

# ELEMENTS OF THE THEORY OF EXTREMA

V.TIKHOMIROV

**Preface**

These are notes of the course 'Elements of the theory of extrema' given by Prof. V. Tikhomirov (Moscow State University) during his stay at the Tinbergen Institute Rotterdam in September 1996. Its central thesis is that almost all concrete extremal problems of interest can be solved in a standard way by means of a few general principles. In the near future this program will be carried out more fully in further publications.

I am grateful to Prof. Tikhomirov for having given this beautiful course and for having made his notes available. I appreciate very much that the Tinbergen Institute made Prof. Tikhomirov's visit possible. Finally I would like to thank Mrs. Olga Gilissen and Arjan Berkelaar for typing the final version of this text.

Jan Brinkhuis.

**Abstract**

The following four questions could be asked about each extremal problem:

1. What are the conditions for an extremum of the problem (necessary, sufficient, necessary and sufficient)?
2. How to describe the evolution of solutions if the problem is perturbed?
3. Does a solution to the problem exist or not?
4. How to find the solution numerically?

The goal of these lecture notes is to discuss some important general approaches to these questions, to give sketches of proofs and to apply the general theory to some concrete problems.

# Contents

<b>1</b>	<b>FOUNDATIONS</b>	<b>5</b>
1.1	Introduction . . . . .	5
1.1.1	Brief historical review . . . . .	5
1.1.2	Formalization of extremal problems . . . . .	7
1.1.3	Main chapters of the theory . . . . .	8
1.2	Base . . . . .	9
1.2.1	Main principles of linear analysis . . . . .	9
1.2.2	Formulation of existence principles . . . . .	12
1.3	Proofs . . . . .	15
1.4	Applications . . . . .	19
1.4.1	Fundamental theorem of algebra . . . . .	19
1.4.2	Inverse mapping . . . . .	20
1.4.3	Global solution of a linear system of differential equations . . . . .	21
1.4.4	Local solution of differential equations . . . . .	21
1.4.5	The Lagrange multiplier rule . . . . .	23
1.4.6	An orthogonal basis of a symmetric matrix . . . . .	24
1.4.7	Separation theorem . . . . .	24

1.5	Tools . . . . .	25
1.5.1	Differential calculus . . . . .	26
1.5.2	Convex analysis . . . . .	28
1.5.3	Nonsmooth analysis . . . . .	32
<b>2</b>	<b>THEORY</b>	<b>35</b>
2.1	Lagrange principle . . . . .	35
2.1.1	Lagrange principle for smooth-convex problems . . . . .	36
2.2	Perturbations . . . . .	39
2.3	Existence, extension, relaxation . . . . .	41
2.3.1	Examples of nonexistence . . . . .	41
2.3.2	Relaxations and extensions . . . . .	45
<b>3</b>	<b>APPLICATIONS</b>	<b>47</b>
3.1	Necessary conditions . . . . .	47
3.1.1	Necessary conditions in the calculus of variations . . . . .	48
3.2	Solutions of special problems . . . . .	53
3.2.1	Classical isoperimetrical problem . . . . .	53
3.2.2	Generalized problem of Euclid . . . . .	53
3.2.3	The problem of Appolonius . . . . .	54
3.2.4	The brachistochrone problem . . . . .	54
3.2.5	Newton's aerodynamical problem. . . . .	55

# Chapter 1

## FOUNDATIONS

### 1.1 Introduction

#### 1.1.1 Brief historical review

The earliest maximum and minimum problems were posed in the distant past. For example, the classical isoperimetrical problem (*among all closed plane curves of a given length find the one that encloses the largest area*) was investigated in the 5th century B.C. Some extremal problems were solved by the greatest ancient mathematicians Euclid, Archimedes and Apollonius. In the Renaissance period many extremal problems were posed and solved (Tartaglia, Viviani, Torricelli, Kepler and others).

For a long time each extremal problem was solved individually. In the 17th century the need to create some general methods became clear, in other words the need for a *theory of extremal problems*.

The theory of extremal problems has a definite beginning. In 1636 Fermat sent a letter to Roberval and Mersenne. In this letter Fermat described his method of finding an extremum of a polynomial function. Now this method is well-known as Fermat's theorem. In its general form the algorithm of Fermat was formulated by Newton and Leibniz. Newton laid the foundation of the differential and integral calculus in his memoir '*On the method of series and fluxions with application to the geometry of curves*', which was written

in 1671 but was published only in 1736. In this memoir Newton generalized Fermat's method (without mentioning Fermat's name) to arbitrary smooth functions.

In 1684 Leibniz published an article in which he also laid the foundation of mathematical calculus. It is interesting, that the title of this first publication begins with the words "*A method for maxima and minima...*", showing the importance of the role of the problems of finding extrema in the formation of modern mathematics.

The research of Fermat, Newton and Leibniz promoted the elements of the method of finding extrema of functions of one variable. It seems, that it would have been natural to study next the extrema of functions of an arbitrary finite number of variables. But this is not what happened. The history of analysis made a kind of jump from one to infinity. This happened when J. Bernoulli posed (in 1696) his famous problem of the brachistochrone (*find the curve, joining two given points, on which a particle, sliding on this path, reaches the end point in the shortest time*). The solution of this problem (given by J. Bernoulli himself, by his brother Ja. Bernoulli, by Leibniz, de l'Hospital and Newton) founded the base of infinite dimensional analysis. For example, the method of Leibniz contained the idea of the so-called *direct methods* in infinite dimensional analysis.

At the end of the 17th and at the beginning of the 18th century some problems similar to the brachistochrone were solved. It led to the foundation of the *Calculus of Variations*. The first work in which all problems analogous to the brachistochrone were considered, was the memoir of Euler (1728).

The investigations of Euler were continued in the 17th century by Lagrange and Legendre and in the 19th century by Gauss, Poisson, Jacobi, Hamilton, Riemann, Kneser, Darboux, Du Bois-Reymond, Mayer, Poincaré, Hilbert and many others.

The theory of the calculus of variations was accomplished in its most principal parts in the first third of our century by Bolza, Bliss, McShane, Graves and others mathematicians.

At the the same period the elements of infinite dimensional calculus were founded by Volterra, Hilbert, Hadamard, Fréchet, Banach, Lyusternik, Graves and many other representatives of the main schools of mathematics

(Italian, German, French, Polish, Russian and American).

At the end of the thirties and the forties Kantorovich, Koopmans, Dantzig, Kuhn, Tucker and others founded the base of the theory of *Convex Programming* (the theory of convex extremal problems). The first work of Kantorovich was published in 1939.

In 1949 Fenchel published a series of articles in which he founded the base of the *Convex Analysis of functions*. This trend was developed by Moreau, Rockafellar, Castaing, Dubovitsky, Milyutin and many others.

At the beginning of the fifties Bellman in the USA and Pontryagin in Russia promoted the further development of the calculus of variations and founded a new branch of this theory called *Optimal Control*. It is interesting to remark, that the first problem of optimal control was posed by Newton in 1687, before the brachistochrone. This is the so-called *aerodynamical problem of Newton: find the solid of revolution of given length and width that is subject to least resistance while moving in a 'rare' medium*.

At last it is necessary to mention that in 1975 Clarke gave the initial impulse to the development of a new chapter of calculus which is now called *Nonsmooth Analysis*. Among mathematicians who took part in investigating problems of nonsmooth analysis are Ioffe, Mordukhovich and many others.

### 1.1.2 Formalization of extremal problems

Extremal problems arising in mathematics, in natural science, or in practical enterprises are usually stated initially without formulae, using the terminology of the field in which they arise (see, for example, our formulations of the classical isoperimetrical problem, the brachistochrone and Newton's aerodynamical problem).

In order to be able to utilize analytical tools it is necessary to apply a translation of the statement of the problem from each specific language to the language of mathematics. Such a translation is called *a formalization*.

To formalize an extremal problem means to describe precisely a functional  $f_0$  ( $f_0 : X \rightarrow \bar{\mathbb{R}}$ ,  $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ ) to be minimized or maximized (together with its domain  $X$ ) and a constraint set  $C \subset X$  such that  $f_0(C) \subset \mathbb{R}$ . The



constraint set is given usually by equalities and inequalities.

We will use the abbreviated notation for a formalized problem:

$$f_0(x) \rightarrow \min(\max); x \in C. \quad (P)$$

Points  $x \in C$  are called *admissible points*. An admissible point  $\hat{x}$  is called an *absolute minimum (maximum)* of the problem (P) if  $f_0(x) \geq f_0(\hat{x})$  ( $f_0(x) \leq f_0(\hat{x})$ ) for every  $x \in C$ . We can change a problem of maximization into a problem of minimization by changing the sign of the functional.

An absolute minimum of an extremal problem is called a *solution of the problem*. Our goal is to find solutions of problems, but at the beginning (as a rule) we find local extrema of these problems. Let  $X$  be a topological space. An admissible point  $\hat{x}$  is called a *local minimum of (P)* if there exists a neighbourhood  $U$  of  $\hat{x}$  such that  $\hat{x}$  is a solution of the problem  $f(x) \rightarrow \min; x \in C \cap U$ .

### 1.1.3 Main chapters of the theory

In subsection 1.1 the main parts of the theory have been mentioned. Now it is possible to describe them more accurately.

*The mathematical programming problem with equality constraints* studies problems (P) with  $X = \mathbb{R}^n$  and  $C = \{x \mid F(x) = 0\}$  with  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

In finite dimensional *convex programming*  $X = \mathbb{R}^n$ ,  $C = \{x \mid f_i(x) \leq 0, 1 \leq i \leq m, x \in A\}$ , where all functions  $f_i, 1 \leq i \leq m$  are convex and  $A$  is a convex set in  $\mathbb{R}^n$ .

In the *calculus of variations* the domains of functionals are *function spaces* ( $C^r, W_p^r, H^r$  and so on) and constraints are defined by *differential equations*. Here is a typical example of a problem of the calculus of variations (the so-called *Lagrange problem*):

$$f_0(x(\cdot)) = \int_{\Delta} L(t, x(t), \dot{x}(t)) dt \rightarrow \min(\max); M(t, x, \dot{x}) = 0, x|_{\partial\Delta} = \xi,$$

where  $\Delta$  is in the one-dimensional case a segment  $[t_0, t_1]$  and in the multidimensional case a domain in  $\mathbb{R}^n$ .

In *optimal control* the unknown variables are divided into two parts: the state (or phase) coordinates ( $x$ ) and the control coordinates ( $u$ ). A typical problem of optimal control is the following:

$$f_0(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} f(t, x(t), u(t)) dt + l(x(t_0), x(t_1)) \rightarrow \min;$$

$$\dot{x}(t) = \varphi(t, x(t), u(t)), u(t) \in U \forall t. \quad (\overline{P})$$

The crucial distinction between this problem and the problems of the calculus of variations consists of the constraint  $u(t) \in U \forall t$ ; this allows for example *inequality constraints in the control coordinates*.

The following four questions could be asked about each extremal problem:

1. What are the conditions for an extremum of the problem (necessary, sufficient, necessary and sufficient)?
2. How to describe the evolution of solutions if the problem is perturbed?
3. Does a solution to the problem exist or not?
4. How to find the solution numerically?

The goal of my lectures is to discuss some important general approaches to these questions.

## 1.2 Base

### 1.2.1 Main principles of linear analysis

The theory of extremal problems is based on functional analysis. We will touch on two parts of functional analysis. Here we discuss some important results of linear analysis and then we will speak about calculus.

#### Duality and the Hahn-Banach theorem

One of the most fundamental ideas not only in mathematics but also in physics and philosophy is the idea of *duality*. There are many dual objects

in functional analysis: topological vector spaces and their duals, subspaces and their annihilators, convex sets and their polars, support functions and their subdifferentials, etc.

It is possible to formulate the following general principle: *every convex subject or phenomenon has two descriptions, a primal one in the original space and a dual one in the conjugate space.*

The majority of such facts follow from the following fundamental result.

**Theorem 1.2.1** (*Hahn-Banach*)

Let  $X$  be a linear (vector) space,  $p : X \rightarrow \mathbb{R}$  a sublinear ( $\Leftrightarrow$  convex positive homogeneous of the first degree and  $p(0) = 0$ ) function,  $X_1$  a subspace of  $X$ . Then every linear functional on  $X_1$  majorized by  $p$  on  $X_1$  can be extended to the whole space  $X$  with the same property

$$\Leftrightarrow \forall \xi' \in X' : \langle \xi', x_1 \rangle \leq p(x_1) \forall x_1 \in X_1 \quad \exists x' \in X' : \langle x', x \rangle \leq p(x) \\ \forall x \in X, x'|_{X_1} = \xi'.$$

It is interesting to mark the dual sense of the theorem. On the one hand the epigraph of an arbitrary sublinear function  $p$  is the *union* of epigraphs of elementary sublinear functions (which are equal to  $p$  on one ray and infinity in all other points;  $p_\xi(x) = p(x)$  if  $x$  lies on the ray  $t\xi, t \geq 0$  and  $\infty$  in all other cases.) On the other hand the epigraph of this function  $p$  is the *intersection* of epigraphs of linear functionals supported to the epigraph of the given sublinear function.

**Inverse function theorem and calculus**

One of the most important results in nonlinear analysis is the inverse function theorem. It can be described as follows: if the derivative of a nonlinear mapping at a certain point has an inverse, then this nonlinear mapping itself has an inverse, locally at that point. Such results are based on the following linear prototype:

**Theorem 1.2.2** (*right inverse theorem*) Let  $X, Y$  be Banach spaces and  $\Lambda$  a continuous surjective linear operator from  $X$  onto  $Y$ . Then, there exists

an operator  $R : Y \rightarrow X$  (not necessarily linear or continuous) and a positive constant  $C$  such that

$$\Lambda Ry = y, \|Ry\| \leq C\|y\| \quad \forall y.$$

This result was originally formulated either as the *inverse operator theorem* (the inverse of a bijective linear continuous operator is continuous) or as the open mapping theorem: *Let  $X, Y$  be Banach spaces and let  $\Lambda : X \rightarrow Y$  be a surjective continuous operator. Then  $\Lambda$  is an open operator ( $\Leftrightarrow$  the image of every open set is an open set).*

The *closed graph theorem* plays an important role in applications of the theory to the calculus of variations and optimal control:

**Theorem 1.2.3** (*Banach-Steinhaus*) *A linear continuous operator from one Banach space to another has a closed graph.*

A different form of the same result is the following: *Let  $\Lambda_n$  be a sequence of continuous linear operators. This sequence tends to the linear continuous operator  $\Lambda$  iff on some dense subset of  $X$   $\Lambda_n x \rightarrow \Lambda x$  and the norms of all operators  $\Lambda_n$  are uniformly bounded.*

### Existence of solutions and the Banach-Alaoglu-Bourbaki principle

Many natural and important variational problems have no solutions (for example, the problem of the rotation surface of the least area). Analyzing such situations Hilbert introduced his famous principle according to which all ‘regular’ problems must have a solution in some extended sense. One of the most beautiful realizations is the idea of ‘embedding’ the initial space into the double dual space in which there exists compactness! This is based on the following result:

**Theorem 1.2.4** (*Banach-Alaoglu-Bourbaki*) *A convex closed bounded subset in the conjugate space of a Banach space is compact in weak\* topology.*

### 1.2.2 Formulation of existence principles

Among the most important results of the general theory of extremal problems are the existence theorems of Lagrange multipliers, of fields of extremals, of solutions and so on. Besides that the investigation of the problems of the calculus of variations and optimal control demands to use existence theorems from the theory of differential equations (existence of a solution of the Cauchy problem, existence of the global solution for linear systems, dependence of solutions on the parameters, and so on). Many of such classical existence theorems can be deduced from a few general principles of existence, which are based on considerations of compactness, contractivity, monotonicity and on other topological ideas.

Most of the concrete theorems of the central part of these lectures are corollaries of the theorems we speak about in this section.

#### I. Principles based on the idea of compactness

##### A) The Weierstrass–Lebesgue compactness principle

*Let  $X$  be a compact topological space and let  $f$  be a continuous (or even a lower-semicontinuous) function on  $X$ . Then there exists a point  $\hat{x}$  in  $X$  at which  $f$  attains its global minimum.*

**B)** *Let  $X$  be a reflexive Banach space and let  $f$  be a proper closed convex coercive function on  $X$  ( $\Leftrightarrow \text{epi } f := \{(x, \alpha) \in X \times \mathbb{R} \mid \alpha \geq f(x)\}$  is a convex, closed proper subset of  $X \times \mathbb{R}$ ,  $f(x) \rightarrow \infty$ , when  $x \rightarrow \infty$  and  $f(x) > -\infty$ ). Then there exists  $\hat{x} \in X$  at which  $f$  attains its absolute minimum.*

#### II. Baire's principle

*Let  $(X, d)$  be a complete metric space and let  $\{A_k\}_{k \in \mathbb{N}}$  be a family of nowhere dense sets in  $X$ . Then there exists a point  $\hat{x} \in X \setminus \cup_{k \in \mathbb{N}} A_k$ .*

#### III. Iterative principles based on the idea of contraction

##### A) Generalized iterative principle

*Let  $(X, d_X)$  be a complete metric space,  $(Y, d_Y)$  a metric space,  $F$  a continuous (or a closed multivalued) mapping from  $X$  to  $Y$ ,  $y \in Y$ ,  $x_0 \in X$ ,  $r > 0$  and let there exist  $M : B_X(x_0, r) \rightarrow X$  and  $\theta$ ,  $0 < \theta < 1$ , such that*

$$a) d_Y(y, F(x_0)) \leq (1 - \theta)r, b) d_X(x, M(x)) \leq d_Y(y, F(x)) \quad \forall x \in B_X(x_0, r),$$

and either

$$c) \quad d_Y(y, F(M(x))) \leq \theta d_X(x, M(x)) \quad \forall x \in B_X(x_0, r),$$

or

$$c') \quad d_Y(y, F(M(x))) \leq \theta d_Y(y, F(x)) \quad \forall x \in B_X(x_0, r)$$

(if  $A$  is a set in the metric space  $Z$ , then  $d_Z(x, A) := \inf_{z \in A} d(x, z)$ .)

Then the iterative sequence  $x_k = M(x_{k-1})$ ,  $k \in \mathbb{N}$  tends to a solution of the equation  $y = F(x)$  (or inclusion  $y \in F(x)$ ).

**Remark 1.** This formulation combines Newton's method of solving equations and Milyutin's ideas of generalization of Lyusternik's theorem. It is obvious that it is sufficient to suppose that in each step there exists an element  $M_k(x_k)$  which satisfies the conditions  $b)$  and  $c)$  (or  $c')$ ).

**B)** Let  $(X, d)$  be a complete metric space,  $r > 0$ ,  $\theta \in \langle 0, 1 \rangle$  and let  $F$  be a mapping from  $B(x_0, r)$  to  $X$  such that

$$(a) \quad d(x_0, F(x_0)) \leq (1 - \theta)r,$$

$$(b) \quad d(F(x), F(F(x))) \leq \theta d(x, F(x)),$$

then the sequence  $x_k = F(x_{k-1})$ ,  $k \in \mathbb{Z}_+$  tends to a fixed point of  $F$ .

**Remark 2.** It is possible to generalize this result by considering a cartesian product of a topological space  $T$  and a complete metric space  $(X, d)$ . Such a generalization is useful in connection with considerations of dependence of solutions on parameters.

### Corollary (Cacciopoli-Banach)

Let  $(X, d)$  be a Banach space and  $F$  a mapping from  $X$  into  $X$  which is contractive, that is

$$\exists \theta \in \langle 0, 1 \rangle: d(F(x), F(y)) \leq \theta d(x, y) \quad \forall x, y \in X.$$

Then the sequence  $x_k = F(x_{k-1})$ ,  $k \in \mathbb{Z}_+$ ,  $x_0$  an arbitrary point, tends to the (unique) fixed point of  $F$ .

**Remark 3.** It is easy to see that if some power of  $F$  is a contractive mapping, then  $F$  has a unique fixed point.

#### IV. Brouwer's and Borsuk's topological principles

**A) (Brouwer)** Let  $F \in C(B^n, \mathbb{R}^n)$ ,  $B^n = \{x \in \mathbb{R}^n \mid |x| \leq 1\}$  such that  $\exists \delta \in \langle 0, 1 \rangle \forall x \in B^n \mid x - F(x) \mid < 1 - \delta$ . Then, for all  $y \in \mathbb{R}^n$  with  $|y| < \delta$  there exists  $\hat{x}$  such that  $y = F(\hat{x})$ .

**B) (Borsuk)** Let  $F \in C(S^n, \mathbb{R}^n)$ ,  $S^n = \{x \in \mathbb{R}^{n+1} \mid |x| = 1\}$  such that  $F(-x) = -F(x) \forall x$ . Then, there exists a solution of the equation  $F(x) = 0$ .

#### V. Monotonicity principles.

The term 'monotonicity' has several meanings which are the same in the one-dimensional case. One meaning is connected with *order*, another with the derivatives of a convex function. We present below three corresponding monotonicity principles.

**A) (Zorn's lemma)** Let  $(\Xi, \mathcal{O})$  be a partially ordered set such that each linearly ordered (that is, totally ordered) subset has an upper bound. Then  $(\Xi, \mathcal{O})$  has a maximal element.

The following principle is a corollary of Zorn's lemma:

**B) (Variational principle of Ekeland)** Let  $(X, d)$  be a complete metric space,  $f$  a lower-semicontinuous function bounded from below, and  $x_0 \in X$  a point such that  $f(x_0) \leq \inf f + \varepsilon$  for some  $\varepsilon > 0$ . Then, for all  $\lambda > 0$  there exists an element  $\hat{x} = \hat{x}(\lambda)$  such that:

$$(a) f(\hat{x}) \leq f(x_0); (b) d(x_0, \hat{x}) \leq \lambda; (c) f(x) + (\varepsilon/\lambda)d(x, \hat{x}) > f(\hat{x}) \forall x \neq \hat{x}$$

(i.e. the perturbed function  $f(\cdot) + (\varepsilon/\lambda)d(\cdot, \cdot, \hat{x})$  attains a strict absolute minimum at the point  $\hat{x}$ ).

**C) (Monotonicity principle of Minty – Brouwer)** A strictly monotone, continuous and coercive mapping  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  has a unique solution of the equation  $F(x) = y$  for each  $y \in \mathbb{R}^n$  (and it can be reached by means of some special iterative procedure).

(The mapping  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is called *monotone (strictly monotone)* if  $\langle F(x_2) - F(x_1), x_2 - x_1 \rangle \geq 0$  ( $\geq \alpha |x_2 - x_1|^2$  for a suitable  $\alpha > 0$ )  $\forall x_1, x_2 \in \mathbb{R}^n$  and *coercive*, if  $\lim_{x \rightarrow \infty} |F(x)| = \infty$ .)

## 1.3 Proofs

### I. The compactness principles.

A) 1. Notation  $\mathcal{L}_c := \{x \mid f(x) \leq c\}$ . By definition of lower-semicontinuity it follows that  $U_n := X \setminus \mathcal{L}_n$ ,  $n \in \mathbb{Z}$  is an open subset of  $X$ . It is clear that  $\cdots \subset U_n \subset U_{n-1} \subset \cdots$  and that  $\{U_n\}_{n \in \mathbb{Z}}$  is an open covering of  $X$ .

Hence (by the definition of compactness) there exists  $m \in \mathbb{Z}$  such that  $X = U_m$ , i.e.  $f$  is bounded from below.

2. Notation:  $\mu := \inf f$ ,  $V_n := X \setminus \mathcal{L}_{\mu+(1/n)}$   $n \in \mathbb{N}$ . By the definition of lower-semicontinuity it follows that  $\{V_n\}_{n \in \mathbb{N}}$  is (if  $f$  does not attain its minimum) an open covering of  $X$ . Hence (by definition of compactness) there exists  $s \in \mathbb{N}$  such that  $X = V_s$ , i.e.  $f > \mu + 1/s$ . But this contradicts the definition of  $\mu$ .  $\square$

B) This result follows immediately from the Weierstrass–Lebesgue compactness principle and the Banach–Alaoglu–Bourbaki compactness theorem, according to which every closed bounded convex set in the conjugate space is compact in the weak\* topology and from the fact that a closed convex function is lower-semicontinuous in the same topology.  $\square$

### II. Baire’s principle.

$B(\xi, r) := \{x \mid d(x, \xi) \leq r\}$  – the ball with center at  $\xi$  and with radius  $r$ .

From the definition of nowhere density  $\exists B_1 := B(x_1, r_1) : B_1 \cap A_1 = \emptyset$

Analogously  $\exists B_2 := B(x_2, r_2)$ ,  $B_2$  included in  $B_1$ ,  $B_2 \cap A_2 = \emptyset$ , and so on; we obtain a sequence of balls each included in its predecessor. Using the lemma on including balls we obtain the result.  $\square$

### III. Iterative principles.



A)  $\triangleleft$  We have:

$$d_X(x_1, x_0) \stackrel{def}{=} d_X(x_0, M(x_0)) \stackrel{(b)}{\leq} d_Y(y, F(x_0)) \stackrel{(a)}{\leq} (1 - \theta)r < r,$$

hence

$$x_1 \in B_X(x_0, r) \tag{i}.$$

Let the elements  $\{x_k\}_{k=1}^n$  belong to  $B_X(x_0, r)$ . Then in case (c) we have:

$$\begin{aligned} d_X(x_{k+1}, x_k) &\stackrel{def}{=} d_X(x_k, M(x_k)) \stackrel{(b)}{\leq} d_Y(y, F(x_k)) \stackrel{(c)}{\leq} \theta d_X(x_k, x_{k-1}) \leq \\ &\stackrel{anal}{\dots} \leq \theta^k d_X(x_1, x_0). \end{aligned}$$

Hence

$$d_X(x_{k+1}, x_0) \leq (\theta^k + \dots + \theta + 1)d_X(x_1, x_0) < r.$$

The case (c') is similar. By induction, the sequence  $(x_k)_{k \in \mathbb{Z}_+}$  is defined for all  $k \in \mathbb{Z}_+$  and from the triangle inequality (TI) we obtain:

$$\begin{aligned} d_X(x_{k+m}, x_k) &\stackrel{TI}{\leq} d_X(x_{k+m}, x_{k+m-1}) + \dots + d_X(x_{k+1}, x_k) \stackrel{(c)}{\leq} \\ &(\theta^{k+m-1} + \dots + \theta^k)d_X(x_1, x_0) \leq \frac{\theta^k}{1 - \theta}d_X(x_1, x_0). \end{aligned}$$

The case (c') is analogous. Hence,  $(x_k)_{k \in \mathbb{Z}_+}$  is a Cauchy sequence which (by the condition that  $X$  is a complete space) tends to some element  $\hat{x}$ . As  $F$  is closed it follows that  $y \in F(\hat{x})$ .  $\square$

The method of proof of theorem B) is analogous to the method used in the proof of theorem A).

#### IV. Topological principles

A)  $\triangleleft$  Brouwer's topological principle of existence is based on the following famous result.

**Brouwer's fixed point theorem.** *Let  $F$  be a continuous mapping from  $B^n$  to itself. Then there exists a fixed point ( $\Leftrightarrow \exists \hat{x} : F(\hat{x}) = \hat{x}$ ).*

*Proof of the fixed point theorem.*

Without loss of generality we may assume that  $F$  is a continuously differentiable function (this follows from the Weierstrass approximation theorem).

Let us suppose that  $F(x) \neq x \quad \forall x \in B^n$ .

For each  $x \in B^n$  let  $G(x)$  be the intersection of  $S^{n-1}$  and the ray with the initial point at  $F(x)$  which passes through  $x$ . In the one-dimensional case  $G(x) = -1$  if  $F(x) > x$  and  $+1$  if  $F(x) < x$ . From the definition of  $G$  it follows that if  $x \in S^{n-1}$  then  $G(x) = x$  (i). It is clear that the map  $G$  is also continuously differentiable. Then in the one-dimensional case ( $N = L \Leftrightarrow$  Newton-Leibniz theorem,  $Vol_1(A)$  is the length of  $A$ ):

$$0 = Vol_1(G(B^1)) \stackrel{def}{=} \int_{B^1} G'(x) dx \stackrel{Id}{=} \int_{B^1} dG \stackrel{N=L}{=} G(1) - G(-1) = 2.$$

Contradiction.

In the  $n$ -dimensional case we use the theorem of Stokes-Poincaré ( $S = P$ ) instead of that of Newton-Leibniz ( $Vol_n(A)$  is the  $n$ -dimensional volume of  $A$ ):

$$0 = Vol_n(G(B^n)) \stackrel{def}{=} \int_{B^n} G'(x) dx \stackrel{Id}{=} \int_{B^n} d(G_1 \wedge dG_2 \wedge \cdots \wedge dG_n) \\ \stackrel{(i)}{=} \int_{B^n} d(x_1 \wedge dx_2 \wedge \cdots \wedge dx_n) \stackrel{S=P}{=} Vol_n(B^n) \neq 0.$$

Contradiction. The fixed point theorem is proved.  $\square$

Brouwer's existence principle follows immediately from the fixed point theorem. We will prove it for  $y = 0$ . The general case is analogous. Let  $F(x) \neq 0 \quad \forall x \in B^n$ , then the mapping  $G_1(x) := -\frac{F(x)}{|F(x)|}$  is a continuous mapping from  $B^n$  to  $B^n$  without a fixed point (because if  $G_1(x) = x$ , then  $|x - F(x)| = |G_1(x) - F(x)| = \left| \frac{F(x)}{|F(x)|} + F(x) \right| > 1$ ). Contradiction.  $\square$

**Remark.** We have given an analytical proof of Brouwer's theorem. Now we will try to explain the topological essence of the proof of this theorem and the theorem of Borsuk. Every (smooth enough) mapping  $F : S^n \rightarrow S^n$  has a special number (the so-called *degree* of this mapping). The degree of a smooth enough mapping  $F : S^n \rightarrow S^n$  is defined to be the number of

inverse images of a point in “general position” where we take orientation into account in the counting of the number of inverse points. It is easy to see that if  $F$  is a continuous mapping from  $B^n$  to itself, then the restriction of the mapping  $x \rightarrow (F(x) - x)/|F(x) - x|$  to the unit sphere has degree one (we suppose that  $F$  has no fixed point). Then we will consider the restriction of this mapping to spheres of radius  $r, 0 < r < 1$ . Then, the degree (it is a continuous but discrete function of  $r$ ) is not changed. But if  $r$  is sufficiently small, then the degree is (evidently) equal to zero. This contradiction leads to a proof of Brouwer’s fixed point theorem.

To prove Borsuk’s theorem we note that the degree of an odd continuous mapping of  $\mathbf{S}^n$  onto itself is an odd number. Also it is necessary to recall that every continuous mapping can be approximated by smooth ones.

### V. Monotonicity principles.

We will not prove Zorn’s lemma; the proof of this result is not connected with mathematical analysis (but we will see that it has many applications to analysis).

B) We said that Ekeland’s principle is a corollary of Zorn’s lemma. But we will give “a constructive” proof of the principle (and from our proof it will be clear how to deduce the result from Zorn’s lemma).

◁ Without loss of generality we may assume that  $\lambda = 1$  and  $f(x) \geq 0$ . The method of the proof is an iterative variant of the bisection procedure.

Let  $z = (\xi, \beta) \in X \times \mathbb{R}_+$ . We denote the set  $\{(x, \alpha) \in X \times \mathbb{R}_+ \mid 0 \leq \alpha \leq \beta, d(x, \xi) \leq \beta - \alpha\}$  by  $S(z)$ . Take  $z_0 = (x_0, f(x_0))$ . If  $S(z_0) \cap \text{epi } f = z_0$ , the problem is solved:  $\hat{x} = x_0$ . If not, then we consider the ball  $B(x_0, f(x_0))$  and  $S_0 := \text{epi } f \cap \Xi_0, \quad \Xi_0 := \{(x, \alpha) \in S(z_0) \mid 0 \leq \alpha \leq f(x_0)/2\}$ .

If  $S_0 \neq \emptyset$  we take  $z_1 = (x_1, f(x_1)) \in S_0$ , the ball  $B(x_1, f(x_1))$  and repeat our procedure. If  $S_0 = \emptyset$  we take  $(x_1, f(x_1)) \in S(z_0) \setminus \Xi_0$ , the ball  $B(x_1, f(x_1) - f(x_0)/2)$  and repeat our procedure. We obtain a sequence of nested balls, whose centers tend to  $\hat{x}$ . It is clear that  $\hat{x}$  is the point we are looking for.  $\square$

C) ◁ First of all we will illustrate the proof of the theorem in the one dimensional case. Then  $F$  is a map from  $\mathbb{R}$  to  $\mathbb{R}$  which is strictly monotonically

increasing and surjective. In this case we can restrict ourselves to a segment  $\Delta_0 = [\xi, \eta]$  such that  $F(\xi) < y$ ,  $F(\eta) > y$ .

Then we use the well-known method of *bisection*.

Let  $x_1$  be the midpoint of the segment  $\Delta_0$ . If  $F(x_1) > y$  then we consider the segment  $\Delta_1 = [\xi, x_1]$ . If  $F(x_1) < y$ , then we consider the segment  $\Delta_1 = [x_1, \eta]$ . Let  $x_2$  be the midpoint of the segment  $\Delta_1$ . Then repeat this step: calculate  $F(x_2)$ , compare it with  $y$ , cut off again the half of the segment which does not contain the solution and so on.

In the general case the problem can be reduced to the case  $y = 0$ . To solve the equation  $F(x) = 0$  we restrict our consideration to an  $n$ -dimensional ball  $\Delta_0$  such that  $|F(x)| \neq 0 \ \forall x \in \mathbb{R} \setminus \Delta_0$  (this is possible by coerciveness) and then instead of the midpoint of the segment we take the center  $x_0$  of the ball. If  $F(x_0) = 0$ , the problem is solved. If  $F(x_0) \neq 0$  we cut off the set  $\{x \mid \langle F(x_0), x - x_0 \rangle > 0\}$  (it is easy to show by definition that in this set there is no solution of the equation  $F(x) = 0$ ). We denote the remainder of the ball by  $\Delta_1$ , then denote the center of gravity of  $\Delta_1$  by  $x_1$  and cut off the set  $\{x \in \Delta_1 \mid \langle F(x), x - x_1 \rangle > 0\}$ , etc.

From the sequence  $(x_k)_{k \in \mathbb{Z}_+}$  it is possible to choose a subsequence which tends to a solution of the equation  $F(x) = 0$ .  $\square$

## 1.4 Applications

### 1.4.1 Fundamental theorem of algebra

**Theorem 1.4.1** (*d'Alembert, Euler, Gauss*)

*A polynomial of degree  $n \geq 1$  with complex coefficients has a complex root.*

**Proof:** Consider the extremal problem:

$$f_0(z) = |p(z)|^2 \rightarrow \min; z \in \mathbf{C}, \quad (P)$$

where  $p(\cdot)$  is a polynomial degree  $n$ . The function  $f_0$  is a coercive, hence by the compactness principle of Weierstrass–Lebesgue a solution  $\hat{z}$  of the

problem (P) exists. Without loss of generality we can assume that  $\hat{z} = 0$  and  $p(z) = a_0 + a_k z^k + \cdots + a_n z^n, a_0 a_k \neq 0$ . Then (that was the main point of d'Alembert's proof) -writing  $\beta = \arg(\bar{a}_0 a_k)$ - one has

$$|P(te^{i\theta})|^2 \stackrel{Id}{=} |a_0|^2 + 2|a_0||a_k|t^k \cos(k\theta + \beta) + O(t^{k+1}) \quad (i).$$

From this expression it follows that 0 is not an absolute minimum in (P) (because it is possible to choose  $\hat{\theta}$  such that  $|P(te^{i\hat{\theta}})|^2 < |a_0|^2$  for small  $t$  - see (i)) Contradiction.  $\square$

### 1.4.2 Inverse mapping

#### Theorem 1.4.2 (Dini)

Let  $\mathcal{U}$  be a neighbourhood of a point  $\hat{x} \in \mathbb{R}^n, F : \mathcal{U} \rightarrow \mathbb{R}^n (F = (f_1, \cdots, f_n))$  a mapping which is strictly differentiable in  $\hat{x}, F(\hat{x}) = \hat{y}$  and the Jacobian  $\det \left( \frac{\partial f_i(\hat{x})}{\partial x_j} \right)_{i,j=1}^n \neq 0$ . Then there exists  $\varepsilon > 0, \delta > 0$  and  $K > 0$  such that for all  $y$  in the ball  $B(\hat{y}, \delta)$  there exists a unique  $x$  in the ball  $B(\hat{x}, \varepsilon)$  such that  $F(x) = y$  and  $|x - \hat{x}| \leq K|y - \hat{y}|$ .

$\triangleleft$  Application of principle III A) leads to the result apart from the uniqueness statement which requires an additional argument. Now we outline the verification of the assumptions of principle III A) in the present case. Without loss of generality we may assume that  $\hat{x} = \hat{y} = 0$  and that  $\|F'^{-1}\| \leq 1$  (we can use shifts and homotheties). Take  $y \in Y$ . Consider the iterative procedure  $x_k = M(x_{k-1}), k \in \mathbb{N}, x_0 = 0$ , where

$$M(x) = x + (F'(0))^{-1}(y - F(x)). \quad (*)$$

Then  $\|x - M(x)\| \leq \|y - F(x)\|$  and if  $\|y\|$  is small enough then for each  $\theta \in (0, 1)$  and for sufficiently small  $r > 0$  one has for all  $x \in B(0, r)$

$$\begin{aligned} \|F(M(x)) - y\| &\stackrel{(*)}{=} \|F(M(x)) - y + y - F(x) - F'(0)(M(x) - x)\| \\ &\leq \theta \|M(x) - x\|. \end{aligned}$$

$\square$

### 1.4.3 Global solution of a linear system of differential equations

**Theorem 1.4.3** *Let  $\Delta = [t_0, t_1]$  be a finite interval of  $\mathbb{R}$ . Let  $A : \Delta \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$  and  $b : \Delta \rightarrow \mathbb{R}^n$  be an integrable matrix function and an integrable vector function. Then the Cauchy problem*

$$\dot{x} = A(t)x + b(t), \quad x(t_0) = x_0 \tag{P}$$

has a unique solution on  $\Delta$ .

**Proof:**  $\triangleleft$  A solution of (P) is equivalent to the integral equation

$$x(t) = x_0 + \int_{t_0}^t A(s)x(s)ds + \int_{t_0}^t b(s)ds.$$

Consider the operator  $F$  from  $C(\Delta, \mathbb{R}^n)$  to  $C(\Delta, \mathbb{R}^n)$  defined by  $(Fx)(t) = x_0 + \int_{t_0}^t A(s)x(s)ds + \int_{t_0}^t b(s)ds$ . The following estimates are evident:

$$\|Fx_1(\cdot) - Fx_2(\cdot)\|_{C(\Delta, \mathbb{R}^n)} \leq |t_1 - t_0| \int_{t_0}^{t_1} \|A(s)\| ds \|x_1(\cdot) - x_2(\cdot)\|_{C(\Delta, \mathbb{R}^n)},$$

$$\|F^2x_1(\cdot) - F^2x_2(\cdot)\|_{C(\Delta, \mathbb{R}^n)} \leq \frac{|t_1 - t_0|^2}{2!} \int_{t_0}^{t_1} \|A(s)\| ds \|x_1(\cdot) - x_2(\cdot)\|_{C(\Delta, \mathbb{R}^n)},$$

... ..

$$\|F^kx_1(\cdot) - F^kx_2(\cdot)\|_{C(\Delta, \mathbb{R}^n)} \leq \frac{|t_1 - t_0|^k}{k!} \int_{t_0}^{t_1} \|A(s)\| ds \|x_1(\cdot) - x_2(\cdot)\|_{C(\Delta, \mathbb{R}^n)}.$$

Applying the principle of Cacciopoli-Banach (see the remark after the theorem of C.-B.), we obtain the result.  $\square$

### 1.4.4 Local solution of differential equations

**Theorem 1.4.4** *(Cauchy, Picard)*

*Let the function  $f : G \rightarrow \mathbb{R}^n$  be defined on an open set  $G$  of  $\mathbb{R} \times \mathbb{R}^n$  and let it satisfy the following conditions:*

- (a) for any  $x$  the function  $t \rightarrow f(t, x)$  defined on the section  $G_x = \{t \mid (t, x) \in G\}$  is measurable and locally integrable (that is, integrable on any finite interval contained in  $G_x$ );
- (b) for any  $t$  the function  $x \rightarrow f(t, x)$  defined on the section  $G_t = \{x \mid (t, x) \in G\}$  is differentiable;
- (c) for any compact set  $\mathcal{K} \subset G$  there exists a locally integrable function  $k(\cdot)$  such that  $\|f_x(t, x)\| \leq k(t) \quad \forall (t, x) \in \mathcal{K}$ .

Let  $\mathcal{K} \subset G$  be a compact set. Then there exist  $\varepsilon > 0$  and  $\delta > 0$  such that for any point  $(\hat{t}, \hat{x}) \in \mathcal{K}$  and for  $(t_0, x_0)$  satisfying the inequalities  $|t_0 - \hat{t}| < \delta$ ,  $|x_0 - \hat{x}| < \varepsilon$  Cauchy's problem  $\dot{x} = f(t, x)$ ,  $x(t_0) = x_0$  has a unique solution  $t \rightarrow X(t, t_0, x_0)$  on the closed interval  $[\hat{t} - \delta, \hat{t} + \delta]$  and this solution is jointly continuous with respect to its arguments.

**Proof:** We will prove only the existence part of the theorem (without the continuity part; for the complete result see [ATF], p. 119). Consider the mapping  $F$  defined by

$$(Fx)(t) = \hat{x} + \int_{\hat{t}}^t f(\tau, x(\tau)) d\tau.$$

It is easy to show that for a small  $\beta > 0$  it is possible to find  $\delta > 0$  such that

$$\|Fx(\cdot)\|_{C([\hat{t}-\delta, \hat{t}+\delta])} \leq \beta(1 - \theta),$$

$$\|Fx_1(\cdot) - Fx_2(\cdot)\|_{C([\hat{t}-\delta, \hat{t}+\delta])} \stackrel{def}{=} \max_{t \in [\hat{t}-\delta, \hat{t}+\delta]} \left| \int_{\hat{t}}^t (f(s, x_1(s)) - f(s, x_2(s))) ds \right|$$

$$\begin{aligned} & \leq \int_{\hat{t}-\delta}^{\hat{t}+\delta} k(s) ds \|x_1(\cdot) - x_2(\cdot)\| \leq \theta \|x_1(\cdot) - x_2(\cdot)\|, \end{aligned}$$

for  $\theta \in (0, 1)$ . Applying III B) we obtain the result.  $\square$

### 1.4.5 The Lagrange multiplier rule

**Theorem 1.4.5** (Lagrange)

Consider the problem

$$f_0(x) \rightarrow \text{extr}; f_i(x) = 0, 1 \leq i \leq m. \quad (P)$$

Here the  $f_i : U \rightarrow \mathbb{R}, 0 \leq i \leq m$  are continuously differentiable functions defined in a neighbourhood  $U \subset \mathbb{R}^n$  of a point  $\hat{x}$  which is a local extremum of (P) and where  $m \leq n$ . Then there exist Lagrange multipliers  $(\lambda_0, \lambda), \lambda = (\lambda_1, \dots, \lambda_m)$  (not all zero) such that

$$\mathcal{L}_x(\hat{x}, \lambda_0, \lambda) = 0 \Leftrightarrow \sum_{i=0}^m \lambda_i f'_i(\hat{x}) = 0 \quad (*)$$

where  $\mathcal{L}(x, \lambda_0, \lambda) := \lambda_0 f_0(x) + \sum_{i=1}^m \lambda_i f_i(x)$ .

**Proof:** The following alternative takes place: either the vectors  $f'_i(\hat{x}), 0 \leq i \leq m$  are linearly dependent (I) or not (II). In the first case (\*) is satisfied by definition. We will show that the second case contradicts the assumptions of the theorem.

According to a well-known theorem of linear algebra in the case (II) it is possible to choose  $m+1$  variables (without loss of generality let these be the first  $m+1$  variables) such that the mapping

$$F(x_1, \dots, x_{m+1}) := (f_0(x_1, \dots, x_{m+1}, \hat{x}_{m+2}, \dots, \hat{x}_n), \dots, f_m(x_1, \dots, x_{m+1}, \hat{x}_{m+2}, \dots, \hat{x}_n))$$

satisfies the assumptions of the inverse theorem (see theorem 1.4.2). But then there exist  $\varepsilon > 0, \delta > 0$  and  $K > 0$  such that for all  $\eta \in ]0, \delta]$  there exists  $(x_1, \dots, x_{m+1}) \in \mathbb{R}^{m+1}$  such that  $F(x_1, \dots, x_{m+1}) = (-\eta, 0, \dots, 0)$  and  $\|(x_1, \dots, x_{m+1}) - (\hat{x}_1, \dots, \hat{x}_{m+1})\| \leq K\eta$ . But this means that  $\hat{x}$  is not a local minimum in (P). Contradiction.  $\square$



### 1.4.6 An orthogonal basis of a symmetric matrix

Let  $A = (a_{ij})_{i,j=1}^n$  be a symmetric matrix ( $a_{ij} = a_{ji}$ ) and let

$$Q(x) = \langle Ax, x \rangle = \sum_{i,j=1}^n a_{ij}x_i x_j$$

be the quadratic form corresponding to the matrix  $A$ .

**Theorem 1.4.6** (*Jacobi*)

*There exists in  $\mathbb{R}^n$  an orthonormal basis  $\{f_i\}_{i=1}^n$  of eigenvectors of  $A$  in which the form  $Q$  is represented as*

$$Q(x) = \sum_{j=1}^n \lambda_j \langle x, f_j \rangle^2.$$

**Proof:** Consider the extremal problem

$$\langle Ax, x \rangle \rightarrow \max; |x|^2 = 1 \tag{P_1}$$

A solution  $f_1$  of  $(P_1)$  exists (by the compactness principle IA). From the Lagrange multiplier rule it follows immediately that  $Af_1 = \lambda_1 f_1$ . Consider the problem

$$\langle Ax, x \rangle \rightarrow \max; |x|^2 = 1, \langle x, f_1 \rangle = 0. \tag{P_2}$$

The solution  $f_2$  exists (by IA) and from the Lagrange multiplier rule we readily find that  $Af_2 = \lambda_2 f_2$ , and so on. As a result we obtain an orthonormal basis  $\{f_i\}_{i=1}^n$  and  $(*)$  follows from the definitions.  $\square$

### 1.4.7 Separation theorem

**Theorem 1.4.7** (*Minkowski*)

*Let  $C$  be a closed convex set in  $\mathbb{R}^n$  and  $x_0$  a point in  $\mathbb{R}^n$  which does not belong to  $C$ . Then there exists a hyperplane which separates  $x_0$  and  $C$ .*

**Proof:** The function  $f(x) := \|x - x_0\|$  is coercive and continuous, hence (by I A) the problem  $f(x) \rightarrow \min; x \in C$  has a solution  $\hat{x}$ . The hyperplane perpendicular to the vector  $x_0 - \hat{x}$  passing through the point  $\hat{x}$  is a desirable one (because if there exists a point  $\bar{x} \in C$  lying on the same side of the hyperplane as  $x_0$  then on the segment  $[\hat{x}, \bar{x}]$  there exists a point whose distance to  $x_0$  is less than  $d_X(x_0, \hat{x})$ ).  $\square$

## 1.5 Tools

At the beginning of this section let us make a little trip to the first floors of an imaginary Museum of Mathematics.

We will begin our excursion from the base, from the foundation of the whole building. According to the point of view of mathematicians of previous generations “the foundation of the whole of mathematics is set theory” (I quote words of A. N. Kolmogorov).

Georg Cantor was the architect and the first builder of the whole structure. The conception of mathematics as a superstructure on set theory was developed by Hilbert, Weyl and many other mathematicians.

N.Bourbaki made an attempt to realize this program and to build Mathematics from its foundations to the highest peaks starting from the ideas of Cantor-Hilbert-Weyl. This plan was not realized (and the majority of modern mathematicians have serious doubts to the possibility of the realization of such a program). But the attempt of Bourbaki played a great role in the history of mathematics.

To put it shortly the conception of Hilbert-Bourbaki can be expressed in the following words: “*Mathematics is the theory of structures*”.

The simplest structures in mathematics are sets either with a few relations or with a system of sets, which define these structures. So the first floor of our imaginary Museum consists of the lodgements in which the simplest fundamental structures are represented. These are algebraic structures (groups, rings, linear spaces and so on), topological spaces, measure spaces etc.

Among the founders of algebraic structures were Galois, Jordan, Kronecker, Cayley. One of the first mathematicians who defined the notion of linear space was Grassman. The primary topological notions (such as closure, density, nowhere density and many others) were introduced by Cantor himself. The modern definition of a topological space as a set with a system of open sets was given by Alexandroff and Urysohn. The definition of measurable space is due to Borel and Lebesgue.

The theory of extremal problems is located on the second floor of our Museum, where the *mixed structures* are located. In the set of such mixed structures there is the class of *linear topological spaces*. A linear topological space is the structure in which the algebraic and topological properties are considered conjointly. One of the most important subclasses of the class of linear topological spaces, is the collection of *linear normed spaces*. This is the real area of the modern theory of extremal problems.

The main tools of the theory are differential, convex and nonsmooth calculus. Now our goal is to familiarize the reader with these tools.

### 1.5.1 Differential calculus

The fundamental notion of differential calculus is the notion of a differential. It is based in turn on the definition of *a derivative*.

For one dimensional functions derivatives were defined by Newton and Leibniz. Newton's definition is close to the modern one (given by Cauchy). Newton possessed the idea of a limit and in fact he defined the derivative of a function as a limit of the ratio of increment of the function over increment of the argument. The definition of Leibniz was more vague. He operated with the term "infinitesimal value", appealing to the intuition. (It is very interesting that now the so-called "Nonstandard analysis" gives the exact and many-sided sense to the idea of infinitesimal value).

Lagrange introduced the notion of derivative in the infinite dimensional case. It is the so-called *first variation of a functional*. This gave Euler the idea to call the whole corresponding direction in the theory of extrema the calculus of variations. Lagrange did not give the exact definition of variation in the modern sense of the word. Such definitions were given by students of Jack

Hadamard, Gateau (the weak derivative) and Fréchet (the strong derivative). Giving his definition (in 1912), Fréchet was sure that it was new even for the finite dimensional case. Specializing his general definition to the finite dimensional case and obtaining the definition which is known now by all mathematicians, Fréchet wrote, that it was the “differential á mon sens”. But this was not quite right, because the correct definitions of derivative and differential of a function of many variables were given by Weierstrass in his lectures at the eighties of the 19th century. These lectures were published in the thirties of our century. The correct definitions of the derivative in the multidimensional case appear also at the beginning of the century in some German and English text-books (Scholz, Young) under the influence of Weierstrass.

In the fifties in the school of Bourbaki the notion of *strict derivative* appeared, which is very convenient in the theory of extremal problems.

*Definition 1.* Let  $X, Y$  be normed spaces, let  $U$  be an open neighbourhood of a point  $\hat{x}$  (we will denote it  $U \in \mathcal{O}(\hat{x}, X)$ ). The mapping  $U \rightarrow Y$  is called

a) *differentiable (in the sense of Fréchet) at the point  $\hat{x}$* , if there exists a linear continuous operator  $\Lambda : X \rightarrow Y$  ( $\Leftrightarrow \Lambda \in \mathcal{L}(X, Y)$ ) such that

$$F(\hat{x} + x) = F(\hat{x}) + \Lambda x + r(x), \quad \|r(x)\|_Y / \|x\|_X \rightarrow 0, \text{ if } \|x\|_X \rightarrow 0.$$

b) *strictly differentiable at  $\hat{x}$*  if it is differentiable in  $\hat{x}$  and if for any  $\varepsilon > 0$  there exists  $\delta > 0$  such that

$$\|F(x_2) - F(x_1) - \Lambda(x_2 - x_1)\|_Y \leq \varepsilon \|x_2 - x_1\|_X \quad \forall (x_1, x_2)$$

such that

$$\|x_1 - \hat{x}\|_X < \delta, \|x_2 - \hat{x}\|_X < \delta.$$

The operator  $\Lambda$  in these definitions is unique. It is called the *derivative (in the sense of Fréchet)* of  $F$  at  $\hat{x}$  and is denoted  $F'(\hat{x})$ . If  $F$  is differentiable at  $\hat{x}$  we write  $F \in D^1(\hat{x}, Y)$ , or simply  $F \in D(\hat{x})$ , and if  $F$  is strictly differentiable at the point  $\hat{x}$  we write  $F \in SD^1(\hat{x}, Y)$ . From the definitions it follows that if  $F \in D^1(\hat{x}, Y)$  then  $F$  is continuous at  $\hat{x}$ , and if  $F \in SD^1(\hat{x}, Y)$  then  $F$  is continuous in a neighbourhood of  $\hat{x}$ .

The following two theorems of differential calculus are the most important for us.

**Theorem 1.5.1** (*On Superposition*)

Let  $X, Y, Z$  be normed spaces,  $U \in \mathcal{O}(\hat{x}, X)$ ,  $F : U \rightarrow Y$ ,  $V \in \mathcal{O}(F(\hat{x}), Y)$ ,  $G : V \rightarrow Z$ ,  $H = G \circ F$ . If  $F \in D^1(\hat{x}, Y)$ ,  $G \in D^1(F(\hat{x}), Z)$  (resp.  $F \in SD^1(\hat{x}, Y)$ ,  $G \in SD^1(F(\hat{x}), Z)$ ), then  $H \in D^1(\hat{x}, Z)$  (resp.  $SD^1(\hat{x}, Z)$ ) and

$$H'(\hat{x}) = G'(F(\hat{x})) \circ F'(\hat{x}).$$

This theorem (due to Leibniz) follows in an elementary way from the definitions.

**Theorem 1.5.2** (*On inverse functions*)

Let  $X$  and  $Y$  be Banach spaces,  $U \in \mathcal{O}(\hat{x}, X)$ ,  $F : U \rightarrow Y$ ,  $F(\hat{x}) = 0$ ,  $F \in SD^1(\hat{x}, Y)$  and  $F'(\hat{x})X = Y$ . Then there exist  $V \in \mathcal{O}(\hat{x}, X)$ ,  $W \in \mathcal{O}(0, Y)$  and a mapping  $\varphi : V \times W \rightarrow X$  and  $K > 0$  such that  $F(\varphi(x, y)) = y$  and  $\|\varphi(x, y)\|_X \leq K\|y - F(x)\|_Y$  for all  $(x, y) \in V \times W$ .

This result follows from the generalized iterative principle and the the right inverse theorem. The idea of the proof is the same as in the finite dimensional case (Dini's theorem).

Let  $\mathcal{M}$  be a subset of a normed space  $X$ . A vector  $x$  is said to be a tangent vector to the set  $\mathcal{M}$  at the point  $\hat{x} \in \mathcal{M}$  if there exists a mapping  $r : [-1, 1] \rightarrow X$  such that  $\hat{x} + \lambda x + r(\lambda) \in \mathcal{M} \quad \forall \lambda \in [-1, 1]$  and  $\|r(\lambda)\|/|\lambda| \rightarrow 0$  if  $\lambda \rightarrow 0$ .

**Corollary 1.5.3** (*Tangent space theorem*)

Under the assumptions of the preceding theorem, the set of all tangent vectors of  $\{x | F(x) = 0\}$  at the point  $\hat{x}$  equals  $\text{Ker}F'(\hat{x})$ .

**1.5.2 Convex analysis**

Convex analysis is a section of mathematics in which convex sets, convex functions and convex extremal problems are studied. Convex analysis is located between geometry and analysis. A geometrical foundation of convex analysis was created in antiquity, but the analytical part of it was developed only in the middle of our century.

Convexity plays an important role in the theory of extremal problems. The following phenomena and facts connected with convexity will be the most essential for us in future: duality, convex calculus, extremal points and decomposition.

Let  $X$  be linear space. A set  $A \subset X$  is called *convex* if for any  $x_1, x_2 \in A$  the whole segment  $[x_1, x_2] := \{x \mid \alpha x_1 + (1 - \alpha)x_2, 0 \leq \alpha \leq 1\}$  belongs to  $A$ .

To a function  $f : X \rightarrow \mathbb{R} \cup \{\infty\}$  the following two sets are associated:

$$\text{dom } f := \{x \in X \mid f(x) < +\infty\},$$

$$\text{epi } f := \{(\alpha, x) \mid \alpha \geq f(x), x \in \text{dom } f\}.$$

The first one is called the *effective domain of  $f$* , the second is called the *epigraph of  $f$* . The function  $f$  is called *convex (resp. closed)*, if its epigraph is a convex (resp. closed) set. A function  $p : X \rightarrow \mathbb{R} \cup \{+\infty\}$  is called *sublinear* if  $p(0) = 0, p(\alpha x) = \alpha p(x) \forall \alpha > 0$  and  $p(x_1 + x_2) \leq p(x_1) + p(x_2)$ , in other words if its epigraph is a convex cone with the origin as top.

There are two important examples of convex functions: an indicator function and a support function. The *indicator function*  $\delta A$  of a set  $A$  is equal to zero in all  $x \in A$  and is equal to  $+\infty$  in all  $x \notin A$ . ( $\delta A$  is convex iff  $A$  is a convex set.) Let  $X$  be a normed space and  $X^*$  its conjugate,  $A$  a non-empty subset of  $X$ . The function  $sA(x^*) := \sup\{\langle x^*, x \rangle \mid x \in A\}$  is said to be the *support function of  $A$* . A support function is a sublinear function.

Let us give some operations with convex functions.

— *sum*:  $(f_1 + f_2)(x) := f_1(x) + f_2(x)$ ;

— *convolution*:  $(f_1 \oplus f_2)(x) := \inf\{f_1(x_1) + f_2(x_2) \mid x_1 + x_2 = x\}$ ;

— *maximum*:  $(f_1 \vee f_2)(x) := \max\{f_1(x), f_2(x)\}$ ;

— *convex hull of minimum*:

$$(f_1(\text{co}\wedge)f_2)(x) :=$$

$$\min\{\alpha f_1(x_1) + (1 - \alpha)f_2(x_2) \mid 0 \leq \alpha \leq 1, \alpha x_1 + (1 - \alpha)x_2 = x\}$$

There are two main notions in convex analysis: conjugate functions and subdifferentials.

If two vectorspaces  $X, \Xi$  and a bilinear pairing  $\langle, \rangle: X \times \Xi \rightarrow \mathbb{R}$  are given then the formula

$$f^*(\xi) = \sup_{x \in X} (\langle \xi, x \rangle - f(x)).$$

associates to each function  $f: X \rightarrow \overline{\mathbb{R}}$  a function  $f^*: X \rightarrow \overline{\mathbb{R}}$  which is called the *conjugate function of  $f$*  (or *Legendre transformation* or *Young-Fenchel transformation* of  $f$ ). Analogously the conjugate function  $g^*: X \rightarrow \overline{\mathbb{R}}$  for a function  $g: \Xi \rightarrow \overline{\mathbb{R}}$  is defined. We can apply this to the natural pairing  $X \times X^* \rightarrow \mathbb{R}: (x, x^*) \rightarrow \langle x^*, x \rangle$  which one has for each normed vector space  $X$ . In the remainder of this section  $X$  is a normed space.

The set  $\partial p$  in  $X^*$  defined by

$$\partial p := \{x^* \in X^* \mid \langle x^*, x \rangle \leq p(x) \quad \forall x \in X\}$$

is said to be the *subdifferential of a sublinear function  $p$*

The following formulae hold true:  $(\delta A)^* = sA$ ;  $N(x) := \|x\| \Rightarrow N^* = \delta B X^*$ . We see that the function conjugate to  $\|\cdot\|$  is the indicator function of the unit ball in the conjugate space.

The set

$$\partial f(\hat{x}) := \{x^* \in X^* \mid \langle x^*, x - \hat{x} \rangle \leq f(x) - f(\hat{x}) \quad \forall x \in X\}.$$

is called the *subdifferential of  $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$  at  $\hat{x}$* . The following formula holds:  $\partial N(0) = B X^*$  and  $x \neq 0 \Rightarrow \partial N(x) = \{x^* \mid \|x^*\| = 1, \langle x^*, x \rangle = \|x\|\}$ .

Now we formulate the most important theorems in convex analysis.

**Theorem 1.5.4** (*Fenchel-Moreau*)

A function  $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$  is equal to its second conjugate  $f^{**}$  iff it is convex and closed.

**Theorem 1.5.5** (*Moreau-Rockafellar*)

Let  $f_i: X \rightarrow \mathbb{R} \cup \{+\infty\}$ ,  $i = 1, 2$  be convex functions and let there exist a point, where one of them is finite and the other is continuous. Then  $\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x)$  for all  $x \in X$ .

**Theorem 1.5.6** (*Formula of Dubovitsky-Milyutin*)

Let functions  $f_i : X \rightarrow \mathbb{R} \cup \{+\infty\}$ ,  $i = 1, 2$  be continuous in a point. If  $f_1(x) = f_2(x)$  then

$$\partial(f_1 \vee f_2)(x) = (\partial f_1(\text{co}\wedge)\partial f_2)(x).$$

**Theorem 1.5.7** (*Decomposition theorem*)

Let  $T$  be a compact topological space, let  $F : T \times \mathbb{R}^n \rightarrow \mathbb{R}$  be a mapping such that  $F(\cdot, x) : T \rightarrow \mathbb{R}$  is an upper semicontinuous function for every  $x$  and  $F(t, \cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  is convex for every  $t \in T$  and let  $m := \inf_x \max_{t \in T} F(t, x) > -\infty$ . Then, there exists a constant  $r \in \mathbb{N}$ ,  $r \leq n + 1$  and  $r$  points  $\{\tau_i\}_{i=1}^r$ ,  $\tau_i \in T$  such that  $m = \inf_x \max_{1 \leq i \leq r} F(\tau_i, x)$ .

We now present the proof of the theorem of Fenchel-Moreau.

**Proof:** Necessity. Let  $f^{**} = f$ . Then by definition

$$f(x) \stackrel{\text{def}}{=} \sup_{x^*} (\langle x^*, x \rangle - f^*(x^*)),$$

that is,  $\text{epi } f$  is the intersection of the epigraphs of the affine functions  $a(\cdot; x^*) : X \rightarrow \mathbb{R}$  which are defined by  $a(x; x^*) = \langle x^*, x \rangle - f^*(x^*)$  for all  $x \in X, x^* \in X^*$  with  $f^*(x^*) \in \mathbb{R}$ . Hence  $f$  is a convex closed function.

*Sufficiency.* Let  $f$  be a closed convex function. Then, by a separation theorem, the epigraph of  $f$  is the intersection of all closed halfspaces in  $X \times \mathbb{R}$  which contain it. As at least one of them has to be the epigraph of an affine function on  $X$  (that is, not vertical) it can be deduced that this intersection does not increase if we omit all halfspaces which are not epigraphs of affine functions (the vertical ones). That is, essentially by definition of  $f^{**}$ , one gets  $f = f^{**}$ .  $\square$

**Remark.** The Fenchel–Moreau theorem expresses the idea of duality we discussed at the beginning of the section. On the one hand a closed convex function  $f$  is a union of elementary functions:

$$\text{epi } f = \cup \{ \text{epi } e(\cdot; x, f(x)) \mid x \in \text{dom } f \}$$

( $e(\cdot; \xi, \alpha) = \alpha$  if  $x = \xi$  and  $\infty$  if  $x \neq \xi$ ) and on the other hand it is the supremum of a family of affine functions.



### 1.5.3 Nonsmooth analysis

The theory of smooth extremal problems is based on differential calculus, the theory of convex problems is based on convex analysis. The development of the theory of extremal problems in our century led to a new branch of calculus which is called nonsmooth analysis.

The first definition of the notion of an analogue of the differential for nonsmooth and nonconvex functions was given by F. Clarke in 1975.

Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a Lipschitzian function on  $\mathbb{R}^n$ . Then, (by the well-known theorem of Rademacher) it has a differential almost everywhere. The convex hull of all limits of gradients of the function  $f$  when points tend to the given point  $\hat{x}$  is called *Clarke's subdifferential of  $f$  at  $\hat{x}$*  and denoted  $\partial_C f(\hat{x})$ .

The definition of Clarke played a great role in the development of nonsmooth calculus. We see that this definition is in fact based on the idea of an approximation of  $f$  which is due to Fermat. In the thirties of our century many mathematicians studied approximative cones of nonlinear objects. The corresponding cones were called *contingency cones*. Such cones were subject of investigations of Bouligan, Hadamard, Kolmogorov and others. The definition of Clarke is closely connected with these considerations.

In the majority of situations Clarke's subdifferential is adequate, but sometimes it is too wide. For example, this subdifferential does not distinguish (at the zero point) the following two functions:  $x \rightarrow |x|$  and  $x \rightarrow -|x|$ . In both cases Clarke's subdifferential is equal to the segment  $[-1, 1]$ .

A fruitful approach to the calculus of nonsmooth mappings was proposed by Mordukhovich and Ioffe.

The approach of Mordukhovich is based on the notion of *normal* to a subset of a finite-dimensional space. He considers limits of normals to the graph of a mapping (from  $\mathbb{R}^n \rightarrow \mathbb{R}^m$ ). Such limits form the so-called *normal cone* to the epigraph of a mapping  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  at a point  $(\hat{x}, F(\hat{x}))$ . The corresponding cones are different for functions  $|x|$  and  $-|x|$ . In the first case it coincides with Clarke's subdifferential, in the second it consists only of two points  $\pm 1$ .

The definition of a subgradient of a function is the following: it is the set of

$x^* \in \mathbb{R}^{n^*}$  such that  $(x^*, -1)$  is proportional to the limits of gradients of the function equal to the distance from a point to the epigraph of the function under consideration.

Ioffe gave another definition of the same notion (which can be realized in the infinite dimensional case). This definition is based on the idea of approximation (the notation of Ioffe's subdifferential is  $\partial_a f(\hat{x})$ ):

$$\partial_a f(\hat{x}) = \limsup_{x \rightarrow \hat{x}} \partial_a^- f(x), \quad \partial_a^- f(x) := \{x^* \mid f'_-(\hat{x}, x) \geq \langle x^*, \hat{x} \rangle\},$$

where  $f'_-(\hat{x}, x) := \liminf_{t \rightarrow +0} t^{-1}(f(\hat{x}+tx) - f(\hat{x}))$  is the so-called *directional derivative in sense of Hadamard*.

The calculus of nonsmooth mappings and nonconvex mappings is based on these definitions and corresponding notions of coderivatives.



## Chapter 2

# THEORY

### 2.1 Lagrange principle

“On peut les réduire à ce principe général. Lors qu’une fonction de plusieurs variables doit être un maximum ou minimum , et qu’il y a entre ces variables une ou plusieurs équations , il suffira d’ajouter à la fonction proposée les fonctions, qui doivent être nulles , multiplier chacune par une quantité indéterminée, et la chercher ensuite le maximum ou minimum comme si les variables étaient indépendantes; les équations qu’on a trouvé, serviront à déterminer toutes les inconnues.”

J.L.Lagrange

The honor of the creation of the general strategy for investigating extremal problems with constraints is due to Lagrange. For all problems he met, Lagrange used a unified approach which was expressed in the words we quoted at the epigraph to this paragraph.

Lagrange’s pivotal idea can be applied to an extremely broad class of extremal problems of a diverse nature. We will try to explain this phenomenon. The first sketch of this explanation is the following.

In the majority of problems where the idea of Lagrange can be realized variables are divided into two parts. Functionals and mappings are smooth

over variables of the first group and are convex over the second one. Such problems we call *smooth-convex problems*. For such problems we prove a necessary condition which fulfils the generalized idea of Lagrange. Existence principles give the possibility to prove existence of Lagrange multipliers which fulfil the generalized idea of Lagrange.

### 2.1.1 Lagrange principle for smooth-convex problems

Let  $X, Y$  be Banach spaces,  $U$  a topological space. Consider the problem:

$$f_0(x) \rightarrow \min, F(x, u) = 0, u \in U, \quad (P)$$

where  $f_0 : W \rightarrow \mathbb{R}, F : W \times U \rightarrow Y$ ,  $W$  is an open subset of  $X$ . We say that a point  $(\hat{x}, \hat{u}) \in W \times U$  with  $F(\hat{x}, \hat{u}) = 0$  is a *strong local minimum of the problem (P)* if there exists  $\delta > 0$  and  $\mathcal{V} \in \mathcal{O}(\hat{u}, U)$  such that for any pair  $(x, u)$  for which  $F(x, u) = 0, u \in \mathcal{V}$  and  $\|x - \hat{x}\| < \delta$  the following inequality holds true:  $f_0(x) \geq f_0(\hat{x})$ .

The function

$$\mathcal{L}((x, u), \lambda, \lambda_0) := \lambda_0 f_0(x) + \langle \lambda, F(x, u) \rangle$$

is called *the Lagrange function of the problem (P)*. The number  $\lambda_0$  and the element  $\lambda \in Y^*$  are called *Lagrange multipliers*.

We call the mapping  $F$  in (P) *smooth-convex* at  $(\hat{x}, \hat{u})$  if  $F(\cdot, u) \in SD(\hat{x}, Y) \forall u \in U$  and  $F(x, U)$  is a convex set in  $Y \forall x \in W$ .

The mapping  $F$  is called *regular at the point  $(\hat{x}, \hat{u})$*  if the space  $F_x(\hat{x}, \hat{u})X$  is closed in  $X$  and has finite codimension in  $Y$  and *totally regular* if  $\text{codim } F_x(\hat{x}, \hat{u})X = 0$  or in other words if  $F_x(\hat{x}, \hat{u})X = Y$ .

**Theorem 2.1.1** (*Lagrange principle for smooth-convex problems*) *Let in the problem (P)  $f_0 \in D(\hat{x})$  and let the mapping  $F$  be smooth-convex and regular at  $(\hat{x}, \hat{u})$ . Then if  $(\hat{x}, \hat{u})$  is a strong local minimum of (P) the Lagrange principle at  $(\hat{x}, \hat{u})$  holds true; if  $F$  is moreover totally regular at  $(\hat{x}, \hat{u})$  then  $\lambda_0 \neq 0$ .*

The Lagrange principle for the problem (P) means that there exist Lagrange multipliers  $\lambda \in Y^*, \lambda_0 \in \mathbb{R}, \lambda_0 \geq 0$  (not equal to zero simultaneously) such

that in the smooth problem (where  $\mathcal{L}((x, u), \lambda, \lambda_0) = \lambda_0 f_0(x) + \langle \lambda, F(x, u) \rangle$ )

$$\mathcal{L}((x, \hat{u}), \lambda, \lambda_0) \rightarrow \min$$

“comme si les variables  $[x]$  étaient indépendantes” Fermat’s theorem holds true:  $\mathcal{L}_x((\hat{x}, \hat{u}), \lambda, \lambda_0) = 0$  (*stationarity condition*) ( $\Leftrightarrow \lambda_0 f'_0(\hat{x}) + (F_x(\hat{x}, \hat{u}))^* \lambda = 0$ ). And in the convex problem

$$\mathcal{L}((\hat{x}, u), \lambda, \lambda_0) \rightarrow \min, u \in U$$

the *minimum principle* is satisfied

$$\min_{u \in U} \mathcal{L}((\hat{x}, u), \lambda, \lambda_0) = \mathcal{L}((\hat{x}, \hat{u}), \lambda, \lambda_0) (\Leftrightarrow \langle \lambda, F(\hat{x}, u) \rangle \geq 0 \quad \forall u \in U).$$

**Remark 2.1.2** *The conditions of this version are too strong for some problem types of interest such as optimal control problems; however the result can be suitable strengthened.*

**Proof:** a) Denote  $\Lambda := F_x(\hat{x}, \hat{u})$ ,  $L_0 := \text{Im} \Lambda = F_x(\hat{x}, \hat{u})X$ ,  $A = F(\hat{x}, U)$ ,  $C := L_0 + A$ . Let  $\pi : Y \rightarrow Y/L_0$  be a canonical projection.

There are two possibilities: (I)  $0 \notin \text{int} \pi C$  (degenerate case) and (II)  $0 \in \text{int} \pi C$  (nondegenerate case). At first we consider (I).

The factor-space  $Z := Y/L_0$  is by the condition of regularity a finite dimensional space. From the finite dimensional separation theorem it follows that there exists a vector  $z^* \in Z^*$  such that  $\langle z^*, z \rangle \geq 0 \quad \forall z \in \pi C$  (i). Denote  $\pi^*$  the conjugate operator  $\pi^* : Z^* \rightarrow Y^*$  and  $\lambda := \pi^*(z^*)$ . It is evident that  $\lambda \neq 0$  (because  $\pi$  is a surjective operator) and then

$$\begin{aligned} \langle \lambda, \Lambda x + F(\hat{x}, u) \rangle &\stackrel{\text{def}}{=} \langle \pi^*(z^*), \Lambda x + F(\hat{x}, u) \rangle \stackrel{\text{Id}}{=} \\ &\langle z^*, \pi(\Lambda x + F(\hat{x}, u)) \rangle \stackrel{\text{(i)}}{\geq} 0 \quad \forall x \in W, u \in U. \end{aligned}$$

From this inequality we obtain that  $\Lambda^* \lambda = 0$  and  $\langle \lambda, F(\hat{x}, u) \rangle \geq 0$ , i.e. the stationarity condition with  $\lambda_0 = 0$  and the minimum principle hold true. The Lagrange principle in the degenerate case is proved.

Consider the nondegenerate case. If  $0 \in \text{int} \pi C$  then there exist  $m \in \mathbb{N}$ ,  $\{z_i\}_{i=1}^m$ ,  $\bar{\alpha}_i > 0$ ,  $1 \leq i \leq m$  such that  $\sum_{i=1}^m \bar{\alpha}_i z_i = 0$ ,  $z_i = \pi F(\hat{x}, v_i)$  and cone

$\{z_i\}_{i=1}^m = Z$ . Then by definition  $\sum_{i=1}^m \bar{\alpha}_i F(\hat{x}, v_i) \in L_0$ , hence there exists an element  $\bar{\xi} \in X$  such that  $\Lambda \bar{\xi} + \sum_{i=1}^m \bar{\alpha}_i F(\hat{x}, v_i) = 0$  (i'). For each  $v_0 \in U$  we define the map  $\phi_{v_0} : X \times \mathbb{R}^{m+1} \rightarrow Y$ ,

$$\phi_{v_0}(x, \alpha) = (1 - \sum_{i=0}^m \alpha_i) F(\hat{x} + x, \hat{u}) + \sum_{i=0}^m \alpha_i F(\hat{x} + x, v_i)$$

for all  $x \in X$  and all  $\alpha = (\alpha_0, \dots, \alpha_m) \in \mathbb{R}^{m+1}$ .

From the superposition theorem it follows that  $\phi_{v_0} \in SD((0,0), Y)$  and  $\phi'_{v_0}(0,0)[x, \alpha] = \Lambda x + \sum_{i=0}^m \alpha_i F(\hat{x}, v_i)$  (ii). Assume  $\Lambda \bar{x} + F(\hat{x}, \bar{v}) = 0$  for some  $(\bar{x}, \bar{v}) \in W \times U$  (iii) and  $\varepsilon > 0$ . Then from (ii) we have  $\phi'_{\bar{v}}(0,0)[\bar{x} + \varepsilon \bar{\xi}, 1, \varepsilon \bar{\alpha}_1, \dots, \varepsilon \bar{\alpha}_m] = \Lambda(\bar{x} + \varepsilon \bar{\xi}) + F(\hat{x}, \bar{v}) + \varepsilon \sum_{i=1}^m \bar{\alpha}_i F(\hat{x}, v_i) = 0$ .

From the tangent space theorem it follows (because  $\phi$  is a totally regular operator!) that there exist  $r(\cdot) : [-1, 1] \rightarrow X$ ,  $\rho_i(\cdot) : [-1, 1] \rightarrow \mathbb{R}$ ,  $0 \leq i \leq m$ ,  $r(t) = o(t)$ ,  $\rho_i(t) = o(t)$  such that

$$\phi_{\bar{v}}(t(\bar{x} + \varepsilon \bar{\xi})r(t), t + \rho_0(t), \varepsilon t \bar{\alpha}_1 + \rho_1(t), \dots, \varepsilon t \bar{\alpha}_m + \rho_m(t)) = 0 \quad \forall t \in [-1, 1].$$

From the condition of convexity (and using that for small  $t > 0$ ,  $t + \rho_0(t) > 0$ ,  $\varepsilon t \bar{\alpha}_i + \rho_i(t) > 0$ !) we obtain that there exists  $\delta > 0$  and a family  $\{u(t)\}_{t \in [-\delta, \delta]}$  such that  $F(\hat{x} + t\bar{x} + \varepsilon t \bar{\xi} + r(t), u(t)) = 0$  for all  $t \in [0, \delta]$ . This means that the pair  $(\hat{x} + t\bar{x} + \varepsilon t \bar{\xi} + r(t), u(t))$ ,  $t \in [0, \delta]$  is admissible and (because  $(\hat{x}, \hat{u})$  is a local minimum) there exists  $\delta_1 > 0$  such that  $f_0(\hat{x} + t\bar{x} + \varepsilon t \bar{\xi} + r(t)) \geq f_0(\hat{x})$ ,  $\forall t \in [0, \delta_1] \Rightarrow \langle f'_0(\hat{x}), \bar{x} \rangle \geq 0$  (iv) (because  $\varepsilon > 0$  is an arbitrary number).

We assume (iii) with  $\bar{v} = \hat{u} (\Leftrightarrow \bar{x} \in \ker \Lambda)$ , then from (iv) we obtain that  $f'_0(\hat{x}) \in (\ker \Lambda)^\perp$ . From the annihilator lemma it follows that there exists  $\lambda_1 \in L_0^*$  such that  $f'_0(\hat{x}) + \Lambda^* \lambda_1 = 0$  (v). Let now  $F(\hat{x}, u) \in L_0$ . Then we can find  $x(u) \in X$  such that  $\Lambda x(u) + F(\hat{x}, u) = 0$  (vi)  $\Rightarrow 0 \stackrel{(iv)}{\leq} \langle f'_0(\hat{x}), x(u) \rangle \stackrel{(v)}{=} -\langle \Lambda^* \lambda_1, x(u) \rangle \stackrel{\text{Id}}{=} -\langle \lambda_1, \Lambda x(u) \rangle \stackrel{(vi)}{=} \langle \lambda_1, F(\hat{x}, u) \rangle$ .

And at last from the separation theorem (we take  $\bar{C} = \text{co}(C \cup \bar{B})$ , where  $\bar{B}$  is a ball with center  $\eta \in L_0$ ,  $\langle \lambda_1, \eta \rangle > 0$ , which does not intersect with  $\{y \mid \langle \lambda_1, y \rangle = 0\}$ , then  $\text{int } \bar{C} \neq \emptyset$  and it is possible to use the separation theorem) there exists  $\lambda \in Y^*$  such that  $\lambda|_{L_0} = \lambda_1$  and  $\langle \lambda, F(x, u) \rangle \geq 0 \quad \forall u$  ( $\Leftrightarrow$  minimum principle). From the first equality we obtain  $\langle \lambda, \Lambda x \rangle = \langle \lambda_1, \Lambda x \rangle \stackrel{\text{Id}}{=} \langle \Lambda^* \lambda_1, x \rangle \stackrel{(v)}{=} -\langle f'_0(\hat{x}), x \rangle \quad \forall x$ . This is the stationarity condition. Thus the

Lagrange principle is proven.  $\square$

## 2.2 Perturbations

In many cases it is convenient to include the primary extremal problem into some family of problems and consider (instead of a problem  $f(x) \rightarrow \min$ ) a family of problems  $F(x, y) \rightarrow \min$  (over  $x$ ) or instead of a problem  $f_0(x) \rightarrow \min, F(x) = 0$  a family of problems  $f_0(x) \rightarrow \min, F(x) = y$  and so on.

Let us consider the last situation (which is going back to Lagrange) in the finite dimensional case (we suppose that  $f_0, F$  are continuously differentiable in a neighbourhood  $W$  of 0):

$$f_0(x) \rightarrow \min, F(x) = 0 \quad (\Leftrightarrow f_i(x) = 0, 1 \leq i \leq m, x \in \mathbb{R}^n). \quad (P_0)$$

This is called the *primal problem* and

$$f_0(x) \rightarrow \min, F(x) = y \quad (P_y)$$

is its *perturbation*. According to the Lagrange principle if  $F'(0)\mathbb{R}^n = \mathbb{R}^m$  and  $0 \in \text{locmin}(P_0)$  there exists a  $\lambda \in \mathbb{R}^{m*}$  such that

$$0 = \mathcal{L}_x(0, \lambda, 1) \Leftrightarrow f'_0(0) + \langle \lambda, F'(0) \rangle = 0 \Leftrightarrow f'_0(0) + (F'(0))^* \lambda = 0.$$

If  $f_0$  and  $F$  belong to  $C^2(W)$  then the following condition is necessary for a local minimum of  $(P_0)$  at the point 0: the operator  $\mathcal{L}_{xx}(0, \lambda, 1)$  restricted to the linear space  $\ker F'(0)$  must be nonnegative. If  $f_0$  and  $F$  belong to  $C^2(W)$  and  $F(0) = 0$ , then the following condition is sufficient for a local minimum of  $(P_0)$  at the point 0:  $\mathcal{L}_x(0, \lambda, 1) = 0$  and  $\mathcal{L}_{xx}(0, \lambda, 1)|_{\ker F'(0)} > 0$ . If  $f_0$  and  $F$  satisfy these assumptions (sufficient for minimality) it is possible to construct a family of extremals depending on  $y$ . More precisely there exists a smooth mapping  $y \rightarrow (x(y), \lambda(y))$  such that  $F(x(y)) = y, \mathcal{L}_x(x(y), \lambda(y), 1) = 0, \mathcal{L}_{xx}(x(y), \lambda(y), 1)|_{\ker F'(x(y))} > 0$  (in some neighbourhood of  $0_{\mathbb{R}^m}$ ).

This fact follows from the implicit function theorem applied to the mapping:  $(x, \lambda) \rightarrow (\mathcal{L}_x(x, \lambda, 1), F(x))$ .

This result has an infinite dimensional generalization.



**Theorem 2.2.1** *In the problem  $(P_0)$  we suppose that  $X, Y$  are Banach spaces,  $W \in \mathcal{O}(\hat{x}, X)$ ,  $f_0 : W \rightarrow \mathbb{R}$ ,  $F : W \rightarrow Y$ ,  $f_0 \in C^2(W)$ ,  $F \in C^2(W, Y)$ ,  $F(\hat{x}) = 0$ ,  $\text{Im}F'(\hat{x}) = Y$  and for certain Lagrange multipliers*

$$\hat{\lambda} \in Y^* \text{ one has } \mathcal{L}_x(\hat{x}, \hat{\lambda}, 1) = 0, \mathcal{L}_{xx}(\hat{x}, \hat{\lambda}, 1)[h, h] \geq \alpha \|h\|^2 \forall h \in \ker F'(\hat{x}).$$

*Then there exists a neighbourhood  $V \in \mathcal{O}(0, Y)$  and a mapping  $y \rightarrow (x(y), \lambda(y))$  of the class  $C^1$  on  $V$  such that  $\mathcal{L}_x(x(y), \lambda(y), 1) = 0$ ,  $F(x(y)) = y$  and for  $S(y) \stackrel{\text{def}}{=} f(x(y))$  the following formula holds:  $S'(0) = -\lambda(0)$ .*

**Proof:** This result can be extended to a wide class of smooth convex problems. For example, for the generalization of this result to problems of mathematical programming see the article of Burzev (Mat. Doklady T. 185, N5, 1994, 79-102).

Being applied to the simplest problem of the calculus of variations:

$$\mathcal{B}(x(\cdot), \tau, \xi) = \int_{t_0}^{\tau} L(t, x(t), \dot{x}(t)) dt \rightarrow \min, x(t_0) = x_0, x(\tau) = \xi, \quad (P_{\tau\xi})$$

where  $L$  is a  $C^2$ -function and  $x(\cdot)$  runs over the  $C^1$ -functions, we obtain the well-known result of Weierstrass about fields of extremals:

$$\{x(\cdot, \lambda)\}_{\lambda \in [-\delta, \delta]}.$$

We consider the  $S$ -function of this perturbation, that is,  $S(\tau, \xi)$  is the value of the problem  $(P_{\tau\xi}) \forall \tau, \xi$ . From the formula  $S'(y) = -\lambda(y)$  we obtain that

$$\frac{\partial S(\tau, \xi)}{\partial \xi} = -\lambda(\tau, \xi) = p(\tau, \xi) = L_{\dot{x}}(\tau, \xi, \dot{x}(\tau; \tau, \xi)),$$

where  $x(\cdot, \tau, \xi)$  is the extremal with  $x(t_0, \tau, \xi) = x_0$ , which passes through  $(\tau, \xi)$ . And  $\frac{\partial S(\tau, \xi)}{\partial \tau} = \frac{d}{d\tau} (\int_0^{\tau} L dt + \lambda(x(\tau) - \xi)) = L(\tau, \xi, \dot{x}(\tau, \tau, \xi)) - p(\tau, \xi) \dot{x}(\tau, \tau, \xi) \stackrel{\text{def}}{=} -H(\tau, \xi, p(\tau, \xi))$  where  $H(\tau, \xi, \rho)$  is the Legendre transformation of the function  $\dot{x} \rightarrow L(\tau, \xi, \dot{x})$  (we suppose that  $L$  is a quasi-regular integrand, see (3.3)). And as a result we obtain the *Jacobi-Hamilton equation*

$$\frac{\partial S(t, x)}{\partial t} + H(t, x, \frac{\partial S(t, x)}{\partial x}) = 0.$$

□

## 2.3 Existence, extension, relaxation

Ich bin überzeugt, daß es möglich sein wird, diese Existenzbeweise durch einen allgemeinen Grundgedanken zu führen, auf den das Dirichletsche Prinzip hinweist, und der uns dann vielleicht in den Stand setzen wird, der Frage näherzutreten, *ob nicht jedes reguläre Variations problem eine Lösung besitzt, sobald hinsichtlich der gegebenen Grenzbedingungen gewisse Annahmen - [...] - erfüllt sind und nötigenfalls der Begriff der Lösung eine sinngemäße Erweiterung erfährt.*

D.Hilbert.

The subject we consider in this section is connected with the 20th problem of Hilbert. Let us ask the question: does an arbitrary extremal problem have a solution? Of course it is easy to construct counterexamples. But let us try to understand the main idea of Hilbert which he expressed in the words of our epigraph.

Usually (and this corresponds literally to the text of Hilbert) the 20th problem is treated in connection with the problem of boundary conditions for elliptic equations. But I want to extend the meaning of Hilbert's word "regular": maybe (in mind) Hilbert treated this word as "natural", given by some vital scientific problem. And he wanted to express the confidence that all such problems have solutions (maybe in some extended sense).

To begin with we discuss the question: what are the main reasons of absence of solutions.

### 2.3.1 Examples of nonexistence

Let us consider some examples of the simplest problem in the calculus of variations:

$$\begin{aligned} \mathcal{J}(x(\cdot)) &= \int_I L(t, x(t), \dot{x}(t)) dt \rightarrow \min, \\ I &= [0, 1], \quad x(0) = x_0, \quad x(1) = x_1. \end{aligned} \tag{1}$$

We have given here the functional and constraints but it is also necessary

to define the domain of the functional. In the textbooks at the beginning of the century usually the space  $C^1(I)$  of continuously differentiable functions (with two topologies  $C^1$  and  $C$  which lead to the notions weak and strong extremum) was considered. But later it was understood that it is more fruitful to search for extrema in wider spaces.

That is the reason to consider our problem in the “widest” space  $W_1^1$  of absolute continuous functions  $x(\cdot)$  such that  $x(\cdot)$  and its derivative belong to  $L_1$ :

$$W_1^1(I) = \{x(\cdot) \in AC(I) | x(\cdot) \in L_1(I) \text{ and } \dot{x}(\cdot) \in L_1(I)\}$$

with norm defined by  $\|x(\cdot)\|_{W_1^1(I)} = \|x(\cdot)\|_{L_1(I)} + \|\dot{x}(\cdot)\|_{L_1(I)}$ .

What are the reasons which prevent the existence of a solution in  $W_1^1(I)$ ?

**Example 1** (Bolza: nonconvexity of the function  $\dot{x} \rightarrow L(t, x, \dot{x})$  (i.e. non-quasi-regularity of the integrand):

$$J_1(x(\cdot)) = \int_I ((\dot{x}^2 - 1)^2 + x^2) dt \rightarrow \inf, \quad x(0) = x(1) = 0.$$

It is clear that

$$J_1(x(\cdot)) > 0 \quad \forall x(\cdot) \in W_1^1(I), \quad x(t) \not\equiv 0.$$

On the other hand

$$\bar{x}(t) \equiv 0 \Rightarrow J_1(\bar{x}(\cdot)) = 0.$$

But if we take the sequence

$$x_n(t) = \int_0^t u_n(\tau) d\tau, \quad u_n(t) = \operatorname{sgn} \sin 2\pi nt, \quad n \in \mathbb{N}$$

it is evident that

$$x_n(\cdot) \xrightarrow{n \rightarrow \infty} 0$$

uniformly and at the same time

$$|\dot{x}_n(t)| = 1 \quad a.e.$$

and consequently

$$J_1(x_n(\cdot)) \xrightarrow{n \rightarrow \infty} 0.$$

That means that the value of the problem is equal to zero but solutions are absent. The reason of this phenomenon is nonconvexity of the function  $\dot{x} \rightarrow (\dot{x}^2 - 1)^2$ .

**Example 2** (Weierstrass: not enough increasing of the integrand):

$$J_2(x(\cdot)) = \int_I t^2 \dot{x}^2 dt \rightarrow \inf, x(0) = 0, x(1) = 1.$$

This is a famous example of Weierstrass. By means of this example Weierstrass explained the insufficiency of Riemann's arguments connected with the Dirichlet principle.

We see that

$$J_2(x(\cdot)) > 0 \quad \forall x(\cdot) \in W_1^1(I), x(0) = 0, x(1) = 1.$$

But if we take

$$x_n(t) = \begin{cases} nt, & 0 \leq t \leq 1/n, \\ 1, & t \geq 1/n, \end{cases} \quad n \in \mathbb{N}$$

we obtain

$$J_2(x_n(\cdot)) \xrightarrow{n \rightarrow \infty} 0.$$

And again the value of the problem is zero but solutions are absent. The reason is absence of increasing of the integrand in the point  $t = 0$ .

**Example 3** (harmonic oscillator: unboundness functional from below):

$$J_3(x(\cdot)) = \int_0^T (\dot{x}^2 - x^2) dt \rightarrow \inf, T > \pi, x(0) = x(T) = 0.$$

Here if we consider the sequence

$$x_n(t) = n \sin(\pi t/T), \quad n \in \mathbb{N},$$

it is easy to show that

$$J_3(x_n(\cdot)) \xrightarrow{n \rightarrow \infty} -\infty,$$

hence solutions are absent.

Let us return to the quotation of Hilbert.

“The general principle” of the proof of the existence theorems in the calculus of variations is without any doubt the principle of compactness. It consists of two components semicontinuity of functionals and compactness of constraints. Absence of semicontinuity can sometimes be eliminated by means of relaxation. This is based on Lebesgue’s observation that the functional

$$J(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt$$

is lower semicontinuous if the function  $z \rightarrow L(t, x, z)$  is convex for all  $(t, x)$  ( $L : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}$ ). Later it was understood that lower-semicontinuity of such functionals is actually the same as convexity of the functions  $z \rightarrow L(t, x, z)$ . This will be discussed in the next section.

Consider the problem

$$J(x(\cdot)) = \int_T (L(t, x(t), \dot{x}(t))) dt \rightarrow \inf, \quad x|_{\partial T} = \xi \quad (P)$$

on the class of functions  $x(\cdot) \in W_1^1(T, \mathbb{R}^m)$  where  $T$  is a bounded domain in  $\mathbb{R}^n$  with smooth boundary and  $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{nm} \rightarrow \mathbb{R}$ .

The first result in this subject was obtained by Tonelli (in 1920). He proved that in the case  $n = m = 1$  if  $L \in C^2(\mathbb{R}^3)$ , lower-semicontinuity of the functional  $J$  takes place iff all functions  $L(t, x, \cdot)$  are convex. Now we give one of the modern generalizations of this result.

Let for all  $p \in [1, \infty]$  the normed space  $W_p^1(I)$  be defined by  $W_p^1(I) = \{x(\cdot) \in AC(I) | x(\cdot) \in L_0(I) \text{ and } \dot{x}(\cdot) \in L_p(I)\}$  with norm

$$\|x(\cdot)\|_{W_p^1(I)} = \|x(\cdot)\|_{L_0(I)} + \|\dot{x}(\cdot)\|_{L_p(I)}.$$

**Theorem 2.3.1** (Marchelini-Sbordone: about semicontinuity.) *Let in the problem (P)  $T$  be a bounded domain in  $\mathbb{R}^n$  with smooth boundary, let the*

$$\text{integrand } L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

satisfy the Carathéodory condition (i. e. it is measurable as a function of  $t$  for all  $(x, z)$ , continuous as a function of  $(x, z)$  for almost all  $t$  and satisfies the inequalities:

$$0 \leq L(t, x, z) \leq \varphi(t, |x|, |z|),$$

where  $\varphi$  is locally  $L_1$  as a function of  $t$  for all  $(|x|, |z|)$  and nondecreasing as a function of  $(|x|, |z|)$ .

Then the functional  $J$  is weakly lower-semicontinuous relatively weak convergence of  $x_n(\cdot)$  to  $x(\cdot)$  in  $W_p^1(I)$ ,  $1 \leq p \leq \infty$  iff  $L(t, x, \cdot)$  is a convex function for almost all  $(t, x)$ .

Functionals  $J$  with integrands convex as functions of  $z$  are called *quasi-regular integrands*.

### 2.3.2 Relaxations and extensions

In 1930 N. Bogolyubov proved the following result: let in the problem (P)  $n = m = 1$  and let the integrand  $L$  be a smooth function, satisfying Tonelli's condition

$$L(t, x, \xi) \geq C_1 |\xi|^{1+\delta} \quad (\delta > 0).$$

Then for any function  $x(\cdot)$  there exists a sequence  $(x_m(\cdot))_m$  of continuously differentiable functions (it is possible also to suppose that  $x(t_i) = x_m(t_i)$ ,  $i = 0, 1$ ,  $m \in \mathbb{N}$ ) converging to  $x(\cdot)$  uniformly on  $T = [t_0, t_1]$  such that

$$\liminf_{m \rightarrow \infty} I(x_m(\cdot)) \leq \tilde{I}(x(\cdot)),$$

where  $\tilde{I}(x(\cdot)) = \int_{t_0}^{t_1} \tilde{L}(t, x(t), \dot{x}(t)) dt$  and  $\tilde{L}$  is the -"convexification" of  $L$  as a function of the last argument ( i.e.  $\tilde{L} = L_z^{**}$  is the second conjugate of  $z \rightarrow L(t, x, z)$ .)

In other words the functional  $\tilde{I}$  is the lower-semicontinuous extension of the functional  $I$ .

Similar extensions can be proved for the arbitrary problem of optimal control. For example, consider the problem  $(\bar{P})$  in Section 1.1.3. For this problem the extended problem has the form:

$$\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt + \ell(x(t_0), x(t_1)) \rightarrow \min, \quad \dot{x} \in Q(t, x), \quad (\tilde{P})$$

where  $L$  is a quasi-regular integrand and  $\{Q(t, s)\}_{(t,s)}$  is a family of convex subsets of  $\mathbb{R}^n$ . If  $Q$  is a continuous and compact-valued mapping we have existence of solutions for the problem, so all such problems have a solution in some extended sense. For the problem  $(\tilde{P})$  it is possible to generalize the concept of perturbation and to obtain an analog of the Jacobi–Hamilton–Bellman equation.

## Chapter 3

# APPLICATIONS

### 3.1 Necessary conditions

In this section we formulate the Lagrange principle for many concrete classes of extremal problems. Let us begin by generalizing the initial fact of the theory of extrema.

**Theorem** (Fermat)

Let  $X$  be a normed space,  $W \in \mathcal{O}(\hat{x}, X)$ ,  $f : W \rightarrow \mathbb{R}$ . If  $\hat{x}$  is a local minimum of  $f$  and  $f \in D(\hat{x})$ , then

$$f'(\hat{x}) = 0.$$

**Proof:** This theorem is an immediate corollary of the definition of Fréchet's derivative and the one dimensional case of Fermat's theorem.  $\square$

**Remark 1** The first idea of such a result was expressed by Fermat in 1636.

**Theorem 3.1.1** (*Lagrange multiplier rule for mathematical programming*)

Let  $X$  and  $Y$  be Banach spaces,  $W \in \mathcal{O}(\hat{x}, X)$ ,  $f_i : W \rightarrow \mathbb{R}, 0 \leq i \leq m$ ,  $F : W \rightarrow Y$ . If  $\hat{x}$  is a local minimum of the problem

$$f_0(x) \rightarrow \min; \quad f_i(x) \leq 0, \quad 1 \leq i \leq m, \quad F(x) = 0 \quad (P_1)$$



where  $f_0 \in D(\hat{x})$ ,  $f_i \in SD(\hat{x}) \forall i \in \{1, \dots, m\}$ ,  $F \in SD(\hat{x})$  and  $F'(\hat{x})X$  is a closed subspace of  $Y$ , then the Lagrange principle for  $(P_1)$  holds true. This means that there exist Lagrange multipliers  $(\lambda_i)_{i=0}^m, \lambda \in Y^*$  not equal to zero simultaneously, such that the following relations hold true

$$\mathcal{L}_x = 0 \Leftrightarrow \sum_{i=0}^m \lambda_i f'_i(\hat{x}) + (F'(\hat{x}))^* \lambda = 0 \quad (\text{stationarity condition}),$$

$$\lambda_i \geq 0. \quad \forall i \in \{0, 1, \dots, m\} \quad (\text{nonnegativity conditions}),$$

$$\lambda_i f_i(\hat{x}) = 0 \quad \forall i \in \{1, \dots, m\} \quad (\text{condition of complementary slackness}).$$

**Proof:** This theorem is an immediate corollary of Theorem 2.1.1 a) if  $F'(\hat{x})X = Y$  (it is necessary to define  $\mathcal{F}(x, u) = (f_1(x) + u_1, \dots, f_m(x) + u_m, F(x))$ ,  $\mathcal{F} : X \times \mathbb{R}_+^m \rightarrow \mathbb{R}^m \times Y$  and to use of the lemma on closedness). If  $F'(\hat{x})X$  is a proper subspace in  $Y$ , then there exists a nontrivial  $\lambda \in Y^*$ ,  $\lambda \in (\text{Im} F'(\hat{x}))^\perp$ ; this leads to the Lagrange principle with  $\lambda_0 = \dots = \lambda_m = 0$ .  $\square$

**Remark 2** This result (for the finite dimensional case and for equality constraints) was formulated by Lagrange (1797, see the epigraph to the previous chapter). The infinite dimensional version of Lagrange's multiplier rule is due to Lyusternik (1934), Graves and others. Problems with inequalities in finite dimensional cases were considered by Karush (1939), John (1948), and others.

Let us now consider problems of the calculus of variations.

### 3.1.1 Necessary conditions in the calculus of variations

**Theorem 3.1.2** (Necessary conditions in unconstrained Bolza problems)  
Consider the Bolza problem

$$\mathcal{B}(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt + \ell(x(t_0), x(t_1)) \rightarrow \min, \quad (P_2)$$

where  $\hat{x}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ ,  $W \in \mathcal{O}(\{(t, \hat{x}(t), \dot{\hat{x}}(t)) | t \in [t_0, t_1]\}, \mathbb{R}^{2n+1})$ ,  $L \in C^1(W)$ ,  $V \in \mathcal{O}((\hat{x}(t_0), \hat{x}(t_1)), \mathbb{R}^{2n})$ ,  $\ell \in C^1(V)$ . If  $\hat{x}(\cdot)$  is a local minimum

of  $(P_2)$  in  $C^1([t_0, t_1], \mathbb{R}^n)$ , then the Euler equation

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0, \quad \forall t \in [t_0, t_1]$$

and the transversality conditions

$$\widehat{L}_{\dot{x}}(t_i) = (-1)^i \widehat{\ell}_i, \quad i = 0, 1 \quad (\widehat{\ell}_i := \ell_{x(t_i)}(\widehat{x}(t_0), \widehat{x}(t_1)))$$

hold true. ( $\widehat{L}_{\dot{x}}(t) = \widehat{L}_{\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$  and so on).

**Proof:** Let us prove the theorem in the case  $n = 1$  (the general case is analogous). From Fermat's theorem we obtain that

$$\mathcal{B}'(\widehat{x}(\cdot)) = 0 \Rightarrow \int_{t_0}^{t_1} (a(t)x(t) + b(t)\dot{x}(t))dt + \widehat{\ell}_0 x(t_0) + \widehat{\ell}_1 x(t_1) = 0 \quad \forall x(\cdot),$$

where  $a(t) = \widehat{L}_x(t)$ ,  $b(t) = \widehat{L}_{\dot{x}}(t)$ . From this equality we obtain by means of standard methods the result of the theorem (for details see ATF).  $\square$

**Theorem 3.1.3** (Necessary conditions in Lagrange problems)

Consider problem  $(P_3)$ :

$$\begin{aligned} \mathcal{B}_0(x(\cdot), u(\cdot)) &= \int_{t_0}^{t_1} f_0(t, x(t), u(t))dt + \psi_0(x(t_0), x(t_1)) \rightarrow \min; \\ \dot{x}(t) &= \varphi(t, x(t), u(t)) \\ \mathcal{B}_i(x(\cdot), u(\cdot)) &= \\ &= \int_{t_0}^{t_1} f_i(t, x(t), u(t))dt + \psi_i(x(t_0), x(t_1)) \begin{cases} \leq 0, & 1 \leq i \leq m' \\ = 0, & m' + 1 \leq i \leq m, \end{cases} \end{aligned}$$

where  $f_i : W \rightarrow \mathbb{R}$ ,  $\varphi : W \rightarrow \mathbb{R}^n$ ,  $\psi_i : V \rightarrow \mathbb{R}$ ,  $W$  is a neighbourhood of the graph  $\{(t, \widehat{x}(t), \widehat{u}(t))\} | t \in [t_0, t_1]\}$  in  $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r$  and  $V$  is a neighbourhood of the point  $(\widehat{x}(t_0), \widehat{x}(t_1))$  in  $\mathbb{R}^{2n}$ , where  $\widehat{x}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ ,  $\widehat{u}(\cdot) \in C([t_0, t_1], \mathbb{R}^r)$ . If  $(\widehat{x}(\cdot), \widehat{u}(\cdot))$  is a local minimum of  $(P_3)$  in  $C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$ , then the Lagrange principle holds true. This means that there

exists Lagrange multipliers  $(\lambda_i)_{i=0}^{m'}$  in  $\mathbb{R}$ ,  $p(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ , not all zero, such that  $\lambda_i \geq 0$ ,  $0 \leq i \leq m'$ ,  $\lambda_i \mathcal{B}_i(\hat{x}(\cdot), \hat{u}(\cdot)) = 0$ ,  $1 \leq i \leq m'$ ,

$$-\frac{d}{dt} \hat{L}_{\dot{x}}(t) + \hat{L}_x(t) = 0, \quad \forall t \in [t_0, t_1],$$

$$\hat{L}_u(t) = 0 \quad \forall t \in [t_0, t_1],$$

$$\hat{L}_{\dot{x}}(t_i) = (-1)^i \hat{\ell}_i, \quad i = 0, 1,$$

where

$$L(t, x, u, \dot{x}) = \langle p(t), \dot{x} - \varphi(t, x, u) \rangle + \sum_{i=0}^m \lambda_i f_i(t, x, u),$$

and

$$\ell(x(t_0), x(t_1)) = \sum_{i=1}^m \lambda_i \psi_i(x(t_0), x(t_1)).$$

**Proof:** This result is a corollary of Theorem 3.1.2. It is necessary to use the existence theorem for linear differential equations (from this theorem it follows that the mapping  $(x(\cdot), u(\cdot)) \rightarrow \dot{x}(t) - \varphi(t, x(t), u(t))$  is a totally regular mapping) and the lemma on closedness (from which it follows that the mapping  $F(x(\cdot), u(\cdot), v) = (\mathcal{B}_1 + v_1, \dots, \mathcal{B}_{m'} + v_{m'}, \mathcal{B}_{m'+1}, \dots, \mathcal{B}_m, \dot{x} - \varphi)$  is regular) and at last it is necessary to apply Theorem 3.1.2.  $\square$

**Remark 3** The problem  $(P_3)$  is called *a Lagrange problem in Pontryagin's form*. The Lagrange principle for such problems was deduced by Lagrange in his works devoted to classical mechanics. It was proved by Mayer, Hilbert, and many other mathematicians. From theorem 3.1.3 most classical results quoted in the majority of books on calculus of variations follow. For example, Euler equations for the simplest problem of calculus of variations (1744), isoperimetrical problems (1744), problems with higher derivatives (Euler-Poisson equation), and so on.

**Theorem 3.1.4** (Lagrange principle for Lyapunov's problems)  
Consider problem  $(P_4)$

$$\mathcal{J}_0(u(\cdot)) + \varphi_0(\xi) \rightarrow \min;$$

$$\mathcal{J}_i(u(\cdot)) + \varphi_i(\xi) \leq 0, \quad 1 \leq i \leq m'$$

$$\mathcal{J}_i(u(\cdot)) + \varphi_i(\xi) = 0, \quad m' + 1 \leq i \leq m, \quad \xi \in A, \quad u(t) \in U \quad \forall t \in \Delta,$$

where  $\mathcal{J}_i(u(\cdot)) = \int_{\Delta} f_i(t, u(t)) dt$ ,  $\Delta = [t_0, t_1] \subset \mathbb{R}$ ,  $f_i : \mathbb{R} \times \mathbb{R}^r \rightarrow \mathbb{R}^n$  are continuous functions,  $\varphi_i : \Xi \rightarrow \mathbb{R}$  convex functions for  $1 \leq i \leq m'$  and affine functions for  $m' + 1 \leq i \leq m$ , defined on the linear space  $\Xi$ ,  $A$  is a convex set and  $U$  is an arbitrary set in  $\mathbb{R}^r$ ,  $u(\cdot)$  is a measurable function on  $\Delta$  such that  $t \rightarrow f_i(t, u(t)) \in L_1(\Delta)$ ,  $u(t) \in U$  a.e.. If  $(\hat{u}(\cdot), \hat{\xi})$  is an absolute minimum in  $(P_4)$ , then there exist Lagrange multipliers  $\{\lambda_i\}_{i=0}^m$ , not all zero, such that

$$\begin{aligned} \lambda_i &\geq 0, \quad 0 \leq i \leq m', \\ \lambda_i(\mathcal{F}_i(\hat{u}(\cdot)) + \varphi_i(\hat{\xi})) &= 0, \quad 1 \leq i \leq m', \\ \min_{u \in U} \sum_{i=0}^m \lambda_i f_i(t, u) &= \sum_{i=0}^m \lambda_i f_i(t, \hat{u}(t)) \quad \text{a.e.}, \\ \min_{\xi \in A} \sum_{i=0}^m \lambda_i \varphi_i(\xi) &= \sum_{i=0}^m \lambda_i \varphi_i(\hat{\xi}). \end{aligned}$$

**Proof:** This result follows directly from Theorem 2.1.1 if we use Lyapunov's theorem (for details see [ATF]).  $\square$

**Remark 4** In the case  $\mathcal{J}_i = 0$  for all  $i$  we obtain the Kuhn-Tucker theorem (1951) for convex programming. Moreover the Lagrange principle for optimal control problems which are linear in the phase coordinate:

$$\mathcal{B}_0(x(\cdot), u(\cdot)) \Rightarrow \min,$$

$$\mathcal{B}_i(x(\cdot), u(\cdot)) \leq 0, \quad 1 \leq i \leq m',$$

$$\mathcal{B}_i(x(\cdot), u(\cdot)) = 0, \quad m' + 1 \leq i \leq m,$$

$$\dot{x} = A(t)x + F(t, u(t)), \quad u(t) \in U \quad \forall t \in \Delta,$$

where

$$\mathcal{B}_i(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} (\langle a_i(t), x(t) \rangle + f_i(t, u(t))) dt + \langle \gamma_{i0}, x(t_0) \rangle + \langle \gamma_{i1}, x(t_1) \rangle,$$

can be reduced to theorem 3.1.4 (see [ATF]).

**Theorem 3.1.5** (Lagrange principle for the problem of optimal control with distributed parameters)

Consider the problem

$$\begin{aligned} \mathcal{B}_0(x(\cdot), u(\cdot)) &\rightarrow \min, \\ \mathcal{B}_i(x(\cdot), u(\cdot)) &\leq 0 \quad (1 \leq i \leq m') = 0 \quad (m' + 1 \leq i \leq m), \\ x(t) &= \int_T K(t, \tau) \varphi(\tau, x(\tau), u(\tau)) d\tau, \quad u(t) \in U, \end{aligned} \quad (P_5)$$

$T$  is an open set in  $\mathbb{R}^n$  with smooth boundary,

$$\mathcal{B}_i(x(\cdot), u(\cdot)) = \int_T f_i(t, x(t), u(t)) dt, \quad 0 \leq i \leq m,$$

$f_i : W \rightarrow \mathbb{R}$ , functions  $\in C^1(W)$  in  $x$  and continuous in all arguments, where  $W$  is a neighbourhood of the graph  $\{t, \hat{x}(t), \hat{u}(t) | t \in T\}$ . The operator  $\mathcal{F}(x(\cdot), u(\cdot)) = x(t) - \int_T K(t, \tau) \varphi(\tau, x(\tau), \hat{u}(\tau)) d\tau$  is a regular operator. Then, if  $(\hat{x}(\cdot), \hat{u}(\cdot))$  is a local minimum in the space  $C^1(T, \mathbb{R}^n) \times C(T, \mathbb{R}^r)$  the Lagrange principle holds true.

**Proof:** We will not prove this result in its general form, but only in the simplest case of the problem of optimal control:

$$\begin{aligned} \int_{t_0}^{t_1} g(t, x(t), u(t)) dt &\rightarrow \min, \\ \dot{x}(t) &= \varphi(t, x(t), u(t)), \quad x(t_0) = x_0, \\ u(t) &\in U \end{aligned} \quad (P'_5)$$

In this case if  $(\hat{x}(\cdot), \hat{u}(\cdot))$  is a local minimum of  $(P'_5)$  there exists an absolutely continuous vector function  $p(\cdot) : [t_0, t_1] \rightarrow \mathbb{R}^n$  such that  $p(t_1) = 0$ ,  $-\dot{p}(t) = (\varphi_X(t, \hat{x}(t), \hat{u}(t)))^* p(t) - g_X(t, \hat{x}(t), \hat{u}(t))$  and  $\max_{u \in U} \langle p(t), \varphi(t, \hat{x}(t), u) \rangle = \langle p(t), \varphi(t, \hat{x}(t), \hat{u}(t)) \rangle$  a.e. (Pontrijagin's maximum principle).

Theorem 3.1.5 follows from a strengthened version of Theorem 2.1.1. To do this one has to verify weak approximative convexity - a weak convexity condition. This is done by means of the following construction of mixing of control functions (we explain it with a one-dimensional example, where  $T = [0, 1]$ ). Given  $0 < \alpha < 1$ ,  $n \in \mathbb{N}$ ,  $u_i(\cdot)$   $i = 1, 2$  measurable functions. We put

$$u_n(t, \alpha) = \begin{cases} u_1(t), & t \in \cup_{i=0}^{n-1} \Delta_{1i}, \\ u_2(t), & t \in \cup_{i=0}^{n-1} \Delta_{2i}, \end{cases}$$

where  $\Delta_{1i} = [\frac{i}{n}, \frac{i+\alpha}{n}]$ ,  $\Delta_{2i} = [\frac{i+\alpha}{n}, \frac{i+1}{n}]$ . □

## 3.2 Solutions of special problems

We begin from some problems which we discussed in the introduction.

### 3.2.1 Classical isoperimetrical problem

*Among all closed plane curves of a given length find one that encloses the largest area.*

A solution of this problem being formulated as a problem of optimal control see in ATF. Here we solve it on the base of convex analysis. It is easy to show that the solution must be considered only among convex figures.

Let  $A$  be a plane convex figure,  $l(A)$  the perimeter of  $A$ ,  $S(A)$  the area of  $A$ , (BM) Brunn-Minkowski's inequality, (S) Steiner's formula, (B) the unit ball in  $\mathbb{R}^2$ . Then

$$\pi t^2 + tl(A) + S(A) \stackrel{(S)}{=} S(A + tB) \stackrel{(BM)}{\geq} (S^{1/2}(A) + S^{1/2}(A))^2 \stackrel{Id}{=} 4S(A) + 2tS^{1/2}(A) + \pi t^2$$

$$S(A) + 2S^{1/2}(A)\pi^{1/2} + \pi t^2 \Rightarrow l^2(A)/4\pi \geq S(A),$$

and if  $A = B$  then  $l^2(B) = 4\pi S(A)$ . □

### 3.2.2 Generalized problem of Euclid

In the "Elements" of Euclid a solution of an extremal problem about an inscribed triangle is given. It can be formalized as follows:

$$f(x) = x(a - x) \rightarrow \max, \quad 0 \leq x \leq a.$$

Fermat illustrated his method (in a letter to Roberval) solving a problem with the same formalization (*find the right triangle of maximum area with a given sum of its sides*). The problem of minimization of a quadratic function played a great role in the history of mathematics. Consider the following general situation:

Let  $\mathcal{H}$  be a Hilbert space,  $A : \mathcal{H} \rightarrow \mathcal{H}$  a symmetric operator for which there exists  $\alpha > 0$  with  $\langle Ax, x \rangle \geq \alpha \langle x, x \rangle$  for all  $x \in \mathcal{H}$ . Consider the convex problem without constraints:

$$\frac{\langle Ax, x \rangle}{2} - \langle b, x \rangle \rightarrow \min .$$

From Fermat's theorem it follows that each solution of the problem is stationary; one readily sees that there is a unique stationary point  $\hat{x} = A^{-1}b$ . One can finish the analysis of the problem in two different ways. Either, one remarks that by the compactness principle a solution exists. Or one remarks that this is a convex problem and so each stationary point is a solution.

### 3.2.3 The problem of Apollonius

*Find the distance between a given point and a given ellipsoid in  $\mathbb{R}^n$ .*

Formalization:

$$f_0(x) = \sum_{i=1}^n (x_i - \xi_i)^2 \rightarrow \min, \quad f_1(x) = \sum_{i=1}^n (x_i/a_i)^2 \leq 1.$$

From the compactness principle it follows that a solution of the problem exists. This problem is a smooth (and also a convex) problem of mathematical programming. The Lagrange principle leads to the equations  $x_i - \xi_i + \lambda x_i/a_i^2 = 0 \Rightarrow x_i = \xi_i a_i^2 / (a_i^2 + \lambda)$  (i). Hence it is necessary to solve the equation

$$1 = \sum_{i=1}^n (x_i/a_i)^2 = \sum_{i=1}^n \xi_i^2 a_i^2 / (a_i^2 + \lambda)^2.$$

This gives, using (i), the - unique - solution of the problem.

### 3.2.4 The brachistochrone problem

Introduce an  $(x, y)$ -coordinate system in a vertical plane such that the  $x$ -axis is horizontal and the  $y$ -axis is directed downward. According to Galileo's law the velocity  $ds/dt$  of a particle is equal to  $\sqrt{2gy}$ . We can now determine

the time of reaching the endpoint if the particle slides along the graph of a function  $y(\cdot)$  from beginpoint  $(x_0, y_0)$  to endpoint  $(x_1, y_1)$ :

$$T(y(\cdot)) = \frac{1}{\sqrt{2g}} \int_{x_0}^{x_1} \frac{\sqrt{1+y'^2(x)}}{\sqrt{y(x)}} dx.$$

We obtain the following formalization of the problem:

$$T(y(\cdot)) \rightarrow \min; y(x_i) = y_i, \quad i = 1, 2.$$

It is the simplest problem of the calculus of variations. The Lagrange principle in this case consists of the Euler equation. The integrand does not contain the independent variable  $x$ , hence the Euler equation has the integral  $y(1+y'^2) = C$ . Solving this differential equation we obtain the family of cycloids:  $x = \frac{C}{2}(\tau - \sin \tau) + C_1$ ,  $y = \frac{C}{2}(1 - \cos \tau)$ ,  $0 \leq \tau \leq 2\pi$ . It is easy to show that if  $y_1 < y_2$  then there exists a unique cycloid with the given boundary conditions and that a family of cycloids covers the strip  $x_1 \leq x \leq x_2$ ,  $y \geq 0$ . It follows from this fact and from Weierstrass formula (see IT, p.313), that the cycloid with given boundary conditions is the solution of the problem.  $\square$

### 3.2.5 Newton's aerodynamical problem.

About the formalization of this problem and the history of its solution see ATF. One of the formalizations is the following:

$$\int_0^T \frac{tdt}{1+u^2} \rightarrow \min, \quad x(0) = 0, \quad x(T) = \xi, \quad u \geq 0$$

( $T$  and  $\xi$  are fixed).

This problem is a standard problem of optimal control. Applying the Lagrange principle we obtain:

$$\mathcal{L} = \int_0^T L dt + \mu_0 x(0) + \mu_1 (x(T) - \xi), \quad L = \frac{\lambda_0 t}{1+u^2} + p(\dot{x} - u).$$

The Euler equation gives  $p = \text{const} = p_0$ , the minimum principle consists of the inequality  $\frac{\lambda_0 t}{1+u^2} - p_0 u \geq \frac{\lambda_0 t}{1+\hat{u}^2(t)} - p_0 \hat{u}(t)$  ( $i$ ) ( $\hat{u}(\cdot)$  is an optimal control function). It is easy to show that  $\lambda_0 \neq 0$ . If  $\lambda_0 = 1$  then from ( $i$ ) it follows



that  $p_0 < 0$  and  $\hat{u}(t)$  is identically zero on the segment  $[0, -2p_0]$ . On the segment  $[-2p_0, T]$  we find  $\hat{u}(t)$  from the equation  $-p_0 = \frac{2ut}{(1+u^2)^2}$ . Hence we obtain the following family of curves which satisfy the necessary conditions:

$$\hat{x}(t, p_0) = 0, t \in [0, -2p_0], \hat{x}(t, p_0) = (-p_0/2)(\ln u + u^2 + 3/4u^4) + 7/8p_0,$$

$$t = (-p_0/2)(1/u + 2u + u^3), \quad t \geq -2p_0.$$

It is easy to show that there exists only one curve from the family which satisfies the given boundary conditions. This curve is the solution of the problem because in Lyapunov-type problems the necessary conditions are sufficient ones.  $\square$

### References

- ATF** Alexeev V., Tikhomirov V., Fomin S.: *Optimal Control*, Plenum, 1987.
- IT** Ioffe A.D., Tikhomirov V., *Theory of Extremal Problems*, North-Holland, 1979.
- T** Tikhomirov V.: *Fundamental Principles of the Theory of Extremal Problems*, Wiley, 1986.