ALEXEI A. MILYUTIN

Calculus of variations and optimal control [1]

1 Statement of the problem

We consider the following optimal control problem:

$$\text{to minimize} \quad J(x_0, t_0;\, x_1, t_1) \tag{1}$$

under the constraints

$$F(x_0, t_0;\, x_1, t_1) \le 0, \qquad K(x_0, t_0;\, x_1, t_1) = 0\,, \tag{2}$$

$$\dot{x} = f(x, u, t), \qquad (x, u, t) \in Q\,, \tag{3}$$

where $t \in [t_0, t_1]$, $x_0 = x(t_0)$, $x_1 = x(t_1)$, Q is an open set, J, F, K and f are continuously differentiable.

It is necessary to find the interval $[t_0, t_1]$, a bounded measurable function $u(t)$ and an absolutely continuous function $x(t)$ on this interval, satisfying the constraints (2) and (3) and giving minimum to the functional (1).

This statement corresponds to Dubovitskii-Milyutin canonical form of optimal control problems. Therefore we call this problem also *canonical*, or briefly CP. We refer to (3) as the *control system*, and to the finite-dimensional problem (1) and (2) as *endpoint problem* (EP). Note that all problems of the classical calculus of variations, for example, those in the book of A.Bliss [1], can be rewritten in the CP form.

We investigate CP as an optimal control problem, but we also use the ideas and methods of the calculus of variations. The paper is devoted to the fragment of the theory of a strong minimum for the canonical problem CP and related directly to the calculus of variations.

2 The canonical Hamiltonian

The *Pontryagin function* for the canonical problem has the form

$$H(\psi, x, u, t) = \langle \psi, f(x, u, t) \rangle\,,$$

where $\langle \psi, f \rangle$ is a scalar product. Next, from the optimal control viewpoint the following function arises naturally:

$$\mathcal{H}(\psi, x, t) = \max_u H(\psi, x, u, t)\,,$$

[1] This work was supported by Russian Foundation for Basic Research, grants 97-01-00135 and 96-15-96072

where maximum is taken over the set $\{u|\,(x,u,t)\in Q\}$.

Denote by $\mathrm{dom}\,\mathcal{H}$ a domain of $\mathcal{H}$, that is the set of all the triples (ψ,x,t) such that the maximum of H in this formula attains. We do not assume that $\mathrm{dom}\,\mathcal{H}$ is an open set. We call $\mathcal{H}$ a *canonical Hamiltonian.* It is closely connected with extremum, since it is part of necessary extremum conditions. This Hamiltonian is not present in the books on the calculus of variations, seemingly, because the calculus of variations problems originaly were not formulated in the canonical form. As we shall see, missing of this form influenced the theory of a strong minimum, and perhaps the whole calculus of variations, which was not revised practically since its foundation.

3 The canonical Hamilton-Jacobi equation and upper approximation of the attainable set

The Hamilton-Jacobi equation has the form:

$$S_t+\mathcal{H}(S_x,x,t)=0\,,\qquad (S_x,x,t)\in\mathrm{dom}\,\mathcal{H}\,.$$

We call it a *canonical Hamilton-Jacobi equation*, since we use the canonical Hamiltonian for its definition, and that is a distinction of this equation from that in the calculus of variations. Thus, we use the term "canonical" for concepts and notions which are different from those in the classical calculus of variations.

Let $S(x,t)$ be a continuously differentiable solution of the canonical equation, defined on an open set G . The smoothness assumptions for the solutions could be weakened, but we should not give the exact assumptions, because our aim is to present only the general idea of application of the theory.

Let $(x(t),u(t)\,|\,t\in[t_0,t_1])$ be a trajectory of the control system, contained in G , that is

$$(x(t),t)\in G\quad\text{for all } t\in[t_0,t_1]\,.$$

Then it is easily follows from definitions that the derivative of the function $S(x(t),t)$ with respect to t is nonpositive on the interval $[t_0,t_1]$:

$$\frac{dS}{dt}\le 0\,.$$

Hence

$$S(x(t_1),t_1)-S(x(t_0),t_0)\le 0\,.$$

Denote by $A(G)$ the *attainable set* of the control system in G , defined as the set of quadruples $(x_0,t_0;\,x_1,t_1)$ such that there exists a trajectory $(x(t),u(t)\,|\,t\in[t_0,t_1])$ of the control system, joining in G the points (x_0,t_0) and (x_1,t_1) , that is

$$x(t_0)=x_0,\qquad x(t_1)=x_1\qquad (x(t),t)\in G\quad \forall t\in[t_0,t_1]\,.$$

Then we have: the condition

$$(x_0,t_0;\,x_1,t_1)\in A(G)$$

implies the inequality

$$S(x_1,t_1) - S(x_0,t_0) \le 0\,. \tag{4}$$

Consequently, the set of quadruples $(x_0,t_0;\,x_1,t_1)$, satisfying inequality (4), contains the attainable set $A(G)$ of the control system, and thus every solution of Hamilton-Jacobi canonical equation provides for an *upper approximation*, or *upper estimation* of the attainable set of the control system. This upper estimation is given by the inequality (4). It was missing in the calculus of variations for at least two reasons: in the first place, because it presupposes a representation of the problem in the canonical form, and secondly, in the canonical form it holds only for the canonical Hamiltonian. The estimation leads at once to the sufficient conditions of a strong minimum.

4 Canonical theory of a strong minimum

Consider the canonical problem with additional constraints

$$(x(t),t) \in G \quad \forall t \in [t_0,t_1]\,, \qquad (x_0,t_0;\,x_1,t_1) \in \mathcal{G}\,,$$

where G and $\mathcal{G}$ are open sets such that $\mathcal{G} \subset G \times G$. We denote this new problem by $CP(G,\mathcal{G})$.

Next, assume that we have some family

$$\{S_\mu(x,t) \,|\, \mu \in M\}$$

of the solutions of the canonical Hamilton-Jacobi equation, defined on G . Consider the following endpoint problem:

$$\text{to minimize} \quad J(x_0,t_0;\,x_1,t_1)$$

under the constraints

$$F(x_0,t_0;\,x_1,t_1) \le 0, \qquad K(x_0,t_0;\,x_1,t_1) = 0\,,$$
$$S_\mu(x_1,t_1) - S_\mu(x_0,t_0) \le 0 \qquad \forall \mu \in M\,,$$
$$(x_0,t_0;\,x_1,t_1) \in \mathcal{G}\,.$$

Denote it by $EP(M,\mathcal{G})$. Then the admissible set of this problem contains the attainable set of the problem $CP(G,\mathcal{G})$. Hence we obtain two following assertions:

(i) $$\inf J[CP(G,\mathcal{G})] \ge \inf J[EP(M,\mathcal{G})]\,,$$

where $\inf J[\cdot]$ *is the infimum of the functional in the problem given in the square brackets.*

(ii) If $(\hat{x}_0,\hat{t}_0;\,\hat{x}_1,\hat{t}_1) \in A(G)$ *is a solution of the problem* $EP(M,\mathcal{G})$ *, then any trajectory* $(x(t),u(t)\,|\,t \in [t_0,t_1])$ *joining in* G *the points* $(\hat{x}_0,\hat{t}_0)$ *and* $(\hat{x}_1,\hat{t}_1)$ *is a solution of the problem* $CP(G,\mathcal{G})$ *, and consequently, a point of a strong minimum in the canonical problem* CP *.*

These conditions are new both for the calculus of variations and the optimal control. They are the basis of the theory which we shall call canonical.

Next, we will give a nontrivial example.

5 Example 1

A control system has the form

$$\dot{x} = u\,, \quad \dot{y} = \langle R(x), u\rangle\,, \quad \dot{z} = \varphi(x,u)\,, \quad x \in \Omega\,, \quad u \in R^n\,, \quad y,z \in R^1\,,$$

where $\Omega \subset R^n$ is an open set. Thus, we have

$$(x,y,z,u,t) \in Q = \Omega \times R^1 \times R^1 \times R^n \times R^1\,.$$

Assume that $R(x)$ is continuous on Ω and φ is continuous on $\Omega \times R^n$. Moreover, we assume that $\varphi(x,u) > 0$ if $u \neq 0$ and $\varphi(x,u)$ is convex and positive homogenious function in u.

Since we assume only the continuity in x we cannot use the maximum principle and hence cannot define the extremals in this control system. At the same time the canonical theory holds, and we shall use it for a given system.

The Pontryagin function here has the form:

$$H = \langle \psi_x + \psi_y R(x), u\rangle + \psi_z \varphi(x,u)\,.$$

Set

$$\theta(x,v) = \max_u \langle v, u\rangle\,,$$

where maximum is taken over the set $\{u \,|\, \varphi(x,u) \leq 1\}$. Then it is easy to prove that

$$(\psi_x, \psi_y, \psi_z, x, y, z, t) \in \operatorname{dom} \mathcal{H}$$

if and only if

$$\theta(x, \psi_x + \psi_y R(x)) \leq -\psi_z\,.$$

Moreover,

$$\mathcal{H} = 0 \qquad \text{on} \quad \operatorname{dom} \mathcal{H}\,.$$

Hence $S_t = 0$ for each solution S of the canonical Hamilton-Jacobi equation.

We look for a solution of the form

$$S(x,y,z) = f(x) - y - \beta z\,,$$

where f is continuously differentiable on Ω and $\beta \geq 0$. Then $S(x,y,z)$ is a solution of the canonical Hamilton-Jacobi equation on $\Omega \times R^1 \times R^1$ if and only if

$$\theta(x, f'(x) - R(x)) \leq \beta \qquad \text{on} \quad \Omega\,.$$

Put

$$\beta(f) = \sup_\Omega \theta(x, f'(x) - R(x))\,.$$

Then

$$S_f(x,y,z) = f(x) - y - \beta(f) z$$

is a solution on $\Omega \times R^1 \times R^1$. Thus, we have constructed a family of solutions $\{S_f\}$ of canonical Hamilton-Jacobi equation for the control system.

Now we set up an optimal control problem for the given control system.

Problem A : To minimize the endpoint functional

$$J = \frac{y_1 - y_0}{z_1 - z_0}$$

on the set of all trajectories of the control system, satisfying the condition

$$x_0 = x_1 \, .$$

In optimal control theory it is important to have a complete system of lower bounds for the infimum of J in the Problem A . Hence this example is not a simple illustration of application of the canonical theory. We shall find this system of lower bounds, using the family $\{S_f\}$.

Now let us formulate an endpoint

Problem B : To minimize

$$J = \frac{y_1 - y_0}{z_1 - z_0}$$

under the constraints

$$x_0 = x_1 \, ,$$

$$S_f(x_1, y_1, z_1) - S_f(x_0, y_0, z_0) \leq 0 \quad \forall f \, .$$

For this problem the infimum of the functional can be calculated directly. It has the form:

$$\inf J[B] = - \inf_f \beta(f) \, .$$

According to the canonical theory of a strong minimum we have

$$\inf J[A] \geq \inf J[B] \, .$$

Hence

$$\inf J[A] \geq -\beta(f) \quad \forall f \, .$$

It can be proved that this system of lower bounds is complete, i.e. the following equality holds:

$$\inf J[A] = \inf J[B] \, .$$

6 Field theory

What is the relationship between the canonical theory of a strong minimum and the classical theory of a strong minimum in the calculus of variations? It is well known that the source of the sufficient conditions of a strong minimum in the calculus of variations is a *field of extremals* satisfying the Weierstrass condition. Each field of extremals generates a *field function.* If the field consists of extremals satisfying the Weierstrass condition, then the field function can be added to the solution of the canonical Hamilton–Jacobi equation. For extremal embedded in the field, the correspoding point in the endpoint problem $EP(M, \mathcal{G})$ with a singleton M turned out to be a stationary one. In the case of the simplest problem of the calculus of variations and in some similar problems this guarantees a strong minimum, and hence makes the role of the field somewhat more significant. But in the general case it is not so. Thus, the sufficient conditions of a strong minimum in the calculus of variations could be percieved as a very special case of the canonical conditions of a strong minimum, given above. The canonical theory is much broader than the classical one. We have just seen it in Example 1, where there were no extremals at all, and hence there were no field, but we obtained the solutions of the canonical Hamilton–Jacobi equation and thus solved the problem.

We can conclude now, that in the calculus of variations the emphasis was not on a proper point. A field of extremals is not a central notion of the theory of a strong minimum at all, but simply one of the methods to find a solution of the canonical Hamilton–Jacobi equation.

The theorem on the line integral preservation for some families of extremals is one of the most important in the classical field theory. This theorem was proved, in the calculus of variations, for a family of extremals which are not presupposed to satisfy the Weierstrass condition. Fairly strong assumptions of smoothness with respect to time variable and the parameter were made for this family.

The extremals satisfying the Weierstrass condition is a different matter. For a family of such extremals the line integral pserervation theorem could be proved under much weaker assumptions. In particular, we can avoid the assumptions of smoothness.

This allows us to deal with a much broader set of the fields and consequently with much broader set of the solutions of the canonical Hamilton–Jacobi equation which could be used in the problem.

In fact, we have two types of the fields in the calculus of variations: satisfying and not satisfying the Weierstrass condition, and each type has its own field theory. This fact was not mentioned in the classical theory, and it is very important for extention of the theory to optimal control.

Let us consider one more interesting example in which the field is formed by nonsmooth extremals.

7 An example of linear control system

First we shall give a definition of extremal of the canonical control system (3), satisfying the Weierstrass condition, or the maximum principle. It is a triple

$$(\psi(t), x(t), u(t) \,|\, t \in [t_0, t_1])\,,$$

where $(x(t), u(t))$ is a trajectory of a control system, $\psi(t)$ is an adjoint variable, satisfying the adjoint equation

$$-\dot{\psi} = H_x(\psi(t), x(t), u(t), t),$$

such that the maximality conditions

$$(\psi(t), x(t), t) \in \operatorname{dom} \mathcal{H} \quad \text{on} \quad [t_0, t_1]$$

$$\mathcal{H}(\psi(t), x(t), t) = H(\psi(t), x(t), u(t), t) \quad \text{a.e. on} \quad [t_0, t_1]$$

are fulfilled. We call these extremals *Pontryagin extremals*, because they defined just as it was done in the optimal control theory. The field function for the Pontryagin extremals is a solution of the canonical Hamilton-Jacobi equation.

The example is as follows. The control system has the form

$$\dot{x} = A(t)x + B(t)u\,, \qquad u \in U\,,$$

where $A(t)$ and $B(t)$ are matrices, U is a compact set. For this system we have

$$H = \langle \psi, A(t)x \rangle + \langle \psi, B(t)u \rangle\,,$$

$$\mathcal{H} = \langle \psi, A(t)x \rangle + \varphi(\psi B(t))\,,$$

where

$$\varphi(v) = \max_{u \in U} \langle v, u \rangle\,.$$

The adjoint equation has the form

$$-\dot{\psi} = \psi A(t)\,.$$

We shall construct a family of fields of Pontryagin extremals. Let $\psi(t)$ be an arbitrary solution of the adjoint equation. Let us find $u(t)$ from the conditions

$$u(t) \in U\,, \qquad \langle \psi(t), B(t)u(t) \rangle = \varphi(\psi(t)B(t))\,.$$

Let $x(t, \xi)$ be a solution of the Cauchy problem

$$\dot{x} = A(t)x + B(t)u(t)\,, \qquad x(0) = \xi\,.$$

Then for each ξ the triple $(\psi(t), x(t, \xi), u(t))$ is a Pontryagin extremal of the control system, and the family

$$(\psi(t), x(t, \xi), u(t))$$

is a field of Pontryagin extremals in all space of the variables x, t. The field function has the form

$$S^{\psi} = \langle \psi(t), x \rangle + f(t)\,.$$

From the canonical Hamilton-Jacobi equation it follows that

$$\dot{f} = -\varphi(\psi(t)B(t))\,.$$

Hence

$$S^{\psi} = \langle \psi(t), x \rangle - \int_0^t \varphi(\psi(t')B(t'))\, dt'\,.$$

One can prove that this family of the fields is complete in the following sense:

$$(x_0, t_0;\, x_1, t_1) \in A(R^n \times R^1)$$

if and only if

$$S^{\psi}(x_1, t_1) - S^{\psi}(x_0, t_0) \leq 0 \qquad \forall \psi\,.$$

Here n is the dimension of x, and $A(R^n \times R^1)$ is the attainable set of the control system.

8 Conclusion

It is usual to praise the calculus of variations, especially now that we selebrate the 300 years anniversary of its foundation. But, in our opinion, we need not shut our eyes to its shotcomings, then we can better appreciate its undeniable achievements. We just have seen that the theory of a strong minimum in the classical calculus of variations is not quite adequate. We mean that the sufficient conditions, based on the solutions of the canonical Hamilton-Jacobi equation were missing as well as the canonical Hamiltonian. Next, the field theory is not complete, because the theory of the field of extremals satisfying the Weierstrass condition was not derived. It caused some insufficiency of the results. No doubt, it was a price paid for a noncanonical form of representation of the problems in the calculus of variations. But also could be that the great creators of the calculus of variatios simply did not appreciate enough the significance of extremals, satisfying the Weierstrass condition, and the authors of the books on the calculus of variatios, even those who were familiar with the optimal control, followed them directly. At the same time the important notions, which could help to derive successfully the theory of such extremals, has long been used in the calculus of variations. Even though it created a powerful tool, the calculus of variations did not use all its potentials.

Together with N. P. Osmolovskii we tried to overcome this gap. Using the ideas both of the optimal control and of the calculus of variations we have derived a theory of Pontryagin's extremals. This theory is contained in our book "Calculus of Variations and Optimal Control" which is being published now by the American Mathematical Society.

9 On a conjecture of Clarke

F.H. Clarke proposed a conjecture, that in the case of Lipschitz continuous compact valued inclusion, defined in an open set, the solution of any optimal control problem can be complemented with an adjoint component up to a solution to a Hamiltonian system. This conjecture can be considered as an attempt to extend the area of application of the classical Pontryagin Maximum Principle (MP). Probably, such an extension is the most general, if one would like to preserve a relationship with the classical Pontryagin MP for the compact admissible control set.

Both the above requirements - the Lipschitz continuity and the definiteness in an open set - turn out to be essential. If one allows for a Lipschitz inclusion to be defined not in an open set, but, say, in a closed set, then the above conjecture fails to be true, which is demonstrated by optimal control problems with state constraints.

On the other hand, if one gives up the Lipschitz continuity of the inclusion, relaxing it just to Hölder continuity, then the conjecture fails to be true even in case when the inclusion is defined in an open set. This last assertion is not so evident as the first one, and has to be proved. In what follows we give an example of an inclusion with a single "Hölder point", for which the Clarke's conjecture is not true, and moreover, the adjoint variable $\psi(t)$ is discontinuous (which does not follow from the Hölder continuity of the Hamiltonian).

1. A Hölder continuous inclusion

Let us define the function $\varepsilon : \ R \to R, \quad \varepsilon(v) = \frac{1}{6}[(v-1)^+]^2$. On the plane $x = (x_1, x_2)$ we set for convenience $z = 1 - x_1, \ \ \zeta = 1 - x_2,$ and define three open sets:

$$Q_1 = \{x \,|\, z > 0, \ \sqrt{z} > \zeta \,\}, \qquad Q_2 = \{x \,|\, \zeta > 0, \ \zeta^2 > z \,\}, \qquad Q_3 = \{x \,|\, z < 0, \ \zeta < 0 \,\}.$$

Let us consider the control System A:

$$\dot{x} = -e_2 + u, \quad e_2 = (0,1),$$

where u is bounded in each set Q_i as follows:

$$|u| \le \sqrt{z + \varepsilon^2(x_2 - 1)} - \varepsilon(x_2 - 1), \quad \text{i.e.} \quad 2\varepsilon|u| + u^2 + x_1 \le 1, \quad \text{if} \quad x \in Q_1,$$

$$|u| \le \zeta, \quad \text{if } x \in Q_2, \qquad\qquad u = 0, \quad \text{if } x \in Q_3.$$

Thus, in fact we have an inclusion for $\dot{x}$. On the boundaries of Q_1, $Q_2\, Q_3$ we define the inclusion by continuity. Denote the obtained inclusion by inclusion A. Obviously, inclusion A has compact and convex values, and is Hölder continuous ; moreover, it is Lipschitz continuous everywhere except the point $x_* = (1,1)$.

For this inclusion A we consider the following Problem A:

$$J = x_1(T) \to \max, \qquad \text{subject to constraints:}$$

$$z(0) = z_*, \quad \zeta(T) = \zeta_* \quad \text{— fixed values, such that} \quad z_* > 0, \quad 0 < \zeta_* < 1.$$

The time interval [0,T] is not fixed.

2. System B.

To solve Problem A we first consider the following auxiliary System B:

$$\dot{x} = -e_2 + u, \quad |u| + x_2 \le 1.$$

This system coincides with System A in the set Q_2. We have $\dot{x}_2 = -1 + u_2$, whence $\dot{\zeta} = 1 - u_2$, so $\dot{\zeta} = 1 - \mu\zeta$, where $|\mu| \le 1$ along any trajectory of System B. Observe that $\dot{\zeta} > 0$ if $\zeta < 1$, and that any trajectory containing a point with $\zeta < 1$ must begin from the boundary $\zeta = 0$.

For System B we consider the following Problem B:

$$J = x_1(T) - x_1(0) \to \max, \qquad \zeta(0) = 0, \quad \zeta(T) = \zeta_*.$$

Since here $|u|$ is bounded, the problem has a solution, and due to the above observation, it is an extremal starting from the boundary $\zeta = 0$. The Pontryagin function here is $H = -\psi_2 + \psi u$, and the maximality condition gives:

$$\psi u = |\psi| \cdot |u| = |\psi| \cdot \zeta, \qquad \psi = \lambda \frac{u}{|u|}, \quad \text{whence } \lambda = |\psi|.$$

The adjoint equation is $\dot{\psi} = \lambda e_2$ i.e. $\dot{\psi} = |\psi| e_2$. From here one can easily get, that

$$|\psi| = a\cosh(t+\tau),\ a > 0, \quad \psi_2 = a\sinh(t+\tau), \quad \psi_1 = \theta a, \quad \text{where } \theta = +-1.$$

Let us find the state component of the extremal. We have

$$\dot{\zeta} = 1 - u_2 = 1 - \frac{\psi_2}{|\psi|}\zeta = 1 - \tanh(t+\tau)\zeta,$$

whence ζ, satisfying initial condition $\zeta(0) = 0$, is expressed as

$$\zeta(t,\tau) = \tanh(t+\tau) - \frac{\sinh\tau}{\cosh(t+\tau)}. \tag{2.1}$$

Then we have $\dot{x} = \theta \cdot \zeta(t,\tau)/\cosh(t+\tau)$. Since $\zeta(t,\tau) > 0$ if $t > 0$, and we seek for the maximum of $x_1(T) - x_1(0)$, we must take $\theta = 1$. Therefore,

$$x_1(t) - x_1(0) = \left. -\frac{1}{\cosh(t'+\tau)}\right|_0^t - \left.\sinh\tau \cdot \tanh(t'+\tau)\right|_0^t. \tag{2.2}$$

The above formulas, defining the extremal, are borrowed from the theory of optimal control problems with general mixed state-control constraints, developed by A.Ja.Dubovitskii and A.A.Milyutin.

Now let us find the extremal solving Problem B. Denote by $\hat{t}(\tau)$ the root of equation $\zeta(\hat{t}(\tau),\tau) = \zeta_*$. Obviously, $\hat{t}(\tau)$ exists and unique for any τ, and it is positive. Putting $p = \hat{t}(\tau) + \tau$, we get from (2.1):

$$\zeta_* \cosh p - \sinh p = -\sinh\tau. \tag{2.3}$$

Denote $\xi(\tau) = x_1(\hat{t}(\tau)) - x_1(0)$. From (2.1) and (2.2) we get

$$\xi(\tau) = -\frac{1}{\cosh p} - \sinh\tau \cdot \tanh p + \frac{1}{\cosh\tau} + \frac{\sinh^2\tau}{\cosh\tau} = -\frac{1}{\cosh p} - \sinh\tau \cdot \tanh p + \cosh\tau.$$

Plugging here $-\sinh\tau$ from (2.3) yields $\xi(\tau) = -\cosh p + \zeta_* \sinh p + \cosh\tau$, and so

$$-\cosh p + \zeta_* \sinh p = \xi(\tau) - \cosh(\tau). \tag{2.4}$$

Considering (2.3), (2.4) as a linear system with respect to $\cosh p$ and $\sinh p$, we get $(\zeta_*^2 - 1)\cosh p = \alpha + \zeta_*\beta$, $\quad (\zeta_*^2 - 1)\sinh p = \zeta_*\alpha + \beta$, where

$$\alpha = \xi(\tau) - \cosh\tau, \qquad \beta = -\sinh\tau. \tag{2.5}$$

Since $\cos p > 0$, (2.5) implies that

$$\alpha + \zeta_* \beta < 0. \tag{2.6}$$

Besides, using the identity $\cosh^2 p - \sinh^2 p = 1$, we get from (2.5):

$$(\zeta_*^2 - 1)^2 = (1 - \zeta_*^2)(\alpha^2 - \beta^2), \quad \text{whence} \quad \alpha^2 - \beta^2 = 1 - \zeta_*^2.$$

Plugging here the expressions for α and β from (2.5) yields

$$\xi^2 - 2\xi \cdot \cosh \tau + \zeta_*^2 = 0, \quad \text{hence} \quad \xi = \cosh \tau + - \sqrt{\cosh^2 \tau - \zeta_*^2}.$$

But due to (2.6) $\xi < \cosh \tau + \zeta_* \cdot \sinh \tau \le \cosh \tau + \zeta_* |\sinh \tau|$. From here it follows easily that

$$\xi(\tau) = \cosh \tau - \sqrt{\cosh^2 \tau - \zeta_*^2} = \frac{\zeta_*^2}{\cosh \tau + \sqrt{\cosh^2 \tau - \zeta_*^2}}.$$

One can obviously observe that here the maximum ξ is attained at $\tau = 0$, i.e.

$$\xi_{max} = 1 - \sqrt{1 - \zeta_*^2}, \tag{2.7}$$

and the corresponding trajectory is expressed by the formulas:

$$\zeta^0(t) = \tanh t, \qquad x_1^0(t) - x_1^0(0) = 1 - \frac{1}{\cosh t}. \tag{2.8}$$

Now we can pass to the solution of Problem A.

3. Solution of Problem A

Let $\sigma = \{x(t), u(t) \,|\, t \in [0, T]\}$ be an admissible trajectory in Problem A. Since $z_* > 0$, the initial point $x(0)$ belongs either to Q_1 or to Q_2, or to their common boundary D, except the point x_*. Obviously, trajectory $x(t)$ cannot get into the set Q_3; moreover, it cannot as well get in the set $z = 0$, $x_2 > 1$. Besides, $x(t)$ cannot have contacts with the set $\zeta > 0$, $x_1 > 1$. Thus, $x(t)$ can lie in Q_1, Q_2 or pass through D.

Denote by T' the minimal time instant τ, such that $x(t) \in Q_2$ on $(\tau, T]$. A'priori the following cases are possible: 1) there are no such τ, 2) $T' = 0$, 3) $0 < T' < T$.

In case 1 any trajectory $x(t) \in Q_1 \cup D$, so $x_1(T) \le 1$, and then

$$x_1(T) \le 2 - \sqrt{1 - \zeta_*^2}. \tag{3.1}$$

We will show that the value of Problem A satisfies the equality.

In case 2 from $z_* > 0$ it follows $x_1(0) < 1$. Then $\zeta(0) > 0$, because otherwise $\zeta(0) = 0$ implies $x(0) \in Q_1$, which contradicts with the definition of case 2.

Putting $u = 0$, let us continue the trajectory σ backwards up to the intersection with D. We get a trajectory σ', defined on interval $[T_1, T]$, $T_1 < 0$, such that $x_1(T_1) = x_1(0)$, $\zeta(T_1) = 0$. The σ' is obviously a trajectory of System B, satisfying the constraints of Problem B, whence $x_1(T) - x_1(0) \le 1 - \sqrt{1 - \zeta_*^2}$. But

$$x_1(T) < 2 - \sqrt{1 - \zeta_*^2}. \tag{3.2}$$

Let us pass to the case 3. Obviously, $x(T') \in D$. If $x(T') \ne x_*$, then $x_1(T') < 1$, $\zeta(T') > 0$, and, repeating the arguments of the case 2, we again can obtain (3.2).

Finally, consider the possibility $x(T') = x_*$. If trajectory $\sigma(t)$ coincides on $[T', T]$ with $\sigma^0(t - T')$, then by virtue of (2.7) and (2.8) we get (3.2), and so this trajectory gives the maximal value of $x_1(T)$. Thus, we only need that, starting from Q_1, the trajectory comes at x_*. Indeed, this is possible. Putting $u_2 = 0$, $u_1 = \sqrt{z + \varepsilon^2} - \varepsilon$, we get $\dot z = -(\sqrt{z + \varepsilon^2} - \varepsilon)$, $\dot x_2 = -1$, and putting $x_2 - 1 = \eta$, we arrive to equation

$$\frac{dz}{d\eta} = \sqrt{z + \varepsilon^2(\eta)} - \varepsilon(\eta).$$

It can be easily seen, that $z(\eta) = \eta^3/3$ is a solution to this equation on the interval $0 \le \eta \le 1/6$, and after that z increases up to $+\infty$. We can use, in particular, this solution. The Hamiltonian $\mathcal{H}$ for System A is equal to

$$\mathcal{H} = \begin{cases} -\psi_2 + |\psi|(\sqrt{z + \varepsilon^2(x_2 - 1)} - \varepsilon(x_2 - 1)), & \text{if } x \in Q_1, \\ -\psi_2 + |\psi|\zeta, & \text{if } x \in Q_2, \\ -\psi_2, & \text{if } x \in Q_3. \end{cases} \tag{5}$$

Because until hiting the point x_* the solution has $u_2 = 0$, it has $\psi_2 = 0$ too, and so until hiting x_* it has $\psi(t) = 0$. But since $\mathcal{H}$ is bounded in Q_2, it is Lipschitz continuous there, and so the solution to Hamiltonian system is unique, hence after x_* we as well get $\psi(t) = 0$.

Thus, solutions to Problem A cannot be accompanied by nontrivial absolutely continuous costate variables, satisfying jointly the Hamiltonian system. Therefore, in this example the Clarke's conjecture fails to be true. The correct answer here is that at the moment of passing the point x_* the costate variable ψ has a jump in the direction $e_1 = (1, 0)$. After that the state and costate components satisfy the Hamiltonian system. This result follows from the general theory of optimal control problems with mixed constraints. This theory is soon to appear in a book by A.A.Milyutin (in Russian).