**V. M. Tikhomirov**
Moscow State University, Moscow, Russia

## *External problems of approximation and recovery*

*Key words:* extremal problems, approximation, recovery, principles of extremum, extremal properties of polynomial, principles of convex analysis

# 1. Introduction

This report is connected with a plan to write (in collaboration with G. G. Magaril-Il'yaev and K. Yu. Osipenko) a paper devoted to the topic expressed in the title. Extremal problems of approximation and recovery are considered here as a test-field for the general theory of extremal problems.

**Extremal problems and their formalization.**

Extremal problems arising in mathematics, in natural science, or in practical enterprises, are stated initially without formulae, using the terminology of fields in which they arise. Thus if we want to investigate an extremal problem by mathematical tools, it is necessary to *formalize the problem* (i. e. to translate it into the mathematical language).

For the extremal problem:

FIND AN EXTREMUM (I. E. MAXIMUM OR MINIMUM) OF A FUNCTION $f : X \to \mathbb{R} \cup \{\pm\infty\}$, WITH RESPECT TO ALL $x \in X$ THAT BELONG TO A CONSTRAINT $C \subset X$

we use the following notation:

$$f(x) \to \text{extr}, \quad x \in C. \tag{$P_0$}$$

If $C = X$, the problem $(P_0)$ is called *the problem without constraints.*

The general theory of extremal problems is designed to create methods and principles for solution of concrete problems. In this report the theory of extremal problems will be applied to the problems of approximation and recovery.

**Examples of extremal problems of approximation theory and their formalizations**

**Example 1.1. (Polynomials of the least deviation.)** *Find a polynomial of degree $n$ with leading coefficient 1 of the least deviation from zero in the space of continuous functions on the segment* $[-1,1]$.

*Formalization:*

$$f(x) = \max_{t\in[-1,1]} |t^n - \sum_{k=1}^{n} x_k t^{k-1}| \to \min, \quad x = (x_1, \dots, x_n).$$

One can see that this is a finite-dimensional problem without constraints. Solution of this problem is the so called *Tchebyshev polynomial* $T_n(t) = 2^{n-1}\cos n \arccos t$ (P. Tchebychev, 1854).

**Example 1.2.** *Find the norm of derivative of trigonometric polynomials of degree $n$ if one considers them as a subspace of* $C([-\pi,\pi])$. This problem was solved by S. Bernstein (1912).

*Formalization.* The question can be reduced to the following extremal problem:

$$\dot{x}(0) \to \max, \ \ \|x(\cdot)\|_{C([-\pi.\pi])} \le 1, \ x(t) = \sum_{k=0}^{n}(x_k \cos kt + x_{k+n}\sin kt).$$

This is a finite-dimensional problem with inequality constraints.

## 2. Extremal problems of approximation and recovery

**Main classes of extremal problems in approximation theory** are the following:

**1) Approximation of an individual element by a fixed approximating set**

Let $X$ be a normed space, $A \subset X$ be an approximative set, $x \in X \setminus A$. The problem of approximation of the element $x$ by the set $A$ in the space $X$ is posed as follows:

$$\|x - \xi\|_X \to \min, \ \xi \in A. \qquad (P_1)$$

The value of the problem, i. e. *the distance from $x$ to $A$ in $X$* is denoted $d(x) = d(x, A, X)$; a solution $\widehat{\xi} \in \operatorname{argmin}(P_1)$ is called *an element of the best approximation.*

**Example 1.1.** *Problems on the best approximation of functions by polynomials in the uniform norm* are problems $(P_1)$ where $X = C([t_0, t_1])$, $A$ is the space $\mathcal{P}_n$ of algebraic polynomials of degree $n$, $x(\cdot) \in C([t_0, t_1])$ (Tchebyshev (1854))[1].

**2) Extremal proprieties of polynomials**

Let $X$ be a normed space, $L$ be a subspace of $X$ (in our case it will be a subspace of polynomials), $l$ be a linear functional on $L$, $\Lambda : L \to X$ be a linear operator. It is required to solve problems:

$$l(x) \to \max, \quad \|x\|_X \leq 1, \; x \in L, \tag{$P_{2a}$}$$

$$\|\Lambda x\|_X \to \max, \; \|x\|_X \leq 1, \; x \in L. \tag{$P_{2b}$}$$

**Example 2.1.** Let $p(\cdot) \in \mathcal{P}_n$ be an algebraical polynomial of degree $n$ not exceed unity on the interval $[-1, 1]$. What is the maximal value it may take at a point $\tau$ outside this interval? This problem is called the *extrapolation problem for polynomials.* (Tchebyshev (1886)).

This is the problem $(P_{2a})$, where $X = C([-1, 1])$, $L = \mathcal{P}_n$, $l(x(\cdot)) = x(\tau)$, $\tau \in \mathbb{R} \setminus [-1, 1]$.

**Example 2.2.** Let $x(\cdot) \in \mathcal{P}_n$ be a trigonometrical polynomial not exceed unity on the interval $[-\pi, \pi]$. What is the maximal value may take the norm of $\dot{x}(\cdot)$? This problem is called the *problem of inequality for derivatives for trigonometrical polynomials.* (Bernstein (1912)) [2]).

This is the problem $(P_{2b})$, where $X = C([-\pi, \pi])$, $L = \mathcal{T}_n$, $\Lambda(x(\cdot)) = \dot{x}(\cdot)$.

**3) Landau – Kolmogorov problems of inequalities for derivatives of smooth functions on the line $\mathbb{R}$ or the half-line $\mathbb{R}_+$**

Let $T$ be $\mathbb{R}$ or $\mathbb{R}_+$. The problem of inequalities for derivatives consists of finding the least constant $K$ in the relation

$$\|x^{(k)}\|_{L_q(T)} \leq K \|x\|^{\alpha}_{L_p(T)} \|x^{(n)}\|^{1-\alpha}_{L_r(T)}, \tag{$P_3$}$$

where $n \in \mathbb{N}$, $k \in \mathbb{N}$ or $0$, $k < n$, $1 \leq p, q, r \leq \infty$, $\alpha = \frac{n-k-r^{-1}+q^{-1}}{n-r^{-1}+p^{-1}}$. Denote the best constant by $K_T(k, n, p, q, r)$.

**Example 3.1.** If (in $(P_3)$) $T = \mathbb{R}_+$, $k = 1$, $n = 2$, $p = q = r = \infty$ , then $K_{\mathbb{R}_+}(1, 2, \infty, \infty, \infty) = 2$ (*Landau*(1913))[3]

**Example 3.2.** If (in $(P_3)$) $T = \mathbb{R}, 1 \leq k \leq n-1, n \geq 2, p = q = r = \infty$, then $K_{\mathbb{R}}(k, n, \infty, \infty, \infty) = \frac{K_{n-k}}{K_n^{(n-k)/n}}$, where $K_r = \frac{4}{\pi} \sum\limits_{j \in \mathbb{N}} \frac{(-1)^{(i+1)(r+1)}}{(2j-1)^{(r+1)}}$ (Kolmogorov 1938) [4].

**4) Problems on deviation of functional classes from polynomials**

Let $X$ be a normed space, $A \subset X$ be an approximative set, $C \subset X$ be a class of elements. Consider the problem:

$$d(x, A, X) \to \max,\ x \in C. \tag{$P_4$}$$

The value of the problem is denoted by $d(C, A, X)$. It is called a *deviation of $C$ from $A$.*

**Example 4.1.** If (in $(P_4)$) $X = C(\mathbb{T})$, $C = W^r_\infty(\mathbb{R})$, $A = \mathcal{T}_{n-1}$, we obtain the problem of finding of deviation of Sobolev class $W^r_\infty(\mathbb{R})$ from the space of trigonometrical polynomials in the space $C(\mathbb{R})$: $d(W^r_\infty(\mathbb{R}), \mathcal{T}_{n-1}, C(\mathbb{T})) = \frac{K_r}{n^r}$ (Favard (1936))[5].

**5) Problems of the best tool of approximation**

Let $X$ be a normed space, $\mathcal{A} = \{A\}$ be a class of approximative sets, $C \subset X$ be a class of elements. Consider the problem:

$$d(C, A, X) \to \min,\ A \in \mathcal{A}. \tag{$P_5$}$$

If $\mathcal{A} = \{L_n\}$ is the set of all $n$-dimensions subspaces of $X$, then the value of the problem $(P_5)$ is denoted $d_n(C, X)$ and is called *Kolmogorov $n$-width* [6].

**Example 5.1.** Kolmogorov $n$-widths of Sobolev class $W^r_2(\mathbb{T})$ in $L_2(\mathbb{T})$: $d_{2k}(W^r_2(\mathbb{T}), L_2(\mathbb{T})) = d_{2k-1}(W^r_2(\mathbb{T}), L_2(\mathbb{T})) = \frac{1}{k^r},\ k \in \mathbb{N}$ [6].

**6) Extremal problems of recovery** [7]

Let $C$ be a class of elements, $(Z, d)$ be a metric space and $f : C \to Z$. It is required to "recover" an element $f(x)$ (or the whole mapping $f$) from a certain "information" $y \in I(x)$ about $x$, where $I$ is a multivalued information operator $I : C \to Y$ (where $Y$ is some set). A mapping $m\colon I(C) \to Z$ is called *a method of recovery.* We denote the problem of recovery of $f$ on $C$ from the information $I$ by $(f, C, I)$. The simplest problem of such type is the problem of finding the quantity

$$e(f, C, I, m) := \sup_{x\in C, y\in I(x)} d(f(x), m(y)), \tag{1}$$

which is called the *error of this method of recovery.* The value $E(f, C, I) = \inf\limits_{m\colon I(C)\to Z} e(f(x), C, I, m)$ is called *the error of the problem* $(f, C, I)$. Any method $\widehat{m}$ such that $e(f(x), C, I, \widehat{m}) = E(f, C, I)$ is called *an optimal recovery method* , and we write in this case $f(x) \simeq \widehat{m}(y)(y \in I(x))$.

## 3. On principles of the theory of extremum

### A) The Lagrange principle for necessary conditions

*Here we only sketch these procedures and it will be easy to apply them, but one can reduce them to this general principle: If a function of several variables should be maximum or minimum, and there are between these variables one or several equations, then it will suffice to add to the proposed function the functions that should be zero, each multiplied by an undetermined quantity and then to look for the maximum or the minimum as if the variables were independent; the equations that one will find, combined with the given equations, will serve to determine all the unknowns.*

J.-L. Lagrange [8]

To solve concrete problems we will further use a single general idea which we call the *Lagrange principle for necessary conditions in the theory of extrema.*

I formulate it as follows: **to solve an extremal problem with constrains it is reasonable to construct the Lagrange function of the problem, and then to write down the necessary conditions in the similar problem of the extremum of the Lagrange function "as if the variables were independent" (in Lagrange's own words) and finally investigate the relations which were obtained.**

The fruitfulness of this principle as a tool for solving of concrete problems will be repeatedly demonstrated further.

### B) On some principles and phenomena in Convex Analysis. (See [9])

Convexity plays an important role in the theory of extremal problems. One of the most important special features of convexity is the *duality principle*, according to which *every convex set, function or problem has two descriptions: a primal one in the origin space and a dual one in the conjugate space.*

First of all we illustrate the principle of duality on an example of functions. The dual object for a convex function $f : X \to \mathbb{R}\cup\{\infty\}$ is its *conjugate function* $f^*(x^*) = \sup\{x \in X \mid \langle x^*, x\rangle - f(x)\}$. $f^*(x^*) = \sup\{x \in X \mid \langle x^*, x\rangle - f(x)\}$. The function $f^{**}(x) = \sup\{x^* \in X^* \mid \langle x^*, x\rangle - f^*(x^*)\}$ is called *the second conjugate of f.*

The theorem of duality of convex functions can be formulated in the following form:

**Fenchel – Moreau theorem**. *Let* $f : X \to \mathbb{R} \cup \{+\infty\}$*; then* $f = f^{**}$ *iff* $f$ *is a convex closed function.*

Let $f$ be a convex function on $X$. The set $\partial f(\widehat{x}) = \{x^* \in X^* \mid f(x) - f(\widehat{x}) \geq \langle x^*, x - \widehat{x}\rangle\}$ is called the subdifferential of $f$ at $\widehat{x}$. The following formulae of subdifferential calculus hold:

**1. Moreau – Rockafellar formula.** Let $f_1 : X \to \overline{\mathbb{R}}$, be a convex function continuous at a point $x \in X$ where $|f_2(x)| < \infty$. Then $\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x)$.

**2. Dubovitskii – Milyutin formula**. Let $f_i : X \to \overline{\mathbb{R}}$, $i = 1, 2$, be convex functions continuous at a point $x \in X$ and $f_1(x) = f_2(x)$. Then $\partial \max(f_1, f_2)(x) = \mathrm{co}(\partial f_1(x) \cup \partial f_2(x))$.

The following phenomenon of convexity plays an important role in the theory of extrema: *integration has a deep connection with convexity*.

In finite dimensional case this phenomenon can be illustrated as follows:

**Lyapounov theorem**. *Let* $\Delta$ *be a segment in* $\mathbb{R}$, $p(\cdot) = (p_1(\cdot), \dots, p_n(\cdot))$ *be an integrable vector-function. Then the set* $M = \{x \in \mathbb{R}^n \mid \int_A p(t)dt,\ A \in \mathcal{A}\}$, *where* $\mathcal{A}$ *is the* $\sigma$*-algebra of all Lebesgue measurable sets, is a convex compact in* $\mathbb{R}^n$*.* (See [10]).

Another phenomenon of convexity is a phenomenon of "cleaning": *searching for minimax of a family of finite dimensional convex functions, it is possible to restrict oneself by a finite subfamily of the functions.*

This phenomenon has the following exact formulation:

**Theorem of V. Levin**. *Let* $T$ *be a compact topological space and* $F : T \times \mathbb{R}^n \to \mathbb{R}$ *be a mapping such that* $F(\cdot, x)$ *is an upper semicontinuous function for every* $x$ *and* $F(t, \cdot)$ *be convex for every* $t \in T$ *and* $m := \inf_x \max_{t \in T} F(t, x) < \infty$. *Then there exists a number* $r \in \mathbb{N}$, $r \leq n + 1$ *and* $r$ *points* $\{\tau_i\}_{i=1}^r, \tau_i \in T$ *such that* $m = \inf_x \max_{1 \leq i \leq r} F(\tau_i, x)$ (*decomposition theorem*). (see [9]).

## 4. Application of the general theory to solution of concrete problems of approximation and recovery

In all cases we realize the following plan of investigation:

1. Formulation of a problem
2. Application of principles (Lagrange, duality or cleaning)

3. Investigation of the relations obtained in 2.

4. Formulation of the final result

**1. Tchebyshev problem on the best approximation of functions by polynomials in the uniform norm.**

*1. Formalization*:

$$f(x(\cdot)) \to \min, \quad f(x) = \max_{t\in[a,b]} |x(t) - \sum_{0\le k\le n+1} x_k t^{k-1}|, \ x \in \mathbb{R}^{n+2}. \qquad (1)$$

This is a convex problem without constraints.

*2. Application of the cleaning principle.* As it follows from the cleaning principle, if $\widehat{x} = (\widehat{x}_1, \dots, \widehat{x}_{n+2}) \leftrightarrow \widehat{p}(t) = \sum_{0\le k\le n+1} \widehat{x}_k t^{k-1}$ is a solution of the problem and $d$ is its value, we can restrict ourselves by the problem of minimization of $g(x) = \max_{1\le i\le r} |x(\tau_i) - \sum_{0\le k\le n+1} x_k \tau_i^{k-1}|, \ r \le n+2$ with $|x(\tau_i) - \widehat{p}(\tau_i)| = d$.

*3. Investigation.* From the Fermat theorem for convex functions it follows: $0 \in \partial f(\widehat{x})$. From the formula of Dubovitskii – Milyutin we obtain that there exist $\alpha_i \ge 0, \ \sum_{i=1}^{r} \alpha_i = 1$ such that $\sum_{i=1}^{r} \alpha_i \mathrm{sgn}(x(\tau_i) - \widehat{p}(\tau_i))\tau_i^j = 0, \ 0 \le j \le n$. From this immediately follows that $r = n+2$ and $((x(\tau_i) - \widehat{p}(\tau_i))(x(\tau_{i+1}) - \widehat{p}(\tau_{i+1})) < 0$. We proved that $n+2$-alternance holds.

*4. Formulation of the final result.*

**Theorem 1 (of Tchebyshev on alternance, 1854).** *If* $\widehat{p}(\cdot) \in \mathrm{absmin}(1)$ *then there exists* $(n+2)$*-alternance (i. e.* $n+2$ *points* $a \le \tau_1 < \tau_2 < \ldots < \tau_{n+2} \le b$ *such that* $|x(\tau_i) - \widehat{p}(\tau_i)| = \|x(\tau_i) - \widehat{p}(\tau_i)\|_{C([t_0,t_1])}, \ 1 \le i \le n+2$ *and* $(x(\tau_i) - \widehat{p}(\tau_i))(x(\tau_{i+1}) - \widehat{p}(\tau_{i+1})) < 0, \ 1 \le i \le n+1$*).*

It is very easy to show that alternance is a sufficient condition for the solution.

Analogously many generalizations of the Tchebyshev's theorem (which were due to S. Bernstein, A. Kolmogorov, S. Zukhovitski, M. Krein, S. Stechkin and others) can be proved.

It immediately follows from Theorem 1 that $T_{n\infty}(t) = 2^{n-1} \cos n \arccos t$ is a polynomial of degree $n$ with leading coefficient 1 of the least deviation in $C([-1,1])$.

## 2. Bernstein inequality for trigonometrical polynomials

*1. Formalization.*

$$f_0(x) \to \max, \ f_1(x) \le 1, \tag{2}$$

where $x = (x_0, \ldots, x_{2n}) \leftrightarrow x(t) = \sum\limits_{0 \le k \le n} x_k \cos kt + x_{n+k} \sin kt$, $f_0(x) = \dot{x}(0)$, $f_1(x) = \max\limits_{\pi \le t \le \pi} |x(t)|$. This is the problem of convex programming.

*2. Lagrange principle.* According to the Lagrange principle, there exists a number $\lambda$ such that a solution $\widehat{x}$ of (2) is a solution of the problem $-f_0(x) + \lambda f_1(x) \to \min$ without constraints.

*3. Investigation.* Applying the cleaning principle, Fermat theorem and formula of Dubovitskii – Milyutin we obtain that there exists $r \le 2n+2$, $\{\tau_i\}_{i=1}^r$, $\tau_i \in [-\pi, \pi)$, $\{\alpha_i\}_{i=1}^r$, $\alpha_i > 0$, $\sum\limits_{i=1}^r \alpha_i = 1$ such that $\dot{x}(0) = \sum\limits_{i=1}^r \alpha_i \mathrm{sgn} x(\tau_i) x(\tau_i) = 0$ for all trigonometric polynomials $x(\cdot)$ of degree $n$. From this it is easy to deduce that $r = 2n$.

Thus the polynomial $\widehat{x}(\cdot)$ attains its extrema in $2n$ points, i. e. satisfies the equation $\dot{x}^2 = n^2(1 - x^2)$, After integrating we obtain that $\widehat{x}(t) = \sin nt$.

*4. Formulation of the final result.*

**Theorem 2. (Bernstein inequality)** *The following inequality holds:*

$$\|\dot{x}(\cdot)\|_{C(\mathbb{T})} \le n \|x(\cdot)\|_{C(\mathbb{T})} \ \forall x(\cdot) \in \mathcal{T}_n.$$

## 3. Landau – Kolmogorov inequalities on the line and the half-line

*1. Formalization.* We consider only one problem:

$$x(0) \to \max, \ \int\limits_{\mathbb{R}_+} x^2(t)dt \le 1, \ \mathrm{Var}\ \dot{x}(\cdot) \le 1. \tag{3}$$

We reformulate this problem in a style of Optimal Control (however, with a nonstandard constraint for control function):

$$x(0) \to \max, \ \int_{\mathbb{R}_+} x(t)^2 dt \le 1, \ \dot{x} = u, \ \mathrm{Var}\, u(\cdot) \le 1. \tag{3$'$}$$

*2. Lagrange principle.* We will apply the Lagrange principle to this problem *"heuristically"*. The Lagrange function here has the form: $-x(0) + \int_{\mathbb{R}_+} (\lambda x^2(t) + p(t)(\dot{x}(t) - u(t)))dt$.

Application of the Lagrange's idea leads to the following identity: $x(0) = \int_{\mathbb{R}_+}(2\lambda\widehat{x}(t)x(t) + p(t)\dot{x}(t))dt\ \forall x(\cdot)$ $(i)$ and condition of minimum: $\min_{\mathrm{Var}u(\cdot)\le 1} - \int_{\mathbb{R}_+} p(t)u(t)dt = -\int_{\mathbb{R}_+} p(t)\widehat{u}(t)dt$.

*3. Investigation.* These relations give possibility to define all unknowns: $\widehat{x}(t) = (T-t)_+$, $\dot{p}(t) = 2\lambda\widehat{x}(t) \Rightarrow p(t) = -\lambda(T-t)_+^2$, $p(0) = -1 \Rightarrow \lambda = \frac{1}{T^2}$, $\|\widehat{x}(\cdot)\|_{L_2(\mathbb{R}_+)} = 1 \Rightarrow T = 3^{\frac{1}{3}}$. Thus the value of the problem which is equal to $\widehat{x}(0) = T = 3^{\frac{1}{3}}$. But it is a heuristical solution. Now we must prove that we found the real solution of the problem.

Substituting into $(i)$ the values of $T$, $p(\cdot)$ and $\lambda$ which were defined heuristically, one obtains the identity $x(0) = \int_{\mathbb{R}_+}(\frac{2}{T^2}(T-t)_+x(t)+(1-\frac{t}{T})_+^2\dot{x}(t))dt$ $(ii)$. This identity one can be checked directly. Substitution into $(ii)$ the function $\widehat{x}(\cdot)$ gives us estimate from below ($S \ge 3^{\frac{1}{3}}$) and application the Cauchy – Bounyakovskii inequality leads to the estimate $S \le 3^{\frac{1}{3}}$.

*4. Formulation of the final result.*

**Theorem 3. (Magaril-Il'yaev inequality, 1983.)** *The following exact inequality holds true*

$$\|x(\cdot)\|_{C^b(\mathbb{R}_+)} \le 3^{\frac{1}{3}}\|x(\cdot)\|^{\frac{2}{3}}_{L_2(\mathbb{R}_+)}\|\dot{x}(\cdot)\|^{\frac{1}{3}}_{Var(\mathbb{R}_+)}.$$

## 4. Favard problem on deviation of the Sobolev class $W^r_\infty(\mathbb{T})$ from trigonometrical polynomials in the uniform norm

*1. Formalization.* One can reduce the problem to the following:

$$x(0) \to \max,\ \int_{\mathbb{T}} x(t)\cos kt \vee \sin kt dt = 0,\ 0 \le k \le n,\ |x^{(r)}(t)| \le 1. \quad (4)$$

This problem can be investigated both by means of **the Lagrange principle and by duality principle** in the Convex Analysis.

*2. Application of the duality principle.* The dual problem to the problem (4) is the problem

$$\|B_r(\cdot) - p(\cdot)\|_{L_1([-\pi,\pi])} \to \min,\ p(\cdot) \in \mathcal{T}_n, \quad (i)$$

where $B_r(\cdot)$ is the Green function of the operator $\frac{d^r}{dt^r}$ in the periodic case (i. e. $2\pi$-periodical solution of equation $\frac{d^r B(t)}{dt^r} = \delta(t) - \frac{1}{2\pi}$; the function $B(\cdot)$ is called the *Bernoulli function*), $B_r(t) = \frac{1}{\pi}\sum\limits_{k\in\mathbb{N}} k^{-r}\cos(kt - \frac{\pi r}{2})$.

*3. Investigation.* The problem $(i)$ is the problem on the best approximation of functions by polynomials. From Rolle theorem it follows that the trigonometric polynomial $\widehat{p}(\cdot)$ which interpolates $B_r(\cdot)$ at the roots of $\cos\cdot$ in the case when $r$ is even and at the roots of $\cos\cdot$ in the case when $r$ is even has no other roots besides those of cosine and sine correspondingly. Thus the polynomial $B_r(\cdot)-\widehat{p}(\cdot)$ has the least deviation from zero. Thus $\|B_r(\cdot)-\widehat{p}(\cdot)\|_{L_1([-\pi,\pi])} = |\int_{-\pi}(B_r(t)-\widehat{p}(t)\mathrm{sign}\cos nt \vee \sin ntdt| = |\int_{-\pi} B_r(t)\cos nt \vee \sin ntdt|$. The last integral was calculated by Euler. Its value is equal to $\frac{K_r}{n^r}$.

We proved

**Theorem 4 (Favard – Akhiezer – Krein (1936 – 1937)).** *The deviation of the class $W^r_\infty(\mathbb{T})$ from the space $\mathcal{T}_n$ is equal to $\frac{K_r}{n^r}$.*

## 5. A. Kolmogorov problem on the best tools of approximation.

1. The problem: *calculate $d_n(W^r_2(\mathbb{T}), L_2(\mathbb{T}))$.*

This problem can be reduced to solution of two problems: a) Bernstein problem on derivatives of trigonometrical polynomials in $L_2$-norm and Favard problem of deviation of Sobolev class $W^r_2(\mathbb{T})$ from trigonometrical polynomials. Both problems are trivial from the point of view of the general theory of extremal problems.

The result was formulated in the p.1.

2. The problem: *Calculate $d_n(H^\omega(\mathbb{T}), C(\mathbb{T})))$, where $H^\omega(\mathbb{T}) = \{x(\cdot) \in C([a,b]) \mid |x(t+\tau)-x(t)| \le \omega(|\tau|)\}$ with an increasing concave function $\omega(\cdot)$,* $\omega(0)=0$, one can reduce to the following extremal problem:

$$\int_0^1 y(t)x(t)dt \to \max,\ \ |x(t)-x(\tau)| \le \omega(|t-\tau|). \tag{5}$$

It is neither the problem of mathematical programming, nor the problem of Optimal Control. But the Lagrange principle for this problem still holds.

This problem is the convex problem and in this case the duality principle can be used. The duality principle also is a corollary of the Lagrange principle.

According to the duality principle, there exists a measure $d\mu(t,\tau)$ such that the dual problem has the form ($[0,1]=I$):

$$\int\limits_{I\times I} \omega(|t-\tau|)d\mu(t,\tau) \to \min,\ \int\limits_I d\mu(t,\tau) = dy(t),\ \int\limits_I d\mu(t,\tau) = dy(\tau)$$

It is nothing else but the so called Monge – Kantorovich duality. In our case it has a very beautiful interpretation, which leads to the solution of the problem.

**6. Optimal recovery of linear functionals**

Consider the particular case $(x^*, C, I)$ of the general problem. This is the problem of recovery $x^* \in X^*$ on a class $C \subset X$ using the information $y \in F(x)$, where $X$ and $Y$ are normed spaces, $X^*$ is the space conjugate to $X$ and $F : C \to Y$.

Any function $m : F(C) \mapsto \mathbb{R}$ we call a method of recovery. The error of this method of recovery in the $(x^*, C, F)$ problem is given by

$$e(x^*, C, F, m) = \sup_{x \in C,\, y \in F(x)} |\langle x^*, x\rangle - m(y)|.$$

We denote

$$E(x^*, C, F, m) = \inf(x_*, C, F, m),$$

where the lower bound is taken over all methods of recovery.

**Theorem. (Magaril-Ilyaev – Osipenko (1991.))** *Let $X, Y$ be normed linear spaces, $\mathrm{gr}F = \{(x, y) \in C \times Y \mid y \in F(x)\}$ be a convex centrally symmetrical set, and $x^* \in X^*$. Then $E(x^*, C, F)$ is the solution of the following problem*:

$$\sup\{\langle x^*, x\rangle \mid x \in C,\ y \in F(x), y^* \in Y^*\}. \qquad (P_1^*)$$

This result gives possibility to apply the duality principle in convex analysis to the problems of recovery of linear functionals.

**Example 1. Polynomials of the least deviation from zero and the problem of the best recovery of moments.**

$$1 \le p < \infty \Rightarrow E(m_n, BL_p([-1,1]), F_{n\,\mathrm{mom}}) = \|T_{np'}(\cdot)\|_{L_{p'}([-1,1])};$$

$$E(m_n, B\mathrm{Var}([-1,1]), F_{n\,\mathrm{mom}}) = \|T_{n\infty}(\cdot)\|_{C([-1,1])} = 2^{-(n-1)}.$$

This result is a direct corollary of the principle of duality.

**Particular cases:**

**1.1.** $E(m_n, \mathrm{Var}([-1,1]), F_{n\,\mathrm{mom}}) = \|T_{n\infty}\|_{C([-1,1])} = 2^{-(n-1)}$,
$m_n \simeq \sum\limits_{k=0}^{n-1} \lambda_{kn\infty} m_k.$

**1.2.** $E(m_n, BL_2([a,b], \rho(\cdot)), F_{n\,\mathrm{mom}}) = \|T_{n\rho}(\cdot)\|_{L_2([a,b],\rho)}$,
$m_n \simeq \sum\limits_{k=0}^{n-1} \lambda_{k\rho} m_k.$

**1.3.** $E(m_n, L_1([-1,1]), F_{n\,\mathrm{mom}}) = \|T_{n\infty}\|_{L_1([-1,1])} = 2^{-(n-1)}$,
$m_n \simeq \sum\limits_{k=0}^{n-1} \lambda_{kn1} m_k.$

### 7. Recovery problems connected with inequalities for derivatives

In the problem of recovery of the value $x(\cdot) \in \mathcal{P}_n$ at a point $\tau \in \mathbb{R} \setminus [-1,1]$ using a continuous function $y(\cdot)$ for which $\|x(\cdot) - y(\cdot)\|_{C([-1,1])} \leq \delta$, the error $E(\delta(\tau), \mathcal{P}_n, \delta BC([-1,1]))$ is equal to $\delta|T_n(\tau)|$, and the optimal method is the value at $\tau$ of the interpolation Lagrange's polynomial, which interpolates $y(\cdot)$ at the alternance points of $T_n(\cdot)$ .

## References

[1] *P. Tchebyshev,* Thèorie des méchanismes connus sous le nom de parallélogrammes. Mém. Acad. Sci. Péterb., VII (1954)

[2] *S. Bernstein,* Sur l'ordre de la meilleure approximation des fonctions continue par les polynomes de degré donn. Mémoires publiés par la classe de Sc. Acad. de Belgique (2), 4, 1 – 103

[3] *E. Landau,* Einige Ungleichungen für zweimal differenzierbare Funktionen. Proc. London Math. Soc., 13, (1913), 43 – 49

[4] *A. Kolmogorov,* Une gènéralisation dt l'inégalité dt M. J. Hadamard entre les bornes supérieures des dérivées successeives d'une fonction. C. R. Acad. Sci., Paris, 207, (1938), 764 – 765

[5] *J. Favard,* Sur l'approximation des fonctions périodiques par les polynômes trigonometrique. C.R. Acad Sci., Paris 203, (1936), 1122 – 1124

[6] *A. Kolmogorov,* Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse. Ann. of Math., 37, (1936), 107 – 110

[7] *S. Smolyak,* On optimal recovery of functions and functionals. PhD thesis, Moscow State University, 1965

[8] *J-L. Lagrange,* Théorie des fonctions analytiques, Paris, 1797.

[9] *G. G. Magaril-Ilyaev, V. M. Tikhomirov,* Convex Analysis: Theory and applications. Providence, AMS, 2003.

[10] *A. A. Lyapunov,* On completely additive vector functions. Izv. Akad. Nauk SSSR, Ser. Mat.,4 (1940), 465 – 478

*Control Theory, 2-regular Mappings*