

# TRANSFORMATIONS OF WELL-POSED VARIATIONAL PROBLEMS

MATHEMATICS

1966

SovietRxiv

---

View the original and related papers at <https://sovietrxiv.org/items/ru-196601.54553>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

**Abstract**

**Full Text**

UDC 519.31/33

**MATHEMATICS**

A. D. IOFFE

## TRANSFORMATIONS OF WELL-POSED VARIATIONAL PROBLEMS

*(Presented by Academician A. N. Kolmogorov on 27 VIII 1965)*

1. Let a differential equation be given

$$dy/dt = f(t, y, u) \tag{1}$$

and a functional

$$I(y) = \int_a^b F(t, y, u) dt.$$

Here  $t$  is a scalar;  $y$  is an  $n$ -dimensional vector;  $u$  is a point of the metric compact set  $U$  (the space of controls). The functions  $f(t, y, u)$  and  $F(t, y, u)$  are continuous on  $[a, b] \times R^n \times U$  and continuously differentiable with respect to  $y$ . By  $D$  we denote the set of solutions of (1) passing through the points  $(a, y_0)$  and  $(b, y_1)$  (solutions of (1) are defined as in <sup>(1)</sup>).

**Problem A.** It is required to find  $\inf_D I(y)$ . It is assumed that  $D \neq \emptyset$ .

We shall say that the control  $u(t)$ ,  $t \in [a, b]$ , belongs to the class  $\mathfrak{A}(\xi)$  if the corresponding solution of (1), issuing from  $(a, y_0)$ , passes at  $t = b$  through the point  $\xi \in R^n$ . Controls for which there is no solution of (1) defined on  $[a, b]$  will be assigned to the class  $\mathfrak{A}(\infty)$ . If controls are regarded as elements of the space  $S$  of measurable mappings of  $[a, b]$  into  $U$  (convergence in  $S$  is convergence in measure), then the Hausdorff distance between classes of controls turns the set of classes into a metric space. The mapping

$$\mathfrak{A}(\xi) \rightarrow \xi \tag{2}$$

of this space into  $R^n$  is one-to-one and continuous if  $\xi \neq \infty$ .

**Definition 1.** We shall say that problem A is posed **correctly** if the mapping inverse to (2) is continuous at the point  $\xi = y_1$ .

In what follows it is assumed that problem A is posed correctly.

Let  $f(t, y, U)$  be the image of  $U$  under the mapping  $u \rightarrow f(t, y, u)$ , and let  $Q(t, y) = \text{co } f(t, y, U)$ . Let

$$\tilde{F}(t, y, q) = \sup((q, z) + w)$$

on the set

$$\{z \in R^n, w \in R^1 : (z, f(t, y, u)) + w - F(t, y, u) \leq 0, \forall u \in U\}.$$

**Problem B.** Determine

$$\inf \int_a^b \tilde{F}(t, y, \dot{y}) dt$$

on the set

$$D_0 = \{y(t) \in C_{[a,b]}^n, y(a) = y_0, y(b) = y_1, \dot{y}(t) \in Q(t, y)\}.$$

**Theorem 1.** *If problem A is posed correctly, then problems A and B are equivalent. More precisely, the following assertions are valid:*

- a)  $D_0$  is the closure of  $D$  in  $C_{[a,b]}^n$ ;
- b)

$$\inf_D I(y) = \inf_{D_0} \int_a^b \tilde{F}(t, y, \dot{y}) dt;$$

- c) if a sequence  $y_m(t) \in D$  minimizes  $I(y)$  in  $D$ , then it also is minimizing for problem B. Conversely, for any sequence minimizing

$$\int_a^b \tilde{F}(t, y, \dot{y}) dt$$

in  $D_0$ , there exists a sequence confinal with it in  $C_{[a,b]}^n$  that minimizes  $I(y)$  in  $D$ .

- d) if  $D$  is compact in  $C_{[a,b]}^n$ , then  $\inf \int_a^b \tilde{F}(t, y, \dot{y}) dt$  is attained at some  $\bar{y}(t) \in D_0$ , and in  $D$  there exists a sequence  $y_m(t) \rightarrow \bar{y}(t)$  such that

$$\lim_{m \rightarrow \infty} I(y_m) = \int_a^b \tilde{F}(t, \bar{y}(t), \dot{\bar{y}}(t)) dt.$$

2. Let us consider two examples.

- 1) Let  $\psi(x)$  be defined on some set  $X \subset R^n$ . We shall say that  $\varphi(x)$  is the convex envelope for  $\psi(x)$  on  $X$ , if  $\varphi(x)$  is defined on  $\text{co} X$ , is convex,  $\varphi(x) \leq \psi(x)$  on  $X$ , and any function having these properties does not exceed  $\varphi(x)$ .

Let  $U \subset R^n$  and  $f(t, y, u) \equiv u$ ; then  $Q(t, y) = \text{co} U$  and  $\tilde{F}(t, y, u)$  is the convex envelope for  $F(t, y, u)$  on  $U$ . Indeed, let  $\varphi(t, y, q)$  be defined on  $Q$ , convex in  $q$ , and  $\varphi(t, y, u) \leq F(t, y, u)$  on  $U$ .  $(z, u) + w$  is a hyperplane defined by the conditions:  $(z, q) + w = \varphi(t, y, q)$ ,  $(z, u) + w \leq \varphi(t, y, u)$ ,  $u, q \in Q$ . Then

$$\varphi(t, y, q) \leq \sup_{z, w} \{(z, q) + w\} = \tilde{F}(t, y, q).$$

Let us note that a similar result can be obtained otherwise from (2):

- 2) Let  $n = 1$ ,  $U = [-1, 1]$ ,  $F(t, y, u) = M(t, y)u + P(t, y)$ ,  $M(t, y) > 0$ ,  $Q(t, y) = \text{co} f(t, y, U) = [\alpha(t, y), \beta(t, y)]$ ;  $\varphi(t, y, u)$  and  $\psi(t, y, u)$  are the convex and concave envelopes for  $f(t, y, u)$  on  $U$ . Suppose that for some  $t, y$ ,  $\varphi'_u(t, y, -1) < 0$ ,  $\psi'_u(t, y, -1) > 0$ .  $[-1, \gamma]$  and  $[-1, \delta]$  are intervals on which  $\varphi$  and  $\psi$  are monotone. Then

$$[\alpha(t, y), f(t, y, -1)] = \varphi(t, y, [-1, \gamma]); \quad [f(t, y, -1), \beta(t, y)] = \psi(t, y, [-1, \delta])$$

and

$$\tilde{F}(t, y, q) = \begin{cases} M(t, y)\psi^{-1}(t, y, q) + P(t, y), & f(t, y, -1) \leq q \leq \beta(t, y), \\ M(t, y)\varphi^{-1}(t, y, q) + P(t, y), & \alpha(t, y) \leq q \leq f(t, y, -1). \end{cases}$$

If  $\varphi'_u(t, y, -1) \geq 0$  for all  $t, y$ , then  $\tilde{F}(t, y, q) = M(t, y)\psi^{-1}(t, y, q) + P(t, y)$ , and problem  $A$  can be reduced to the form

$$\min \int_a^b F(t, y, u) dt, \quad y(a) = y_0, \quad y(b) = y_1, \quad \dot{y} = \psi(t, y, u).$$

3. The main idea of the proof of Theorem 1 is the following. We shall call  $y(t)$  a **generalized solution** of equation (1) if there exists a sequence of solutions of (1) converging uniformly to  $y(t)$ . The totality of generalized solutions passing through  $(a, y_0)$  and  $(b, y_1)$  will be denoted by  $B$ . Let  $y(t) \in B$ , and let  $y_m(t) \rightarrow$

$y(t)$  be a sequence of solutions of (1), and let  $A(y)$  be the set of such sequences. Define on  $B$  the functional

$$J(y) = \inf_{A(y)} \lim_{m \rightarrow \infty} I(y_m). \quad (3)$$

The well-posedness of problem  $A$  ensures the validity of assertions a), b), c) of Theorem 1 if  $D_0$  is replaced by  $B$ , and  $\int_a^b \tilde{F}(t, y, \dot{y}) dt$  by  $J(y)$ . It turns out that always

$$B = D_0, \quad J(y) = \int_a^b \tilde{F}(t, y, \dot{y}) dt. \quad (4)$$

$J(y)$  is lower semicontinuous on  $B$ . Therefore, if  $D$  is compact, then the lower bound (3) is attained, and from (4) the validity of d) follows.

4. The proof of the equalities (4) rests on certain facts of measure theory, which may be of independent interest.

Let  $U$  be a metric compactum,  $G$  a bounded domain in  $R^n$ ; let  $S_U$  and  $S_G$  be, respectively, the  $\sigma$ -algebras of Borel subsets of  $U$  and measurable subsets of  $G$ , and let  $V(U)$  be the space of measures defined on  $S_U$ .

**Definition 2.**  $\nu(x)$  is a **regular mapping**  $G \rightarrow V(U)$  if it is defined on a set of full measure in  $G$  and

$$\nu(x)(E) \geq 0$$

for all  $E \in S_U$  and almost all  $x \in G$ ,

$$\nu(x)(U) = 1$$

almost everywhere on  $G$ ,

$$\nu(x)(E)$$

is measurable as a function of  $x$  for any  $E \in S_U$ .

Let  $u(x)$  be a measurable mapping  $G \rightarrow U$ . Denote

$$\eta(E\Delta) = \mu\{x \in \Delta, u(x) \in E\},$$

where  $\mu$  is Lebesgue measure on  $G$ ,  $\Delta \in S_G$ .

**Theorem 2.** For any regular mapping  $\nu(x)$ ,  $G \rightarrow V(U)$ , there exists a sequence of measurable mappings  $u_m(x)$ ,  $G \rightarrow U$ , such that:

a)

$$\lim_{m \rightarrow \infty} \eta_m(E\Delta) = \int_{\Delta} \nu(x)(E) dx$$

for all  $E \in S_U$ ,  $\Delta \in S_G$ ;

b) if  $f(x, u)$  is measurable, continuous in  $u$ ,  $|f(x, u)| \leq \varphi(x) \in L_2(G)$ , then  $f(x, u_m(x))$  converge weakly in  $L_2(G)$  to

$$\int_U f(x, u) \nu(x)(du).$$

**Theorem 3.** From any sequence of measurable mappings  $G \rightarrow U$  one can choose a subsequence related by condition b) of the preceding theorem to some regular mapping  $G \rightarrow V(U)$ .

The proof of these theorems uses the following two lemmas.

**Lemma 1.** Let  $\lambda$  be a nonnegative finite measure on  $S_U$ . There exists a measurable mapping  $u(x)$ ,  $G \rightarrow U$ , such that for any  $E \in S_U$

$$\frac{\lambda E}{\lambda U} = \frac{\mu\{x \in G, u(x) \in E\}}{\mu G}.$$

**Lemma 2.** Let  $P$  be an arbitrary countable subset of  $S_U$ . From any sequence of measurable mappings  $G \rightarrow U$  one can choose a subsequence  $u_{m_k}(x)$  in such a way that everywhere on  $P \times S_G$  there exists the limit

$$\xi(E\Delta) = \lim_{k \rightarrow \infty} \eta_{m_k}(E\Delta).$$

The function  $\xi(E\Delta)$  is finitely additive for each  $\Delta$ , countably additive and absolutely continuous with respect to  $\mu$  for each  $E$ .

5. Theorems 2 and 3 make it possible, instead of sequences of solutions of (1), to consider regular mappings  $[a, b] \rightarrow V(U)$ .

**Theorem 4.** In order that  $y(t)$  ( $y(a) = y_0$ ,  $y(b) = y_1$ ) be contained in  $B$ , it is necessary and sufficient that there exist a regular mapping  $\nu(t)$ ,  $[a, b] \rightarrow V(U)$ , for which  $y(t)$  is a solution of the equation

$$\frac{dy}{dt} = \int_U f(t, y, u) \nu(t)(du). \quad (5)$$

If  $y(t) \in B$ , then

$$J(y) = \min \int_a^b \int_U F(t, y(t), u) \nu(t)(du) dt. \quad (6)$$

The minimum is taken over all  $\nu(t)$  satisfying (5) for  $y = y(t)$ .

If there exists at least one  $\nu(t)$  satisfying, for  $y = y(t)$ , (5), (6) and concentrated on a set of full measure at a single point, then  $y(t) \in D$  and  $I(y) = J(y)$ . In the opposite case a sliding regime is realized along  $y(t)$ . A sufficient condition for the absence of a sliding regime is contained in the following

**Corollary.** If the image of  $U$  in  $R^{n+1}$  under the mapping

$$u \rightarrow \begin{cases} f(t, y, u), \\ F(t, y, u) \end{cases}$$

is convex for all  $t, y$ , then  $D = B$ , and any minimum for  $J(y)$  in  $B$  is a minimum for  $I(y)$  in  $D$ .

For  $U \subset R^k$  this result was obtained by A. F. Filippov (3).

Denote by

$$\Phi(t, y, q) = \inf \int_U F(t, y, u) \lambda(du)$$

over the set of nonnegative Borel measures  $\lambda$  for which the equalities

$$\int_U f(t, y, u) \lambda(du) = q, \quad \lambda(U) = 1. \quad (7)$$

hold.

If this set is nonempty, then the infimum is attained.

Let  $y(t) \in B$  be fixed, and let  $\nu(t)$  satisfy (5) together with  $y(t)$ .

**Theorem 5.** In order that

$$\int_a^b \int_U F(t, y, u) \nu(t)(du) dt = J(y),$$

it is necessary and sufficient that, for almost all  $t$ ,

$$\int_U F(t, y(t), u) \nu(t)(du) = \Phi(t, y(t), \dot{y}(t)).$$

The determination of  $\Phi(t, y, q)$  under conditions (7) is, as is easy to see, a linear programming problem over a space of measures.

The dual problem is: find

$$\tilde{F}(t, y, q) = \sup((q, z) + w) \quad (8)$$

subject to the condition

$$(z, f(t, y, u)) + w - F(t, y, u) \leq 0$$

for all  $u \in U$ .

Let

$$Q(t, y) = \{q \in R^n, \tilde{F}(t, y, q) < \infty\}$$

be the set of those  $q$  for which problem (8) is solvable. If  $q \in Q(t, y)$ , then  $\tilde{F}(t, y, q) = \Phi(t, y, q)$ . It is not hard to see that  $Q(t, y)$  is closed and that  $\tilde{F}(t, y, q)$  is continuous in its domain of definition. Moreover, since

$$((\alpha q_1 + (1-\alpha)q_2), z) + w = \alpha((q_1, z) + w) + ((q_2, z) + w)(1-\alpha) \leq \alpha \tilde{F}(t, y, q_1) + (1-\alpha) \tilde{F}(t, y, q_2),$$

$Q(t, y)$  is convex and  $\tilde{F}(t, y, q)$  is convex in  $q$ .

Obviously,  $f(t, y, U) \subset Q(t, y)$ , and therefore also  $\text{co } f(t, y, U) \subset Q(t, y)$ . The reverse inclusion is easily obtained with the aid of the Hahn–Banach theorem, which completes the proof of Theorem 1.

6. If problem A is posed incorrectly, then Theorem 1 is false. Let us consider an example. Let

$$I(y) = \int_0^1 (y^1)^2 dt, \quad \frac{dy^1}{dt} = u, \quad \frac{dy^2}{dt} = (y^1)^2 - u^2 \left(1 + \frac{(y^1)^4}{\varepsilon^2}\right), \quad (9)$$

$$U = [-1, 1], \quad y(0) = (0, 0), \quad y(1) = (0, -1).$$

It is clear that if  $y(t) \in D$ , then

$$\|y^1\|_C \geq \varepsilon. \quad (10)$$

On the other hand, there exists a solution of (9) passing through the points  $(0, 0, 0)$  and  $(1, 0, -1 + \delta)$  and such that  $\|y^1\|_C$  is arbitrarily small. The problem is posed incorrectly. By virtue of (10),

$$\inf_D I(y) \geq \varepsilon^3.$$

Further,  $y(t) = (0, -t) \in D_0$ , and therefore

$$\inf_D I(y) = 0.$$

In particular, V. F. Krotov's optimality principle (<sup>4</sup>) is inapplicable to incorrectly posed problems.

I take this opportunity to express my deep gratitude to V. M. Tikhomirov for numerous discussions and helpful advice.

Received  
26 VII 1965

## CITED LITERATURE

- <sup>1</sup> L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *Mathematical Theory of Optimal Processes*, Moscow, 1961.
- <sup>2</sup> R. V. Gamkrelidze, DAN, 143, No. 6, 1243 (1962).
- <sup>3</sup> A. F. Filippov, *Vestn. MGU*, No. 2, 25 (1959).
- <sup>4</sup> V. F. Krotov, Dissertation, Moscow, 1963.

*Note: Figure translations are in progress. See original paper for figures.*

*Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.*