f_i(x) f_j(x) dx$ .

If the functions in this system are normalized, then we obtain the orthonormal system of functions

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\sin kx}{\sqrt{\pi}}, \frac{\cos kx}{\sqrt{\pi}}, \dots \right\}.$$

A trigonometric series is a series of the form

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \sin kx + b_k \cos kx),$$

where  $a_0, a_k, b_k$  ( $k = 1, 2, \dots$ ) are the coefficients of the trigonometric series.

The sum of a converging trigonometric series is a periodic function with a period of  $2\pi$ , since the functions  $\sin kx$  and  $\cos kx$  are periodic with a period  $2\pi$ .

If the coefficients of the trigonometric series are found by the formulas

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx; \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx; \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx,$$

Then, the trigonometric series is called the Fourier series in the orthogonal system  $\{1, \sin x, \cos x, \sin 2x, \cos 2x, \dots, \sin kx, \cos kx, \dots\}$ .

**Theorem 5. Dirichlet's Theorem.** *If the original function  $f(x)$  is periodic with period  $2\pi$ , piecewise monotone and bounded on the interval  $[-\pi, \pi]$ , then the Fourier series for this function converges at all points. At the points of continuity of the function, the sum of the series is equal to the value of the function at these points. If  $x_0$  is the point of discontinuity of the original function, then the sum of the series at this point is equal to the half-sum of the one-sided limits, i.e.,*

$$(f(x_0 - 0) + f(x_0 + 0))/2.$$

The proof of the theorem is omitted.

Example. Let a periodic function  $f(x)$  with period  $2\pi$  be defined as

$$f(x) = \begin{cases} 0, & \forall x \in (-\pi, 0]; \\ x, & \forall x \in (0, \pi]. \end{cases}$$

Find the coefficients of the Fourier series

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx = \frac{1}{\pi} \left( \int_{-\pi}^0 0 dx + \int_0^{\pi} x dx \right) = \frac{1}{\pi} \cdot \frac{\pi^2}{2} = \frac{\pi}{2};$$

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx = \frac{1}{\pi} \int_0^{\pi} x \sin kx dx = \frac{1}{\pi} \left( -\frac{x \cos kx}{k} I_0^{\pi} + \int_0^{\pi} \cos kx dx \right) = -\frac{1}{k} \cos kx;$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx = \frac{1}{\pi} \int_0^{\pi} x \cos kx dx = \frac{1}{\pi} \left( \frac{x \sin kx}{k} I_0^{\pi} - \frac{1}{k} \int_0^{\pi} \sin kx dx \right) = \frac{1}{\pi k^2} (\cos kx - 1).$$

Then, the Fourier series for the original function will be written as follows

$$f(x) = \frac{\pi}{4} + \sin x - \frac{2}{\pi} \cos x - \frac{1}{2} \sin 2x + \frac{1}{3} \sin 3x - \frac{2}{9\pi} \cos 3x + \dots$$

The graphs of the original piecewise linear function (thickened line) and its approximations by several successive partial sums of the Fourier series (thin lines) on the segment  $[-\pi, \pi]$  are shown in Figure 2.

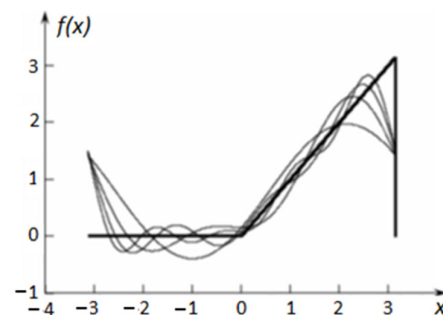


Figure 2. Graphs of the original function and its approximations.

As you can see in Figure 2, the approximation error is large enough and, as will be shown in the next chapter, in the vicinity of the discontinuity points, even with an infinite increase in the number of terms in the Fourier series; the error, understood as the difference between the values of the original function and its approximation, does not tend to zero.

For an odd function, the coefficients of the Fourier series are found by the formulas

$$a_0 = 0; a_k = \frac{2}{\pi} \int_0^{\pi} f(x) \sin kx dx; b_k = 0.$$

for an even function, these coefficients can be found as follows

$$a_0 = \frac{2}{\pi} \int_0^{\pi} f(x) dx; a_k = 0; b_k = \frac{2}{\pi} \int_0^{\pi} f(x) \cos kx dx.$$

In the case of a periodic function with a period of  $2l$  by changing the variable, one can always go to a function with a period of  $2\pi$ . In this case, the Fourier series for a function with a period  $2l$  will have the form

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left( a_k \sin \frac{k\pi}{l} + b_k \cos \frac{k\pi}{l} \right),$$

where the coefficients are found by the formulas

$$a_0 = \frac{1}{l} \int_{-l}^l f(x) dx; a_k = \frac{1}{l} \int_{-l}^l f(x) \sin \frac{k\pi}{l} x dx; b_k = \frac{1}{l} \int_{-l}^l f(x) \cos \frac{k\pi}{l} x dx.$$

If a piecewise-monotone non-periodic function is given, the values of which are of interest to us only on a certain interval  $[a, b]$ , then to expand this function in a Fourier series, we can use a periodic function with a period  $2l \geq b - a$ , which coincides with the original function on the interval  $[a, b]$ .

## 2.6. Function Approximation Using Splines

A spline is a function composed of parts of polynomials that form a basis [45–49]. The polynomials  $1, x, x^2, \dots$  are usually taken as a basis. In the general case, the functions forming the basis may not be polynomials, but in the overwhelming majority of cases, so-called polynomial splines are constructed, the basic functions of which are precisely polynomials.

Some advantages of approximating the original function using splines can be pointed out:

1. Stability of splines with respect to outliers and bursts;
2. Good convergence of the approximation method;
3. Ease of implementation on computers using well-developed mathematical methods such as the sweep method.

There are other advantages to this kind of approximation.

Let's consider the basic idea of spline-function approximation.

Let the segment  $[a, b]$  by points  $x_0, x_1, x_2, \dots, x_i, x_{i+1}, \dots, x_{n-1}, x_n$ , be divided into partial segments, so that

$$a = x_0 < x_1 < x_2 < \dots < x_i < x_{i+1} < \dots < x_{n-1} < x_n = b.$$

It is said that a grid is given on the segment  $[a, b]$ . In addition, let  $P_k$  be the set of all polynomials of degree at most  $k$ , and  $C^{(k)}[a, b]$  is the set of all continuous functions defined on the segment  $[a, b]$  and having continuous derivatives on this segment up to the  $k$ -th order inclusive.

**Definition 8.** The function  $S_k(x)$  is called a spline of degree  $k$  of a defect  $d$  with a given grid if the following conditions are met:

1.  $S_k(x) \in P_k, \forall x \in [x_i, x_{i+1}], i = \overline{0, n-1}$ ;
2.  $S_k(x) \in C^{(k-d)}[a, b]$ .

Let some initial function  $f(x)$  defined on the segment  $[a, b]$  be given. A spline can be used to approximate this function by putting  $S_k(x_i) = f(x_i), \forall i$ . In this case, the grid nodes are called approximation nodes.

Consider an example of drawing up a parabolic spline, that is, a spline of second degree.

Let there be a function  $y(x)$ , which at the nodes  $x = 0, 1, 2$  has the values 5, 2, 3, respectively.

For this function on the segment  $[0, 2]$  we construct a spline of the form

$$S_2(x) = \begin{cases} a_1x^2 + b_1x + c_1, & \forall x \in [0, 1]; \\ a_2x^2 + b_2x + c_2, & \forall x \in [1, 2]. \end{cases}$$

From the condition  $S_2(x_i) = y(x_i)$  we obtain

$$\begin{cases} c_1 = 5; \\ a_1 + b_1 + c_1 = 2; \\ a_2 + b_2 + c_2 = 2; \\ 4a_2 + 2b_2 + c_2 = 3. \end{cases}$$

From the continuity condition for the first derivative, we find

$$2a_1 + b_1 = 2a_2 + b_2.$$

In total, we obtained five equations, but we also have six unknowns. Additional necessary equations are usually derived from some considerations, most often associated with boundary conditions. Putting, for example,  $S_2''(0) = 0$ , we obtain the missing condition  $b_1 = 0$ .

Solving all the obtained equations in the system, we find the values of all coefficients  $a_1 = -3$ ,  $a_2 = 7$ ,  $b_1 = 0$ ,  $b_2 = -20$ ,  $c_1 = 5$ ,  $c_2 = 15$ .

The desired spline will be written as follows

$$S_2(x) = \begin{cases} -3x^2 + 5, & \forall x \in [0, 1); \\ 7x^2 - 20x + 15, & \forall x \in [1, 2]. \end{cases}$$

The spline plot is shown in Figure 3.

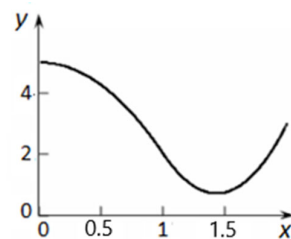


Figure 3. Parabolic spline plot.

Many practical examples, especially those related to mechanics, consider cubic splines, that is, splines of the third degree. Such splines allow not only the first derivative to be continuous, but also the second order derivative. In this case, it is possible to simulate the laws of motion with continuous speeds and accelerations. However, since we are primarily interested in piecewise linear functions, we will consider splines of the first degree, or, in other words, linear splines. The graphs of such splines will be continuous broken lines. For such splines, the following conditions will be met:

1.  $S_1(x) \in P_1, \forall x \in [x_i, x_{i+1}], i = \overline{0, n-1}$ ;
2.  $S_1(x) \in C[a, b]$ ;
3.  $S_1(x_i) = y(x_i), \forall i$ .

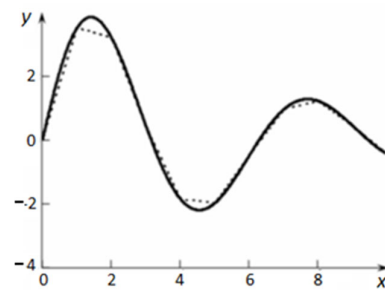
Example. The original function on the segment  $[0, 10]$  is defined by the expression

$$y = \frac{5 \sin x}{\sqrt[3]{2 + \exp(3x)}}.$$

Taking the points at which  $x = 0, 1, 2, \dots, 10$ , as the nodes of the approximation, we construct a linear spline, which will have the form

$$S(x) = \begin{cases} 3.507x, & x \in [0, 1); \\ -0.314x + 3.821, & x \in [1, 2); \\ -2.778x + 8.749, & x \in [2, 3); \\ -2.284x + 7.266, & x \in [3, 4); \\ -0.116x - 1.404, & x \in [4, 5); \\ 1.499x - 9.123, & x \in [5, 6); \\ 1.44x - 9.123, & x \in [6, 7); \\ 0.251x - 0.798, & x \in [7, 8); \\ -0.785x + 7.483, & x \in [8, 9); \\ -0.8787x + 8.402, & x \in [9, 10]. \end{cases}$$

The graphs of the original function (solid line) and its approximating linear spline (dashed line) are shown in Figure 4.



**Figure 4.** An example of approximating a function by a linear spline.

The error in approximating the original functions using linear splines can be quite large. Nevertheless, in some cases, approximation by linear splines may be more preferable than approximation by splines of higher degrees, for example, due to simpler expressions for linear splines. In addition, the monotonicity of the values of the origin of the specified function for a linear spline is preserved, which may not be the case for splines of higher degrees.

To reduce the error, the number of nodal points can be increased, but at the same time, having a simple structure at each partial section, the linear spline acquires a cumbersome appearance as a whole, which is clearly seen in the example considered. In addition, already the first derivative for a linear spline is not continuous. This drawback often prevents linear spline functions from being used to solve practical problems. For example, the study of the dynamics of motion of various objects involves the use of velocities and accelerations, which are derivatives of the angles of rotation and displacement. Discontinuities in the functions for velocities and accelerations create uncertainties and inconsistencies between mathematical models and real processes. A way out of the situation can be the methods of approximation of piecewise linear splines, considered in the subsequent chapters and paragraphs of the review.

### 2.7. Least Squares Method: Linear Regression

The approximating function  $F(x)$  by the least squares method [50,51] is determined from the condition of the minimum sum of squared deviations ( $\xi_i$ ) of the calculated approximating function from a given array of experimental data. This criterion of the least squares method is written as the following expression:

$$\sum_{i=1}^N \xi_i^2 = \sum_{i=1}^N (F(x_i) - y_i)^2 \rightarrow \min,$$

where  $F(x_i)$  are the values of the calculated approximating function  $F(x)$  at the nodal points  $x_i$ ;  $y_i$  is a given array of experimental data at nodal points  $x_i$ .

This method can be useful when dealing with a large amount of information.

As an example, consider the method for determining the approximating function, which is given as a linear relationship [52]. In accordance with the least squares method, the condition for the minimum sum of squared deviations is written as follows:

$$S = \sum_{i=1}^N \xi_i^2 = \sum_{i=1}^N (a_0 + a_1 \cdot x_i - y_i)^2 \rightarrow \min,$$

where  $x_i, y_i$  are coordinates of nodal points of the table;  $a_0, a_1$  are unknown coefficients of the approximating function, which is given as a linear dependence.

The necessary condition for the existence of a minimum of a function is the equality to zero of its partial derivatives with respect to unknown variables. Then, we obtain:

$$\begin{cases} \frac{\partial S}{\partial a_0} = 2 \cdot \sum_{i=1}^N (a_0 + a_1 \cdot x_i - y_i) = 0; \\ \frac{\partial S}{\partial a_1} = 2 \cdot \sum_{i=1}^N (a_0 + a_1 \cdot x_i - y_i) \cdot x_i = 0. \end{cases}$$

After some transformations we have:

$$\begin{cases} a_0 \cdot N + a_1 \cdot \sum_{i=1}^N x_i = \sum_{i=1}^N y_i; \\ a_0 \cdot \sum_{i=1}^N x_i + a_1 \cdot \sum_{i=1}^N x_i^2 = \sum_{i=1}^N y_i \cdot x_i. \end{cases}$$

Solving the resulting system of linear equations, we find the coefficients of the approximating function:

$$\begin{cases} a_0 = \frac{\sum_{i=1}^N y_i \cdot \sum_{i=1}^N x_i^2 - \sum_{i=1}^N (y_i \cdot x_i) \cdot \sum_{i=1}^N x_i}{N \cdot \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}; \\ a_1 = \frac{N \cdot \sum_{i=1}^N (y_i \cdot x_i) - \sum_{i=1}^N y_i \cdot \sum_{i=1}^N x_i}{N \cdot \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}. \end{cases}$$

These coefficients are used to construct a linear approximating function according to the criterion of the minimum sum of squares of the approximating function from the given tabular values representing the experimental data.

Example. Suppose we have initial data (Table 1).

**Table 1.** Initial data.

$x_i$	17.28	17.05	18.30	18.80	19.20	18.50
$y_i$	537	534	550	555	560	552

Using the above formulas, we find the pair of regression coefficients:  $a_0 = 328.3$ ,  $a_1 = 12.078$ .

Then, the regression equation will take the form

$$y = 12.078 \cdot x + 328.3.$$

## 2.8. Hermite Interpolation

When constructing the Hermite interpolation polynomial, it is required that not only that its values coincide with the tabular data in the nodes, but also the values of its derivatives are in a certain order [53,54].

Let us now look at the nodes  $x_i \in [a, b]$ ,  $i = 0, 1, \dots, m$ , among which there are no coinciding ones, the values of the function  $f(x_i)$  and its derivatives  $f^{(j)}(x_i)$ ,  $j = 1, 2, \dots, N_i - 1$  up to  $N_i - 1$  order are given. In this case, the numbers  $N_i$  are called the multiplicity of the node  $x_i$ . At each point  $x_i$ , thus, the following values  $N_i$  are given:  $f(x_i), f'(x_i), f''(x_i), \dots, f^{(N_i-1)}(x_i)$ . In total, the values  $N_0 + N_1 + \dots + N_m$  are known on the entire set of nodes  $x_0 + x_1 + \dots + x_m$ , which makes it possible to raise the question of constructing a polynomial  $H_n(x)$  of order  $n = N_0 + \dots + N_m - 1$ , satisfying the requirements:

$$H_n^{(j)}(x_i) = f^{(j)}(x_i), \quad i = 0, 1, \dots, m, \quad j = 1, 2, \dots, N_i - 1.$$

Such a polynomial is called the Hermite interpolation polynomial for the function  $f(x)$ . It is proved [] that the Hermite interpolation polynomial exists and is unique.



The construction of the Hermite polynomial in the general case for an arbitrary number of nodes and their multiplicity leads to rather cumbersome expressions and is rarely used. Therefore, we confine ourselves to one example.

Example. Construct the Hermite interpolation polynomial for the function  $f(x)$  in the case when at all interpolation nodes  $x_i \in [a, b]$ ,  $i = 0, 1, \dots, m$ , the values of the function  $f(x_i) = f_i$  and its first derivative  $f'(x_i) = f'_i$ .

In this case  $N_i = 2$ ,  $i = 0, 1, \dots, m$ , therefore, the degree of the polynomial  $H_n(x)$  is  $2m + 1$ . We write the original polynomial in the form:

$$H_{2m+1}(x) = \sum_{i=0}^m (f_i + \alpha_i(x - x_i)) \cdot \frac{(x - x_0)^2 \cdot \dots \cdot (x - x_{i-1})^2 \cdot (x - x_{i+1})^2 \cdot \dots \cdot (x - x_m)^2}{(x_i - x_0)^2 \cdot \dots \cdot (x_i - x_{i-1})^2 \cdot (x_i - x_{i+1})^2 \cdot \dots \cdot (x_i - x_m)^2}.$$

When calculating the derivative of the polynomial  $H_{2m+1}(x)$  at the node  $x_i$ , it should be taken into account that all terms of the sum, except for the term corresponding to the node itself, provide zero contribution to the derivative at this point, so we obtain

$$H'_{2m+1}(x) = f_i \cdot \left( \frac{2}{x_i - x_0} + \dots + \frac{2}{x_i - x_{i-1}} + \frac{2}{x_i - x_{i+1}} + \frac{2}{x_i - x_m} \right) + \alpha_i = f'_i.$$

Therefore, we obtain  $\alpha_i = f'_i - 2f_i A_i$ , where the numbers  $A_i$  are determined by the formula

$$A_i = \frac{1}{x_i - x_0} + \dots + \frac{1}{x_i - x_{i-1}} + \frac{1}{x_i - x_{i+1}} + \frac{1}{x_i - x_m}.$$

Thus, the solution to this problem is the Hermite polynomial

$$H_{2m+1}(x) = \sum_{i=0}^m (f_i + (f'_i - 2f_i A_i) \cdot (x - x_i)) \cdot \frac{(x - x_0)^2 \cdot \dots \cdot (x - x_{i-1})^2 \cdot (x - x_{i+1})^2 \cdot \dots \cdot (x - x_m)^2}{(x_i - x_0)^2 \cdot \dots \cdot (x_i - x_{i-1})^2 \cdot (x_i - x_{i+1})^2 \cdot \dots \cdot (x_i - x_m)^2}.$$

## 2.9. Lebesgue Functions and Lebesgue Constant in Polynomial Interpolation

The Lebesgue constant is a valuable numerical tool for linear interpolation because it provides a measure of how close the interpolation of a function is to the best polynomial approximation of a function. Many publications [55–59] have been devoted to finding optimal interpolation points in the sense that these points lead to the minimum Lebesgue constant for interpolation problems on the interval  $[-1, 1]$ .

**Definition 9.** Let  $\Omega_n = \{x_i\}_{i=1}^n$  be a grid on  $[a, b]$ .

Function  $\Lambda_n(x) = \Lambda_n(x, x_1, x_2, \dots, x_n) = \sum_{i=1}^n |l_i(x)|$  is called the Lebesgue function, and the Lebesgue constant is the number

$$\Lambda_n = \Lambda_n(x_1, x_2, \dots, x_n) = \max_{x \in [a, b]} \Lambda_n(x, x_1, x_2, \dots, x_n).$$

Here  $l_1(x), l_2(x), \dots, l_n(x)$  is some basis in the linear (vector) space of functions of dimension  $n$ .

The statements are true [59]:

1.  $1 \leq \Lambda_n(x) \leq \Lambda_n$  for any  $x \in [a, b]$ .
2. The value of  $\Lambda_n$  does not depend on  $[a, b]$ , but depends only on the relative position of the nodes on it.

Let us pose the question: to what extent is the method of interpolation of a function by an algebraic polynomial inferior in accuracy to the best possible method of approximating a function by an algebraic polynomial of the same degree?

Let  $P_{n-1}$  be an algebraic polynomial, an approximation of the function  $f$ , obtained by some method. Thus, each method has its own polynomial  $P_{n-1}$ . The value  $|f(x) - P_{n-1}(x)|$

determines the approximation error at a point  $x \in [a, b]$ , and the number  $\|f - P_{n-1}\|_\infty = \max_{x \in [a, b]} |f(x) - P_{n-1}(x)|$  is the maximum error of this method.

**Definition 10.** An algebraic polynomial  $L_{n-1}$  is called a polynomial of best uniform approximation if  $\|f - L_{n-1}\|_\infty = \min_{P_{n-1}} \|f - P_{n-1}\|_\infty$ . The solution to this problem exists and is uniquely determined. The value  $E_n(f) = \|f - L_{n-1}\|_\infty$  is called the error of the best uniform approximation.

Let's make the following remarks:

1. If  $F_{n-1}$  is an approximation of  $f$  obtained by some method (for example,  $F_{n-1}$  is an interpolation polynomial), then  $\|f - F_{n-1}\|_\infty \geq E_n(x)$ .
2.  $E_n(x) \rightarrow 0$  as  $n \rightarrow \infty$  for any function  $f$  continuous on  $[a, b]$ . This follows directly from the Weierstrass' theorem.

**Theorem 6.** The estimates  $E_n(f) \leq \|f - R_{n-1}\|_\infty \leq (1 + \Lambda_n)E_n(f)$  are valid.

**Proof.** Let  $L_{n-1}$  be a polynomial of the best uniform approximation of  $f$ . Since the interpolation polynomial is unique,  $L_{n-1}(x) = \sum_{i=1}^n L_{n-1}(x_i)l_i(x)$ . Therefore,

$$\begin{aligned} |f(x) - R_{n-1}(x)| &= |f(x) - L_{n-1}(x) + L_{n-1}(x) - R_{n-1}(x)| \\ &\leq |f(x) - L_{n-1}(x)| + \sum_{i=1}^n |L_{n-1}(x_i) - f(x_i)|l_i(x) \\ &\leq (1 + \Lambda_n(x))\|f - L_{n-1}\|_\infty \leq (1 + \Lambda_n)E_n(f). \end{aligned}$$

□

The lower estimate is valid by the definition of  $E_n(f)$ . It follows from the upper estimate that the interpolation polynomial  $R_{n-1}(x)$  is less accurate than the best uniform approximation by a maximum of  $1 + \Lambda_n$  times in accuracy.

Let's pose the second question: how sensitive is the interpolation polynomial to the error in setting the function?

Let the approximate values  $\tilde{f}(x_i)$  be known at the interpolation nodes instead of the exact values of  $f(x_i)$  with an error  $\epsilon(x_i)$  not exceeding  $\epsilon$ :  $|\epsilon(x_i)| = |f(x_i) - \tilde{f}(x_i)| \leq \epsilon$ . Thus, instead of  $R_{n-1}(x)$  the perturbed polynomial  $\tilde{R}_{n-1}(x)$  will be constructed from the values of  $\tilde{f}(x_i)$ . Of practical interest is the deviation of  $\tilde{R}_{n-1}$  from  $f$ .

**Theorem 7.** The estimate is  $|f(x) - \tilde{R}_{n-1}(x)| \leq |f(x) - R_{n-1}(x)| + \epsilon\Lambda_n$ .

**Proof.** Obviously,

$$|R_{n-1}(x) - \tilde{R}_{n-1}(x)| = \left| \sum_{i=1}^n (f(x_i) - \tilde{f}(x_i))l_i(x) \right| \leq \epsilon\Lambda_n.$$

Therefore,

$$|f(x) - \tilde{R}_{n-1}(x)| \leq |f(x) - R_{n-1}(x)| + |R_{n-1}(x) - \tilde{R}_{n-1}(x)| \leq |f(x) - R_{n-1}(x)| + \epsilon\Lambda_n.$$

□

It follows from estimate that the larger  $\Lambda_n$ , the more sensitive the interpolation procedure to the error in setting the function.

Important conclusions follow from the estimates obtained [59]:

1. The smaller the Lebesgue constant  $\Lambda_n$ , and the smoother the function, the better both in terms of accuracy and the sensitivity of interpolation to the error of setting the function;
2. If the sequence of grids  $\Omega_n = \{x_i\}_{i=1}^n$  satisfies the condition  $\Lambda_n E_n(f) \rightarrow 0$  as  $n \rightarrow \infty$ , then  $|f(x) - R_{n-1}(x)| \rightarrow 0$  as  $n \rightarrow \infty$  uniformly in  $x$  (in this case one speaks of the convergence of the interpolation process);
3. During calculations, the following picture can be observed: the error  $|f(x) - \tilde{R}_{n-1}(x)|$  as  $n$  increases, it first decreases and then begins to increase.

The value of  $\Lambda_n$  depends on the choice of nodes  $\Omega_n$ .

A detailed consideration of issues related to the significance of the Lebesgue constant, moduli of smoothness, selection of optimal nodes, weighted polynomial interpolation is given in fundamental works [1–4].

### 3. New Methods of Approximation of Piecewise-Linear and Generalized Functions

This section of the review describes new methods for approximating piecewise-linear functions, especially, piecewise constant functions, and generalized functions, a comparative analysis of the proposed and existing methods for approximating such functions by analytical dependences based on Fourier series is carried out. In addition, the issues of convergence and error of the proposed methods are studied, numerous examples and applications are considered.

The general idea of using a repeating procedure, which gives a more accurate result with each subsequent application, is the basis, for example, of the mathematical theory of deep learning. Deep learning is a type of machine learning using multi-layer neural networks that learn on their own on a large dataset. Artificial intelligence with deep learning itself finds an algorithm for solving the original problem, learns from its mistakes, and after each iteration of training gives a more accurate result.

#### 3.1. Disadvantages of Approximating Piecewise Linear Functions by Fourier Series

To simplify calculations, when working with piecewise linear and generalized functions, in many cases they resort to approximation methods. Replacing piecewise linear functions with more regular  $C^k$  functions allows you not to worry about tracking and matching the values of process variables at the boundaries of the sections, which greatly simplifies the calculations. In some cases, algebraic polynomials are used to approximate piecewise linear functions. Another of the most widely used methods for approximating piecewise linear functions is the expansion of these functions using Fourier series  $f = \sum_{k=1}^{\infty} c_k \varphi_k$ , where  $\{\varphi_1, \varphi_2, \dots, \varphi_n, \dots\}$  is orthogonal system in the functional Hilbert space  $L_2[-\pi, \pi]$  of measurable functions with Lebesgue integrable squares,  $f \in L_2[-\pi, \pi]$ ,  $c_k = (f, \varphi_k) / \|\varphi_k\|^2$ . The trigonometric system of  $2\pi$ -periodic functions  $\{1, \sin nx, \cos nx, n \in N\}$  is often taken as the orthogonal system.

As for the approximation of continuous functions by polynomials or Fourier series, we can discuss the uniform convergence of the approximating functions based on the Weierstrass theorems.

However, for discontinuous piecewise linear functions, the Weierstrass theorems do not hold. Therefore, when approximating such functions, problems may arise that cause negative consequences when solving applied problems. For example, the use of Fourier series, along with positive properties, has certain disadvantages. With a relatively small number of terms in the Fourier series used for the expansion of piecewise-linear functions, the approximating function has a pronounced wavy character even within one rectilinear section of the piecewise-linear function, which leads to a sufficiently large approximation error. To approximate continuous functions. Curves 1 and 2 in Figure 5 illustrate this drawback.

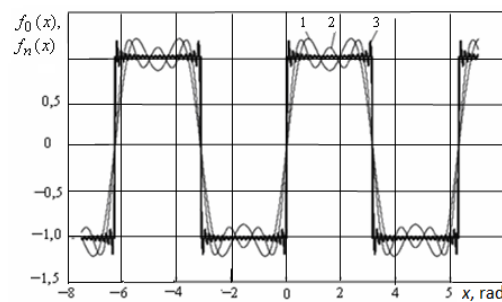


Figure 5. Approximation errors using Fourier series expansion.

Moreover, even for a large number of terms in the expansion using the Fourier series, there are characteristic jumps of the approximating function in the vicinity of the discontinuity points  $O_\delta(x_0)$  of the original function. For such points  $\sup_{x \in O_\delta(x_0)/\{x_0\}} |f(x) - S_n(x)| \xrightarrow{n \rightarrow \infty} A \neq 0$ , where  $S_n(x)$  is the partial sum of the Fourier series [60].

For example, for the function

$$f_0(x) = \text{sign}(\sin x) \quad (1)$$

with rectangular pulses, the point  $x = \pi/m$ , where  $m = 2[(n+1)/2]$ ,  $[A]$  is the integer part of the number  $A$ , is the maximum point of the partial sum  $S_n(f_0)$  of the trigonometric Fourier series [61], moreover  $S_n(f_0, \pi/m) \xrightarrow{n \rightarrow \infty} \frac{2}{\pi} \int_0^\pi \frac{\sin t}{t} dt \approx 1.17898$ . That is, the magnitude of the absolute error  $|f_0(\pi/m) - \lim_{n \rightarrow \infty} S_n(f_0, \pi/m)| > 0.178$ , and the relative error is more than 17%, regardless of the number of terms in the partial sum of the Fourier series. Notice, that  $x = \pi/m \xrightarrow{n \rightarrow \infty} 0+0$ .

In Figure 5, curve 3 corresponds to the graph of the approximating function  $f_n(x) = \sum_{n=1}^{20} c_n \varphi_n$  and illustrates the increased approximation error in the vicinity of the discontinuity points of the original Function (1). This is the manifestation of the so-called Gibbs effect, and with an increase in the number of harmonics, the Gibbs effect does not disappear, which leads to extremely negative consequences of using the approximating function. Figure 6 shows a graph of the partial sum  $S_{20}(f_0)$  of the trigonometric series on the segment  $[-\pi, \pi]$ , illustrating the manifestation of the Gibbs effect.

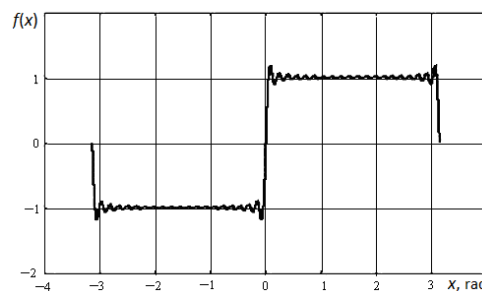


Figure 6. Manifestation of the Gibbs effect.

The most unpleasant thing is that the Gibbs effect is general in nature, it manifests itself for any function  $f \in L_2[a, b]$ , that has bounded variation on a segment  $[a, b]$ , with an isolated breakpoint  $x_0 \in (a, b)$ . For such functions, the following condition is satisfied [61]

$$\lim_{n \rightarrow \infty} S_n(f, x_0 + \pi/m) = f(x_0 + 0) + \frac{d}{2} \cdot \left( \frac{2}{\pi} \int_0^\pi \frac{\sin t}{t} dt - 1 \right),$$

where  $d = f(x_0 + 0) - f(x_0 - 0)$ .

Let us show that the absolute  $\Delta = \Delta(x)$  and conditional  $\delta = \delta(x)$  approximation errors in the vicinity of the discontinuity points can be arbitrarily large.

Really,

$$\begin{aligned} \lim_{n \rightarrow \infty} \Delta(x_0 + \pi/m) &= \lim_{n \rightarrow \infty} |S_n(f, x_0 + \pi/m) - f(x_0 + \pi/m)| = \left| \lim_{n \rightarrow \infty} S_n(x_0 + \pi/m) - \lim_{n \rightarrow \infty} f(x_0 + \pi/m) \right| \\ &= \left| f(x_0 + 0) + \frac{d}{2} \cdot \left( \frac{2}{\pi} \int_0^\pi \frac{\sin t}{t} dt - 1 \right) - f(x_0 + 0) \right| = \left| \frac{d}{2} \cdot \left( \frac{2}{\pi} \int_0^\pi \frac{\sin t}{t} dt - 1 \right) \right| = \Delta(d). \end{aligned}$$

For each  $d$  there exists a function  $f = f_d$  satisfying the previous conditions. The property  $\Delta(d) \rightarrow +\infty$  has to be understood in this way. The function  $\Delta(d)$  is infinitely large since

$$\forall M > 0 \exists d = d^*(M) > 0 \forall d : |d| > d^* \Rightarrow \Delta(d^*) = \left| \frac{d^*}{2} \cdot \left( \frac{2}{\pi} \int_0^\pi \frac{\sin t}{t} dt - 1 \right) \right| > M.$$

$$\text{As } d^* \text{ you can take, for example, } \left[ 2M\pi / \left( 2 \int_0^\pi \frac{\sin t}{t} dt - \pi \right) \right] + 1.$$

For the relative error  $\delta(x) = \Delta(x)/|f(x)|$  the proof is similar. Moreover, even with a fixed value  $d \in \mathbf{R} (d \neq 0)$  for any  $M > 0$  you can choose a function  $f(x) \in L_2[a, b]$ , for which  $\delta(x_0 + 0, d) = \Delta(x_0 + 0, d)/|f(x_0 + 0)| > M$ . As such a function, for example, one can take a function  $f$ , or which  $|f(x_0 + 0)| < \Delta(x_0 + 0, d)/M$ ,  $f(x_0 + 0) \neq 0$ .

Note that even on the set of continuous functions  $C[-\pi, \pi]$  the Fourier series, as is known [31], does not yet have to converge at every point.

The existence of the Gibbs effect leads to extremely negative consequences of using a partial sum of a trigonometric series as an approximating function for solving problems of mathematical modeling, for example, when studying periodic movements of technical systems, distortions in signal transmission, etc.

The approximation error is especially striking when using Fourier series for generalized functions, for example,  $\delta$ —function or, in other words, Dirac function. This function is widely used to describe the density of a point mass, the density of a point charge, quantum theory, concentrated loads, instantaneous impulse processes, shock effects, the intensity of a point heat source, diffusion processes in semiconductors, etc.

Generalized functions were introduced in connection with the problems of physics and mathematics that appeared in the twentieth century and required a new understanding of the concept of a function. Generalized singular functions are very different from regular functions. It is known that the  $\delta$ —function is not a function in the usual sense of this word; rather, it is determined by a functional, and informally by the expression

$$\delta(x) = \begin{cases} +\infty, & x = 0, \\ 0, & \forall x \neq 0, \end{cases}$$

$$\text{moreover } \int_{-\infty}^{+\infty} \delta(x) dx = 1.$$

Generalized functions were introduced in mathematics by Sobolev and Schwartz [62–65]. Generalized functions have become a key tool in much of PDE theory and form a huge part of analysis.

Let us provide a mathematically more precise definition of a generalized function.

**Definition 11.** A generalized function in the sense of Sobolev-Schwartz is any linear continuous functional on the space of basic functions [23].

Thus: (1) the generalized function  $f$  is a functional on the set of basic functions  $D$  [23], that is, each support of the piecewise continuous function [23]  $\varphi \in D$  is associated with a (complex) number  $(f, \varphi)$ ; (2) a generalized function  $f$  is a linear functional on  $D$ , that

is, if  $\varphi, \psi \in D$  and  $\lambda, \mu$  are complex numbers, then  $(f, \lambda\varphi + \mu\psi) = \lambda(f, \varphi) + \mu(f, \psi)$ ; (3) the generalized function  $f$  is a continuous functional on  $D$ , that is, if  $\varphi_k \rightarrow \varphi$ ,  $k \rightarrow \infty$  in  $D$ , then  $(f, \varphi_k) \rightarrow (f, \varphi)$ ,  $k \rightarrow \infty$ .

A very intuitive graph of  $\delta$ -function is shown in Figure 7.

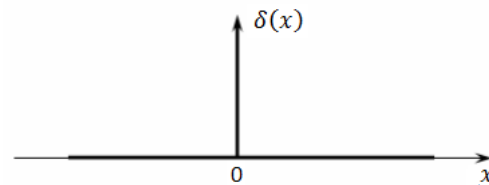


Figure 7. A very intuitive graph of  $\delta$ -function.

For the convenience of using analytical research methods, the delta function is decomposed into a Fourier series.

We introduce a sequence of step functions of the form

$$\delta_n(x) = \begin{cases} n/2, & \forall x \in [-1/n, 1/n], \\ 0, & \forall x \notin [-1/n, 1/n]. \end{cases}$$

The functions of this sequence have graphs corresponding to the graph of the step function shown in Figure 8.

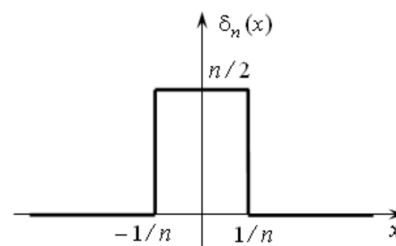


Figure 8. Graph of step function.

It is easy to see that for any  $n$  the area of the figure under the graph of such a step function is equal to one.

For the function, the graph of which is shown in Figure 8, we find the values of the coefficients of the Fourier series on the segment  $[-\pi, \pi]$ :

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{+\pi} \delta_n(x) dx = \frac{1}{\pi} \int_{-1/n}^{1/n} \frac{n}{2} dx = \frac{1}{\pi};$$

$$a_k = 0, \text{ due to the parity of the function;}$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} \delta_n(x) \cos kx dx = \frac{1}{\pi} \int_{-1/n}^{1/n} \frac{n}{2} \cos kx dx = \frac{1}{\pi} \cdot \frac{n}{2} \cdot \frac{2}{n} \cos(kx^*) = \frac{\cos(kx^*)}{\pi},$$

$$x^* \in [-1/n, 1/n],$$

by virtue of the theorem on the mean value of a definite integral.

Equalities with  $\delta(x)$  have to be meant in the sense of limits of sequences distributions.

The equalities with  $\delta(x)$  have to be meant in the sense of limits of sequences distributions. In the theory of generalized functions, the limit of a sequence of distributions is the distribution that sequence approaches. The distance, suitably quantified, to the limiting distribution can be made arbitrarily small by selecting a distribution sufficiently far along the sequence. This notion generalizes a limit of a sequence of functions; a limit as a distribution may exist when a limit of functions does not.



Given a sequence of distributions  $f_n$ , its limit  $f$  is the distribution given by  $f(\varphi) = \lim_{n \rightarrow \infty} f_n(\varphi)$  for each test function  $\varphi$ , provided that distribution exists.

Since the delta function  $\delta(x) = \lim_{n \rightarrow \infty} \delta_n(x)$  and noticing that  $x^* \xrightarrow[n \rightarrow \infty]{} 0$ , we find  $b_k = \frac{1}{\pi}$ .

Consequently, the expansion of the delta function in a Fourier series on the interval  $[-\pi, \pi]$  has the form

$$\delta(x) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{k=1}^{\infty} \cos(kx).$$

For a finite series, we have an approximate relation

$$\delta(x) \approx \frac{1}{2\pi} + \frac{1}{\pi} \sum_{k=1}^n \cos(kx).$$

This approximate equality is only informal because the point value  $\delta(x)$  has no meaning in Sobolev-Schwartz theory. Generalized function can be considered as set-theoretical maps (in a non-Archimedean ring of scalars) [24].

The graph of the approximation of the delta function by the Fourier series is shown in Figure 9.

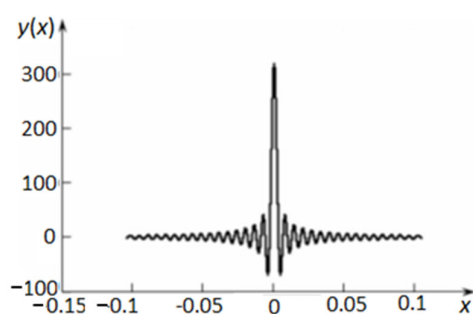


Figure 9. Partial sum plot of Fourier series.

Comparison of the graphs (Figures 7 and 9) shows that even with a significant number of harmonics (in our case  $n = 1000$ ), the approximation error is very large. The minimum value of the constructed approximation is negative and is  $-69.182$ . Moreover, with an infinite increase in the number of terms in the approximating Fourier series, the minimum value of its sum tends to  $-\infty$  (Figure 10), which fully corresponds to the assertion proved in this section about the possible infinitely large error in approximation using the Fourier series.

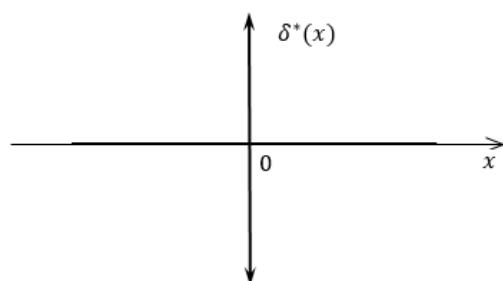
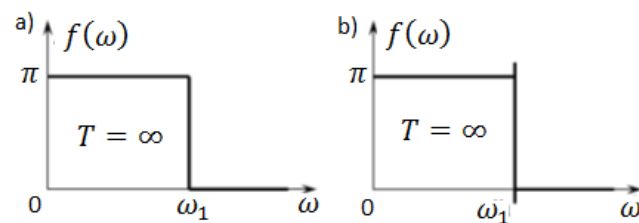


Figure 10. A very intuitive graph of approximation of the delta function by Fourier series.

The existence of the Gibbs effect in the approximation of functions by trigonometric expressions also makes the proof of some important theorems critical. In particular, in the theory of signal transmission, the classical sampling Nyquist-Shannon-Kotelnikov theorem is widely used. When proving the theorem [66] to approximate functions, it uses the so-called integral sine determined by the expression.

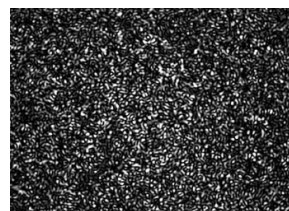
On the basis of the integral sine Kotelnikov V.A. to prove the theorem [66] builds a function  $Si(T(\omega + \omega_1)) - Si(T(\omega - \omega_1))$ , where  $\omega$  is argument,  $T$ ,  $\omega_1$  are some parameters. At the same time, he claims that with increasing  $T$  this function tends to the limits shown in Figure 11a, that is, we quote literally, is equal to zero at  $\omega > \omega_1$  and equal to  $\pi$  at  $\omega < \omega_1$ .



**Figure 11.** Limit function graphs in the classical sampling Nyquist-Shannon-Kotelnikov theorem.

In fact, this is not the case. The graph of the limiting function will have the form shown in Figure 11b. That is, for any, even arbitrarily large but finite values of the parameter  $T$ , there will always be those  $\omega < \omega_1$ , for which the values of the function constructed by Kotelnikov will be different from  $\pi$ , and there will always be those  $\omega > \omega_1$ , for which its values will be different from zero. Moreover, it is important to note that the indicated difference with increasing  $T$  does not tend to zero, but tends to some number other than zero, approximately equal to 0.281, that is, constituting a sufficiently large value. Therefore, the classical sampling Nyquist-Shannon-Kotelnikov theorem requires a careful revision.

In the practice of creating images, the noted errors lead to a speckle effect, which manifests itself in the spotting of such images, their increased graininess (Figure 12). The speckle effect is the result of the interference of many waves of the same frequency, having different phases and amplitudes, which add up to the resulting wave, the amplitude and, therefore, the intensity of which changes randomly. It seems to the viewer that the image is covered with frequent, small spots, which, of course, degrades the quality of these images. Such disadvantages in signal transmission led to signal distortions, which can be significant.



**Figure 12.** Example of speckle effect (<https://bigenc.ru/physics/text/4246597> accessed on 17 July 2022).

The described shortcomings clearly indicate the need to develop new, more efficient methods for approximating piecewise linear functions.

### 3.2. Description of New Methods of Approximation of Piecewise-Linear Functions and Their Convergence

To eliminate the noted shortcomings, S. Aliukov [5,6,11,14,16] proposed new methods for approximating piecewise-linear functions, based, like the Fourier series, on the use of trigonometric expressions, but in the form of recursive functions.

To explain these methods, consider, for example, step Function (1) in more detail. This function is often used for an example of the application of Fourier series and therefore it is convenient to take this function for a comparative analysis of the traditional Fourier series expansion and the proposed method.

The expansion of Function (1) in a Fourier series has all the above-described disadvantages. To eliminate them, it is proposed to approximate the original step function by a sequence of recursive periodic functions

$$\{f_n(x) \mid f_n(x) = \sin((\pi/2) \cdot f_{n-1}(x)), f_1(x) = \sin x; n - 1 \in \mathbb{N}\} \subset C^\infty[-\pi, \pi] \quad (2)$$

The graphs of the original function (thickened line) and its five successive approximations in this case have the form (Figure 13).

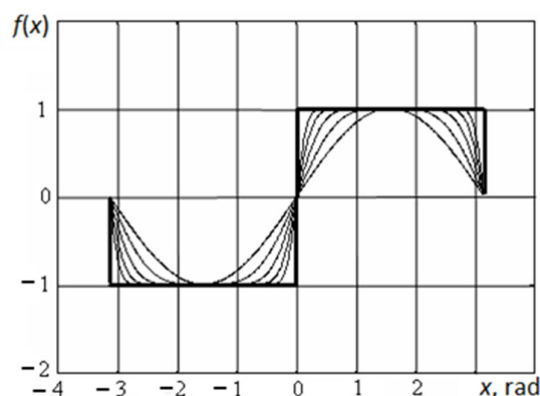


Figure 13. Graphs of the original function and five of its successive approximations.

As you can see, even with relatively small values when using the iterative procedure (2), the graph of the approximating function approximates the original Function (1) quite well. In this case, the approximating functions obtained using the proposed method are free from the drawbacks of expansion in Fourier series. The Gibbs effect is completely absent.

Let us note some features of the proposed approximating iteration procedure.

Note that the functions  $f_n(x)$  and  $f_0(x)$  are odd and periodic with a period  $2\pi$ . The functions  $f_n(x + \pi/2)$  and  $f_0(x + \pi/2)$  are periodic even. Therefore, it is sufficient to consider the sequence of approximating functions (2) on an interval  $[0, \pi/2]$ .

**Theorem 8.** The sequence of functions  $f_n(x)$  converges to the original function  $f_0(x)$ , and the convergence is pointwise, but not uniform.

**Proof.** At points  $x = 0$  and  $x = \pi/2$  we have  $f_n(x) - f_0(x) = 0, \forall n \in \mathbb{N}$ . Therefore, at these points  $f_n(x) \xrightarrow{n \rightarrow \infty} f_0(x)$ .

Since  $\sin x > (2/\pi) \cdot x, \forall x \in (0, \pi/2)$ , then the condition  $f_n(x) = \sin((\pi/2) \cdot f_{n-1}(x)) > f_{n-1}(x) > \dots > f_1(x) > 0$  is satisfied for any  $x \in (0, \pi/2)$ . Then, the sequence  $f_n(x), \forall x \in (0, \pi/2)$  is positive, increasing, and bounded, and therefore has a finite limit, which we denote by  $\lim_{n \rightarrow \infty} f_n(x) = A \in \mathbb{R}$ . We obtain  $A = \lim_{n \rightarrow \infty} \sin((\pi/2) \cdot f_{n-1}(x)) = \sin((\pi/2) \cdot \lim_{n \rightarrow \infty} f_{n-1}(x)) = \sin((\pi/2) \cdot A)$ , whence we find that  $A = 0$  or  $A = 1$ . Since the sequence is positive and increasing, then  $A = 1 = f_0(x)$ . Then, on the considered interval  $f_n(x) \xrightarrow{n \rightarrow \infty} f_0(x)$ . Taking into account the previously made conclusion about the convergence of the sequence at the points  $x = 0$  and  $x = \pi/2$ , we conclude that  $f_n(x) \xrightarrow{n \rightarrow \infty} f_0(x), \forall x \in [0, \pi/2]$ . This convergence is only pointwise and not uniform since the function  $f_0(x)$  is not continuous on the segment  $[0, \pi/2]$ .  $\square$

**Theorem 9.** In the space of measurable functions  $L_1[0, \pi/2]$  and in the Hilbert space  $L_2[0, \pi/2]$  the sequence of approximating functions  $f_n(x)$  converges in the norm to the original function  $f_0(x)$ .

**Proof.** We introduce a sequence of minorants with respect to a sequence  $f_n(x)$  of functions

$$\{\eta_n(x) \mid \eta_n(x) = (2/\pi) \cdot \arctg(n\pi); n \in \mathbb{N}\} \subset C^\infty[0, \pi/2].$$

It can be shown that  $f_n(x) \geq \eta_n(x)$ ,  $\forall n \in \mathbb{N}$ ,  $\forall x \in [0, \pi/2]$ . Note that the measure of the set of discontinuity points of the function  $f_0(x)$  is equal to zero. Then, taking into account the non-negative sign and boundedness of the functions  $f_n(x)$  and  $\eta_n(x)$  and on the segment under consideration, in space  $L_1[0, \pi/2]$  we obtain

$$\|f_0(x) - f_n(x)\| = \int_0^{\pi/2} (1 - f_n(x))dx \leq \int_0^{\pi/2} (1 - \eta_n(x))dx = \frac{\pi}{2} - \operatorname{arctg} \frac{\pi n}{2} + \frac{1}{\pi n} \cdot \ln(1 + (\pi n)^2/4).$$

Since  $\lim_{n \rightarrow \infty} \left( \frac{\pi}{2} - \operatorname{arctg} \frac{\pi n}{2} + \frac{1}{\pi n} \cdot \ln(1 + (\pi n)^2/4) \right) = 0$ , then  $\|f_0(x) - f_n(x)\| \xrightarrow{n \rightarrow \infty} 0$ .

Similarly, one can prove that the sequence  $f_n(x)$  converges in the norm to a function  $f_0(x)$  in the space  $L_2[0, \pi/2]$ .  $\square$

Thus, the sequence of approximating functions  $f_n(x)$  in spaces  $L_1[-\pi, \pi]$  and  $L_2[-\pi, \pi]$  is fundamental. In space  $C[-\pi, \pi]$  the sequence  $f_n(x)$  is not fundamental.

The function  $f_1(x)$  will be called initial (or angular). Instead of sine, we can use another (not necessarily periodic) function as an initial function. Note that when using the iterative procedure (2) and under the condition  $|f_1(x)| < 2$  we obtain  $\lim_{n \rightarrow \infty} f_n(x) = \operatorname{sign}(f_1(x))$ . In this case, any step function can be approximated. Indeed, for the step function

$$f(x) = \begin{cases} h, & x \in (x_1, x_2), \\ 0, & x \notin (x_1, x_2) \end{cases} \quad (3)$$

take the initial function in the form  $f_1(x) = \exp(1 - (ax + b)^2) - 1$ . From the condition  $f_1(x_1) = f_1(x_2) = 0$  we find  $a = 2/(x_1 - x_2)$ ;  $b = (x_1 + x_2)/(x_2 - x_1)$ . For these values of the coefficients  $a$  and  $b$  the sequence

$$\{f_n(x) \mid f_n(x) = (h/2) \cdot (1 + \sin \varphi_n(x)), \varphi_n(x) = (\pi/2) \cdot \sin \varphi_{n-1}, \varphi_1(x) = (\pi/2) \cdot f_1(x), n-1 \in \mathbb{N}\}$$

converges to the step function  $f(x)$ . Then, any step function with values  $h_i$  on the intervals  $(x_{1i}, x_{2i})$  can be approximated by the sum of similar sequences  $\sum_{i=1}^k \{f_n(x)\}_i$ .

The proved Theorem 2 is of a general nature and is valid for an arbitrary step function. Therefore, for example, an arbitrary periodic step function can be represented as a linear combination  $f(x) = \sum_{i=1}^k h_i \cdot f_{0i}(x)$ ,  $h_i \in \mathbb{R}$ , shifted in phase and along the ordinate axis functions  $f_{0i}(x) = \operatorname{sign}(\sin(l_i x - x_i))$ ,  $l_i, x_i \in \mathbb{R}$ . According to the proved theorem, in the spaces  $L_1[-\pi, \pi]$  and  $L_2[-\pi, \pi]$  we have the convergence  $\|f_{0i}(x) - f_{ni}(x)\| \xrightarrow{n \rightarrow \infty} 0, \forall i$ , therefore the function  $f_n(x) = \sum_{i=1}^k h_i \cdot f_{ni}(x)$  converges in the norm to the function  $f(x)$ , since

$$\|f(x) - f_n(x)\| = \left\| \sum_{i=1}^k h_i \cdot f_{0i}(x) - \sum_{i=1}^k h_i \cdot f_{ni}(x) \right\| \leq \sum_{i=1}^k |h_i| \cdot \|f_{0i}(x) - f_{ni}(x)\| \xrightarrow{n \rightarrow \infty} 0.$$

### 3.3. Approximation Error

To estimate the error of approximation (2), we use the relation

$$\varphi_n(x) \leq f_n(x) \leq \psi_n(x)$$

(Figure 14), where

$$\psi_n(x) = (\pi/2)^{n-1} \cdot x, x \in \pi/2, n \in \mathbb{N}.$$

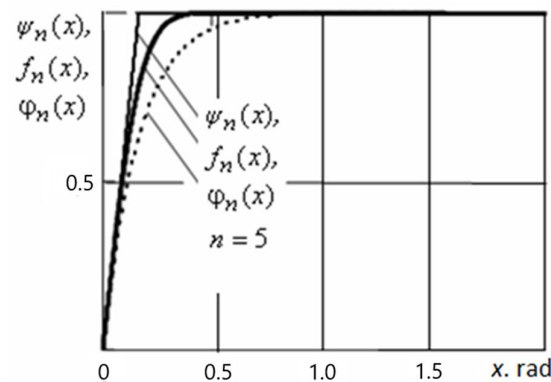


Figure 14. Graphs of limiting functions.

Functions  $\varphi_n(x)$  and  $\psi_n(x)$  are constructed from the condition of equality of derivatives at zero  $\varphi'_n(x) = \psi'_n(x) = f'_n(x)$ , which allows one to obtain a narrow interval for estimating the approximation error.

In space  $L_1[0, \pi/2]$  the estimates for the absolute and relative errors are, respectively

$$\frac{(2/\pi)^{n-1}}{2} \leq \|f_n(x) - f_0(x)\|_{L_1[0, \pi/2]} \leq \left(\frac{2}{\pi}\right)^{n-1} \cdot \left(1 - \exp\left(-\left(\frac{\pi}{2}\right)^n\right)\right);$$

$$\frac{(2/\pi)^n}{2} \leq \frac{\left(\|f_n(x) - f_0(x)\|_{L_1[0, \pi/2]}\right)}{\left(\frac{\pi}{2}\right)} \leq \left(\frac{2}{\pi}\right)^n \cdot \left(1 - \exp\left(-\left(\frac{\pi}{2}\right)^n\right)\right).$$

For space  $L_2[0, \pi/2]$ , these estimates take, respectively, the form

$$\left((2/\pi)^{n-1}/3\right)^{1/2} \leq \|f_n(x) - f_0(x)\|_{L_2[0, \pi/2]} \leq \left(\left(\frac{2}{\pi}\right)^{n-1} \cdot \frac{\left(1 - \exp\left(-2\left(\frac{\pi}{2}\right)^n\right)\right)}{2}\right)^{1/2};$$

$$\left((2/\pi)^n/3\right)^{1/2} \leq \frac{\left(\|f_n(x) - f_0(x)\|_{L_2[0, \pi/2]}\right)}{\left(\frac{\pi}{2}\right)^{1/2}} \leq \left(\left(\frac{2}{\pi}\right)^n \cdot \frac{\left(1 - \exp\left(-2\left(\frac{\pi}{2}\right)^n\right)\right)}{2}\right)^{1/2}.$$

The graphs of the upper and lower estimates of the relative error  $\delta$  depending on  $n \in N$  for the space  $L_1[0, \pi/2]$  (curves 1) and space  $L_2[0, \pi/2]$  (curves 2) are shown in Figure 15.

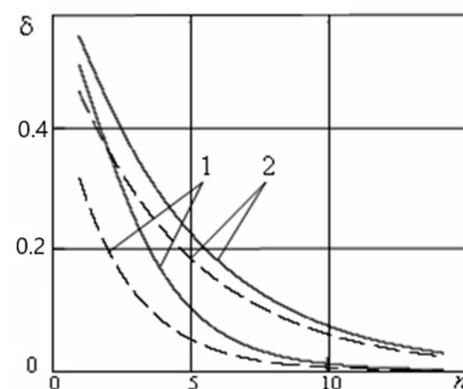


Figure 15. Graphs of estimates of the relative error.

Considering the approximation of the step function  $f(x)$  (3), we assumed that its position and height are precisely known. In real problems, the parameters are usually set approximately. Let, for example, the initial parameters are specified with absolute errors

$$|\hat{x}_1 - x_1| = \Delta x_1 \in [0, \Delta^* x_1), |\hat{x}_2 - x_2| = \Delta x_2 \in [0, \Delta^* x_2), |\hat{h} - h| = \Delta h \in [0, \Delta^* h),$$

where  $\Delta^* x_1 = \sup \Delta x_1$ ,  $\Delta^* x_2 = \sup \Delta x_2$ ,  $\Delta^* h = \sup \Delta h$ ,  $\hat{x}_1$ ,  $\hat{x}_2$ ,  $\hat{h}$  are the approximate values of the parameters. Consider the step function (3) on the segment  $[c, d]$ , for which  $[x_1 - \Delta^* x_1, x_2 + \Delta^* x_2] \subset [c, d]$ . In this case, in spaces  $L_1[c, d]$ ,  $L_2[c, d]$ ,  $M[c, d]$ , where  $M[c, d]$  is the set of functions bounded on an interval  $[c, d]$  with a metric  $\rho(f^{(1)}(x), f^{(2)}(x)) = \sup_{x \in [c, d]} |f^{(1)}(x) - f^{(2)}(x)|$ , for the absolute error of approximation in the norm, we obtain, respectively, the estimates

$$\begin{aligned} \|\Delta f\| &< \sup_{\Delta x_1} \sup_{\Delta x_2} \sup_{\Delta h} \lim_{n \rightarrow \infty} \|f(x) - f_n(x)\|_{L_1[c, d]} = (|h| + \Delta^* h) \cdot (\Delta^* x_1 + \Delta x_2) + (x_2 - x_1) \cdot \Delta^* h; \\ \|\Delta f\| &< \sup_{\Delta x_1} \sup_{\Delta x_2} \sup_{\Delta h} \lim_{n \rightarrow \infty} \|f(x) - f_n(x)\|_{L_2[c, d]} = \sqrt{(|h| + \Delta^* h)^2 \cdot (\Delta^* x_1 + \Delta^* x_2) + (x_2 - x_1) \cdot \Delta^* h^2}; \\ \|\Delta f\| &< \sup_{\Delta x_1} \sup_{\Delta x_2} \sup_{\Delta h} \lim_{n \rightarrow \infty} \|f(x) - f_n(x)\|_{M[c, d]} = |h| + \Delta^* h. \end{aligned}$$

As we can see from the estimates obtained, the approximation error does not accumulate, which is a positive side of the proposed method.

Since in practice, as a rule, we only know the approximate values of the parameters and measurement errors, it is better to express the upper estimates for the absolute approximation error in the form

$$\begin{aligned} \|\Delta f\|_{L_1[c, d]} &< (|\hat{h}| + 2\Delta^* h) \cdot (\Delta^* x_1 + \Delta^* x_2) + (\hat{x}_2 - \hat{x}_1 + \Delta^* x_1 + \Delta^* x_2) \cdot \Delta^* h; \\ \|\Delta f\|_{L_2[c, d]} &< (|\hat{h}| + 2\Delta^* h)^2 \cdot (\Delta^* x_1 + \Delta^* x_2) + (\hat{x}_2 - \hat{x}_1 + \Delta^* x_1 + \Delta^* x_2) \cdot \Delta^* h^2; \\ \|\Delta f\|_{M[c, d]} &< |\hat{h}| + 2\Delta^* h. \end{aligned}$$

Let us return to function (1) and its approximation using sequence (2) in the space of bounded functions  $M[0, \pi]$ .

Let  $\Delta = |f_0(x) - f_n(x)| \in [0, 1]$  be the absolute error of approximation.

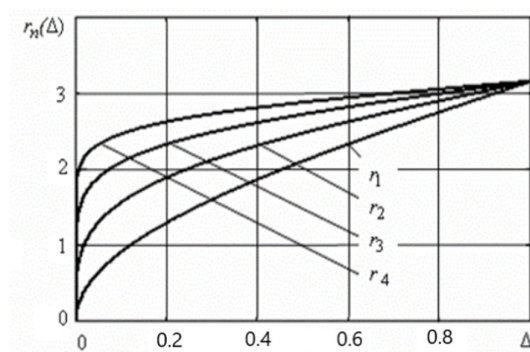
Let's write down a sequence  $\left\{ r_n \mid r_n = \max_{\Delta} \max_{x_1, x_2 \in [0, \pi] : f_n(x_1) = f_n(x_2)} |x_2 - x_1| \right\}$  of maximum metrics. From the equation  $f_n(x) = 1 - \Delta$  we obtain that this sequence can be represented as

$$\{r_n \mid r_n = \pi - 2\arcsin \lambda_n, \lambda_n = (2/\pi) \cdot \arcsin \lambda_{n-1}, \lambda_1 = 1 - \Delta, n-1 \in \mathbb{N}\}.$$

It can be proved similarly to the proof of Theorem 1 that the sequence  $r_n(\Delta) \xrightarrow{n \rightarrow \infty} r^*(\Delta) = \begin{cases} \pi, & \Delta \in (0, 1], \\ 0, & \Delta = 0, \end{cases}$  where the convergence on an interval  $[0, 1]$  is pointwise but not uniform. It is important to note that the sequence  $\{r_n\}$  also converges to a step function.

The graphs of the first few functions in the sequence are shown in Figure 16.





**Figure 16.** The lengths of the gaps with an approximation error not exceeding  $\Delta$ .

As we can see in Figure 16, the length of the gap at which the approximation error does not exceed  $\Delta$ , sharply increases with an increase  $n$  in the region of sufficiently small error values  $\Delta$ . This fact speaks of the fast convergence of the proposed method and is its positive feature.

For a quantitative assessment of the change in the length of this interval, we derive an approximate dependence for the function  $\Delta r(n, \Delta) = r_n - r_{n-1}$ . For this purpose, we use the ratio  $r_n - r_{n-1} = 2(x_{n-1} - x_n)$ , where  $x_n = \arcsin((2/\pi) \cdot x_{n-1})$ ,  $x_1 = \arcsin(1 - \Delta)$ . Then,  $r_n - r_{n-1} = 2(x_{n-1} - \arcsin((2/\pi) \cdot x_{n-1}))$ . Expanding  $\arcsin((2/\pi) \cdot x_{n-1})$  in a Maclaurin series and taking into account the sufficiently small values  $x_{n-1}$ , we obtain approximately  $r_n - r_{n-1} \approx (2/\pi) \cdot (\pi - 2) \cdot x_{n-1}$ . Then,  $r_n - r_{n-1} \approx (2/\pi)^{n-1} \cdot (\pi - 2) \cdot \arcsin(1 - \Delta)$ .

Let us indicate some properties of the proposed approximation (2).

**Property 1.** The maximum value of the difference in the lengths of the intervals  $r_n - r_{n-1}$  does not depend on  $n$  and is found by the ratio:

$$\max_{\Delta \in [0, 1]} (r_n - r_{n-1}) = \sqrt{\pi^2 - 4} - 2\arcsin\sqrt{1 - 4/\pi^2}, \quad n-1 \in \mathbb{N}.$$

**Proof.** Based on the previously obtained relation  $r_n - r_{n-1} = 2(x_{n-1} - \arcsin((2/\pi) \cdot x_{n-1}))$ ,  $n-1 \in \mathbb{N}$ , we find the derivative

$$\frac{d(r_n - r_{n-1})}{d\Delta} = -2^{n-1} \cdot \left( \sqrt{\pi^2 - 4x_{n-1}^2} - 2 \right) / \sqrt{\prod_{i=1}^{n-1} (\pi^2 - 4x_i^2)(1 - (1 - \Delta)^2)}.$$

The points  $x_{n-1} = x_{n-2} = \dots = x_1 = \pi/2$  are the minimum points at which  $r_n - r_{n-1} = 0$ . In case  $\Delta = 1$  we also have  $r_n - r_{n-1} = 0$ . The points  $x_{n-1} = \sqrt{(\pi^2/4) - 1}$  are maximum points and are independent of  $n$ . Then, we obtain that

$$\max_{\Delta \in [0, 1]} (r_n - r_{n-1}) = \sqrt{\pi^2 - 4} - 2\arcsin\sqrt{1 - 4/\pi^2}.$$

□

For reference, we point out that  $\max_{\Delta \in [0, 1]} (r_n - r_{n-1}) \approx 0.661$ .

**Property 2.** The maximum value of the difference between the values of the functions  $f_n(x) - f_{n-1}(x)$  does not depend on  $n$  and is found by the ratio:

$$\max_{x \in [0, \pi]} (f_n(x) - f_{n-1}(x)) = (\sqrt{\pi^2 - 4} - 2\arccos(2/\pi))/\pi, \quad n-1 \in \mathbb{N}.$$

The proof is similar to the proof of Property 1.

Moreover, for reference, we point out that  $\max_{x \in [0, \pi]} (f_n(x) - f_{n-1}(x)) \approx 0.211$ .

Property 2 shows that the sequence of approximating functions  $f_n(x)$  (2) does not converge in the Cauchy sense, that is, it is not fundamental, since  $\exists \varepsilon > 0 \forall n^* \in \mathbb{N} \exists n, m > n^*$ , that  $\max_{x \in [0, \pi]} |f_n(x) - f_m(x)| > \varepsilon$ . As  $\varepsilon$  you can take, for example, the number 0,1, setting  $m = n^* + 1$ ,  $n = n^* + 2$ .

The obtained relations can be used to estimate the approximation error in solving applied problems.

### 3.4. Generalized Functions and Their Approximation by a Sequence of Recursive Functions

Generalized functions [67] became widespread in the 20th century, when new problems in physics and mathematics led to an urgent need to expand the definition of a function.

Let  $Y$  be a linear space whose elements are functions in the sense of the usual definition.

If there is a rule according to which a certain number is assigned to each function  $y \in Y$ , then it is said that on the set  $Y$  there is a given functional. We denote the functional by  $I : Y \rightarrow R$ , or more simply  $I(y)$ .

A functional is called linear if the condition  $I(\alpha y_1 + \beta y_2) = \alpha I(y_1) + \beta I(y_2)$ ,  $\forall y_1, y_2 \in Y, \forall \alpha, \beta \in R$  holds.

A functional is called continuous if the condition  $y_n \Rightarrow y^*$  implies the fulfillment of the condition  $I(y_n) \rightarrow I(y^*)$ ,  $\forall y_n, y^* \in Y$ .

We will consider functions on the set  $R$ .

A function  $\varphi(x)$  is said to be compactly supported if it equals zero outside a finite interval  $[a, b]$ , and the boundaries of the interval depend on  $\varphi(x)$ . Any continuous compactly supported function is called basic. The set of basic functions will be denoted by  $C_0$ .

Let the function  $f(x)$  be ordinary in the sense of the definition, and it is continuous except, perhaps, a finite number of discontinuity points, and bounded on any finite interval.

We define the functional by the integral  $I(\varphi) = \int_{-\infty}^{+\infty} f(x)\varphi(x)dx$ , which for any basic function  $\varphi(x)$  will be finite. A functional of this kind is called a regular functional.

**Definition 12.** A linear functional  $\xi : C_c^\infty(\mathbb{R}) \rightarrow \mathbb{R}$  is continuous, if for every convergent sequence  $\{f_n\}_{n=1}^\infty$  of functions  $f_n \in C_c^\infty(\mathbb{R})$  we have  $\lim_{n \rightarrow \infty} \langle \xi, f_n \rangle = \langle \xi, \lim_{n \rightarrow \infty} f_n \rangle$ . We use the notation  $\langle \xi, f \rangle$  instead of  $\xi(f)$ .

**Definition 13.** A generalized function is any linear continuous functional  $I(\varphi)$ , defined on the set  $C_0$ , having the properties

1.  $I(\alpha \varphi_1 + \beta \varphi_2) = \alpha I(\varphi_1) + \beta I(\varphi_2)$ ,  $\forall \varphi_1, \varphi_2 \in C_0, \forall \alpha, \beta \in R$ ;
2.  $I(\varphi_n) \rightarrow I(\varphi)$ , if  $\varphi_n \Rightarrow \varphi$  in  $C_0$ .

Not every generic function is regular. A generic function that cannot be represented by the integral  $I(\varphi) = \int_{-\infty}^{+\infty} f(x)\varphi(x)dx$ , is called singular. An example of a singular generalized function is the function  $I(\varphi) = \varphi(0)$ . This function is called the  $\delta$ -function or the Dirac function.

Using the proposed methods, one can also approximate singular generalized functions, for example, a  $\delta$ -function.

The meaning of singular generalized functions can be understood based on their approximations, perceiving the generalized function as the limit of some approximating sequence of ordinary functions. For example, as noted, the delta function can be viewed as the limit of a sequence of step functions (Figure 8). However, the use of a sequence of step functions does not allow adequate representation of the derivatives of the delta function, which, in turn, are also generalized functions. The problem is that step functions have discontinuity points at which they are not mathematically differentiable. Therefore,

to represent the derivatives of a delta function, it is necessary to use an approximating sequence of analytic functions with derivatives of any order.

The expression used for approximation in this case can be of the form [11]

$$f(x) = \cos(A(A(\dots A(x)))), \text{ where } A(x) = \frac{\pi}{2} \sin(x).$$

In particular, Figure 17 shows the graph of the function

$$f(x) = 310 \cos(A(x)))))))))))))))))).$$

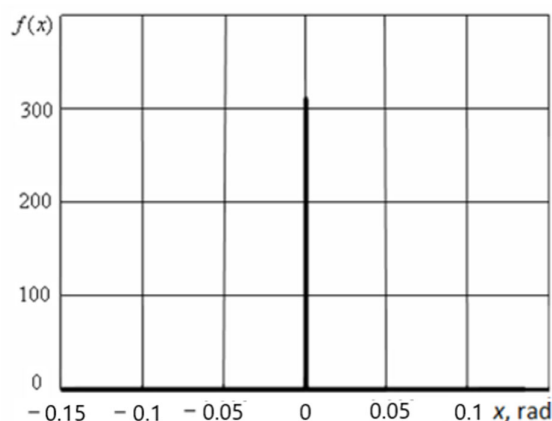


Figure 17. Graph of approximation  $\delta$ -functions.

Comparing the graphs in Figures 7, 9 and 17, we note that the proposed approximation methods give a much more accurate approximation of the  $\delta$ -function than the Fourier series. Moreover, the accuracy of the approximation can be increased to an arbitrarily large degree by increasing the number of nested functions. The height of the approximation peak (amplitude) can be determined by the integral condition in the definition of the  $\delta$ -function.

To determine the height of the approximation peak, we use the fact that the  $\delta$ -function is the derivative of the Heaviside function, or the unit jump function, which is defined as

$$H(x) = \begin{cases} 1, & \forall x > 0; \\ 0, & \forall x < 0. \end{cases}$$

The Heaviside function can be approximated by a sequence of functions of the form  $H_n(x) = 0.5(1 + f_n(x))$ , where the sequence of functions  $f_n(x)$  is defined by relation (2) and is considered on the interval  $[-\pi/2, \pi/2]$ .

For example, Figure 18 shows graphs of three successive approximations

$$\begin{aligned} H_9(x) &= 0.5(1 + \sin(A(A(A(A(A(A(A(A(A(A(A(x))))))))))), \\ H_{10}(x) &= 0.5(1 + \sin(A(A(A(A(A(A(A(A(A(A(A(A(A(A(A(x))))))))))), \\ H_{11}(x) &= 0.5(1 + \sin(A(A(A(A(A(A(A(A(A(A(A(A(A(A(A(A(A(A(A(x))))))))))))), \end{aligned}$$

where  $A(x) = \frac{\pi}{2} \sin x$ .

The thickness of the graph in Figure 18 increases as the number of the approximating dependence increases.

Finding the first derivatives of the approximations of the Heaviside function, we obtain successive approximations  $\frac{dH_9(x)}{dx}$ ,  $\frac{dH_{10}(x)}{dx}$  and  $\frac{dH_{11}(x)}{dx}$  for the delta function. Their graphs are shown in Figure 19.

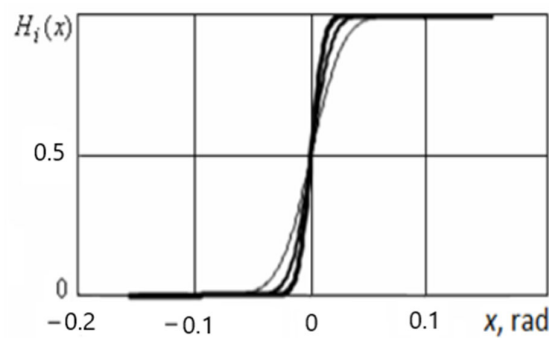


Figure 18. Graphs of approximations of the Heaviside function.

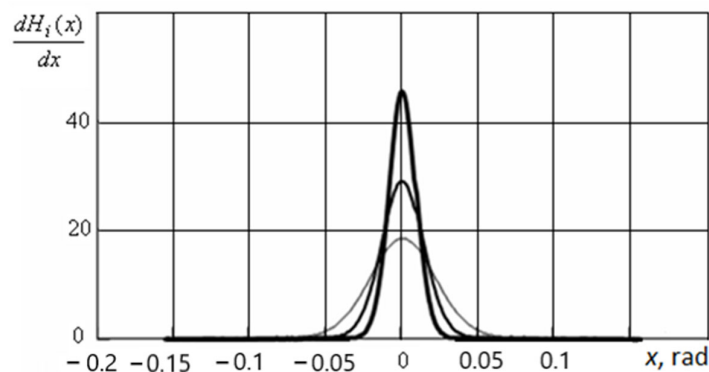


Figure 19. Graphs of approximations  $\delta$ -functions.

With a sufficiently large number of nested functions, we obtain an approximating function  $\frac{dH_{18}(x)}{dx}$ , the graph of which was obtained using the MathCAD computer program and is shown in Figure 17.

Differentiating the approximating functions of the considered sequence  $H_n(x) = 0.5(1 + f_n(x))$ , we obtain

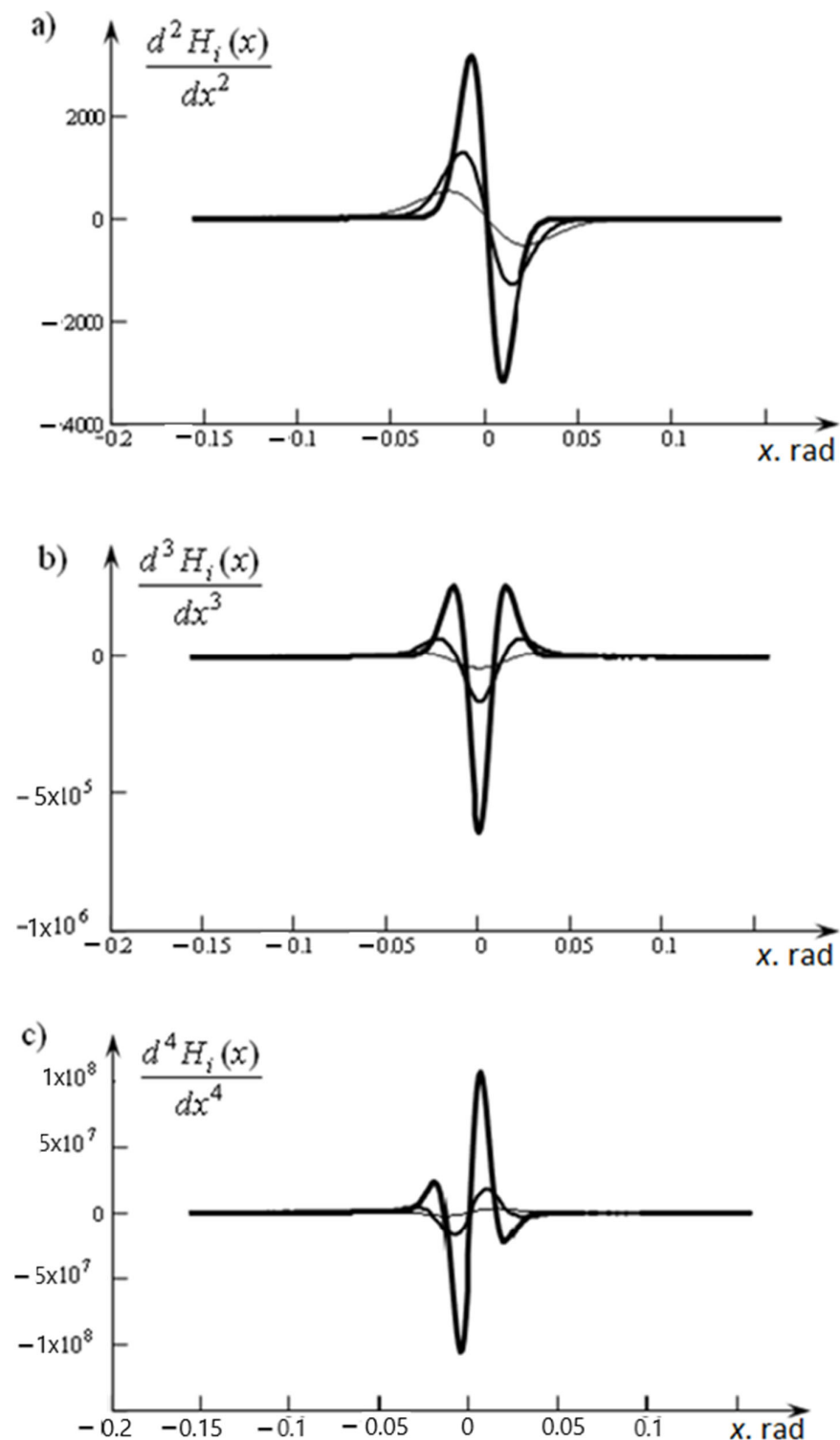
$$\frac{dH_n(x)}{dx} = \frac{\pi^{n-1}}{2^n} \prod_{k=1}^{n-1} \cos\left(\frac{\pi}{2} f_k(x)\right) \cdot \cos x.$$

Substituting into the resulting expression for the derivatives  $x = 0$ , taking into account the parity of the  $\delta$ -function, we find the value for the peak height  $A_n$  of the approximating functions  $H_n(x)$

$$A_n = \frac{\pi^{n-1}}{2^n}$$

### 3.5. Approximation of Derivatives of Generalized Functions: Comparison of Approximation Methods

Since we approximated the generalized functions with analytical functions, we can differentiate these approximating functions and find them to obtain approximations of the derivatives of the generalized functions with any degree of accuracy. For example, similarly to how it was conducted in the previous section, we can build graphs of approximations of the derivatives of the  $\delta$ -function. Figure 20 shows graphs of successive approximations of the first, second and third derivatives of the  $\delta$ -function.



**Figure 20.** Graphs of approximations of the derivatives of the  $\delta$ -function.

Derivatives of higher orders can be found in the same way. The plotted graphs give a good idea of the behavior of the derivatives of the  $\delta$ -function. By mentally increasing the number of the approximating function [11], according to the graphs (Figure 20), it is possible to continue the traced tendencies of changes in the approximations and to present the limiting positions of the sequences of functions approximating the derivatives of the  $\delta$ -function. This will help to improve the understanding of generalized functions that are derivatives of the  $\delta$ -function, to use them not just as an abstract mathematical apparatus,

but to consciously understand their structure, even if they are written in limiting form. This approach can also be used to better understand other generic functions and their behavior.

The  $\delta$ -function can also be approximated by other continuously differentiable functions, for example, such

$$\begin{aligned}\delta(x, \alpha) &= \frac{\alpha}{\pi(\alpha^2 x^2 + 1)}, \alpha \rightarrow \infty, \\ \delta(x, \alpha) &= \frac{\alpha}{\sqrt{\pi}} \exp(-\alpha^2 x^2), \alpha \rightarrow \infty, \\ \delta(x, \alpha) &= \frac{\alpha}{\pi} \cdot \frac{\sin(\alpha x)}{\alpha x}, \alpha \rightarrow \infty,\end{aligned}$$

for which  $\lim_{\alpha \rightarrow \infty} \delta(x, \alpha) = 0 (x \neq 0)$  and  $\lim_{\alpha \rightarrow \infty} \int_{-\infty}^{+\infty} \delta(x, \alpha) dx = 1$ .

The disadvantage of approximating the  $\delta$ -function using the third of these functions is a big deviation from the  $\delta$ -function since this function has not only positive, but also negative values. The graphs of such a function correspond to the graph shown in Figure 10. Moreover, the sequence of negative values is not limited from below, that is, the error can be arbitrarily large.

As for the approximation using the first two functions, they allow for approximating the periodic delta function only as a sum  $\delta(x) = \sum_{k=-\infty}^{+\infty} \delta(x - 2\pi k)$ , which can be inconvenient for practical use, while the approximating functions according to the proposed method are periodic in nature and allow approximating the periodic delta function without any additional constructions. By a periodic delta function, we mean a generalized function whose argument values, in which the function is not equal to zero, repeat periodically.

An example is the graph of the function

$$f(x) = \frac{\pi^{13}}{2^{14}} \cos(A(A(A(A(A(A(A(A(A(A(A(A(A(x)))))))))))))),$$

where  $A(x) = \frac{\pi}{2} \sin(x)$ , is shown in Figure 21.

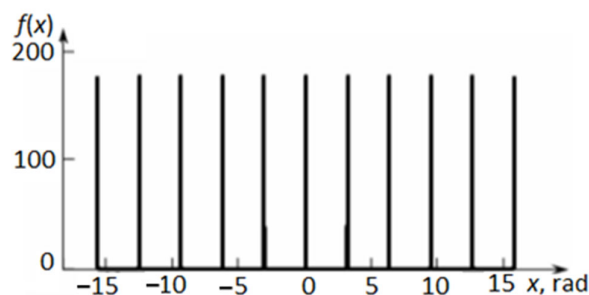


Figure 21. Graph of the function that approximates a periodic  $\Delta$ -function.

The constructed function  $f(x)$  can be used to approximate the distribution function of a discrete random variable using the relation  $F(x) = P \int_{t_0}^x f(x) dx$ , where  $P$  is a parameter determined from the properties of the distribution function. An example of a distribution function constructed in this way is shown in Figure 22 [5,6,14].



Figure 22. An example of approximation of the distribution function.

These techniques can be used in the context of mathematical models represented by ODEs or PDEs. For example, in the case of systems with a variable structure, mathematical models are often presented in the form of several systems of differential equations in sections. In this case, problems arise in constructing a solution to equations during a cycle, periodic solutions, the need to track the transitions of the system from section to section and coordinate solutions at the boundaries of sections. Similar problems arise when solving differential equations with piecewise linear and impulse characteristics. The use of the developed approximation methods makes it possible to overcome these problems. In particular, the authors applied the developed methods in studies, the results of which are published in the following articles [6–9,13,68–70] and others.

#### 4. Practical Application and Examples of Using the Developed Approximation Methods

This section of the review discusses various possibilities of using the developed methods for approximating piecewise linear and generalized functions. The examples given are characterized by specific content, a more complete reflection of physical reality. In reality, even fast-flowing processes occur, albeit in a short, but not zero time. For example, in reality there cannot be an instantaneous change in the speed of a material object with a non-zero mass, since such an instantaneous change would require an infinite amount of energy.

##### 4.1. Application of New Methods of Approximation in Problems of Structural Mechanics

The proposed approximation methods, similarly to the Fourier series, are universal. They can be applied in a wide variety of fields of science and technology. One of the many possible examples of the use of these methods in practice is their application in problems of resistance of materials in the broad sense of this term and in problems of structural mechanics in particular.

When calculating building structures such as trusses, beams, etc. it is necessary to take into account the action of distributed and concentrated loads [71]. Loads, as a rule, have piecewise linear characteristics and Fourier series are widely used for their approximation.

Consider, for example, a beam on which a uniformly distributed load with an intensity  $p$  acts on a section of length  $2c$  (Figure 23). Other parameters are directly given in Figure 23.

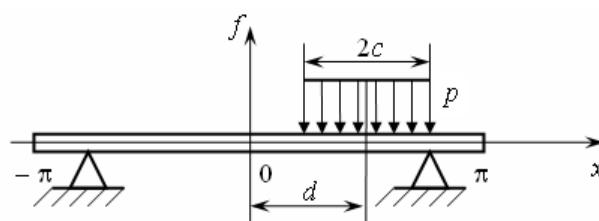
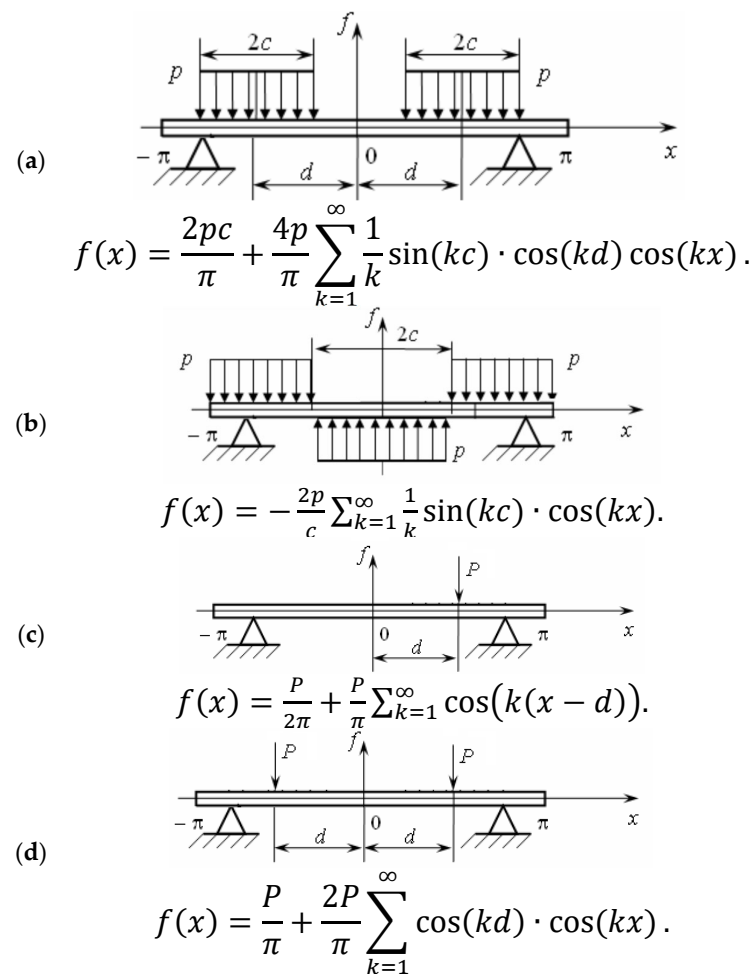


Figure 23. Example of a uniformly distributed load.

By transforming the argument, without losing the generality of reasoning, we can always consider the load function  $f = f(x)$  only on the interval  $[-\pi, \pi]$ . In this case, the expansion of the load in a Fourier series in our case has the form

$$f(x) = \frac{pc}{\pi} + \frac{2p}{\pi} \sum_{k=1}^{\infty} \frac{1}{k} \sin(kc) \cdot \cos(k(x-d)).$$

Figure 24 shows some other types of loads and their representations in the form of Fourier series.



**Figure 24.** Variants of loads and their approximation by Fourier series.

Similarly, if necessary, you can write down the corresponding Fourier series for other symmetric and asymmetric types of loading of building structures.

The decomposition of loads (acting forces) in Fourier series will have all the disadvantages noted in Section 2.1. Therefore, to eliminate these disadvantages, one can use the proposed approximation methods.

For the load shown in Figure 24a, for  $c = \pi/6$ ,  $d = \pi/2$ ,  $p = 4$  the Fourier series expansion will have the form

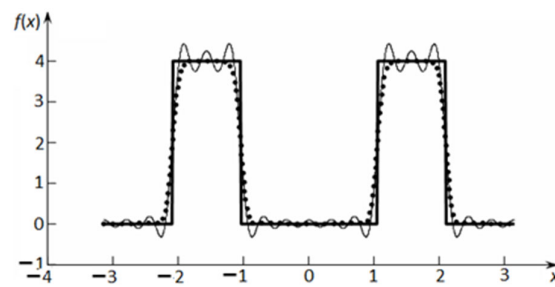
$$f(x) = \frac{4}{3} + \frac{\pi}{16} \sum_{k=1}^{\infty} \sin\left(k\frac{\pi}{6}\right) \cos\left(k\frac{\pi}{6}\right) \cos(kx).$$

Approximating function by the proposed method is found by the expression

$$f(x) = 2 - 2 \sin\left(\frac{\pi}{2} \sin\left(\frac{\pi}{2} \sin(\dots \sin(0.5 + \cos 2x))\right)\right).$$

Figure 25 shows the graphs of the original function and its approximating functions. Here, the graph of the original function is highlighted with a thickened stepped line. The graph of the approximating function, constructed using the Fourier series for  $k = 1, 2, \dots, 17$ , is shown by a thin line. The dotted line corresponds to the graph of the approximating function constructed by the proposed method for four nestings.



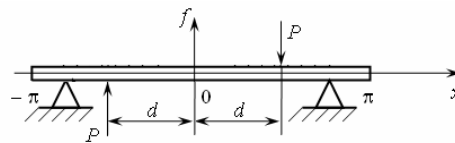


**Figure 25.** Graphs of the initial load and approximating functions.

From Figure 25 we see that, despite the fact that for the expansion in the Fourier series we used seventeen harmonic terms ( $k = 17$ ), and in the approximation by the proposed method we used only four compositions, the last approximating function approximates the original function much better. Moreover, one could graphically show that a further, arbitrarily large increase in the number of harmonic terms in the Fourier series does not save the situation. The reason for this lies in the aforementioned Gibbs effect.

The advantages of approximating the load using the proposed methods are clearly manifested for concentrated actions on a beam. For example, for the case shown in Figure 26 under the action of two equal in magnitude, but oppositely directed forces applied at symmetrically located points, the expansion in a Fourier series has the form

$$f(x) = \frac{2P}{\pi} \sum_{k=1}^{\infty} \sin(kd) \cdot \sin(kx).$$

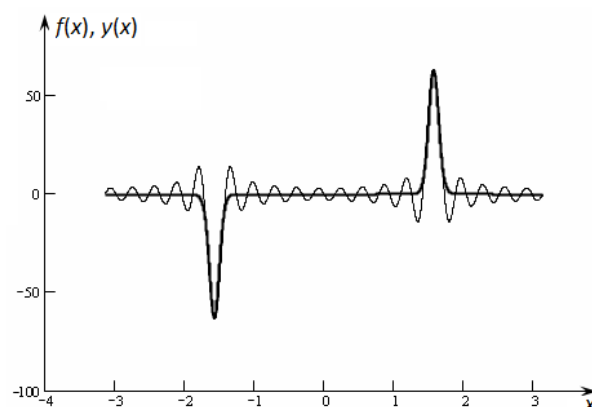


**Figure 26.** Scheme of action of concentrated forces on a beam.

For  $d = \pi/2$ ,  $P = 10$  we approximate this loading scheme by the finite sum of the Fourier series  $f(x) = \frac{2P}{\pi} \sum_{k=1}^{20} \sin(kd) \cdot \sin(kx)$  and function

$$y(x) = \cos\left(\frac{\pi}{2} \sin\left(\frac{\pi}{2} \sin\left(\frac{\pi}{2} \sin\left(\frac{\pi}{2} \sin\left(\frac{\pi}{2} \cos(x)\right)\right)\right)\right)\right) \cdot 40x.$$

Figure 27 shows the graphs of the obtained approximating functions. The thin line corresponds to the approximation using the Fourier series, the thickened line corresponds to the approximation by the proposed method. As you can see, the results of the approximation clearly speak in favor of the proposed method. We have proposed an approach that allows us to model the load using an analytical function that allows finding derivatives of any order. This function is a single analytical expression for the entire domain of definition and is not composite across sections. This approach can describe a given design load circuit with any degree of accuracy.



**Figure 27.** Graphs of approximating functions.

#### 4.2. Application of New Approximation Methods for Modeling Diffusion Processes in Semiconductor Materials

Based on the developed methods for approximating piecewise linear and generalized functions, modified models of diffusion processes in semiconductor materials were created [70,72–76]. These models make it possible to provide a more accurate description of diffusion processes in comparison with the known models.

In the modified models, the possibilities of using a new method of continuous approximation of step functions based on the use of trigonometric expressions in the form of recursive functions are considered. Calculations are performed for the classical model of diffusion of minority charge carriers generated by a wide electron beam in a two-layer semiconductor material.

To quantitatively describe the phenomenon of diffusion of nonequilibrium minority charge carriers generated in a semiconductor by an external energy action, the following two approaches are usually used:

- (1) a model of collective motion of minority charge carriers [77,78], according to which diffusion of nonequilibrium minority charge carriers from any microvolume of a semiconductor is influenced by other electrons or holes from other microregions of the material. Mathematically, this is expressed in the fact that the differential diffusion equation as a function of generation of minority charge carriers (usually written in the right-hand side of the differential equation) includes a function that describes the dependence on the coordinates of the density of minority charge carriers generated per unit time in the target. This model is successfully used to quantitatively describe the diffusion processes of minority charge carriers generated by a wide electron beam in homogeneous semiconductors, for which the right-hand side of the differential equation is a continuous function of coordinates. The use of a wide electron beam makes it possible to neglect the edge effects and to solve the one-dimensional problem of heat and mass transfer;
- (2) the model of independent sources, according to which the diffusion of nonequilibrium minority charge carriers from any microvolume of the semiconductor is not influenced by other electrons or holes from other microregions of the material. Mathematically, this is expressed in the fact that first the diffusion equation is solved for each of the point sources of minority charge carriers, after which, by integrating over the volume occupied by the sources of minority charge carriers, their concentration in the semiconductor is found as a result of their diffusion. The idea of this approach is borrowed from classical work [79]. This model was previously used to quantitatively describe the processes of one-dimensional diffusion of minority charge carriers generated by a wide electron beam in inhomogeneous and multilayer planar structures, for which the distribution of electrophysical parameters of materials over depth has break points of the first kind [80,81].

A modification of the first model was proposed, allowing it to be used to simulate the diffusion of minority charge carriers in a two-layer material. The possibility of using this model to solve such a problem appears if, instead of piecewise constant coefficients (electrophysical parameters) of the differential equation of diffusion of minority charge carriers we use their new approximations based on trigonometric expressions in the form of recursive functions [5]. Note that the approximating functions are continuous and analytical, and therefore, at the boundary of the layers, they correspond to a greater extent than step functions to the dependence of the values of real electrophysical parameters on the coordinate [82].

Within the framework of the considered mathematical model, in the case of one-dimensional diffusion into the final semiconductor, the concentration of minority charge carriers in depth is found as a solution to the differential equation [68,70]

$$D(z) \frac{d^2 \Delta p(z)}{dz^2} - \frac{\Delta p(z)}{\tau(z)} = -\rho(z)$$

with boundary conditions

$$D_1 \frac{d\Delta p(z)}{dz} \Big|_{z=0} = v_{s1} \Delta p(0), \Delta p(l) = 0.$$

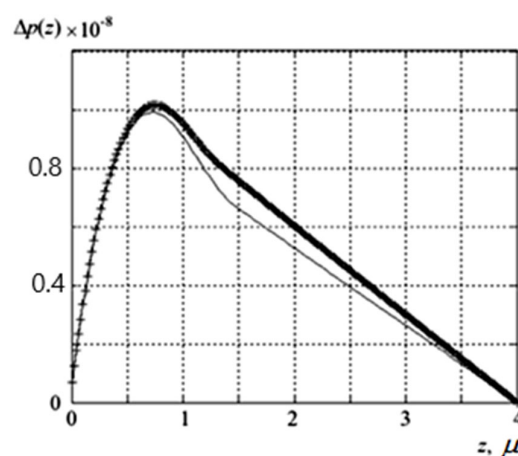
In the modified models in the equation, instead of piecewise-constant coefficients (electrophysical parameters), their new approximations are used, based on the use of trigonometric expressions in the form of recursive functions [5]:

$$\left\{ f_n(z) \mid f_n(z) = \left( \frac{H}{2} \right) (1 + \sin(\varphi_n(z))), \varphi_n(z) = \left( \frac{\pi}{2} \right) \sin(\varphi_{n-1}(z)), \right. \\ \left. \varphi_1(z) = \left( \frac{\pi}{2} \right) f_1(z), n-1 \in N \right\}.$$

As an initial function, a function of the form is taken

$$f_1(z) = \exp(1 - (az + b)^2) - 1.$$

Figure 28 shows the results of calculations carried out using the mathematical package Matlab (MathWorks, Inc., Natick, MA, USA) for the parameters characteristic of the semiconductor structure “epitaxial GaAs film—single-crystal GaAs substrate”.



**Figure 28.** The concentration of minority charge carriers obtained using  $n = 5$  recursive trigonometric functions (the graph is marked with crosses), and using  $n = 11$  functions (the graph is marked with a solid line), as well as the concentration of minority charge carriers, calculated accurately analytically using piecewise constant coefficients (electro-physical parameters) of the differential equation of diffusion of minority charge carriers (the graph is marked with a dashed line).

The following estimates were obtained for the relative error between the exact analytical solution of the problem and the numerical one for  $n = 5$ :

$$\Delta = \frac{\max_{0 \leq i \leq n} |\delta p(z_i) - \delta p_i|}{\max_{0 \leq i \leq n} |\delta p(z_i)|} \cdot 100\% = 9.66\%$$

and for  $n = 11$ :  $\Delta = 0.37\%$ .

Thus, for  $n = 5$ , the influence of the approximations on the simulation result is visible, and for  $n = 11$ , the error in the results is rather small, which indicates the convergence of the approximating procedure used.

The authors of the modified model note [68,70] that the described model allows, with a relatively small number of recursive functions (up to 11), to estimate the concentration of minority charge carriers generated by a wide electron beam in a semiconductor target with an accuracy sufficient for practical use. The model subsequently makes it possible to relatively easily take into account the features of real semiconductor structures (the number and nature of the layers, the space charge region, possibly the energy distribution of electrons that occurs during the interaction of the primary beam with the target, etc.), which makes it promising for quantitative descriptions of the processes of one-dimensional diffusion of minority charge carriers in inhomogeneous and multilayer planar structures.

Note that modified models have been developed for various other examples describing diffusion processes in semiconductor technology [69,70,72–76].

## 5. Conclusions

As the review has shown, the developed methods for approximating periodic and non-periodic piecewise-linear and generalized functions are undoubtedly promising and have a number of advantages. These methods are characterized by fast convergence and low approximation errors. Similarly to the Fourier series, these methods can be based on the use of well-studied trigonometric functions that have good implementation in applied computer programs. While retaining the positive qualities of the Fourier series in this respect, the new methods are devoid of their drawbacks and can be widely used in solving applied problems. The developed methods are characterized by the complete absence of negative consequences of the Gibbs effect. There is also no wavelike character of the approximation on straight sections of the original piecewise-linear function, even with a small number of nested functions used for approximation.

The proved theorems and properties concerning the convergence and error of the developed approximation methods have confirmed all of the positive aspects of these methods.

The developed methods are illustrated by a large number of theoretical and practical examples taken from a wide variety of areas, so we can confidently speak about the universality of these methods. Despite the wide variety of examples given, these examples by no means exhaust all possible areas of application of new methods. The developed methods can find wide application in the field of signal transmission, the theory of automatic control and regulation systems, mathematical models of technical systems of variable structure, systems with discontinuous characteristics, systems with distributed and concentrated loads and influences, the theory of quantum and mathematical physics, and many other areas. These methods make it possible to find analytical functions for approximating singular generalized functions, and therefore, using differential calculus methods, construct approximations for the derivatives of generalized functions, thereby helping to achieve a better understanding of the meaning of generalized functions when applied to real applied problems. In other words, the developed methods also perform an epistemological function.

Numerical tests carried out on the basis of various practical applications have convincingly shown the correctness of the theoretical studies and the assumptions made.

Note also that the proposed approximating functions are continuous and analytical and even more than step functions correspond to real processes, since in reality even jump-like processes occur, albeit for short, but not zero time intervals. So, for example, an

instantaneous change in the speed of a material object requires infinite energy, which is impossible to implement in practice. Concentrated impact in reality does not occur at a point, but is a distributed impact in some small neighborhoods of this point. These realities are fully consistent with the approximating functions obtained using the considered methods based on recursive sequences. In addition, we note that the widespread approximating functions based on splines, for example, are not analytical. Smoothing abrupt changes in the function, the proposed approximation methods bring mathematical models closer to reality, contributing to a deeper understanding of the laws of the world around us.

**Author Contributions:** Conceptualization, S.A.; Methodology, S.A.; Validation, S.A. and K.O.; Formal Analysis, S.A.; Investigation, S.A.; Resources, S.A. and A.A.; Data Collection, S.A.; Writing-Original Draft Preparation, S.A.; Writing-Review & Editing, A.A.; Visualization, S.A.; Supervision, S.A.; Project Administration, S.A., K.O. and A.A.; Funding Acquisition, S.A. and A.A. All of the authors contributed significantly to the completion of this review, conceiving and designing the review, writing and improving the paper. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** There are no data applicable in this study.

**Acknowledgments:** The authors thank South Ural State University (SUSU) for supporting.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. DeVore, R.A.; Lorentz, G.G. *Constructive Approximation*; Springer-Verlag: Berlin, Germany, 1993.
2. Lorentz, G.G.; von Golitschek, M.; Makovoz, Y. *Constructive Approximation: Advanced Problems*; Springer: Berlin, Germany, 1996.
3. Mastroianni, G.; Milovanović, G.V. *Interpolation Processes. Basic Theory and Applications*; Springer-Verlag: Berlin, Germany, 2008.
4. Von Neumann, J.; Taub, A.H. (Eds.) *Theory of Games, Astrophysics, HydroDynamics and Meteorology*. In *Method in the Physical Sciences*; Collected Works Volume VI; Pergamon Press: Oxford, UK, 1961.
5. Alyukov, S.V. Approximation of step functions in problems of mathematical modeling. *Math. Modeling J. Russ. Acad. Sci.* **2011**, *23*, 75–88. [\[CrossRef\]](#)
6. Alyukov, S.V. Modeling of dynamic processes with piecewise-linear characteristics. *Proceedings of universities. Appl. Nonlinear Dyn.* **2011**, *19*, 27–34.
7. Alyukov, S.V. Improvement of mathematical models of inertial torque transformers. *Bull. Mech. Eng.* **2010**, *7*, 3–10.
8. Alyukov, S.V. Improved Models of Inertial Torque Transformers. *Russ. Eng. Res.* **2010**, *30*, 4–15. [\[CrossRef\]](#)
9. Alyukov, S.V. Relay type freewheel mechanisms. *Heavy Mech. Eng.* **2010**, *12*, 34–37.
10. Dubrovskaya, O.A.; Dubrovsky, S.A.; Dubrovsky, A.F.; Alyukov, S.V. On the analytical representation of the elastic-dissipative characteristics of the car suspension. *Bull. Sib. State Automob. Acad.* **2010**, *16*, 23–26.
11. Alyukov, S.V. Approximation of generalized functions and their derivatives. *Questions of atomic science and technology. Ser. Math. Model Phys. Process.* **2013**, *2*, 57–62.
12. Osintsev, K.V.; Alyukov, S.V. Mathematical modeling of discontinuous gas-dynamic flows using a new approximation method. *Materials Science. Energy* **2020**, *26*, 41–55.
13. Alyukov, S.V. *Dynamics of Inertial Continuously Variable Automatic Transmissions*; INFRA-M: Moscow, Russia, 2020; p. 251.
14. Alabugin, A.; Aliukov, S.; Khudyakova, T. Models and Methods of Formation of the Foresight-Controlling Mechanism. *Sustainability* **2022**, *14*, 9899. [\[CrossRef\]](#)
15. Alyukov, S.V. Scientific Basis of Inertial Continuous Transmissions of Increased Loading Capacity. Ph.D. Thesis, South Ural State University, Chelyabinsk, Russia, 2014; p. 369.
16. Aliukov, S.; Buleca, J. Comparative Multidimensional Analysis of the Current State of European Economies Based on the Complex of Macroeconomic Indicators. *Mathematics* **2022**, *10*, 847. [\[CrossRef\]](#)
17. Alyukov, S.V. Approximation of electrocardiograms with help of new mathematical methods. *Comput. Math. Model.* **2018**, *29*, 59–70. [\[CrossRef\]](#)
18. Nikitin, A.V.; Shishlakov, V.F. *Parametric Synthesis of Nonlinear Automatic Control Systems*; SPb, SPbGUAP: St Petersburg, Russia, 2003; p. 358.
19. Meltzer, D. On the expressibility of piecewise linear continuous functions as the difference of two piecewise linear convex functions. *Math. Program. Study* **1986**, *29*, 118–134.
20. Baskakov, S.I. *Radio Circuits and Signals: Textbook for Universities*, 3rd ed.; Moscow Higher School: Moscow, Russia, 2000; p. 462.



21. Rahman, M. *Applications of Fourier Transforms to Generalized Functions*; Dalhousie University: Halifax, NS, Canada, 2011; 192p.
22. Balakrishnan, V. Generalized Functions. In *Mathematical Physics*; Springer: Cham, Switzerland, 2020; pp. 29–41. [\[CrossRef\]](#)
23. Vladimirov, V.S. Methods of the Theory of Generalized Functions. In *Analytical Methods and Special Functions*; Taylor&Francis: Abingdon, UK, 2002; p. 328.
24. Grosser, M.; Kunzinger, M.; Oberguggenberger, M.; Steinbauer, R. *Geometric Theory of Generalized Functions with Applications to General Relativity*; Kluwer Academic Publishers: Amsterdam, The Netherlands, 2001; 505p.
25. Popov, E.P. *Theory of Nonlinear Systems of Automatic Regulation and Control: Textbook Allowance*, 2nd ed.; Science, C., Ed.; Springer: Berlin/Heidelberg, Germany, 1988; p. 256.
26. Achieser, N.I. *Theory of Approximation*; Dover Publications: Dover, UK, 2004; 320p.
27. Varga, R. *Functional Analysis and Approximation Theory in Numerical Analysis*; Publishing House: Mir, Russia, 1974; p. 124.
28. Lyusternik, L.A.; Sobolev, V.I. *A Short Course in Functional Analysis*; Moscow Higher School: Moscow, Russia, 1982; p. 271.
29. Von Neumann, J. *Selected Works on Functional Analysis*; Nauka: Moscow, Russia, 1987; Volume 1–2.
30. Banach, S. Sur les opérations dans les ensembles abstraits et leur applications aux équations intégrales. *Fundam. Math.* **1922**, *3*, 133–181. [\[CrossRef\]](#)
31. Kirk, W.A. Fixed points of asymptotic contractions. *J. Math. Anal. Appl.* **2003**, *277*, 645–650. [\[CrossRef\]](#)
32. Jleli, M.; Samet, B. A new generalization of the Banach contraction principle. *J. Inequal Appl.* **2014**, *2014*, 38. [\[CrossRef\]](#)
33. Romaguera, S. On the Correlation between Banach Contraction Principle and Caristi's Fixed Point Theorem in b-Metric Spaces. *Mathematics* **2022**, *10*, 136. [\[CrossRef\]](#)
34. Choudhury, B.S.; Chakraborty, P. Local generalizations of Banach's contraction mapping principle. *J. Anal.* **2022**, *30*, 1131–1142. [\[CrossRef\]](#)
35. Kokkotas, K. *Interpolation, Extrapolation & Polynomial Approximation*; Eberhard Karls University of Tübingen: Tübingen, Germany, 2019; p. 38.
36. Lasser, R.; Mache, D.; Obermaier, J. On approximation methods by using orthogonal polynomial expansions. In *Advanced Problems in Constructive Approximation*; Springer: Basel, Switzerland, 2003; pp. 95–107.
37. Ditzian, Z.; Jiang, D. Approximation of Functions by Polynomials in  $C[-L, 1]$ . *Can. J. Math.* **1992**, *44*, 924–940. [\[CrossRef\]](#)
38. Khurmonov, K.M.; Shabozov, M.S. Jackson–Stechkin type inequalities between the best joint polynomials approximation and a smoothness characteristic in Bergman space. *Vladikavkaz. Mat. Zh.* **2022**, *24*, 109–120.
39. Trigub, R.M. On the Approximation of Functions by Polynomials and Entire Functions of Exponential Type. *Ukr. Math. J.* **2019**, *71*, 333–341. [\[CrossRef\]](#)
40. Gasquet, C.; Witomski, P. *Fourier Analysis and Applications: Filtering, Numerical Computation, Wavelets*; Springer: New York, NY, USA, 1999; Volume 30.
41. Levin, D. Reconstruction of Piecewise Smooth Multivariate Functions from Fourier Data. *Axioms* **2020**, *9*, 88. [\[CrossRef\]](#)
42. Fischer, J.V. Four Particular Cases of the Fourier Transform. *Mathematics* **2018**, *6*, 335. [\[CrossRef\]](#)
43. Rahman, M. *Applications of Fourier Transforms to Generalized Functions*; WIT Press: Southampton, UK, 2011.
44. Brandwood, D. *Fourier Transforms in Radar and Signal Processing*; Artech House, Inc.: Norwood, MA, USA, 2003.
45. Howard, R.M. Dual Taylor Series, Spline Based Function and Integral Approximation and Applications. *Math. Comput. Appl.* **2019**, *24*, 35. [\[CrossRef\]](#)
46. Ezhov, N.; Neitzel, F.; Petrovic, S. Spline Approximation—Part 1: Basic Methodology. *J. Appl. Geod.* **2018**, *12*, 139–155. [\[CrossRef\]](#)
47. Ezhov, N.; Neitzel, F.; Petrovic, S. Spline Approximation, Part 2: From Polynomials in the Monomial Basis to B-splines—A Derivation. *Mathematics* **2021**, *9*, 2198. [\[CrossRef\]](#)
48. Schumaker, L. *Spline Functions: Basic Theory*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2007.
49. Abdallah, A.K. An Approximation Method of Spline Functions. *Am. J. Math. Stat.* **2015**, *5*, 311–315. [\[CrossRef\]](#)
50. Usmonov, M.T. Least square method. *Sci. Educ. Sci. J.* **2021**, *2*, 54–65.
51. Hirica, I.; Udriste, C.; Pripoe, G.; Tevy, I. Least Squares Approximation of Flatness on Riemannian Manifolds. *Mathematics* **2020**, *8*, 1757. [\[CrossRef\]](#)
52. Kumari, K.; Yadav, S. Linear regression analysis study. *J. Pract. Cardiovasc. Sci.* **2018**, *4*, 33–36. [\[CrossRef\]](#)
53. Khasyanov, R. Hermite interpolation on a simplex. *Izv. Saratov Univ. Math. Mech. Inform.* **2018**, *18*, 316–327. [\[CrossRef\]](#)
54. Berriochoa, E.; Cachafeiro, A.; García Rábade, H.; García-Amor, J.M. Mechanical Models for Hermite Interpolation on the Unit Circle. *Mathematics* **2021**, *9*, 1043. [\[CrossRef\]](#)
55. Ibrahimoglu, B.A. Lebesgue functions and Lebesgue constants in polynomial interpolation. *J. Inequal. Appl.* **2016**, *2016*, 93. [\[CrossRef\]](#)
56. Smith, S.J. Lebesgue constants in polynomial interpolation. *Ann. Math. Inform.* **2006**, *33*, 109–123.
57. Dencker, P.; Erb, W.; Kolomoitsev, Y.; Lomako, T. Lebesgue constants for polyhedral sets and polynomial interpolation on Lissajous-Chebyshev nodes. *J. Complex* **2017**, *43*, 93. [\[CrossRef\]](#)
58. Stenger, F.; El-Sharkawy, H.; Baumann, G. *The Lebesgue Constant for Sinc Approximations*; Birkhäuser: Cham, Switzerland, 2014; pp. 319–335. [\[CrossRef\]](#)
59. Dautov, R.; Timerbaev, M. Numerical methods. In *Approximation of Functions*; Kazan Federal University: Kazan, Russia, 2021; p. 123.
60. Helmbert, G. The Gibbs phenomenon for Fourier interpolation. *J. Approx. Theory* **1994**, *78*, 41–63. [\[CrossRef\]](#)

61. Zhuk, V.V.; Natanson, G.I. *Trigonometric Fourier Series and Elements of Approximation Theory*; Publishing House Leningrad University: Saint Petersburg, Russia, 1983; p. 188.
62. Gjestland, F.J. *Distributions, Schwartz Space and Fractional Sobolev Spaces*; Norwegian University of Science and Technology: Trondheim, Norway, 2013.
63. Montillet, J.P. Sobolev Spaces, Schwartz Spaces, and a Definition of the Electromagnetic and Gravitational Coupling. *J. Mod. Phys.* **2017**, *8*, 1700–1722. [\[CrossRef\]](#)
64. Bhattacharyya, P.K. *Distributions: Generalized Functions with Applications in Sobolev Spaces*; De Gruyter: Berlin, Germany; Boston, MA, USA, 2012. [\[CrossRef\]](#)
65. Becnel, J.; Sengupta, A. The Schwartz Space: Tools for Quantum Mechanics and Infinite Dimensional Analysis. *Mathematics* **2015**, *3*, 527–562. [\[CrossRef\]](#)
66. Kotelnikov, V.A. On the bandwidth of “ether” and wire in telecommunications. *Adv. Phys. Sci. J.* **2006**, *7*, 762–770.
67. Gel’fand, I.M.; Vilenkin, N.Y. *Generalized Functions, Vols I–VI*; Academic Press: Cambridge, MA, USA, 1964. (In Russian)
68. Alabugin, A.; Aliukov, S.; Osintsev, K. Approximation Methods for Analysis and Formation of Mechanisms for Regulating Heat and Mass Transfer Processes in Heat Equipment Systems. *Int. J. Heat Technol.* **2020**, *38*, 45–58. [\[CrossRef\]](#)
69. Seregina, E.V.; Stepovich, M.A.; Makarenkov, A.M.; Filippov, M.N.; Platoshin, E.V. On a Modified Model of the Collective Motion of Minority Charge Carriers in a Two-Layer Semiconductor. In Proceedings of the XXIII International Scientific and Technical Conference on Photoelectronics and Night Vision Devices, Yundola, Russia, 13–15 November 2014; pp. 427–430.
70. Seregina, E.V.; Stepovich, M.A.; Makarenkov, A.M. On modification of the model of diffusion of non-basic charge carriers in semiconductor materials based on the use of recursive trigonometric functions, and estimation of the stability of the power of the solution. *Surf. X-ray Synchrotron Neutron Res.* **2014**, *9*, 72.
71. Kiselev, V.A. Construction Mechanics: Special Course. Dynamics and Stability of Structures. In *Textbook for Universities*, 3rd ed.; Stroyizdat Publisher: Moscow, Russia, 1980; p. 616.
72. Seregina, E.V.; Stepovich, M.A.; Makarenkov, A.M. Modified model of minority charge carrier diffusion in semiconductor materials. *Adv. Appl. Phys.* **2013**, *1*, 544–547.
73. Seregina, E.V.; Stepovich, M.A.; Makarenkov, A.M. On the modification of a model of minority charge-carrier diffusion in semiconductor materials based on the use of recursive trigonometric functions and the estimation of the stability of solutions for the modified model. *J. Surf. Investig. X-ray Synchrotron Neutron Tech.* **2014**, *8*, 922–925. [\[CrossRef\]](#)
74. Seregina, E.V.; Stepovich, M.A.; Makarenkov, A.M.; Filippov, M.N.; Platoshin, E.V. About the possibility of using the trigonometric expressions in the form of recursive functions for solving the diffusion equation with discontinuous coefficients. *Appl. Phys.* **2015**, *1*, 5–10.
75. Seregina, E.V.; Stepovich, M.A.; Makarenkov, A.M.; Filippov, M.N. About the possibility of using recursive trigonometric functions for calculating the distribution of non-equalable non-basic charge carriers in two-floor water. *Surf. X-ray Synchrotron Neutron Res.* **2015**, *9*, 70.
76. Seregina, E.V.; Makarenkov, A.M.; Stepovich, M.A.; Filippov, M.N. On the possibility of using recursive trigonometric functions to calculate the distribution of nonequilibrium minority charge carriers in a two-layer semiconductor material. *J. Surf. Investig. X-ray Synchrotron Neutron Tech.* **2015**, *9*, 929–933. [\[CrossRef\]](#)
77. Wittry, D.B.; Kyser, D.F. Measurement of Diffusion Lengths in Direct-Gap Semiconductors by Electron-Beam Excitation. *J. Appl. Phys.* **1967**, *38*, 375. [\[CrossRef\]](#)
78. Rao-Sahib, T.S.; Wittry, D.B. Measurement of Diffusion Lengths in p-Type Gallium Arsenide by Electron Beam Excitation. *J. Appl. Phys.* **1969**, *40*, 3745. [\[CrossRef\]](#)
79. Van Roosbroeck, W. Injected Current Carrier Transport in a Semi-Infinite Semiconductor and the Determination of Lifetimes and Surface Recombination Velocities. *J. Appl. Phys.* **1955**, *26*, 380. [\[CrossRef\]](#)
80. Stepovich, M.A.; Snopova, M.G.; Khokhlov, A.G. Usage of model of independent sources for calculation of distribution of minority carriers generated in a two-layer semiconductor by a electron beam. *Appl. Phys.* **2004**, *3*, 61.
81. Burylova, I.V.; Petrov, V.I.; Snopova, M.G.; Stepovich, M.A. Mathematical simulation of distribution of minority charge carriers, generated in multy-layer semiconducting structure by a wide electron beam. *Phys. Technol. Semicond.* **2007**, *41*, 458.
82. Baek, D.H.; Kim, S.B.; Schroder, D.K. Epitaxial silicon minority carrier diffusion length by photoluminescence. *J. Appl. Phys.* **2008**, *104*, 054503. [\[CrossRef\]](#)