Sbornik: *Mathematics* **0**:0 1–000 © 1809 RAS(DoM) and LMS

Matematicheskii Sbornik **0**:0 000–000 [Константин Юрьевич Осипенко] DOI 10.4213/sm8330

Optimal recovery of linear operators in non-Euclidean metrics

K. Yu. Osipenko

Abstract. The paper is concerned with recovery problems of operators from noisy information in non-Euclidean metrics. A number of general theorems is put forward and applied to recovery problems of functions and their derivatives from noisy Fourier transform. In some cases, a family of optimal methods is found, from which the methods requiring the least amount of original information are singled out.

Bibliography: 25 titles.

Keywords: optimal recovery, linear operator, the Fourier transform.

Introduction

Given a linear operator Λ acting from a linear space X into a normed linear space Z, the general problem of optimal recovery of Λ on a set $W \subset X$ from noisy values of another linear operator $I\colon X \to Y$, where Y is a normed linear space, may be stated as the problem of finding, for a fixed $\delta \geqslant 0$, first, the quantity

$$E(\Lambda, W, I, \delta) = \inf_{m\colon Y \to Z} \sup_{\substack{x \in W,\ y \in Y \\ \|Ix - y\|_Y \leqslant \delta}} \|\Lambda x - m(y)\|_Z, \tag{0.1}$$

known as the *optimal recovery error*, and second, a mapping (a method), called an *optimal recovery method*, on which the infimum in (0.1) is attained. Here, δ describes the noise level in the original information.

In the simplest case, when Λ is a linear functional, Y is a finite-dimensional space, and $\delta = 0$, this problem was posed by Smolyak [1]. In particular, he proved that, for a convex centrally-symmetric set W, there is a linear method among the optimal recovery methods. This result and the formulation of the problem itself were published only in his Candidate Thesis and are not widely available. This topic was brought to attention by Bakhvalov [2], who initiated further studies in this direction. As a result, some optimal recovery methods for specific problems were put forward and the original formulation of the problem was extended to the complex case and to the case when the original information is given with noise (see [3]–[5]).

Subsequently, much research has been devoted to extensions of the original formulation of this problem (see [6]–[14], and the references given therein). The final

This research was carried out with the financial support of the Russian Foundation for Basic Research (grant nos. 13-01-12447, 14-01-00456).

AMS 2010 *Mathematics Subject Classification*. Primary 49N30, 35Q93.

(in a sense) result for linear functionals (namely, necessary and sufficient conditions for the existence of an optimal linear method) was put forward in [8].

The case when a set W, on which the operator Λ is recovered, is also given by some linear operator,

$$W = \{x \in X : \|I_1 x\|_{Y_1} \leqslant \delta_1\},$$

has also frequently engaged the attention. Here, $I_1 \colon X \to Y_1$ and Y_1 is a normed linear space. A general result on the existence of an optimal linear method in the case when Y, Y_1 and Z are Hilbert spaces was obtained in the paper [14], which also contains the first concrete results on the recovery of linear operators. This topic was further developed in [15]–[17] on the basis of methods residing in general principles of the theory of extreme values. However, all these works made crucial use of the Euclidean structure of spaces under study.

The optimal recovery problem of linear operators is closely connected with approximation of these operators by operators with bounded norms (S. B. Stechkin's problem). There is an intimate connection between the approximation errors in these problems, as well as between the corresponding extremal operators—this often helps one to simultaneously solve these problems (see [18], [19]). In turn, in some particular cases both these problems result in sharp inequalities for derivatives—this topics has received extensive treatment.

A number of exact solutions to Stechkin's problem and sharp constants in inequalities for derivatives in the operator case (here a metric is not uniform, in which the operator or the derivative is estimated, for otherwise the problem reduces to the functional case) was also obtained for non-Euclidean metrics (exact solutions are conveniently tabulated in [18]). Nevertheless, only few optimal recovery methods of operators are known explicitly in the non-Euclidean case. One of such examples is given in [6], Theorem 12 on p. 45 (later, we will consider this example in detail, obtaining for it, as a corollary to our general results, a family of optimal methods). Another example of construction of an optimal method in the non-Euclidean case was proposed in [20].

The aim of this paper is to obtain a number of general results on recovery of linear operators in the non-Euclidean case.

§ 1. The general formulation

Let T be a nonempty set, Σ be the σ-algebra of subsets of T, and μ be a nonnegative σ-additive measure on Σ. We let $L_p(T,\Sigma,\mu)$ (or $L_p(T,\mu)$, for brevity) denote the space of all classes of Σ-measurable functions with values in $\mathbb{R}$ or $\mathbb{C}$, for which

$$\|x(\,\cdot\,)\|_{L_p(T,\mu)} = \left(\int_T |x(t)|^p\,d\mu\right)^{\frac{1}{p}} < \infty, \qquad 1 \leqslant p < \infty,$$
$$\|x(\,\cdot\,)\|_{L_\infty(T,\mu)} = \operatorname*{vrai\,sup}_{t\in T} |x(t)| < \infty, \qquad p = \infty.$$

We set

$$\mathscr{W} = \bigl\{x(\,\cdot\,) \in L_p(T,\mu) : \|\varphi(\,\cdot\,)x(\,\cdot\,)\|_{L_r(T,\mu)} < \infty\bigr\},$$
$$W = \bigl\{x(\,\cdot\,) \in \mathscr{W} : \|\varphi(\,\cdot\,)x(\,\cdot\,)\|_{L_r(T,\mu)} \leqslant 1\bigr\},$$

where $1 \leqslant p, r \leqslant \infty$, and $\varphi(\,\cdot\,)$ is some function on T.

For an operator $\Lambda x(\,\cdot\,) = \psi(\,\cdot\,)x(\,\cdot\,)$, $\Lambda\colon \mathscr{W} \to L_q(T,\mu)$, $1 \leqslant q \leqslant \infty$, where $\psi(\,\cdot\,)$ is some function on T, we consider the recovery problem of Λ on the class W from the function $x(\,\cdot\,) \in W$, which is known with errors on some subset of T. More precisely, we shall assume that, for each function $x(\,\cdot\,) \in W$, one knows the function $y(\,\cdot\,) \in L_p(T_0,\mu)$, $T_0 \subset T$, such that $\|x(\,\cdot\,) - y(\,\cdot\,)\|_{L_p(T_0,\mu)} \leqslant \delta$, $\delta \geqslant 0$. It is required to recover $\Lambda x(\,\cdot\,)$ from the function $y(\,\cdot\,)$.

As recovery methods we consider all possible mappings $m\colon L_p(T_0,\mu) \to L_q(T,\mu)$. The error of a method m is defined as

$$e(p,q,r,m) = \sup_{\substack{x(\,\cdot\,)\in W,\ y(\,\cdot\,)\in L_p(T_0,\mu)\\ \|x(\,\cdot\,)-y(\,\cdot\,)\|_{L_p(T_0,\mu)}\leqslant\delta}} \|\Lambda x(\,\cdot\,) - m(y)(\,\cdot\,)\|_{L_q(T,\mu)}.$$

The quantity

$$E(p,q,r) = \inf_{m\colon L_p(T_0,\mu)\to L_q(T,\mu)} e(p,q,r,m)$$

is known as the optimal recovery error, and a method on which this infimum is attained is called optimal.

It is easily checked that

$$E(p,q,r) \geqslant \sup_{\substack{x(\,\cdot\,)\in W\\ \|x(\,\cdot\,)\|_{L_p(T_0,\mu)}\leqslant\delta}} \|\Lambda x(\,\cdot\,)\|_{L_q(T,\mu)}. \tag{1.1}$$

Indeed, let $x(\,\cdot\,) \in W$, $\|x(\,\cdot\,)\|_{L_p(T_0,\mu)} \leqslant \delta$, and let $m\colon L_p(T_0,\mu) \to L_q(T,\mu)$ be an arbitrary recovery method. Since $x(\,\cdot\,) \in W$ and $-x(\,\cdot\,) \in W$, we have

$$\begin{aligned} 2\|\Lambda x(\,\cdot\,)\|_{L_q(T,\mu)} &\leqslant \|\Lambda x(\,\cdot\,) - m(0)(\,\cdot\,)\|_{L_q(T,\mu)} + \|-\Lambda x(\,\cdot\,) - m(0)(\,\cdot\,)\|_{L_q(T,\mu)} \\ &\leqslant 2e(p,q,r,m). \end{aligned}$$

It follows that, for any method m,

$$e(p,q,r,m) \geqslant \sup_{\substack{x(\,\cdot\,)\in W\\ \|x(\,\cdot\,)\|_{L_p(T_0,\mu)}\leqslant\delta}} \|\Lambda x(\,\cdot\,)\|_{L_q(T,\mu)}.$$

Now the required inequality follows by taking the infimum on the left over all methods.

The extremal problem emerging on the right of (1.1), known as the *dual* problem, may be written as

$$\|\psi(\,\cdot\,)x(\,\cdot\,)\|_{L_q(T,\mu)} \to \max, \qquad \|x(\,\cdot\,)\|_{L_p(T_0,\mu)} \leqslant \delta, \quad \|\varphi(\,\cdot\,)x(\,\cdot\,)\|_{L_r(T,\mu)} \leqslant 1. \tag{1.2}$$

For $T_0 = T \subset \mathbb{R}^n$ and $q = 1$ (the constraint $q = 1$ is immaterial, because changing $y(t) = |x(t)|^q$ reduces the case $q \leqslant p$, $q \leqslant r$ to the one in question), problem (1.2) was examined in [21] in connection with Stechkin's problem.

The emphasis in the present paper is on the construction of optimal recovery methods for an operator Λ. Under this approach, problem (1.2) is studied with the help of the Lagrange function, which enables one in a number of cases (when some two of p, q and r coincide) to obtain explicit expressions both for the value of problem (1.2) and for an optimal recovery method in terms of Lagrange multipliers. In § 6 we consider the cases when these multipliers may be explicitly calculated.

In the present paper, the general scheme for constructing optimal recovery methods is as follows. First we solve the dual problem (1.2) or estimate its value; then a method or a family of methods is constructed whose the error is estimated by the same quantity. As a result, in all cases to be considered below, the quantity $E(p,q,r)$ coincides with the value of problem (1.2); that is, inequality (1.1) becomes an equality. Each time an optimal method on the set T_0 is sought in the form $\alpha(\,\cdot\,)\psi(\,\cdot\,)y(\,\cdot\,)$, where the function $\alpha(\,\cdot\,)$ plays the role of some filter.

To start with, we give one straightforward result (resembling the sufficient conditions in the Kuhn–Tucker theorem), which will be required in solving the extremal problem (1.2).

Let $f_j\colon A\to\mathbb{R}$, $j=0,1,\dots,n$, be functions defined on some set A. Consider the extremal problem

$$f_0(x)\to\max,\qquad f_j(x)\leqslant 0,\qquad j=1,\dots,n,\quad x\in A, \tag{1.3}$$

and write down its Lagrange function

$$\mathscr{L}(x,\lambda)=-f_0(x)+\sum_{j=1}^{n}\lambda_j f_j(x),\qquad \lambda=(\lambda_1,\dots,\lambda_n).$$

Lemma 1. *Assume that there exist $\widehat{\lambda}_j\geqslant 0$, $j=1,\dots,n$, and an element $\widehat{x}\in A$, admissible for problem* (1.3), *such that*

(a) $\min\limits_{x\in A}\mathscr{L}(x,\widehat{\lambda})=\mathscr{L}(\widehat{x},\widehat{\lambda})$, $\quad\widehat{\lambda}=(\widehat{\lambda}_1,\dots,\widehat{\lambda}_n)$,

(b) $\sum\limits_{j=1}^{n}\widehat{\lambda}_j f_j(\widehat{x})=0$.

Then $\widehat{x}$ is an extremal element for problem (1.3).

Proof. Indeed, for any element $x\in A$, admissible for problem (1.3),

$$-f_0(x)\geqslant\mathscr{L}(x,\widehat{\lambda})\geqslant\mathscr{L}(\widehat{x},\widehat{\lambda})=-f_0(\widehat{x}).$$

§ 2. The case $r = q$

Assume that $1\leqslant q<p<\infty$, $r=q$.

Theorem 1. *Let $1\leqslant q<p<\infty$ and $\delta>0$. Assume that $\widehat{\lambda}_2>0$ satisfies the condition*

$$\begin{aligned}\biggl(\int_{T_0}\bigl(|\psi(t)|^q-\widehat{\lambda}_2|\varphi(t)|^q\bigr)_+^{\frac{p}{p-q}}\,d\mu(t)\biggr)^{\frac1p}&\\ =\delta\biggl(\int_{T_0}|\varphi(t)|^q\bigl(|\psi(t)|^q-\widehat{\lambda}_2|\varphi(t)|^q\bigr)_+^{\frac{q}{p-q}}\,d\mu(t)\biggr)^{\frac1q}&>0\end{aligned}\tag{2.1}$$

(*functions $\varphi(\,\cdot\,)$ and $\psi(\,\cdot\,)$ are assumed to be such that the integrals in* (2.1) *exist*) *and that $|\psi(t)|^q-\widehat{\lambda}_2|\varphi(t)|^q\leqslant 0$ for almost all $t\notin T_0$. Then*

$$E(p,q,q)=\Bigl(\frac{p}{q}\,\widehat{\lambda}_1\delta^p+\widehat{\lambda}_2\Bigr)^{\frac1q},$$

where

$$\widehat{\lambda}_1 = \frac{q}{p}\,\delta^{q-p}\biggl(\int_{T_0}\bigl(|\psi(t)|^q - \widehat{\lambda}_2|\varphi(t)|^q\bigr)_+^{\frac{p}{p-q}}\,d\mu(t)\biggr)^{\frac{p-q}{p}}.$$

Moreover, the method

$$\widehat{m}(y)(t) = \begin{cases} \biggl(1 - \widehat{\lambda}_2\dfrac{|\varphi(t)|^q}{|\psi(t)|^q}\biggr)_+\psi(t)y(t), & t \in T_0,\\ 0, & t \notin T_0, \end{cases} \tag{2.2}$$

is optimal.

Proof. 1. *Lower estimate.* The extremal problem (1.2) (for numerical convenience, we raise the quantity to be maximized in the qth power) is as follows:

$$\begin{gathered} \int_T |\psi(t)x(t)|^q\,d\mu(t) \to \max,\\ \int_{T_0}|x(t)|^p\,d\mu(t) \leqslant \delta^p, \qquad \int_T |\varphi(t)x(t)|^q\,d\mu(t) \leqslant 1. \end{gathered} \tag{2.3}$$

The Lagrange function for this problem reads as

$$\mathscr{L}\bigl(x(\,\cdot\,),\lambda_1,\lambda_2\bigr) = \int_T L\bigl(t,x(t),\lambda_1,\lambda_2\bigr)\,d\mu(t),$$

where

$$L(t,x,\lambda_1,\lambda_2) = \begin{cases} -|\psi(t)x|^q + \lambda_1|x|^p + \lambda_2|\varphi(t)x|^q, & t \in T_0,\\ -|\psi(t)x|^q + \lambda_2|\varphi(t)x|^q, & t \notin T_0. \end{cases}$$

We take $\widehat{x}(\,\cdot\,)$ so as to minimize $L(t,x(t),\lambda_1,\lambda_2)$ for each t. It is easily checked that

$$\widehat{x}(t) = \begin{cases} \biggl(\dfrac{q}{p\widehat{\lambda}_1}\bigl(|\psi(t)|^q - \widehat{\lambda}_2|\varphi(t)|^q\bigr)_+\biggr)^{\frac{1}{p-q}}, & t \in T_0,\\ 0, & t \notin T_0. \end{cases}$$

As a result,

$$\mathscr{L}\bigl(x(t),\widehat{\lambda}_1,\widehat{\lambda}_2\bigr) \geqslant \mathscr{L}\bigl(\widehat{x}(t),\widehat{\lambda}_1,\widehat{\lambda}_2\bigr).$$

From the definition of $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$, we have

$$\int_{T_0}|\widehat{x}(t)|^p\,d\mu(t) = \delta^p, \qquad \int_T |\varphi(t)\widehat{x}(t)|^q\,d\mu(t) = 1. \tag{2.4}$$

By Lemma 1, $\widehat{x}(\,\cdot\,)$ is a solution of problem (2.3). Hence, the value of this problem is

$$\int_T |\psi(t)\widehat{x}(t)|^q\,d\mu(t).$$

The equality

$$-q|\psi(t)\widehat{x}(t)|^q + p\widehat{\lambda}_1|\widehat{x}(t)|^p + q\widehat{\lambda}_2|\varphi(t)\widehat{x}(t)|^q = 0,$$

easily follows from the definition of $\widehat{x}(\,\cdot\,)$. Integrating this equality over the set T, this gives

$$\int_T |\psi(t)\widehat{x}(t)|^q\,d\mu(t) = \frac{p}{q}\widehat{\lambda}_1\delta^p + \widehat{\lambda}_2.$$

Finally, from (1.1),

$$E(p,q,q) \geqslant \left(\frac{p}{q}\widehat{\lambda}_1\delta^p + \widehat{\lambda}_2\right)^{\frac{1}{q}}.$$

2. *Upper estimate.* We set

$$\alpha(t) = \begin{cases} \left(1 - \lambda_2\dfrac{|\varphi(t)|^q}{|\psi(t)|^q}\right)_+, & t \in T_0,\\ 0, & t \notin T_0. \end{cases}$$

To estimate the error of method (2.2) we need to find the value of the extremal problem:

$$\begin{gathered} \int_{T_0} |\psi(t)|^q \left|x(t) - \alpha(t)y(t)\right|^q d\mu(t) + \int_{T\setminus T_0} |\psi(t)x(t)|^q\, d\mu(t) \to \max,\\ \int_{T_0} |x(t) - y(t)|^p\, d\mu(t) \leqslant \delta^p, \qquad \int_T |\varphi(t)x(t)|^q\, d\mu(t) \leqslant 1. \end{gathered} \tag{2.5}$$

Taking $z(\,\cdot\,) = x(\,\cdot\,) - y(\,\cdot\,)$, we rewrite problem (2.5) as follows:

$$\begin{gathered} \int_{T_0} |\psi(t)|^q \left|(1-\alpha(t))x(t) + \alpha(t)z(t)\right|^q d\mu(t) + \int_{T\setminus T_0} |\psi(t)x(t)|^q\, d\mu(t) \to \max,\\ \int_{T_0} |z(t)|^p\, d\mu(t) \leqslant \delta^p, \qquad \int_T |\varphi(t)|^q\, |x(t)|^q\, d\mu(t) \leqslant 1. \end{gathered} \tag{2.6}$$

Clearly, the value of this problem agrees with that of the problem

$$\begin{gathered} \int_T |\psi(t)|^q \big((1-\alpha(t))v(t) + \alpha(t)u(t)\big)^q\, d\mu(t) \to \max,\\ \int_{T_0} u^p(t)\, d\mu(t) \leqslant \delta^p, \qquad \int_T |\varphi(t)|^q v^q(t)\, d\mu(t) \leqslant 1,\\ u(t), v(t) \geqslant 0 \quad \text{for almost all} \quad t \in T. \end{gathered} \tag{2.7}$$

The Lagrange function for this problem reads as

$$\mathscr{L}_1\big(u(\,\cdot\,), v(\,\cdot\,), \lambda_1, \lambda_2\big) = \int_T L_1\big(t, u(t), v(t), \lambda_1, \lambda_2\big)\, d\mu(t),$$

where

$$L_1(t,u,v,\lambda_1,\lambda_2) = \begin{cases} -|\psi(t)|^q\big((1-\alpha(t))v + \alpha(t)u\big)^q & \\ \qquad +\lambda_1 u^p + \lambda_2|\varphi(t)|^q v^q, & t \in T_0,\\ -|\psi(t)|^q v^q + \lambda_2|\varphi(t)|^q v^q, & t \notin T_0. \end{cases}$$

We set $\lambda_1 = \widehat{\lambda}_1$, $\lambda_2 = \widehat{\lambda}_2$. Hence, for $\alpha(t) > 0$,

$$\frac{\partial L_1}{\partial v} = q\widehat{\lambda}_2|\varphi(t)|^q\big(v^{q-1} - ((1-\alpha(t))v + \alpha(t)u)^{q-1}\big).$$

Consequently, for $\alpha(t) > 0$ and any fixed $u > 0$, the minimum of the function $L_1(t,u,v,\widehat{\lambda}_1,\widehat{\lambda}_2)$ for $v \in (0,+\infty)$ is attained for $v = u$. If now $\alpha(t) = 0$, then

$L_1(t,u,v,\widehat{\lambda}_1,\widehat{\lambda}_2)\geqslant 0$. So, we have for all $u(t),v(t)\geqslant 0$

$$
\begin{aligned}
\mathscr{L}_1\bigl(u(\,\cdot\,),v(\,\cdot\,),\widehat{\lambda}_1,\widehat{\lambda}_2\bigr)&\geqslant\int_{T_0}L_1\bigl(u(\,\cdot\,),u(\,\cdot\,),\widehat{\lambda}_1,\widehat{\lambda}_2\bigr)\,d\mu(t)\\
&=\int_{T_0}L\bigl(t,u(\,\cdot\,),\widehat{\lambda}_1,\widehat{\lambda}_2\bigr)\,d\mu(t)\geqslant\int_{T_0}L\bigl(t,\widehat{x}(\,\cdot\,),\widehat{\lambda}_1,\widehat{\lambda}_2\bigr)\,d\mu(t)\\
&=\mathscr{L}_1\bigl(\widehat{x}(\,\cdot\,),\widehat{x}(\,\cdot\,),\widehat{\lambda}_1,\widehat{\lambda}_2\bigr).
\end{aligned}
$$

Taking into account (2.4) it follows that the functions $\widehat{u}(\,\cdot\,)=\widehat{v}(\,\cdot\,)=\widehat{x}(\,\cdot\,)$ are solutions of problem (2.7). Hence,

$$e^q(p,q,q,\widehat{m})=\int_T|\psi(t)\widehat{x}(t)|^q\,d\mu(t)=\frac{p}{q}\widehat{\lambda}_1\delta^p+\widehat{\lambda}_2\leqslant E^q(p,q,q).$$

It follows that the method $\widehat{m}$ is optimal and the optimal recovery error is as claimed.

§ 3. The case $q = p$

Assume that $1\leqslant p<r<\infty$, $q=p$ and $T_0=T$.

Theorem 2. *Let $1\leqslant p<r<\infty$ and $\delta>0$. Assume that $\widehat{\lambda}_1>0$ satisfies the condition*

$$
\begin{aligned}
&\biggl(\int_T|\varphi(t)|^{\frac{pr}{p-r}}\bigl(|\psi(t)|^p-\widehat{\lambda}_1\bigr)_+^{\frac{p}{r-p}}\,d\mu(t)\biggr)^{\frac1p}\\
&\qquad=\delta\biggl(\int_T|\varphi(t)|^{\frac{pr}{p-r}}\bigl(|\psi(t)|^p-\widehat{\lambda}_1\bigr)_+^{\frac{r}{r-p}}\,d\mu(t)\biggr)^{\frac1r}>0
\end{aligned}
\tag{3.1}
$$

(*functions $\varphi(\,\cdot\,)$ and $\psi(\,\cdot\,)$ will be assumed to be such that the integrals in* (3.1) *exist*). *Then*

$$E(p,p,r)=\biggl(\widehat{\lambda}_1\delta^p+\frac{r}{p}\widehat{\lambda}_2\biggr)^{\frac1p},$$

where

$$\widehat{\lambda}_2=\frac{p}{r}\,\delta^{p-r}\biggl(\int_T|\varphi(t)|^{\frac{pr}{p-r}}\bigl(|\psi(t)|^p-\widehat{\lambda}_1\bigr)_+^{\frac{p}{r-p}}\,d\mu(t)\biggr)^{\frac{r-p}{p}}.$$

Moreover, the method

$$\widehat{m}(y)(t)=\alpha(t)\psi(t)y(t),\tag{3.2}$$

where

$$\alpha(t)=1-\biggl(1-\frac{\widehat{\lambda}_1}{|\psi(t)|^p}\biggr)_+,$$

is optimal.

Proof. The proof is much like that of the previous theorem.

1. *Lower estimate.* Problem (1.2) (for numerical convenience, we raise the quantity to be maximized in the pth power) is as follows:

$$
\begin{gathered}
\int_T|\psi(t)x(t)|^p\,d\mu(t)\to\max,\\
\int_T|x(t)|^p\,d\mu(t)\leqslant\delta^p,\qquad\int_T|\varphi(t)x(t)|^r\,d\mu(t)\leqslant 1.
\end{gathered}
\tag{3.3}
$$

The Lagrange function for this problem reads as

$$\mathscr{L}\big(x(\,\cdot\,),\lambda_1,\lambda_2\big)=\int_T L\big(t,x(t),\lambda_1,\lambda_2\big)\,d\mu(t),$$

where

$$L(t,x,\lambda_1,\lambda_2)=-|\psi(t)x|^p+\lambda_1|x|^p+\lambda_2|\varphi(t)x|^r.$$

We again choose $\widehat{x}(\,\cdot\,)$ so as to minimize $L(t,x(t),\lambda_1,\lambda_2)$ for each t. We have

$$\widehat{x}(t)=\left(\frac{p(|\psi(t)|^p-\widehat{\lambda}_1)_+}{r\widehat{\lambda}_2|\varphi(t)|^r}\right)^{\frac{1}{r-p}}.$$

As a result,

$$\mathscr{L}\big(x(t),\widehat{\lambda}_1,\widehat{\lambda}_2\big)\geqslant\mathscr{L}\big(\widehat{x}(t),\widehat{\lambda}_1,\widehat{\lambda}_2\big).$$

By the definition of $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$,

$$\int_T|\widehat{x}(t)|^p\,d\mu(t)=\delta^p,\qquad\int_T|\varphi(t)\widehat{x}(t)|^r\,d\mu(t)=1.\tag{3.4}$$

From Lemma 1 it follows that $\widehat{x}(\,\cdot\,)$ is a solution of problem (3.3). Consequently, the value of this problem is

$$\int_T|\psi(t)\widehat{x}(t)|^p\,d\mu(t).$$

The equality

$$-p|\psi(t)\widehat{x}(t)|^q+p\widehat{\lambda}_1|\widehat{x}(t)|^p+r\widehat{\lambda}_2|\varphi(t)\widehat{x}(t)|^r=0,$$

follows easily from the definition of $\widehat{x}(\,\cdot\,)$. Integrating this equality over the set T, this gives

$$\int_T|\psi(t)\widehat{x}(t)|^p\,d\mu(t)=\widehat{\lambda}_1\delta^p+\frac{r}{p}\,\widehat{\lambda}_2.$$

Finally, from (1.1) we have

$$E(p,p,r)\geqslant\left(\widehat{\lambda}_1\delta^p+\frac{r}{p}\widehat{\lambda}_2\right)^{\frac{1}{p}}.$$

2. *Upper estimate.* To estimate the error of method (3.2) one needs to find the value of the following extremal problem:

$$\begin{gathered}\int_T|\psi(t)|^p\,\big|x(t)-\alpha(t)y(t)\big|^p\,d\mu(t)\to\max,\\ \int_T|x(t)-y(t)|^p\,d\mu(t)\leqslant\delta^p,\qquad\int_T|\varphi(t)x(t)|^r\,d\mu(t)\leqslant 1.\end{gathered}\tag{3.5}$$

Setting $z(\,\cdot\,)=x(\,\cdot\,)-y(\,\cdot\,)$, we rewrite problem (3.5) in the form

$$\begin{gathered}\int_T|\psi(t)|^p\,\big|(1-\alpha(t))x(t)+\alpha(t)z(t)\big|^p\,d\mu(t)\to\max,\\ \int_T|z(t)|^p\,d\mu(t)\leqslant\delta^p,\qquad\int_T|\varphi(t)x(t)|^r\,d\mu(t)\leqslant 1.\end{gathered}\tag{3.6}$$

Clearly, the value of this problem is the same as the value of the problem

$$\begin{gathered} \int_T |\psi(t)|^p \big((1-\alpha(t))v(t)+\alpha(t)u(t)\big)^p \, d\mu(t) \to \max, \\ \int_T u^p(t)\,d\mu(t) \leqslant \delta^p, \qquad \int_T |\varphi(t)|^r v^r(t)\,d\mu(t) \leqslant 1, \\ u(t), v(t) \geqslant 0 \quad \text{for almost all} \quad t\in T. \end{gathered} \tag{3.7}$$

The Lagrange function for this problem is as follows:

$$\mathscr{L}_1\big(u(\,\cdot\,), v(\,\cdot\,), \lambda_1, \lambda_2\big) = \int_T L_1\big(t, u(t), v(t), \lambda_1, \lambda_2\big)\,d\mu(t),$$

where

$$L_1(t,u,v,\lambda_1,\lambda_2) = -|\psi(t)|^p\big((1-\alpha(t))v+\alpha(t)u\big)^p + \lambda_1 u^p + \lambda_2 |\varphi(t)|^r v^r.$$

We set $\lambda_1=\widehat{\lambda}_1$, $\lambda_2=\widehat{\lambda}_2$. Hence, for $|\psi(t)|^p>\widehat{\lambda}_1$,

$$\frac{\partial L_1}{\partial u} = p\widehat{\lambda}_1\big(u^{p-1}-\big((1-\alpha(t))v+\alpha(t)u\big)^{p-1}\big).$$

As a result, for $|\psi(t)|^p>\widehat{\lambda}_1$ and any fixed $v>0$, the minimum of the function $L_1(u,v,\widehat{\lambda}_1,\widehat{\lambda}_2)$ for $u\in(0,+\infty)$ is attained for $u=v$. If now $|\psi(t)|^p\leqslant\widehat{\lambda}_1$, then $\alpha(t)=1$ and $L_1(u,v)\geqslant 0$. We set $T_1=\{t\in T: |\psi(t)|^p>\widehat{\lambda}_1\}$. Therefore, for all $u(t), v(t)\geqslant 0$,

$$\begin{aligned} \mathscr{L}_1\big(u(\,\cdot\,), v(\,\cdot\,), \widehat{\lambda}_1, \widehat{\lambda}_2\big) &\geqslant \int_{T_1} L_1\big(t, u(\,\cdot\,), u(\,\cdot\,), \widehat{\lambda}_1, \widehat{\lambda}_2\big)\,d\mu(t) \\ &= \int_{T_1} L\big(t, u(\,\cdot\,), \widehat{\lambda}_1, \widehat{\lambda}_2\big)\,d\mu(t) \geqslant \int_{T_1} L\big(t, \widehat{x}(\,\cdot\,), \widehat{\lambda}_1, \widehat{\lambda}_2\big)\,d\mu(t) \\ &= \mathscr{L}_1\big(\widehat{x}(\,\cdot\,), \widehat{x}(\,\cdot\,), \widehat{\lambda}_1, \widehat{\lambda}_2\big). \end{aligned}$$

Taking (3.4) into account we see that the functions $\widehat{u}(\,\cdot\,)=\widehat{v}(\,\cdot\,)=\widehat{x}(\,\cdot\,)$ are solution of problem (3.7). Hence,

$$e^p(p,p,r,\widehat{m}) = \int_T |\psi(t)\widehat{x}(t)|^q\,d\mu(t) = \widehat{\lambda}_1\delta^p + \frac{r}{p}\widehat{\lambda}_2 \leqslant E^p(p,p,r).$$

Thus, the method $\widehat{m}$ is optimal and the optimal recovery error is as claimed.

§ 4. The case $r = p$

Assume that $1\leqslant q<p=r<\infty$. Let $\chi_0(\,\cdot\,)$ be the characteristic function of the set T_0:

$$\chi_0(t) = \begin{cases} 1, & t\in T_0, \\ 0, & t\notin T_0. \end{cases}$$

Theorem 3. *Let $1\leqslant q<p<\infty$ and $\delta>0$. Assume that $\widehat{\lambda}_2>0$ satisfies the condition*

$$\int_{T_0}\bigg(\frac{|\psi(t)|^q}{1+\widehat{\lambda}_2|\varphi(t)|^p}\bigg)^{\frac{p}{p-q}}\,d\mu(t) = \delta^p\int_T|\varphi(t)|^p\bigg(\frac{|\psi(t)|^q}{\chi_0(t)+\widehat{\lambda}_2|\varphi(t)|^p}\bigg)^{\frac{p}{p-q}}\,d\mu(t) > 0 \tag{4.1}$$

(functions $\varphi(\,\cdot\,)$ and $\psi(\,\cdot\,)$ are assumed to be such that the integrals in (4.1) *exist). Then*

$$E(p,q,p)=(\widehat{\lambda}_1\delta^p+\widehat{\lambda}_1\widehat{\lambda}_2)^{\frac{1}{q}},$$

where

$$\lambda_1=\delta^{q-p}\biggl(\int_{T_0}\biggl(\frac{|\psi(t)|^q}{1+\widehat{\lambda}_2|\varphi(t)|^p}\biggr)^{\frac{p}{p-q}}\,d\mu(t)\biggr)^{\frac{p-q}{p}}.$$

Moreover, the method

$$\widehat{m}(y)(t)=\begin{cases}\dfrac{\psi(t)}{1+\widehat{\lambda}_2|\varphi(t)|^p}y(t), & t\in T_0,\\ 0, & t\notin T_0,\end{cases}$$

is optimal.

Proof. 1. *Lower bound.* Problem (1.2) (as above, we are dealing with an equivalent problem) is as follows:

$$\begin{gathered}\int_T|\psi(t)x(t)|^q\,d\mu(t)\to\max,\\ \int_{T_0}|x(t)|^p\,d\mu(t)\leqslant\delta^p,\qquad \int_T|\varphi(t)x(t)|^p\,d\mu(t)\leqslant 1.\end{gathered}\tag{4.2}$$

We set

$$\widehat{x}(t)=\widehat{\lambda}_1^{-\frac{1}{p-q}}\biggl(\frac{|\psi(t)|^q}{\chi_0(t)+\widehat{\lambda}_2|\varphi(t)|^p}\biggr)^{\frac{1}{p-q}}.$$

It is easily seen from the definition of $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$ that

$$\int_{T_0}|\widehat{x}(t)|^p\,d\mu(t)=\delta^p,\qquad \int_T|\varphi(t)\widehat{x}(t)|^p\,d\mu(t)=1.$$

Hence, $\widehat{x}(\,\cdot\,)$ is an admissible function for problem (4.2). As a result, the value of this problem is at most

$$\int_T|\psi(t)\widehat{x}(t)|^q\,d\mu(t).$$

The equality

$$|\psi(t)\widehat{x}(t)|^q=\widehat{\lambda}_1|\widehat{x}(t)|^p\chi_0(t)+\widehat{\lambda}_1\widehat{\lambda}_2|\varphi(t)\widehat{x}(t)|^p,$$

is an easy consequence of the definition of $\widehat{x}(\,\cdot\,)$. Integrating this equality over the set T, this gives

$$\int_T|\psi(t)\widehat{x}(t)|^q\,d\mu(t)=\widehat{\lambda}_1\delta^p+\widehat{\lambda}_1\widehat{\lambda}_2.$$

Finally, from (1.1) we have

$$E(p,q,p)\geqslant(\widehat{\lambda}_1\delta^p+\widehat{\lambda}_1\widehat{\lambda}_2)^{\frac{1}{q}}.$$

2. *Upper estimate.* An optimal recovery method will be sought in the form

$$\widehat{m}(y)(t)=\begin{cases}\alpha(t)\psi(t)y(t), & t\in T_0,\\ 0, & t\notin T_0.\end{cases}$$

To estimate the error of this method we need to find the value of the following extremal problem:

$$\begin{gathered}\int_{T_0} |\psi(t)|^q \left|x(t)-\alpha(t)y(t)\right|^q d\mu(t) + \int_{T\setminus T_0} |\psi(t)x(t)|^q \, d\mu(t) \to \max, \\ \int_{T_0} |x(t)-y(t)|^p \, d\mu(t) \leqslant \delta^p, \qquad \int_T |\varphi(t)x(t)|^p \, d\mu(t) \leqslant 1.\end{gathered} \tag{4.3}$$

Using Hölder's inequality,

$$\begin{gathered}|(1-\alpha(t))x(t)+\alpha(t)(x(t)-y(t))|^q \leqslant h(t)\big(\widehat{\lambda}_2|\varphi(t)x(t)|^p + |x(t)-y(t)|^p\big)^{\frac{q}{p}}, \\ h(t) = \left(\frac{|1-\alpha(t)|^{p'}}{\widehat{\lambda}_2^{p'/p}|\varphi(t)|^{p'}} + |\alpha(t)|^{p'}\right)^{\frac{q}{p'}}, \qquad \frac{1}{p}+\frac{1}{p'}=1.\end{gathered} \tag{4.4}$$

Hence, the value of problem (4.3) is estimated by

$$\int_T f(t)g(t)\, d\mu(t), \tag{4.5}$$

where

$$f(t) = \begin{cases} |\psi(t)|^q h(t), & t\in T_0, \\ \dfrac{|\psi(t)|^q}{\widehat{\lambda}_2^{\frac{q}{p}}|\varphi(t)|^q}, & t \in T\setminus T_0, \end{cases}$$

$$g(t) = \begin{cases} \big(\widehat{\lambda}_2|\varphi(t)x(t)|^p + |z(t)|^p\big)^{\frac{q}{p}}, & t\in T_0, \\ \widehat{\lambda}_2^{\frac{q}{p}}|\varphi(t)x(t)|^q, & t \in T\setminus T_0. \end{cases}$$

Let

$$\alpha(t) = \frac{1}{1+\widehat{\lambda}_2|\varphi(t)|^p}.$$

Then

$$f(t) = \frac{|\psi(t)|^q}{(\chi_0(t)+\widehat{\lambda}_2|\varphi(t)|^p)^{\frac{q}{p}}}.$$

An application of Hölder's inequality to (4.5) gives the bound

$$\left(\int_T |f(t)|^{s'}\, d\mu(t)\right)^{\frac{1}{s'}} \left(\int_T |g(t)|^{s}\, d\mu(t)\right)^{\frac{1}{s}},$$

where $1/s+1/s'=1$. Taking $s=p/q$, we have, for this bound,

$$\begin{aligned}&\left(\int_T |f(t)|^{\frac{p}{p-q}}\, d\mu(t)\right)^{\frac{p-q}{p}} \left(\int_T |g(t)|^{\frac{p}{q}}\, d\mu(t)\right)^{\frac{q}{p}} \\ &\qquad \leqslant (\delta^p+\widehat{\lambda}_2)^{\frac{q}{p}} \left(\int_T \frac{|\psi(t)|^{\frac{qp}{p-q}}}{(\chi_0(t)+\widehat{\lambda}_2|\varphi(t)|^p)^{\frac{q}{p-q}}}\, d\mu(t)\right)^{\frac{p-q}{p}} \\ &\qquad = (\widehat{\lambda}_1\delta^p+\widehat{\lambda}_1\lambda_2)^{\frac{q}{p}} \left(\int_T |\psi(t)\widehat{x}(t)|^q\, d\mu(t)\right)^{\frac{p-q}{p}} = \widehat{\lambda}_1\delta^p+\widehat{\lambda}_1\widehat{\lambda}_2.\end{aligned}$$

Therefore,

$$e^q(p,q,p,\widehat{m}) \leqslant \widehat{\lambda}_1\delta^p+\widehat{\lambda}_1\widehat{\lambda}_2 \leqslant E^q(p,q,p).$$

§ 5. The case $r = q = p$

Assume now that $1 \leqslant p < \infty$, $r = q = p$.

Theorem 4. *Let $1 \leqslant p < \infty$. Assume that there exist $\lambda_1, \lambda_2 \geqslant 0$ such that the value of the extremal problem*

$$\begin{gathered}\int_T |\psi(t)x(t)|^p\,d\mu(t) \to \max,\\ \int_{T_0} |x(t)|^p\,d\mu(t) \leqslant \delta^p, \qquad \int_T |\varphi(t)x(t)|^p\,d\mu(t) \leqslant 1,\end{gathered} \tag{5.1}$$

is not smaller than $\lambda_1\delta^p + \lambda_2$ and that, for almost all $t \in T$,

$$-|\psi(t)|^p + \lambda_1\chi_0(t) + \lambda_2|\varphi(t)|^p \geqslant 0, \tag{5.2}$$

where $\chi_0(\,\cdot\,)$ is the characteristic function of the set T_0. Then

$$E(p,p,p) = (\lambda_1\delta^p + \lambda_2)^{\frac{1}{p}}.$$

Moreover, each of the methods

$$\widehat{m}(y)(t) = \begin{cases} \alpha(t)\psi(t)y(t), & t \in T_0,\\ 0, & t \notin T_0,\end{cases} \tag{5.3}$$

is optimal, where, for $\delta > 0$, $\lambda_1, \lambda_2 > 0$, a function $\alpha(\,\cdot\,)$ satisfies the condition

$$\begin{cases} |\psi(t)|^{p'}\left(\dfrac{|1-\alpha(t)|^{p'}}{\lambda_2^{p'/p}|\varphi(t)|^{p'}} + \dfrac{|\alpha(t)|^{p'}}{\lambda_1^{p'/p}}\right) \leqslant 1, \quad 1 < p < \infty, \quad \dfrac{1}{p} + \dfrac{1}{p'} = 1,\\ \dfrac{|\psi(t)(1-\alpha(t))|}{\lambda_2|\varphi(t)|} \leqslant 1, \quad \dfrac{|\psi(t)\alpha(t)|}{\lambda_1} \leqslant 1, \quad p = 1 \end{cases} \tag{5.4}$$

for almost all $t \in T_0$ (in particular, for $\alpha(t) = \widehat{\alpha}(t) = \lambda_1(\lambda_1 + \lambda_2|\varphi(t)|^p)^{-1}$ method (5.3) *is optimal), and for $1 \leqslant p < \infty$, $\delta = 0$, the condition*

$$|\psi(t)|\,|1-\alpha(t)| \leqslant \lambda_2^{\frac{1}{p}}|\varphi(t)|. \tag{5.5}$$

For $\lambda_1 = 0$ the method $\widehat{m}(y)(\,\cdot\,) = 0$ is optimal, and for $\lambda_2 = 0$, method (5.3) *with $\alpha(\,\cdot\,) = 1$ is optimal.*

Proof. From the hypotheses of the theorem and inequality (1.1) it follows that

$$E(p,p,p) \geqslant (\lambda_1\delta^p + \lambda_2)^{\frac{1}{p}}.$$

Let $\delta > 0$ and $\lambda_1\lambda_2 > 0$. To estimate methods of form (5.3) we write down the extremal problem

$$\begin{gathered}\int_{T_0} |\psi(t)|^p\,\bigl|x(t) - \alpha(t)y(t)\bigr|^p\,d\mu(t) + \int_{T\setminus T_0} |\psi(t)x(t)|^p\,d\mu(t) \to \max,\\ \int_{T_0} |x(t) - y(t)|^p\,d\mu(t) \leqslant \delta^p, \qquad \int_T |\varphi(t)x(t)|^p\,d\mu(t) \leqslant 1.\end{gathered} \tag{5.6}$$

Similarly to (4.4), we have

$$\bigl|(1-\alpha(t))x(t) + \alpha(t)(x(t) - y(t))\bigr|^p \leqslant h_p(t)\bigl(\lambda_2|\varphi(t)x(t)|^p + \lambda_1|x(t) - y(t)|^p\bigr),$$

where

$$h_p(t)=\begin{cases}\left(\dfrac{|1-\alpha(t)|^{p'}}{\lambda_2^{p'/p}|\varphi(t)|^{p'}}+\dfrac{|\alpha(t)|^{p'}}{\lambda_1^{p'/p}}\right)^{p/p'}, & 1<p<\infty,\\ \max\left\{\dfrac{|1-\alpha(t)|}{\lambda_2|\varphi(t)|},\dfrac{|\alpha(t)|}{\lambda_1}\right\}, & p=1.\end{cases}$$

Hence, letting

$$S(\alpha(\,\cdot\,))=\operatorname*{vrai\,sup}_{t\in T_0}|\psi(t)|^p h_p(t)$$

and taking into account that $S(\alpha(\,\cdot\,))\leqslant 1$ by the hypotheses of the theorem and since inequality (5.2) for $t\in T\setminus T_0$ reads as $|\psi(t)|^p\leqslant\lambda_2|\varphi(t)|^p$, we have

$$\begin{aligned}\int_{T_0}|\psi(t)|^p\left|x(t)-\alpha(t)y(t)\right|^p d\mu(t)+\int_{T\setminus T_0}|\psi(t)x(t)|^p\,d\mu(t)\\ \leqslant S(\alpha(\,\cdot\,))\int_{T_0}\left(\lambda_2|\varphi(t)x(t)|^p+\lambda_1|x(t)-y(t)|^p\right)d\mu(t)\\ +\lambda_2\int_{T\setminus T_0}|\varphi(t)x(t)|^p\,d\mu(t)\leqslant\lambda_1\delta^p+\lambda_2.\end{aligned}$$

Consequently, it follows that

$$e(p,p,p,m)\leqslant(\lambda_1\delta^p+\lambda_2)^{\frac{1}{p}}\leqslant E(p,p,p).$$

It remains to show that the set of functions $\alpha(\,\cdot\,)$ satisfying conditions (5.4) is nonempty. Setting

$$\widehat{\alpha}(t)=\frac{\lambda_1}{\lambda_1+\lambda_2|\varphi(t)|^p},\tag{5.7}$$

we have

$$S(\widehat{\alpha}(\,\cdot\,))(t)=\operatorname*{vrai\,sup}_{t\in T_0}\left(\frac{|\psi(t)|^p}{\lambda_1+\lambda_2|\varphi(t)|^p}\right).$$

Inequality (5.2) shows that $S(\widehat{\alpha}(\,\cdot\,))(t)\leqslant 1$.

For $\delta=0$ we need to estimate the value of the extremal problem

$$\int_{T_0}|\psi(t)|^p\left|x(t)-\alpha(t)x(t)\right|^p d\mu(t)+\int_{T\setminus T_0}|\psi(t)x(t)|^p\,d\mu(t)\to\max,$$

$$\int_T|\varphi(t)x(t)|^p\,d\mu(t)\leqslant 1.$$

Using condition (5.5), this gives

$$|\psi(t)|^p\left|x(t)-\alpha(t)x(t)\right|^p\leqslant\lambda_2|\varphi(t)x(t)|^p.$$

Since $|\psi(t)|^p\leqslant\lambda_2|\varphi(t)|^p$ for $t\in T\setminus T_0$, we have

$$\int_{T_0}|\psi(t)|^p\left|x(t)-\alpha(t)x(t)\right|^p d\mu(t)+\int_{T\setminus T_0}|\psi(t)x(t)|^p\,d\mu(t)\leqslant\lambda_2.$$

If $\lambda_1=0$, then $|\psi(t)|^p\leqslant\lambda_2|\varphi(t)|^p$ for almost all $t\in T$, and so, for the method $\widehat{m}(y)(\,\cdot\,)=0$,

$$\int_T|\psi(t)x(t)|^p\,d\mu(t)\leqslant\lambda_2.$$

If $\lambda_2 = 0$, then $|\psi(t)|^p \leqslant \lambda_1$ for almost all $t \in T_0$ and $\psi(t) = 0$ for almost all $t \in T \setminus T_0$. Hence,

$$\int_{T_0} \bigl|\psi(t)x(t) - \psi(t)y(t)\bigr|^p \, d\mu(t) + \int_{T\setminus T_0} |\psi(t)x(t)|^p \, d\mu(t) \\ \leqslant \lambda_1 \int_{T_0} |x(t) - y(t)|^p \, d\mu(t) \leqslant \lambda_1 \delta^p.$$

The action of the so-obtained optimal methods can be regarded by the action of the operator being recovered multiplied by some function, the latter may be looked upon as a filter or a smoothing multiplier. For example, for (5.7) the method is as follows:

$$\widehat{m}(y)(t) = \begin{cases} \dfrac{\lambda_1 \psi(t)}{\lambda_1 + \lambda_2 |\varphi(t)|^p} \, y(t), & t \in T_0, \\ 0, & t \notin T_0, \end{cases}$$

the function

$$\frac{\lambda_1}{\lambda_1 + \lambda_2 |\varphi(t)|^p}$$

may be regarded as such a filter.

We shall be concerned with the sets for which the use of filtration may be dispensed with (in other words, this multiplier may be taken to be 1). Besides, we shall be interested in how the original set, on which a noisy information about the function is given, may be reduced without increasing the optimal recovery error. In other words, our aim is to find all the sets on which we may put $\alpha(t) = \psi(t)$ and $\alpha(t) = 0$.

We set

$$T^0 = \bigl\{t \in T_0 : |\psi(t)| > \lambda_2^{\frac{1}{p}} |\varphi(t)|\bigr\}, \qquad T^1 = \bigl\{t \in T^0 : |\psi(t)| \leqslant \lambda_1^{\frac{1}{p}}\bigr\}.$$

Corollary 1. *For $\delta > 0$, $\lambda_1, \lambda_2 > 0$, the methods*

$$\widehat{m}(y)(t) = \begin{cases} \psi(t)y(t), & t \in T^1, \\ \alpha(t)\psi(t)y(t), & t \in T^0 \setminus T^1, \\ 0, & t \in T \setminus T^0, \end{cases}$$

with $\alpha(\,\cdot\,)$ satisfying conditions (5.4) *are optimal. For $\delta = 0$, the methods*

$$\widehat{m}(y)(t) = \begin{cases} \psi(t)y(t), & t \in T^0, \\ 0, & t \in T \setminus T^0, \end{cases}$$

are optimal.

From Corollary 1 it follows that there exist optimal methods that use the given noisy information only on the set T^0. In other words, information on the set $T \setminus T^0$ becomes superfluous in the sense that it does not decrease the optimal recovery error.

§ 6. Optimal recovery of functions from noisy Fourier transform

Let S be the Schwartz space of rapidly decreasing C^∞-functions on $\mathbb{R}$, S' be the corresponding space of distributions, and let $F\colon S'\to S'$ be the Fourier transform. We let $\mathscr{F}_p$ denote the space of distributions from S' for which

$$\|x(\,\cdot\,)\|_p=\begin{cases}\biggl(\displaystyle\int_{\mathbb{R}}|Fx(\xi)|^p\,d\xi\biggr)^{\frac1p}<\infty, & 1\leqslant p<\infty,\\ \operatorname*{vrai\,sup}\limits_{\xi\in\mathbb{R}}|Fx(\xi)|, & p=\infty.\end{cases}$$

We set

$$\mathscr{F}_p^n=\bigl\{x(\,\cdot\,)\in S' : \|x^{(n)}(\,\cdot\,)\|_p<\infty\bigr\},\qquad F_p^n=\bigl\{x\in\mathscr{F}_p^n : \|x^{(n)}(\,\cdot\,)\|_p\leqslant 1\bigr\}.$$

Assume that the Fourier transform of a function $x(\,\cdot\,)\in F_r^n\cap\mathscr{F}_p$ is known on the interval $\Delta_\sigma=(-\sigma,\sigma)$, $0<\sigma\leqslant\infty$, to within $\delta>0$ in the metric of $L_p(\Delta_\sigma)$. In other words, one knows a function $y(\,\cdot\,)\in L_p(\Delta_\sigma)$ such that $\|Fx(\,\cdot\,)-y(\,\cdot\,)\|_{L_p(\Delta_\sigma)}\leqslant\delta$. How one should best use this information to recover the kth derivative of the function in the metric $\mathscr{F}_q$, $0\leqslant k<n$? By recovery methods here we mean all possible mappings $m\colon L_p(\Delta_\sigma)\to\mathscr{F}_q$. The error of a method is, by definition, the quantity

$$e_{pqr}(m)=\sup_{\substack{x(\,\cdot\,)\in F_r^n\cap\mathscr{F}_p,\ y(\,\cdot\,)\in L_p(\Delta_\sigma)\\ \|Fx(\,\cdot\,)-y(\,\cdot\,)\|_{L_p(\Delta_\sigma)}\leqslant\delta}}\|x^{(k)}(\,\cdot\,)-m(y)(\,\cdot\,)\|_q.$$

The optimal recovery error is defined as follows:

$$E_{pqr}=\inf_{m\colon L_p(\Delta_\sigma)\to\mathscr{F}_q}e_{pqr}(m).$$

A method on which this lower bound is attained is called optimal.

It is readily checked that this problem is a particular case of the above general problem with $T=\mathbb{R}$, $T_0=\Delta_\sigma$, $\psi(\xi)=(i\xi)^k$, $\varphi(\xi)=(i\xi)^n$. We now proceed to apply the results obtained above to the problem in question.

We start with the case $1\leqslant r=q<p<\infty$. Let

$$B=B\biggl(\frac{k+1/q-1/p}{(n-k)(1-q/p)},\frac{2-q/p}{1-q/p}\biggr),$$

where $B(\,\cdot\,,\cdot\,)$ is the Euler B-function and

$$\widehat{\sigma}_\delta=\biggl(\biggl(\frac{q}{2B}\biggr)^{\frac1q-\frac1p}\frac{(n-k)^{\frac2q-\frac1p}}{\delta(k+1/q-1/p)^{\frac1q}}\biggr)^{\frac{1}{n+1/q-1/p}}.$$

Theorem 5. *Let $k,n\in\mathbb{Z}$, $0\leqslant k<n$, $1\leqslant q<p<\infty$, $\delta>0$ and $\sigma\geqslant\widehat{\sigma}_\delta$. Then*

$$E_{pqq}=\biggl(\frac{n+1/q-1/p}{k+1/q-1/p}\biggr)^{\frac1q}\widehat{\sigma}_\delta^{\,-(n-k)}.$$

Moreover, the method $\widehat{m}(y)(\,\cdot\,)=F^{-1}Y_y(\,\cdot\,)$ with

$$Y_y(\xi)=\begin{cases}(i\xi)^k\biggl(1-\biggl(\dfrac{|\xi|}{\widehat{\sigma}_\delta}\biggr)^{q(n-k)}\biggr)y(\xi), & |\xi|<\widehat{\sigma}_\delta,\\ 0, & |\xi|\geqslant\widehat{\sigma}_\delta\end{cases}$$

is optimal.

Proof. Consider equation (2.1), in which we put $\widehat{\lambda}_2 = s^{-q(n-k)}$, where $s \leqslant \sigma$ is be to specified later. We have

$$\left(\int_{-s}^{s}\left(|\xi|^{kq} - \frac{|\xi|^{nq}}{s^{q(n-k)}}\right)^{\frac{p}{p-q}} d\xi\right)^{\frac{1}{p}} = \delta\left(\int_{-s}^{s}|\xi|^{nq}\left(|\xi|^{kq} - \frac{|\xi|^{nq}}{s^{q(n-k)}}\right)^{\frac{q}{p-q}} d\xi\right)^{\frac{1}{q}}.$$

Changing $\xi = su$, this gives

$$2^{\frac{1}{p}} s^{\frac{1}{p}}\left(\int_0^1 u^{\frac{kpq}{p-q}}(1-u^{q(n-k)})^{\frac{p}{p-q}}\,du\right)^{\frac{1}{p}}$$
$$= 2^{\frac{1}{q}} s^{\frac{1}{q}+n}\left(\int_0^1 u^{nq+\frac{kq^2}{p-q}}(1-u^{q(n-k)})^{\frac{q}{p-q}}\,du\right)^{\frac{1}{q}}.$$

To specialize to B-functions, we write $t = u^{(n-k)q}$. As a result,

$$\left(\frac{2s}{q(n-k)}\right)^{\frac{1}{p}} B^{\frac{1}{p}}\left(\frac{k+1/q-1/p}{(n-k)(1-q/p)}, \frac{2-q/p}{1-q/p}\right)$$
$$= \delta s^n \left(\frac{2s}{q(n-k)}\right)^{\frac{1}{q}} B^{\frac{1}{q}}\left(\frac{k+1/q-1/p}{(n-k)(1-q/p)}+1, \frac{2-q/p}{1-q/p}-1\right).$$

Using the well-known equality

$$B(a+1,b-1) = \frac{a}{b-1}B(a,b),$$

we find that $s = \widehat{\sigma}_\delta$. The proof is completed by application of Theorem 1.

Note that the case with $\sigma < \widehat{\sigma}_\delta$ requires more sophisticated analysis. For $q = 2$ the corresponding analysis was given in the paper [20], which also puts forward the above theorem with $q = 2$.

Taking into account the remark at the end of § 5, it follows from Theorem 5 that if a function is such that

$$\int_{\mathbb{R}} |x(\xi)|^p\,d\xi \leqslant \delta^p, \qquad \int_{\mathbb{R}} |\xi|^{qn}|x(\xi)|^q\,d\xi \leqslant 1,$$

then it satisfies the sharp inequality

$$\int_{\mathbb{R}} |\xi|^{kq}\,|x(\xi)|^q\,d\xi \leqslant C\delta^{\frac{(n-k)q}{n+1/q-1/p}},$$

where $0 \leqslant k < n$, $1 \leqslant q < p < \infty$, and

$$C = \frac{n+1/q-1/p}{k+1/q-1/p}\left(\left(\frac{2B}{q}\right)^{\frac{1}{q}-\frac{1}{p}}\frac{(k+1/q-1/p)^{\frac{1}{q}}}{(n-k)^{\frac{2}{q}-\frac{1}{p}}}\right)^{\frac{k+1/q-1/p}{n+1/q-1/p}}.$$

From this we get the following sharp inequality

$$\|\xi^k x(\xi)\|_{L_q(\mathbb{R})} \leqslant C\|x(\xi)\|_{L_p(\mathbb{R})}^{\frac{n-k}{n+1/q-1/p}}\ \|\xi^n x(\xi)\|_{L_q(\mathbb{R})}^{\frac{(n-k)q}{n+1/q-1/p}}. \tag{6.1}$$

Considering even functions $x(\,\cdot\,)$, we have the inequality

$$\|\xi^k x(\xi)\|_{L_q(\mathbb{R}_+)} \leqslant C_1\|x(\xi)\|_{L_p(\mathbb{R}_+)}^{\frac{n-k}{n+1/q-1/p}}\ \|\xi^n x(\xi)\|_{L_q(\mathbb{R}_+)}^{\frac{(n-k)q}{n+1/q-1/p}}, \tag{6.2}$$

where

$$C_1 = \frac{n+1/q-1/p}{k+1/q-1/p}\left(\left(\frac{B}{q}\right)^{\frac{1}{q}-\frac{1}{p}}\frac{(k+1/q-1/p)^{\frac{1}{q}}}{(n-k)^{\frac{2}{q}-\frac{1}{p}}}\right)^{\frac{k+1/q-1/p}{n+1/q-1/p}},$$

which is a particular case of Carlson–Bellman–Levin's inequality [22] (this inequality was carried over to the multidimensional case in [23]).

Now let $1 \leqslant p = q < r < \infty$, $1 \leqslant k < n$ and $\sigma = +\infty$. We set

$$\widehat{s}_\delta = \left(\left(\frac{2B_1}{p}\right)^{\frac{1}{p}-\frac{1}{r}}\frac{(n-k-1/p+1/r)^{\frac{1}{p}}}{\delta k^{\frac{2}{p}-\frac{1}{r}}}\right)^{\frac{1}{n-1/p+1/r}},$$

where

$$B_1 = B\left(\frac{n-k-1/p+1/r}{k(1-p/r)}, \frac{2-p/r}{1-p/r}\right).$$

Theorem 6. *Let $k, n \in \mathbb{N}$, $1 \leqslant k < n$, $1 \leqslant p < r < \infty$, $\sigma = +\infty$ and $\delta > 0$. Then*

$$E_{ppr} = \left(\frac{n-1/p+1/r}{n-k-1/p+1/r}\right)^{\frac{1}{p}}\widehat{s}_\delta^k\delta.$$

Moreover, the method $\widehat{m}(y)(\,\cdot\,) = F^{-1}Y_y(\,\cdot\,)$ with

$$Y_y(\xi) = \begin{cases}(i\xi)^k y(\xi), & |\xi| < \widehat{s}_\delta,\\ (i\xi)^k\left(\dfrac{\widehat{s}_\delta}{|\xi|}\right)^{kp} y(\xi), & |\xi| \geqslant \widehat{s}_\delta.\end{cases}$$

is optimal.

Proof. Consider equation (3.1), in which we put $\widehat{\lambda}_1 = s^{kp}$, where s will be specified later. We have

$$\left(\int_{|\xi|\geqslant s}|\xi|^{\frac{npr}{p-r}}(|\xi|^{kp}-s^{kp})^{\frac{p}{r-p}}\,d\xi\right)^{\frac{1}{p}} = \delta\left(\int_{|\xi|\geqslant s}|\xi|^{\frac{npr}{p-r}}(|\xi|^{kp}-s^{kp})^{\frac{r}{r-p}}\,d\xi\right)^{\frac{1}{r}}.$$

Changing $\xi = s/u$, we arrive at the equation

$$2^{\frac{1}{p}}s^{\frac{1}{p}}\left(\int_0^1 u^{\frac{pr(n-k)}{r-p}+kp-2}(1-u^{kp})^{\frac{p}{r-p}}\,du\right)^{\frac{1}{p}}$$
$$= 2^{\frac{1}{r}}s^{\frac{1}{r}+n}\left(\int_0^1 u^{\frac{pr(n-k)}{r-p}-2}(1-u^{kp})^{\frac{r}{r-p}}\,du\right)^{\frac{1}{r}}.$$

To specialize to B-functions, we write $t = u^{kp}$. As a result, we have

$$\left(\frac{2s}{kp}\right)^{\frac{1}{p}}B^{\frac{1}{p}}\left(\frac{n-k-1/p+1/r}{k(1-p/r)}+1, \frac{2-p/r}{1-p/r}-1\right)$$
$$= \delta s^n\left(\frac{2s}{kp}\right)^{\frac{1}{r}}B^{\frac{1}{r}}\left(\frac{n-k-1/p+1/r}{k(1-p/r)}, \frac{2-p/r}{1-p/r}\right).$$

Consequently $s = \widehat{s}_\delta$. It remains to invoke Theorem 2.

Now we consider the case $1 \leqslant q < p = r < \infty$.

Theorem 7. *Let $k,n\in\mathbb{Z}$, $0\leqslant k<n$, $1\leqslant q<p<\infty$, $\sigma<+\infty$, $\delta>0$ and let $\widehat{a}$ be such that*

$$\int_0^\sigma(1-\delta^p\xi^{np})\left(\frac{\xi^{kq}}{\widehat{a}+\xi^{np}}\right)^{\frac{p}{p-q}}d\xi=\frac{\delta^p(1/q-1/p)}{\sigma^{\frac{n-k-1/q+1/p}{1/q-1/p}}(n-k-1/q+1/p)}.$$

Then

$$E_{pqp}=\left(2\int_0^\sigma\frac{\xi^{\frac{kpq}{p-q}}}{(1+\widehat{a}\xi^{np})^{\frac{p}{p-q}}}\,dt\right)^{\frac{1}{q}-\frac{1}{p}}(\widehat{a}+\delta^{-p})^{\frac{1}{q}}\delta.$$

Moreover, the method $\widehat{m}(y)(\,\cdot\,)=F^{-1}Y_y(\,\cdot\,)$ with

$$Y_y(\xi)=\begin{cases}\dfrac{\widehat{a}(i\xi)^k}{\widehat{a}+|\xi|^{np}}y(\xi), & |\xi|<\sigma,\\ 0, & |\xi|\geqslant\sigma\end{cases}$$

is optimal.

Proof. In our setting equation (4.1) reads as

$$\int_0^\sigma\left(\frac{\xi^{kq}}{1+\widehat{\lambda}_2\xi^{np}}\right)^{\frac{p}{p-q}}d\xi=\delta^p\int_0^\sigma\xi^{np}\left(\frac{\xi^{kq}}{1+\widehat{\lambda}_2\xi^{np}}\right)^{\frac{p}{p-q}}d\xi$$
$$+\frac{\delta^p(1/q-1/p)}{\widehat{\lambda}_2^{\frac{p}{p-q}}\sigma^{\frac{n-k-1/q+1/p}{1/q-1/p}}(n-k-1/q+1/p)}.$$

Letting $a=\widehat{\lambda}_2^{-1}$, we obtain the equation $f(a)=0$, where

$$f(a)=\int_0^\sigma(1-\delta^p\xi^{np})\left(\frac{\xi^{kq}}{a+\xi^{np}}\right)^{\frac{p}{p-q}}d\xi-\frac{\delta^p(1/q-1/p)}{\sigma^{\frac{n-k-1/q+1/p}{1/q-1/p}}(n-k-1/q+1/p)}.$$

It is readily checked that $f(a)\to+\infty$ as $a\to0$ and

$$f(a)\to-\frac{\delta^p(1/q-1/p)}{\sigma^{\frac{n-k-1/q+1/p}{1/q-1/p}}(n-k-1/q+1/p)}$$

as $a\to+\infty$. The function $f(\,\cdot\,)$ is continuous, and hence there exists $\widehat{a}>0$ at which $f(\widehat{a})=0$. It remains to apply Theorem 3.

For $\sigma=+\infty$ equation (4.1) may be solved explicitly. In this case we have the following result.

Theorem 8. *Let $k,n\in\mathbb{Z}$, $0\leqslant k<n$, $1\leqslant q<p<\infty$, $\delta>0$ and $\sigma=+\infty$. Then*

$$E_{pqp}=\left(\frac{n}{n-K}\right)^{\frac{1}{q}}\left(\frac{2B_2}{np}\right)^{\frac{1}{q}-\frac{1}{p}}\left(\frac{n-K}{K}\right)^{\frac{K}{np}}\delta^{1-\frac{K}{n}},$$

where

$$K=k+\frac{1}{q}-\frac{1}{p},\qquad B_2=B\left(\frac{\frac{K}{n}q/p}{1-q/p},\frac{1-\frac{K}{n}q/p}{1-q/p}\right).$$

Moreover, the method $\widehat{m}(y)(\,\cdot\,)=F^{-1}Y_y(\,\cdot\,)$ with

$$Y_y(\xi)=\frac{(i\xi)^k}{1+\frac{K}{n-K}\delta^p|\xi|^{np}}\,y(\xi)$$

is optimal.

We finally consider the case $1 < p = q = r < \infty$. If $\sigma = +\infty$ and $\delta = 0$, then the recovery problem becomes vacuous, because all the information about the function is available, so we exclude this case hereafter. Given $k \geqslant 1$, we set

$$\widehat{\sigma} = \begin{cases} \left(\dfrac{n}{k}\right)^{\frac{1}{p(n-k)}} \delta^{-\frac{1}{n}}, & \delta > 0, \\ +\infty, & \delta = 0, \end{cases}$$

$$\lambda_1 = \begin{cases} \dfrac{n-k}{n}\delta^{-\frac{pk}{n}}, & \sigma \geqslant \widehat{\sigma}, \\ \sigma^{kp}\left(\dfrac{k}{n}\right)^{\frac{k}{n-k}} \dfrac{n-k}{n}, & \sigma < \widehat{\sigma}, \end{cases} \qquad \lambda_2 = \begin{cases} \dfrac{k}{n}\delta^{\frac{p(n-k)}{n}}, & \sigma \geqslant \widehat{\sigma}, \\ \sigma^{-p(n-k)}, & \sigma < \widehat{\sigma}. \end{cases}$$

Next, given $k = 0$, we take $\lambda_1 = 1$,

$$\lambda_2 = \begin{cases} \sigma^{-pn}, & \sigma < +\infty, \\ 0, & \sigma = +\infty. \end{cases}$$

Theorem 9. *Let $k, n \in \mathbb{Z}$, $0 \leqslant k < n$ and $1 \leqslant p < \infty$. Then*

$$E_{ppp} = (\lambda_1 \delta^p + \lambda_2)^{\frac{1}{p}}.$$

Moreover, the methods $\widehat{m}(y)(\,\cdot\,) = F^{-1}Y_y(\,\cdot\,)$ are optimal, where

$$Y_y(\xi) = \begin{cases} \alpha(\xi)(i\xi)^k y(\xi), & |\xi| < \sigma, \\ 0, & |\xi| \geqslant \sigma, \end{cases}$$

and, for $\delta > 0$, $\lambda_2 > 0$, the function $\alpha(\,\cdot\,)$ satisfies the condition,

$$\begin{cases} \dfrac{|1-\alpha(\xi)|^{p'}}{\lambda_2^{\frac{p'}{p}}|\xi|^{p'(n-k)}} + \dfrac{|\xi^k \alpha(\xi)|^{p'}}{\lambda_1^{\frac{p'}{p}}} \leqslant 1, \quad 1 < p < \infty, \quad \dfrac{1}{p} + \dfrac{1}{p'} = 1, \\ \dfrac{|1-\alpha(\xi)|}{\lambda_2|\xi|^{n-k}} \leqslant 1, \quad \dfrac{|\xi^k\alpha(\xi)|}{\lambda_1} \leqslant 1, \quad p = 1 \end{cases} \tag{6.3}$$

for almost $\xi \in \Delta_\sigma$, and for $1 \leqslant p < \infty$, $\delta = 0$, the condition

$$|1-\alpha(\xi)| \leqslant \lambda_2^{\frac{1}{p}}|\xi|^{n-k}.$$

For $\lambda_2 = 0$ the method $\widehat{m}(y)(\,\cdot\,) = F^{-1}y(\,\cdot\,)$ is optimal.

Proof. In our setting the extremal problem (5.1) reads as

$$\begin{gathered} \int_{\mathbb{R}} |\xi|^{kp}\,|Fx(\xi)|^p\,d\xi \to \max, \\ \int_{\Delta_\sigma} |Fx(\xi)|^p\,d\xi \leqslant \delta^p, \qquad \int_{\mathbb{R}} |\xi|^{np}\,|Fx(\xi)|^p\,d\xi \leqslant 1. \end{gathered} \tag{6.4}$$

In the (u,v)-plane consider the curve $v = u^{k/n}$ is defined parametrically by

$$\begin{cases} u = |\xi|^{np}, \\ v = |\xi|^{kp}, \end{cases} \qquad \xi \in \mathbb{R}. \tag{6.5}$$

Let $k \geqslant 1$, $\delta > 0$. The line $v = \lambda_1 + \lambda_2 u$, where

$$\lambda_1 = \frac{n-k}{n}\delta^{-\frac{pk}{n}}, \qquad \lambda_2 = \frac{k}{n}\delta^{\frac{p(n-k)}{n}},$$

is tangent to this curve at $u = \delta^{-p}$. The curve (6.5) is concave, and hence, for all $\xi \in \mathbb{R}$,

$$-|\xi|^{kp} + \lambda_1 + \lambda_2 |\xi|^{np} \geqslant 0.$$

If $\sigma \geqslant \widehat{\sigma}$, then, for all $\xi \in \mathbb{R}$,

$$-|\xi|^{kp} + \lambda_1 \chi_\sigma(\xi) + \lambda_2 |\xi|^{np} \geqslant 0, \tag{6.6}$$

where $\chi_\sigma(\,\cdot\,)$ is the characteristic function of Δ_σ.

For sufficiently small $\varepsilon > 0$, we consider the function $x_\varepsilon(\,\cdot\,)$ such that

$$Fx_\varepsilon(\xi) = \begin{cases} \dfrac{\delta}{\varepsilon^{1/p}}, & \xi \in (\delta^{-\frac{1}{n}} - \varepsilon, \delta^{-\frac{1}{n}}), \\ 0, & \xi \notin (\delta^{-\frac{1}{n}} - \varepsilon, \delta^{-\frac{1}{n}}). \end{cases}$$

Hence,

$$\int_{\Delta_\sigma} |Fx_\varepsilon(\xi)|^p \, d\xi = \delta^p,$$

and further,

$$\int_{\mathbb{R}} |\xi|^{np} \, |Fx_\varepsilon(\xi)|^p \, d\xi = \frac{\delta^p}{\varepsilon} \int_{\delta^{-\frac{1}{n}}-\varepsilon}^{\delta^{-\frac{1}{n}}} \xi^{np} \, d\xi \leqslant 1.$$

So, the function $x_\varepsilon(\,\cdot\,)$ is admissible in (6.4), and therefore,

$$E_{ppp}^p \geqslant \int_{\mathbb{R}} |\xi|^{kp} \, |Fx_\varepsilon(\xi)|^p \, d\xi = \frac{\delta^p}{\varepsilon} \int_{\delta^{-\frac{1}{n}}-\varepsilon}^{\delta^{-\frac{1}{n}}} \xi^{kp} \, d\xi.$$

Making $\varepsilon \to 0$, this gives

$$E_{ppp} \geqslant \delta^{1-\frac{k}{n}} = (\lambda_1 \delta^p + \lambda_2)^{\frac{1}{p}}.$$

For $k \geqslant 1$, $\delta \geqslant 0$ and $\sigma < \widehat{\sigma}$, the tangent to curve (6.5) at the point

$$u = \sigma^{np} \left(\frac{k}{n}\right)^{\frac{n}{n-k}}$$

is given by $v = \lambda_1 + \lambda_2 u$, where

$$\lambda_1 = \sigma^{kp} \left(\frac{k}{n}\right)^{\frac{k}{n-k}} \frac{n-k}{n}, \qquad \lambda_2 = \sigma^{-p(n-k)}.$$

Since curve (6.5) is concave and since $\lambda_2 |\xi|^{pn} > |\xi|^{pk}$ for $|\xi| > \sigma$, inequality (6.6) is satisfied (with new λ_1 and λ_2). For sufficiently small $\varepsilon_1, \varepsilon_2 > 0$, consider the function $x_{\varepsilon_1,\varepsilon_2}(\,\cdot\,)$ such that

$$Fx_{\varepsilon_1,\varepsilon_2}(\xi) = \begin{cases} \dfrac{\delta}{\varepsilon_1^{1/p}}, & \xi \in (\xi_0, \xi_0 + \varepsilon_1), \\ c^{\frac{1}{p}}, & \xi \in (\sigma, \sigma + \varepsilon_2), \\ 0, & \xi \notin (\xi_0, \xi_0 + \varepsilon_1) \cup (\sigma, \sigma + \varepsilon_2), \end{cases}$$

where

$$\xi_0 = \sigma\left(\frac{k}{n}\right)^{\frac{1}{p(n-k)}},$$

and c will be specified later. We have

$$\int_{\Delta_\sigma} |Fx_{\varepsilon_1,\varepsilon_2}(\xi)|^p \, d\xi = \delta^p,$$

and moreover,

$$\int_{\mathbb{R}} |\xi|^{np} \, |Fx_{\varepsilon_1,\varepsilon_2}(\xi)|^p \, d\xi = \frac{\delta^p}{\varepsilon_1}\int_{\xi_0}^{\xi_0+\varepsilon_1} \xi^{np} \, d\xi + c\int_\sigma^{\sigma+\varepsilon_2} \xi^{np} \, d\xi.$$

For $\varepsilon_1 \to 0$ and $\delta > 0$

$$\frac{\delta^p}{\varepsilon_1}\int_{\xi_0}^{\xi_0+\varepsilon_1} \xi^{np} \, d\xi \to \delta^p \xi_0^{np} = \delta^p \sigma^{np}\left(\frac{k}{n}\right)^{\frac{n}{n-k}} < \delta^p \widehat{\sigma}^{np}\left(\frac{k}{n}\right)^{\frac{n}{n-k}} = 1.$$

Hence, setting

$$c = \frac{1 - \dfrac{\delta^p}{\varepsilon_1}\displaystyle\int_{\xi_0}^{\xi_0+\varepsilon_1} \xi^{np} \, d\xi}{\displaystyle\int_\sigma^{\sigma+\varepsilon_2} \xi^{np} \, d\xi}$$

for $\delta \geqslant 0$, this gives

$$\int_{\mathbb{R}} |\xi|^{np} \, |Fx_{\varepsilon_1,\varepsilon_2}(\xi)|^p \, d\xi = 1.$$

As a result, the function $x_{\varepsilon_1,\varepsilon_2}(\,\cdot\,)$ is admissible in (6.4), and hence,

$$\begin{aligned} E_{ppp}^p &\geqslant \int_{\mathbb{R}} |\xi|^{kp} \, |Fx_{\varepsilon_1,\varepsilon_2}(\xi)|^p \, d\xi \\ &= \frac{\delta^p}{\varepsilon_1}\int_{\xi_0}^{\xi_0+\varepsilon_1} \xi^{kp} \, d\xi + \left(1 - \frac{\delta^p}{\varepsilon_1}\int_{\xi_0}^{\xi_0+\varepsilon_1} \xi^{np} \, d\xi\right) \frac{\displaystyle\int_\sigma^{\sigma+\varepsilon_2} \xi^{kp} \, d\xi}{\displaystyle\int_\sigma^{\sigma+\varepsilon_2} \xi^{np} \, d\xi}. \end{aligned}$$

Making ε_1 and ε_2 tend to zero, this establishes

$$E_{ppp} \geqslant (\lambda_1 \delta^p + \lambda_2)^{\frac{1}{p}}.$$

Let $k = 0$ and $\sigma < +\infty$. Then $\lambda_1 = 1$, $\lambda_2 = \sigma^{-pn}$. It is readily checked that inequality (6.6) also holds in this case. The same function $x_{\varepsilon_1,\varepsilon_2}$ (with $\xi_0 = 0$) is easily seen to be admissible in problem (6.4). Hence,

$$E_{ppp}^p \geqslant \int_{\mathbb{R}} |Fx_{\varepsilon_1,\varepsilon_2}(\xi)|^p \, d\xi = \delta^p + \left(1 - \frac{\delta^p}{\varepsilon_1}\int_0^{\varepsilon_1} \xi^{np} \, d\xi\right) \frac{\varepsilon_2}{\displaystyle\int_\sigma^{\sigma+\varepsilon_2} \xi^{np} \, d\xi}.$$

Making ε_1 and ε_2 to zero, we see that

$$E_{ppp} \geqslant (\delta^p + \sigma^{-pn})^{\frac{1}{p}} = (\lambda_1 \delta^p + \lambda_2)^{\frac{1}{p}}.$$

It remains to consider the case when $k = 0$ and $\sigma = +\infty$. Now $\lambda_1 = 1$, $\lambda_2 = 0$. It is clear that (6.6) holds. For sufficiently small $\varepsilon > 0$, consider the function $x_\varepsilon(\,\cdot\,)$ such that

$$Fx_\varepsilon(\xi) = \begin{cases} \dfrac{\delta}{\varepsilon^{1/p}}, & \xi \in (0, \varepsilon), \\ 0, & \xi \notin (0, \varepsilon). \end{cases}$$

Hence,

$$\int_{\mathbb{R}} |Fx_\varepsilon(\xi)|^p \, d\xi = \delta^p,$$

and

$$\int_{\mathbb{R}} |\xi|^{np} \, |Fx_\varepsilon(\xi)|^p \, d\xi = \frac{\delta^p}{\varepsilon} \int_0^\varepsilon \xi^{np} \, d\xi \to 0$$

as $\varepsilon \to 0$. This shows that, for sufficiently small $\varepsilon > 0$, the function $x_\varepsilon(\,\cdot\,)$ is admissible in problem (6.4). Hence,

$$E_{ppp}^p \geqslant \int_{\mathbb{R}} |Fx_\varepsilon(\xi)|^p \, d\xi = \delta^p = \lambda_1 \delta^p + \lambda_2.$$

Now the assertion of Theorem 9 follows from Theorem 4.

From Theorems 6, 8 and 9 one may get sharp Carlson-type inequalities as in (6.1) and (6.2).

Similarly to Corollary 1 we have the following result.

Corollary 2. *For $k \geqslant 1$, $\delta > 0$ the methods $\widehat{m}(y)(\,\cdot\,) = F^{-1}Y_y(\,\cdot\,)$ are optimal, where*

$$Y_y(\xi) = \begin{cases} (i\xi)^k y(\xi), & |\xi| \leqslant \theta\sigma_0, \\ \alpha(\xi)(i\xi)^k y(\xi), & \theta\sigma_0 < |\xi| < \sigma_0, \\ 0, & |\xi| \geqslant \sigma_0, \end{cases}$$

$$\theta = \left(\frac{n-k}{n}\right)^{\frac{1}{kp}} \left(\frac{k}{n}\right)^{\frac{1}{p(n-k)}}, \qquad \sigma_0 = \min\{\sigma, \widehat{\sigma}\},$$

and $\alpha(\,\cdot\,)$ satisfies condition (6.3). *For $\delta = 0$ or $k = 0$ the method $\widehat{m}(y)(\,\cdot\,) = F^{-1}Y_y(\,\cdot\,)$ with*

$$Y_y(\xi) = \begin{cases} (i\xi)^k y(\xi), & |\xi| \leqslant \sigma, \\ 0, & |\xi| \geqslant \sigma \end{cases}$$

is optimal.

For $p = 2$ the corresponding recovery problem of the derivative on the Sobolev class $W_2^n = F_2^n$ was examined in [16] (see also [24]).

§ 7. The discrete case

If $T = \mathbb{N}$ and $\mu(\{j\}) = 1$, then the corresponding space $L_p(T, \mu)$, $1 \leqslant p < \infty$, agrees with the space l_p, which consists of the vectors $x = (x_1, x_2, \ldots)$ such that

$$\|x\|_{l_p} = \left(\sum_{j=1}^\infty |x_j|^p\right)^{\frac{1}{p}} < \infty.$$

We set

$$\mathscr{W}_p = \left\{ x \in l_p : \sum_{j=1}^{\infty} |\nu_j x_j|^p < \infty \right\}, \qquad W_p = \left\{ x \in \mathscr{W}_p : \sum_{j=1}^{\infty} |\nu_j x_j|^p \leqslant 1 \right\}.$$

Assume we are given an operator $\Lambda\colon \mathscr{W}_p \to l_p$,

$$\Lambda x = (\mu_1 x_1, \mu_2 x_2, \dots),$$

where the sequence $|\mu_j|/|\nu_j|$ is bounded for sufficiently large j (this condition implies that $\Lambda x \in l_p$ for all $x \in \mathscr{W}_p$).

Let us consider the problem of optimal recovery of the values of an operator Λ on the set W_p from noisy information on the coordinates $x_1, \dots, x_N$. More precisely, it is assumed that, for any $x \in W_p$, one knows the vector $y = (y_1, \dots, y_N)$ such that $\|I^N x - y\|_{l_p^N} \leqslant \delta$, where $I^N x = (x_1, \dots, x_N)$ and

$$\|x\|_{l_p^N} = \left(\sum_{j=1}^{N} |x_j|^p \right)^{\frac{1}{p}}.$$

The problem is to recover, with the best possible accuracy, the value of Λx from the given vector y.

Here, by recovery methods we mean all possible mappings $m\colon l_p^N \to l_p$. In accordance with the general formulation of the problem, the error of a method is defined by

$$e_p(m) = \sup_{\substack{x \in W_p,\ y \in l_p^N \\ \|I^N x - y\|_{l_p^N} \leqslant \delta}} \|\Lambda x - m(y)\|_{l_p}.$$

The quantity

$$E_p = \inf_{m\colon l_p^N \to l_p} e_p(m)$$

is the optimal recovery error, and a method on which this lower bound is attained is called optimal.

We shall assume that $\nu_j \neq 0$ for all $j \geqslant N+1$. We set

$$\lambda = \sup_{j \geqslant N+1} \frac{|\mu_j|^p}{|\nu_j|^p},$$
$$M = \operatorname{co}\bigl\{(0,0) \cup \{(|\nu_j|^p, |\mu_j|^p)\}_{j \in \mathbb{N}}\bigr\} + \{(t, t\lambda) \mid t \geqslant 0\},$$

where $\operatorname{co}\Omega$ is the convex hull of a set Ω. Consider the function $\theta(\,\cdot\,)$ on $[0,\infty)$ defined by $\theta(t) = \max\{x \mid (t,x) \in M\}$. Clearly, $\theta(\,\cdot\,)$ is a concave broken line (see Fig. 1).

Рис. 1

Theorem 10. *For all $\delta > 0$*

$$E_p = \delta\theta^{\frac{1}{p}}(\delta^{-p}).$$

Let $\delta > 0$ and let δ^{-p} lies in the interval of $\mathbb{R}_+$ on which $\theta(\,\cdot\,)$ is given by the equation $\theta(t) = \lambda_1 + \lambda_2 t$. If $\lambda_1, \lambda_2 > 0$, then, for all α_j, $1 \leqslant j \leqslant N$, satisfying the condition

$$\begin{cases} |\mu_j|^{p'}\left(\dfrac{|1-\alpha_j|^{p'}}{|\nu_j|^{p'}\lambda_2^{\frac{p'}{p}}} + \dfrac{|\alpha_j|^{p'}}{\lambda_1^{\frac{p'}{p}}}\right) \leqslant 1, \quad 1 < p < \infty, \ \dfrac{1}{p} + \dfrac{1}{p'} = 1, \\ \dfrac{|\mu_j(1-\alpha_j)|}{|\nu_j|\lambda_2} \leqslant 1, \quad \dfrac{|\mu_j\alpha_j|}{\lambda_1} \leqslant 1, \quad p = 1, \end{cases} \tag{7.1}$$

the methods

$$\widehat{m}(y) = (\alpha_1\mu_1 y_1, \ldots, \alpha_N\mu_N y_N, 0, \ldots) \tag{7.2}$$

are optimal.

If $\lambda_1 = 0$, then $\widehat{m}(y) = 0$ is an optimal method, and if $\lambda_2 = 0$, then $\widehat{m}(y) = (\mu_1 y_1, \ldots, \mu_N y_N, 0, \ldots)$ is an optimal method. If $\delta = 0$, then $E_p = \lambda^{\frac{1}{p}}$ and all methods (7.2), in which

$$|\mu_j(1-\alpha_j)| \leqslant |\nu_j|\lambda^{\frac{1}{p}},$$

are optimal.

Proof. In this setting the extremal problem (5.1) is as follows:

$$\sum_{j=1}^{\infty}|\mu_j x_j|^p \to \max, \qquad \sum_{j=1}^{N}|x_j|^p \leqslant \delta^p, \quad \sum_{j=1}^{\infty}|\nu_j x_j|^p \leqslant 1. \tag{7.3}$$

According to the definition of the set M and the function $\theta(\,\cdot\,)$, if on some interval the broken line $\theta(\,\cdot\,)$ is given by the equation $\theta(t) = \lambda_1 + \lambda_2 t$, then $\lambda_1 \geqslant 0$, $\lambda_2 \geqslant \lambda$, and moreover,

$$-|\mu_j|^p + \lambda_1\chi_j + \lambda_2|\nu_j|^p \geqslant 0$$

for all $j \in \mathbb{N}$, where

$$\chi_j = \begin{cases} 1, & 1 \leqslant j \leqslant N, \\ 0, & j > N. \end{cases}$$

Let $0 < |\nu_{s_1}|^p < \cdots < |\nu_{s_k}|^p$ be the arguments of the bend points of the broken line $\theta(\,\cdot\,)$. Assume that $|\nu_{s_{l-1}}| \leqslant \delta^{-1} < |\nu_{s_l}|$ for some l, $1 < sl \leqslant k$, and that the function $\theta(\,\cdot\,)$ for $|\nu_{s_{l-1}}|^p \leqslant t \leqslant |\nu_{s_l}|^p$ is as follows: $\theta(t) = \lambda_1 + \lambda_2 t$. We set $\widehat{x}_j = 0$, $j \neq s_{l-1}, s_l$,

$$\widehat{x}_{s_{l-1}} = \left(\frac{\delta^p|\nu_{s_l}|^p - 1}{|\nu_{s_l}|^p - |\nu_{s_{l-1}}|^p}\right)^{\frac{1}{p}}, \qquad \widehat{x}_{s_l} = \left(\frac{1 - \delta^p|\nu_{s_{l-1}}|^p}{|\nu_{s_l}|^p - |\nu_{s_{l-1}}|^p}\right)^{\frac{1}{p}}.$$

Since

$$|\widehat{x}_{s_{l-1}}|^p + |\widehat{x}_{s_l}|^p = \delta^p, \qquad |\nu_{s_{l-1}}\widehat{x}_{s_{l-1}}|^p + |\nu_{s_l}\widehat{x}_{s_l}|^p = 1,$$

$\widehat{x} = (\widehat{x}_1, \widehat{x}_2, \ldots)$ is an admissible element in problem (7.3). Consequently, the value of this problem is estimated from below by the quantity

$$|\mu_{s_{l-1}}|^p\,|\widehat{x}_{s_{l-1}}|^p + |\mu_{s_l}|^p\,|\widehat{x}_{s_l}|^p = \lambda_1\delta^p + \lambda_2. \tag{7.4}$$

Assume now that $\delta^{-1} < \nu_{s_1}$. If $\nu_j = 0$ for some $j \in \mathbb{N}$, then, for s_0 satisfying $\theta(0) = |\mu_{s_0}|^p$, $\nu_{s_0} = 0$, equality (7.4) is proved by the arguments similar to those used above in the case $l = 1$; here, λ_1 and λ_2 are such that $\theta(t) = \lambda_1 + \lambda_2 t$ on the closed interval $[0, |\nu_{s_1}|^p]$. If $\nu_j > 0$ for all $j \in \mathbb{N}$, then $\lambda_1 = 0$ and $\theta(t) = \lambda_2 t$ for $t \in [0, |\nu_{s_1}|^p]$. So, we set $\widehat{x}_j = 0$, $j \neq s_1$, $\widehat{x}_{s_1} = 1/\nu_{s_1}$. It is easily checked that $\widehat{x}$ is an admissible element for problem (7.3). Hence, its value is estimated from below by the quantity

$$|\mu_{s_1}|^p \, |\widehat{x}_{s_1}|^p = \lambda_2.$$

Now assume that $\delta \leqslant |\nu_{s_k}|^{-1}$ and that $\theta(t) = \lambda_1 + \lambda_2 t$ on the interval $[|\nu_{s_k}|^p, +\infty)$. Clearly, $\lambda_2 = \lambda$, because $|\nu_{s_k}|^p$ is the last bend of $\theta(\,\cdot\,)$. From the definition of λ it follows that, for any $\varepsilon > 0$, there exists $s_{k+1} \geqslant N + 1$ such that

$$\frac{|\mu_{s_{k+1}}|^p}{|\nu_{s_{k+1}}|^p} > \lambda - \varepsilon. \tag{7.5}$$

We set $\widehat{x}_j = 0$, $j \neq s_k, s_{k+1}$, and choose $\widehat{x}_{s_k}$ and $\widehat{x}_{s_{k+1}}$ so as to have

$$|\widehat{x}_{s_k}|^p = \delta^p, \qquad |\nu_{s_k} \widehat{x}_{s_k}|^p + |\nu_{s_{k+1}} \widehat{x}_{s_{k+1}}|^p = 1.$$

Clearly, $\widehat{x}$ is an admissible element for problem (7.3). Hence, its value is estimated from below by the quantity

$$\begin{aligned} |\mu_{s_k}|^p \, |\widehat{x}_{s_k}|^p + |\mu_{s_{k+1}}|^p \, |\widehat{x}_{s_{k+1}}|^p &= \lambda_1 \delta^p + \lambda_2 - \left(\lambda - \frac{|\mu_{s_{k+1}}|^p}{|\nu_{s_{k+1}}|^p} \right) (1 - \delta^p |\nu_{s_k}|^p) \\ &\geqslant \lambda_1 \delta^p + \lambda_2 - \varepsilon (1 - \delta^p |\nu_{s_k}|^p). \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, the value of problem (7.3) is estimated from below by the quantity $\lambda_1 \delta^p + \lambda_2$.

Assume now that $\theta(\,\cdot\,)$ has no bends. In this case, $\theta(\,\cdot\,)$ is the line $\theta(t) = \lambda_1 + \lambda_2 t$ in $\mathbb{R}_+$. Clearly, now $\lambda_2 = \lambda$. To build the vector $\widehat{x}$ we consider two cases. Assume first that there exists s_0 such that $\nu_{s_0} = 0$ and $\theta(0) = |\mu_{s_0}|^p$. We set $\widehat{x}_j = 0$, $j \neq s_0, s_{k+1}$, and choose $\widehat{x}_{s_0}$ and $\widehat{x}_{s_{k+1}}$ so as to have

$$|\widehat{x}_{s_0}|^p = \delta^p, \qquad |\nu_{s_{k+1}} \widehat{x}_{s_{k+1}}|^p = 1$$

(here, s_{k+1} is the same as in (7.5)). Consequently, the value of problem (7.3) is estimated from below by the quantity

$$|\mu_{s_0}|^p \, |\widehat{x}_{s_0}|^p + |\mu_{s_{k+1}}|^p \, |\widehat{x}_{s_{k+1}}|^p = \lambda_1 \delta^p + \lambda_2 - \left(\lambda - \frac{|\mu_{s_{k+1}}|^p}{|\nu_{s_{k+1}}|^p} \right) > \lambda_1 \delta^p + \lambda_2 - \varepsilon.$$

Next, if $\theta(0) = 0$ (in this case, $\lambda_1 = 0$), we set $\widehat{x}_j = 0$, $j \neq s_{k+1}$, $\widehat{x}_{s_{k+1}} = 1/\nu_{s_{k+1}}$. Now the value of the problem is estimated by

$$|\mu_{s_{k+1}}|^p \, |\widehat{x}_{s_{k+1}}|^p = \lambda_2 + \left(\lambda - \frac{|\mu_{s_{k+1}}|^p}{|\nu_{s_{k+1}}|^p} \right) > \lambda_2 - \varepsilon.$$

The required estimate $\lambda_1 \delta^p + \lambda_2$ follows in this and the remaining cases, because $\varepsilon > 0$ is arbitrary.

Finally, application of Theorem 4 completes the proof.

We note that at the bend points $(|\nu_{s_l}|^p, |\mu_{s_l}|^p)$, $l = 1, \dots, k$, of the broken line $\theta(\,\cdot\,)$ there are may tangent lines to the set M. It is easily checked that for $\delta = |\nu_{s_l}|^{-1}$ for each such a tangent line $\lambda_1 + \lambda_2 t$ with $\lambda_1, \lambda_2 > 0$, all methods (7.2) for which α_j, $j = 1, \dots, N$, satisfy condition (7.1), are optimal.

Corollary 3. *Assume that δ^{-p} lies in the interval $\mathbb{R}_+$, on which $\theta(t) = \lambda_1 + \lambda_2 t$ with $\lambda_1, \lambda_2 > 0$. We set*

$$D = \bigl\{j : |\mu_j|^p > |\nu_j|^p \lambda_2,\ 1 \leqslant j \leqslant N\bigr\}, \qquad D_0 = \bigl\{j \in D : |\mu_j|^p \leqslant \lambda_1\bigr\}.$$

Then the methods

$$\{\widehat{m}(y)\}_j = \begin{cases} \mu_j y_j, & j \in D_0, \\ \alpha_j \mu_j y_j, & j \in D \setminus D_0, \\ 0, & \mathbb{N} \setminus D, \end{cases}$$

where α_j satisfies conditions (7.1), are optimal.

From Corollary 3 it follows that the valuable information about a given noisy vector x consists of the coordinates x, whose numbers are contained in the set D. Moreover, for numbers from the set D_0 one may apply a method that acts on the coordinates exactly as the operator being recovered. Figure 2 illustrates the form of the sets D and D_0.

Рис. 2

In the case of recovery of the derivative from noisy Fourier coefficients, the above recovery problem was solved in the Euclidean case in [15], and in a more general setting, in [25], but again in the Euclidean framework (in both cases no family of methods was provided).

We now apply the results of Theorem 10 to the example from [6] (Theorem 12 on p. 45). In this example $\mu_j = 1$, $j \in \mathbb{N}$ (in fact, [6] considers sequences on $\mathbb{Z}$, but this is immaterial). We set

$$B = \min_{1 \leqslant j \leqslant N} |\nu_j|, \qquad A = \inf_{j > N} |\nu_j|, \qquad C = \frac{B}{A}$$

(it may be assumed that $A > 0$, for otherwise $E_p = +\infty$). For the quantity λ introduced above, we have $\lambda = 1/A^p$.

The next result follows from Theorem 10.

Theorem 11. 1. *For $\delta = 0$, all methods (7.2), in which*

$$|1 - \alpha_j| \leqslant \frac{|\nu_j|}{A}, \qquad j = 1, \dots, N,$$

are optimal.

2. *For $C = 0$, $\theta(t) = 1 + t/A^p$ and all methods (7.2), in which*

$$\begin{cases} \dfrac{A^{p'}}{|\nu_j|^{p'}} |1 - \alpha_j|^{p'} + |\alpha_j|^{p'} \leqslant 1, & 1 < p < \infty, \quad \dfrac{1}{p} + \dfrac{1}{p'} = 1, \\ |1 - \alpha_j| \leqslant \dfrac{|\nu_j|}{A}, \quad |\alpha_j| \leqslant 1, & p = 1, \end{cases}$$

are optimal.

3. *For* $0 < C < 1$ *and* $\delta^{-1} \geqslant B$, $\theta(t) = 1 + (t - B^p)/A^p = 1 - C^p + t/A^p$ *and all methods* (7.2), *in which*

$$\begin{cases} \dfrac{A^{p'}}{|\nu_j|^{p'}}|1-\alpha_j|^{p'} + \dfrac{|\alpha_j|^{p'}}{(1-C^p)^{p'/p}} \leqslant 1, & 1 < p < \infty, \quad \dfrac{1}{p} + \dfrac{1}{p'} = 1, \\ |1-\alpha_j| \leqslant \dfrac{|\nu_j|}{A}, \quad |\alpha_j| \leqslant 1 - C, & p = 1, \end{cases} \tag{7.6}$$

are optimal.

4. *For* $C \geqslant 1$, $\theta(t) = t/A^p$, *and for* $0 < C < 1$ *and* $\delta^{-1} \leqslant B$, $\theta(t) = t/B^p$, *and the method* $\widehat{m}(y) = 0$ *is optimal.*

In [6] the following values of the coefficients α_j were found:
1, 2. $\alpha_j = 1$;
3. $\alpha_j = 1 - C^p$;
4. $\alpha_j = 0$.

Let us examine for which j one may put $\alpha_j = 0$ and $\alpha_j = 1$. We set

$$N_1 = \bigl\{j \in \{1, \dots, N\} : |\nu_j| < A\bigr\}.$$

Corollary 4. 1. *For* $\delta = 0$ *or* $C = 0$, *method* (7.2) *with*

$$\alpha_j = \begin{cases} 1, & j \in N_1, \\ 0, & j \notin N_1 \end{cases}$$

is optimal.

2. *For* $0 < C < 1$ *and* $\delta^{-1} \geqslant B$, *all methods* (7.2), *in which* α_j *satisfy condition* (7.6) *for* $j \in N_1$ *and* $\alpha_j = 0$ *for* $j \notin N_1$, *are optimal.*

I am greatly indebted to the referees for a number of valuable remarks, and in particular, for drawing attention to the connection of the problems under study with sharp inequalities of Carlson's type.

Bibliography

[1] С. А. Смоляк, *Об оптимальном восстановлении функций и функционалов от них*, Дисс. … канд. физ.-матем. наук, МГУ, М. 1965.

[2] Н. С. Бахвалов, "Об оптимальности линейных методов приближения операторов на выпуклых классах функций", *Ж. вычисл. матем. и матем. физ.* **11**:4 (1971), 1014–1018; англ. пер.:, N. S. Bakhvalov, "On the optimality of linear methods for operator approximation in convex classes of functions", *U.S.S.R. Comput. Math. Math. Phys.* **11**:4 (1971), 244–249.

[3] К. Ю. Осипенко, "Оптимальная интерполяция аналитических функций", *Матем. заметки* **12**:4 (1972), 465–476; англ. пер.:, K. Yu. Osipenko, "Optimal interpolation of analytic functions", *Math. Notes* **12**:4 (1972), 712–719.

[4] А. Г. Марчук, К. Ю. Осипенко, "Наилучшее приближение функций, заданных с погрешностью в конечном числе точек", *Матем. заметки* **17**:3 (1975), 359–368; англ. пер.:, A. G. Marchuk, K. Yu. Osipenko, "Best approximation of functions specified with an error at a finite number of points", *Math. Notes* **17**:3 (1975), 207–212.

[5] К. Ю. Осипенко, “Наилучшее приближение аналитических функций по информации об их значениях в конечном числе точек”, *Матем. заметки* **19**:1 (1976), 29–40; англ. пер.:, K. Yu. Osipenko, “Best approximation of analytic functions from information about their values at a finite number of points”, *Math. Notes* **19**:1 (1976), 17–23.

[6] C. A. Micchelli, T. J. Rivlin, “A survey of optimal recovery”, *Optimal estimation in approximation theory*, Proc. Internat. Sympos. (Freudenstadt, 1976), Plenum, New York 1977, pp. 1–54.

[7] В. В. Арестов, “Наилучшее восстановление операторов и родственные задачи”, *Сборник трудов Всесоюзной школы по теории функций* (Душанбе, август 1986 г.), Тр. МИАН СССР, **189**, Наука, М. 1989, с. 3–20; англ. пер.:, V. V. Arestov, “Optimal recovery of operators and related problems”, *Proc. Steklov Inst. Math.* **189**:4 (1990), 1–20.

[8] Г. Г. Магарил-Ильяев, К. Ю. Осипенко, “Об оптимальном восстановлении функционалов по неточным данным”, *Матем. заметки* **50**:6 (1991), 85–93; англ. пер.:, G. G. Magaril-Il'yaev, K. Yu. Osipenko, “Optimal recovery of functionals based on inaccurate data”, *Math. Notes* **50**:6 (1991), 1274–1279.

[9] Дж. Фр. Трауб, Х. Вожьняковский, *Общая теория оптимальных алгоритмов*, Мир, М. 1983, 382 с.; пер. с англ.:, J. F. Traub, H. Woźniakowski, *A general theory of optimal algorithms*, ACM Monograph Series, Academic Press, Inc., New York–London 1980, xiv+341 pp.

[10] L. Plaskota, *Noisy information and computational complexity*, Cambridge Univ. Press, Cambridge 1996, xii+308 pp.

[11] K. Yu. Osipenko, *Optimal recovery of analytic functions*, Nova Science Publ., Inc., New York 2000, 220 pp.

[12] Г. Г. Магарил-Ильяев, В. М. Тихомиров, *Выпуклый анализ и его приложения*, 2-е изд., Эдиториал УРСС, М. 2003, 176 с.; англ. пер. 1-го изд.:, G. G. Magaril-Il'yaev, V. M. Tikhomirov, *Convex analysis: theory and applications*, Transl. Math. Monogr., vol. 222, Amer. Math. Soc., Providence, RI 2003, viii+183 pp.

[13] G. G. Magaril-Il'yaev, V. M. Tikhomirov, *Convex analysis: theory and applications*, Transl. Math. Monogr., vol. 222, Amer. Math. Soc., Providence, RI 2003, viii+183 pp.

[14] A. A. Melkman, C. A. Micchelli, “Optimal estimation of linear operators in Hilbert spaces from inaccurate data”, *SIAM J. Numer. Anal.* **16**:1 (1979), 87–105.

[15] Г. Г. Магарил-Ильяев, К. Ю. Осипенко, “Оптимальное восстановление функций и их производных по коэффициентам Фурье, заданным с погрешностью”, *Матем. сб.* **193**:3 (2002), 79–100; англ. пер.:, G. G. Magaril-Il'yaev, K. Yu. Osipenko, “Optimal recovery of functions and their derivatives from Fourier coefficients prescribed with an error”, *Sb. Math.* **193**:3 (2002), 387–407.

[16] Г. Г. Магарил-Ильяев, К. Ю. Осипенко, “Оптимальное восстановление функций и их производных по приближенной информации о спектре и неравенства для производных”, *Функ. анализ и его прил.* **37**:3 (2003), 51–64; англ. пер.:, G. G. Magaril-Il'yaev, K. Yu. Osipenko, “Optimal recovery of functions and their derivatives from inaccurate information about the spectrum and inequalities for derivatives”, *Funct. Anal. Appl.* **37**:3 (2003), 203–214.

[17] К. Ю. Осипенко, “Неравенство Харди–Литтлвуда–Полиа для аналитических функций из пространств Харди–Соболева”, *Матем. сб.* **197**:3 (2006), 15–34; англ. пер.:, K. Yu. Osipenko, “The Hardy–Littlewood–Pólya inequality for analytic functions in Hardy–Sobolev spaces”, *Sb. Math.* **197**:3 (2006), 315–334.

[18] В. В. Арестов, “Приближение неограниченных операторов ограниченными и родственные экстремальные задачи”, *УМН* **51**:6(312) (1996), 89–124; англ. пер.:, V. V. Arestov, “Approximation of unbounded operators by bounded operators and related extremal problems”, *Russian Math. Surveys* **51**:6 (1996), 1093–1126.

[19] Г. Г. Магарил-Ильяев, К. Ю. Осипенко, “Оптимальное восстановление значений функций и их производных по неточно заданному преобразованию Фурье”, *Матем. сб.* **195**:10 (2004), 67–82; англ. пер.:, G. G. Magaril-Il’yaev, K. Yu. Osipenko, “Optimal recovery of values of functions and their derivatives from inaccurate data on the Fourier transform”, *Sb. Math.* **195**:10 (2004), 1461–1476.

[20] Г. Г. Магарил-Ильяев, К. Ю. Осипенко, “Как наилучшим образом восстановить функцию по неточно заданному спектру?”, *Матем. заметки* **92**:1 (2012), 59–67; англ. пер.:, G. G. Magaril-Il’yaev, K. Yu. Osipenko, “How best to recover a function from its inaccurately given spectrum?”, *Math. Notes* **92**:1 (2012), 51–58.

[21] В. В. Арестов, “Приближение линейных операторов и родственные экстремальные задачи”, *Приближение функций и операторов*, Сборник статей, Тр. МИАН СССР, **138**, 1975, с. 29–42; англ. пер.:, V. V. Arestov, “Approximation of linear operators, and related extremal problems”, *Proc. Steklov Inst. Math.* **138** (1977), 31–44.

[22] В. И. Левин, “Точные константы в неравенствах типа Карлсона”, *Докл. АН СССР* **59**:4 (1948), 635–639.

[23] Ф. И. Андрианов, “Многомерные аналоги неравенства Карлсона и его обобщений”, *Изв. вузов. Матем.*, 1967, № 1, 3–7.

[24] Г. Г. Магарил-Ильяев, К. Ю. Осипенко, “Неравенство Харди–Литтлвуда–Полиа и восстановление производных по неточной информации”, *Докл. РАН* **438**:3 (2011), 300–302; англ. пер.:, G. G. Magaril-Il’yaev, K. Yu. Osipenko, “The Hardy–Littlewood–Pólya inequality and the reconstruction of derivatives from inaccurate data”, *Dokl. Math.* **83**:3 (2011), 337–339.

[25] Н. Д. Выск, К. Ю. Осипенко, “Оптимальное восстановление решения волнового уравнения по неточным начальным данным”, *Матем. заметки* **81**:6 (2007), 803–815; англ. пер.:, N. D. Vysk, K. Yu. Osipenko, “Optimal reconstruction of the solution of the wave equation from inaccurate initial data”, *Math. Notes* **81**:6 (2007), 723–733.

K. Yu. Osipenko (Konstantin Yu. Osipenko)
Moscow State Aviation Technological University;
Institute for Information Transmission Problems,
Russian Academy of Sciences, Moscow
E-mail: `kosipenko@yahoo.com`

Received 17.01.2014 and 01.07.2014