OPTIMAL RECOVERY AND EXTREMUM THEORY

GEORGII G. MAGARIL-IL'YAEV, KONSTANTIN YU. OSIPENKO, and VLADIMIR M. TIKHOMIROV[1]

Abstract. In this paper optimal recovery problems of linear functionals on classes of smooth and analytic functions on the basis of linear information are considered from the general viewpoint of extremum theory. A general result about the connection of optimal recovery method with Lagrange multipliers of some convex extremal problem is applied to the analysis of classical recovery problems on the generalized Sobolev, Hardy, and Hardy–Sobolev classes.

Keywords: optimal recovery, Lagrange principle, Hardy spaces.

2000 Mathematics Subject Classification: 41A46, 30D55.

1 Setting of the problem and general theory

Let X and Y be real or complex linear spaces and x' a linear functional on X. It is required to recover x' (as exactly as possible) on elements from some set (class) $A \subset X$ using the information $y = Fx$ where $F\colon A \to Y$ is a linear operator which is called an *information operator*. Any function $\varphi\colon F(A) \to K$ where $K = \mathbb{R}$ or $\mathbb{C}$ we call a *method of recovery of x' on A from the information F*. The error of recovery is given by

$$e(x', A, F, \varphi) = \sup_{x \in A} |\langle x', x\rangle - \varphi(Fx)|.$$

The value

$$E(x', A, F) = \inf_{\varphi} e(x', A, F, \varphi) \tag{1}$$

where the infimum is taken over all functions $\varphi\colon F(A) \to K$ is called the *error of optimal recovery*. Any method $\widehat{\varphi}$ for which the infimum in (1) is attained we call an *optimal recovery method*.

Examples of such recovery problems are the problem of best integration methods (it is required to recover an integral of a function from some class using information about values of the function and its derivatives at a fixed system of points), the problem of recovery of a function value or a value of its derivative at some given point using information about the Fourier coefficients, Taylor coefficients, or values of the function at some other points, etc.

The method of the solution of the optimal recovery problems which we propose in this paper is based on the following concepts. In problem (1), for a convex and balanced set A, among optimal methods of recovery there exists a linear method. Thus the infimum in (1) may be taken over linear functionals on Y. In other words, the value $E(x', A, F,)$ is the value of the following convex problem

$$\sup_{x \in A} |\langle x', x\rangle - \langle y', Fx\rangle| \to \min, \quad y' \in Y'$$

[1]This research was carried out with the financial support of the Russian Foundation for Basic Research (grants №99-01-01181, №02-01-00386), Federal Program on Support of Leading Scientific Schools (grant №00–15–96109), INTAS-97-1050, and Research Grant 12513 of the Royal Swedish Acad. of Sci.

(where Y' is the algebraic dual of Y), which is dual to another convex problem (see [13, p. 61]; here for definiteness X and Y are complex linear spaces)

$$\mathrm{Re}\langle x', x\rangle \to \max, \quad Fx = 0, \quad x \in A, \tag{2}$$

which we call an *associated problem* to (1). Denote by

$$\mathcal{L}(x, \lambda, \lambda_0) = \lambda_0 \,\mathrm{Re}\langle x', x\rangle + \mathrm{Re}\langle \lambda, Fx\rangle$$

the Lagrange function of the problem (2) where $\lambda_0 \le 0$ and $\lambda \in Y'$ are the Lagrange multipliers. If there exists a solution to (2) then it follows from the general theory of extremum that the Lagrange multipliers are connected with the solution of the dual problem, i.e., with the optimal method of recovery. The explicit assertions are contained in the following theorem.

Theorem 1 (the Lagrange principle for optimal recovery problems). *Let X and Y be real or complex linear spaces, A a convex balanced subset of X, and $F\colon X \to Y$ a linear operator. Then the admissible in (2) point $\widehat{x}$ is a solution of this problem if and only if there exists the Lagrange multiplier $\widehat{\lambda} \in Y'$ for which*

$$\min_{x \in A} \mathcal{L}(x, \widehat{\lambda}, -1) = \mathcal{L}(\widehat{x}, \widehat{\lambda}, -1). \tag{3}$$

In this case

$$\langle x', x\rangle \approx \langle \widehat{\lambda}, Fx\rangle$$

is an optimal method of recovery in (1) *and*

$$E(x', A, F) = \mathrm{Re}\langle x', \widehat{x}\rangle.$$

Proof: We use the following algebraic version of the separation theorem: *Let C be a convex subset of a real linear space X, $\mathrm{icr}\, C \ne \emptyset$[2], and $x_0 \notin \mathrm{icr}\, C$. Then there exists $x' \in X'$, $x' \ne 0$, such that*

$$\inf_{x \in C} \langle x', x\rangle \ge \langle x', x_0\rangle$$

and $\langle x', x\rangle > \langle x', x_0\rangle$ for all $x \in \mathrm{icr}\, C$ (see [13, p. 38]).

1. Necessity. Let $\widehat{x}$ be a solution of (2). Suppose first that $\mathrm{Re}\langle x', \widehat{x}\rangle = 0$. We show that in this case there exists a $\widehat{\lambda} \in Y'$ such that

$$\mathrm{Re}\langle x', x\rangle = \mathrm{Re}\langle \widehat{\lambda}, Fx\rangle \tag{4}$$

for all $x \in A$. From here evidently follows (3). Note that since A is balanced $\mathrm{Re}\langle x', x\rangle = 0$ for all admissible x. Define the functional l on the subspace $F(\mathrm{span}\, A)$ by the equality $l(y) = \mathrm{Re}\langle x', x\rangle$ where $x \in F^{-1}(y)$. This definition is well-defined. Indeed, let $x_1, x_2 \in F^{-1}(y)$. Since A is balanced it is absorbing in $\mathrm{span}\, A$ and therefore there exists an $\alpha > 0$ such that $\alpha(x_1 - x_2) \in A$. It is clear that $\alpha(x_1 - x_2) \in F^{-1}(0)$ and consequently $\alpha(x_1 - x_2)$ is an admissible element in (2). Thus $\mathrm{Re}\langle x', \alpha(x_1 - x_2)\rangle = 0$, that is $\mathrm{Re}\langle x', x_1\rangle = \mathrm{Re}\langle x', x_2\rangle$. It is easy to verify that l is a linear functional. Denote by $\widehat{\lambda}$ any of its extensions on the all Y. It is obvious that (4) is fulfilled with this $\widehat{\lambda}$.

[2] $\mathrm{icr}\, C$ is the set of algebraic relative interior points of C. If $\mathrm{aff}\, C = \overline{x} + L_C$ (where $\overline{x} \in C$ and L_C is a subspace of X) is an affine hull of C, i.e., a minimal linear manifold containing C, then $x_0 \in \mathrm{icr}\, C$ if for any $x \in L_C$ there exists $\varepsilon = \varepsilon(x) > 0$ such that $[x_0, x_0 + \varepsilon x] \subset C$.

Assume that $\operatorname{Re}\langle x', \widehat{x}\rangle \neq 0$. Denote by $Y_{\mathbb{R}}$ a real linear space of elements from Y with multiplication only by the real numbers. Consider the set

$$C = \{\, (\alpha, y) \in \mathbb{R} \times Y_{\mathbb{R}} \mid \alpha = \operatorname{Re}\langle x', x\rangle,\ y = Fx,\ x \in A \,\}.$$

It is easy to see that C is a convex balanced set and, in particular, $(0,0) \in \operatorname{icr} C$. It is also easy to verify that $(\operatorname{Re}\langle x', \widehat{x}\rangle, 0) \notin \operatorname{icr} C$. Then by the separation theorem there exist $(\widehat{\lambda}_0, \widehat{\lambda}_{\mathbb{R}}) \in \mathbb{R} \times Y'_{\mathbb{R}}$ not all equal to zero such that

$$\widehat{\lambda}_0 \alpha + \langle \widehat{\lambda}_{\mathbb{R}}, y\rangle \geq \widehat{\lambda}_0 \operatorname{Re}\langle x', \widehat{x}\rangle, \quad \forall (\alpha, y) \in C, \tag{5}$$

and

$$\widehat{\lambda}_0 \alpha + \langle \widehat{\lambda}_{\mathbb{R}}, y\rangle > \widehat{\lambda}_0 \operatorname{Re}\langle x', \widehat{x}\rangle, \quad \forall (\alpha, y) \in \operatorname{icr} C. \tag{6}$$

Since $(0,0) \in \operatorname{icr} C$ it follows from (6) that $\widehat{\lambda}_0 \neq 0$. It is clear that $(\operatorname{Re}\langle x', \widehat{x}\rangle, 0) \in C$ and consequently $2^{-1}(\operatorname{Re}\langle x', \widehat{x}\rangle, 0) \in C$. Substituting this in (5) we have that $2\widehat{\lambda}_0 \leq \widehat{\lambda}_0$, that is $\widehat{\lambda}_0 < 0$. Let us assume that $\widehat{\lambda}_0 = -1$.

Denote by $\widehat{\lambda}$ an element from Y' for which

$$\langle \widehat{\lambda}_{\mathbb{R}}, y\rangle = \operatorname{Re}\langle \widehat{\lambda}, y\rangle \tag{7}$$

for all $y \in Y$. Let $x \in A$. Then

$$(\operatorname{Re}\langle x', x\rangle, Fx) \in C \tag{8}$$

and we have

$$\begin{aligned}
\mathcal{L}(x, \widehat{\lambda}, -1) &= -\operatorname{Re}\langle x', x\rangle + \operatorname{Re}\langle \widehat{\lambda}, Fx\rangle \overset{(7)}{=} -\operatorname{Re}\langle x', x\rangle + \langle \widehat{\lambda}_{\mathbb{R}}, Fx\rangle \\
&\overset{(5),(8)}{\geq} -\operatorname{Re}\langle x', \widehat{x}\rangle \overset{F\widehat{x}=0}{=} -\operatorname{Re}\langle x', \widehat{x}\rangle + \operatorname{Re}\langle \widehat{\lambda}, F\widehat{x}\rangle = \mathcal{L}(\widehat{x}, \widehat{\lambda}, -1).
\end{aligned}$$

2. Sufficiency. Let (3) be fulfilled and x be an admissible point in (2). Then

$$-\operatorname{Re}\langle x', x\rangle = -\operatorname{Re}\langle x', x\rangle + \operatorname{Re}\langle \widehat{\lambda}, Fx\rangle \overset{(3)}{\geq} -\operatorname{Re}\langle x', \widehat{x}\rangle + \operatorname{Re}\langle \widehat{\lambda}, F\widehat{x}\rangle = -\operatorname{Re}\langle x', \widehat{x}\rangle,$$

i.e., $\widehat{x}$ is a solution of (2).

Let us prove the second assertion of the theorem. Since A is balanced (3) may be rewritten as follows

$$\max_{x \in A} |\langle x', x\rangle - \langle \widehat{\lambda}, Fx\rangle| = \operatorname{Re}\langle x', \widehat{x}\rangle. \tag{9}$$

Hence

$$E(x', A, F) \leq \operatorname{Re}\langle x', \widehat{x}\rangle. \tag{10}$$

Let us show that, in fact, we have here equality. Assume $x \in A$ and $Fx = 0$. Since $-x \in A$ for any method φ we have

$$\begin{aligned}
&2\operatorname{Re}\langle x', x\rangle \leq 2|\langle x', x\rangle| = |\langle x', x\rangle - \varphi(0) + \varphi(0) + \langle x', x\rangle| \\
&\quad \leq |\langle x', x\rangle - \varphi(0)| + |\langle x', -x\rangle - \varphi(0)| \leq 2 \sup_{\substack{x \in A \\ Fx = 0}} |\langle x', x\rangle - \varphi(0)| \leq 2 \sup_{x \in A} |\langle x', x\rangle - \varphi(Fx)|.
\end{aligned}$$

This means that the reverse inequality to (10) holds and thus

$$E(x', A, F) = \operatorname{Re}\langle x', \widehat{x}\rangle.$$

It follows from this equality and (9) that $\widehat{\lambda}$ is an optimal method of recovery. ■

Remark 1 *The existence of a linear optimal method of recovery in the problem* (1) *was discovered for the first time by Smolyak [21] for the real case and convex centrally-symmetric set A with* $\dim \operatorname{span} F(A) < \infty$*. The generalization of this result and corresponding literature may be found in [12]. Dual methods for the solution of the problem* (1) *were used by many authors (see [5], [14], [24], [15], [4]) but the exact connection between the problems* (1) *and* (2)*, which is that the optimal method of recovery is none other than the Lagrange multiplier in* (2)*, was apparently used for the first time in [13].*

Remark 2 *The constraints on A in* (2) *may be also described by a system of equalities and/or inequalities. In this case some of them may be included in the Lagrange function (with corresponding multipliers). But an optimal method is always the Lagrange multiplier at the constraints related to the information operator. The proof of this fact (which is more general only in appearance) is just the same as that of Theorem 1.*

In Theorem 1 the so-called Lagrange principle for convex extremal problems with constraints defined by equalities and inclusions is confirmed. This principle is in the fact that if a problem has a solution then there exist Lagrange multipliers such that this solution is the absolute minimum of the Lagrange function on the set of the remaining constraints (not included in the Lagrange function). In the examples below we use this principle as an heuristic method. Namely, using equality (3) we extract the information about what $\widehat{x}$ and $\widehat{\lambda}$ must be to satisfy the equality (3). After that we use the sufficiency of this condition and find the solution of (2) and the optimal recovery method. Sometimes, instead of using the sufficiency, it is easy to verify the optimality of the obtained method directly.

In the next section using Theorem 1 we prove a general result about the optimal recovery of functions from classes defined by the convolution with some kernels on the basis of information about Fourier coefficients. In Section 3 we apply this result to classes defined by the convolution with cyclic variation diminishing kernels. We list there several well-known results which are particular cases of the considered problem.

In Section 4 using Theorem 1 we obtain optimal recovery algorithms for Hardy classes. These results are known, but we point them out in order to demonstrate the general method from Theorem 1. In Sections 5 and 6 we obtain some new results related to optimal recovery methods from Hardy–Sobolev classes.

2 Optimal recovery of function values from Fourier coefficients

Let $r \in \mathbb{N}$ and $1 \le p \le \infty$. Denote by $W_p^r(\mathbb{T})$ the Sobolev class of functions $x(\cdot)$ defined on the unit circle $\mathbb{T}$ (realized as the interval $[-\pi, \pi]$ with identified endpoints) whose the $(r-1)$st derivative is absolutely continuous and $\|x^{(r)}(\cdot)\|_{L_p(\mathbb{T})} \le 1$. In 1936 Favard proved that for all $n \in \mathbb{N}$ and for all functions $x(\cdot) \in W_\infty^r(\mathbb{T})$ such that

$$\int_{\mathbb{T}} x(t) \cos kt \, dt = \int_{\mathbb{T}} x(t) \sin kt \, dt = 0, \quad k = 0, 1, \dots, n-1,$$

the following exact inequality

$$\|x(\cdot)\|_{C(\mathbb{T})} \le \frac{K_r}{n^r} \tag{11}$$

holds. The numbers K_r (known as the Favard constants) are defined by

$$K_r = \frac{4}{\pi}\sum_{j=0}^{\infty}\frac{(-1)^{j(r+1)}}{(2j+1)^{r+1}}, \quad r\in\mathbb{Z}_+.$$

Note that the case when $n=r=1$ was previously considered by H. Bohr and therefore (11) is usually called the *Bohr–Favard inequality.*

It is obvious that the problem of the exact constant in (11) is equivalent (in view of shift-invariance of the norm) to the following:

$$x(0)\to\max, \quad a_0=\ldots=a_{n-1}=b_1=\ldots=b_{n-1}=0, \quad x(\cdot)\in W_\infty^r(\mathbb{T}),$$

where $a_0,\ldots,b_{n-1}$ are the Fourier coefficients of $x(\cdot)$. This problem has the form (2) and hence it relates to the optimal recovery problem of a function value at the point 0 on the class $W_\infty^r(\mathbb{T})$ from Fourier coefficients. The same problem is closely related to the problem of deviation of the class $W_\infty^r(\mathbb{T})$ from the space of trigonometric polynomials $\mathcal{T}_{n-1}$ of degree at most $n-1$. Beginning from the Favard's result a lot of papers were devoted to these subjects.

The recovery problem of a function value at some given point from the Fourier coefficients on the class defined as the convolution of a real kernel $K(\cdot)$ with functions from the unit ball of $L_p(\mathbb{T})$ involves many particular cases. More precisely, let $K(\cdot)\in L_{p'}(\mathbb{T})$ $(1/p+1/p'=1)$ and

$$\alpha_k=\frac{1}{\pi}\int_{\mathbb{T}}K(t)\cos kt\,dt, \quad k\in\mathbb{Z}_+, \quad \beta_k=\frac{1}{\pi}\int_{\mathbb{T}}K(t)\sin kt\,dt, \quad k\in\mathbb{N},$$

be the Fourier coefficients of $K(\cdot)$. Assume that $\alpha_k^2+\beta_k^2\neq 0$ $(\beta_0=0)$ with the exception of a finite (possibly empty) set $Q\subset\mathbb{Z}_+$. Set $\mathcal{T}_Q=\operatorname{span}\{\cos kt,\sin kt,\ k\in Q\}$ and

$$\mathcal{W}_p^K(\mathbb{T},Q)=\Big\{\,x(\cdot)\mid x(\cdot)=y(\cdot)+\frac{1}{\pi}\int_{\mathbb{T}}K(\cdot-t)u(t)\,dt,\ y(\cdot)\in\mathcal{T}_Q,\ u(\cdot)\in\mathcal{T}_Q^{\perp},\ u(\cdot)\in L_p(\mathbb{T})\,\Big\},$$

where $\mathcal{T}_Q^{\perp}$ is the annihilator of $\mathcal{T}_Q$. It is clear that $\mathcal{W}_p^K(\mathbb{T},Q)$ is a subspace of the space $C(\mathbb{T})$ of continuous functions on $\mathbb{T}$. The corresponding convolution class is the set

$$W_p^K(\mathbb{T},Q)=\{x(\cdot)\in\mathcal{W}_p^K(\mathbb{T},Q)\mid \|u(\cdot)\|_{L_p(\mathbb{T})}\le 1\}.$$

For instance, in the case of the Sobolev class $W_p^r(\mathbb{T})$ we have $Q=\{0\}$, $K(\cdot)=B_r(\cdot)$ where

$$B_r(t)=\sum_{k=1}^{\infty}\frac{\cos(kt-\pi r/2)}{k^r}$$

is the Bernoulli kernel.

Consider the problem of optimal recovery of a function $x(\cdot)$ at a point $\theta\in\mathbb{T}$ on the class $W_p^K(\mathbb{T},Q)$ from the Fourier coefficients

$$a_k=\frac{1}{\pi}\int_{\mathbb{T}}x(t)\cos kt\,dt, \quad k=0,1,\ldots,n-1, \quad b_k=\frac{1}{\pi}\int_{\mathbb{T}}x(t)\cos kt\,dt, \quad k=1,\ldots,n-1.$$

In accordance with the general notation we have $X=\mathcal{W}_p^K(\mathbb{T},Q)$, $A=W_p^K(\mathbb{T},Q)$, $Y=\mathbb{R}^{2n-1}$, $Fx(\cdot)=\operatorname{Four}_n x(\cdot)=(a_0,a_1,\ldots,a_{n-1},b_1,\ldots,b_{n-1})$, and $\langle x',x(\cdot)\rangle=x(\theta)$.

Note that if $\{0,1,\ldots,n-1\}\setminus Q\neq\emptyset$, then it is easy to check that the error of optimal recovery equals $+\infty$ and hence any method is optimal. Therefore we assume that $Q\subset\{0,1,\ldots,n-1\}$. Put $Q'=\{0,1,\ldots,n-1\}\setminus Q$.

For a normed linear space X, $x \in X$, and a nonempty subset A of X denote by $d(x, A, X)$ the deviation from x to A in the metric of X.

We say that a function $K(\cdot) \in L_1(\mathbb{T})$ satisfies the *Favard γ-property* (for a fixed $n \in \mathbb{N}$) if there exists a polynomial $\widehat{q}(\cdot) \in \mathcal{T}_{n-1}$ and a number $\gamma \in [0, \pi/n)$ such that the function $(K(t) - \widehat{q}(t)) \sin n(t+\gamma)$ is nonnegative or nonpositive for almost all $t \in \mathbb{T}$. If $K(\cdot)$ is a continuous function, then $\widehat{q}(\cdot)$ may be found as a polynomial which interpolates $K(\cdot)$ at the zeros of $\sin n(\cdot + \gamma)$.

The following theorem holds.

Theorem 2 (on optimal recovery from Fourier coefficients). *Let $1 < p \le \infty$ and*

$$\widehat{p}(t) = \frac{A_0}{2} + \sum_{k=1}^{n-1} (A_k \cos kt + B_k \sin kt)$$

be a polynomial of the best approximation of $K(\cdot)$ by $\mathcal{T}_{n-1}$ in the metric $L_{p'}(\mathbb{T})$. Then

$$x(\theta) \approx \widehat{\mu}_0 a_0 + \sum_{k=1}^{n-1} (\widehat{\mu}_k(\theta) a_k + \widehat{\nu}_k(\theta) b_k),$$

where $\widehat{\mu}_0 = 1/2$ if $0 \in Q$ and $\widehat{\mu}_0 = A_0/(2\alpha_0)$ if $0 \notin Q$; $\widehat{\mu}_k(\theta) = \cos k\theta$, $\widehat{\nu}_k(\theta) = \sin k\theta$ if $k \in Q \setminus \{0\}$ and

$$\widehat{\mu}_k(\theta) = \frac{(\alpha_k A_k + \beta_k B_k) \cos k\theta + (\alpha_k B_k - \beta_k A_k) \sin k\theta}{\alpha_k^2 + \beta_k^2},$$
$$\widehat{\nu}_k(\theta) = \frac{(\beta_k A_k - \alpha_k B_k) \cos k\theta + (\alpha_k A_k + \beta_k B_k) \sin k\theta}{\alpha_k^2 + \beta_k^2}$$

if $k \in Q'$, is an optimal method of recovery of $x(\theta)$ on the class $W_p^K(\mathbb{T}, Q)$ from Fourier coefficients. Moreover,

$$E(x(\theta), W_p^K(\mathbb{T}, Q), \mathrm{Four}_n) = \frac{1}{\pi} d\left(K(\cdot), \mathcal{T}_{n-1}, L_{p'}(\mathbb{T})\right).$$

If $p = \infty$ and $K(\cdot)$ satisfies the Favard γ-property, then

$$E(x(\theta), W_\infty^K(\mathbb{T}, Q), \mathrm{Four}_n) = \frac{1}{\pi} \left| \int_{\mathbb{T}} K(t) \operatorname{sign} \sin n(t+\gamma)\, dt \right|.$$

Let us formulate a corollary from this theorem related to the deviation of $W_p^K(\mathbb{T}, Q)$ from the subspace of trigonometric polynomials. Recall that for a normed linear space X and nonempty subsets A and C of X the value

$$d(C, A, X) = \sup_{x \in C} d(x, A, X)$$

is called the *deviation of C from A in the metric X.*

The value

$$d^L(W_p^K(\mathbb{T}, Q), \mathcal{T}_{n-1}, L_p(\mathbb{T})) = \inf_{\Lambda} \sup_{x(\cdot) \in W_p^K(\mathbb{T}, Q)} \|x(\cdot) - \Lambda x(\cdot)\|_{L_p(\mathbb{T})},$$

where the infimum is taken over all linear operators $\Lambda \colon \mathcal{W}_p^K(\mathbb{T}, Q) \to \mathcal{T}_{n-1}$, characterizes a best linear approximation of $W_p^K(\mathbb{T}, Q)$ by trigonometric polynomials from $\mathcal{T}_{n-1}$. An operator Λ for which the infimum is attained is called an *extremal method.*

Obviously,

$$d(W_p^K(\mathbb{T}, Q), \mathcal{T}_{n-1}, L_p(\mathbb{T})) \le d^L(W_p^K(\mathbb{T}, Q), \mathcal{T}_{n-1}, L_p(\mathbb{T})).$$

Corollary 1 *If $K(\cdot)$ satisfies the Favard γ-property, then*

$$d\left(W^K_\infty(\mathbb{T},Q),\mathcal{T}_{n-1},C(\mathbb{T})\right)=d^L\left(W^K_\infty(\mathbb{T},Q),\mathcal{T}_{n-1},C(\mathbb{T})\right)=E(x(\theta),W^K_\infty(\mathbb{T},Q),\mathrm{Four}_n)$$

and the operator $\widehat{\Lambda}$ which associates $x(\cdot)\in\mathcal{W}^K_\infty(\mathbb{T},Q)$ with the polynomial

$$\widehat{\mu}_0a_0+\sum_{k=1}^{n-1}(\widehat{\mu}_k(\theta)a_k+\widehat{\nu}_k(\theta)b_k)$$

is extremal.

Proof of Theorem 2: The problem associated with the considered problem of the optimal recovery has the form

$$x(\theta)\to\max,\quad a_0=\ldots=a_{n-1}=b_1=\ldots=b_{n-1}=0,\quad x(\cdot)\in W^K_p(\mathbb{T}). \tag{12}$$

Its Lagrange function is

$$\begin{aligned}\mathcal{L}(x(\cdot),\mu_0,\mu_1,\ldots,\mu_{n-1},\nu_1,\ldots,\nu_{n-1},\lambda_0)&=\lambda_0x(\theta)+\frac{\mu_0}{\pi}\int_{\mathbb{T}}x(t)\,dt\\&+\frac{1}{\pi}\sum_{k=0}^{n-1}\int_{\mathbb{T}}(\mu_k\cos kt+\nu_k\sin kt)x(t)\,dt,\end{aligned}\tag{13}$$

where λ_0, μ_k, $k=0,1\ldots,n-1$, and ν_k, $k=1,\ldots,n-1$, are the Lagrange multipliers.

Further we argue heuristically. Namely, we set $\lambda_0=-1$ and use the Lagrange principle formally. It allows us to understand how the solution of (12) is organized and what form the Lagrange multipliers have (which determine an optimal method of recovery according to Theorem 1). After that we use the sufficient conditions of Theorem 1.

Let $\widehat{x}(\cdot)$ be a solution of the problem (12). Then (according to the Lagrange principle) there exist such numbers $\widehat{\mu}_k$, $k=0,1\ldots,n-1$, and $\widehat{\nu}_k$, $k=1,\ldots,n-1$, that the function $\mathcal{L}(x(\cdot),\widehat{\mu}_0,\widehat{\mu}_1,\ldots,\widehat{\mu}_{n-1},\widehat{\nu}_1,\ldots,\widehat{\nu}_{n-1},-1)$ attains its absolute minimum on $W^r_p(\mathbb{T})$ at the point $\widehat{x}(\cdot)$ (for simplicity we do not indicate that $\widehat{\mu}_k$ and $\widehat{\nu}_k$ depend on θ).

For definiteness we assume that $0\in Q$. Substitute in $\mathcal{L}$ instead of $x(\cdot)$ its representation in terms of $u(\cdot)$ and

$$y(\cdot)=\gamma_0+\sum_{k\in Q\setminus\{0\}}(\gamma_k\cos kt+\delta_k\sin kt)$$

(the coefficients γ_k and δ_k are uniquely determined by $x(\cdot)$ since they are the corresponding Fourier coefficients of $x(\cdot)$). Denote by $\widehat{\gamma}_0$, $\widehat{\gamma}_k$, $\widehat{\delta}_k$, $k\in Q\setminus\{0\}$, and $\widehat{u}(\cdot)$ the coefficients and the function which correspond to $\widehat{x}(\cdot)$. By means of simple calculations we obtain that the function

$$\begin{aligned}&-\gamma_0-\sum_{k\in Q\setminus\{0\}}(\gamma_k\cos k\theta+\delta_k\sin k\theta)+2\widehat{\mu}_0\gamma_0+\sum_{k\in Q\setminus\{0\}}(\widehat{\mu}_k\gamma_k+\widehat{\nu}_k\delta_k)\\&\quad+\frac{1}{\pi}\int_{\mathbb{T}}\Big(-K(\theta-t)+\sum_{k\in Q'}((\widehat{\mu}_k\alpha_k+\widehat{\nu}_k\beta_k)\cos kt+(\widehat{\nu}_k\alpha_k-\widehat{\mu}_k\beta_k)\sin kt)\Big)u(t)\,dt\end{aligned}\tag{14}$$

attains its absolute minimum on the set

$$\gamma_0,\gamma_k,\delta_k\in\mathbb{R},\quad \frac{1}{\pi}\int_{\mathbb{T}}u(t)\begin{matrix}\cos kt\\ \sin kt\end{matrix}dt=0,\ k\in Q,\quad \|u(\cdot)\|_{L_p(\mathbb{T})}\le 1 \tag{15}$$

at the point $(\{\widehat{\gamma}_0, \widehat{\gamma}_k, \widehat{\delta}_k\}_{k\in Q\setminus\{0\}}, \widehat{u}(\cdot))$. Hence $\widehat{\mu}_0 = 1/2$ and $\widehat{\mu}_k = \cos k\theta$, $\widehat{\nu}_k = \sin k\theta$, $k \in Q \setminus \{0\}$.

The problem (14)–(15) is a problem of type (12) (the minimization of a linear functional on a convex balanced set). Its Lagrange function may be written obviously (we set the multiplier at the minimizing functional equals 1 and do not include the last constraint in (15)). Then according to the Lagrange principle there exist such $\widehat{c}_0$, $\widehat{c}_k, \widehat{d}_k$, $k \in Q \setminus \{0\}$, that the function

$$\frac{1}{\pi}\int_{\mathbb{T}}\Bigg(-K(\theta-t) + \widehat{c}_0 + \sum_{k\in Q\setminus\{0\}}(\widehat{c}_k\cos kt + \widehat{d}_k \sin kt) \\ + \sum_{k\in Q'}((\widehat{\mu}_k\alpha_k + \widehat{\nu}_k\beta_k)\cos kt + (\widehat{\nu}_k\alpha_k - \widehat{\mu}_k\beta_k)\sin kt)\Bigg)u(t)\,dt \quad (16)$$

attains the absolute minimum on the unite ball of $L_p(\mathbb{T})$ at the point $\widehat{u}(\cdot)$. If we denote by $L(\cdot)$ the multiplier preceding $u(\cdot)$ under the integral sign, then it is clear that

$$\widehat{u}(\cdot) = -\|L(\cdot)\|^{1-p'}_{L_{p'}(\mathbb{T})}|L(\cdot)|^{p'-1}\operatorname{sign} L(\cdot).$$

Note that $\widehat{u}(\cdot) \in \mathcal{T}^{\perp}_{n-1}$. It follows from (15) (when $k \in Q$) and the fact that for $k \in Q'$, $\alpha_k^2 + \beta_k^2 \neq 0$ and therefore the vanishing of Fourier coefficients of $\widehat{x}(\cdot)$ implies the vanishing of corresponding Fourier coefficients of $\widehat{u}(\cdot)$. Then in accordance with the criterion of the best approximation in $L_{p'}(\mathbb{T})$ we obtain that the polynomial

$$\overline{p}(t) = \widehat{c}_0 + \sum_{k\in Q\setminus\{0\}}(\widehat{c}_k\cos kt + \widehat{d}_k\sin kt) + \sum_{k\in Q'}((\widehat{\mu}_k\alpha_k + \widehat{\nu}_k\beta_k)\cos kt + (\widehat{\nu}_k\alpha_k - \widehat{\mu}_k\beta_k)\sin kt)$$

must be the best approximation polynomial for the function $t \to K(\theta - t)$ by the subspace $\mathcal{T}_{n-1}$ in the metric $L_{p'}(\mathbb{T})$.

Now we shall apply sufficient conditions. Let $\widehat{p}(\cdot)$ be the polynomial mentioned in the statement of the theorem. Then $\widehat{p}(\theta - \cdot)$ is the polynomial of the best approximation of $K(\theta-\cdot)$ by the subspace $\mathcal{T}_{n-1}$ in the metric $L_{p'}(\mathbb{T})$. Choose multipliers $\widehat{c}_0$, $\widehat{c}_k$, $\widehat{d}_k$, $k \in Q\setminus\{0\}$, and $\widehat{\mu}_k, \widehat{\nu}_k$, $k \in Q'$, so that $\overline{p}(\cdot) = \widehat{p}(\theta - \cdot)$. We obtain just the same formulae for these coefficients which are given in the theorem. With these Lagrange multipliers the polynomial $\overline{p}(\cdot)$ is in fact the polynomial of the best approximation of $K(\cdot)$ by the subspace $\mathcal{T}_{n-1}$ in the metric $L_{p'}(\mathbb{T})$. From the criterion of the best approximation it follows that the function

$$\widehat{u}(\cdot) = -\|\widehat{L}(\cdot)\|^{1-p'}_{L_{p'}(\mathbb{T})}|\widehat{L}(\cdot)|^{p'-1}\operatorname{sign}\widehat{L}(\cdot),$$

where $\widehat{L}(\cdot)$ is $L(\cdot)$ with just defined Lagrange multiplier, is orthogonal to $\mathcal{T}_{n-1}$ and evidently $\|\widehat{u}(\cdot)\|_{L_p(\mathbb{T})} = 1$. Hence $\widehat{u}(\cdot)$ is admissible in (15). Put $\widehat{\gamma}_0 = \widehat{\gamma}_k = \widehat{\delta}_k = 0$, $\widehat{\mu}_0 = 1/2$, $\widehat{\mu}_k = \cos k\theta$, and $\widehat{\nu}_k = \sin k\theta$, $k \in Q \setminus \{0\}$. Then since $\widehat{u}(\cdot)$ is a solution of (16) (with corresponding multipliers) by Theorem 1 the point $(\{\widehat{\gamma}_0, \widehat{\gamma}_k, \widehat{\delta}_k\}_{k\in Q\setminus\{0\}}, \widehat{u}(\cdot))$ is a solution of the problem (14)–(15). This is equivalent to the fact that the function

$$\widehat{x}(\cdot) = \frac{1}{\pi}\int_{\mathbb{T}} K(\cdot - \tau)\widehat{u}(\tau)\,d\tau$$

($\widehat{y}(\cdot) = 0$ since $\widehat{\gamma}_0 = \widehat{\gamma}_k = \widehat{\delta}_k = 0$, $k \in Q \setminus \{0\}$) gives the minimum of the Lagrange function (13) with $\lambda_0 = -1$ and Lagrange multipliers defined above. Since $\widehat{u}(\cdot) \in \mathcal{T}^{\perp}_{n-1}$, by the same arguments as above the Fourier coefficients of $\widehat{x}(\cdot)$ vanish for $k \in Q'$ and it means that $\widehat{x}(\cdot)$ is admissible in (12). Then by Theorem 1 it is a solution of this problem and the Lagrange multipliers define an optimal method. The case when $0 \in Q'$ is considered analogously. The first part of the theorem is proved.

Further,

$$\widehat{x}(\theta) = \frac{1}{\pi}\int_{\mathbb{T}} K(\theta - t)\widehat{u}(t)\,dt = \frac{1}{\pi}\int_{\mathbb{T}}(K(\theta - t) - \widehat{p}(\theta - t))\widehat{u}(t)\,dt$$
$$= \frac{1}{\pi}\|K(\cdot) - \widehat{p}(\cdot)\|_{L_{p'}(\mathbb{T})} = \frac{1}{\pi}d(K(\cdot), \mathcal{T}_{n-1}, L_{p'}(\mathbb{T})),$$

that is the quantity in the right-hand side is the value of the problem (12). Hence and from Theorem 1 the second assertion of the theorem follows.

Let $p = \infty$ and $K(\cdot)$ satisfies the Favard γ-property. Put

$$\overline{u}(t) = \operatorname{sign}(K(t) - \widehat{q}(t)),$$

where $\widehat{q}(\cdot)$ is from the definition of the Favard γ-property. Then

$$\overline{u}(t) = \varepsilon \operatorname{sign}\sin n(t + \gamma), \quad \varepsilon = 1 \text{ or } -1,$$

almost everywhere. Since it is clear that $\overline{u}(\cdot) \in \mathcal{T}_{n-1}^{\perp}$, from the criterion of the best approximation it follows that $\widehat{q}(\cdot)$ is the best approximation polynomial of $K(\cdot)$ by the subspace $\mathcal{T}_{n-1}$ in the metric $L_1(\mathbb{T})$. Consequently,

$$d(K(\cdot), \mathcal{T}_{n-1}, L_1(\mathbb{T})) = \int_{\mathbb{T}} |K(t) - \widehat{q}(t)|\,dt = \varepsilon\int_{\mathbb{T}}(K(t) - \widehat{q}(t))\operatorname{sign}\sin n(t + \gamma)dt$$
$$= \left|\int_{\mathbb{T}} K(t)\operatorname{sign}\sin n(t + \gamma)\,dt\right|.$$

Together with the previous equality this proves the last assertion of the theorem. ■

Proof of Corollary 1: The upper bound. Set

$$\kappa = \left|\int_{\mathbb{T}} K(t)\operatorname{sign}\sin n(t + \gamma)\,dt\right|.$$

Let $x(\cdot) \in W_p^K(\mathbb{T}, Q)$. Then taking into account the last assertion of Theorem 2 we have

$$\|x(\cdot) - \widehat{\Lambda}x(\cdot)\|_{C(\mathbb{T})} = \max_{\theta\in\mathbb{T}} |x(\theta) - \widehat{\Lambda}x(\theta)| \le \max_{\theta\in\mathbb{T}} E(x(\theta), W_p^K(\mathbb{T}, Q), \mathrm{Four}_n) = \frac{\kappa}{\pi}.$$

Hence
$$d^L(W_\infty^K(\mathbb{T}, Q), \mathcal{T}_{n-1}, C(\mathbb{T})) \le E(x(\theta), W_\infty^K(\mathbb{T}, Q), \mathrm{Four}_n) = \frac{\kappa}{\pi}.$$

The lower bound. Consider the function

$$\overline{x}(\cdot) = \frac{1}{\pi}\int_{\mathbb{T}} K(\cdot - \tau)\operatorname{sign}\sin n(\tau - \gamma)\,d\tau.$$

Clearly, $\overline{x}(\cdot) \in W_\infty^K(\mathbb{T}, Q)$. This function may be rewritten as follows

$$\overline{x}(t) = \frac{1}{\pi}\int_{\mathbb{T}} K(\tau)\operatorname{sign}\sin n(\tau + \gamma - t)\,d\tau.$$

It is easily seen that

$$\overline{x}\left(\frac{k\pi}{n}\right)=\frac{(-1)^k}{\pi}\int_{\mathbb{T}}K(\tau)\,\mathrm{sign}\sin n(\tau+\gamma)\,d\tau=\varepsilon(-1)^k\frac{\kappa}{\pi},\quad \varepsilon=1 \text{ or } -1,$$

and in view of the fact that $|\overline{x}(t)|\le\kappa/\pi$ the function $\overline{x}(\cdot)$ has $2n$-alternance on the period. By the Chebyshev alternance theorem the trivial polynomial is its best approximation polynomial by the subspace $\mathcal{T}_{n-1}$ in $C(\mathbb{T})$. Therefore,

$$d(W^K_\infty(\mathbb{T},Q),\mathcal{T}_{n-1},C(\mathbb{T}))\ge d(\overline{x}(\cdot),\mathcal{T}_{n-1},C(\mathbb{T}))=\|\overline{x}(\cdot)\|_{C(\mathbb{T})}=\frac{\kappa}{\pi}.\quad\blacksquare$$

3 Cyclic variation diminishing kernels

Denote by $\mathcal{K}(Q)$ the set of kernels $K(\cdot)\in L_1(\mathbb{T})$ for which for all $y(\cdot)\in\mathcal{T}_Q$ and all $u(\cdot)\in L_\infty(\mathbb{T})$ such that $u(\cdot)\perp\mathcal{T}_Q$ and $u(\cdot)\ne 0$ the inequality

$$S(y(\cdot)+(K*x)(\cdot))\le S(u(\cdot))$$

holds, where $S(u(\cdot))$ is the number of sign changes of $u(\cdot)$ on the period and

$$(K*u)(\cdot)=\frac{1}{\pi}\int_{\mathbb{T}}K(\cdot-t)u(t)\,dt.$$

For a function $u(\cdot)\in C(\mathbb{T})$ denote by $\mathrm{dist}\,u(\cdot)$ the length of the largest subinterval of $\mathbb{T}$ containing no zeros of $u(\cdot)$. Denote by $\mathcal{K}(Q,\delta)$ the class of kernels $K(\cdot)\in L_1(\mathbb{T})$ for which for all $u(\cdot)\in L_\infty(\mathbb{T})$ and $y(\cdot)\in\mathcal{T}_Q$ such that $u(\cdot)\perp\mathcal{T}_Q$, $u(\cdot)\ne 0$, and $\mathrm{dist}(y(\cdot)+(K*u)(\cdot))<\delta$ the inequality

$$S(y(\cdot)+(K*u)(\cdot))\le S(u(\cdot))$$

holds, and moreover, if $(K*u)(\cdot)\in C^2(\mathbb{T})$, then

$$Z_2(y(\cdot)+(K*u)(\cdot))\le S(u(\cdot)),$$

where $Z_2(u(\cdot))$ is the number of zeros of $u(\cdot)$ when multiple zeros are counted twice and intervals on which the function vanishes identically are discarded. Assume as before that $\alpha_k^2+\beta_k^2\ne 0$, $k\notin Q$, where α_k and β_k are the Fourier coefficients of $K(\cdot)$.

Suppose that

$$K_j(\cdot)\in\mathcal{K}(Q_j,\delta_j),\quad j=1,\dots,k,\qquad K_0(\cdot)\in\mathcal{K}(Q_0).\tag{17}$$

Set

$$K(\cdot)=(K_k*\ldots*K_1*K_0)(\cdot),\quad Q=\bigcup_{j=0}^{k}Q_j,\quad \delta=\min_{1\le j\le k}\delta_j.\tag{18}$$

Consider some particular cases of the classes $W^K_\infty(\mathbb{T},Q)$ for such kernels.

1. Let $k=0$, $Q_0=\emptyset$. The kernels from the set $\mathcal{K}(\emptyset)$ are called *cyclic variation diminishing kernels* or CVD-kernels. The corresponding classes

$$W^K_\infty(\mathbb{T},\emptyset)=\{\,x(\cdot)\mid x(\cdot)=(K*u)(\cdot),\ \|u(\cdot)\|_{L_\infty(\mathbb{T})}\le 1\,\}$$

were studied in [20]. In particular, the kernel

$$K_\beta(t) = \frac{1}{2} + \sum_{m=1}^{\infty} \frac{\cos mt}{\cosh m\beta}$$

is a CVD-kernel and the corresponding class $W_\infty^{K_\beta}(\mathbb{T}, \emptyset)$ coincides with the class h_∞^β which is the set of real, 2π-periodic functions $f(\cdot)$ that can be analytically continued to the strip $S_\beta = \{z \in \mathbb{C} \mid |\operatorname{Im} z| < \beta\}$ so that $|\operatorname{Re} f(z)| \le 1$ in this strip.

2. Let $P(D)$ be a differential polynomial of degree r with constant real coefficients

$$P(D) = D^r + a_{r-1}D^{r-1} + \ldots + a_0, \quad D = \frac{d}{dt}.$$

Set

$$K_P(t) = \frac{1}{2} \sum_{\substack{m \in \mathbb{Z} \\ P(im) \ne 0}} \frac{e^{imt}}{P(im)}.$$

For $Q = \{m \in \mathbb{Z}_+ \mid P(im) = 0\}$ the class $W_\infty^{K_P}(\mathbb{T}, Q)$ coincides with the generalized Sobolev class which is the set of 2π-periodic functions $x(\cdot)$ with $(r-1)$st derivative absolutely continuous and satisfying the condition

$$\|(P(D)x)(\cdot)\|_{L_\infty(\mathbb{T})} \le 1.$$

In particular, for $P(D) = D^r$ this class coincides with the standard Sobolev class $W_\infty^r(\mathbb{T})$.

In the general case a polynomial $P(D)$ can be represented in the following form

$$P(D) = \prod_{j=1}^{k} P_j(D), \tag{19}$$

where $P_j(D)$ are differential polynomials with real coefficients of degrees at most 2. It follows from [16] that $K_{P_j}(\cdot) \in \mathcal{K}(Q_j, \delta_j)$ where

$$Q_j = \{\, m \in \mathbb{Z}_+ \mid P_j(im) = 0 \,\}, \qquad \delta_j = \pi / h(P_j(\cdot)), \tag{20}$$

and $h(P_j(\cdot))$ is the largest imaginary part of the zeros of the polynomial $P_j(\cdot)$.

3. For a differential polynomial $P(D)$ with real coefficients let $h_{\infty,\beta}^P$ be the class of 2π-periodic real-valued functions $f(\cdot)$ that can be analytically continued to the strip S_β satisfying the condition $|\operatorname{Re}(P(D)f)(z)| \le 1$ for all $z \in S_\beta$. Then $h_{\infty,\beta}^P = W_\infty^{K_P}(\mathbb{T}, Q)$ where (using the notation (19), (20))

$$K_P(\cdot) = (K_{P_k} * \ldots * K_{P_1} * K_\beta)(\cdot), \quad Q = \bigcup_{j=1}^{k} Q_j.$$

Set

$$h_n(t) := \operatorname{sign} \sin nt.$$

Lemma 1 *Assume that a kernel $K(\cdot)$ satisfies* (17) *and* (18)*. Then for all*

$$n > \max\{ \sup_{j \in Q} j,\ 2\pi/\delta \} \tag{21}$$

$K(\cdot)$ satisfies the Favard γ-property where γ defined by the condition

$$(K * h_n)(\gamma) = -\|(K * h_n)(\cdot)\|_{L_\infty(\mathbb{T})}. \tag{22}$$

Proof: Consider the problem of optimal recovery of $x(\cdot)$ at the zero on the class $W_p^K(\mathbb{T},Q)$ from the Fourier coefficients of this function $a_0,\dots,a_{n-1},b_1,\dots,b_{n-1}$. The associated problem has the form

$$x(0)\to\max,\quad a_0=\ldots=a_{n-1}=b_1=\ldots=b_{n-1}=0,\quad x(\cdot)\in W_\infty^K(\mathbb{T},Q).$$

It follows from [19] that under the conditions (21), (22) the function $(K*h_n)(\cdot-\gamma)$ is a solution of this problem. Assume for definiteness that $0\in Q$. Similarly to the proof of Theorem 2 we obtain that there exist $\widehat{c}_k$, $\widehat{d}_k$, $k\in Q$, and $\widehat{\mu}_k$, $\widehat{\nu}_k$, $k\in Q'$, such that the function

$$\frac{1}{\pi}\int_{\mathbb{T}}\Bigg(-K(-t)+\widehat{c}_0+\sum_{k\in Q\setminus\{0\}}(\widehat{c}_k\cos kt+\widehat{d}_k\sin kt)$$
$$+\sum_{k\in Q'}((\widehat{\mu}_k\alpha_k+\widehat{\nu}_k\beta_k)\cos kt+(\widehat{\nu}_k\alpha_k-\widehat{\mu}_k\beta_k)\sin kt)\Bigg)u(t)\,dt$$

attains the absolute minimum on the unit ball of $L_\infty(\mathbb{T})$ at the point $\widehat{u}(\cdot)=h_n(\cdot+\gamma)$. Denoting by $L(\cdot)$ the multiplier preceeding $u(\cdot)$ under the integral sign we obtain that $\widehat{u}(\cdot)=-\operatorname{sign}L(\cdot)$. Changing the variable t on $-t$ we get the existence of polynomial $P(\cdot)\in\mathcal{T}_{n-1}$ such that $\operatorname{sign}(K(t)-P(t))=-h_n(t+\gamma)$. ■

Thus for the classes $W_\infty^K(\mathbb{T},Q)$ with kernels $K(\cdot)$ satisfying the conditions (17) and (18) the assertions of Theorem 2 (for $p=\infty$) and Corollary 1 hold.

We list several well-known results which are particular cases of the assertions proved here. The inequality (11) obtained by Favard [7] was used in [8] (and also independently by Akhiezer and M. Krein [3]) to prove the equality

$$d(W_\infty^r(\mathbb{T}),\mathcal{T}_{n-1},C(\mathbb{T}))=\frac{K_r}{n^r},\quad r\in\mathbb{N}. \tag{23}$$

The class $W_\infty^r(\mathbb{T})$ is defined by the convolution with the Bernoulli kernel which satisfies the Favard γ-property for $\gamma=0$ if r is odd and $\gamma=\pi/(2n)$ if r is even. For this class the problem of optimal recovery from Fourier coefficients was solved by Bojanov [5] who proved that

$$E(x(\theta),W_\infty^r(\mathbb{T}),\mathrm{Four}_n)=\frac{K_r}{n^r}.$$

The result of Favard–Akhiezer–Krein (23) was developed in several directions. Partially this was elucidated in [1]. Note the own result of Akhiezer [2]

$$d(h_\infty^\beta,\mathcal{T}_{n-1},C(\mathbb{T}))=\frac{4}{\pi}\sum_{m=0}^\infty\frac{(-1)^m}{(2m+1)\cosh(2m+1)n\beta}$$

and M. Krein [11]

$$d(\Gamma_\infty^\rho,\mathcal{T}_{n-1},C(\mathbb{T}))=\frac{4}{\pi}\arctan\rho^n,$$

where Γ_∞^ρ is the class of functions $x(\cdot)$ represented in the form $x(\cdot)=u(\rho,\cdot)$, $0<\rho<1$, with functions $u(r,t)$, $0\le r<1$, $t\in\mathbb{T}$, harmonic in the unit ball and satisfying there the condition $|u(r,t)|\le 1$. The class Γ_∞^ρ coincides with the class $W_\infty^{P_\rho}(\mathbb{T},\emptyset)$ where

$$P_\rho(t)=\frac{1}{2}\,\frac{1-\rho^2}{1-2\rho\cos t+\rho^2}$$

is the Poisson kernel which satisfies the Favard $\pi/(2n)$-property.

The problem of generalization of the Favard–Akhiezer–Krein result for fractional r was open for a long time. This problem was solved by Dzyadyk [6] and Sun Yongsheng [22], [23]. It turns out that the Bernoulli kernel with fractional $r \geq 1$ also satisfies the Favard γ-property with γ defined by the condition

$$\sum_{m=0}^{\infty} \frac{\cos((2m+1)\gamma - \pi r/2)}{(2m+1)^r} = 0.$$

The following result holds:

$$d(W_\infty^r(\mathbb{T}), \mathcal{T}_{n-1}, C(\mathbb{T})) = \frac{4}{\pi n^r} \left| \sum_{m=0}^{\infty} \frac{\sin((2m+1)\gamma - \pi r/2)}{(2m+1)^r} \right|.$$

4 The Hardy spaces

Now we consider optimal recovery problems for classes of analytic functions. First we give some definitions. Denote by $\mathcal{H}_p(D)$, $1 \leq p \leq \infty$, the Hardy space, i.e., the set of functions $f(\cdot)$ analytic in the unit disk $D = \{z \in \mathbb{C} \mid |z| < 1\}$ and satisfying

$$\sup_{0<r<1} \frac{1}{2\pi} \int_{\mathbb{T}} |f(re^{it})|^p \, dt = A_p^p < \infty, \quad 1 \leq p < \infty,$$
$$\sup_{z \in D} |f(z)| = A_\infty < \infty, \quad p = \infty.$$

Every function $f(\cdot) \in \mathcal{H}_p(D)$ associates with the unique function $\tilde{f}(\cdot) \in L_p(\partial D)$ (∂D is the boundary of D) by the rule:

$$\tilde{f}\left(e^{it}\right) = \lim_{r \to 1} f\left(re^{it}\right)$$

for almost all t. Moreover, $\|\tilde{f}(\cdot)\|_{L_p(\partial D)} = A_p$ and for all $z \in D$ the Cauchy formula

$$f^{(k)}(z) = \frac{k!}{2\pi i} \int_{\partial D} \frac{\tilde{f}(\zeta)}{(\zeta - z)^{k+1}} \, d\zeta, \quad k \in \mathbb{Z}_+, \tag{24}$$

holds.

The subset of $L_p(\partial D)$ that consists of all such functions $\tilde{f}(\cdot)$ is the set of those functions from $L_p(\partial D)$ for which

$$\frac{1}{2\pi i} \int_{\partial D} \tilde{f}(\zeta)\zeta^k \, d\zeta = 0, \quad k \in \mathbb{N}. \tag{25}$$

For simplicity we shall use the same notation for $f(\cdot) \in \mathcal{H}_p(D)$ and its boundary values.

The set

$$H_p(D) = \{f(\cdot) \in \mathcal{H}_p(D) \mid \|f(\cdot)\|_{L_p(\partial D)} \leq 1\}$$

we call the *Hardy class*.

We consider the following recovery problem: recover a value of $f(\cdot)$ at a point $\tau \in D$ on the Hardy class $H_p(D)$ from the information

$$f(z_1), f'(z_1), \ldots, f^{(k_1-1)}(z_1), \ldots, f(z_n), f'(z_n), \ldots, f^{(k_n-1)}(z_n),$$

where $z_1, \dots, z_n$ are distinct points from the disk D. We are interested in an optimal method of recovery.

In accordance with the general notation here $X = \mathcal{H}_p(D)$, $Y = \mathbb{C}^N$, $N = k_1 + \ldots + k_n$, $\operatorname{Re}\langle x', f(\cdot)\rangle = \operatorname{Re} f(\tau)$, $A = H_p(D)$, and $F\colon A \to \mathbb{C}^N$,

$$Ff(\cdot) = \left(f(z_1), f'(z_1), \dots, f^{(k_1-1)}(z_1), \dots, f(z_n), f'(z_n), \dots, f^{(k_n-1)}(z_n)\right).$$

The associated problem has the form

$$\operatorname{Re} f(\tau) \to \max, \quad f^{(k)}(z_j) = 0,\ j = 1, \dots n,\ k = 0, 1, \dots, k_j - 1, \quad f(\cdot) \in H_p(D). \tag{26}$$

In contrast to the real case where a solution of associated problem and optimal method were obtained simultaneously, here we first find a solution of (26) directly and then use it to obtain an optimal recovery method.

Since every function $f(\cdot) \in H_p(D)$ for which

$$f(z_1) = \ldots = f^{(k_1-1)}(z_1) = \ldots = f(z_n) = \ldots = f^{(k_n-1)}(z_n) = 0$$

may be represented in the form $f(z) = B(z)g(z)$ where $g(\cdot) \in H_p(D)$ and

$$B(z) = \prod_{j=1}^{n} \left(\frac{z - z_j}{1 - \overline{z}_j z}\right)^{k_j},$$

it suffices to find the extremum in the problem

$$\operatorname{Re} g(\tau) \to \max, \quad g(\cdot) \in H_p(D).$$

Evidently, for $p = \infty$ the function $\widehat{g}(z) \equiv 1$ is extremal. We prove that for $1 \le p < \infty$ the function $\widehat{g}(z) = (1 - |\tau|^2)^{1/p}(1 - \overline{\tau} z)^{-2/p}$ is extremal. By the residue theorem we have

$$g(\tau) = \frac{1}{2\pi i} \int_{\partial D} \frac{g(\zeta)(1 - |\tau|^2)^{(p-2)/p}}{(\zeta - \tau)(1 - \overline{\tau}\zeta)^{(p-2)/p}}\, d\zeta = \frac{1}{2\pi} \int_{\mathbb{T}} \frac{g(e^{it})(1 - |\tau|^2)^{(p-2)/p}}{(1 - \tau e^{-it})(1 - \overline{\tau} e^{it})^{(p-2)/p}}\, dt. \tag{27}$$

Applying the Hölder inequality to the last integral we obtain that

$$|g(\tau)| \le (1 - |\tau|^2)^{-1/p} \tag{28}$$

for all $g(\cdot) \in H_p(D)$. Moreover, it follows from (27) that

$$(1 - |\tau|^2)^{-1/p} = \widehat{g}(\tau) = \|\widehat{g}(\cdot)\|^p_{L_p(\partial D)} (1 - |\tau|^2)^{-1/p}.$$

Thus, $\widehat{g}(\cdot) \in H_p(\partial D)$ and for this function (28) turns to equality. Consequently, the function

$$\widehat{f}(z) = e^{-i \arg B(\tau)} B(z) \widehat{g}(z)$$

is extremal in (26).

In accordance with (24) and (25) the problem (26) may be rewritten as follows

$$\operatorname{Re} \frac{1}{2\pi i} \int_{\partial D} \frac{f(\zeta)}{\zeta - \tau}\, d\zeta \to \max, \quad \frac{k!}{2\pi i} \int_{\partial D} \frac{f(\zeta)}{(\zeta - z_j)^{k+1}}\, d\zeta = 0, \quad j = 1, \dots, n,$$
$$k = 0, 1, \dots, k_j - 1, \quad \frac{1}{2\pi i} \int_{\partial D} f(\zeta)\zeta^m\, d\zeta = 0, \quad m \in \mathbb{Z}_+, \quad \|f(\cdot)\|_{L_p(\partial D)} \le 1. \tag{29}$$

It is a convex problem. We apply the Lagrange principle to it noting that the set A (see (2)) is defined here by a countable number of equalities and one inequality. We include the constraints of equality type in the Lagrange function by "natural" way without giving more precise descriptions since, as it was said, we apply the Lagrange principle heuristically. We note only that an optimal recovery method is defined by multipliers at the constraints related to the information operator (see Remark 2).

The Lagrange function of the problem (29) is

$$\mathcal{L} = \operatorname{Re}\frac{1}{2\pi}\int_{\mathbb{T}}\left(\frac{-1}{e^{it}-\tau}+\sum_{j=1}^{n}\sum_{k=0}^{k_j-1}\frac{\mu_{jk}k!}{(e^{it}-z_j)^{k+1}}+\sum_{m\geq 0}\lambda_m e^{imt}\right)f(e^{it})e^{it}\,dt,$$

where $\mu_{jk},\lambda_m\in\mathbb{C}$, $1\leq j\leq n$, $0\leq k\leq k_j-1$, $m\geq 0$. By the Lagrange principle there exist such $\widehat{\mu}_{jk},\widehat{\lambda}_m\in\mathbb{C}$, $1\leq j\leq n$, $0\leq k\leq k_j-1$, $m\geq 0$, that $\mathcal{L}$ attains its minimum at the point $\widehat{f}(\cdot)$ on the set $\{f(\cdot)\in L_p(\partial D)\mid \|f(\cdot)\|_{L_p(\partial D)}\leq 1\}$. Hence it follows that for $z=e^{it}$ and $1<p<\infty$

$$L(z)=C\overline{z}\overline{\widehat{f}(z)}|\widehat{f}(z)|^{p-2}=\frac{C_1\overline{B(z)}}{(z-\tau)(1-\overline{\tau}z)^{(p-2)/p}}, \tag{30}$$

where $L(z)$ is the expression in parentheses under the integral sign in the Lagrange function. Since $\overline{B(z)}=B^{-1}(z)$ for $z=e^{it}$, we have

$$-\frac{1}{z-\tau}+\sum_{j=1}^{n}\sum_{k=0}^{k_j-1}\frac{\widehat{\mu}_{jk}k!}{(z-z_j)^{k+1}}+\sum_{m\geq 0}\widehat{\lambda}_m z^m=\frac{C_1}{B(z)(z-\tau)(1-\overline{\tau}z)^{(p-2)/p}}. \tag{31}$$

The function in the right-hand side of this equality is analytic in the disk D with the exception of points $\tau,z_1,\ldots,z_n$ where it has poles. If we multiply the both sides of (31) by $z-\tau$ and substitute $z=\tau$, we get $C_1=-B(\tau)(1-|\tau|^2)^{(p-2)/p}$. In a similar way we obtain

$$\widehat{\mu}_{jk}=\frac{B(\tau)(1-|\tau|^2)^{(p-2)/p}}{k!(k_j-k-1)!}\left(\frac{(1-\overline{z}_jz)^{k_j}}{\omega_j(z)(\tau-z)(1-\overline{\tau}z)^{(p-2)/p}}\right)^{(k_j-k-1)}\Bigg|_{z=z_j},$$

where

$$\omega_j(z)=\prod_{\substack{m=1\\ m\neq j}}^{n}\left(\frac{z-z_m}{1-\overline{z}_mz}\right)^{k_m}.$$

We prove now that the method defined by $\widehat{\mu}_{jk}$ (which are well-defined for all $1\leq p\leq\infty$; for $p=\infty$ all expressions involving p are understood as the limit values as $p\to\infty$) is optimal. Indeed, for all $1\leq p\leq\infty$ the equality (31) holds with some $\widehat{\lambda}_m$, $m\geq 0$ (we do not need explicit expressions for them). Let $1\leq p<\infty$. Then for all $f\in H_p(D)$ taking into account the last equality of (30) and applying the Hölder inequality we obtain

$$\begin{aligned}
&\left|f(\tau)-\sum_{j=1}^{n}\sum_{k=0}^{k_j-1}\widehat{\mu}_{jk}f^{(k)}(z_j)\right|\\
&\qquad=\left|\frac{1}{2\pi i}\int_{\partial D}\left(\sum_{m\geq 0}\widehat{\lambda}_m z^m+\frac{B(\tau)}{(1-|\tau|^2)^{1/p}}\overline{z}\overline{\widehat{f}(z)}|\widehat{f}(z)|^{p-2}\right)f(z)\,dz\right|\\
&\qquad\qquad=\frac{|B(\tau)|}{(1-|\tau|^2)^{1/p}}\frac{1}{2\pi}\left|\int_{\mathbb{T}}\overline{\widehat{f}(e^{it})}|\widehat{f}(e^{it})|^{p-2}f(e^{it})\,dt\right|\\
&\qquad\qquad\qquad\leq\frac{|B(\tau)|}{(1-|\tau|^2)^{1/p}}\|\widehat{f}(\cdot)\|_{L_p(\partial D)}^{p-1}\|f(\cdot)\|_{L_p(\partial D)}\leq|\widehat{f}(\tau)|.
\end{aligned}$$

For $p=\infty$ using the fact that the integral of the Poisson kernel equals 1 we have

$$\left|f(\tau)-\sum_{j=1}^{n}\sum_{k=0}^{k_j-1}\widehat{\mu}_{jk}f^{(k)}(z_j)\right| = \left|\frac{1}{2\pi i}\int_{\partial D}\left(\sum_{m\geq 0}\widehat{\lambda}_m z^m + \frac{B(\tau)(1-|\tau|^2)}{B(z)(z-\tau)(1-\overline{\tau}z)}\right)f(z)\,dz\right|$$
$$= |B(\tau)|\frac{1}{2\pi}\left|\int_{\mathbb{T}}\frac{1-|\tau|^2}{|1-\overline{\tau}e^{it}|^2}f(e^{it})\,dt\right| \leq |B(\tau)| = |\widehat{f}(\tau)|.$$

It follows from Theorem 1 that the error of optimal recovery equals $|\widehat{f}(\tau)|$. Thus it is proved that the method

$$f(\tau)\approx\sum_{j=1}^{n}\sum_{k=0}^{k_j-1}\widehat{\mu}_{jk}f^{(k)}(z_j)$$

is optimal. In particular, for one point $z_1=0$ with the multiplicity n (that is we consider the problem of optimal recovery from Taylor coefficients at the zero) for $p=\infty$ we have

$$f(\tau)\approx\sum_{j=0}^{n-1}\tau^j(1-|\tau|^{2(n-j)})\frac{f^{(j)}(0)}{j!}. \tag{32}$$

The optimality of these methods was obtained, without using the Lagrange principle, in [17], [18] ($p=\infty$), and [9] ($1\leq p<\infty$).

5 The Hardy–Sobolev spaces

For $p=\infty$ we can obtain a more general result rather than formula (32).

Denote by $H^r_\infty(D)$ the *Hardy–Sobolev* class which is the set of all functions analytic in the unit disk D for which $|f^{(r)}(z)|\leq 1$, $z\in D$. Consider the problem of optimal recovery of $f(\cdot)\in H^r_\infty(D)$ at a point τ (without loss of generality we may assume that $\tau\in(0,1)$) from the information $f(0), f'(0),\dots,f^{(n+r-1)}(0)$.

From the equality

$$f(z)=S_{r-1}(z)+\int_0^z\frac{(z-\xi)^{r-1}}{(r-1)!}f^{(r)}(\xi)\,d\xi,\quad S_{r-1}(z)=\sum_{k=0}^{r-1}\frac{f^{(k)}(0)}{k!}z^k,$$

and the Cauchy formula for $f^{(r)}(\xi)$ we have

$$f(z)=S_{r-1}(z)+\frac{1}{2\pi}\int_{\mathbb{T}}\sum_{k=0}^{\infty}\frac{k!}{(k+r)!}z^{k+r}e^{-ikt}f^{(r)}(e^{it})\,dt. \tag{33}$$

Thus the associated problem can be written as follows

$$\operatorname{Re}\left(S_{r-1}(\tau)+\frac{1}{2\pi}\int_{\mathbb{T}}\sum_{k=0}^{\infty}\frac{k!}{(k+r)!}\tau^{k+r}e^{-ikt}f^{(r)}(e^{it})\,dt\right)\to\max,$$
$$f(0)=\ldots=f^{(r-1)}(0)=0,\quad \frac{k!}{2\pi}\int_{\mathbb{T}}f^{(r)}(e^{it})e^{-ikt}\,dt=0,\ k=0,\dots,n-1,$$
$$\frac{1}{2\pi}\int_{\mathbb{T}}f^{(r)}(e^{it})e^{imt}\,dt=0,\ m\in\mathbb{N},\quad \|f^{(r)}(\cdot)\|_{L_\infty(\partial D)}\leq 1.$$

We prove that the function

$$\widehat{f}(z) = \frac{n!}{(n+r)!} z^{n+r}$$

is extremal in this problem. Assume that there exists a function $f_0(\cdot) \in H^r_\infty(D)$ for which $f_0(0) = \ldots = f^{(n+r-1)}(0) = 0$ and $|f_0(\tau)| > \widehat{f}(\tau)$. Without loss of generality we may assume that $f_0(\tau) > 0$. Since

$$g_0(z) = \frac{f_0(z) + \overline{f_0(\overline{z})}}{2}$$

has the same properties as $f_0(\cdot)$ and is real on the real axis, we assume from the very beginning that the function $\widehat{f}(\cdot)$ is real on the real axis. Put

$$F(z) := \widehat{f}(z) - \rho f_0(z), \quad \rho = \frac{\widehat{f}_0(\tau)}{f_0(\tau)}.$$

The function $F(\cdot)$ has at least $n+r+1$ zeros on the interval $(-1,1)$ taking into account multiplicities. Consequently, by Rolle's theorem $F^{(r)}(\cdot)$ has at least $n+1$ zeros on this interval. For $z \in \partial D$ we have

$$|\widehat{f}^{(r)}(z) - F^{(r)}(z)| = \rho |f_0^{(r)}(z)| \le \rho < 1 \equiv |\widehat{f}^{(r)}(z)|.$$

Since $\widehat{f}^{(r)}(\cdot)$ has exactly n zeros in D (counting multiplicities) Rouche's theorem implies that the function $F^{(r)}(\cdot)$ must have the same number of zeros. This contradiction proves that the function $\widehat{f}(\cdot)$ is extremal.

Now let us write out the Lagrange function of the considered problem

$$\mathcal{L} = \mathrm{Re}\left(-S_{r-1}(\tau) + \sum_{k=0}^{r-1} \mu_k f^{(k)}(0) \right.$$
$$\left. + \frac{1}{2\pi} \int_{\mathbb{T}} \left(-\sum_{k=0}^{\infty} \frac{k!}{(k+r)!} \tau^{k+r} e^{-ikt} + \sum_{k=0}^{n-1} \mu_{k+r} k! e^{-ikt} + \sum_{m=1}^{\infty} \lambda_m e^{imt} \right) f^{(r)}(e^{it})\, dt \right),$$

where $\mu_k, \lambda_m \in \mathbb{C}$, $0 \le k \le n+r-1$, $m \in \mathbb{N}$. According to the Lagrange principle there exist such $\widehat{\mu}_k, \widehat{\lambda}_m \in \mathbb{C}$ that $\mathcal{L}$ attains its minimum at the point $\widehat{f}(\cdot)$ on the set of functions for which $\|f^{(r)}(\cdot)\|_{L_\infty(\partial D)} \le 1$.

Clearly,

$$\widehat{\mu}_k = \frac{1}{k!} \tau^k, \quad k = 0, \ldots, r-1.$$

Denoting by $L(\cdot)$ the expression in parentheses under the integral sign we have

$$L(e^{it}) = -\overline{\widehat{f}^{(r)}(e^{it})} |L(e^{it})|.$$

Consider the Fourier-series expansion of $|L(e^{it})|$

$$|L(e^{it})| = \sum_{k=-\infty}^{\infty} a_k e^{ikt}.$$

Taking into account the fact that it is a real function we have $a_{-k} = \overline{a}_k$. Thus

$$-\sum_{k=0}^{\infty} \frac{k!}{(k+r)!} \tau^{k+r} e^{-ikt} + \sum_{k=0}^{n-1} \widehat{\mu}_{k+r} k! e^{-ikt} + \sum_{m=1}^{\infty} \widehat{\lambda}_m e^{imt}$$
$$= -e^{-int} \left(a_0 + \sum_{k=1}^{\infty} a_k e^{ikt} + \sum_{k=1}^{\infty} \overline{a}_k e^{-ikt} \right).$$

Consequently,

$$\overline{a}_k = \frac{(n+k)!}{(n+k+r)!}\tau^{n+k+r},$$
$$a_k = \frac{(n-k)!}{(n-k+r)!}\tau^{n-k+r} - (n-k)!\widehat{\mu}_{n-k+r}, \quad k = 1, \dots, n.$$

Hence we obtain

$$\widehat{\mu}_{n+r} = \frac{\tau^{m+r}}{(m+r)!}\left(1 - \frac{(m+r)!}{m!}\frac{(2n-m)!}{(2n-m+r)!}\tau^{2(n-m)}\right).$$

By the direct verification (similar to the case described above) it can be proved that the method

$$f(\tau) \approx \sum_{k=0}^{r-1} \frac{f^{(k)}(0)}{k!}\tau^k + \sum_{k=r}^{n+r-1}\left(1 - \frac{k!}{(k-r)!}\frac{(2n+r-k)!}{(2n+2r-k)!}|\tau|^{2(n+r-k)}\right)\frac{f^{(k)}(0)}{k!}\tau^k \tag{34}$$

is optimal.

Thus we have proved the following theorem.

Theorem 3 (on optimal recovery from Taylor coefficients). *Let $r \in \mathbb{Z}_+$, $n \in \mathbb{Z}$, and $\tau \in D$. Then the method (34) is optimal method of recovery on the class $W^r H_\infty(D)$ from the information*

$$\mathrm{Tay}_{n+r}\, f(\cdot) = (f(0), f'(0), \dots, f^{(n+r-1)}(0)).$$

Moreover,

$$E(f(\tau), W^r H_\infty, \mathrm{Tay}_{n+r}) = \frac{n!}{(n+r)!}|\tau|^{n+r}.$$

For $r = 1$ this result was obtained by Newman (see [14, p. 42]).

6 Optimal recovery from the values at the equidistant system of points on a circle

Consider now the problem of optimal recovery of a value $f(\tau)$, $f(\cdot) \in H^1_\infty(D)$, $\tau \in D$, from the values $f(\tau_j)$, $j = 0, \dots, n-1$, where $\{\tau_j\}$ is the system of equidistant points on the circle of the radius $0 < \rho < 1$: $\tau_j = \rho e^{ij2\pi/n}$. From [10] it follows that the function

$$\widehat{f}(z) := \varepsilon\frac{z^n - \rho^n}{n}, \quad \varepsilon = e^{-i\arg(\tau^n - \rho^n)},$$

is extremal in the associated problem

$$\mathrm{Re}\, f(\tau) \to \max, \quad f(\tau_j) = 0, \quad j = 0, \dots, n-1, \quad f \in W^1 H_\infty(D).$$

In view of (33) this problem may be rewritten as follows

$$\mathrm{Re}\left(f(0) + \frac{1}{2\pi}\int_{\mathbb{T}} \sum_{k=0}^{\infty} \frac{\tau^{k+1}}{k+1} e^{-ikt} f'(e^{it})\, dt\right) \to \max,$$
$$f(0) + \frac{1}{2\pi}\int_{\mathbb{T}} \sum_{k=0}^{\infty} \frac{\tau_j^{k+1}}{k+1} e^{-ikt} f'(e^{it})\, dt = 0,\ j = 0, \dots, n-1,$$
$$\frac{1}{2\pi}\int_{\mathbb{T}} f'(e^{it}) e^{imt}\, dt = 0,\ m \in \mathbb{N}, \quad \|f'(\cdot)\|_{L_\infty(\partial D)} \le 1.$$

The Lagrange function of this problem is

$$\mathcal{L} = \mathrm{Re}\left(-f(0) + \sum_{j=0}^{n-1} \mu_j f(0)\right.$$
$$\left. + \frac{1}{2\pi}\int_{\mathbb{T}} \left(-\sum_{k=0}^{\infty} \frac{\tau^{k+1}}{k+1} e^{-ikt} + \sum_{j=0}^{n-1} \mu_j \sum_{k=0}^{\infty} \frac{\tau_j^{k+1}}{k+1} e^{-ikt} + \sum_{m=1}^{\infty} \lambda_m e^{imt}\right) f'(e^{it})\,dt\right),$$

where $\mu_j, \lambda_m \in \mathbb{C}$, $0 \le j \le n-1$, $m \in \mathbb{N}$. According to the Lagrange principle there exist such $\widehat{\mu}_j, \widehat{\lambda}_m \in \mathbb{C}$ that $\mathcal{L}$ attains its minimum at the point $\widehat{f}(\cdot)$ on the set of functions for which $\|f'(\cdot)\|_{L_\infty(\partial D)} \le 1$.

It is clear that

$$\sum_{j=0}^{n-1} \widehat{\mu}_j = 1.$$

Denoting by $L(\cdot)$ the expression in parentheses under the integral sign we have

$$L(e^{it}) = -\overline{\widehat{f}'(e^{it})}|L(e^{it})|.$$

Consider the Fourier-series expansion of $|L(e^{it})|$

$$|L(e^{it})| = \sum_{k=-\infty}^{\infty} a_k e^{ikt}.$$

Taking into account the fact that it is a real function we have $a_{-k} = \overline{a}_k$. Thus

$$\varepsilon e^{i(n-1)t}\left(\sum_{k=0}^{\infty} \nu_k e^{-ikt} - \sum_{m=1}^{\infty} \widehat{\lambda}_m e^{imt}\right) = a_0 + \sum_{k=1}^{\infty} a_k e^{ikt} + \sum_{k=1}^{\infty} \overline{a}_k e^{-ikt},$$

where

$$\nu_k = \frac{1}{k+1}\left(\tau^{k+1} - \sum_{j=0}^{n-1} \widehat{\mu}_j \tau_j^{k+1}\right).$$

Hence $a_k = \varepsilon \nu_{n-k-1}$, $\overline{a}_k = \varepsilon \nu_{n+k-1}$, $k = 1, \dots, n-1$.

Assume for simplicity that $\tau \in (-1, 1)$. Put $\alpha = \tau/\rho$. Then

$$\frac{1}{n-k}\left(\alpha^{n-k} - \sum_{j=0}^{n-1} \widehat{\mu}_j \xi_j^{-k}\right) = \frac{\rho^{2k}}{n+k}\left(\alpha^{n+k} - \sum_{j=0}^{n-1} \overline{\widehat{\mu}}_j \overline{\xi}_j^{k}\right),$$

where $\xi_j = e^{ij2\pi/n}$, $j = 0, \dots, n-1$. Putting

$$b_k := \sum_{j=0}^{n-1} \widehat{\mu}_j \xi_j^{-k}$$

we have

$$\sum_{j=0}^{n-1} \overline{\widehat{\mu}}_j \overline{\xi}_j^{k} = \sum_{j=0}^{n-1} \overline{\widehat{\mu}}_j (\overline{\xi}_j)^{-(n-k)} = \overline{b}_{n-k}.$$

Thus we obtain the system

$$\frac{1}{n-k}(\alpha^{n-k} - b_k) = \frac{\rho^{2k}}{n+k}(\alpha^{n+k} - \overline{b}_{n-k}), \quad k = 1, \dots, n-1.$$

Hence

$$\begin{cases} \dfrac{1}{n-k} b_k - \dfrac{\rho^{2k}}{n+k} \overline{b}_{n-k} = \dfrac{\alpha^{n-k}}{n-k} - \dfrac{\alpha^{n+k}}{n+k} \rho^{2k}, \\ \dfrac{\rho^{2(n-k)}}{2n-k} b_k - \dfrac{1}{k} \overline{b}_{n-k} = \dfrac{\alpha^{2n-k}}{2n-k} \rho^{2(n-k)} - \dfrac{\alpha^k}{k}. \end{cases}$$

Consequently,

$$b_k = \frac{\alpha^{n-k}(1 - q_k \alpha^n \rho^{2n}) - p_k \alpha^k \rho^{2k}(1 - \alpha^n)}{1 - q_k \rho^{2n}}, \quad k = 1, \dots, n-1,$$

where

$$p_k = \frac{n-k}{n+k}, \quad q_k = \frac{k}{2n-k} p_k.$$

Since $b_0 = 1$ we have

$$\widehat{\mu}_j = \frac{1}{n} \sum_{k=0}^{n-1} b_k \xi_k^j.$$

As above, the direct verification leads to the fact that the method

$$f(\tau) \approx \frac{1}{n} \sum_{j=0}^{n-1} \left(\sum_{k=0}^{n-1} b_k \xi_k^j \right) f(\tau_j)$$

is optimal for the considered problem.

References

[1] Achiezer, N. I. [Akhiezer], *Theory of Approximation*, Nauka, Moscow, 1965; English transl. of 1st ed., Frederick Ungar, New York, 1956.

[2] Akhiezer, N. I., On the best approximation of analytic functions (in Russian), *Dokl. Akad. Nauk SSSR* **15** (1937), 241–244.

[3] Akhiezer, N. I. and M. G. Krein, On the best approximation by trigonometric sums of differentiable periodic functions (in Russian), *Dokl. Akad. Nauk SSSR* **15** (1937), 107–112.

[4] Arestov, V. V. (1989), Best recovery of operators and related problems, *Trudy Mat. Inst. Steklov.* **189** (1989), 3–20; English transl. in *Proc. Steklov Inst. Math.* **189** (1989).

[5] Bojanov, B. D., Best reconstruction of differentiable periodic functions from their Fourier coefficients, *Serdica B'lg. Mat. Spis.* **2** (1976), 300–304.

[6] Dzyadyk, V. K., On the best approximation for classes of periodic functions defined by kernels which are integrals of absolutly monotonic functions (in Russian), *Izv. Akad. Nauk SSSR. Ser. Mat.* **23** (1959), 933–950.

[7] Favard, J., Sur l'approximation des fonctions periodiques par des polinômes trigonométriques, *C. R. Acad. Sci.* **203** (1936), 1122–1124.

[8] Favard, J., Sur les meilleurs procédés d'approximation de certaines classes de fonctions par les polinômes trigonométriques, *Bull. Sci. Math.* **61** (1937), 209–224; 243–256.

[9] Fisher, S. D. and C. A. Micchelli, The n-width of sets of analytic functions, *Duke Math. J.* **47** (1980), 789–801.

[10] Horwitz, A., and D. J. Newman, An extremal problem for analytic functions with prescribed zeros and rth derivative in H^∞, *Trans. Amer. Math. Soc.* **295** (1986), 699–713.

[11] Krein, M. G. (1937), On the theory of the best approximation of periodic functions (in Russian), *Dokl. Akad. Nauk SSSR* **15** (1937), 245–251.

[12] Magaril-Il'yaev, G. G. and K. Yu. Osipenko, Optimal recovery of functionals based on inaccurate data, *Mat. Zametki* **50** (1991), 85–93; English transl. in *Math. Notes* **50** (1991), 1274–1279.

[13] Magaril-Il'yaev, G. G. and V. M. Tikhomirov, *Convex Analysis and its Applications* (in Russian), Editorial URSS, Moscow, 2000.

[14] Micchelli, C. A. and T. J. Rivlin, A survey of optimal recovery, in *Optimal Estimation in Approximation Theory*, C. A. Micchelli, T. J. Rivlin (eds.), Plenum Press, New York, 1977, 1–54.

[15] Micchelli, C. A. and T. J. Rivlin, Lectures on optimal recovery, Lecture Notes in Math. vol. 1129, Springer–Verlag, Berlin, 1985, 21–93.

[16] Nguen Tkhi Tkh'eu Khoa [Nguen Thi Thieu Hoa], Some extremal problems on classes of functions defined by linear differential operators, *Mat. Sb.* **180** (1989), 1355–1395; English transl. in *Math. USSR-Sb.* **68** (1989).

[17] Osipenko, K. Yu., Optimal interpolation of analytic functions, *Mat. Zametki* **12** (1972), 465–476; English transl. in *Math. Notes* **12** (1972), 712–719.

[18] Osipenko, K. Yu., Best approximation of analytic functions from their values at a finite number of points, *Mat. Zametki* **19** (1976), 29–40; English transl. in *Math. Notes* **19** (1976), 17–23.

[19] Osipenko, K. Yu., On the precise values of n-widths for classes defined by cyclic variation diminishing operators, *Mat. Sb.* **188** (1997), 113–126; English transl. in *Sbornic: Mathematics* **188** (1997), 1371–1383.

[20] Pinkus, A., *n-Widths in Approximation Theory*, Springer-Verlag, Berlin, 1985.

[21] Smolyak, S. A., On optimal restoration of functions and functionals of them (in Russian), dissertation, Moscow State University, 1965.

[22] Sun Yongsheng, On the best approximation of periodic differentiable functions by trigonometric polynomials (in Russian), *Izv. Akad. Nauk SSSR. Ser. Mat.* **23** (1959), 67–92.

[23] Sun Yongsheng, On the best approximation of periodic differentiable functions by trigonometric polynomials (in Russian), *Izv. Akad. Nauk SSSR. Ser. Mat.* **25** (1961), 143–152.

[24] Traub, J. F. and H. Woźniakowski, *A General Theory of Optimal Algorithms*, Academic Press, New York, 1980.

Georgii G. Magaril-Il'yaev georg@magaril.mccme.ru

Department of Higher Mathematics, Moscow State Institute of Radio Engineering, Electronics and Automation (Technology University), pr. Vernadskogo 78, 117454 Moscow, Russia

Konstantin Yu. Osipenko konst@osipenko.mccme.ru

Department of Higher Mathematics, MATI — Russian State Technological University, Orshanskaya 3, 121552 Moscow, Russia

Vladimir M. Tikhomirov tikh@tikhomir.mccme.ru

Department of General Control Problems, Moscow State University, Vorob'evy Gory, 119899 Moscow, Russia