ISSN 0081-5438, Proceedings of the Steklov Institute of Mathematics, 2016, Vol. 293, pp. 194–208. © Pleiades Publishing, Ltd., 2016.
Original Russian Text © G.G. Magaril-Il'yaev, K.Yu. Osipenko, 2016, published in Trudy Matematicheskogo Instituta imeni V.A. Steklova, 2016, Vol. 293, pp. 201–216.

Exactness and Optimality of Methods for Recovering Functions from Their Spectrum

G. G. Magaril-Il'yaev [a,b,c] **and K. Yu. Osipenko** [a,b,d]

Received November 9, 2015

Abstract—Optimal methods are constructed for recovering functions and their derivatives in a Sobolev class of functions on the line from exactly or approximately defined Fourier transforms of these functions on an arbitrary measurable set. The methods are exact on certain subspaces of entire functions. Optimal recovery methods are also constructed for wider function classes obtained as the sum of the original Sobolev class and a subspace of entire functions.

DOI: 10.1134/S0081543816040143

1. INTRODUCTION

The question that was originally addressed by the authors is as follows. Given a class of smooth functions on the line for each of which its Fourier transform on some set is known (in general, approximately), is it possible to construct a method for recovering these functions and/or their derivatives that would be exact on a given subspace and would be the best in a sense?

This question arose as follows. An important characteristic of any quadrature formula is the maximum dimension of the subspace of algebraic or trigonometric polynomials on which this formula is exact. In this sense, optimal formulas are given by Gaussian quadratures (see, e.g., [1]). In the 1950s, there appeared studies devoted to the problem of finding the best quadrature formulas on function classes (Kolmogorov–Nikol'skii quadrature formulas, see [2]). The problem goes back to Kolmogorov's research on finding optimal methods for approximating function classes (see [3]); in [2], Nikol'skii wrote about this even more definitely: "In this section, we give a solution to one of the problems set up by A.N. Kolmogorov."

The problem of Kolmogorov–Nikol'skii quadratures served as a starting point for the general problem of finding optimal methods for recovering linear functionals and operators on classes of elements from inaccurate information on the elements themselves. There is a rather extensive literature devoted to this subject. We point out only the publications [4–12], which can be considered as source publications in this field.

It turns out that among optimal methods for recovering functions and their derivatives on a Sobolev class on the line, there are methods that are exact on some subspaces of entire functions of exponential type. Moreover, these methods are also optimal on a wider class than the original one. Namely, they are optimal on the class obtained as the sum of the original class and the subspace on which these methods are exact.

In this connection, there arises a general problem of constructing methods that are exact on a given subspace of entire functions and are optimal on the sum of the Sobolev class with this

[a] Faculty of Mechanics and Mathematics, Moscow State University, Moscow, 119991 Russia.
[b] Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences, Bol'shoi Karetnyi per. 19, str. 1, Moscow, 127051 Russia.
[c] Peoples' Friendship University of Russia, ul. Miklukho-Maklaya 6, Moscow, 117198 Russia.
[d] Moscow Aviation Institute (National Research University), Volokolamskoe sh. 4, Moscow, 125993 Russia.
E-mail addresses: magaril@mech.math.msu.su (G.G. Magaril-Il'yaev), kosipenko@yahoo.com (K.Yu. Osipenko).

subspace. In the present study, we solve this problem and, as a corollary, find methods for recovering functions and their derivatives that are optimal on the Sobolev class and are exact on a maximally wide subspace of entire functions. In other words, we try to combine two approaches: the one going back to Gauss, which is based on constructing methods exact on subspaces, and the one going back to Kolmogorov, which is based on the constructing methods optimal on a given class.

Note also that similar problems were studied in the authors' earlier publications [13–15].

2. STATEMENTS OF PROBLEMS AND FORMULATIONS OF RESULTS

Let F be the Fourier transform in $L_2(\mathbb{R})$. If $x(\cdot)\in L_2(\mathbb{R})$, then it is convenient to assume that the function $Fx(\cdot)$ is defined on $\mathbb{R}$ with the Lebesgue measure divided by 2π. Denote the norm of a function $y(\cdot)$ in the space of square integrable functions on $\mathbb{R}$ with such a measure by $\|y(\cdot)\|_{\widehat{L}_2(\mathbb{R})}$, i.e.,

$$\|y(\cdot)\|_{\widehat{L}_2(\mathbb{R})}=\left(\frac{1}{2\pi}\int\limits_{\mathbb{R}}|y(\xi)|^2\,d\xi\right)^{1/2}.$$

Let n be a positive integer and $\mathcal{W}_2^n(\mathbb{R})$ be the Sobolev space of functions $x(\cdot)\in L_2(\mathbb{R})$ such that their $(n-1)$th derivative is locally absolutely continuous and $x^{(n)}(\cdot)\in L_2(\mathbb{R})$.

Let, next, W be a subset (class) of functions in $\mathcal{W}_2^n(\mathbb{R})$ and A be a measurable subset of the real line. Assume that for every function $x(\cdot)\in W$, its Fourier transform on A is known either exactly or approximately, i.e., a function $y(\cdot)\in\widehat{L}_2(\mathbb{R})$ is known such that $\|Fx(\cdot)-y(\cdot)\|_{\widehat{L}_2(\mathbb{R})}\le\delta$ for some $\delta>0$.

Given this information, we want to recover (in the best possible way) the functions $x(\cdot)\in W$ and their derivatives up to order $n-1$ inclusive in the $L_2(\mathbb{R})$ metric.

Prior to formulating the problem exactly, we introduce some notation. Let $I_A\colon\mathcal{W}_2^n(\mathbb{R})\to\widehat{L}_2(A)$ be the mapping whose value on a function $x(\cdot)\in\mathcal{W}_2^n(\mathbb{R})$ is the restriction $Fx(\cdot)|_A$ of the function $Fx(\cdot)$ to A, and let $I_A^\delta\colon\mathcal{W}_2^n(\mathbb{R})\to\widehat{L}_2(A)$ be the multivalued mapping defined as

$$I_A^\delta x(\cdot)=\bigl\{y(\cdot)\in\widehat{L}_2(A)\colon\ \|I_Ax(\cdot)-y(\cdot)\|_{\widehat{L}_2(A)}\le\delta\bigr\}.$$

If we formally set $\delta=0$ here, then we get $I_A^0=I_A$; thus, available information on the function $x(\cdot)\in W$ (depending on whether its Fourier transform is known exactly or approximately) is described by a function $y(\cdot)\in I_A^\delta x(\cdot)$, where $\delta\ge0$.

It is clear that any recovery method for the kth $(0\le k\le n-1)$ derivative of a function of class W in the $L_2(\mathbb{R})$ metric from the above information is a mapping $\varphi\colon\widehat{L}_2(A)\to L_2(\mathbb{R})$. By definition, the error of this method is the quantity

$$e(D^k,W,I_A^\delta,\varphi)=\sup_{x(\cdot)\in W,\ y(\cdot)\in I_A^\delta x(\cdot)}\bigl\|x^{(k)}(\cdot)-\varphi(y(\cdot))(\cdot)\bigr\|_{L_2(\mathbb{R})},$$

where D^k denotes the operator of k-fold differentiation (D^0 is the identity operator).

By the problem of *optimal recovery* of the kth $(0\le k\le n-1)$ derivative of a function of class W in the metric of $L_2(\mathbb{R})$ from the above information we mean the problem of finding the quantity

$$E(D^k,W,I_A^\delta)=\inf_{\varphi\colon\ \widehat{L}_2(A)\to L_2(\mathbb{R})}e(D^k,W,I_A^\delta,\varphi),$$

which is called the *error of optimal recovery*, and methods $\widehat{\varphi}$ for which the lower bound is attained, i.e., for which

$$E(D^k,W,I_A^\delta)=e(D^k,W,I_A^\delta,\widehat{\varphi}).$$

Below these methods are called *optimal recovery methods*.

For short, we will refer to the problem formulated as the (D^k,W,I_A^δ)-*problem*.

Along with optimal recovery methods, we will study exact methods. A method $\varphi\colon \widehat{L}_2(A)\to L_2(\mathbb{R})$ is said to be *exact on a set* $L\subset \mathcal{W}_2^n(\mathbb{R})$ if $x^{(k)}(\cdot)=\varphi(I_Ax(\cdot))(\cdot)$ for all $x(\cdot)\in L$. The following proposition shows that the optimality and exactness of a method are not independent concepts.

Proposition 1. *If $\widehat{\varphi}$ is an optimal linear method in the (D^k,W,I_A^δ)-problem that is exact on a set $L\subset \mathcal{W}_2^n(\mathbb{R})$ containing zero, then it is also optimal in the $(D^k,W+L,I_A^\delta)$-problem and, in addition, $E(D^k,W,I_A^\delta)=E(D^k,W+L,I_A^\delta)$.*

If $\widehat{\varphi}$ is a linear method with finite error in the $(D^k,W+L,I_A^\delta)$-problem, where L is a subspace in $\mathcal{W}_2^n(\mathbb{R})$, then it is exact on L.

Proof. Let $x(\cdot)\in W+L$ and $x(\cdot)=x_1(\cdot)+x_2(\cdot)$, where $x_1(\cdot)\in W$ and $x_2(\cdot)\in L$, and let $y(\cdot)\in L_2(A)$ be such that $\|I_Ax(\cdot)-y(\cdot)\|_{L_2(A)}\le\delta$. Set $y_1(\cdot)=y(\cdot)-I_Ax_2(\cdot)$. It is clear that $y_1(\cdot)\in L_2(A)$, and

$$\|I_Ax_1(\cdot)-y_1(\cdot)\|_{L_2(A)}\le\delta \tag{2.1}$$

since $I_Ax_1(\cdot)-y_1(\cdot)=I_Ax(\cdot)-y(\cdot)$. The linearity and exactness of $\widehat{\varphi}$ on L imply the equality

$$\big\|x^{(k)}(\cdot)-\widehat{\varphi}(y(\cdot))(\cdot)\big\|_{L_2(\mathbb{R})}=\big\|x_1^{(k)}(\cdot)-\widehat{\varphi}(y_1(\cdot))(\cdot)\big\|_{L_2(\mathbb{R})}. \tag{2.2}$$

In view of (2.1), the expression on the right-hand side of (2.2) is not greater than $e(D^k,W,I_A^\delta,\widehat{\varphi})$, which is equal to $E(D^k,W,I_A^\delta)$ because the method $\widehat{\varphi}$ is optimal. Therefore, taking the supremum over all $x(\cdot)$ and $y(\cdot)$ on the left-hand side of (2.2), we obtain

$$e(D^k,W+L,I_A^\delta,\widehat{\varphi})\le E(D^k,W,I_A^\delta).$$

Hence we have (because $W\subset W+L$)

$$E(D^k,W,I_A^\delta)\le E(D^k,W+L,I_A^\delta)\le e(D^k,W+L,I_A^\delta,\widehat{\varphi})\le E(D^k,W,I_A^\delta).$$

Consequently, $\widehat{\varphi}$ is an optimal method in the $(D^k,W+L,I_\sigma^\delta)$-problem, and we have the equality $E(D^k,W,I_A^\delta)=E(D^k,W+L,I_A^\delta)$.

Now, let $\widehat{\varphi}$ be a linear method with finite error in the $(D^k,W+L,I_A^\delta)$-problem, where L is a subspace of $\mathcal{W}_2^n(\mathbb{R})$. Suppose that there exists an element $x_0(\cdot)\in L$ such that

$$\big\|x_0^{(k)}(\cdot)-\widehat{\varphi}(I_Ax_0(\cdot))(\cdot)\big\|_{L_2(\mathbb{R})}=c>0.$$

Then $\lambda x_0(\cdot)\in L$ for any $\lambda>0$. Hence,

$$e(D^k,W+L,I_A^\delta,\widehat{\varphi})\ge\lambda c,$$

which contradicts the fact that the error of the method $\widehat{\varphi}$ is finite. $\square$

It follows from this proposition that if one seeks methods with simple structure (for example, linear) that are exact on some subspaces and, in addition, possess some optimality properties, then it is quite natural to set up the problem of finding optimal methods on classes of the form $W+L$.

We implement this in the case when W is a Sobolev class of functions, i.e.,

$$W_2^n(\mathbb{R})=\big\{x(\cdot)\in\mathcal{W}_2^n(\mathbb{R})\colon\ \|x^{(n)}(\cdot)\|_{L_2(\mathbb{R})}\le1\big\},$$

and $L=\mathcal{B}_{\sigma,2}(\mathbb{R})$ is the space of entire functions of exponential type σ.

Recall that if $\sigma>0$, then $\mathcal{B}_{\sigma,2}(\mathbb{R})$ is the subspace in $L_2(\mathbb{R})$ formed by the restrictions of entire functions of exponential type σ to $\mathbb{R}$. As is well known, $x(\cdot)\in\mathcal{B}_{\sigma,2}(\mathbb{R})$ if and only if the support of $Fx(\cdot)$ belongs to the interval $\Delta_\sigma=[-\sigma,\sigma]$. By definition, $\mathcal{B}_{0,2}(\mathbb{R})=\{0\}$.

If $x(\cdot)\in\mathcal{B}_{\sigma,2}(\mathbb{R})$, then $x^{(m)}(\cdot)\in\mathcal{B}_{\sigma,2}(\mathbb{R})$ for all $m\in\mathbb{N}$ (by Bernstein's inequality for entire functions of exponential type); therefore, in particular, $\mathcal{B}_{\sigma,2}(\mathbb{R})\subset\mathcal{W}_2^n(\mathbb{R})$.

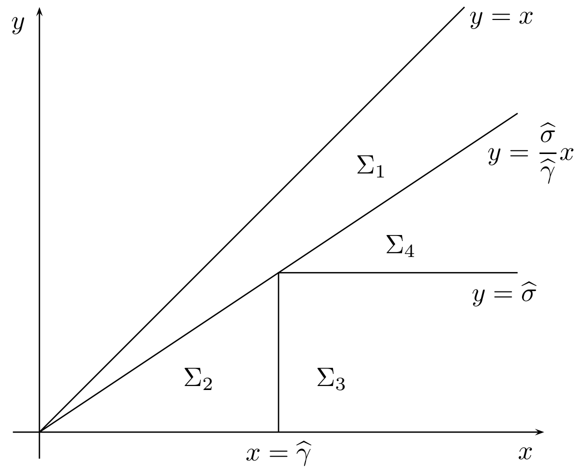

Fig. 1.

Prior to formulating a theorem, we introduce some notations. For a measurable set A on the real line, let

$$\gamma_A = \sup\bigl\{a \ge 0\colon\ \mathrm{mes}(A \cap [-a,a]) = 2a\bigr\}.$$

Let $1 \le k \le n-1$ and $\delta > 0$. Introduce the notations

$$\widehat{\gamma} = \Bigl(\frac{n}{k}\Bigr)^{1/(2(n-k))}\delta^{-1/n}, \qquad \widehat{\sigma} = \Bigl(\frac{n-k}{n}\Bigr)^{1/(2k)}\delta^{-1/n}$$

and consider the following four domains in the plane $\mathbb{R}^2$:

$$\Sigma_1 = \Bigl\{(x,y) \in \mathbb{R}^2\colon\ 0 < \frac{\widehat{\sigma}}{\widehat{\gamma}}x \le y \le x\Bigr\}, \qquad \Sigma_2 = \Bigl\{(x,y) \in \mathbb{R}^2\colon\ 0 \le y \le \frac{\widehat{\sigma}}{\widehat{\gamma}}x,\ 0 < x \le \widehat{\gamma}\Bigr\},$$

$$\Sigma_3 = \bigl\{(x,y) \in \mathbb{R}^2\colon\ x \ge \widehat{\gamma},\ 0 \le y \le \widehat{\sigma}\bigr\}, \qquad \Sigma_4 = \Bigl\{(x,y) \in \mathbb{R}^2\colon\ \widehat{\sigma} \le y \le \frac{\widehat{\sigma}}{\widehat{\gamma}}x\Bigr\}.$$

These domains are shown in Fig. 1.

Next, for every set A and number $\sigma \ge 0$, we define a pair of numbers $\lambda_1 = \lambda_1(A,\sigma)$ and $\lambda_2 = \lambda_2(A,\sigma)$ by the rule

$$(\lambda_1,\lambda_2) = \begin{cases} \bigl(\sigma^{2k}, \gamma_A^{-2(n-k)}\bigr), & (\gamma_A,\sigma) \in \Sigma_1, \\ \Bigl(\Bigl(\dfrac{\widehat{\sigma}}{\widehat{\gamma}}\gamma_A\Bigr)^{2k}, \gamma_A^{-2(n-k)}\Bigr), & (\gamma_A,\sigma) \in \Sigma_2, \\ \bigl(\widehat{\sigma}^{\,2k}, \widehat{\gamma}^{\,-2(n-k)}\bigr), & (\gamma_A,\sigma) \in \Sigma_3, \\ \Bigl(\sigma^{2k}, \Bigl(\dfrac{\widehat{\gamma}}{\widehat{\sigma}}\sigma\Bigr)^{-2(n-k)}\Bigr), & (\gamma_A,\sigma) \in \Sigma_4, \end{cases} \tag{2.3}$$

as well as a set $\Xi(A,\sigma)$ of measurable functions $\theta(\cdot)$ on $A \setminus \Delta_\sigma$ such that $|\theta(\xi)| \le 1$ for a.e. $\xi \in A \setminus \Delta_\sigma$.

Theorem 1. *Let $0 \le k \le n-1$, A be a measurable subset of $\mathbb{R}$, $\delta \ge 0$, and $\sigma \ge 0$. In this case,*

(1) *if $\sigma > \gamma_A$ or $\sigma = \gamma_A = 0$, then*

$$E\bigl(D^k, W_2^n(\mathbb{R}) + \mathcal{B}_{\sigma,2}(\mathbb{R}), I_A^\delta\bigr) = +\infty; \tag{2.4}$$

(2) *if* $k \geq 1$, $\delta > 0$, $\gamma_A > 0$, *and* $\sigma \leq \gamma_A$, *then*

$$E\big(D^k, W_2^n(\mathbb{R}) + \mathcal{B}_{\sigma,2}(\mathbb{R}), I_A^\delta\big) = \sqrt{\lambda_1\delta^2 + \lambda_2}$$

and, for every function $\theta(\cdot) \in \Xi(A,\sigma)$, *the method*

$$\widehat{\varphi}_\theta(y(\cdot))(t) = \frac{1}{2\pi}\int_{-\sigma}^{\sigma}(i\xi)^k y(\xi)e^{i\xi t}\,dt + \frac{1}{2\pi}\int_{A\setminus\Delta_\sigma}(i\xi)^k a_\theta(\xi)y(\xi)e^{i\xi t}\,dt, \tag{2.5}$$

where

$$a_\theta(\xi) = \frac{\lambda_1 + \theta(\xi)|\xi|^{n-k}\sqrt{\lambda_1\lambda_2}\sqrt{\lambda_1 + \lambda_2\xi^{2n} - \xi^{2k}}}{\lambda_1 + \lambda_2\xi^{2n}}, \tag{2.6}$$

is optimal;

(3) *if* $k \geq 1$, $\delta = 0$, $\gamma_A > 0$, *and* $\sigma \leq \gamma_A$, *then*

$$E\big(D^k, W_2^n(\mathbb{R}) + \mathcal{B}_{\sigma,2}(\mathbb{R}), I_A^0\big) = \gamma_A^{-(n-k)}$$

and, for every function $\theta(\cdot) \in \Xi(A,\sigma)$, *the method*

$$\widehat{\varphi}_\theta(y(\cdot))(t) = \frac{1}{2\pi}\int_{-\sigma}^{\sigma}(i\xi)^k Fx(\xi)e^{i\xi t}\,dt + \frac{1}{2\pi}\int_{A\setminus\Delta_\sigma}(i\xi)^k\left(1 + \theta(\xi)\left|\frac{\xi}{\gamma_A}\right|^{n-k}\right)Fx(\xi)e^{i\xi t}\,dt$$

is optimal;

(4) *if* $k = 0$, $\gamma_A > 0$, *and* $\sigma \leq \gamma_A$, *then*

$$E\big(D^0, W_2^n(\mathbb{R}) + \mathcal{B}_{\sigma,2}(\mathbb{R}), I_A^\delta\big) = \sqrt{\delta^2 + \gamma_A^{-2n}}$$

and, for every function $\theta(\cdot) \in \Xi(A,\sigma)$, *the method*

$$\widehat{\varphi}_\theta(y(\cdot))(t) = \frac{1}{2\pi}\int_{-\sigma}^{\sigma} y(\xi)e^{i\xi t}\,dt + \frac{1}{2\pi}\int_{A\setminus\Delta_\sigma}\frac{\gamma_A^{2n} + \theta(\xi)\xi^{2n}}{\gamma_A^{2n} + \xi^{2n}}y(\xi)e^{i\xi t}\,dt$$

is optimal.

Before proving this theorem, we make a number of remarks.

The set A on which information on the approximate Fourier transform is defined may be "large enough," and among the optimal methods (2.5) there may be those that do not employ all the available information. Naturally, the question arises as to whether there are optimal methods that use less information. More precisely, how much can one reduce the set A without increasing the error of optimal recovery? In terms of the function $a_\theta(\cdot)$ (which we consider as a smoothing factor), this means that we are interested in the sets where one can set $a_\theta(\cdot) = 0$.

We also wonder whether it is possible to take the smoothing factor equal to one on a wider set $[-\sigma_0, \sigma_0]$, where $\sigma_0 \geq \sigma$. In this case, the corresponding optimal method will be exact on the wider space $\mathcal{B}_{\sigma_0,2}(\mathbb{R})$ and, hence, in view of Proposition 1, will be optimal on the wider class $W_2^n(\mathbb{R}) + \mathcal{B}_{\sigma_0,2}(\mathbb{R})$.

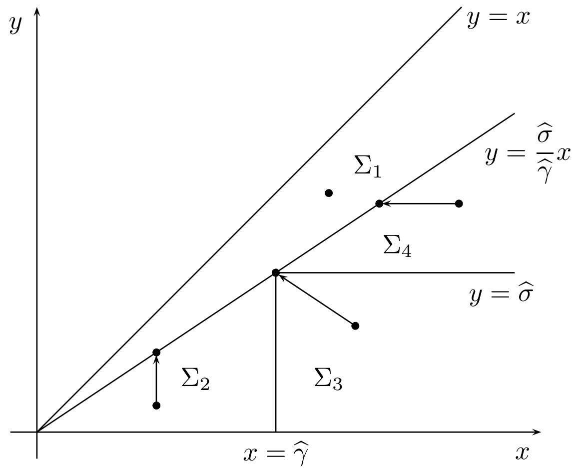

Fig. 2.

In the following corollary to Theorem 1, we set

$$(\sigma_0,\gamma_0)=\begin{cases}(\sigma,\gamma_A), & (\gamma_A,\sigma)\in\Sigma_1,\\ \left(\dfrac{\widehat{\sigma}}{\widehat{\gamma}}\gamma_A,\gamma_A\right), & (\gamma_A,\sigma)\in\Sigma_2,\\ (\widehat{\sigma},\widehat{\gamma}), & (\gamma_A,\sigma)\in\Sigma_3,\\ \left(\sigma,\dfrac{\widehat{\gamma}}{\widehat{\sigma}}\sigma\right), & (\gamma_A,\sigma)\in\Sigma_4.\end{cases}$$

Corollary 1. *Let $0\le k\le n-1$, A be a measurable subset of $\mathbb{R}$, $\delta\ge 0$, $\gamma_A>0$, and $0\le\sigma\le\gamma_A$. In this case,*

(1) *if $k\ge 1$ and $\delta>0$, then, for all $\theta(\cdot)\in\Xi(A,\sigma_0)$, the methods*

$$\widehat{\varphi}_\theta(y(\cdot))(t)=\frac{1}{2\pi}\int\limits_{|\xi|\le\sigma_0}(i\xi)^k y(\xi)e^{i\xi t}\,dt+\frac{1}{2\pi}\int\limits_{\sigma_0\le|\xi|\le\gamma_0}(i\xi)^k a_\theta(\xi)y(\xi)e^{i\xi t}\,dt$$

with the functions $a_\theta(\cdot)$ defined in (2.6) are optimal in the $(D^k,W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R}),I_A^\delta)$-problem and exact on the subspace $\mathcal{B}_{\sigma_0,2}(\mathbb{R})$;

(2) *if $k=0$ or $\delta=0$, then the method*

$$\widehat{\varphi}(y(\cdot))(t)=\frac{1}{2\pi}\int\limits_{|\xi|\le\gamma_A}(i\xi)^k y(\xi)e^{i\xi t}\,dt$$

is optimal in the $(D^k,W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R}),I_A^\delta)$-problem and exact on the subspace $\mathcal{B}_{\gamma_A,2}(\mathbb{R})$.

In case (1), the transition from the point (σ,γ_A) to the point (σ_0,γ_0) for each of the domains Σ_j, $j=1,2,3,4$, is schematically illustrated in Fig. 2.

Let us indicate the form of the optimal methods in the original $(D^k,W_2^n(\mathbb{R}),I_A^\delta)$-problem that are exact on the subspaces $\mathcal{B}_{\sigma,2}(\mathbb{R})$.

Corollary 2. *Let $0\le k\le n-1$, A be a measurable subset of $\mathbb{R}$, $\delta\ge 0$, and $\gamma_A>0$. In this case,*

(1) *if $k\ge 1$ and $\delta>0$, then, for all $\theta(\cdot)\in\Xi(A,\sigma_0)$, the methods*

$$\widehat{\varphi}_\theta(y(\cdot))(t)=\frac{1}{2\pi}\int\limits_{|\xi|\le\widetilde{\sigma}}(i\xi)^k y(\xi)e^{i\xi t}\,dt+\frac{1}{2\pi}\int\limits_{\widetilde{\sigma}\le|\xi|\le\widetilde{\gamma}}(i\xi)^k a_\theta(\xi)y(\xi)e^{i\xi t}\,dt,$$

where $\widetilde{\gamma}=\min\{\gamma_A,\widehat{\gamma}\}$, $\widetilde{\sigma}=(\widehat{\sigma}/\widehat{\gamma})\widetilde{\gamma}$, *and the functions* $a_\theta(\cdot)$ *are defined in* (2.6), *are optimal in the* $(D^k, W_2^n(\mathbb{R}), I_A^\delta)$*-problem and exact on the subspace* $\mathcal{B}_{\widetilde{\sigma},2}(\mathbb{R})$;

(2) *if* $k=0$ *or* $\delta=0$, *then the method*

$$\widehat{\varphi}(y(\cdot))(t)=\frac{1}{2\pi}\int\limits_{|\xi|\le\gamma_A}(i\xi)^k y(\xi)e^{i\xi t}\,dt$$

is optimal in the $(D^k, W_2^n(\mathbb{R}), I_A^\delta)$*-problem and exact on the subspace* $\mathcal{B}_{\gamma_A,2}(\mathbb{R})$.

3. PROOFS

Proof of Theorem 1. We begin with a lower estimate for $E(D^k, W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R}), I_A^\delta)$. Consider the problem

$$\big\|x^{(k)}(\cdot)\big\|_{L_2(\mathbb{R})}\to\max,\quad \|Fx(\cdot)\|_{\widehat{L}_2(A)}\le\delta,\quad \big\|Fx^{(n)}(\cdot)\big\|_{\widehat{L}_2(\mathbb{R}\setminus\Delta_\sigma)}\le 1,\quad x(\cdot)\in\mathcal{W}_2^n(\mathbb{R}),\tag{3.1}$$

where $\Delta_\sigma=[-\sigma,\sigma]$. Let us show that the value of this problem, i.e., the supremum of the functional to be maximized under the indicated constraints, is not greater than $E(D^k, W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R}), I_A^\delta)$.

As a preliminary step, we show that $x(\cdot)\in\mathcal{W}_2^n(\mathbb{R})$ belongs to $W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R})$ if and only if $\|Fx^{(n)}(\cdot)\|_{\widehat{L}_2(\mathbb{R}\setminus\Delta_\sigma)}\le 1$. Indeed, if $x(\cdot)\in W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R})$, then $x(\cdot)=x_1(\cdot)+x_2(\cdot)$, where $x_1(\cdot)\in W_2^n(\mathbb{R})$ and $x_2(\cdot)\in\mathcal{B}_{\sigma,2}(\mathbb{R})$. By the Plancherel theorem (since $Fx_2(\cdot)$ is concentrated on the interval Δ_σ), we have

$$\big\|Fx^{(n)}(\cdot)\big\|^2_{\widehat{L}_2(\mathbb{R}\setminus\Delta_\sigma)}=\big\|Fx_1^{(n)}(\cdot)\big\|^2_{\widehat{L}_2(\mathbb{R}\setminus\Delta_\sigma)}=\frac{1}{2\pi}\int\limits_{\mathbb{R}\setminus\Delta_\sigma}\xi^{2n}|Fx_1(\xi)|^2\,d\xi\le\frac{1}{2\pi}\int\limits_{\mathbb{R}}\xi^{2n}|Fx_1(\xi)|^2\,d\xi$$
$$=\big\|x_1^{(n)}(\cdot)\big\|^2_{L_2(\mathbb{R})}\le 1.$$

Conversely, let $x(\cdot)\in\mathcal{W}_2^n(\mathbb{R})$ and $\|Fx^{(n)}(\cdot)\|_{\widehat{L}_2(\mathbb{R}\setminus\Delta_\sigma)}\le 1$. Denote by $x_2(\cdot)$ the function in $L_2(\mathbb{R})$ with the Fourier transform $Fx_2(\cdot)=\chi_\sigma(\cdot)Fx(\cdot)$, where $\chi_\sigma(\cdot)$ is the characteristic function of the interval Δ_σ. Then clearly $x_2(\cdot)\in\mathcal{B}_{\sigma,2}(\mathbb{R})$. Set $x_1(\cdot)=x(\cdot)-x_2(\cdot)$. It is obvious that $x_1(\cdot)\in\mathcal{W}_2^n(\mathbb{R})$, and by the Plancherel theorem (since $Fx_1(\cdot)=0$ on Δ_σ) we have

$$\big\|x_1^{(n)}(\cdot)\big\|^2_{L_2(\mathbb{R})}=\frac{1}{2\pi}\int\limits_{\mathbb{R}\setminus\Delta_\sigma}\xi^{2n}|Fx_1(\xi)|^2\,d\xi=\frac{1}{2\pi}\int\limits_{\mathbb{R}\setminus\Delta_\sigma}\xi^{2n}|Fx(\xi)|^2\,d\xi=\big\|Fx^{(n)}(\cdot)\big\|^2_{\widehat{L}_2(\mathbb{R}\setminus\Delta_\sigma)}\le 1;$$

i.e., $x(\cdot)=x_1(\cdot)+x_2(\cdot)\in W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R})$.

Taking into account the remark made, we now prove that $E(D^k, W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R}), I_A^\delta)$ is not less than the value of problem (3.1). Let $x_0(\cdot)$ be an admissible function in (3.1) (i.e., $x_0(\cdot)$ satisfies the constraints of the problem); then it is obvious that the function $-x_0(\cdot)$ is also admissible and for any $\varphi\colon L_2(A)\to L_2(\mathbb{R})$ ($\varphi(0)(\cdot)$ is the value of the mapping φ on the zero function) we have

$$2\|x_0^{(k)}(\cdot)\|_{L_2(\mathbb{R})}\le\big\|x_0^{(k)}(\cdot)-\varphi(0)(\cdot)\big\|_{L_2(\mathbb{R})}+\big\|-x_0^{(k)}(\cdot)-\varphi(0)(\cdot)\big\|_{L_2(\mathbb{R})}$$
$$\le 2\sup_{\substack{x(\cdot)\in\mathcal{W}_2^n(\mathbb{R})\\ \|Fx(\cdot)\|_{\widehat{L}_2(A)}\le\delta,\ \|Fx^{(n)}(\cdot)\|_{\widehat{L}_2(\mathbb{R}\setminus\Delta_\sigma)}\le 1}}\big\|x^{(k)}(\cdot)-\varphi(0)(\cdot)\big\|_{L_2(\mathbb{R})}$$
$$=2\sup_{\substack{x(\cdot)\in W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R})\\ \|Fx(\cdot)\|_{\widehat{L}_2(A)}\le\delta}}\big\|x^{(k)}(\cdot)-\varphi(0)(\cdot)\big\|_{L_2(\mathbb{R})}$$
$$\le 2\sup_{x(\cdot)\in W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R}),\ y(\cdot)\in I_A^\delta x(\cdot)}\big\|x^{(k)}(\cdot)-\varphi(y(\cdot))(\cdot)\big\|_{L_2(\mathbb{R})}.$$

Taking the supremum over all admissible functions in (3.1) on the left and to the infimum over all methods φ on the right, we obtain what was required.

Now, we proceed directly to the proof of the assertions of the theorem.

1. In case (1), let, first, $\sigma > \gamma_A$. By the definition of γ_A, in the set $[-\sigma, -\gamma_A] \cup [\gamma_A, \sigma]$ there exists a subset D of positive measure such that $D \cap A = \varnothing$. Let $c > 0$ and a function $x_c(\cdot)$ be such that $Fx_c(\cdot) = c$ on D and $Fx_c(\cdot) = 0$ outside D. It is clear that $x_c(\cdot)$ is admissible in problem (3.1) and (by the Plancherel theorem)

$$\|x_c^{(k)}(\cdot)\|^2_{L_2(\mathbb{R})} = \frac{c^2}{2\pi} \int\limits_D \xi^{2k}\, d\xi.$$

The number c can be arbitrarily large; therefore, equality (2.4) is proved.

Suppose that $\sigma = \gamma_A = 0$. In this case, $\operatorname{mes}(A \cap [-\varepsilon, \varepsilon]) < 2\varepsilon$ for any $\varepsilon > 0$. Hence, the measure of the set $\Omega_\varepsilon = \{(\mathbb{R} \setminus A) \cap [-\varepsilon, \varepsilon]\}$ is positive. Consider a function $x_\varepsilon(\cdot)$ such that

$$Fx_\varepsilon(\xi) = \begin{cases} \left(\int\limits_{\Omega_\varepsilon} \xi^{2n}\, d\xi \right)^{-1/2}, & \xi \in \Omega_\varepsilon, \\ 0, & \xi \notin \Omega_\varepsilon. \end{cases}$$

This function is admissible in problem (3.1), and

$$\|x_\varepsilon^{(k)}(\cdot)\|^2_{L_2(\mathbb{R})} = \frac{\int_{\Omega_\varepsilon} \xi^{2k}\, d\xi}{\int_{\Omega_\varepsilon} \xi^{2n}\, d\xi} = \frac{\int_{\Omega_\varepsilon} \xi^{2n} \xi^{-2(n-k)}\, d\xi}{\int_{\Omega_\varepsilon} \xi^{2n}\, d\xi} \ge \varepsilon^{-2(n-k)},$$

which implies (since ε is arbitrary) that the value of the functional to be maximized in (3.1) can be made arbitrarily large.

2. In case (2), we first show that the following estimate is valid in each of the domains Σ_j, $j = 1, 2, 3, 4$:

$$E\big(D^k, W_2^n(\mathbb{R}) + \mathcal{B}_{\sigma,2}(\mathbb{R}), I_A^\delta\big) \ge \sqrt{\lambda_1 \delta^2 + \lambda_2}. \tag{3.2}$$

Let $(\gamma_A, \sigma) \in \Sigma_1$. By the definition of γ_A, for every positive integer m, there exists a subset D_m of positive measure in the set $[-\gamma(A) - 1/m, -\gamma(A)] \cup [\gamma(A), \gamma(A) + 1/m]$ such that $A \cap D_m = \varnothing$. Let m be such that $1/m < \sigma$. For every such m, consider a function $x_m(\cdot)$ such that

$$Fx_m(\xi) = \begin{cases} \delta\sqrt{2\pi m}, & \sigma - \dfrac{1}{m} \le \xi < \sigma, \\ \sqrt{2\pi}\left(\gamma_A + \dfrac{1}{m}\right)^{-n} (\operatorname{mes} D_m)^{-1/2}, & \xi \in D_m, \\ 0 & \text{otherwise}. \end{cases}$$

The functions $x_m(\cdot)$ are admissible in problem (3.1). Indeed, applying the Plancherel theorem and the definition of $x_m(\cdot)$, we have

$$\|Fx_m(\cdot)\|^2_{\widehat{L}_2(A)} = \frac{1}{2\pi} \int\limits_A |Fx_m(\xi)|^2\, d\xi = \frac{1}{2\pi} \delta^2 2\pi m \frac{1}{m} = \delta^2 \tag{3.3}$$

and

$$\|Fx_m^{(n)}(\cdot)\|^2_{\widehat{L}_2(\mathbb{R}\setminus\Delta_\sigma)}=\frac{1}{2\pi}\int\limits_{|\xi|\ge\sigma}\xi^{2n}|Fx_m(\xi)|^2\,d\xi=\frac{1}{2\pi}2\pi\biggl(\gamma_A+\frac{1}{m}\biggr)^{-2n}(\operatorname{mes}D_m)^{-1}\int\limits_{D_m}\xi^{2n}\,d\xi$$
$$\le\biggl(\gamma_A+\frac{1}{m}\biggr)^{-2n}(\operatorname{mes}D_m)^{-1}\biggl(\gamma_A+\frac{1}{m}\biggr)^{2n}\operatorname{mes}D_m=1.$$

Next,

$$\|Fx_m^{(k)}(\cdot)\|^2=\frac{1}{2\pi}\int\limits_{\mathbb{R}}\xi^{2k}|Fx_m(\xi)|^2\,d\xi=\delta^2m\int\limits_{\sigma-1/m}^{\sigma}\xi^{2k}\,d\xi+\biggl(\gamma_A+\frac{1}{m}\biggr)^{-2n}(\operatorname{mes}D_m)^{-1}\int\limits_{D_m}\xi^{2k}\,d\xi$$
$$\ge\delta^2m\biggl(\sigma-\frac{1}{m}\biggr)^{2k}\frac{1}{m}+\biggl(\gamma_A+\frac{1}{m}\biggr)^{-2n}(\operatorname{mes}D_m)^{-1}\gamma_A^{2k}\operatorname{mes}D_m$$
$$=\delta^2\biggl(\sigma-\frac{1}{m}\biggr)^{2k}+\biggl(\gamma_A+\frac{1}{m}\biggr)^{-2n}\gamma_A^{2k}.$$

As $m\to\infty$, the expression on the right-hand side tends to $\sigma^{2k}\delta^2+\gamma_A^{-2(n-k)}=\lambda_1\delta^2+\lambda_2$, which is obviously not greater than the value of problem (3.1). However, by what has been proved, this value is not greater than $E(D^k,W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R}),I_A^\delta)$; hence inequality (3.2) is proved in this case.

Let $(\gamma_A,\sigma)\in\Sigma_2$. Set

$$\xi_0=\biggl(\frac{k}{n}\biggr)^{1/(2(n-k))}\gamma_A.$$

Notice that

$$\sigma\le\frac{\widehat{\sigma}}{\widehat{\gamma}}\gamma_A=\biggl(\frac{n-k}{n}\biggr)^{1/(2k)}\xi_0<\xi_0,\qquad \xi_0^{2n}\le\biggl(\frac{k}{n}\biggr)^{n/(n-k)}\widehat{\gamma}^{\,2n}=\delta^{-2}.$$

Let m be such that $\sigma<\xi_0-1/m$. For every such m, consider a function $x_m(\cdot)$ such that

$$Fx_m(\xi)=\begin{cases}\delta\sqrt{2\pi m}, & \xi_0-\dfrac{1}{m}\le\xi<\xi_0,\\[2mm] \dfrac{\sqrt{2\pi(1-\delta^2\xi_0^{2n})}}{(\gamma_A+1/m)^n\sqrt{\operatorname{mes}D_m}}, & \xi\in D_m,\\[2mm] 0 & \text{otherwise}.\end{cases}$$

Equalities (3.3) remain valid. In addition,

$$\|Fx_m^{(n)}(\cdot)\|^2_{\widehat{L}_2(\mathbb{R}\setminus\Delta_\sigma)}=\frac{1}{2\pi}\int\limits_{|\xi|\ge\sigma}\xi^{2n}|Fx_m(\xi)|^2\,d\xi$$
$$=\delta^2m\int\limits_{\xi_0-1/m}^{\xi_0}\xi^{2n}\,d\xi+\bigl(1-\delta^2\xi_0^{2n}\bigr)\biggl(\gamma_A+\frac{1}{m}\biggr)^{-2n}(\operatorname{mes}D_m)^{-1}\int\limits_{D_m}\xi^{2n}\,d\xi$$
$$\le\delta^2\xi_0^{2n}+\bigl(1-\delta^2\xi_0^{2n}\bigr)\biggl(\gamma_A+\frac{1}{m}\biggr)^{-2n}(\operatorname{mes}D_m)^{-1}\biggl(\gamma_A+\frac{1}{m}\biggr)^{2n}\operatorname{mes}D_m=1.$$

Thus, the functions $x_m(\cdot)$ are admissible in problem (3.1).

Next, we have

$$\|Fx_m^{(k)}(\cdot)\|^2 = \frac{1}{2\pi}\int\limits_{\mathbb{R}} \xi^{2k}|Fx_m(\xi)|^2\,d\xi$$

$$= \delta^2 m \int\limits_{\xi_0-1/m}^{\xi_0} \xi^{2k}\,d\xi + (1-\delta^2\xi_0^{2n})\left(\gamma_A+\frac{1}{m}\right)^{-2n}(\operatorname{mes} D_m)^{-1}\int\limits_{D_m}\xi^{2k}\,d\xi$$

$$\ge \delta^2\left(\xi_0-\frac{1}{m}\right)^{2k} + (1-\delta^2\xi_0^{2n})\left(\gamma_A+\frac{1}{m}\right)^{-2n}\gamma_A^{2k}.$$

As $m\to\infty$, the expression on the right-hand side tends to

$$\delta^2\xi_0^{2k} + (1-\delta^2\xi_0^{2n})\gamma_A^{-2(n-k)} = \lambda_1\delta^2+\lambda_2.$$

Hence, for the same reasons as above, inequality (3.2) is also valid in this case.

Let $(\gamma_A,\sigma)\in\Sigma_3$. Set

$$\xi_1 = \delta^{-1/n}.$$

In this case,

$$\gamma_A \ge \widehat{\gamma} > \xi_1, \qquad \sigma \le \widehat{\sigma} < \xi_1.$$

Let m be such that $\sigma < \xi_1 - 1/m$. For every such m, consider a function $x_m(\cdot)$ such that

$$Fx_m(\xi) = \begin{cases} \delta\sqrt{2\pi m}, & \xi_1-\dfrac{1}{m}\le\xi<\xi_1,\\ 0 & \text{otherwise.}\end{cases}$$

One can easily verify that the functions $x_m(\cdot)$ are admissible in problem (3.1). Next,

$$\|Fx_m^{(k)}(\cdot)\|^2 = \frac{1}{2\pi}\int\limits_{\mathbb{R}}\xi^{2k}|Fx_m(\xi)|^2\,d\xi = \delta^2 m\int\limits_{\xi_1-1/m}^{\xi_1}\xi^{2k}\,d\xi \ge \delta^2\left(\xi_1-\frac{1}{m}\right)^{2k}.$$

The expression on the right-hand side tends to $\delta^2\xi_1^{2k} = \lambda_1\delta^2+\lambda_2$ as $m\to\infty$. Hence, inequality (3.2) is satisfied in this case as well.

Finally, let $(\gamma_A,\sigma)\in\Sigma_4$. Set

$$\xi_2 = \left(\frac{n-k}{n}\right)^{-1/(2k)}\sigma.$$

It is obvious that $\xi_2>\sigma$. On the other hand,

$$\xi_2 \le \left(\frac{n-k}{n}\right)^{-1/(2k)}\frac{\widehat{\sigma}}{\widehat{\gamma}}\gamma_A = \left(\frac{k}{n}\right)^{1/(2(n-k))}\gamma_A < \gamma_A.$$

Note also that

$$\xi_2^{-2n} \le \left(\frac{n-k}{n}\right)^{n/k}\widehat{\sigma}^{\,-2n} = \delta^2.$$

Let m be such that $1/m < \sigma < \xi_2 - 1/m$. For every such m, consider a function $x_m(\cdot)$ such that

$$Fx_m(\xi) = \begin{cases} \sqrt{2\pi m\big(\delta^2 - \xi_2^{-2n}\big)}, & \sigma - \dfrac{1}{m} \le \xi < \sigma, \\ \xi_2^{-n}\sqrt{2\pi m}, & \xi_2 - \dfrac{1}{m} \le \xi \le \xi_2, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\|Fx_m(\cdot)\|^2_{\widehat{L}_2(A)} = \frac{1}{2\pi}\int\limits_A |Fx_m(\xi)|^2\,d\xi = \delta^2 - \xi_2^{-2n} + \xi_2^{-2n} = \delta^2$$

and

$$\big\|Fx_m^{(n)}(\cdot)\big\|^2_{\widehat{L}_2(\mathbb{R}\setminus\Delta_\sigma)} = \frac{1}{2\pi}\int\limits_{|\xi|\ge\sigma} \xi^{2n}|Fx_m(\xi)|^2\,d\xi = \xi_2^{-2n} m \int\limits_{\xi_2-1/m}^{\xi_2} \xi^{2n}\,d\xi \le 1.$$

Thus, the functions $x_m(\cdot)$ are admissible in problem (3.1). In addition,

$$\big\|Fx_m^{(k)}(\cdot)\big\|^2 = \frac{1}{2\pi}\int\limits_{\mathbb{R}} \xi^{2k}|Fx_m(\xi)|^2\,d\xi = m\big(\delta^2 - \xi_2^{-2n}\big)\int\limits_{\sigma-1/m}^{\sigma} \xi^{2k}\,d\xi + \xi_2^{-2n} m \int\limits_{\xi_2-1/m}^{\xi_2} \xi^{2k}\,d\xi$$

$$\ge \big(\delta^2 - \xi_2^{-2n}\big)\left(\sigma - \frac{1}{m}\right)^{2k} + \xi_2^{-2n}\left(\xi_2 - \frac{1}{m}\right)^{2k}.$$

The expression on the right-hand side tends to

$$\big(\delta^2 - \xi_2^{-2n}\big)\sigma^{2k} + \xi_2^{-2(n-k)} = \lambda_1\delta^2 + \lambda_2$$

as $m \to \infty$. Hence, inequality (3.2) is satisfied in this case as well.

Let us proceed to estimating $E(D^k, W_2^n(\mathbb{R}) + \mathcal{B}_{\sigma,2}(\mathbb{R}), I_A^\delta)$ from above and to constructing optimal recovery methods. We will seek such methods among the mappings $\widehat{\varphi}_a\colon \widehat{L}_2(A) \to L_2(\mathbb{R})$ that are represented in terms of Fourier transforms as

$$F\widehat{\varphi}_a(y(\cdot))(\xi) = (i\xi)^k a(\xi) y(\xi), \qquad \xi \in \mathbb{R},$$

where the function $a(\cdot) \in L_\infty(\mathbb{R})$ is such that $F\widehat{\varphi}_a(y(\cdot))(\cdot) \in L_2(\mathbb{R})$.

Let us estimate the error of such a method, which is by definition (see also the remark at the beginning of the proof) equal to the value of the following problem:

$$\begin{gathered} \big\|x^{(k)}(\cdot) - \widehat{\varphi}_a(y(\cdot))(\cdot)\big\|_{L_2(\mathbb{R})} \to \max, \\ \|Fx(\cdot) - y(\cdot)\|_{\widehat{L}_2(A)} \le \delta, \quad y(\cdot) \in \widehat{L}_2(A), \quad \big\|Fx^{(n)}(\cdot)\big\|_{\widehat{L}_2(\mathbb{R}\setminus\Delta_\sigma)} \le 1, \quad x(\cdot) \in \mathcal{W}_2^n(\mathbb{R}). \end{gathered} \tag{3.4}$$

Passing to the Fourier images in the functional to be maximized, by the Plancherel theorem we find that the squared value of problem (3.4) is equal to the value of the following problem:

$$\begin{gathered} \frac{1}{2\pi}\int\limits_A \big|(i\xi)^k Fx(\xi) - (i\xi)^k a(\xi) y(\xi)\big|^2\,d\xi + \frac{1}{2\pi}\int\limits_{\mathbb{R}\setminus A} \xi^{2k}|Fx(\xi)|^2\,d\xi \to \max, \\ \frac{1}{2\pi}\int\limits_A |Fx(\xi) - y(\xi)|^2\,d\xi \le \delta^2, \qquad y(\cdot) \in \widehat{L}_2(A), \qquad \frac{1}{2\pi}\int\limits_{|\xi|\ge\sigma} \xi^{2n}|Fx(\xi)|^2\,d\xi \le 1. \end{gathered} \tag{3.5}$$

Notice that on the admissible pairs $(x(\cdot), y(\cdot))$ in this problem, where $x(\cdot) \in \mathcal{B}_{\sigma,2}(\mathbb{R})$ and $y(\cdot) = Fx(\cdot)$, the functional has the form

$$\frac{1}{2\pi} \int\limits_{\Delta_\sigma} \xi^{2k} |Fx(\xi)|^2 |1 - a(\xi)|\, d\xi.$$

Therefore, if the function $a(\cdot)$ is not equal to one almost everywhere on Δ_σ, then the value of problem (3.5) (and, hence, the value of problem (3.4)) is equal to infinity, because $\mathcal{B}_{\sigma,2}(\mathbb{R})$ is a linear space; i.e., the error of the method with such $a(\cdot)$ is infinite, and this case is of no interest to us.

Let $a(\cdot) \equiv 1$ on $A \cap \Delta_\sigma$. We estimate the functional maximized in (3.5) from above by representing it as a sum of three terms,

$$I_1 = \frac{1}{2\pi} \int\limits_{A\cap\Delta_\sigma} \left|(i\xi)^k Fx(\xi) - (i\xi)^k y(\xi)\right|^2 d\xi, \qquad I_2 = \frac{1}{2\pi} \int\limits_{A\setminus\Delta_\sigma} \left|(i\xi)^k Fx(\xi) - (i\xi)^k a(\xi) y(\xi)\right|^2 d\xi,$$

$$I_3 = \frac{1}{2\pi} \int\limits_{\mathbb{R}\setminus A} \xi^{2k} |Fx(\xi)|^2\, d\xi.$$

Let us show that

$$I_1 \le \frac{\lambda_1}{2\pi} \int\limits_{A\cap\Delta_\sigma} |Fx(\xi) - y(\xi)|^2\, d\xi \tag{3.6}$$

in all the domains Σ_i, $i = 1, 2, 3, 4$.

Indeed, the inequality

$$I_1 \le \frac{\sigma^{2k}}{2\pi} \int\limits_{A\cap\Delta_\sigma} |Fx(\xi) - y(\xi)|^2\, d\xi$$

is obvious. Since $\sigma^{2k} = \lambda_1$ in Σ_1 and Σ_4, inequality (3.6) holds for these domains. If $(\gamma_A, \sigma) \in \Sigma_2$, then

$$\lambda_1 = \left(\frac{\widehat{\sigma}}{\widehat{\widehat{\gamma}}}\gamma_A\right)^{2k} \ge \sigma^{2k},$$

and if $(\gamma_A, \sigma) \in \Sigma_3$, then

$$\lambda_1 = \widehat{\sigma}^{\,2k} \ge \sigma^{2k},$$

so estimate (3.6) is valid for all the domains.

Now, let us estimate I_2. Applying the Cauchy–Bunyakovsky–Schwarz inequality, we have

$$\begin{aligned}\left|(i\xi)^k Fx(\xi) - (i\xi)^k a(\xi) y(\xi)\right|^2 &= \xi^{2k}\left|(1 - a(\xi))Fx(\xi) + a(\xi)(Fx(\xi) - y(\xi))\right|^2 \\ &\le \xi^{2k}\left(\frac{|1 - a(\xi)|^2}{\lambda_2 \xi^{2n}} + \frac{|a(\xi)|^2}{\lambda_1}\right)\left(\lambda_2 \xi^{2n} |Fx(\xi)|^2 + \lambda_1 |Fx(\xi) - y(\xi)|^2\right). \end{aligned}\tag{3.7}$$

Set

$$S_a = \operatorname*{ess\,sup}_{\xi\in A\setminus\Delta_\sigma} \xi^{2k}\left(\frac{|1 - a(\xi)|^2}{\lambda_2 \xi^{2n}} + \frac{|a(\xi)|^2}{\lambda_1}\right). \tag{3.8}$$

Then, integrating (3.7), we obtain the following estimate for I_2:

$$I_2 \le S_a \left(\frac{1}{2\pi} \int\limits_{A\setminus\Delta_\sigma} \left(\lambda_2 \xi^{2n} |Fx(\xi)|^2 + \lambda_1 |Fx(\xi) - y(\xi)|^2\right) d\xi\right). \tag{3.9}$$

Now, let us show that I_3 can be estimated in all the domains Σ_i, $i=1,2,3,4$, as

$$I_3 \le \frac{\lambda_2}{2\pi}\int\limits_{\mathbb{R}\setminus A}\xi^{2n}|Fx(\xi)|^2\,d\xi. \tag{3.10}$$

Indeed, since $|\xi|>\gamma_A$ for a.e. $\xi\in\mathbb{R}\setminus A$ (by the definition of γ_A), it follows that

$$I_3=\frac{1}{2\pi}\int\limits_{\mathbb{R}\setminus A}\xi^{-2(n-k)}\xi^{2n}|Fx(\xi)|^2\,d\xi \le \frac{\gamma_A^{-2(n-k)}}{2\pi}\int\limits_{\mathbb{R}\setminus A}\xi^{2n}|Fx(\xi)|^2\,d\xi. \tag{3.11}$$

Since $\gamma_A^{-2(n-k)}=\lambda_2$ in Σ_1 and Σ_2, inequality (3.10) holds in these domains. If $(\gamma_A,\sigma)\in\Sigma_3$, then

$$\lambda_2=\widehat{\gamma}^{\,-2(n-k)}\ge\gamma_A^{-2(n-k)},$$

and if $(\gamma_A,\sigma)\in\Sigma_4$, then $\sigma\le\widehat{\sigma}\,\widehat{\gamma}^{\,-1}\gamma_A$; therefore,

$$\lambda_2=\left(\frac{\widehat{\gamma}}{\widehat{\sigma}}\sigma\right)^{-2(n-k)}\ge\gamma_A^{-2(n-k)}.$$

Thus, estimate (3.10) is valid in all the domains.

If we assume that the function $a(\cdot)$ is such that $S_a\le 1$, then, summing inequalities (3.6), (3.9), and (3.10), we obtain the following estimate for the functional in problem (3.5):

$$\lambda_1\frac{1}{2\pi}\int\limits_A|Fx(\xi)-y(\xi)|^2\,d\xi+\lambda_2\frac{1}{2\pi}\int\limits_{A\setminus\Delta_\sigma}\xi^{2n}|Fx(\xi)|^2\,d\xi+\lambda_2\frac{1}{2\pi}\int\limits_{\mathbb{R}\setminus A}\xi^{2n}|Fx(\xi)|^2\,d\xi$$
$$=\lambda_1\frac{1}{2\pi}\int\limits_A|Fx(\xi)-y(\xi)|^2\,d\xi+\lambda_2\frac{1}{2\pi}\int\limits_{|\xi|\ge\sigma}\xi^{2n}|Fx(\xi)|^2\,d\xi\le\lambda_1\delta^2+\lambda_2,$$

which means that

$$e\big(D^k,W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R}),I_A^\delta,\widehat{\varphi}_a\big)\le\sqrt{\lambda_1\delta^2+\lambda_2}.$$

Comparing this with (3.2), we see that $\widehat{\varphi}_a$ is an optimal method in the $(D^k,W_2^n(\mathbb{R})+\mathcal{B}_{\sigma,2}(\mathbb{R}),I_A^\delta)$-problem.

Now, we show that functions $a(\cdot)$ for which $S_a\le 1$ do exist. First (completing the square), notice that the condition $S_a\le 1$ is equivalent to the fact that the inequality

$$\left|a(\xi)-\frac{\lambda_1}{\lambda_1+\lambda_2\xi^{2n}}\right|^2\le\frac{\xi^{2(n-k)}\lambda_1\lambda_2(\lambda_1+\lambda_2\xi^{2n}-\xi^{2k})}{\lambda_1+\lambda_2\xi^{2n}}$$

holds for a.e. $\xi\in A\setminus\Delta_\sigma$. If the function $\xi\mapsto f(\xi)=\lambda_1+\lambda_2\xi^{2n}-\xi^{2k}$ is nonnegative on $A\setminus\Delta_\sigma$, then such $a(\cdot)$ obviously exist and are described by equality (2.6). Let us check that $f(\cdot)$ is nonnegative on $A\setminus\Delta_\sigma$.

One can easily verify that the minimum value of this function on the whole real axis is

$$C=\lambda_1-\frac{n-k}{n}\left(\frac{k}{n\lambda_2}\right)^{k/(n-k)}.$$

Let us show that $C\ge 0$ in each of the domains Σ_j, $j=1,2,3,4$.

Let $(\gamma_A, \sigma) \in \Sigma_1$. Then

$$\sigma^{2k} \geq \frac{\widehat{\sigma}^{\,2k}}{\widehat{\gamma}^{\,2k}} \gamma_A^{2k}.$$

By the definition of λ_1 and λ_2, this inequality can be rewritten in Σ_1 as

$$\lambda_1 \geq \frac{n-k}{n} \left(\frac{k}{n}\right)^{k/(n-k)} \lambda_2^{-k/(n-k)},$$

which implies that $C \geq 0$. We can easily verify by direct substitution that $C = 0$ for the domains Σ_j, $j = 2, 3, 4$.

3. In case (3), by analogy with the proof of case (2) for the domain Σ_1, we obtain the lower bound

$$E\big(D^k, W_2^n(\mathbb{R}) + \mathcal{B}_{\sigma,2}(\mathbb{R}), I_A^0\big) \geq \gamma_A^{-(n-k)}.$$

For the upper bound, applying the same arguments as in the proof of the upper bound in case (2), we arrive at the following problem:

$$\begin{gathered} \frac{1}{2\pi} \int\limits_A \big|(i\xi)^k Fx(\xi) - (i\xi)^k a(\xi) Fx(\xi)\big|^2\, d\xi + \frac{1}{2\pi} \int\limits_{\mathbb{R}\setminus A} \xi^{2k} |Fx(\xi)|^2\, d\xi \to \max, \\ \frac{1}{2\pi} \int\limits_{|\xi|\geq\sigma} \xi^{2n} |Fx(\xi)|^2\, d\xi \leq 1. \end{gathered} \tag{3.12}$$

Since $a(\cdot) \equiv 1$ on $A \cap \Delta_\sigma$ (otherwise, as has been shown, the error of the method is equal to infinity), the functional to be maximized in (3.12) is represented as a sum of two terms,

$$J_1 = \frac{1}{2\pi} \int\limits_{A\setminus\Delta_\sigma} \xi^{2k} |Fx(\xi)|^2 |1 - a(\xi)|^2\, d\xi, \qquad I_3 = \frac{1}{2\pi} \int\limits_{\mathbb{R}\setminus A} \xi^{2k} |Fx(\xi)|^2\, d\xi.$$

We have

$$J_1 \leq \operatorname*{ess\,sup}_{\xi\in A\setminus\Delta_\sigma} \left(\frac{\gamma_A^{2(n-k)}}{\xi^{2(n-k)}} |1 - a(\xi)|^2\right) \frac{\gamma_A^{-2(n-k)}}{2\pi} \int\limits_{A\setminus\Delta_\sigma} \xi^{2n} |Fx(\xi)|^2\, d\xi.$$

For I_3, estimate (3.11) holds. Therefore, if the inequality

$$\frac{\gamma_A^{2(n-k)}}{\xi^{2(n-k)}} |1 - a(\xi)|^2 \leq 1 \tag{3.13}$$

is satisfied for a.e. $\xi \in A \setminus \Delta_\sigma$, then the functional in (3.12) is estimated by

$$\frac{\gamma_A^{-2(n-k)}}{2\pi} \int\limits_{|\xi|\geq\sigma} \xi^{2n} |Fx(\xi)|^2\, d\xi \leq \gamma_A^{-2(n-k)}.$$

It remains to notice that condition (3.13) is equivalent to

$$a(\xi) = 1 + \theta(\xi) \left|\frac{\xi}{\gamma_A}\right|^{n-k}.$$

4. In case (4), the proof repeats almost word for word the proof of case (2) for the domain Σ_1 (here $\lambda_1 = 1$ and $\lambda_2 = \gamma_A^{-2n}$). $\square$

Proof of Corollary 1. Let us consider only case (1). The condition $S_a \leq 1$ obtained in the proof of Theorem 1 implies that the inequality

$$\frac{|1-a(\xi)|^2}{\lambda_2 \xi^{2n}} + \frac{|a(\xi)|^2}{\lambda_1} \leq \xi^{-2k} \tag{3.14}$$

holds almost everywhere. This implies that for those $\xi \in A \setminus \Delta_\sigma$ for which $|\xi| \geq \lambda_0 = \lambda_2^{-1/(2(n-k))}$, we can set $a(\xi) = 0$.

It follows immediately from the same inequality (3.14) that on the set $\sigma < |\xi| < \sigma_0$, where $\sigma_0 = \lambda_1^{1/(2k)}$, we can take the smoothing factor $a(\cdot)$ equal to one. $\square$

Corollary 2 follows from Corollary 1 for $\sigma = 0$.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project nos. 14-01-00456 and 14-01-00744.

REFERENCES

1. S. M. Nikol'skii, *Quadrature Formulas* (Nauka, Moscow, 1988) [in Russian].
2. S. M. Nikol'skii, "On estimates for approximation by quadrature formulas," Usp. Mat. Nauk **5** (2), 165–177 (1950).
3. A. N. Kolmogorov, "Über die beste Annäherung von Funktionen einer gegebenen Functionenklasse," Ann. Math., Ser. 2, **37**, 107–110 (1936). See also: "On the best approximation of functions of a given function class," in *Selected Works. Mathematics and Mechanics* (Nauka, Moscow, 1985), pp. 186–189 [in Russian].
4. S. A. Smolyak, "On optimal recovery of functions and functionals on them," Cand. Sci. (Phys.–Math.) Dissertation (Mosk. Gos. Univ., Moscow, 1965).
5. K. Yu. Osipenko, "Optimal interpolation of analytic functions," Mat. Zametki **12** (4), 465–476 (1972) [Math. Notes **12**, 712–719 (1972)].
6. C. A. Micchelli and T. J. Rivlin, "A survey of optimal recovery," in *Optimal Estimation in Approximation Theory: Proc. Int. Symp., Freudenstadt, 1976*, Ed. by C. A. Micchelli and T. J. Rivlin (Plenum, New York, 1977), pp. 1–54.
7. A. A. Melkman and C. A. Micchelli, "Optimal estimation of linear operators in Hilbert spaces from inaccurate data," SIAM J. Numer. Anal. **16**, 87–105 (1979).
8. J. F. Traub and H. Woźniakowski, *A General Theory of Optimal Algorithms* (Academic, New York, 1980).
9. C. A. Micchelli and T. J. Rivlin, "Lectures on optimal recovery," in *Numerical Analysis: Proc. SERC Summer School, Lancaster, 1984* (Springer, Berlin, 1985), Lect. Notes Math. **1129**, pp. 21–93.
10. V. V. Arestov, "Optimal recovery of operators, and related problems," Tr. Mat. Inst. im. V.A. Steklova, Akad. Nauk SSSR **189**, 3–20 (1989) [Proc. Steklov Inst. Math. **189**, 1–20 (1990)].
11. G. G. Magaril-Il'yaev and K. Yu. Osipenko, "Optimal recovery of functionals based on inaccurate data," Mat. Zametki **50** (6), 85–93 (1991) [Math. Notes **50**, 1274–1279 (1991)].
12. G. G. Magaril-Il'yaev and V. M. Tikhomirov, "Kolmogorov-type inequalities for derivatives," Mat. Sb. **188** (12), 73–106 (1997) [Sb. Math. **188**, 1799–1832 (1997)].
13. G. G. Magaril-Il'yaev and K. Yu. Osipenko, "Optimal recovery of functions and their derivatives from inaccurate information about the spectrum and inequalities for derivatives," Funkts. Anal. Prilozh. **37** (3), 51–64 (2003) [Funct. Anal. Appl. **37**, 203–214 (2003)].
14. G. G. Magaril-Il'yaev and K. Yu. Osipenko, "On optimal harmonic synthesis from inaccurate spectral data," Funkts. Anal. Prilozh. **44** (3), 76–79 (2010) [Funct. Anal. Appl. **44**, 223–225 (2010)].
15. G. G. Magaril-Il'yaev and K. Yu. Osipenko, "Hardy–Littlewood–Paley inequality and recovery of derivatives from inaccurate data," Dokl. Akad. Nauk **438** (3), 300–302 (2011) [Dokl. Math. **83** (3), 337–339 (2011)].

Translated by I. Nikitin