MATHEMATICS OF COMPUTATION
Volume 70, Number 235, Pages 1131–1153
S 0025-5718(01)01307-2
Article electronically published on March 9, 2001

ON STABLE NUMERICAL DIFFERENTIATION

ALEXANDER G. RAMM AND ALEXANDRA B. SMIRNOVA

ABSTRACT. A new approach to the construction of finite-difference methods is presented. It is shown how the multi-point differentiators can generate regularizing algorithms with a stepsize h being a regularization parameter. The explicitly computable estimation constants are given. Also an iteratively regularized scheme for solving the numerical differentiation problem in the form of Volterra integral equation is developed.

1. INTRODUCTION

The problem of numerical differentiation is known to be ill posed in the sense that small perturbations of the function to be differentiated may lead to large errors in the computed derivative. However, in many applications it is necessary to estimate the derivative of a function given the noisy values of this function. As an example we refer to the analysis of photoelectric response data (see [13, 1970]). The goal of this experiment is to determine the relationship between the intensity of light falling on certain plant cells and their rate of uptake of various substances in order to gain further information about photosynthesis. Rather than measuring the uptake rate directly, the experimenters measure the amount of each substance not absorbed as a function of time, the uptake rate being defined as minus the derivative of this function. As for the other example, one can mention the problem of finding the heat capacity of a gas c_p as a function of temperature T. Experimentally, one measures the heat content

$$q(T) = \int_{T_0}^{T} c_p(\tau)d\tau,$$

and the heat capacity is determined by numerical differentiation.

A number of techniques have been developed for numerical differentiation. They fall into three categories: difference methods, interpolation methods and regularization methods. The first two categories (see, for instance, [3], [15], [24], [25] and others), especially the central difference formula that can be related to both of them, are well known. They have the advantage of simplicity and are considered by many authors to be the ones which yield satisfactory results when the function to be differentiated is given very precisely ([9], [2], [4]). In Sections 1 and 2 of our paper the other view on these methods, based on the works [16]-[23] and [12], is presented (see also [10] where the results and ideas of [16] are used). Namely, it

Received by the editor August 5, 1999.
2000 *Mathematics Subject Classification.* Primary 65D25; Secondary 65D05.
Key words and phrases. Numerical differentiation, noisy data, ill-posed problems, multi-point methods, regularization.

is shown how the multipoint difference schemes may construct stable regularizing algorithms for the process of numerical differentiation with a stepsize h being a regularization parameter. The main points are:

a) h must depend on δ, the level of noise in the initial data, and

b) one has to take into account a priori information about the specific class to which the function to be differentiated belongs.

Most of the regularization procedures [4]-[7], [27], [28] that belong to the third category make use of the variational ([14], [26]) approach for solving ill-posed problems. These methods typically involve writing the derivative as the solution to an integral Volterra equation and then reducing the integral equation to a family of well-posed problems that depend on a regularization parameter. Once an optimal value of this parameter is found, the corresponding well-posed problem is solved to obtain an estimate for the derivative. Unfortunately, the determination of the optimal parameter value is generally a nontrivial task.

In [6] the authors propose the quasi-solution method (see [8]) for regularization, which can be described as follows: find the coefficients of the expansion of f_δ in the Legendre polynomials $P_k(x)$

$$b_k = \frac{2k+1}{2} \int_{-1}^{1} f_\delta(x) P_k(x) dx, \quad k = 1, 2, ..., n+1, \tag{1.1}$$

"choosing n so that $||q_{n+1} - f_\delta|| \le \delta$ and $\sum_{k=0}^{n+1} \frac{2}{2k+1} b_k^2 > 4\delta^2$," where the quotation is from [6]. The function q_{n+1} here is the approximating polynomial and $||f - f_\delta|| \le \delta$. Then the estimate of the derivative of f is given by

$$p_\delta = \sum_{k=1}^{n+1} \frac{b_k}{1 + \lambda k^2 (1+k)^2} P_k(x). \tag{1.2}$$

The existence and the uniform convergence of approximation (1.2) is proved in [6]. Apparently, it is assumed in [6] that f and f_δ are such that the above choice of n is possible. This is not always so: if $||f_\delta|| < 2\delta$, then

$$4\delta^2 > ||f_\delta||^2 \ge \sum_{k=0}^{n+1} \frac{2}{2k+1} b_k^2$$

by Parseval's identity. In this case it is impossible to choose n in such a way that $\sum_{k=0}^{n+1} \frac{2}{2k+1} b_k^2 > 4\delta^2$. If n can be chosen, as suggested in [6], then to determine λ one has to solve the nonlinear equation

$$\sum_{k=1}^{n+1} \frac{2}{2k+1} b_k^2 \frac{\lambda^2 k^4 (k+1)^4}{[1 + \lambda k^2 (k+1)^2]^2} = 4\delta^2. \tag{1.3}$$

In [7] the idea proposed by the authors of [6] is generalized to the case of the kth order derivative, $k > 1$. No results of a numerical experiment based on procedure (1.2)-(1.3) are given in [6] as well as in [7].

One of the first regularization methods, which performed well in practice (according to the experiments illustrated in [2] and [4]) when used with noisy data, can be found in [4]. There a variational approach to a regularization is used to obtain the following family of well-posed optimization problems:

$$(1.4)\qquad ||Af-g||^2+\left(\int_0^1 f\,dx\right)^2+\alpha(||f||^2+||f'||^2)=\min_{f\in W_1^2}.$$

The norm is understood in L^2-sense, and $Af=g$ is the initial Volterra equation. Then (1.4) is reduced to the second kind Fredholm integral equation

$$(1.5)\qquad \alpha f(x)+\int_0^1 K_0(x,u)f(u)du=m(x),\quad 0\le x\le 1.$$

The error contributed by variational regularization (1.4) and trapezoidal discretization of (1.5) is estimated in [4]. From this estimate it is clear that for a given δ the regularization parameter α has a nonzero optimal value. Although a way of choosing the regularization parameter optimally is not offered there, the later paper [2] describes a spectral interpretation of method [4] that allows one to obtain an optimal value of this parameter. However, the applicability of that interpretation is restricted to the cases when the spectrum of the data shows a clear distinction between the signal and the noise.

In the literature there were no results which give error estimates for a stable numerical differentiation algorithm such that the estimate would be suitable for a fixed $\delta>0$ and the estimation constants would be given explicitly, so that the error estimate could be used in practical computations. An exception is the result in [16] generalized and applied in [17]-[23].

In Section 4 we suggest an iteratively regularized scheme for solving a Volterra equation based on the idea of continuous regularization (see [1]), which is an alternative to the variational one. This procedure avoids some of the limitations in a choice of the regularization parameter mentioned above.

In Section 5 the results of the numerical experiment are discussed. The derivative of the test function $f(x)=\sin(\pi x)$, $x\in[0,1]$, was computed in the presence of noise, whose maximum value was 10% of the maximum value of $f(x)$. The dependence of the actual and estimated errors on the regularization parameters is considered. The important practical recommendations on stable numerical differentiation with various a priori information are given.

2. Inequalities for the derivatives

2.1. **The main result.** In this section we investigate and answer Questions 2.1 and 2.2 (see [22]).

Question 2.1. Given $f_\delta\in L^\infty(\mathbb{R})$ and $f\in C^1(\mathbb{R})$ such that inequalities

$$||f_\delta-f||\le\delta,\quad ||f^{(m)}||\le M_m<\infty,\quad m=0,1,$$

hold with some known δ and unknown (or roughly estimated) M_m, can one compute f' stably?

Here

$$(2.1)\qquad ||f||:=\operatorname*{ess\,sup}_{x\in\mathbb{R}}|f(x)|.$$

In other words, does there exist an operator T such that

(2.2) $$\sup_{f\in\mathcal{K}(\delta,M_m)} \|Tf_\delta - f'\| \le \eta(\delta) \longrightarrow 0 \quad \text{as} \quad \delta \to 0,$$

where

(2.3) $$\mathcal{K}(\delta, M_m) := \left\{ f: \ f \in C^1(\mathbb{R}), \ ||f^{(m)}|| \le M_m < \infty, \ ||f_\delta - f|| \le \delta \right\},$$

$m = 0$ or $m = 1$?

Question 2.2. This is similar to Question 1 but now it is assumed that $m = 1 + a$, $0 < a \le 1$.

(2.4) $$||f^{(1+a)}|| := M_{1+a} < \infty,$$

where

$$||f^{(1+a)}|| := \operatorname*{ess\,sup}_{x,y\in\mathbb{R}} \frac{|f'(x) - f'(y)|}{|x-y|^a}.$$

The basic results of this section are summarized in Theorem 2.1.

Theorem 2.1. *There does not exist an operator* $T : L^\infty(\mathbb{R}) \to L^\infty(\mathbb{R})$ *such that inequality* (2.2) *holds if* $m = 0$ *or* $m = 1$. *There exists such an operator if* $m > 1$. *This operator is given by* (2.11) *with* $h = h_{1+a}(\delta)$ *defined in* (2.13). *The error of the corresponding differentiation formula* (2.11) *is presented in* (2.13).

Proof. Consider

(2.5) $$f_1(x) := -\frac{M}{2} x(x - 2h), \quad 0 \le x \le 2h.$$

Extend $f_1(x)$ to the whole real axis in such a way that the norms $||f_1^{(m)}||$, $m = 0, 1, 2,$ are preserved. Define $f_2(x) := -f_1(x)$. Note that

(2.6) $$\sup_{x\in R} |f_k(x)| = \frac{Mh^2}{2}, \quad k = 1, 2.$$

Take

(2.7) $$\frac{Mh^2}{2} = \delta,$$

then for $f_\delta(x) \equiv 0$ it follows the estimate $||f_k - f_\delta|| \le \delta$, $k = 1, 2$. Let $(Tf_\delta)(0) = (T0)(0) := b$. If $||f_k^{(m)}|| \le M_m$, $k = 1, 2$, one has

(2.8)
$$\begin{aligned}\gamma_m(\delta) &:= \inf_T \sup_{f\in\mathcal{K}(\delta,M_m)} ||Tf_\delta - f'|| \ge \inf_T \max_{k=1,2} ||Tf_\delta - f_k'|| \\ &\ge \inf_T \max_{k=1,2} ||Tf_\delta(0) - f_k'(0)|| = \inf_{b\in R} \max\{|b - Mh|, \ |b + Mh|\} = Mh.\end{aligned}$$

By (2.7) $h = \sqrt{\frac{2\delta}{M}}$, $Mh = \sqrt{2\delta M}$.

If $m = 0$, then (2.6) implies that $f_k \in \mathcal{K}(\delta, M_0)$, $k = 1, 2$, with $M_0 := \frac{Mh^2}{2} = \delta$. Since for any fixed $\delta > 0$ and $M_0 = \delta$, the constant M in (2.5) can be chosen arbitrary, inequality (2.8) proves that (2.2) is false in the class $\mathcal{K}(\delta, M_0)$ and in fact $\gamma_0(\delta) \to \infty$ as $M \to \infty$.

Estimate (2.2) is also false in the class $\mathcal{K}(\delta, M_1)$. Indeed,

(2.9) $$||f_1'|| = ||f_2'|| = \sup_{0\le x\le 2h} |M(x-h)| = Mh = \sqrt{2\delta M}.$$

Thus, for given δ, $M_1 > 0$ one can find M such that (2.7) holds and $M_1 = \sqrt{2\delta M}$. For this M the functions $f_k \in \mathcal{K}(\delta, M_1)$, $k = 1, 2$. By (2.8) one obtains

$$\gamma_1(\delta) \geq M_1 > 0 \quad \text{as} \quad \delta \to 0 \tag{2.10}$$

so that (2.2) is false.

Let us assume now that (2.4) holds. Take

$$T_h f_\delta := \frac{f_\delta(x+h) - f_\delta(x-h)}{2h}, \quad h > 0. \tag{2.11}$$

One gets, using the Lagrange formula,

$$\begin{aligned} \|T_h f_\delta - f'\| &= \|T_h(f_\delta - f)\| + \|T_h f - f'\| \\ &\leq \frac{\delta}{h} + \left\| \frac{f(x+h) - f(x-h) - 2hf'(x)}{2h} \right\| \\ &\leq \frac{\delta}{h} + \left\| \frac{[f'(y) - f'(x)]h + [f'(z) - f'(x)]h}{2h} \right\| \\ &\leq \frac{\delta}{h} + M_{1+a} h^a := \varepsilon_{1+a}(\delta, h), \end{aligned} \tag{2.12}$$

where y and z are the intermediate points in the Lagrange formula.

Minimizing the right-hand side of (2.12) with respect to $h \in (0, \infty)$ yields

$$h_{1+a}(\delta) = \left(\frac{\delta}{a M_{1+a}} \right)^{\frac{1}{1+a}}, \quad \varepsilon_{1+a}(\delta) = \frac{a+1}{a^{\frac{a}{1+a}}} \, M_{1+a}^{\frac{1}{1+a}} \, \delta^{\frac{a}{1+a}}, \quad 0 < a \leq 1. \tag{2.13}$$

Theorem 2.1 is proved. □

Remark 2.2. By formula (2.12) one can see that if M_{1+a} is not known a priori, then one chooses $h(\delta) = O(\delta^\gamma)$, $0 < \gamma < 1$, where γ is a constant and $\epsilon_{1+a}(\delta, \delta^\gamma) \to 0$ as $\delta \to 0$. However, since in most cases in practice $\delta > 0$ is fixed, and the choice of $h = O(\delta^\gamma)$ is not quite determined (it gives only asymptotics of h as $\delta \to 0$ and does not determine the constant factor), this choice does not yield a practically computable error estimate for the formula of stable numerical differentiation. This is in sharp contrast with the practically computable error estimate given in (2.13).

2.2. **The best estimate for** $m = 2$. It is significant to mention that $\varepsilon_{1+a}(\delta)$ obtained in (2.13) can be improved if $a = 1$ ($m = 2$). In this subsection we will give the optimal error estimate for functions of one or several variables in the class $\mathcal{K}_n(\delta, M_2)$, $n \geq 1$, defined as follows:

$$\begin{aligned} \mathcal{K}_n(\delta, M_2) := \{ f : \ & f \in C^2(\mathbb{R}^n), \quad |(d^2 f(x)k, k)| \leq M_2 |k|^2, \\ & \forall x, k \in \mathbb{R}^n, \quad ||f_\delta - f|| \leq \delta \}, \end{aligned} \tag{2.14}$$

where

$$(d^2 f(x)k, k) = \sum_{i,j=1}^{n} f_{x_i x_j}(x) k_i k_j, \tag{2.15}$$

$$|k|^2 = \sum_{i=1}^{n} |k_i|^2, \quad ||f|| := \operatorname*{ess\,sup}_{x \in \mathbb{R}^n} |f(x)|.$$

We want to approximate the directional derivative $D_v f(x) = \nabla f(x) \cdot v$, where v is a unit vector. Define

$$T_h f_\delta := \frac{f_\delta(x+hv) - f_\delta(x-hv)}{2h}, \quad h > 0. \tag{2.16}$$

Theorem 2.3 ([17]). *If* $h_2(\delta) = \left(\frac{2\delta}{M_2}\right)^{\frac{1}{2}}$ *and* $\varepsilon_2(\delta) = (2\delta M_2)^{\frac{1}{2}}$, *then*

$$||T_{h_2(\delta)} f_\delta - D_v f|| \leq \varepsilon_2(\delta) \tag{2.17}$$

and

$$\gamma_2(\delta) := \inf_{T \in \mathcal{A}} \sup_{f \in \mathcal{K}_n(\delta, M_2)} ||T f_\delta - D_v f|| = \varepsilon_2(\delta). \tag{2.18}$$

Here $\mathcal{A}$ *is the set of all operators* $T : L^\infty(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n)$.

Proof. For any $x \in \mathbb{R}^n$ and $h > 0$

$$\begin{aligned} ||T_h f_\delta - D_v f|| &\leq ||T_h f_\delta - T_h f|| + ||T_h f - D_v f|| \\ &\leq ||T_h(f_\delta - f)|| + \frac{M_2 h}{2} \\ &\leq \frac{\delta}{h} + \frac{M_2 h}{2}. \end{aligned} \tag{2.19}$$

Consider the right-hand side of (2.19) as a function of h. As before, this function has an absolute minimum $\varepsilon_2(\delta)$ at $h = h_2(\delta)$. Thus $||T_{h_2(\delta)} f_\delta - D_v f|| \leq \varepsilon_2(\delta)$. Let

$$f_1(x) = \frac{1}{2} M_2 \sum_{i=1}^n x_i^2 - M_2 h_2(\delta) \sum_{i=1}^n v_i x_i$$

for x inside the ball centered at $(h_2(\delta)v_1, ..., h_2(\delta)v_n)$ with radius $h_2(\delta)$. Note that $f_1(x) = 0$ on the boundary of this ball, f_1 has an absolute minimum value $-\delta$ at $x = (h_2(\delta)v_1, ..., h_2(\delta)v_n)$, and $|(d^2 f_1(x)k, k)| = M_2|k|^2$. Also one gets

$$D_v f_1(0) = -M_2 h_2(\delta) \sum_{i=1}^n v_i^2 = -M_2 h_2(\delta) = -(2\delta M_2)^{\frac{1}{2}} = -\varepsilon_2(\delta).$$

Continue f_1 on $\mathbb{R}^n$ in such a way that

$$|(d^2 f_1(x)k, k)| = M_2|k|^2, \quad ||f_1|| \leq \delta \quad \forall x \in \mathbb{R}^n. \tag{2.20}$$

The continuation is possible since in the above ball conditions (2.20) are satisfied. Set $f_2 = -f_1$, $f_\delta = 0$. Following the proof of Theorem 2.1 (formula (2.8)) one obtains

$$\varepsilon_2(\delta) \leq \gamma_2(\delta). \tag{2.21}$$

On the other hand,

$$\gamma_2(\delta) \leq \sup_{f \in \mathcal{K}_n(\delta, M_2)} ||T_h f_\delta - D_v f|| \leq \varepsilon_2(\delta). \tag{2.22}$$

From (2.21) and (2.22) one gets (2.18). □

3. Multi-point finite difference formulas

3.1. **Finite-difference methods in the class $\mathcal{K}(\delta, M_m)$, $m > 2$.** Assume that $f(x)$ has derivatives up to order m and $||f^{(m)}|| \le M_m$, $m > 2$. Let $f_\delta \in L^\infty(\mathbb{R})$ be given such that $||f - f_\delta|| \le \delta$, where the norm is calculated by (2.1). Suppose that $m > 2$ is odd and define (see [17], [19])

$$T_h^Q f_\delta := h^{-1} \sum_{j=-Q}^{Q} A_j^Q f_\delta \left(x + \frac{jh}{Q} \right), \tag{3.1}$$

where the numbers A_j^Q $(j = -Q, ..., Q)$ are to be determined. One has

(3.2)

$$\begin{aligned}\left|T_h^Q f_\delta - f'\right| &\le \left|T_h^Q (f_\delta - f)\right| + \left|T_h^Q f - f'\right| \\ &\le \frac{\delta}{h} \sum_{j=-Q}^{Q} \left|A_j^Q\right| + \left| \frac{f(x)}{h} \sum_{j=-Q}^{Q} A_j^Q + f'(x) \left[\sum_{j=-Q}^{Q} \left(\frac{j}{Q}\right) A_j^Q - 1 \right] \right. \\ &\quad \left. + \sum_{p=2}^{m-1} \frac{f^{(p)}(x) h^{p-1}}{p!} \sum_{j=-Q}^{Q} \left(\frac{j}{Q}\right)^p A_j^Q + \frac{h^{m-1}}{m!} \sum_{j=-Q}^{Q} f^{(m)}(\xi) \left(\frac{j}{Q}\right)^m A_j^Q \right|,\end{aligned}$$

$m > 2.$

If one requires the order of smallness (as $h \to \infty$) for $\left\|T_h^Q f_\delta - f'\right\|$ to be minimal, one gets the following system of equations for the coefficients A_j^Q:

$$\sum_{j=-Q}^{Q} \left(\frac{j}{Q}\right)^l A_j^Q = \delta_{1l}, \quad 0 \le l \le 2Q, \quad \delta_{1l} := \begin{cases} 0, & \text{if } \ l \ne 1, \\ 1, & \text{if } \ l = 1. \end{cases} \tag{3.3}$$

This is a linear system of m equations with $2Q + 1$ unknowns. When $m = 2Q + 1$, system (3.3) has a nonsingular (Vandermonde) matrix. So it is uniquely solvable for A_j^Q. For the first few values of Q one obtains (see [19] and [12]):

$$\begin{gathered} A_0^1 = 0, \quad A_{\pm 1}^1 = \pm \frac{1}{2}, \\ A_0^2 = 0, \quad A_{\pm 1}^2 = \pm \frac{4}{3}, \quad A_{\pm 2}^2 = \mp \frac{1}{6}, \\ A_0^3 = 0, \quad A_{\pm 1}^3 = \pm \frac{9}{4}, \quad A_{\pm 2}^3 = \mp \frac{9}{20}, \quad A_{\pm 3}^3 = \pm \frac{1}{20}, \\ A_0^4 = 0, \quad A_{\pm 1}^4 = \pm \frac{16}{5}, \quad A_{\pm 2}^4 = \mp \frac{4}{5}, \end{gathered} \tag{3.4}$$

With this choice of coefficients A_j^Q, from (3.2) one gets

$$\begin{aligned}\left|T_h^Q f_\delta - f'\right| &\le \frac{\delta}{h} \sum_{j=-Q}^{Q} \left|A_j^Q\right| + h^{m-1} \frac{M_m}{m!} \sum_{j=-Q}^{Q} \left| \left(\frac{j}{Q}\right)^m A_j^Q \right| \\ &\equiv \frac{\delta}{h} \alpha_m + h^{m-1} M_m \beta_m.\end{aligned} \tag{3.5}$$

The right-hand side of (3.5) as a function of h has an absolute minimum $\varepsilon_m(\delta)$ when

$$h = h_m(\delta) = \left[\frac{\alpha_m \delta}{(m-1)M_m\beta_m}\right]^{\frac{1}{m}}, \tag{3.6}$$

the coefficients α_m, β_m are defined in (3.5), and

$$\varepsilon = \varepsilon_m(\delta) = m\left[\frac{\alpha_m^{m-1}M_m\beta_m\delta^{m-1}}{(m-1)^{m-1}}\right]^{\frac{1}{m}}. \tag{3.7}$$

Remark 3.1. In [19] and [12] multi-point formulas for the two-dimensional case are also investigated.

3.1.1. *Best possible estimates of f' and f'' in the class $\mathcal{K}(\delta, M_3)$.* It is shown above that the minimal order of smallness for the approximation error $\left\|T_h^Q f_\delta - f'\right\|$ is attained when $m = 2Q+1$ and A_j^Q are the solutions to (3.3).

According to (3.1), (3.4) and (3.6), for $m = 3$ one gets the central difference derivative.

$$T^1_{h_3(\delta)} f_\delta = \frac{f_\delta(x + h_3(\delta)) - f_\delta(x - h_3(\delta))}{2h} \tag{3.8}$$

with

$$h_3(\delta) = \left(\frac{3\delta}{M_3}\right)^{\frac{1}{3}}. \tag{3.9}$$

As a simple consequence of this observation we state the following proposition.

Proposition 3.2. *Among all linear and nonlinear operators $T : L^\infty(\mathbb{R}) \to L^\infty(\mathbb{R})$ the operator $T^1_{h_3(\delta)}$ defined by (3.8), (3.9) gives the best possible estimate of f' in the class of all functions $f \in \mathcal{K}(\delta, M_3)$ (see (2.3)) and*

$$\gamma_3(\delta) := \inf_T \sup_{f\in\mathcal{K}(\delta,M_3)} \|Tf_\delta - f'\| = \frac{3^{\frac{2}{3}}}{2} M_3^{\frac{1}{3}}\delta^{\frac{2}{3}} = \varepsilon_3(\delta). \tag{3.10}$$

Proof. By (3.5)-(3.7) one obtains $\gamma_3(\delta) \le \varepsilon_3(\delta)$. To prove that

$$\gamma_3(\delta) \ge \varepsilon_3(\delta) \tag{3.11}$$

take

$$f_1(t) := -\frac{M_3}{6}t^3 + \varepsilon_3(\delta)t, \quad t \in \left[0, \left(\frac{6\varepsilon_3(\delta)}{M_3}\right)^{\frac{1}{2}}\right]. \tag{3.12}$$

It is clear that $f_1(0) = f_1\left(\left(\frac{6\varepsilon_3(\delta)}{M_3}\right)^{\frac{1}{2}}\right) = 0$ and

$$|f_1^{(3)}(t)| = M_3, \quad f_1'(0) = \varepsilon_3(\delta), \quad |f_1(t)| \le \delta. \tag{3.13}$$

Continue $f_1(t)$ on $\mathbb{R}$ in such a way that (3.13) are satisfied for any $t \in \mathbb{R}$. Let $f_2(t) = f_1(t)$, $f_\delta = 0$. Thus, following the proof of Theorem 2.1 once again (formula (2.8)) one derives (3.11) and completes the proof. □

A similar result can be proved for the second derivative operator in the class $\mathcal{K}(\delta, M_3)$. Introduce

$$R_\delta f_\delta := \frac{f_\delta(x - g_3(\delta)) - 2f_\delta(x) + f_\delta(x + g_3(\delta))}{g_3^2(\delta)}, \tag{3.14}$$

$$g_3(\delta) := 2\left(\frac{3\delta}{M_3}\right)^{\frac{1}{3}}. \tag{3.15}$$

Proposition 3.3. *Among all linear and nonlinear operators $T : L^\infty(\mathbb{R}) \to L^\infty(\mathbb{R})$, the operator R_δ, defined by* (3.14)-(3.15), *gives the best possible estimate of f'' in the class $\mathcal{K}(\delta, M_3)$ and*

$$\beta_3(\delta) := \inf_T \sup_{f \in \mathcal{K}(\delta, M_3)} \|Tf_\delta - f''\| = 3^{\frac{1}{3}} M_3^{\frac{2}{3}} \delta^{\frac{1}{3}} := \alpha_3(\delta). \tag{3.16}$$

Proof. Using the Taylor formula one gets

$$\|R_\delta f_\delta - f''\| \le \frac{4\delta}{g_3^2(\delta)} + \frac{M_3 g_3(\delta)}{3}. \tag{3.17}$$

The function $G(g) := \frac{4\delta}{g^2} + \frac{M_3 g}{3}$ has an absolute minimum $\alpha_3(\delta)$ if $g = g_3(\delta)$ (see formulas (3.16) and (3.15), respectively). Therefore $\beta_3(\delta) \le \alpha_3(\delta)$. To obtain the inequality

$$\beta_3(\delta) \ge \alpha_3(\delta), \tag{3.18}$$

we consider the function

$$f(t) = -\frac{M_3}{6}t^3 + \frac{\alpha_3(\delta)}{2}t^2 - \delta.$$

This function has three real roots: $t_1 < 0 < t_2 < t_3$. One can also check that on the interval $[t_1, t_3]$

$$|f^{(3)}(t)| = M_3, \quad f''(0) = \alpha_3(\delta), \quad |f(t)| \le \delta. \tag{3.19}$$

The last inequality is true because $f(t)$ attains its global maximum on the interval $[t_1, t_3]$ at $\hat{t} = \frac{2\alpha_3(\delta)}{M_3}$, its global minimum on the above interval at $\tilde{t} = 0$, and $|f(\hat{t})| = |f(\tilde{t})| = \delta$.

Suppose that $f_1(t) = f(t)$ for $t \in [t_1, t_3]$ and continue $f_1(t)$ on $\mathbb{R}$ so that (3.19) holds for any $t \in \mathbb{R}$. Let $f_2(t) = -f_1(t)$, $f_\delta(t) \equiv 0$. If one estimates $\beta_3(\delta)$ from below following the idea used in (2.8) then one arrives at (3.18) and the proof is completed. □

3.2. **Finite difference methods in the class $\mathcal{K}(\sigma, M_2)$.** Here we consider the family of the operators (see [17], [19] and [23])

$$T_h^Q f_\sigma := h^{-1} \sum_{j=-Q}^{Q} A_j^Q f_\sigma\left(x + \frac{jh}{Q}\right) \tag{3.20}$$

under the assumption that f is twice differentiable and $||f''|| \le M_2$. We assume also that the approximation $f_\sigma(x) = f(x) + e_\sigma(x)$, where the error $e_\sigma(x)$ is independent and identically distributed with zero mean value and variance σ^2. We will call the class of functions $f \in C^2(\mathbb{R})$ with the above properties $\mathcal{K}(\sigma, M_2)$.

We wish to determine the coefficients A_j^Q which minimize the mean square error

$$E\left[\left(T_h^Q f_\sigma - f'\right)^2\right] = \min \tag{3.21}$$

and

$$E\left[T_h^Q f_\sigma\right] = f', \tag{3.22}$$

where E denotes the mean value. By Taylor's formula one gets

$$\begin{aligned} T_h^Q f_\sigma - f' = h^{-1} \sum_{j=-Q}^{Q} A_j^Q \left[f_\sigma\left(x + \frac{jh}{Q}\right) - f\left(x + \frac{jh}{Q}\right)\right] \\ + \frac{f(x)}{h} \sum_{j=-Q}^{Q} A_j^Q + f'(x)\left[\sum_{j=-Q}^{Q}\left(\frac{j}{Q}\right) A_j^Q - 1\right] \\ + h \sum_{j=-Q}^{Q} \frac{f''(\xi_j)}{2}\left(\frac{j}{Q}\right)^2 A_j^Q. \end{aligned} \tag{3.23}$$

It follows from (3.21)-(3.23) (see [19] and [23]) that

$$\sum_{j=-Q}^{Q} A_j^Q = 0, \quad \sum_{j=-Q}^{Q}\left(\frac{j}{Q}\right) A_j^Q = 1, \tag{3.24}$$

$$\frac{\sigma^2}{h^2} \sum_{j=-Q}^{Q}\left(A_j^Q\right)^2 + h^2\left(\sum_{j=-Q}^{Q}\left|A_j^Q\right|\left(\frac{j}{Q}\right)^2 \frac{M_2}{2}\right)^2 = \min. \tag{3.25}$$

Let us find A_j^Q satisfying conditions (3.24) and the condition

$$\sum_{j=-Q}^{Q}\left(A_j^Q\right)^2 = \min.$$

Then we will minimize (3.25) with respect to h. Using Lagrange multipliers λ and ν, one has

$$2A_j^Q - \lambda - \nu\left(\frac{j}{Q}\right) = 0, \quad -Q \le j \le Q. \tag{3.26}$$

System (3.26) together with (3.24) imply

$$\lambda = 0, \quad \nu = \frac{6Q}{(Q+1)(2Q+1)}.$$

So

$$A_j^Q = \frac{\nu j}{2Q} = \frac{3j}{(Q+1)(2Q+1)} \tag{3.27}$$

and

$$\sum_{j=-Q}^{Q}\left(A_j^Q\right)^2 = \frac{18}{[(Q+1)(2Q+1)]^2} \sum_{j=1}^{Q} j^2 = \frac{3Q}{(Q+1)(2Q+1)}. \tag{3.28}$$

Thus (3.25) becomes

$$\varphi(h) := \frac{\sigma^2 a_Q}{h^2} + h^2 M_2^2 b_Q = \min, \tag{3.29}$$

$$a_Q := \frac{3Q}{(Q+1)(2Q+1)}, \tag{3.30}$$

$$b_Q := \frac{9}{Q^4(Q+1)^2(2Q+1)^2}\left(\sum_{j=1}^{Q} j^3\right)^2 = \frac{9(Q+1)^2}{16(2Q+1)^2}. \tag{3.31}$$

Minimizing with respect to h, we find

$$-2\sigma^2 a_Q h^{-3} + 2h^2 M_2^2 b_Q = 0, \quad h(\sigma) = \left(\frac{\sigma}{M_2}\right)^{\frac{1}{2}} \left(\frac{a_Q}{b_Q}\right)^{\frac{1}{4}}. \tag{3.32}$$

Let $\varphi(h(\sigma)) = \varepsilon_2^2(\sigma)$. Then

$$\varepsilon_2^2(\sigma) = \sigma^2 a_Q \frac{M_2}{\sigma}\left(\frac{b_Q}{a_Q}\right)^{\frac{1}{2}} + M_2^2 b_Q \frac{\sigma}{M_2}\left(\frac{a_Q}{b_Q}\right)^{\frac{1}{2}} = 2\sigma M_2 \sqrt{a_Q b_Q}. \tag{3.33}$$

Thus the standard deviation of the optimal estimate is

$$\sqrt{E\left[\left(T_h^Q f_\sigma - f'\right)^2\right]} \le \sqrt{2\sigma M_2}(a_Q b_Q)^{\frac{1}{4}}. \tag{3.34}$$

Note that

$$(a_Q b_Q)^{\frac{1}{4}} = \left(\frac{27Q(Q+1)}{16(2Q+1)^3}\right)^{\frac{1}{4}} \approx \frac{0.67}{Q^{\frac{1}{4}}}, \quad Q \to \infty. \tag{3.35}$$

Let Q be 16. Then $\varepsilon_2(\sigma) \le \sqrt{2\sigma M_2} \cdot 0.34$. Therefore there is a gain in accuracy compared with (2.17). Estimates (3.34) and (3.35) show that, in principle, it is possible to attain an arbitrary accuracy of the approximation by taking Q sufficiently large.

Remark 3.4. It follows from (3.20), (3.27) that the multipoint differentiator can be written as

$$T_h^Q f_\sigma := \frac{3}{h(Q+1)(2Q+1)} \sum_{j=-Q}^{Q} j f_\sigma\left(x + \frac{jh}{Q}\right), \tag{3.36}$$

or in the equivalent form

$$T_{\tilde{h}}^Q f_\sigma := \frac{\sum_{j=-Q}^{Q} j f_\sigma(x + j\tilde{h})}{2\tilde{h}\sum_{j=-Q}^{Q} j^2}, \quad \tilde{h} := \frac{h}{Q}, \tag{3.37}$$

which coincides with the least squares differentiator derived by Lanczos ([11]) and also investigated by Anderssen and de Hoog ([3]). However, neither in [11] nor in [3] the special choice of h by formula (3.32) was proposed to guarantee a better accuracy of the approximations.

3.2.1. *Construction of finite-difference methods without using the a priori estimate of M_2.* To calculate h by formula (3.32) one needs to know σ^2 and M_2. If the constants σ^2 and M_2 cannot be found a priori, then one can use the following considerations. By (3.29), (3.30) and (3.31) for small h and large $Q \geq 1$ one has

$$E\left[\left(T_h^Q f_\sigma - f'\right)^2\right] \leq \frac{3\sigma^2 Q}{h^2(Q+1)(2Q+1)} + \frac{h^2 M_2^2}{2} \sim \frac{3\sigma^2}{2h^2 Q}, \tag{3.38}$$

which means that the right-hand side of (3.38) is bounded if

$$h^2 Q \longrightarrow \infty \quad \text{as} \quad h \to 0. \tag{3.39}$$

The condition $h \to 0$ guarantees that the term $\frac{h^2 M_2^2}{2}$ neglected in (3.38) is small. If it is assumed that $Q = h^{-p}$, then (3.39) holds if $p > 2$. Therefore, the simple strategy in the case when σ^2 and M_2 are not available is to ignore the choice of h by (3.32), to take h sufficiently small, and then to choose Q in such a way that (3.39) holds. There is no efficient error estimate (with explicitly given estimation constants as in (2.12)-(2.13)) in this case.

4. Continuous regularization and numerical differentiation in the class $\mathcal{G}(\delta)$

4.1. **Statement of the problem.** Suppose that f is the function to be differentiated and z is its unknown derivative. Then z satisfies the following Volterra equation

$$\int_a^x z(s)\,ds = f(x), \quad a \leq x \leq b < \infty, \quad z \in L^2[a,b]. \tag{4.1}$$

Below we assume that $f(a) = 0$. If $f(a) \neq 0$, then (4.1) is inconsistent, but the problem is essentially unchanged since the derivative of a constant is zero.

In this section we make use of the approach developed in [1] for solving nonlinear operator equations

$$A(z) = f, \quad A : H \to H, \tag{4.2}$$

where H is a Hilbert space and the Fréchet derivative of A is not assumed to be boundedly invertible. To deal with the ill-posedness of (4.2), the following Cauchy problem

$$\dot{z}(t) := \frac{dz}{dt} = \Phi(z(t),t), \quad z(0) = z_0 \in H, \quad \Phi : H \times [0,+\infty) \to H, \tag{4.3}$$

is considered and a general theorem on a continuous regularization in form (4.3) is proved (see Theorem 2.4, [1]). As a consequence of this theorem different types of the operators Φ are investigated in [1]. In particular, it is shown that for monotone operators A, one can write Φ as

$$\Phi(z(t),t) := -[A(z(t)) - f + \varepsilon(t)(z(t) - z_0)], \quad \varepsilon(t) > 0. \tag{4.4}$$

Let $\rightharpoonup$ denote weak convergence in a Hilbert space H. To state the main asymptotic property of the solution to (4.3)-(4.4) we will need the following definition.

Definition 4.1. The operator A in a Hilbert space H is said to be monotone if $(A(z) - A(y), z - y) \geq 0$ for all z and y in the domain of A.

Theorem 4.2 ([1]). *If*

1. *problem* (4.2) *has a unique solution* $y \in H$*;*
2. A *is monotone;*
3. A *is continuously Fréchet differentiable and*

$$||A'(x)|| \le N_1, \quad \forall x \in H; \tag{4.5}$$

4. $\varepsilon(t) > 0$ *is continuously differentiable and tends to zero monotonically as* $t \to +\infty$, *and* $\lim_{t\to+\infty} \frac{\dot\varepsilon(t)}{\varepsilon^2(t)} = 0$.

Then Cauchy problem (4.3)-(4.4) *has a unique solution* $z(t)$ *for all* $t \in [0, +\infty)$ *and*

$$\lim_{t\to+\infty} ||z(t) - y|| = 0 \tag{4.6}$$

in the norm of a Hilbert space H.

Remark 4.3. One can see that for linear operator

$$A(z) := \int_a^x z(s)\, ds, \quad H = L^2[a, b], \tag{4.7}$$

conditions 2 and 3 of Theorem 4.2 hold. Indeed,

$$(Ah, h) = \int_a^b \left(\int_a^x h(\tau) d\tau \right) h(x) dx = \frac{1}{2} \int_a^b \left[\left(\int_a^x h(\tau) d\tau \right)^2 \right]' dx \ge 0.$$

Here (,) denotes the inner product in a real Hilbert space $L^2[a, b]$. Also since A is bounded, the constant $N_1 = ||A|| \le (b - a)/\sqrt{2}$. The proof is immediate:

$$||A||^2 = \int_a^b \left[\int_a^x z(s) ds \right]^2 dx \le \int_a^b \int_a^x z^2(s) ds \int_a^x ds\; dx \le ||z||^2 \frac{(b - a)^2}{2}.$$

Finally, note that equation (4.1) is uniquely solvable if f is absolutely continuous and $f' \in L^2[a, b]$, and it is not solvable otherwise.

Let us use Euler's method to solve the Cauchy problem (4.3)-(4.4) numerically.

$$p_{n+1} = p_n - h_n[A(p_n) - f_\delta + \varepsilon_n(p_n - z_0)], \quad n = 0, 1, 2, ..., \tag{4.8}$$

$$p_0 := z_0, \quad \varepsilon_n := \varepsilon(t_n), \quad t_n := \sum_{i=1}^n h_i, \quad h_n > 0, \tag{4.9}$$

where $f_\delta \in H$ and $||f - f_\delta|| \le \delta$ in H-norm.

4.2. **Convergence and stability analysis.** The goal of this subsection is to prove that under the assumptions of Theorem 4.2 and under the assumptions on h_n, ε_n and δ, listed in Theorem 4.5, one gets

$$\lim_{\delta\to 0} ||p_{n(\delta)} - z(t_{n(\delta)})|| = 0. \tag{4.10}$$

The steps of the proof are:

1) we prove inequality (4.18) below,
2) we prove (4.22) for the solution to a difference inequality (4.20),
3) we apply (4.20) and (4.22) to (4.18).

The limit in (4.10) is understood in the sense of H-norm and $n(\delta) \to \infty$ as $\delta \to 0$. Equalities (4.6) and (4.10) together imply that (4.8)-(4.9) generate a stable numerical scheme for solving (4.2), which in particular can be applied in the case of equation (4.1) if $f \in \mathcal{G}(\delta)$, where

$$\mathcal{G}(\delta) := \left\{ f: \ f \text{ is absolutely continuous}, \ f' \in L^2[a,b], \ ||f - f_\delta|| \le \delta \right\}. \tag{4.11}$$

In (4.11) the L^2-norm is taken. Note that if the function $z(t)$ satisfies (4.3)-(4.4), one has

$$\begin{aligned} z(t_{n+1}) &= z(t_n) + h_n \dot{z}(t_n) + \frac{h_n^2}{2}\ddot{z}(\xi_n) \\ &= z(t_n) - h_n[A(z(t_n)) - f + \varepsilon(t_n)(z(t_n) - z_0)] \\ &\quad - \frac{h_n^2}{2}\Big\{ [A'(z(\xi_n)) + \varepsilon(\xi_n) I] \\ &\quad \cdot [A(z(\xi_n)) - f + \varepsilon(\xi_n)(z(\xi_n) - z_0)] + \dot{\varepsilon}(\xi_n)(z(\xi_n) - z_0) \Big\}. \end{aligned} \tag{4.12}$$

Denote

$$z_n := z(t_n), \quad \tilde{z}_n := z(\xi_n), \quad \tilde{\varepsilon}_n := \varepsilon(\xi_n). \tag{4.13}$$

Then

$$\begin{aligned} ||z_{n+1} - p_{n+1}|| \le {} & ||z_n - p_n - h_n[A(z_n) - A(p_n) + \varepsilon_n(z_n - p_n)]|| \\ & + h_n\delta + \frac{h_n^2}{2}\Big\{ ||A'(\tilde{z}_n) + \tilde{\varepsilon}_n I|| \\ & \cdot ||A(\tilde{z}_n) - f + \tilde{\varepsilon}_n(\tilde{z}_n - z_0)|| + |\dot{\varepsilon}(\xi_n)| \cdot ||\tilde{z}_n - z_0|| \Big\}. \end{aligned} \tag{4.14}$$

Suppose that conditions of Theorem 4.2 are fulfilled. Introduce the notation

$$\lambda_n := (N_1 + \tilde{\varepsilon}_n)||A(\tilde{z}_n) - f + \tilde{\varepsilon}_n(\tilde{z}_n - z_0)|| + |\dot{\varepsilon}(\xi_n)| \cdot ||\tilde{z}_n - z_0||. \tag{4.15}$$

Since $A(y) = f$, by Theorem 4.2 one gets

$$\begin{aligned} \lim_{n\to\infty} \lambda_n \le \lim_{n\to\infty} \{ & (N_1 + \tilde{\varepsilon}_n)(N_1||\tilde{z}_n - y|| + \tilde{\varepsilon}_n||\tilde{z}_n - z_0||) \\ & + |\dot{\varepsilon}(\xi_n)| \cdot ||\tilde{z}_n - z_0|| \} = 0. \end{aligned} \tag{4.16}$$

If $h_n\varepsilon_n \le 1$, then from the monotonicity of A one obtains

$$\begin{aligned} ||z_{n+1} - p_{n+1}|| \le {} & \Big\{ ||z_n - p_n||^2 (1 - h_n\varepsilon_n)^2 \\ & - 2h_n(1 - h_n\varepsilon_n)(A(z_n) - A(p_n), z_n - p_n) \\ & + h_n^2 ||A(z_n) - A(p_n)||^2 \Big\}^{\frac{1}{2}} + h_n\delta + \frac{h_n^2\lambda_n}{2} \\ \le {} & \Big\{ ||z_n - p_n||^2 \left(1 - 2h_n\varepsilon_n + h_n^2\varepsilon_n^2 + h_n^2 N_1^2\right) \Big\}^{\frac{1}{2}} + h_n\delta + \frac{h_n^2\lambda_n}{2}. \end{aligned} \tag{4.17}$$

Applying the elementary estimate

$$(a+b)^2 \leq (1+h_n\varepsilon_n)a^2 + \left(1+\frac{1}{h_n\varepsilon_n}\right)b^2$$

to the right-hand side of (4.17) with

$$a := \left\{||z_n - p_n||^2\left(1-2h_n\varepsilon_n + h_n^2\varepsilon_n^2 + h_n^2N_1^2\right)\right\}^{\frac{1}{2}}$$

and

$$b := h_n\delta + \frac{h_n^2\lambda_n}{2},$$

one gets

$$||z_{n+1} - p_{n+1}||^2 \leq \left(1 - h_n\varepsilon_n + ch_n^2\right)||z_n - p_n||^2 + d\frac{h_n(\delta+h_n)^2}{\varepsilon_n}. \tag{4.18}$$

Here we assume that h_n tends to zero monotonically as $n \to \infty$,

$$c := (1+h_0\varepsilon_0)\left(\varepsilon_0^2 + N_1^2\right), \quad d := \left[\max\left\{1,\ \frac{1}{2}\sup_n \lambda_n\right\}\right]^2(1+h_0\varepsilon_0). \tag{4.19}$$

To state the result we need the following lemma on recursive numerical sequences. To make the paper self-contained we include the proof.

Lemma 4.4. *Let the sequence of positive numbers ν_n satisfy the inequality*

$$\nu_{n+1} \leq (1-\alpha_n)\nu_n + \beta_n, \tag{4.20}$$

where

$$0 < \alpha_n \leq 1, \quad \sum_{n=0}^{\infty}\alpha_n = \infty, \quad \lim_{n\to\infty}\frac{\beta_n}{\alpha_n} = 0. \tag{4.21}$$

Then

$$\lim_{n\to\infty}\nu_n = 0. \tag{4.22}$$

Proof. Inequality (4.20) yields

$$\begin{aligned}\nu_{n+1} &\leq (1-\alpha_n)(1-\alpha_{n-1})\nu_{n-1} + (1-\alpha_n)\beta_{n-1} + \beta_n \\ &\leq \dots \leq \prod_{j=0}^{n}(1-\alpha_j)\nu_0 + \sum_{i=0}^{n-1}\beta_i\prod_{j=i+1}^{n}(1-\alpha_j) + \beta_n.\end{aligned} \tag{4.23}$$

By (4.21) $\lim_{n\to\infty}\prod_{j=0}^{n}(1-\alpha_j) = 0$ and $\lim_{n\to\infty}\beta_n = 0$. Thus we have to show that

$$\lim_{n\to\infty}\sum_{i=0}^{n-1}\beta_i\prod_{j=i+1}^{n}(1-\alpha_j) = 0. \tag{4.24}$$

Without loss of generality one may assume that $\beta_n \geq 0$ for any n: if $\beta_n < 0$, then (4.20) holds with $|\beta_n|$ replacing β_n. To get (4.24) we will use the representation

$$\sum_{i=0}^{n-1}\beta_i\prod_{j=i+1}^{n}(1-\alpha_j) = \sum_{i=0}^{m}\beta_i\prod_{j=i+1}^{n}(1-\alpha_j) + \sum_{i=m+1}^{n-1}\beta_i\prod_{j=i+1}^{n}(1-\alpha_j). \tag{4.25}$$

Here m is an arbitrary integer less than n. The first term in the right-hand side of (4.25) can be estimated as

$$\sum_{i=0}^{m}\beta_i\prod_{j=i+1}^{n}(1-\alpha_j)=\sum_{i=0}^{m}\frac{\beta_i}{\alpha_i}\left[\prod_{j=i+1}^{n}(1-\alpha_j)-\prod_{j=i}^{n}(1-\alpha_j)\right]$$
$$\le\zeta\left[\prod_{j=m+1}^{n}(1-\alpha_j)-\prod_{j=0}^{n}(1-\alpha_j)\right] \tag{4.26}$$
$$\le\zeta\exp\left(-\sum_{j=m+1}^{n}\alpha_j\right),$$

where $\frac{\beta_i}{\alpha_i}\le\zeta,\ i=0,1,2,\dots$. Introduce the notation $\varrho_{m,n}:=\max_{m+1\le i\le n-1}\frac{\beta_i}{\alpha_i}$. For the second term in (4.25) one obtains

$$\sum_{i=m+1}^{n-1}\beta_i\prod_{j=i+1}^{n}(1-\alpha_j)\le\varrho_{m,n}\sum_{i=m+1}^{n-1}\left[\prod_{j=i+1}^{n}(1-\alpha_j)-\prod_{j=i}^{n}(1-\alpha_j)\right]$$
$$=\varrho_{m,n}\left[1-\alpha_n-\prod_{j=m+1}^{n}(1-\alpha_j)\right]\le\varrho_{m,n}. \tag{4.27}$$

Therefore

$$\lim_{n\to\infty}\sum_{i=0}^{n-1}\beta_i\prod_{j=i+1}^{n}(1-\alpha_j)\le\limsup_{m\to\infty}\left\{\lim_{n\to\infty}\left[\zeta\exp\left(-\sum_{j=\tilde{n}+1}^{n}\alpha_j\right)+\varrho_{m,n}\right]\right\}$$
$$\le\limsup_{m\to\infty}\left(0+\sup_{m+1\le n}\frac{\beta_n}{\alpha_n}\right)=\limsup_{n\to\infty}\frac{\beta_n}{\alpha_n}=0. \tag{4.28}$$

The last conclusion is a consequence of the last assumption in (4.21). From (4.28), (4.24) and (4.23) one concludes that (4.22) holds. □

Finally we present the convergence theorem.

Theorem 4.5. *Assume that conditions of Theorem 4.2 hold. Let*

1. *δ be the level of noise in* (4.8)*: $||f-f_\delta||\le\delta$;*
2. *$n=n(\delta)$ be chosen in such a way that $\lim_{\delta\to0}n(\delta)=\infty$;*
3. *$h_{n(\delta)}$ tend to zero monotonically as $\delta\to0$, $0<h_{n(\delta)}\varepsilon_{n(\delta)}-ch^2_{n(\delta)}\le1$ with c defined by* (4.19)*;*
4. $\sum_{n=1}^{\infty}h_n\varepsilon_n=\infty,\quad \lim_{\delta\to0}\frac{h_{n(\delta)}}{\varepsilon_{n(\delta)}}=0,\quad \lim_{\delta\to0}\frac{\delta}{\varepsilon_{n(\delta)}}=0.$

Then

$$\lim_{\delta\to0}||p_{n(\delta)}-y||=0 \tag{4.29}$$

in the norm of a Hilbert space H, where $\{p_{n(\delta)}\}$ is defined by (4.8)-(4.9) *and y is the solution to* (4.2).

Proof. Let us take $\nu_n:=||z_n-p_n||^2$, $\alpha_n:=h_n\varepsilon_n-ch_n^2$ and $\beta_n:=d\frac{h_n(\delta+h_n)^2}{\varepsilon_n}$. Inequality (4.29) is an immediate consequence of (4.18) and Lemma 4.4. □

Thus it is shown that under the assumptions of Theorem 4.5, procedure (4.8)-(4.9) can be used for stable numerical differentiation of a function from the class $\mathcal{G}(\delta)$, which is defined in (4.11).

5. Numerical aspects

5.1. **Multi-point formulas in the class $\mathcal{K}(\delta, M_m)$, $m > 2$.** MATLAB programs were written to investigate the practicability of the methods described above. The derivative of the function $f(x) = sin(\pi x)$, $x \in [0,1]$, was computed in the presence of noise functions

$$e_1(x) = \delta x, \quad e_2(x) = \delta cos(3\pi x), \quad e_3(x) = \delta cos(7\pi x) \tag{5.1}$$

and their combinations. The constant δ in (5.1) changes from 0.01 to 0.1.

Our first experiment makes use of formulas (3.1), (3.6) with the numbers A_j^Q ($j = -Q, ..., Q$) defined by (3.4) and $m = 2Q + 1$. The goals of the test are the following:

1. To compare numerical results obtained by (3.1), (3.6) with $Q = 1, 2, 3, 4$; to consider the dependence of estimated error (3.7) and of the computed error on Q (on m). Note that $Q = 1, 2, 3, 4$ correspond to $m = 3, 5, 7, 9$, respectively.
2. To evaluate the dependence of the computed and estimated errors on the perturbations of M_m (see (3.5)-(3.7)), since in practice only the approximate bounds on the L^∞-norm of $f^{(m)}$ might be available.
3. To study the dependence of the results on δ, because the optimal stepsize $h_m(\delta)$ and the estimated error $\varepsilon_m(\delta)$ are functions of δ.

Figure 1

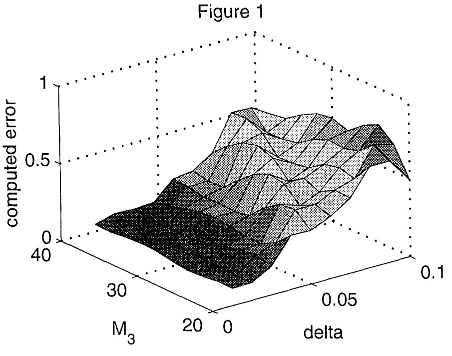


Figure 2

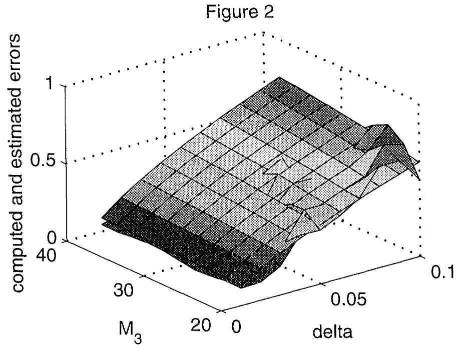


Figure 3

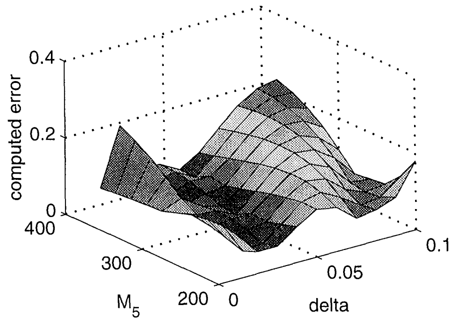


Figure 4

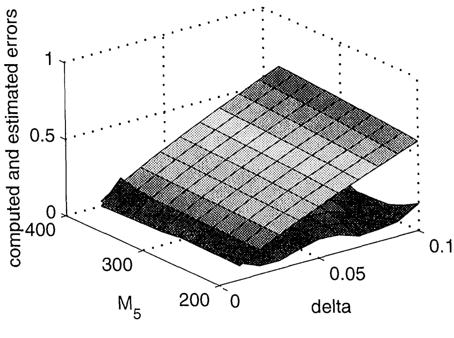


Figure 5

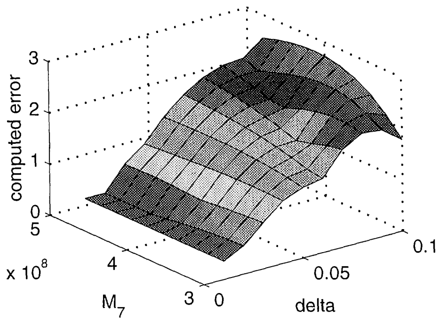


Figure 6

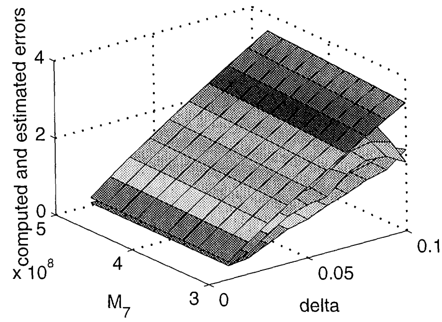


Figure 7

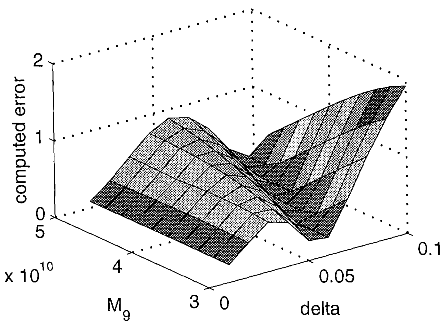
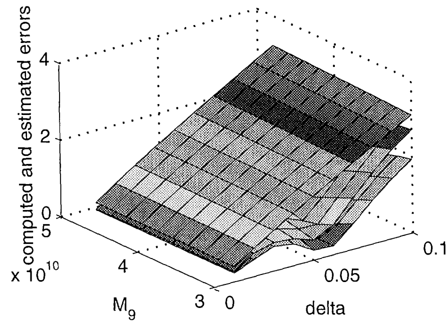


Figure 8



Figures 1 and 2 represent the errors obtained in the process of numerical differentiation by formula (3.1) with $Q = 1$ and $h = h_3(\delta)$ being chosen by (3.6). The function to be differentiated is $f_\delta(x) = f(x) + e_3(x)$. The upper surface on Figure 2 is the estimated error. Note that $f(x) = sin(\pi x)$ belongs to the class $\mathcal{K}(\delta, M_3)$ for $M_3 \geq \pi^3 \approx 31.0063$. So if one uses $20 \leq M_3 < \pi^3$, then $h_3(\delta)$ found by (3.6) is not optimal and cannot guarantee error (3.7). Therefore the computed error in some regions, where $20 \leq M_3 < \pi^3$, is greater then the estimated one.

Figures 3 and 4 illustrate the same experiment with $Q = 2$. There is a little gain in accuracy for both estimated and computed errors. The optimal stepsize $h_5(\delta)$ is approximately three times as big as $h_3(\delta)$.

One can see on Figures 5-8 that for $Q = 3$ $(m = 7)$ and $Q = 4$ $(m = 9)$ the results are not accurate. The reason is that the constant β_m (see (3.5)), which does not depend on f, decreases very fast as $m = 2Q + 1$ changes from 5 to 9:

$$\beta_5 \approx 3.47 \cdot 10^{-3}, \quad \beta_7 \approx 3.07 \cdot 10^{-5}, \quad \beta_9 \approx 1.50 \cdot 10^{-7}.$$

Since β_m appears in the denominator of (3.6), one has to take M_m large; otherwise, $h_m(\delta)$ becomes so big that it cannot be used on the interval $[0, 1]$. Namely, for $M_7 = \pi^7$, $M_9 = \pi^9$ and $\delta = 0.01$, one gets

$$h_7 \approx 0.7185, \qquad h_9 \approx 1.0985.$$

For the same M_7, M_9 and $\delta = 0.1$

$$h_7 \approx 0.9986, \qquad h_9 \approx 1.4188.$$

The ratio $\frac{\alpha_m}{m-1}$ (see (3.6)) also grows as m increases:

$$\frac{\alpha_5}{4} \approx 0.75, \qquad \frac{\alpha_7}{6} \approx 0.92, \qquad \frac{\alpha_9}{8} \approx 1.042.$$

Thus to obtain $h_7(\delta)$ and $h_9(\delta)$ approximately at the same level as $h_3(\delta)$ for $20 \leq M_3 \leq 40$, one needs $M_7 \sim 10^8$ and $M_9 \sim 10^{10}$. The estimated errors $\varepsilon_7(\delta)$, $\varepsilon_9(\delta)$ for these M_7 and M_9 are much greater than $\varepsilon_3(\delta)$: one can compare the upper surfaces on Figures 6, 8 and 2. The computed errors are less than the estimated ones, but even they are almost 100% for $\delta = 0.1$. Figures 5 and 7 show the errors for $f_\delta(x) = f(x) + e_2(x)$. On Figures 6 and 8 the lower surfaces correspond to $f_\delta(x) = f(x) + e_2(x)$ and $f_\delta(x) = f(x) + e_2(x) + e_3(x)$.

The conclusion is: practically, one does not have a good reason to use formulas (3.1), (3.6) for $m > 5$ in many cases. Besides, even for $m = 5$ sometimes the best possible constant M_5 is rather small and together with small β_5 they make $h_5(\delta)$ larger than is appropriate on a specific interval. Therefore one has to increase the constant M_5, which in turn increases $\varepsilon_5(\delta)$ in (3.7). Then one can compare $\varepsilon_5(\delta)$ with $\varepsilon_3(\delta)$ and make a choice between (3.1), (3.6) for $m = 3$ and (3.1), (3.6) for $m = 5$.

5.2. **Multi-point formulas in the class $\mathcal{K}(\delta, M_2)$.** In our second experiment we investigate multi-point differentiator (3.20) with h satisfying (3.32). Here we use the statistical nature of noise assuming zero mean value and varance σ^2. The perturbated function $f_\sigma(x) = sin(\pi x) + \sqrt{2}\,\sigma\, cos(2\pi x)$, $x \in [0,1]$, is to be differentiated. The goal is to consider the dependence of the computed and estimated errors on Q, M_2 and σ.

Figure 9

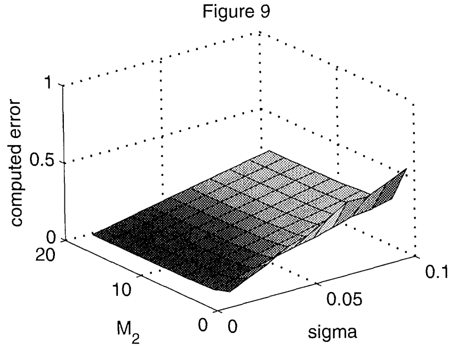


Figure 10

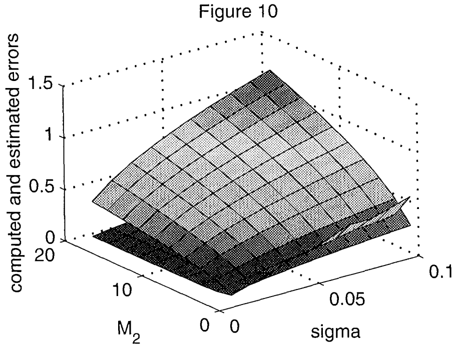


Figure 11

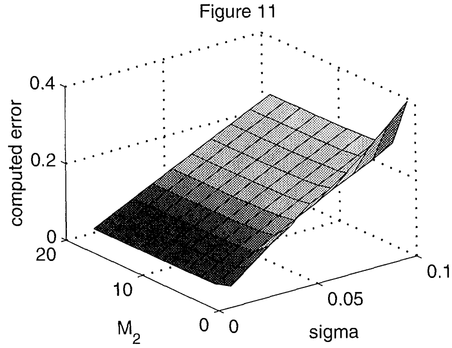


Figure 12

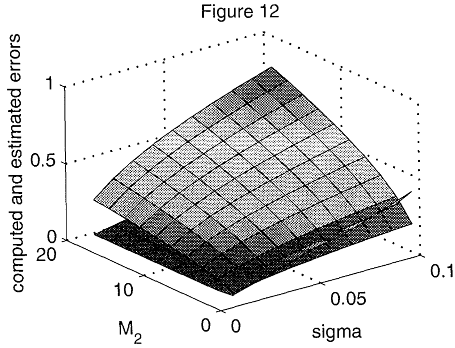


Figure 13

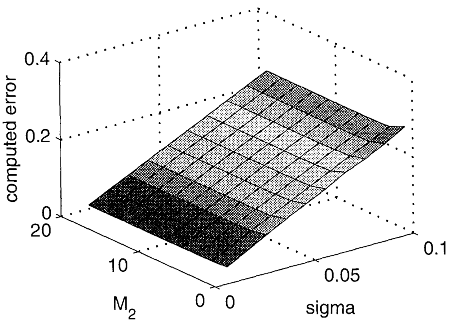


Figure 14

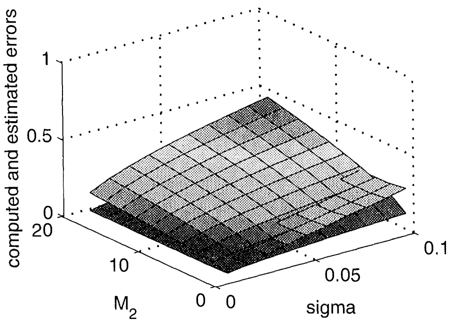


Figure 15

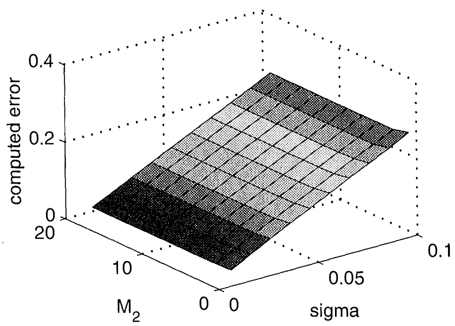
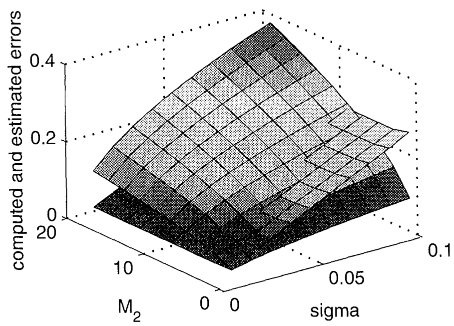


Figure 16



Figures 9 and 10 illustrate the results obtained with $Q = 1$. The function $f(x) = sin(\pi x)$ belongs to the class $\mathcal{K}(\sigma, M_2)$ for $M_2 \geq \pi^2 \approx 9.8696$. That is the reason why the computed errors on Figures 10, 12, 14 and 16 are greater than the estimated ones in some areas, where $0.1 \leq M_2 < \pi^2$.

Figures 11 and 12 correspond to the case $Q = 7$. One can see that both the estimated and the computed errors are less than for $Q = 1$.

Theoretically, by formula (3.34), one can attain an arbitrary accuracy of the approximation by taking Q sufficiently large. However, in practice we do not recommend using $Q > 50$. Indeed, denote $\mu(Q) := \sqrt{2\sigma M_2}(a_Q b_Q)^{\frac{1}{4}}$, where a_Q and b_Q satisfy (3.30) and (3.31), respectively. Then for $\sigma = 0.1$, $M_2 = 20$, one has

$$\mu(50) \approx 0.50843, \quad \mu(75) \approx 0.45981, \quad \mu(100) \approx 0.42808, \quad \mu(150) \approx 0.38698.$$

Thus the accuracy decreases slowly as Q grows. The estimated errors for $Q = 75$ (see Figure 13) and $Q = 150$ (see Figure 15) are almost identical because of the roundoff error. For $Q = 200$, the estimated error is even bigger than the one for $Q = 150$.

5.3. **Numerical solution of regularized Volterra equation.** Our last numerical experiment on reconstruction of the derivative is based on Euler's method (4.8)-(4.9) for solving the Cauchy problem (4.3)-(4.4). The operator A in (4.4) is defined by (4.7) with $a = 0$ and $b = 1$. We assume that the function to be differentiated is

$$f_\delta(x) = sin(\pi x) + \delta cos(3\pi x).$$

As in previous subsections we take $\delta \in [0.01, 0.1]$, which corresponds to the level of noise 1% – 10% of the maximum value of $f(x) = sin(\pi x)$. The choice of the

Figure 17

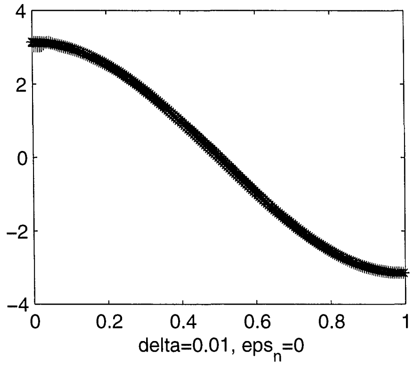


Figure 18

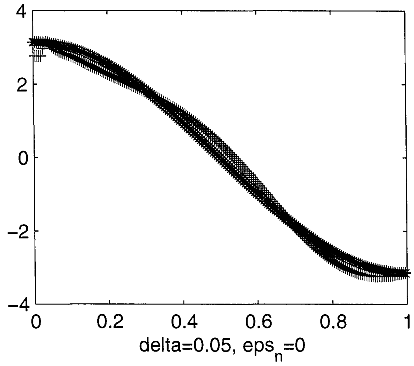


Figure 19

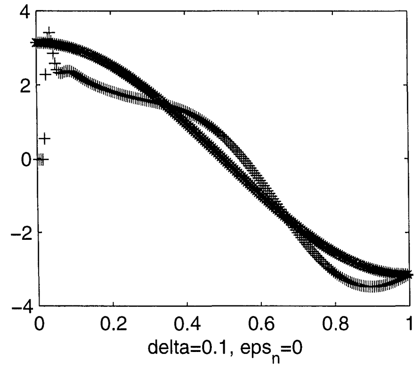
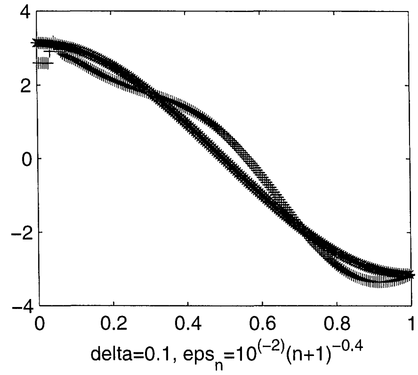


Figure 20



parameters h_n and ε_n in (4.8)-(4.4) is dictated by condition 4 of Theorem 4.2 and by condition 4 of Theorem 4.5. Therefore, we use

$$h_n = h_0\,(n+1)^{-0.5}, \qquad \varepsilon_n = \varepsilon_0\,(n+1)^{-0.4}.$$

Figures 17-20 contain the exact derivative $f'(x) = \pi cos(\pi x)$ and the derivatives computed by (4.8)-(4.9) with $\delta = 0.01$, 0.05 and 0.1. The argument x changes from 0 to 1 along the horizontal axis, and eps_n on Figures 17-20 denotes ε_n. The results in these figures correspond to the initial guess $p_0 = 0$, the number of iterations $N = 10$ and $h_0 = 100$. Actually, any $h_0 \in [0.5, 100]$ provides convergence for $\delta \in [0.01, 0.1]$, the difference is in its rate. If $h_0 = 100$ and $\delta = 0.01$, then the discrepancy after ten iterations is $\sim 10^{-17}$. As one can see from Figures 17 and 18 for the levels of noise 1% and 5% the computed derivatives are found with $\varepsilon_0 = 0$. The use of $\varepsilon_0 \neq 0$ does not give any gain in accuracy here. The reason for such a phenomenon is probably the self-regularization of Volterra equations in the process of their discrete approximation. This means that Volterra integral equations allow one to generate stable numerical methods by application of quadrature formula directly to the initial equation and the following theorem holds.

Theorem 5.1 ([29]). *Let*

$$(5.2) \qquad \int_a^x K(x,s)z(s)\,ds = f_\delta(x), \quad ||f_\delta - f||_{C[a,b]} \leq \delta, \quad a \leq x \leq b,$$

and assume

1. $\min_{a\le x\le b}|K(x,x)|>0$, $K'_x(x,s)\in C(\Delta)$, $K''_{xs}(x,s)\in C(\Delta)$, *where*
$$\Delta := \left\{x, s:\ a\le s\le x\le b\right\};$$
2. $f(x)\in C^{(1)}[a,b]$, $f_\delta(x)\in C[a,b]$, $f(a)=f_\delta(a)=0$.

Then

$$\max_{1\le i\le k}|\hat{y}(x_i)-\tilde{y}_i|\le c_1\tau+c_2\,\frac{\delta}{\tau}. \tag{5.3}$$

Here $\hat{y}(x)$ *is the unique solution to* (5.2); $\tilde{y}=(\tilde{y}_1,\tilde{y}_2,\dots,\tilde{y}_k)$ *is a unique solution to the linear system*

$$\tau\sum_{j=1}^{i}K(x_i,x_j)z(x_j)=f_\delta(x_i),\quad i=1,2,\dots,k, \tag{5.4}$$

$x_i=a+i\tau$, $\tau=\frac{b-a}{k}$; *the constants* c_1, c_2 *do not depend on* τ *and* δ.

The right-hand side of (5.3) tends to 0 as $\delta\to 0$ if $h(\delta)=c\delta^\gamma$, $0<\gamma<1$. Note that in our case the functions $f(x)$ and $f_\delta(x)$ satisfy condition 2 of Theorem 5.1. The fact that the assumption $f(a)=f_\delta(a)=0$ is natural is briefly explained below formula (4.1).

For $\delta=0.1$, the result obtained with $\varepsilon_0=0$ is worse than the result obtained with $\varepsilon_0\in[10^{-3},10^{-2}]$. One can compare Figures 19 and 20. The main difficulty occurs in the neighborhood of 0, and $\varepsilon_0\in[10^{-3},10^{-2}]$ helps to get better approximation.

Scheme (4.8)-(4.9) can be realized without knowledge (and even existence) of the constants M_m, since the regularization parameter ε_n is the function of δ only. This is an advantage of the approach based on (4.8)-(4.9). However, this approach does not give the explicitly computable estimation constants. Such estimates can be found under the additional assumptions on p_0 (see [1]).

References

1. R. G. Airapetyan, A. G. Ramm, A. B. Smirnova, "Continuous methods for solving nonlinear ill-posed problems", *Operator theory and applications*, Fields Institute Communications, 25, Amer. Math. Soc., Providence, RI, 2000, pp. 111-138. CMP 2000:13
2. R. S. Anderssen, P. Bloomfield, *Numerical differentiation procedures for non-exact data,* Numer. Math. **22** (1973/74), 157-182. MR **50:**1470
3. R. S. Anderssen, F. R. de Hoog, *Finite difference methods for the numerical differentiation of non-exact data,* Computing **33** (1984), 259-267. MR **86e:**65032
4. J. Cullum, *Numerical differentiation and regularization,* SIAM J. Numer. Analysis. **8** (1971), 254-265. MR **44:**7747
5. T. F. Dolgopolova, *Finite dimensional regularization in the case of numerical differentiation of periodic functions,* Ural. Gos. Univ. Mat. Zap. **7**(4) (1970), 27-33. MR **44:**6148
6. T. F. Dolgopolova, V. K. Ivanov, *Numerical differentiation,* Comp. Math and Math. Physics. **6**(3) (1966), 570-576. MR **33:**8098
7. Yu. V. Egorov, V. A. Kondrat'ev, *On a problem of numerical differentiation,* Vestnik Moskov. Univ. Ser. I Mat. Mekh. **3** (1989), 80-81. MR **91a:**65048
8. V. K. Ivanov, *On linear problems, which are not well-posed*, Doklady Akad. Nauk SSSR. **145**(2) (1962), 270-272. MR **25:**4357
9. I. Knowles, R. Wallace, *A variational method for numerical differentiation,* Numer. Math. **70** (1995), 91-110. MR **96h:**65031
10. E. V. Kolpakova, *Numerical solution of the problem of reconstructing the derivative,* Differencial'nye Uravnenija i Vychisl. Mat. **6** (1976), 137-143. MR **58:**8159
11. C. Lanczos, *Applied Analysis,* Englewood Cliffs, N.J. Prentice-Hall, 1956. MR **18:**823c

12. T. Miller, A. G. Ramm, *Estimates of the derivatives of random functions II,* J. Math. Anal. Appl. **110** (1985), 429-435. MR **87f:**93076
13. C. K. Pallaghy, U. Luttge, *Light-induced and H^*-ion fluxes and bioelectric phenomena in mesophyll cells of Atriplex Spongiosa,* Zeit. fuer Pflanz. **62** (1970), 417-425.
14. D. L. Phillips, *A technique for the numerical solution of certain integral equations of the first kind,* J. Assoc. Comput. Machinery. **9**(1) (1962), 84-97. MR **24:**B534
15. R. Qu, *A new approach to numerical differentiation and regularization,* Math. Comput. Modelling. **24**(10) (1996), 55-68. MR **98b:**65024
16. A. G. Ramm, *On numerical differentiation,* Izv. Vyssh. Uchebn. Zaved. Mat., **11** (1968), 131-135. MR **40:**5130
17. ______, *Simplified optimal differentiators,* Radiotech.i Electron. **17** (1972), 1325-1328.
18. ______, *On simultaneous approximation of a function and its derivative by interpolation polynomials,* Bull. Lond. Math. Soc. **9** (1977), 283-288. MR **58:**6823
19. ______, *Stable solutions of some ill-posed problems,* Math. Methods Appl. Sci. **3** (1981), 336-363. MR **83i:**65102
20. ______, *Estimates of the derivatives of random functions,* J. Math. Anal. Appl., **102** (1984), 244-250. MR **85m:**93040
21. ______, *Random fields estimation theory,* Longman Scientific and Wiley, New York, 1990. MR **92g:**00012
22. ______, *Inequalities for the derivatives,* Math. Inequal. Appl., **3** (2000), 129-132. MR **2000i:**65026
23. A. G. Ramm, A. I. Katsevich, *The Radon transform and local tomography,* CRC Press, Boca Raton, Florida, 1996. MR **97g:**44009
24. T. J. Rivlin, *Optimally stable Lagrangian numerical differentiation,* SIAM J. Numer. Anal. **12**(5) (1975), 712-725. MR **53:**11968
25. H. E. Salzer, *Some problems in optimally stable Lagrangian differentiation,* Math. Comp. **28** (1974), 1105-1115. MR **51:**4632
26. A. N. Tikhonov, V. Y. Arsenin, *Solutions of ill-posed problems,* John Wiley and Sons, New York, 1977. MR **56:**13604
27. V. V. Vasin, *Regularization of a numerical differentiation problem,* Ural. Gos. Univ. Mat. Zap. **7**(2) (1969), 29-33. MR **43:**5726
28. V. V. Vasin, *The stable evaluation of a derivative in $C(-\infty, \infty)$,* Comp. Math and Math. Physics. **13**(6) (1973), 1383-1389. MR **49:**11807
29. V. V. Vasin, A. L. Ageev, *Ill-posed problems with a priori information*, VSP, Utrecht, 1995. MR **97j:**65100

Department of Mathematics, Kansas State University, Manhattan, Kansas 66506-2602

E-mail address: ramm@math.ksu.edu

Department of Mathematics, Kansas State University, Manhattan, Kansas 66506-2602

E-mail address: matabs@zeus.cs.gsu.edu