Filomat 38:22 (2024), 7949–7970
https://doi.org/10.2298/FIL2422949A

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Recovery of the intermediate derivative of an approximately given function

A. S. Kochurov[a]

[a]*Lomonosov Moscow State University, Russian Federation*

Abstract. The problem of numerical recovery of intermediate derivatives or a differential operator is considered. With this aim, the norm of the leading derivative is estimated, from which optimal difference schemes for intermediate derivatives are obtained. An example of a smooth model function is considered for which the measurement error is simulated using real-world experimental data. The trusted probability of the obtained estimate for the norm of the leading derivative is estimated. Intervals of trusted recovery of intermediate derivatives are estimated.

1. Introduction

In the present paper, we will estimate the uniform norm of the leading derivative for a function x approximately defined on a grid. This problem is important for evaluation of intermediate derivatives of x or of some differential operator of this function. Problems of this kind appear frequently in numerous applied problems (see, for example, [1]–[9], [13], [15]–[18], [21]), in which an estimate of the leading derivative is assumed to be known. However, the problem of estimation of the norm of the leading derivative from approximately given values of a function is a nontrivial problem. The purpose of the present paper is to work out concrete recommendations for delivering estimates of this kind for the norm of the leading derivative.

The majority of ill-posed problems on a class of smooth functions is usually solved according to the following general scheme. In the class of smooth functions of which the original problem is posed, one singles out the subset of functions whose derivative is bounded from above by a fixed constant. After this the original problem is considered only on functions of this subclass. This constraint usually reduces the ill-posed problem to that amenable to standard methods of the theory of functions. However, in general, no estimate of the norm of (say) leading derivative (of order n) is known. This is the principal impediment to application of many remarkable results on optimal recovery of various characteristics.

The present paper is not concerned with new methods of recovery of derivatives. Instead, we attack the main problem for estimation of the norm of the leading derivative. This will make it possible to employ the available methods for optimal recovery of intermediate derivative of order $< n$.

Let $x_0(\cdot)$ be a smooth real function on $\mathbb{R}$. One knows approximate values of $x_0(\cdot)$ obtained from a real-world experiment. We construct an algorithm for estimating the norm of the leading derivative of $x_0(\cdot)$ from

2020 *Mathematics Subject Classification*. Primary 65D25.

Keywords. derivative recovery, recovery of differentiation operator, divided difference, numerical differentiation, intermediate derivative.

Received: 20 December 2023; Revised: 20 February 2024; Accepted: 01 March 2024

Communicated by Miodrag Spalević

Email address: kchrvas@yandex.ru (A. S. Kochurov)

these real data. This makes its possible to numerically recover the values of all intermediate derivatives of $x_0(\cdot)$. The performance of the algorithm developed in the present paper is also tested on artificially generated random noise data, the resulting estimate of the norm of the leading derivative is found to be correct with large probability. Numerous examples of model recovery (a polynomial with noise) are given.

In the simplest case, in such problems, the values of the measured function $x_0(\cdot)$ are known to a certain accuracy δ (the measurement accuracy) at points $\{t_i\} \subset \mathbb{R}$, i.e., for each i, numbers $y_i \in \mathbb{R}$ are defined so that $|y_i - x_0(t_i)| \leq \delta$. The problem is as follows: how one should act in order, for some $k \in \mathbb{N}$, $k \leq n$, to approximately evaluate the derivative $x_0^{(k)}(t)$ at a point t, which is, for example, a points from the system $\{t_i\}$? This problem is a particular case of problem of the recovery theory, and in this problem, a certain terminology, statements of problems, and approaches to their solution have been developed.

These problems were originated and developed in the works of S.B. Stechkin [4], [5], V.V. Arestov [17], V.N. Gabushin [7], Yu.N. Subbotin, L.V. Taikov [8], [9], A.P. Buslaev, V.M. Tikhomirov, and in a large number of works by G.G. Magaril-Il'yaev and K.Yu. Osipenko [13], [14], [15], [16], C. A. Micchelli, T. J. Rivlin [20], [21], and many others. For a survey on this and related topics, see [17], [18].

Let us recall some definitions, statements of problems, and methods of their solutions.

Any mapping Λ of a space sending $\{y_i\}$ into $\mathbb{R}$ is called a method of recovery (in our problem, a method of recovery of the derivative $x_0^{(k)}(t)$). The recovery error on a given vector of measurements $\{y_i\}$ corresponding to a selected method of recovery Λ is defined by

$$|\Lambda(\{y_i\}) - x_0^{(k)}(t)|\,. \tag{1.1}$$

A recovery method, under a nonadaptive approach, should not depend on the choice of $x_0(\cdot)$ and the system $\{y_i\}$, and hence, as an error related to Λ, one uses the "worst" error, which may appear with all possible choices of $x(\cdot)$ and $\{y_i\}$ — this leads to the concept of the error of the method

$$e(W, \Lambda, \delta) = \sup_{x \in W,\ \{y_i\}:\ |y_i - x(t_i)| \leq \delta\ \forall i} |\Lambda(\{y_i\}) - x^{(k)}(t)| \tag{1.2}$$

on the class of functions W. The choice of the class W depends substantially on the construction of the so-called optimal recovery method. A recovery method $\widehat{\Lambda}$ is called optimal (on a class of functions W) if

$$e(W, \widehat{\Lambda}, \delta) = E(W, \delta), \tag{1.3}$$

where

$$E(W, \delta) = \inf_{\Lambda} e(W, \Lambda, \delta) \tag{1.4}$$

is the error of recovery on the class W. For the above problem, as W one uses the class of functions

$$W = \{x(\cdot) \in W_\infty^{n+1}(\mathbb{R}) \ : \ \|x^{(n+1)}\|_\infty \leq \Delta\}, \tag{1.5}$$

where $W_\infty^{n+1}(\mathbb{R})$, $n \in \mathbb{N}$, is the class of functions $x : \mathbb{R} \to \mathbb{R}$ with absolutely continuous nth derivative $x^{(n)}$, essentially bounded $(n+1)$st derivative, and $\Delta > 0$ should be refined later: the smaller Δ, the smaller $E(W, \delta)$, and, on the other hand, if Δ is chosen to be so small, then the original element $x_0(\cdot) \in W_\infty^{n+1}(\mathbb{R})$, whose kth derivative should be recovered, may happen to lie outside W.

So, the parameter $\Delta > 0$ should satisfy the inequality $\Delta \geq \|x_0^{(n+1)}\|$. The equality $\Delta = \|x_0^{(n+1)}\|$ might be the best variant of this choice, however, $\|x_0^{(n+1)}\|$ is unknown. As an alternative to the best choice of Δ, we may choose upper estimates $\|x_0^{(n+1)}\|$ as $\Delta > 0$: if $\Delta \geq \|x_0^{(n+1)}\|$ exceeds $\|x_0^{(n+1)}\|$ only by a not very large number of times, then the optimal method $\widehat{\Lambda} = \widehat{\Lambda}(\Delta)$ will provide a good error of recovery for (1.1).

To summarize: the recovery problem calls for the choice of a constant $\Delta > 0$ and $W = W(\Delta, n)$, and requires finding the recovery error $E(W, \delta)$ and explicit provision of an optimal recovery method. For practical purposes, a satisfactory answer includes only provision of a near optimal recovery method: for example, a method for which the recovery error from (1.2) is not worse than $E(W, \delta)$ in, say, 3 or 5 times.

For some pretty general sets W which are symmetric about zero, evaluation of $E(W,\delta)$ and construction of an optimal (or a near-optimal) recovery method is closely related to the study of a special recovery problem, i.e., to a situation where $x_0(\cdot)$ is the identically zero mapping; in this case, the recovery problem amounts to a simpler one. As already mentioned, the statement of the general recovery problem and its simultaneous study with the above particular case was given in [5]; later, this approach has contributed significantly to many other recovery problems.

Let us write down this particular problem in the context of the recovery problem we are interested in:

$$x^{(k)}(0) \to \sup \quad \|x\|_C := \sup_{t\in\mathbb{R}} |x(t)| \le \delta, \quad \|x^{(n+1)}\|_{L_\infty} := \inf_{M\subset\mathbb{R}:\ \mu(M)=0} \sup_{t\in\mathbb{R}\setminus M} |x^{(n+1)}(t)| \le \Delta\,. \tag{1.6}$$

For $n = 2$, problem (1.6) was solved by J. Hadamard [10], and in the general case, by A. N. Kolmogorov [11]. This problem is a particular case of optimal control problems (a kind of conditional extremum problems). Such problems can be successfully solved by the well known method of Lagrange multipliers, which, for optimal control problems, is augmented by the Pontryagin maximum principle. Such an approach is capable of composing necessary extremum conditions and using them to single out a small set of functions on which a solution may be delivered. However, since in addition the above problem (1.6) is convex, the necessary and sufficient conditions for extremum for this problem coincide. As a result, for small n, problem (1.6) can be successfully attacked by standard methods.

For problem (1.6), its value coincides with $E(W,\delta)$, which was considered in (1.1)–(1.5), and by using the Lagrange multipliers, which appeared in the course of the solution of (1.6), one can construct an optimal method in the general problem, for which $x_0(\cdot)$ is not necessary the identically zero function. Let us give some known examples (see [5])

In the case $k = n = 1$, the formula, which was obtained in the solution of problem (1.6),

$$\dot{x}(0) = \frac{y_1 - y_{-1}}{2h} \tag{1.7}$$

gives the derivative $x^{(1)}(0)$ with error not exceeding $\sqrt{2\delta\Delta}$ if $h = (2\delta/\Delta)^{1/2}$ is an optimal step; here it is assumed that the function $x(\cdot)$ of real variable is known approximately with accuracy $\delta > 0$ at all points of the real line and $|x(h) - y_1| \le \delta$, $|x(-h) - y_{-1}| \le \delta$; we also assume that $\|x^{(2)}\|_{L_\infty} \le \Delta$. The same recovery formula for the first derivative, but with a different optimal h, also holds for $k = 1$, $n = 3$.

In the case $k = n = 2$, the formula

$$\ddot{x}(0) = \frac{y_1 - 2y_0 + y_{-1}}{h^2} \tag{1.8}$$

gives $x^{(2)}(0)$ with error not exceeding $\sqrt[3]{3\delta\Delta^2}$ if $h = (24\delta/\Delta)^{1/3}$ is an optimal step. Here also, as in the previous example, the function $x(\cdot)$ is known approximately at all points $\mathbb{R}$, $|x(h) - y_1| \le \delta$, $|x(0) - y_0| \le \delta$, $|x(-h) - y_{-1}| \le \delta$ and $\|x^{(3)}\|_{L_\infty} \le \Delta$.

It is also known that, in each of the above cases, the above formula is optimal, i.e., the error of its recovery is realized on some explicit admissible function and the error of its recovery at nodes of a uniform grid. In the actual fact, correctness of the estimate of the norm of the leading derivative is verified from "good" recovery (i.e., on a sufficiently large set) of intermediate derivatives (see § 6).

In the case $n > 2$, any optimal recovery formula for any intermediate derivative contains an infinite number of terms (for $n = 2, 3$, it was obtained by Arestov [6]). For $n > 3$, the solution was given by Buslaev [13]. Their results depend essentially on the previous studies by Domar [12] and Arestov.

From these examples (as well as from other solved recovery problems) one can see that the choice of Δ (an estimate for the leading derivative) plays a principal role in the construction of a recovery method of the derivative and estimating the error of this method. Let us consider one approach to obtaining an approximate value of Δ. This approach will be demonstrated on an example of evaluation of the second order derivative, for which, in order to find an optimal h in (1.8), one should estimate $\|x^{(3)}\|_{L_\infty} \le \Delta$.

Experimental data in these examples were provided for the study and evaluation of processing methods by the Laboratory of the Physical Biochemistry at the National Medical Research Center of Hematology

(Russia). These data consist of several tuples of numbers of which each represents approximate values of the function $x(t, s)$ of the variables t on the uniform grid $\{t_i\}$ for $s = s_k$, $k = 1, 2, \dots$. In changing from one tuple to a different one, the values of s_k also change with constant step.

2. Structure of the ratios of divided differences for functions defined with error on a grid

Let $n \in \mathbb{Z}_+$. In the problem of estimation of the largest value of the absolute value of the $(n+1)$st derivative of a function $x(\cdot) : \mathbb{R} \to \mathbb{R}$, we know only approximate values of $\widetilde{x}(t_i) = x(t_i) + \delta(t_i)$ of this function, which differ from $x(t_i)$ by a small quantity $\delta(t_i)$ at some points $t_i \in \mathbb{R}$, $t_{i+1} - t_i = \tau$. In this setting, one cannot guarantee a satisfactory estimate of $\|x^{(n+1)}\|_C$ for an arbitrary sufficiently smooth function $x(\cdot)$ (even if all $\delta(t_i)$ are zero). Hence, in this problem, we will be focused on the case when from evaluation of various difference relations for $x(\cdot)$, in which precise values of $x(t_i)$ are involved, one may still properly judge about the quantity $\|x^{(n+1)}\|_C$, thereby providing the value which is smaller than $\|x^{(n+1)}\|_C$, but which differs from it by a small (of order 1) factor.

Recall the definition of the successive difference of order $(n+1)$ for a function x:

$$\Delta_{j\tau}^{n+1} x(t_i) = \sum_{k=0}^{n+1} (-1)^{n+1-k} C_{n+1}^k x(t_i + kj\tau) = \sum_{k=0}^{n+1} (-1)^{n+1-k} C_{n+1}^k x(t_{i+kj}), \quad C_{n+1}^k = \frac{(n+1)!}{k!(n+1-k)!}.$$

For the functions $\widetilde{x}$, δ we proceed similarly. The linearity in x in the definition implies that

$$\Delta_{j\tau}^{n+1} \widetilde{x}(t_i) = \Delta_{j\tau}^{n+1} x(t_i) + \Delta_{j\tau}^{n+1} \delta(t_i) \tag{2.1}$$

or, more precisely, already for the divided difference

$$\sum_{k=0}^{n+1} \frac{(-1)^{n+1-k}}{(j\tau)^{n+1}} C_{n+1}^k \widetilde{x}(t_{i+kj}) = \sum_{k=0}^{n+1} \frac{(-1)^{n+1-k}}{(j\tau)^{n+1}} C_{n+1}^k x(t_{i+kj}) + \sum_{k=0}^{n+1} \frac{(-1)^{n+1-k}}{(j\tau)^{n+1}} C_{n+1}^k \delta(t_{i+kj}).$$

In addition to the order $(n+1)$ and the function $x(\cdot)$, this definition of $\Delta_{j\tau}^{n+1} x(t_i)$ involves the initial point t_i and the step $j\tau$ at which this difference is evaluated.

It is known (see, for example, [19], p. 231) that if a function $x(\cdot)$ has the continuous $(n+1)$st derivative on the interval $[t_i, t_{i+(n+1)j}]$, then

$$\frac{\Delta_{j\tau}^{n+1} x(t_i)}{(j\tau)^{n+1}} = \frac{1}{(j\tau)^{n+1}} \int\limits_{[0,\, j\tau]^{n+1}} x^{(n+1)}(t_i + \beta_1 + \ldots + \beta_{n+1})\, d\beta_1 \ldots d\beta_{n+1} = x^{(n+1)}(\zeta), \tag{2.2}$$

where ζ is some intermediate point on $[t_i, t_{i+(n+1)j}]$.

If the values of $x(t_i)$ were known precisely, i.e., if $\delta(t_i)$ would be zero for all i, then for estimation of the modulus of the $(n+1)$st derivative of x one could evaluate all possible values of $(j\tau)^{-n-1}|\Delta_{j\tau}^{n+1} x(t_i)|$ from all admissible i, j, take the largest one among them, and assume that the maximum of the absolute value of the required derivative is bounded by this quantity. Here it is worth pointing out again that the step τ for the nodes t_i at which the values of $x(t_i)$ are specified is sufficiently small, so that the sequence $\{x(t_i)\}$ would correctly reflect the characteristic features of the function $x(t)$, and, using which, one would be able to recover, with good precision, the maximum of the absolute value of $(n+1)$st derivative.

We will say that the step $j\tau$ is *acceptable* (for a smooth function $x(\cdot)$ and a set of nodes $x(t_i)$) if, when evaluating the largest absolute value of expressions (2.2) with given j, we will obtain a number that differs from $\|x^{(n+1)}\|$ by a small (of order 1) number of times. Such a decision (whether the step $j\tau$ was acceptable or not) does not depend on the accuracy of measurements, but depends only on the character of the quantity x being measured. This decision is taken in each case individually and is not discussed in the present paper.

If, in place of $x(t_i)$, one knows only approximate values of $\widetilde{x}(t_i)$, then, by evaluating all possible values of $\Delta_{j\tau}^{n+1} \widetilde{x}(t_i)$ from i, j, we get quantities that differ from $\Delta_{j\tau}^{n+1} x(t_i)$ by $\Delta_{j\tau}^{n+1} \delta(t_i)$. We will assume that the values of the function $\delta(t_i) \in [-\delta_0, \delta_0]$ are realizations of some random variable with zero expectation. By definition and from the condition $|\delta(t_i)| \le \delta_0$ we have

$$\frac{1}{(j\tau)^{n+1}}|\Delta_{j\tau}^{n+1}\widetilde{x}(t_i) - \Delta_{j\tau}^{n+1}x(t_i)| \le \frac{\delta_0}{(j\tau)^{n+1}}\sum_{k=0}^{n+1} C_{n+1}^k = \left(\frac{2}{j\tau}\right)^{n+1}\delta_0. \tag{2.3}$$

Since usually the absolute value of the $(n+1)$st derivative of a smooth function $x(\cdot)$ is comparable with the norm of $\|x^{(n+1)}\|$ on a relatively small interval of variation of the independent argument, from (2.2) it follows that by evaluating $\frac{\Delta_{j\tau}^{n+1}\widetilde{x}(t_i)}{(j\tau)^{n+1}}$ for large j, we will obtain values that lie far from $\|x^{(n+1)}\|$ for any choice of the initial point t_i. Hence, to estimate $\|x^{(n+1)}\|$, one should focus on the possibly small values the parameter j, which specifies the step size $j\tau$ for which $\Delta_{j\tau}^{n+1}\widetilde{x}(t_i)$ is calculated. However, in evaluation of $(j\tau)^{-n-1}|\Delta_{j\tau}^{n+1}\widetilde{x}(t_i)|$ one should take j to be so large to reduce the effect of measurement errors on evaluation of these values with increased j — in view of (2.3) the effect (on the estimate of the $(n+1)$st derivative) of the measurement error decreases with the rate at least j^{-n-1}.

We set $b_i := (j\tau)^{-n-1}\Delta_{j\tau}^{n+1}x(t_i)$, $c_i := (j\tau)^{-n-1}\Delta_{j\tau}^{n+1}\widetilde{x}(t_i)$, $r_i := (j\tau)^{-n-1}\Delta_{j\tau}^{n+1}\delta(t_i)$ for brevity, and consider the quantities

$$\frac{c_i}{c_{i+1}} - 1 \tag{2.4}$$

for various i. We note here that all c_i, b_i, r_i and the relations in (2.4) depend on the parameter j: $c_i = c_{i,j}, b_i = b_{i,j}, r_i = r_{i,j}$. Let us illustrate, for example, the ranges of variation of the indexes i and j: j may assume any values 1, 2, 3, ... If the interval, on which the sequence $\{\widetilde{x}(t_i)\}$ is investigated, is $i \in [N_1, N_2]$, then, for a fixed $j \in \mathbb{N}$, the interval of variation i, where the sequence $\{\frac{c_i}{c_{i+1}}\}$ is defined, reduces to $[N_1, N_3 - 1]$, $N_3 = N_2 - j(n+1)$. If it happens that $N_1 > N_3 - 1$, then for such j the sequence $\{\frac{c_i}{c_{i+1}}\}$ cannot be considered.

As already mentioned, regarding the quantities $\delta(t_i)$, which constitute the distortion in $\widetilde{x}(t_i)$, we will assume that, for all i, their effect on $x(t_i)$ is random and is an observation of some random variable.

If, on some interval of variation of i, in

$$\frac{c_i}{c_{i+1}} = \frac{b_i + r_i}{b_{i+1} + r_{i+1}} \tag{2.5}$$

the absolute values of r_i, r_{i+1} are small in comparison with $|b_i|, |b_{i+1}|$ (for example, $|r_s| \le \rho \cdot |b_s|$, $s = i, i+1$, ρ is a small positive number), then the ratios $\frac{c_i}{c_{i+1}}$ differ from $\frac{b_i}{b_{i+1}}$ by the factor θ, and, in the worst case, they are confined in the interval $[\kappa^{-1}, \kappa]$, $\kappa = \frac{1+\rho}{1-\rho}$:

$$\frac{c_i}{c_{i+1}} = \frac{b_i}{b_{i+1}} \cdot \theta. \tag{2.6}$$

Assuming, for the sake of clarity, that b_s are values of some smooth function $b_s = g(t_s)$ at the points t_s, $s = i, i+1$, and assuming that this function preserves its sign on $[t_i, t_{i+1}]$, we write the identity (for $g(t) > 0$ on $[t_i, t_{i+1}]$):

$$b_i/b_{i+1} = \exp(\ln g(t_i) - \ln g(t_{i+1})).$$

Since $\ln g(t_i) - \ln g(t_{i+1}) = (\ln g)'(\zeta) \cdot \tau$, and $\zeta \in [t_i, t_{i+1}]$ is some intermediate point, it is natural to expect that

$$\frac{b_i}{b_{i+1}} - 1 = O(\tau). \tag{2.7}$$

Similar relations may also be written for the case $g(t) < 0$ on $[t_i, t_{i+1}]$.

Remark 2.1. *If the function $g(t)$ changes its sign on $[t_i, t_{i+1}]$, then b_i, b_{i+1} do not seem to be maximal in absolute value, and such an i can be excluded from consideration.*

Remark 2.2. *Since we a priori do not know the value of the derivative in comparison with the error, the logarithmic scale appears to be most representative for representing all possible relations between the evaluated derivative and the measurement error.*

Estimating the quantities from (2.4), we get in view of (2.6), (2.7)

$$\left|\frac{c_i}{c_{i+1}} - 1\right| = \left|\Big(\frac{b_i}{b_{i+1}} - 1\Big)\theta + (\theta - 1)\right| = |\theta - 1| + O(\tau), \quad \theta \in [\kappa^{-1}, \kappa]. \tag{2.8}$$

As a result, for small positive ρ and τ, the values in (2.4) are close to zero precisely for these values of the estimated derivative, where b_i is essentially prevailed over the random component r_i.

This should be manifested most markedly at the points of at which b_i is greatest in absolute value, because near the maximum point the smooth function behaves like a constant function, i.e., $|\frac{b_i}{b_{i+1}} - 1|$ is small, and since $|\theta - 1|$ is most likely the smallest possible. Here it is worth pointing out that because of the continuous dependence (on the index i) the smallness of the quantities in (2.4) should be observed not for separate i, but for entire intervals of variation of i.

If, vice versa, on some interval of variation of i in (2.5), r_i is prevailing in absolute value over b_i, and r_{i+1} is prevailing over b_{i+1}, then, due to the random behavior of r_i, the values of the ratios from (2.5) are also random.

Let us consider various relations between b_i and r_i. To this end, we define the quantities $\frac{c_i}{c_{i+1}} - 1$. For a more complete visual representation of information, we will use not the values of $\frac{c_i}{c_{i+1}}$, but rather the quantities

$$a_i = \begin{cases} c_i/c_{i+1} & \text{if } \left|\frac{c_i}{c_{i+1}}\right| \leq 1, \\ c_{i+1}/c_i & \text{if } \left|\frac{c_i}{c_{i+1}}\right| > 1. \end{cases} \tag{2.9}$$

In the examples that follow, we will use the approximate values $\widetilde{x}(t_i)$ of the function $x(t_i) = x(t_i, s)$ at the points $t_i = i\tau$, $i = 0, \dots, 1484$ with some fixed s. The explicit expression for τ occurs only in the factor multiplying all the difference relations of the form b_i, c_i, r_i, and so for brevity we may assume that $\tau = 1$; this will result in no changes in the relative characteristics of the form a_i. From $\{\widetilde{x}(t_i)\}$ and $n = 2$, $n + 1 = 3$, we evaluate $\Delta_j^{n+1}\widetilde{x}(t_i)$ and $(a_i - 1)$ for various steps j and for the interval $[N_1, N_2] = [0, 1484]$ of variation of i in the sequence $\{\widetilde{x}(t_i)\}$. In Fig. 1, we show the dependence of $(a_i - 1)$ on i for the step $j = 1, 20, 40, 60$.

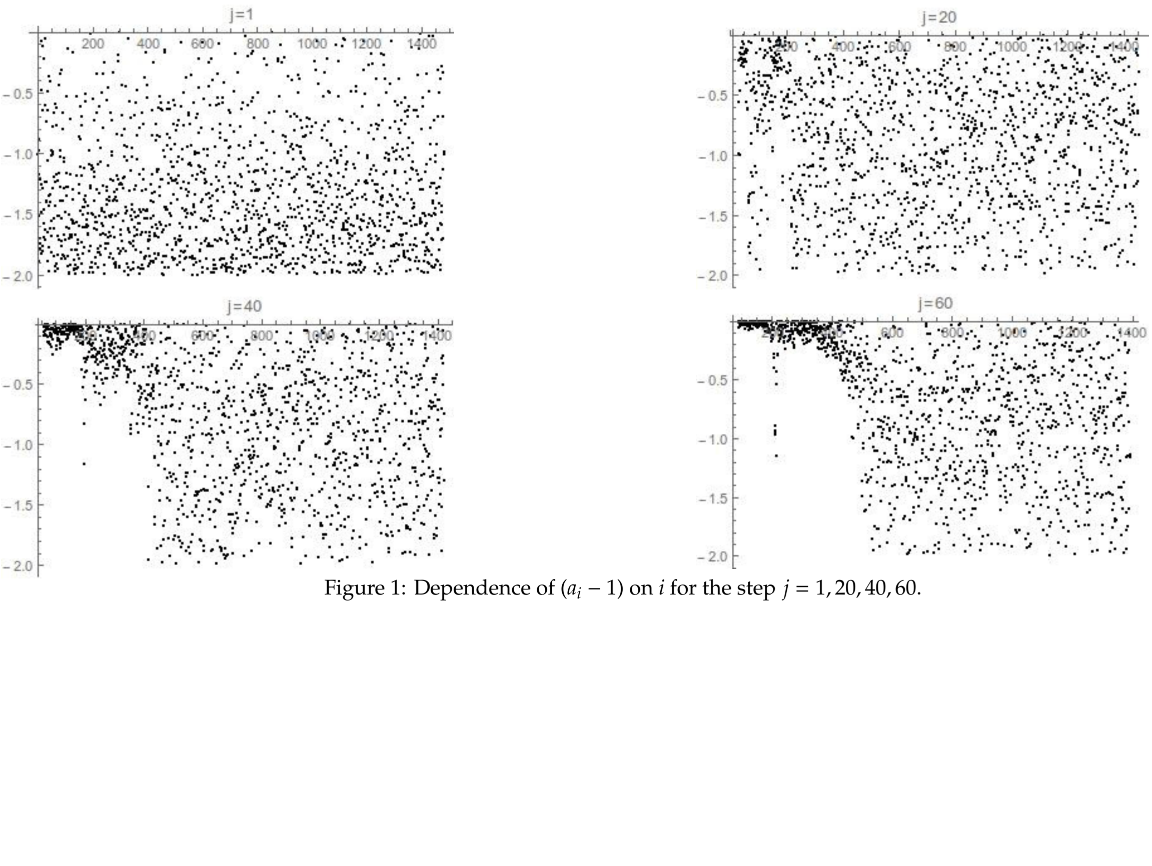

Figure 1: Dependence of $(a_i - 1)$ on i for the step $j = 1, 20, 40, 60$.

It is clear that, for $j = 1$, r_i's prevail over b_i's for all admissible values of i: the graph of the sequence $(a_i - 1)$ is random, and with increasing j in the range $i \in [N_1, N_1 + 400]$ the quantities r_i start to be affected by b_i, since a_i approach 1 (in the graphs, $(a_i - 1)$ tends to 0).

Remark 2.3. *For $j = 40$ and $j = 60$, there is a marked registered departure of $(a_i - 1)$ from zero for numerical evaluation near $i = 125$. This is related to the fact that at some point near $i = 125$ the third derivative of the function $x(\cdot)$ is zero, and hence b_i is close to zero.*

3. Identification of the greatest $(n + 1)$st divided difference as a function of the step

Let us investigate how the behavior of the sequence a_i changes as the step j increases. We will be oriented on the appearance of a j for which one can single out one or several essential intervals of variation of i, where the sequence $\{a_i = a_{i,j}\}$ from (2.9) is close to 1.

In Fig. 1 ($j = 40$ and $j = 60$), on the interval $[1, 400]$, there are closed intervals (possibly of small length) on which the calculated $(1 - a_i)$'s are small in magnitude:

$$(1 - a_i) \ll 1, \quad \text{i.e.} \quad 0 \leq 1 - a_i \leq \epsilon, \quad i \in [1, 400], \tag{3.1}$$

ϵ is much smaller than 1. It is natural to assume that, on these intervals, $|b_i|$ are much larger than $|r_i|$, and hence, $b_i \approx c_i$, and consequently, on these intervals $|b_i| \approx |c_i|$ attain their maximum values among the remaining ones (for the step j chosen for the calculation of c_i). As a result, the maximum of $|c_i|$, as calculated for these intervals, gives an approximate value for $\|x^{(n+1)}\|$ (provided that the step j is *acceptable*, i.e., under the condition that the greatest in absolute value $b_s = (j)^{-n-1}\Delta_j^{n+1}x(t_s)$, as obtained already from all possible s's, is different from $\|x^{(n+1)}\|$, but by a small factor).

Let us consider in more detail the verification of condition (3.1). By definition, $(a_i - 1) \in [-2, 0]$ for any i. The quantity a_i is the ratio of c_i and c_{i+1}, of which each is the sum of a smooth component and a quantity of random nature — these being b_i and r_i for c_i, and, respectively, b_{i+1} and r_{i+1} for c_{i+1}. Recall that $r_i = j^{-n-1}\Delta_j^{n+1}\delta(t_i)$ are obtained as realizations of some random variable ξ formed from δ. Let

$$r_i \in [-R, R], \quad R > 0, \tag{3.2}$$

be the interval of observation of r_i.

Let us consider several possible variants of behavior of the sequence $\{a_i\}$. If the condition $(a_i - 1) > -1$, or equivalently $a_i > 0$, is satisfied on some range of variation of $i \in [i_1, i_2] \subset [N_1, N_3 - 1]$, then by definition

for $\left|\frac{c_i}{c_{i+1}}\right| \leq 1$ we have $a_i = \frac{c_i}{c_{i+1}} > 0$, i.e., c_i and c_{i+1} are of the same sign,

for $\left|\frac{c_{i+1}}{c_i}\right| \leq 1$ we have $a_i = \frac{c_{i+1}}{c_i} > 0$, i.e., c_i and c_{i+1} are of the same sign again.

Hence, for all $i \in [i_1, i_2 + 1]$, $c_i = b_i + r_i$ are of the same sign. The interval $[i_1, i_2 + 1]$ is quite large and r_i appear as realizations of the random variable ξ. Hence on $[i_1, i_2 + 1]$ the quantity $|b_i|$ assumes values that are approximately larger than R (though, perhaps, not for all $i \in [i_1, i_2 + 1]$, and even, perhaps, only for a small number of i). In general, R is not known, but in any case we approximately have $R \leq \max_{i\in[i_1,i_2+1]} |b_i|$. Hence, for any $i_0 \in [i_1, i_2 + 1]$, $|c_{i_0}| \leq |b_{i_0}| + \max_{i\in[i_1,i_2+1]} |b_i|$ holds approximately, i.e.,

$$\max_{i\in[i_1,i_2+1]} |b_i| \geq \max_{i\in[i_1,i_2+1]} |c_i|/2. \tag{3.3}$$

Let us generalize (3.3) for other assumptions on a_i:

$$a_i > \alpha, \quad i \in [i_1, i_2]. \tag{3.4}$$

If $\alpha \in (0, 1)$, then c_i, $i \in [i_1, i_2 + 1]$, also preserve sign, and the best situation occurs if $|r_{i_0}| = |r_{i_0+1}| = R$, $r_{i_0} \cdot r_{i_0+1} < 0$ for some $i_0 \in [i_1, i_2]$. By definition (for $c_i \geq 0$),

$$\alpha(b_{i+1} + r_{i+1}) \leq b_i + r_i \leq \frac{b_{i+1} + r_{i+1}}{\alpha}$$

and hence, both for $r_{i_0} = -r_{i_0+1} = R$, and for $-r_{i_0} = r_{i_0+1} = R$,

$$\frac{1+\alpha}{1-\alpha} \cdot R \leq \max\{b_{i_0}, b_{i_0+1}\} + O(\tau)$$

i.e.,

$$\max_{i\in[i_1,i_2+1]} |b_i| \geq \frac{1+\alpha}{2} \max_{i\in[i_1,i_2+1]} |c_i| + O(\tau). \tag{3.5}$$

This inequality holds also for $\alpha \in (-1, 0)$ (if, as above, there is $i_0 \in [i_1, i_2]$ such that $|r_{i_0}| = |r_{i_0+1}| = R$, $r_0 \cdot r_{i_0+1} < 0$).

It will be convenient, in place of the above condition $|r_{i_0}| = |r_{i_0+1}| = R$, $r_{i_0} \cdot r_{i_0+1} < 0$, we will consider the inequalities $R(1-\beta_2) \leq |r_{i_0}|, |r_{i_0+1}| \leq R(1-\beta_1)$, $r_i r_{i_0+1} < 0$, or $R(1-\beta_2) \leq r_{i_0} \leq R(1-\beta_1)$, or $-R(1-\beta_1) \leq r_{i_0} \leq -R(1-\beta_2)$, where $0 \leq \beta_1 < \beta_2 < 1$. Under these constraints, we will obtain an estimate similar to (3.5). To this end, we will consider the sets

$$\begin{aligned} A_i &:= \{\mathbf{r} = \{r_{N_1}, \ldots, r_{N_3}\} : R(1-\beta_2) \leq |r_i|, |r_{i+1}| \leq R(1-\beta_1),\ r_i r_{i+1} < 0\},\ i \in [N_1, N_3 - 1], \\ B_i &:= \{\mathbf{r} = \{r_{N_1}, \ldots, r_{N_3}\} : -R(1-\beta_1) \leq r_i \leq -R(1-\beta_2)\}, \quad i \in [N_1, N_3], \\ C_i &:= \{\mathbf{r} = \{r_{N_1}, \ldots, r_{N_3}\} : R(1-\beta_2) \leq r_i \leq R(1-\beta_1)\}, \quad i \in [N_1, N_3], \end{aligned} \tag{3.6}$$

where $0 \leq \beta_1 < \beta_2 < 1$. Note that

$$A_i = (C_i \cap B_{i+1}) \cup (C_{i+1} \cap B_i).$$

Let $P(i, i_2) := P(\bar{A}_i \cap \ldots \cap \bar{A}_{i_2})$ be the probability that none of the events A_m, $m = i, \ldots, i_2$, will hold. We also set $S(i) = P(C_i \cap \bar{A}_i \cap \ldots \cap \bar{A}_{i_2})$, $T(i) = P(B_i \cap \bar{A}_i \cap \ldots \cap \bar{A}_{i_2})$, $P(i) = P(i, i_2)$, $i = i_1, \ldots, i_2$. The quantities $P(i_1, i_2)$, as well as the similar quantities

$$Q_1(i_1, i_2) = P(\bar{B}_{i_1} \cap \ldots \cap \bar{B}_{i_2+1}), \quad Q_2(i_1, i_2) = P(\bar{C}_{i_1} \cap \ldots \cap \bar{C}_{i_2+1}),$$

will appear below in the statement of the theorem. We will assume that $P(i_1, i_2)$ (and $Q_1(i_1, i_2)$, $Q_2(i_1, i_2)$) depends only on the size of the interval $[i_1, i_2]$, i.e., for example, $P(i_1, i_2) = P(m, m + i_2 - i_1)$ for all integer $m \in [N_1, N_3 - i_2 + i_1 - 1]$. Let us use this observation to find $P(i_1, i_2)$. To this end, we first, from the available observations $\{x_i\}$, compose the sequence $\{c_i\}$, then compose $\{r_i\}$, and find

$$P(i_1, i_2) = \frac{M}{N_3 - i_2 + i_1 - N_1} \tag{3.7}$$

empirically from the available series $\mathbf{r} = \{r_{N_1}, \ldots, r_{N_3}\}$ by evaluating the number M of elements of the set

$$\{m \in [N_1, N_3 - i_2 + i_1 - 1] : \mathbf{r} \in \bar{A}_i, \quad i = m, \ldots, m + i_2 - i_1\}.$$

We proceed similarly for $Q_1(i_1, i_2)$, $Q_2(i_1, i_2)$. A corresponding example is given in § 5.

Remark 3.1. *Let $\beta_1 = 0$. Given $\mathbf{r} = \{r_{N_1}, \ldots, r_{N_3}\}$, we set*

$$q_i = \frac{1}{2}\Big(1 - \operatorname{sgn}(r_i r_{i+1})\Big) \cdot \min\{|r_i|, |r_{i+1}|\}, \quad i \in [N_1, N_3 - 1],$$

and relate A_i with q_i. If $q_i < R(1-\beta_2)$, then

either $\min\{|r_i|, |r_{i+1}|\} \geq R(1-\beta_2)$ *and* $\operatorname{sgn}(r_i r_{i+1}) = 1$,

or $\min\{|r_i|, |r_{i+1}|\} < R(1-\beta_2)$.

In both cases, $\mathbf{r} \in \bar{A}_i$. If $R(1-\beta_2) \leq q_i$, then $r_i r_{i+1} < 0$ and $R(1-\beta_2) \leq |r_i|, |r_{i+1}|$, i.e., $\mathbf{r} \in A_i$. Hence, the condition $R(1-\beta_2) \leq q_i$ is equivalent to $\mathbf{r} \in A_i$. Hence,

$$P(\bar{A}_{m_1} \cap \ldots \cap \bar{A}_{m_2}) = P\Big(\max_{m_1 \leq i \leq m_2} q_i < R(1-\beta_2)\Big),$$

and so to find $P\Big(\max\limits_{m_1\le i\le m_2} q_i < q\Big)$, *where* $q = R(1-\beta_2)$, *it suffices to construct the empirical distribution function of the random variable*

$$q_s^* = \max_{s\le i\le s+m_2-m_1} q_i,$$

in view of the equality $P(q_s^* < q) = P\Big(\max\limits_{m_1\le i\le m_2} q_i < q\Big)$.

Even though $P(i_1, i_2)$, $Q_m(i_1, i_2)$, $m = 1, 2$, should be evaluated by (3.7) (or similar formulas), the following result provides another interesting method for evaluation of these quantities.

Lemma 3.2. *We set*

$$P\{-R(1-\beta_1) \le \xi \le -R(1-\beta_2)\} = \gamma_1, \quad P\{R(1-\beta_2) \le \xi \le R(1-\beta_1)\} = \gamma_2, \tag{3.8}$$

where $0 \le \beta_1 < \beta_2 < 1$. *Assume that* $0 < \gamma_1 + \gamma_2 < 1$ *and the sequence* $\{r_i\}$ *consists of independent random variables. Then*

$$P(i_1) := P(i_1, i_2) = D_1\lambda_1^{i_2-i_1+2} + D_2\lambda_2^{i_2-i_1+2} + D_3\lambda_3^{i_2-i_1+2}, \tag{3.9}$$

here, λ_i, $i = 1, 2, 3$, *are the roots of the cubic (in* λ*) equation*

$$\begin{aligned}
&(1-\lambda)\lambda^2 + \gamma_1\gamma_2(\gamma_1+\gamma_2-1-\lambda) = 0, \quad 0 < (-\lambda_1), \lambda_2 \le \sqrt{\gamma_1\gamma_2} < \lambda_3 < 1,\\
&D_1 := \frac{\lambda_1^2 - \gamma_1\gamma_2}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} < 0, \quad D_2 := \frac{\lambda_2^2 - \gamma_1\gamma_2}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} \ge 0, \quad D_3 := \frac{\lambda_3^2 - \gamma_1\gamma_2}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} > 0.
\end{aligned}$$

$$Q_s(i_1, i_2) = (1-\gamma_s)^{i_2-i_1+2}, \quad s = 1, 2. \tag{3.10}$$

Proof. Equalities (3.10) follow from independence of r_i. By definition of the events A_i, $i = i_1, \ldots, i_2$,

$$\begin{aligned}
P(i+1) &= P(i) + P(A_i \cap \bar{A}_{i+1} \cap \ldots \cap \bar{A}_{i_2})\\
&= P(i) + P(C_i \cap B_{i+1} \cap \bar{A}_{i+1} \cap \ldots \cap \bar{A}_{i_2}) + P(C_{i+1} \cap B_i \cap \bar{A}_{i+1} \cap \ldots \cap \bar{A}_{i_2})\\
&= P(i) + \gamma_2 T(i+1) + \gamma_1 S(i+1).
\end{aligned}$$

We also have

$$\begin{aligned}
\gamma_2 P(i+1) &= P(C_i \cap (A_i \cup \bar{A}_i) \cap \bar{A}_{i+1} \cap \ldots \cap \bar{A}_{i_2})\\
&= S(i) + \gamma_2 P(B_{i+1} \cap \bar{A}_{i+1} \cap \ldots \cap \bar{A}_{i_2}) = S(i) + \gamma_2 T(i+1)
\end{aligned}$$

and

$$\begin{aligned}
\gamma_1 P(i+1) &= P(B_i \cap (A_i \cup \bar{A}_i) \cap \bar{A}_{i+1} \cap \ldots \cap \bar{A}_{i_2})\\
&= T(i) + \gamma_1 P(C_{i+1} \cap \bar{A}_{i+1} \cap \ldots \cap \bar{A}_{i_2}) = T(i) + \gamma_1 S(i+1).
\end{aligned}$$

Hence

$$\begin{pmatrix} P(i) \\ S(i) \\ T(i) \end{pmatrix} = \Gamma \cdot \begin{pmatrix} P(i+1) \\ S(i+1) \\ T(i+1) \end{pmatrix}, \quad \Gamma = \begin{pmatrix} 1 & -\gamma_1 & -\gamma_2 \\ \gamma_2 & 0 & -\gamma_2 \\ \gamma_1 & -\gamma_1 & 0 \end{pmatrix}.$$

The solution of this recurrence relation can be written as the sum

$$d_1 Z_1 \lambda_1^{i_2-i} + d_2 Z_2 \lambda_2^{i_2-i} + d_3 Z_3 \lambda_3^{i_2-i} \tag{3.11}$$

in eigenvectors Z_1, Z_2, Z_3 of the matrix Γ with some coefficients d_1, d_2, d_3, λ_m is the eigenvalue corresponding to the eigenvector Z_m, $m = 1, 2, 3$. The characteristic polynomial of Γ is as follows:

$$\varphi(\lambda) = (1-\lambda)\lambda^2 + \gamma_1^2\gamma_2 + \gamma_1\gamma_2^2 - \gamma_1\gamma_2(\lambda + \lambda + 1 - \lambda) = (1-\lambda)\lambda^2 + \gamma_1\gamma_2(\gamma_1 + \gamma_2 - 1 - \lambda).$$

Since $0 < \gamma_1 + \gamma_2 < 1$, we have $\varphi(0) < 0$. Further $\varphi(-\sqrt{\gamma_1\gamma_2}) \geq 0$, and hence one of the roots (say, λ_1) of the characteristic polynomial lies in the interval $[-\sqrt{\gamma_1\gamma_2}, 0)$. In addition, $\varphi(1) < 0$, and, for $\lambda = \sqrt{\gamma_1\gamma_2} \in (0, 1)$,

$$\varphi(\sqrt{\gamma_1\gamma_2}) = \gamma_1\gamma_2\left(\gamma_1 + \gamma_2 - 2\sqrt{\gamma_1\gamma_2}\right) \geq 0.$$

Hence, for $\gamma_1 \neq \gamma_2$, the equation $\varphi(\lambda) = 0$ has two distinct roots $\lambda_2 < \sqrt{\gamma_1\gamma_2} < \lambda_3$ on $(0, 1)$. If $\gamma_1 = \gamma_2 = \gamma$, then $\varphi(\gamma) = 0$,

$$\begin{aligned} \varphi(\lambda) = (1-\lambda)\lambda^2 + \gamma^2(2\gamma - 1 - \lambda) = -(\lambda - \gamma)(2\gamma^2 - \gamma + (\gamma - 1)\lambda + \lambda^2), \\ -(2\gamma^2 - \gamma + (\gamma - 1)\gamma + \gamma^2) = -2\gamma(2\gamma - 1) > 0, \end{aligned}$$

and, as above, the equation $\varphi(\lambda) = 0$ has two distinct roots $\lambda_2 = \sqrt{\gamma_1\gamma_2} < \lambda_3 < 1$ on $(0, 1)$.

We can assume that

$$Z_m = \begin{pmatrix} -\gamma_1\gamma_2 + \lambda_m^2 \\ -\gamma_1\gamma_2 + \gamma_2\lambda_m \\ -\gamma_1\gamma_2 + \gamma_1\lambda_m \end{pmatrix}, \quad m = 1, 2, 3.$$

Let us find d_1, d_2, d_3 from (3.11). For $i = i_2$, we have

$$\begin{aligned} &P(i_2) = P(\bar{A}_{i_2}) = 1 - 2\gamma_1\gamma_2, \\ &\gamma_2 = P(C_{i_2} \cap (A_{i_2} \cup \bar{A}_{i_2})) = S(i_2) + \gamma_2\gamma_1, \\ &\gamma_1 = T(i_2) + \gamma_1\gamma_2, \end{aligned}$$

and so, for d_1, d_2, d_3, we have the system

$$d_1Z_1 + d_2Z_2 + d_3Z_3 = \begin{pmatrix} 1 - 2\gamma_1\gamma_2 \\ \gamma_2 - \gamma_1\gamma_2 \\ \gamma_1 - \gamma_1\gamma_2 \end{pmatrix}.$$

Consequently,

$$d_1 = \frac{\lambda_1^2}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)}, \quad d_2 = \frac{\lambda_2^2}{(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3)}, \quad d_3 = \frac{\lambda_3^2}{(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)}. \tag{3.12}$$

Let us substitute d_1, d_2, d_3 into (3.11). The first coordinate of the resulting vector is $P(i_1)$, which is the right-hand side of (3.9). □

We will choose β_1, β_2 so that $P(i_1, i_2)$ (and $Q_s(i_1, i_2)$, $s = 1, 2$) will be small, for example, < 0.1. Then the probability of the complementary event is at least 0.9, and hence, the event A_{i_0} (B_{i_0} or C_{i_0}) holds, with this probability, for some $i_0 \in [i_1, i_2]$ ($i_0 \in [i_1, i_2 + 1]$).

Theorem 3.3. *Let $0 \leq \beta_1 < \beta_2 < 1$, and let $P(i_1, i_2)$ be the probability of the event $\{\bar{A}_{i_1} \cap \ldots \cap \bar{A}_{i_2}\}$.*

1) If $a_i \geq \alpha > 0$ for all $i \in [i_1, i_2] \subset [N_1, N_2]$, $\beta_1 = 0$, $\beta := \beta_2$, then

$$\max_{i \in [i_1, i_2 + 1]} |b_i| + O(\tau) \geq \frac{(1+\alpha)(1-\beta)}{2 - \beta(1+\alpha)} \max_{i \in [i_1, i_2 + 1]} |c_i| \tag{3.13}$$

with probability $1 - P(i_1, i_2)$;

2) *if all* c_i, $i_1 \le i \le i_2+1$, *are of the same sign (in particular, if* $a_i > 0$ *for all* $i \in [i_1, i_2] \subset [N_1, N_2]$*), and* $s = 1 \vee 2$ *is such that* $(-1)^{s+1} c_i \ge 0$, $\beta_1 = 0$, $\beta := \beta_2$, *then*

$$\max_{i \in [i_1, i_2+1]} |b_i| \ge \frac{1-\beta}{2-\beta} \max_{i \in [i_1, i_2+1]} |c_i| + \frac{\sigma}{2-\beta}, \quad \sigma := \min_{i \in [i_1, i_2+1]} |c_i|. \tag{3.14}$$

with probability $1 - Q_s(i_1, i_2)$;

3) *if* $\alpha < 0$, $a_i \ge \alpha$ *for all* $i \in [i_1, i_2] \subset [N_1, N_2]$, $0 \le \beta_1 < \beta_2 < 1$, $(1 - \beta_2 + \alpha(1 - \beta_1)) > 0$, *then*

$$\max_{i \in [i_1, i_2+1]} |b_i| \ge \frac{1 - \beta_2 + \alpha - \alpha\beta_1}{2 - \beta_2 - \alpha\beta_1} \max_{i \in [i_1, i_2+1]} |c_i|. \tag{3.15}$$

with probability $1 - P(i_1, i_2)$.

Proof. Let us verify 1). If the condition $(a_i - 1) > -1 + \alpha$ ($\Leftrightarrow$ $a_i > \alpha$) is satisfied on some interval of variation of the index $i \in [i_1, i_2] \subset [N_1, N_3 - 1]$, $\alpha > 0$, then, for all $i \in [i_1, i_2 + 1]$, $c_i = b_i + r_i$ are also of the same sign. In addition, assume, for example, that all $c_i = b_i + r_i$ are nonnegative. For each $i \in [i_1, i_2]$, we have two possibilities:

either $b_{i+1} + r_{i+1} \ge b_i + r_i \ge \alpha(b_{i+1} + r_{i+1})$,

or, vice versa, $b_i + r_i \ge b_{i+1} + r_{i+1} \ge \alpha(b_i + r_i)$

and, hence, in any case,

$$b_i + r_i \ge \alpha(b_{i+1} + r_{i+1}), \quad b_{i+1} + r_{i+1} \ge \alpha(b_i + r_i). \tag{3.16}$$

Assume that one of the events A_{i_0} for some fixed $i_0 \in \{i_1, \dots, i_2\}$ is satisfied. Then from the inequalities (3.16) we have

$$b_{i_0}(1-\alpha) + O(\tau) \ge \alpha r_{i_0+1} - r_{i_0}, \quad b_{i_0+1}(1-\alpha) + O(\tau) \ge \alpha r_{i_0} - r_{i_0+1}$$

and therefore,

$$(1+\alpha) \cdot R \cdot (1-\beta) \le (1-\alpha) \cdot \max\{b_{i_0}, b_{i_0+1}\} + O(\tau), \quad R \cdot (1-\beta) \le \frac{1-\alpha}{1+\alpha} \max_{i \in [i_1, i_2+1]} b_i + O(\tau).$$

Hence

$$\max_{i \in [i_1, i_2+1]} |c_i| \le \max_{i \in [i_1, i_2+1]} |b_i| + R \le \Big(1 + \frac{1-\alpha}{(1+\alpha)(1-\beta)}\Big) \max_{i \in [i_1, i_2+1]} |b_i| + O(\tau),$$

which implies the required inequality (3.13). This estimate holds with probability $1 - P(i_1, i_2)$.

Let us prove assertion 3). Assume now that on some interval of variation of $i \in [i_1, i_2] \subset [N_1, N_2]$, the condition $a_i > \alpha$ is satisfied and $\alpha < 0$, $0 \le \beta_1 < \beta_2 < 1$, $(1 - \beta_2 + \alpha(1 - \beta_1)) > 0$. Assume also that some event A_{i_0}, $i_0 \in [i_1, i_2]$ is satisfied. Hence, if c_{i_0} and c_{i_0+1} are of the same sign (for example, these values are negative), then

$$b_{i_0} \le -r_{i_0}, \quad b_{i_0+1} \le -r_{i_0+1}, \quad |r_{i_0}| \ge R \cdot (1-\beta_2), \ |r_{i_0+1}| \ge R \cdot (1-\beta_2), \ r_{i_0} r_{i_0+1} < 0,$$

that is,

$$\max_{i \in [i_1, i_2+1]} |b_i| \ge \max\{|b_{i_0}|, |b_{i_0+1}|\} \ge R \cdot (1-\beta_2). \tag{3.17}$$

We have the same result also for positive c_{i_0} and c_{i_0+1}. If c_{i_0} and c_{i_0+1} are of different sign (for example, $c_{i_0} > 0$ and $c_{i_0+1} < 0$), then we consider two possible cases.

I) Let $r_{i_0} \le 0$, $r_{i_0+1} \ge 0$. Then

$$0 < c_{i_0} = b_{i_0} + r_{i_0}, \quad c_{i_0+1} = b_{i_0+1} + r_{i_0+1} < 0, \quad |r_{i_0}|, |r_{i_0+1}| \ge R \cdot (1-\beta_2),$$

and so (3.17) holds.

II) Let $r_{i_0} > 0$, $r_{i_0+1} < 0$. By definition of a_{i_0} (see (2.9)), it follows that either $0 < c_{i_0} = b_{i_0} + r_{i_0} < \alpha c_{i_0+1} = \alpha(b_{i_0+1} + r_{i_0+1})$ or $\alpha c_{i_0} = \alpha(b_{i_0} + r_{i_0}) < c_{i_0+1} = b_{i_0+1} + r_{i_0+1} < 0$.

Hence, we have either a) $-b_{i_0} + \alpha b_{i_0+1} > r_{i_0} - \alpha r_{i_0+1}$, or b) $\alpha r_{i_0} - r_{i_0+1} < b_{i_0+1} - \alpha b_{i_0}$. The inequality in case a) can be rewritten as

$$(b_{i_0+1} - b_{i_0}) + (-1+\alpha)b_{i_0+1} > r_{i_0} - \alpha r_{i_0+1} \quad \text{or} \quad \alpha(b_{i_0+1} - b_{i_0}) + (-1+\alpha)b_{i_0} > r_{i_0} - \alpha r_{i_0+1}.$$

We have $\alpha < 0$, and so one of the numbers $(b_{i_0+1} - b_{i_0})$ and $\alpha(b_{i_0+1} - b_{i_0})$ is negative. Hence, so that one of the inequalities $(-1+\alpha)b_{i_0+1} > r_{i_0} - \alpha r_{i_0+1}$ or $(-1+\alpha)b_{i_0} > r_{i_0} - \alpha r_{i_0+1}$ holds, and, therefore,

$$(1-\alpha)\max_{i\in[i_1,i_2+1]}|b_i| \geq R\cdot(1-\beta_2) - \alpha(-R(1-\beta_1)) = R\cdot(1-\beta_2+\alpha(1-\beta_1)). \tag{3.18}$$

The inequality in case b) can be rewritten as

$$\alpha(b_{i_0+1} - b_{i_0}) + (1-\alpha)b_{i_0+1} > \alpha r_{i_0} - r_{i_0+1} \quad \text{or} \quad (b_{i_0+1} - b_{i_0}) + (1-\alpha)b_{i_0} > \alpha r_{i_0} - r_{i_0+1},$$

and so, at least one of the inequalities $(1-\alpha)b_{i_0+1} > \alpha r_{i_0} - r_{i_0+1}$ or $(1-\alpha)b_{i_0} > \alpha r_{i_0} - r_{i_0+1}$ holds, and, therefore, estimate (3.18). The remaining case of signs $c_{i_0} < 0$ and $c_{i_0+1} > 0$ is investigated similarly and also gives (3.18).

Hence, if, for all $i \in [i_1, i_2] \subset [N_1, N_3 - 1]$, $(a_i - 1) > -1 + \alpha$ and $\alpha < 0$, $(1 - \beta_2 + \alpha(1-\beta_1)) > 0$, then

$$\max_{i\in[i_1,i_2+1]}|c_i| \leq \max_{i\in[i_1,i_2+1]}|b_i| + R \leq \max_{i\in[i_1,i_2+1]}|b_i|\Big(1 + \frac{1-\alpha}{1-\beta_2+\alpha(1-\beta_1)}\Big),$$

$$\max_{i\in[i_1,i_2+1]}|b_i| \geq \frac{1-\beta_2+\alpha-\alpha\beta_1}{2-\beta_2-\alpha\beta_1}\max_{i\in[i_1,i_2]}|c_i|. \tag{3.19}$$

This estimate also holds with probability $1 - P(i_1, i_2)$.

Let us prove assertion 2). Assume that all $c_i = b_i + r_i$ are nonnegative for all $i \in [i_1, i_2 + 1]$. Hence, for $i \in [i_1, i_2 + 1]$,

$$b_i + r_i \geq \sigma := \min_{z\in[i_1,i_2+1]}|c_z|, \tag{3.20}$$

$$1 - Q_1(i_1, i_2) = P(B_{i_1} \cup \ldots \cup B_{i_2+1}) = P\Big(\min_{i\in[i_1,i_2+1]} r_i \leq -R(1-\beta)\Big) \leq P\Big(R(1-\beta) \leq \max_{i\in[i_1,i_2+1]} b_i - \sigma\Big),$$

i.e., for all $z \in [i_1, i_2 + 1]$,

$$|c_z| \leq |b_z| + |r_z| \leq |b_z| + (\max_{i\in[i_1,i_2+1]}|b_i| - \sigma)/(1-\beta),$$

$$\max_{i\in[i_1,i_2+1]}|b_i| \geq \frac{1-\beta}{2-\beta}\Big(\max_{i\in[i_1,i_2+1]}|c_i| + \frac{\sigma}{1-\beta}\Big)$$

with probability $\geq 1 - Q_1(i_1, i_2)$. If $c_i \leq 0$, $i \in [i_1, i_2 + 1]$, then the above arguments apply with c_i replaced by $(-c_i)$. □

As a result, if, for some $\alpha \in (-1, 1)$ and for "sufficiently" large interval $[i_1, i_2]$, we have $a_i > \alpha$, $i \in [i_1, i_2]$, we may expect that the estimate

$$\max_{i\in[i_1,i_2]}|b_i| \geq \frac{1+\alpha}{2}\max_{i\in[i_1,i_2]}|c_i| \tag{3.21}$$

holds approximately.

Each of the estimates from Theorem 3.3 is a lower estimate for the maximum of the absolute value of the $(n+1)$st probability (with the corresponding probability). Based on these results, we give some recommendations for choosing Δ from (1.5) for use in formulas for derivative recovery (like (1.7), (1.8)):

– for each step j we define the sequences $\{c_i\} = \{c_{i,j}\}$ and $\{a_i\} = \{a_{i,j}\}$,

– from $\{c_i\}$ one defines the sequence $\{r_i\}$ for distortions associated with evaluation of difference relations; from $\{r_i\}$ one constructs the empirical distribution function of the random variable ξ (see (3.2)),

– using $\{a_i\}$, one chooses α and $[i_1, i_2]$ to satisfy (3.4),

– for given $\{j, i_1, i_2, \alpha\}$, the right-hand side of (3.21) can be used as a lower estimate for the norm of the derivative, because as a rule this estimate is the correct one. We maximize this right-hand side over all $\{j, i_1, i_2, \alpha\}$ and take the resulting value as Δ in recovery formulas of the form (1.7) or (1.8). Here, it is important to exclude the interval $[i_1, i_2]$ of small length,

– in order to use Theorem 3.3, we choose β_1, β_2, and then evaluate $P(i_1, i_2)$ (or $Q_s(i_1, i_2)$, $s = 1, 2$). Next, if $P(n_1, n_2)$ is not small (say, > 0.1), then one should increase β_1, decrease β_2, and then repeat these calculations,

– if $P(n_1, n_2)$ (or $Q_s(i_1, i_2)$, $s = 1, 2$) is small (say, < 0.1), then we use the right-hand of (3.13), (3.14) or (3.15) (depending on α) to estimate from below the norm of the derivative,

– as above, we maximize the norm of the derivative over all $\{j, i_1, i_2, \alpha\}$, and take as Δ in recovery formulas of the form (1.7) or (1.8).

So, regardless of whether the step was acceptable or not, if (3.1) is satisfied on some visible interval (possibly of small length) of variation of the argument i, then, by evaluating the maximum of $|c_i|$ along this interval, we obtain a lower estimate for $\|x^{(n+1)}\|$ (as a corollary of the Theorem).

If there are no intervals on the entire interval $[N_1, N_2]$ of consideration, and if the step j remains acceptable, then the value

$$\max_i |c_i| \tag{3.22}$$

can be looked upon (up to a small factor > 1) as one of the upper estimates for $\|x^{(n+1)}\|$ (note that the maximum in (3.22) can differ substantially from $\|x^{(n+1)}\|$ (exceed it considerably)).

If, in the course of the above calculations, it was found that the upper and lower estimates of the derivative differ by a not very large factor and if the "acceptability" conditions for the size of the step j are satisfied for the upper estimate of the derivative, then in what follows one may use some average of these two estimates as a value of $\|x^{(n+1)}\|$ — this not quite accurate replacement of $\|x^{(n+1)}\|$ by a quantity that differs from it by a not very large factor will not change much the accuracy of recovery of the derivative $x^{(n)}$: here it is important that the error should not differ by a very large factor.

If, in the course of all calculations, one succeeds in obtaining a lower estimate for the derivative $x^{(n+1)}$ only for a very large step j, then as the value of $\|x^{(n+1)}\|$ one may use any of the upper estimates (3.22) — in this case, the quantity $\|x^{(n+1)}\|$ remains unknown, but with an available upper estimate. The same can be said in this case also about the recovery error of the derivative.

In § 4, we give an example of estimation of the uniform norm of the 3rd derivative. To this end, from a given function (a polynomial of degree 5), we form a sequence of its values on a uniform grid consisting of 1485 knots. This sequence is augmented with a random perturbing sequence. Next, using inequalities (3.21), we estimate the absolute value of the modulus of the 3rd derivative. This value is compared with the exact value of the norm of the 3rd derivative of the original function (a polynomial of degree 5).

In § 5, we consider the same function on the same grid, but in this case the perturbed values of the function are obtained using random variables.

4. Model case for the errors $\{\delta_i\}$ obtained from experimental data

In this section, we consider the errors of values of a function from a real experiment. In this case, we find estimate (3.21) of the norm of the 3rd derivative for a model function (a given polynomial).

Assume that as a result of an experiment we have obtained approximate values $\tilde{x}(t_i)$ of some function $x(\cdot)$ at some points $\{t_i\}$ from its domain. We also assume that the same random factors were acting in the course of the experiment and lead to the error in the function. We choose some value of the parameter k_0, and, from the approximate values of $\tilde{x}$, we construct the new sequence $\{\widetilde{x}(t_i) - \frac{1}{2k_0+1}\sum_{k=-k_0}^{k_0}\widetilde{x}(t_{i+k})\}$ such that

$$\widetilde{x}(t_i) - \frac{1}{2k_0+1}\sum_{k=-k_0}^{k_0}\widetilde{x}(t_{i+k}) = \frac{2}{2k_0+1}\sum_{k=1}^{k_0}\Big(\widetilde{x}(t_i) - \frac{\widetilde{x}(t_{i+k}) + \widetilde{x}(t_{i-k})}{2}\Big)$$

$$= \frac{2}{2k_0+1}\sum_{k=1}^{k_0}\Big(\delta(t_i) - \frac{\delta(t_{i+k}) + \delta(t_{i-k})}{2}\Big) + O(\tau^2) = \delta(t_i) - \frac{1}{2k_0+1}\sum_{k=-k_0}^{k_0}\delta(t_{i+k}) + O(\tau^2), \tag{4.1}$$

because $\Big(x(t_i) - \frac{x(t_{i+k})+x(t_{i-k})}{2}\Big) = O(\tau^2)$ (here, k_0 is a small parameter controlling the number of points around t_i which are involved in the evaluation of the mean). From the assumption that $\delta(t_i)$ are observations of some random variable with zero mean, one may assume that the values in (4.1) are approximately equal to $\delta(t_i)$ even for small k_0.

Taking this into account, we consider the sequence thus constructed as a random perturbation $\{\delta(t_i)\}$, which affected $x(\cdot)$ during measurements at the points $\{t_i\}$. We add this perturbation to the model function at the points $\{t_i\}$. As a model function, we take the polynomial p, and will study the error of recovery of the norm of the leading derivative on this sequence $\{\widetilde{p}(t_i)\}$. The polynomial p is subject to the only condition that, on the domain of the function x, the range of p would be approximately equal to the smallest interval containing all values of the sequence $\tilde{x}(t_i)$. Let us give two graphs of the difference (4.1) for $k_0 = 2$ and $k_0 = 5$ (see Fig. 2).

Figure 2: The difference (4.1) for $k_0 = 2$ and $k_0 = 5$.

These dependences were obtained for the graphs of experimental data for $\widetilde{x}(t_i)$ (see Fig. 3):

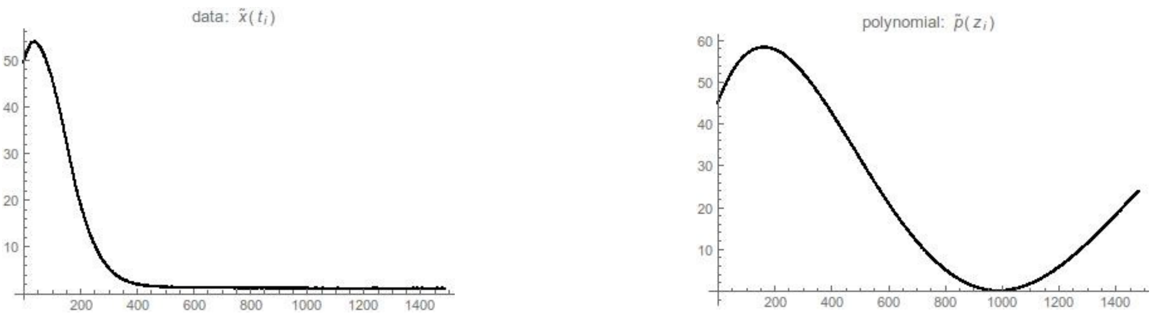

Figure 3: The original data and the model function.

Using the example of the sequence $\widetilde{p}(z_i)$, $i = 0, \dots, 1484$, (see Fig. 3) constructed from the polynomial

$$p(z) = 45\left((z-1)^5 - 2(z-1)^4 - (z-1)^3 + 3(z-1)^2\right), \quad z \in [0, 1.5], \tag{4.2}$$

and which has similar characteristics with the sequence $\widetilde{x}(t_i)$, let us see how the estimate of the norm of the third derivative depends on the choice of the interval in (3.4) (by a characteristic of a sequence we will mean the difference between its greatest and smallest values; in this example, this difference is approximately 60).

In Fig. 4, we show the graphs of the analogue of the sequence $(a_i - 1)$, which is constructed for $\{\widetilde{p}(t_i)\}$ and for $j = 5, 40, 60, 80, 100, 140$.

Remark 4.1. *The successive differences from* (2.1), (2.2) *can be naturally looked upon as the derivative of order* $(n+1)$ *taken at the central point of the interval* $[t_i, t_{i+(n+1)j}]$ *(due to symmetry considerations, the order of smallness the difference between the precise value of the derivative at this point and* b_i *is greater by one that the difference between the derivative at the neighboring points and* b_i*). To get values with increased order of accuracy not only at the central point* $[t_i, t_{i+(n+1)j}]$*, but also at different points (see* (4.3)*), one should replace the successive differences from* (2.1), (2.2) *by different one. So, if one considers the point* $t + m\tau$ *of the interval* $[t, t+4\tau]$ *in the problem of approximation of the third derivative, one should replace the successive difference used in* (2.1), (2.2) *by*

$$(m - 5/2)x(t) + (9 - 4m)x(t+\tau) + 6(m-2)x(t+2\tau) + (7-4m)x(t+3\tau) + (m - 3/2)x(t+4\tau). \qquad (4.3)$$

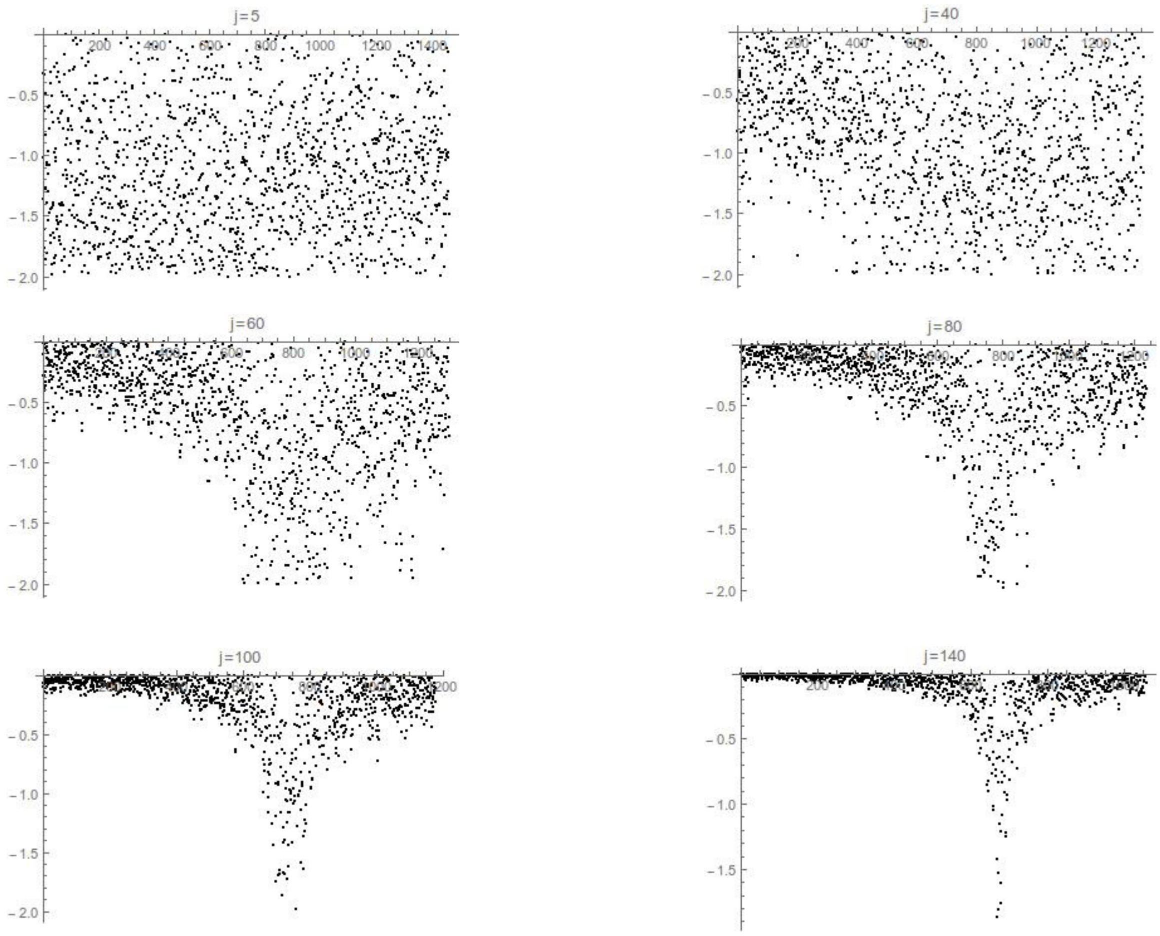

Figure 4: The sequence $(a_i - 1)$ for $\widetilde{p}(z_i)$, for $j = 5, 40, 60, 80, 100, 140$.

In the graph of the dependence $(a_i - 1)$ on i (Fig. 4), which corresponds to $j = 40$, one can single out the interval $i \in [55, 201]$, on which the quantities $(a_i - 1)$, as evaluated for $\widetilde{p}(z_i)$, exceed $(-1 - 0.4)$, i.e., for this interval, one may take the parameter $\alpha = -0.4$. By calculating the largest value among $|c_i|$, $i \in [55, 202]$ and finding the difference relation corresponding to this value, and after multiplying this relation by $\frac{1+\alpha}{2} = 0.3$ (in accordance with (3.21)), we get a lower estimate for the norm of the third derivative, which is approximately 2633.24. If on the interval $i \in [55, 202]$ one evaluates the maximum of the absolute value of the third order difference relations for precisely given values of the polynomial $p(\cdot)$, we get approximately 3740.17. For the function $p(\cdot)$ itself, the maximum of the absolute value of the third derivative is attained at zero and is equal to 4590 (on the interval $[0, 0.8]$ the absolute value decreases). So, we have approximately obtained a lower estimate for the third derivative of the function on the interval $i \in [0, 1359]$, which (see (4.2)) for the chosen values $k_0 = 2$, $j = 40$ corresponds to the interval $z \in [0.06, 1.44] \subset [0, 1.5]$: the value of the third central difference for $i = 0$ with step 40τ corresponds approximately to the third derivative at the central point of the interval $[0, 120\tau]$, i.e., approximately at the point 0.06, and, similarly, for the rightmost possible value $i = 1359 = 1484 - 3j - 2k_0 - 1$ (at the point 1.44). The absolute value of the third derivative of the function $p(\cdot)$ at the point $z = 0.06$ is 4146.12.

With $j = 60$, similar calculations give $i \in [0, 122]$, $\alpha = 0.35$, $\frac{1+\alpha}{2} = 0.675$; the lower estimate from approximate (exact) values of the polynomial is 3434.02 (3925.8, respectively). The interval on which the calculations were carried out is $i \in [0, 1299]$, which corresponds to $z \in [0.09, 1.41] \subset [0, 1.5]$. The absolute value of the third derivative at the point $z = 0.09$ is 3931.47.

For other values of the parameter j, the results of estimation of the derivative are as follows:

$j = 80$, for $i \in [0, 332]$, $\alpha = 0.55$, $\frac{1+\alpha}{2} = 0.775$, the lower estimate from approximate (exact) values of the polynomial is 3561.57 (3715.48, respectively); the calculations were carried out for $i \in [0, 1239]$, which corresponds to $z \in [0.12, 1.38] \subset [0, 1.5]$. The absolute value of the derivative for $z = 0.12$ is 3721.68.

$j = 100$, for $i \in [0, 320]$, $\alpha = 0.82$, $\frac{1+\alpha}{2} = 0.91$, the estimate from approximate (exact) values of the polynomial is 3504.55 (3510.69, respectively); the calculations were carried out for $i \in [0, 1179]$, which corresponds to $z \in [0.15, 1.35] \subset [0, 1.5]$. The absolute value of the derivative for $z = 0.15$ is 3516.75.

$j = 140$, for $i \in [0, 148]$, $\alpha = 0.93$, $\frac{1+\alpha}{2} = 0.965$, the estimate from approximate (exact) values of the polynomial is 3042.28 (3117.75, respectively) the calculations were carried out for $i \in [0, 1059]$, which corresponds to $z \in [0.21, 1.29] \subset [0, 1.5]$. The absolute value of the derivative for $z = 0.21$ is 3121.47.

The above analysis was carried out in the case where the approximate values of the derivative were calculated via the central difference relations. As a result, the analysis did not cover the points near the end-points of $[0, 1.5]$. To analyse this case also, we replace everywhere the central difference by the shifted difference from (4.3).

Figure 5 shows the analogues of the sequence $(a_i - 1)$ obtained if the successive difference is replaced by the shifted difference (4.3) for $m = 0$ for the values j of the step $j\tau$ ($j = 60, 80, 100, 140$):

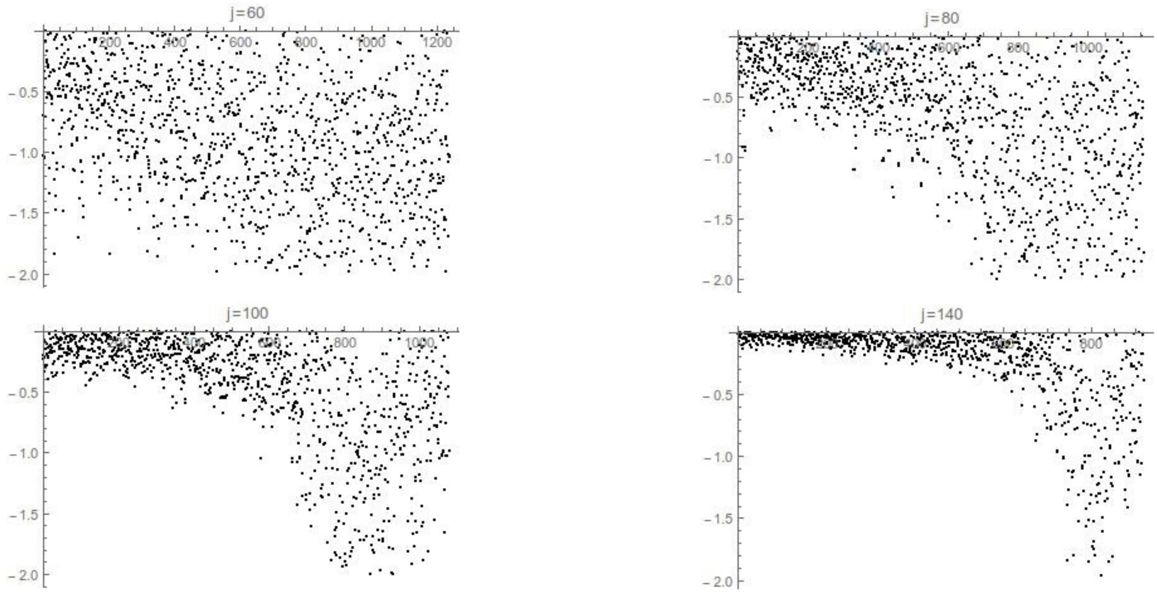

Figure 5: The sequence $(a_i - 1)$ for the shifted difference (4.3).

For the first graph constructed for $j = 60$ (see Fig. 5): on the interval $i \in [0, 32]$, $\alpha = -0.33$, $\frac{1+\alpha}{2} = 0.335$, the lower estimate from approximately (precisely) given values of the polynomial is 2947.85 (4555.1, respectively). For the interval $i \in [63, 108]$, we get $\alpha = -0.16$, $\frac{1+\alpha}{2} = 0.42$, and the lower estimate, as obtained from approximately given values, is 4641.13, which exceeds $\|p\| = 4590$. The range on which this estimate was conducted is the interval $i \in [0, 1239]$, which corresponds to $z \in [0, 1.25] \subset [0, 1.5]$.

For $j = 80$, on the interval $i \in [0, 331]$, $\alpha = 0.05$, $\frac{1+\alpha}{2} = 0.5025$. The lower estimate from approximate (exact) values is 4268.75 (4527.96, respectively). The range on which this estimate was conducted is the interval $i \in [0, 1159]$, which corresponds to $z \in [0, 1.17] \subset [0, 1.5]$.

For $j = 100$, on the interval $i \in [0, 343]$, $\alpha = 0.53$, $\frac{1+\alpha}{2} = 0.765$. The lower estimate from approximate (exact) values is 4444.94 (4493.06, respectively). The range on which this estimate was conducted is the interval $i \in [0, 1079]$, which corresponds to $z \in [0, 1.09] \subset [0, 1.5]$.

5. Estimating the norm of the leading derivative for random distribution of divided differences $\{r_i\}$

In this section, we consider the same model function (a polynomial) as in § 4 subject to the same noise. The norm of the leading derivative is estimated via Theorem 3.3. We search for a sufficiently large interval $[i_1, i_2] \ni i$ on which the infimum $\alpha - 1$ of $a_i - 1$ (see (2.9)) is greatest. As distinct from § 4, here we will use the parameters β_1, β_2 chosen from the condition that the estimate from Theorem 3.3 would be satisfied with probability ≥ 0.9. It turns out, that the estimate from § 4 is sufficiently close to that obtained in this section.

In order to use Theorem 3.3 to estimate the norm of the leading derivative, we find approximative values of the events considered in § 3. As in § 4, c_i is the sum of b_i (the value of a given smooth function at t_i) and r_i (the ith observation of the random variable ξ which is independent of i and having zero mean). In analogy with § 4, we single out the random component r_i that specifies the distortions involved in evaluations of b_i: for a fixed j, consider the difference between c_i and some average value over these values:

$$c_i - \frac{1}{2k_0 + 1} \sum_{k=-k_0}^{k_0} c_{i+k} = (b_i + r_i) - \frac{1}{2k_0 + 1} \sum_{k=-k_0}^{k_0} (b_{i+k} + r_{i+k}) = r_i - \frac{1}{2k_0 + 1} \sum_{k=-k_0}^{k_0} r_{i+k} + O(\tau^2) \tag{5.1}$$

(see also (4.1)), which is approximately r_i.

In Fig. 6 we show the graph of the empirical distribution function for ξ from the observations r_i of this functions, which are calculated by (5.1) with $j = 40$.

For the values of i close to 1, a small effect of quantities from $O(\tau^2)$ from (5.1) is manifested, and hence the distribution function was evaluated for the indexes starting from $i = 100$ (see Fig. 6).

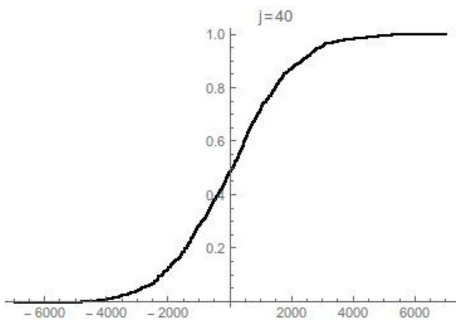

Figure 6: The distribution function of r_i.

In this case, R is approximately 5372. Let us use this function for estimates (3.15) for the model function $\widetilde{p}$. Let $i \in [55, 201]$ and $\alpha = -0.4$ (see p. 7963). To use Theorem 3.3, we specify the desired value of the empirical probability $1 - P(i_1, i_2)$ (for example, 0.9). Then, for each $0 \leq \beta_1 < \beta_2 < 1$, from the resulting distribution function we can find γ_1 and γ_2, etc., and further, using Lemma 3.2, evaluate $1 - P(i_1, i_2)$. This quantity grows with increasing β_2 (or with decreasing β_1). Hence, for each sufficiently large β_2, there is β_1 such that $1 - P(i_1, i_2) \approx 0.9$. For such β_2, β_1, let us estimate the derivative from the right-hand side of (3.15). By varying β_2 we find the greatest possible estimate of the norm of the derivative. Let $\beta_1 \approx 0.0.427$, $\beta_2 \approx 0.63$. Then $\gamma_1 \approx 0.0919905$, $\gamma_2 \approx 0.0911975$, and (3.19) holds with empirical probability $\approx 90\%$. In this case, inequality (3.19) is the estimate for the norm of the 3rd derivative:

$$\max_{i\in[55,202]} |b_i| \geq \frac{1 - \beta_2 + \alpha - \alpha\beta_1}{2 - \beta_2 - \alpha\beta_1} \max_{i\in[55,202]} |c_i| \approx 802.09.$$

This estimate is obtained assuming that $\{r_i\}$ are independent. If $1 - P(i_1, i_2)$ is evaluated via (3.7) without this independence assumption, we can chose $\beta_1 \approx 0.385$, $\beta_2 \approx 0.585$, which gives

$$\max_{i\in[55,202]} |b_i| \geq 945.44.$$

If $[55, 201]$ is reduced to $[55, 113] \ni i$, then we can chose $\alpha = 0$, and σ from (3.14) is also 0. In this case, we can apply estimate (3.14). Since $c_i \geq 0$ for $i \in [55, 114]$, we set $\beta = 0.42$ in Theorem 3.3, 2). Then, assuming that $\{r_i\}$ are independent, we have, with probability 0.9,

$$\max_{i\in[55,114]} |b_i| \geq \frac{1-\beta}{2-\beta} \cdot \max_{i\in[55,114]} |c_i| \approx 3222.11\,.$$

If $1 - Q_1(i_1, i_2)$ is evaluated without this independence assumption, then by a formula similar to (3.7), we have $\beta = 0.435$, which gives a slightly smaller estimate for $\max_{i\in[55,114]} |b_i|$.

Remark 5.1. *To refine $P(i_1, i_2)$, $Q_1(i_1, i_2)$, $Q_2(i_1, i_2)$, which were calculated by* (3.7) *(and similar formulas), we used not only $\{r_{i,k^*}\}$, which were obtained from $\{x(t_i, s_k)\}_i$ for some $k = k^*$, but also for all k (see the introduction).*

We repeat the evaluations of the probability for $j = 60$ (the central difference case) with the use of inequality (3.14). For the empirical distribution function r_i, we find $R = 2010$. On the interval $i \in [0, 436]$, we have $\alpha = 0.213$, $\sigma = 451.38$; for $i \in [0, 347]$, we have $\alpha = 0.257$ and $\sigma = 667.775$. In order that the probability of the estimate would exceed 0.9 (for independent $\{r_i\}$), we take $\beta = 0.175$ in the first case, which gives the estimate ≥ 2547.13 for the norm 3rd derivative, and in the second case, we get $\beta = 0.2075$ and the estimate ≥ 2621.79. Evaluating the probability by (3.7), we get $\beta = 0.4351$ and $\beta = 0.4628$, respectively, in these two cases.

For the shifted difference, for $m = 0$ and $j = 60$, following § 4, we chose $i \in [63, 108]$ with $\alpha = -0.16$. Let $\beta_1 \approx 0.418$, $\beta_2 \approx 0.749$. Then, assuming that $\{r_i\}$ are independent, we get $\gamma_1 \approx 0.158033$, $\gamma_2 \approx 0.175593$, and with probability ≈ 0.9, using (3.15), we obtain the lower estimate 1323.81 for the norm of the 3rd derivative. If, for example, $\beta_1 = 0.0$, $\beta_2 \approx 0.6$, then the probability is 0.378, and the norm of the 3rd derivative is estimate from below by 1894.34; for $\beta_1 = 0.0$, $\beta_2 \approx 0.1$, the upper estimate is 4303.81, which is close to the norm of the 3rd derivative, but here the probability is only 0.00013856. Similar estimates can also be obtained using (3.7).

For the shifted difference, for $m = 0$ and $j = 140$, using (3.13) and employing Lemma 3.2, we have, with probability close to 0.9, the lower estimate 4206.72 for the norm of the 3rd derivative (on [1,195] with $\alpha = 0.845$ and $\beta_1 = 0.0$, $\beta_2 \approx 0.52$). If (3.14) is used, we have the lower estimate 3241.2 with probability close to 0.9 .

Remark 5.2. *Optimal formulas capable of recovering the derivative from inexactly given data frequently indicate that it is possible to consider only the recovery by values of the function on a sufficiently rare set of nodal points at which these values are given. In addition, if, for example, approximate values of the function are known also at different points, then these values will not be used. This happens, for example, in the problem related to* (1.6) *and to the recovery formulas* (1.7), (1.8). *Another situation occurs if one does not know the value of the parameter Δ that specifies the class of functions on which the recovery problem of the derivative is solved. In this case, the greater the* *number of nodes at which the approximate values of the function are known, the better, for example, the recovered characteristics of the random variables related to observation of the values of* $\delta(t_i)$ *in* (4.1) *and the values of* r_i *in* (5.1). *It seems that a finer number of nodes will also allow one to obtain, via* (3.13), (3.14), (3.15) *or* (3.21), *more precise estimates of the derivative in the uniform norm.*

6. Analysis of divided differences for intermediate derivatives for ascertaining intervals of their satisfactory recovery

Having solved the problem of approximate estimate of the leading derivative, one may consider formulas of the form (1.7), (1.8) (and similar ones) for a direct recovery of intermediate derivatives. These formulas involve the optimal step $j_{opt} \approx (24\delta/\Delta)^{1/3}$ (see (1.8)), using which one estimates the difference relation that provides an answer in our problem. At this stage it is worth formulating and analyzing the ratios a_i already obtained with the use of the successive differences for the intermediate derivative and the step j_{opt}. Let us use in (1.8), for example, the above estimate 4206.72 as Δ and take $\delta = 0.15$ (see Fig. 2). This gives $j_{\text{opt}} = 94$, whereas the step of recovery, as calculated from the exact value of the norm of the third derivative, is 91. Let us give the graph of $a_i - 1$ composed from the second central differences with the step j_{opt} thus obtained (see also Remark 3)

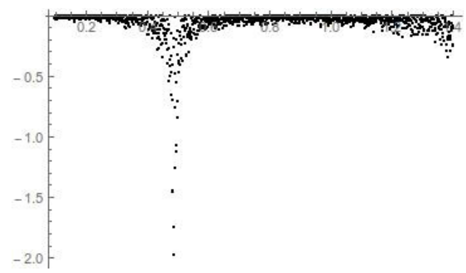

Figure 7: The sequence $(a_i - 1)$ composed from the second central differences.

Using the following two graphs, one may compare the result of approximate evaluation of the derivative (the left panel) with its exact value (the right panel).

Figure 8: The result of approximate evaluation of $p''(z)$.

In Fig. 7, the surge at $i = 485$ corresponds to the value of the argument $\frac{1.5}{1484} \cdot 485$, which is approximately $0.49 \approx 0.5$. In Fig. 8, this corresponds to the zero of the second derivative of the polynomial p. This is why $a_i - 1$ behaves chaotically for $i \approx 485$, because the random perturbation near the zero of the second derivative of the polynomial p strongly dominates. Figure 7 shows that the recovery of the second derivative is correct everywhere except the zero of the second derivative of the polynomial p. This supports correctness of our estimate of the norm of the leading derivative.

Recall that both graphs are formed on the interval $[0.1\,, 1.4] \subset [0, 1.5]$, which appears because the second derivative of the second order central difference was used for the recovery — for points lying near the endpoints of the interval $[0, 1.5]$ such a recovery is impossible — in this case, difference relations of different kind should be taken. Replacing the second order central difference by

$$(2-m)x(t) + (3m-5)x(t+j\tau) + (4-3m)x(t+2j\tau) + (m-1)x(t+3j\tau), \tag{6.1}$$

and considering the corresponding difference relation, we get an approximation of the second derivative of second order of accuracy at the point $t + mj\tau$, $m \in [0, 3]$. In addition, to secure the continuity of the passage of the successive difference in (6.1) from $m = 0$ at the point t to $m = 1$ at the point $t + j\tau$ from the system of points $t + s\tau$, $s = 0, 1, \dots, j$, one may use the successive differences with various values of $m = m_s$. For example, at the point $t + s\tau$, $s = 0, 1, \dots, j$, such m_s is equal to s/j. The sequences c_i and other sequences constructed in this way will consist of the differences with the variable parameter m. With increased step j, the number of c_i's constructed not from the central difference will increase, but the principal required property of the sequence $\{c_i\}$ (the linearity with respect to the functional parameter $x(\cdot)$) will be preserved. Note that in general the estimates from § 3 need not hold because of the m-dependent (see (6.1)) realizations of r_i linear combinations of the random variable ξ; however, this is irrelevant on this stage of smoothing of the recovery result.

Analyzing the original experimental data, we choose, with the purpose of estimating the third derivative, the sequence a_i, which is calculated for the multiplicity step $j = 40$ of the central difference. Choosing the domain $i \in [0, 130]$ with value $\alpha = 0.74$, and taking $\beta = 0.57$, we get, with probability 0.9, the estimate 54653.3 for the uniform norm of the third derivative of the function under consideration (such estimate was obtained in conditional units along the axis of the independent argument i). Using this estimate and (1.8), let us find the optimal step $j_{\mathrm{opt}} = 41$. Let us show the graph of $a_i - 1$ obtained from the second central differences with step j_{opt} for the original experimental data

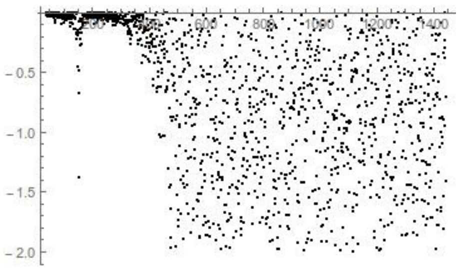

Figure 9: The graph of $a_i - 1$ (from the second central differences for the original experimental data).

Figure 9 shows that the recovered values of the second derivative for $i \geq 450$ are very noisy and their interrelations are random, while for $i < 450$, the recovered values prevail over the random component.

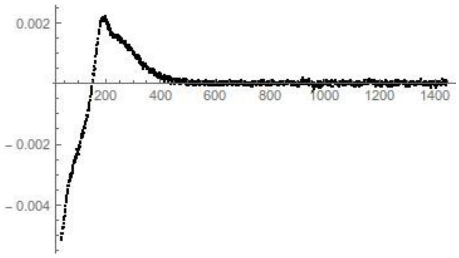

Figure 10: Recovered values of the second derivative.

7. Example of recovery of the differential operator

Analysis performed for the available set of experimental data can be carried out also for the recovery of some linear functional constructed from a two-dimensional (or, in general, of dimension > 2) set of data.

For example, to analyze in this way the possibility of recovery of the operator

$$\frac{\partial x}{\partial t} - \frac{\partial^2 x}{\partial s^2}$$

via the difference relations

$$c_{k,i} = \frac{\widetilde{x}(t_{k+j_1}, s_i) - \widetilde{x}(t_{k-j_1}, s_i)}{2j_1\tau} - \frac{\widetilde{x}(t_k, s_{i+j_2}) - 2\widetilde{x}(t_k, s_i) - \widetilde{x}(t_k, s_{i-j_2})}{(j_2 h)^2}$$

from the value of the step of the difference relations j_1 and j_2 for the variables t and s, respectively, we compose

$$a_{k,i} = \min\{a'_{k,i}, a''_{k,i}\}, \tag{7.1}$$

$$\text{where} \quad a'_{k,i} = \begin{cases} c_{k,i}/c_{k+1,i} & \text{if } \left|\frac{c_{k,i}}{c_{k+1,i}}\right| \le 1, \\ c_{k+1,i}/c_{k,i} & \text{if } \left|\frac{c_{k,i}}{c_{k+1,i}}\right| > 1, \end{cases} \quad a''_{k,i} = \begin{cases} c_{k,i}/c_{k,i+1} & \text{if } \left|\frac{c_{k,i}}{c_{k,i+1}}\right| \le 1, \\ c_{k,i+1}/c_{k,i} & \text{if } \left|\frac{c_{k,i}}{c_{k,i+1}}\right| > 1. \end{cases}$$

Let

$$b_{k,i} = \frac{x(t_{k+j_1}, s_i) - x(t_{k-j_1}, s_i)}{2j_1\tau} - \frac{x(t_k, s_{i+j_2}) - 2x(t_k, s_i) - x(t_k, s_{i-j_2})}{(j_2 h)^2}, \qquad r_{k,i} := c_{k,i} - b_{k,i}.$$

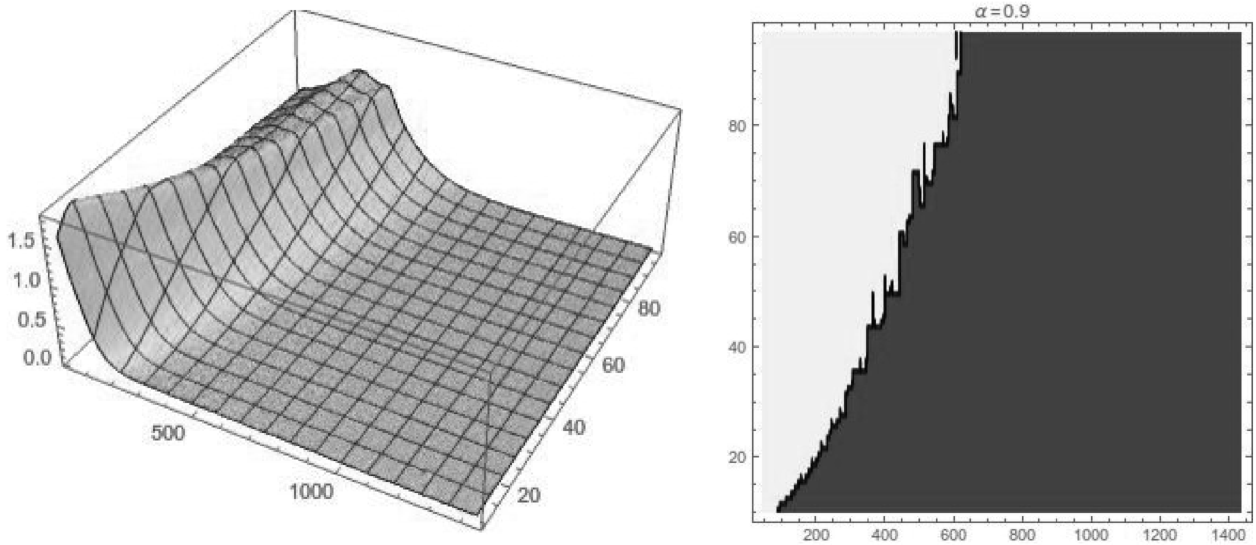

Figure 11: Left panel: The recovery of the values of the operator $\frac{\partial x}{\partial t} - \frac{\partial^2 x}{\partial s^2}$ on measurement data ($k \in [1, 111]$, $i \in [1, 1485]$). Right panel: the bright domain on the left corresponds to strong domination of the recovered values of the operator over noise, that is, $a_{k,i} - 1 \ge -0.1$ for $j_1 = 5$, $j_2 = 40$.

One may assume, for example, that for this connected component of the range of variation of $\{k, i\}$, where $a_{k,i} \ge 0.9$, the values $b_{k,i}$ are prevailing over $r_{k,i}$, and the "optimistic" inequality $|r_{k,i}| \le \rho \cdot |b_{k,i}|$ is satisfied. As a result, we can write

$$\frac{c_{k,i}}{c_{k,i+1}} = \frac{b_{k,i} + r_{k,i}}{b_{k,i+1} + r_{k,i+1}} = \frac{b_{k,i}}{b_{k,i+1}} \cdot \theta_1 = \theta_1 + O(h), \quad \frac{c_{k,i}}{c_{k+1,i}} = \theta_2 + O(\tau),$$

$\theta_1, \theta_2 \in [\kappa^{-1}, \kappa]$, $\kappa = \frac{1+\rho}{1-\rho}$. Then the above assumption about the acceptability of the "optimistic" inequality will correspond to the condition $a_{k,i} - 1 \ge -0.1$ for $\kappa^{-1} - 1 = -0.1$, i.e., $\rho = 1/19$.

Acknowledgment. The author expresses his deep gratitude to V. M. Tikhomirov, I. G. Tsar'kov, A. R. Alimov, A. S. Demidov, and L. V. Lokutsievskiy for their attention and help. The figures were drawn using Wolfram Mathematica. The paper was published with the financial support of the Ministry of Education and Science of the Russian Federation as part of the program of the Moscow Center for Fundamental and Applied Mathematics under the agreement no. 075-15-2019-1621.

References

[1] T. Wei, M. Li, "High order numerical derivatives for one-dimensional scattered noisy data", Applied Mathematics and Computation. 175, Issue 2, P. 1744–1759, (2006)

[2] A.G. Ramm, A. Smirnova "Stable numerical differentiation: When is it possible?", J. KSIAM. 7, No 1, P. 47–61 (2003)

[3] A.G. Ramm, A.B. Smirnova, "On stable numerical differentiation", Mathematics of Computation. 70, No 235, P. 1131–1153 (2001)

[4] S. B. Stechkin, "Inequalities between norms of derivatives of an arbitrary function", Acta Sci. Math. 26 (No. 3–4), 225–230 (1965).

[5] S. B. Stechkin, "Best approximation of linear operators", Mat. Zametki 1 (2), 137–148 (1967) [Math. Notes 1 (1–2), 91–99 (1967) (1968)].

[6] V. N. Gabushin, "Exact constants in inequalities between norms of derivatives of functions", Mat. Zametki, 4:2 (1968), 221–232; Math. Notes, 4:2 (1968), 624–630

[7] V. N. Gabushin, "Exact constants in inequalities between norms of derivatives of functions", Mat. Zametki, 4:2 (1968), 221–232; Math. Notes, 4:2 (1968), 624–630

[8] Yu. N. Subbotin, L. V. Taikov, "Best approximation of a differentiation operator in L^2-space", Mat. Zametki, 3:2 (1968), 157–164; Math. Notes, 3:2 (1968), 106–109

[9] L. V. Taikov, "Kolmogorov-type inequalities and the best formulas for numerical differentiation", Mat. Zametki, 4:2 (1968), 233–238; Math. Notes, 4:2 (1968), 631–634.

[10] J. Hadamard Sur le module maximum d'une fonction et de ses dérivées // Bull. Soc. math. France. 1914. V. 41. P. 68–72.

[11] A. N. Kolmogorov, "On inequalities between upper bounds of the successive derivatives of an arbitrary function on an infinite interval", Uchen. Zap. Moskov. Gos. Univ. 30 (1939), 3–13; English transl. in Amer. Math. Soc. Transl. (1) 2 (1962).

[12] Y. Domar, An extremal problem related to Kolmogoroff's inequality for bounded functions, // Arkiv for Mat., 7 (1968), p. 433–441.

[13] A. P. Buslaev, "Approximation of a differentiation operator", Mat. Zametki, 29:5 (1981), 731–742; Math. Notes, 29:5 (1981), 372–378.

[14] A. P. Buslaev, G. G. Magaril-Il'yaev, V. M. Tikhomirov, "Existence of extremal functions in inequalities for derivatives", Mat. Zametki, 32:6 (1982), 823–834; Math. Notes, 32:6 (1982), 898–904.

[15] G. G. Magaril-Il'yaev, K. Yu. Osipenko, and V. M. Tikhomirov, "Optimal recovery and the theory of extremum," Dokl. Ross. Akad. Nauk, 379, No. 2, 161–164 (2001); English transl. in Russian Acad. Sci. Dokl. Math., 64 (2001).

[16] G. G. Magaril-Il'yaev and K. Yu. Osipenko, "Optimal reconstruction of functionals from inaccurate data," Mat. Zametki 50 (6), 85–93 (1991) [Math. Notes 50 (5–6), 1274–1279 (1991)].

[17] V. V. Arestov, "Approximation of unbounded operators by bounded operators and related extremal problems", Uspekhi Mat. Nauk, 51:6(312) (1996), 89–124; Russian Math. Surveys, 51:6 (1996), 1093–1126

[18] V. V. Arestov, R. R. Akopyan, "Stechkin's problem on the best approximation of an unbounded operator by bounded ones and related problems", Trudy Inst. Mat. i Mekh. UrO RAN, 26, no. 4, 2020, 7–31

[19] K. I. Babenko, Principles of Numerical Analysis (Regular and Chaotic Dynamics, Moscow, 2002) [in Russian].

[20] C. A. Micchelli, T. J. Rivlin, A survey of optimal recovery // Optimal estimation in approximation theory, ed. C. A. Micchelli, T. J. Rivlin. New York: Plenum Press, 1977, p. 1–54.

[21] C. A. Micchelli, T. J. Rivlin, Lectures on optimal recovery // Numerical analysis, Proc. SERC Summer Sch., Lancaster/Engl., 1984. Berlin: Springer–Verlag, 1984. P. 21–93. (Lecture Notes in Math. V. 1129.)