Quadratic Sufficient Conditions for Minimality of Abnormal Sub-Riemannian Geodesics

A. V. Dmitruk

Contents

Introduction

The present paper aims to obtain quadratic sufficient minimality conditions for abnormal trajectories in the problem on length minimizing curves in the sub-Riemannian metrics. More precisely, we consider the trajectories that are subjected to (or, as one also says, are tangent to) some distribution on which some submetric is given; by a submetric, we understand an arbitrary positive sublinear functional on the distribution. In particular, it can be a sub-Riemannian metric. (Thus, the class of metrics that is studied in the paper is substantially wider than the class of sub-Riemannian metrics which is indicated in the title, but we have left this term in the title as a recognizable "key word.")

First, we deal with the following problem. Let a "singular" (i.e., a strictly abnormal) trajectory of some submetric that connects two given point be given. What would be quadratic sufficient conditions for this trajectory to be really the shortest one (with respect to a given submetric) among all other curves from some neighborhood of the given trajectory, which connect the same two points? This situation fits well into the framework of general class of optimal control problems that are linear in control, for which, in the case where the examined trajectory is singular [9], sufficient conditions of some quadratic order for the strong minimum and for the so-called Pontryagin minimum are already known. A direct use of these general results allows one to obtain quadratic sufficient conditions for the strong minimality and for the Pontryagin minimality of singular trajectories for distributions and manifolds of arbitrary dimensions.

Moreover, these conditions can be subjected to further (rather nontrivial) transformations, as a result of which we will arrive at some final form of these sufficient conditions. These final conditions exactly coincide with the quadratic sufficient conditions for rigidity that were obtained by A. A. Milyutin [15] not so long ago. Thus, the quadratic sufficient conditions for the strong (or the Pontryagin) minimum for a singular trajectory turn out to be identical to the quadratic sufficient conditions for the rigidity (respectively, for the Pontryagin rigidity) of this trajectory. We stress once more that all these results will be first obtained for singular trajectories.

On the other hand, it is known that the rigid trajectory of a given distribution should not necessarily be a singular one for arbitrarily chosen metric on this distribution. In what follows, we consider the case of an arbitrary quadratically rigid trajectory of some distribution, and, using the method that was proposed by A. A. Milyutin, we make the "reverse" passage, i.e., from the conditions for rigidity, we pass to the conditions for the strong minimum in an arbitrary submetric. As a result, we prove the following statement.

Main Theorem 1. *If a trajectory is quadratically rigid, then it yields a strong minimum of the distance between two given points in any strictly convex submetric* *(although this trajectory can become a nonsingular one).*

(Here we give a somewhat simplified statement; actually, a weaker property of the submetric than its strict convexity is required; see Assumption A3 and Theorem 8.1 below.)

The Pontryagin minimum is considered in a similar way with due regard for the fact that here the rigidity and the minimality depend on the choice of basis of the distribution. In this case, we prove the following statement.

Main Theorem 2. *If, in some basis, a trajectory is quadratically rigid in the Pontryagin sense, then, in any submetric from the sheaf corresponding to this basis, it yields the minimum of distance between two given points with respect to the* $W_{1,1}$ *-norm, i.e., with respect to the* L_1 *-norm in the velocities.*

(The exact definitions of all notions are given in Sec. 13, and the formulation of the statement is given in Theorem 14.1.)

These theorems were announced by the author in [10]; they generalize and strengthen the sufficient minimality conditions obtained in [13, 16, 17, 1] in the following directions: (a) only two-dimensional distributions and only sub-Riemannian metrics were considered previously; (b) the conditions obtained in [19, 13, 16, 17] ensure the minimum only for small segments of the curve; (c) the conditions obtained in [1] do not ensure the strong minimum; they guarantee only the $W_{1,1}$ -minimum; (d) the conditions from [1] contain somewhat extra strong requirements that can be essentially relaxed.

Note here that the problems of the rigidity and minimality of trajectories subjected to distributions have been recently studied very intensively. Not pretending to the completeness, note in this connection [3, 19, 13, 16, 17, 1, 15].

The following circumstance is worth noting. In spite of the fact that the problem of finding the shortest curve is, by its definition, an extremum problem, basically only the Pontryagin maximum principle, i.e., a necessary condition of the first order, has been used until recently, out of the whole theory of extremal problems, for obtaining minimality conditions in the above problem in the case of abnormal trajectories. All further conditions have been obtained principally not by the application of some general conditions to the given particular class of problems, but by the construction of special and rather complex structures facilitating the analysis of just this class of problems. Such a situation cannot be considered as satisfactory from the point of view of the theory of extremal problems. In the present paper, we attempt to apply the quadratic minimality conditions of singular trajectories that were previously obtained for the general class of linear optimal control problems [5, 6, 7, 8] (the results of these works which are necessary for us are summarized in [9]) to the given class of problems. This attempt proved to be successful in the sense that we have managed to obtain more general and more strong conditions than the already known ones. The quadratic *sufficient* conditions turned out to be most full and complete; namely these conditions will be the subject of the present paper. (In [15], the general quadratic minimality conditions were applied, also successfully, for obtaining quadratic conditions for rigidity; to this end, the concept of rigidity was first reduced to the concept of minimum in some optimal control problem.)

The author is grateful to A. A. Milyutin for intensive discussion and useful advices, and also to N. N. Petrov for noting the given class of problems as an object of possible application of previously obtained results. The work supported by the Russian Foundation for Basic Research, projects 97-01-00135 and 96-15-96072.

Part I
Sufficient Conditions for Strong Minimality of Singular Trajectories

1 Statement of the Problem on the Curve of Minimum Length in the Form of an Optimal Control Problem

Since all our considerations are made in a neighborhood of some given curve that is assumed to have no self-intersections, one can always assume that all is performed in an ordinary n-dimensional space, in order not to shade the heart of the matter by additional constructions related to a manifold.

Thus, assume that on an open connected set $\mathcal{M}$ in $\mathbb{R}^n$, there are given k twice smooth vector fields $r_0(x), \ldots, r_{k-1}(x)$ that are linearly independent at every point of $\mathcal{M}$. (It is convenient to adopt the enumeration $0, 1, \ldots, k-1$; this will be made clear in the sequel.) These fields define a so-called distribution

$$\Gamma(x) = \mathrm{Lin}\{r_0(x), \ldots, r_{k-1}(x)\} \tag{1.1}$$

of dimension k on the set $\mathcal{M}$. In the case where $\mathcal{M}$ is contractible (and a neighborhood of any trajectory that is of interest for us can always be considered to be such), the converse can also be stated: if a k-dimensional twice smooth (i.e., having twice smooth basis in a neighborhood of each point of $\mathcal{M}$) distribution $\Gamma(x)$ is given on $\mathcal{M}$, then there exist k twice smooth vector fields defined on the whole set $\mathcal{M}$, that form a basis of $\Gamma(x)$ at each point of $\mathcal{M}$, i.e., satisfy relation (1.1). We do not require that $\Gamma(x)$ should be bracket generating.

Definition 1.1. A real function $q(x,\bar{x})$ of arguments $x \in \mathcal{M}$ and $\bar{x} \in \Gamma(x)$ is called a submetric on $\Gamma(x)$ if, for any fixed $x \in \mathcal{M}$, it is a positive sublinear function of $\bar{x}$, i.e., if $q(x,\bar{x}) > 0$ for any nonzero $\bar{x} \in \Gamma(x)$, $q(x,\alpha\bar{x}) = \alpha\, q(x,\bar{x}) \quad \forall \alpha \geq 0$, and $q(x,\bar{x}_1+\bar{x}_2) \leq q(x,\bar{x}_1) + q(x,\bar{x}_2) \quad \forall \bar{x}_1, \bar{x}_2 \in \Gamma(x)$.

We assume a priori that the function q is continuous in $(x,\bar{x})$; some assumptions on its "twice smoothness" will be made in the sequel.

In the particular case when $q^2(x,\bar{x}) = (R(x)\bar{x},\bar{x})$ is a positive definite quadratic form in $\bar{x} \in \Gamma(x)$, the function q is called a sub-Riemannian metric, and it is said that the distribution $\Gamma(x)$ together with the submetric $q(x,\bar{x})$ define the the sub-Riemannian structure on $\mathcal{M}$. Thus, sub-Riemannian metric is simply a Euclidean metric on $\Gamma(x)$. Without loss of generality, one can assume that $R(x)$ is a matrix of dimension $n \times n$, i.e., that the Euclidean metric is given for all $\bar{x} \in \mathbb{R}^n$, but we will consider it only for $\bar{x} \in \Gamma(x)$. We will also assume that $R(x)$ is twice smooth. A more general, but still particular case of a submetric is the case of a sub-Finsler metric where q is symmetric with respect to $\bar{x}$: $q(x,-\bar{x}) = q(x,\bar{x})$.

Since each $\bar{x} \in \Gamma(x)$ can be represented in the form $\bar{x} = \sum u_i r_i(x)$, where $u = (u_0,\ldots,u_{k-1}) \in \mathbb{R}^k$, one can define a function

$$\varphi(x,u) = q(x,\bar{x}) = q(x, \sum u_i r_i(x)), \tag{1.2}$$

which specifies the submetric q in a given basis $r_0(x),\ldots,r_{k-1}(x)$.

In the case of a sub-Riemannian metric we have

$$(R(x)\bar{x},\bar{x}) = \sum_{ij} u_i u_j (R(x) r_i(x), r_j(x)) = (C(x)u,u),$$

where $C(x)$ is a matrix of dimension $k \times k$ with entries $C_{ij}(x) = (R(x)r_i(x), r_j(x))$, and the length of the vector $\bar{x}$ is expressed through its coefficients in the basis $r_0(x),\ldots,r_{k-1}(x)$ in the following way:

$$\|\bar{x}\| = q(x,\bar{x}) = \sqrt{(C(x)u,u)}.$$

If the basis is orthonormal, i.e., if $C(x) = E$ is the identity matrix, then

$$\|\bar{x}\| = |u|_2 = \sqrt{u_0^2 + \ldots + u_{k-1}^2}.$$

In the set of $\mathcal{M}$, we fix two points a and b and consider all possible absolutely continuous curves $x(t)$, $t \in [0,T]$ (generally the segment is its own for each curve), that connect these two points: $x(0)=a$, $x(T)=b$, and are "tangent" to the distribution $\Gamma(x)$, i.e., $\dot{x}(t) \in \Gamma(x(t))$ for almost all t. The curves that are tangent to Γ are called Γ-admissible or simply admissible. They can be also defined as solutions to the differential equation

$$\dot{x}(t) = \sum u_i(t) r_i(x(t)), \tag{1.3}$$

where all $u_i \in L_1[0,T]$.

For each admissible curve, its length is defined by the relation

$$J(x(\cdot)) = \int_0^T q(x,\dot{x})\,dt, \tag{1.4}$$

or, by using the representation (1.2),

$$J(x,u) = \int_0^T \varphi(x,u)\,dt.$$

In the case of a sub-Riemannian metric, we have

$$J = \int_0^T \sqrt{(R(x)\dot{x},\dot{x})}\,dt = \int_0^T \sqrt{(Q(x)u,u)}\,dt.$$

Since the length of a curve does not depend on the parametrization of this curve (this corresponds to the positive homogeneity of first degree of the function q of $\bar{x}$), one can consider all admissible curves on one and the same closed interval $[0,T]$.

It is required to find a curve of minimum length among all admissible curves that connect two given points a and b. Thus, we have the following optimal control problem.

Problem (I)**.**

$$J = \int_0^T \varphi(x,u)\,dt \longrightarrow \min, \quad \dot{x} = \sum u_i r_i(x), \quad x(0)=a, \quad x(T)=b.$$

We share the opinion (see, e.g., [13]), that when dealing with extremal problem in the class of Γ-admissible curves, it is more convenient to deal with the distribution represented in the form of the control system (1.3), rather than in the form of zeros of differential forms, which is more usual for the differential geometers. (The latter corresponds to the dual method for defining the distribution $\Gamma(x)$ as a set of those $\bar{x} \in \mathbb{R}^n$ for which $(\tilde{r}_j(x),\bar{x})=0$, where $\tilde{r}_j(x)$, $j=1,\dots,n-k$, is a basis in the complement to $\Gamma(x)$.)

2 Basic Notions and Assumptions

Now let $\hat{x}(t)$ be one of the admissible curves. The question is: what are necessary and sufficient conditions for this curve be of minimum length among all admissible curves from some its neighborhood (and connecting the same two points)? The present paper is devoted to the study of namely this question.

We need some assumptions on the curve $\hat{x}$ and on the behavior of the submetric q in a neighborhood of this curve.

Assumption A1. *The curve $\hat{x}(t)$, $t \in [0,T]$, is three times smooth, has the nonzero derivative, and does not have self-intersections.*

(The latter requirement is actually not essential; one can easily eliminate it by introducing an additional coordinate; however we accept this requirement in order not to complicate the presentation.)

Let $\hat{\chi}$ denote the image of the curve $\hat{x}(t)$, $t \in [0,T]$ in the space $\mathbb{R}^n$.

Definition 2.1. We say that the submetric $q(x,\bar{x})$ has a twice smooth support hyperplane in a neighborhood of the curve $\hat{x}(t)$ if, in a neighborhood of the set $\hat{\chi}$, there exist a twice smooth $(k-1)$-dimensional subspace $\Gamma_0(x) \subset \Gamma(x)$ and a twice smooth nonvanishing vector field $r_0(x) \in \Gamma(x)$ such that the affine hyperplane $r_0(x)+\Gamma_0(x)$ has no common points with the relative interior of the hodograph $F(x) = \{\bar{x} \in \Gamma(x) \mid q(x,\bar{x}) \le q(x, r_0(x))\}$, and, in addition, $\frac{d}{dt}\,\hat{x}(t) = r_0(\hat{x}(t)) \quad \forall t$.

Given a basis in $\Gamma(x)$, this property means that in a neighborhood of the set $\hat{\chi}$, there exist twice smooth nonvanishing vector-functions $l(x), v(x) \in \mathbb{R}^k$ such that if $u \in \mathbb{R}^k$ and $\varphi(x,u) \le \varphi(x,v(x))$, then $(l(x),u) \le (l(x),v(x))$ and, in addition, $v(\hat{x}(t)) = \hat{u}(t) \quad \forall t$.

Assumption A2. The submetric q has a twice smooth support hyperplane in a neighborhood of the curve $\hat{x}(t)$.

For example, it is easy to show that any three times smooth metric q on $\mathcal{M} \times \mathbb{R}^n$ or any twice smooth Riemannian metric, being restricted to any twice smooth distribution $\Gamma(x)$, has a twice smooth support hyperplane in a neighborhood of any admissible curve $\hat{x}(t)$ satisfying Assumption A1. In what follows, we will omit the words "twice smooth" for brevity.

Definition 2.2. A support hyperplane $\Gamma_0(x)$ for the submetric q in a neighborhood of the curve $\hat{x}(t)$ is called a strictly support hyperplane if, for any x from a neighborhood $\hat{\chi}$, the affine hyperplane $r_0(x)+\Gamma_0(x)$ has the unique common point $r_0(x)$ with the hodograph $F(x)$.

Given a basis in $\Gamma(x)$, this property means that in a neighborhood of the set $\hat{\chi}$, there exist twice smooth nonvanishing vector-functions $l(x), v(x) \in \mathbb{R}^k$ for which, in addidion to the above properties, the following property holds: if $u \in \mathbb{R}^k$, $\varphi(x,u) \leq \varphi(x, v(x))$, and $(l(x), u) = (l(x), v(x))$, then $u = v(x)$.

Obviously, if a submetric is strictly convex (i.e., if the sublinear function $q(x, \bar{x})$ is strictly convex in $\bar{x}$ $\forall x$), then any support hyperplane will automatically be strictly support hyperplane. In particular, this property holds for any sub-Riemannian metric.

We will also need the following strengthening of Assumption A2 (for obtaining sufficcient conditions for the strong minimality).

Assumption A3. The submetric q has a twice smooth *strict* support hyperplane in a neighborhood of the curve $\hat{x}(t)$.

Note that Assumptions A2 and A3 hold regardless of the choice of parametrization of the curve $\hat{x}(t)$. Any sub-Riemannian metric satisfies Assumption A3.

Thus, we study admissible curve $\hat{x}(t), \hat{u}(t)$ in Problem (I) and suppose for the time being that Assumptions A1 and A3 hold. It will be convenient for us to reformulate slightly the problem. Problem (I) is obviously equivalent to (and is usually stated as) the following time-optimal control problem:

Problem (T) **.**

$$\dot{x} = \sum u_i r_i(x), \qquad x(0) = a, \quad x(T) = b,$$

$$\varphi(x,u) \leq 1, \qquad J(x,u) = T \longrightarrow \min .$$

But it will be more convenient for us to state it as the following problem on a fixed time interval $[0, T]$:

Problem (Z) **.**

$$\dot{x} = z \sum u_i r_i(x), \qquad \dot{z} = 0, \tag{2.1}$$

$$x(0) = a, \quad x(T) = b, \tag{2.2}$$

$$z > 0, \qquad \varphi(x,u) \leq 1, \qquad J = z(0) \longrightarrow \min .$$

Here z is an additional state variable; it is a constant that bounds the velocity of the motion:

$$||\dot{x}|| = q(x, \dot{x}) \;\leq\; z \cdot \varphi(x,u) \;\leq z,$$

and the problem consists in steering the point a to the point b in a given time T with the minimally possible upper bound of the velocity. In this problem, the state variables $(x,z) \in \mathbb{R}^{n+1}$ and all the curves are considered in the open set $x \in \mathcal{M}$, $z > 0$. The control is still $u = (u_0, \dots, u_{k-1}) \in \mathbb{R}^k$. Obviously, Problems (T) and (Z) are equivalent.

Thus, namely Problem (Z) will be considered from now on. We denote $w = (z, x, u)$ and study, respectively, a trajectory $\hat{w} = (\hat{z}, \hat{x}, \hat{u})$.

Recall that according to the classical calculus of variations (CCV), one says that a trajectory $\hat{w}$ yields the *weak minimum* in Problem (Z) if it is a local minimum point with respect to the norm $\|w\| = |z| + \|x\|_C + \|u\|_\infty$; and that it yields the *strong minimum* if it is a local minimum point with respect to the seminorm $\|w\|' = |z|+\|x\|_C$ for free u .

Lemma 2.1. *A trajectory $\hat{w}$ yields a strong (strictly strong) minimum in Problem (Z) if and only if the curve $\hat{x}(t)$ is of minimum (strictly minimum) length among all admissible curves connecting the points a, b and lying in a neighborhood of the set $\hat{\chi}$.*

This means that the concept of strong minimum in Problem (Z) is in agreement with the concept of local minimum of the length of the curve in the usual geometric sense. The proof is given in Appendix A.

We will also consider the following type of the minimum, which is intermediate between the weak and strong ones.

Definition 2.3. We will say that $\hat{w}$ is a *Pontryagin minimum* (briefly, Π -*minimum*) point in Problem (Z) if, for any N, there exists an $\varepsilon > 0$ such that $\hat{w}$ is a minimum point in Problem (Z) on the set

$$|z - \hat{z}| + \|x - \hat{x}\|_C < \varepsilon, \qquad \|u - \hat{u}\|_1 < \varepsilon, \qquad \|u - \hat{u}\|_\infty \leq N.$$

It is obvious that for any submetric, the Π -minimum is simply the minimum with respect to the norm $\|w\|_1 = |z| + \|x\|_C + \|u\|_1$.

Further, note that there is a mixed constraint in Problem (Z) : $\varphi(x,u) \leq 1$. Let us introduce the admissible control set (the hodograph or the unit ball of the submetric in the given basis):

$$U(x) = \{u \,|\, \varphi(x,u) \leq 1\}, \tag{2.3}$$

which in general depends on x. Thus, the mixed constraint $\varphi(x,u) \leq 1$ is equivalent to the inclusion $u \in U(x)$.

In the case of a sub-Riemannian metric, one can assume that the vectors $r_0(x), \dots, r_{k-1}(x)$ form an orthonormal basis in $\Gamma(x)$; then $\varphi(x,u) = |u|$, and thereby the mixed constraint is transformed into the classical Pontryagin constraint on the control of the form $|u| \leq 1$, i.e., $u \in U$ with a constant set U. As far as the general case is concerned, one cannot eliminate the dependence of U on x .

3 Maximum Principle for Problem (Z)

Let $\hat{w}$ be a Π-minimum point in Problem (Z). Then the first order necessary condition, i.e., the Pontryagin Maximum Principle, holds for this point. Let us write this condition first for the case of a sub-Riemannian metric and then show what will change in the general case.

The Maximum Principle (MP) means that there exist Lipschitzian functions $\psi_z(t)$ and $\psi_x(t)$ of dimensions n and 1, respectively, and numbers $\alpha_0 \geq 0$, β_0, and β_T such that the tuple (which is called the tuple of Lagrangian multipliers) $\lambda = (\psi_z(\cdot), \psi_x(\cdot), \alpha_0, \beta_0, \beta_T)$ is nontrivial ($\lambda \neq 0$); to this tuple, there correspond

the Pontryagin function $H[\lambda](z,x,u) = \psi_x \cdot z \sum u_i r_i(x) + \psi_z \cdot 0$

and the endpoint Lagrange function $l[\lambda](x_0, x_T) = \alpha_0 z(0) + \beta_0 x(0) + \beta_T x(T)$,

for which the following relations hold along the trajectory $\hat{w}$:

the adjoint equations

$$\dot{\psi}_x = -H_x[\lambda] = -\hat{z}\psi_x \sum \hat{u}_i r_i'(\hat{x}), \tag{3.1}$$

$$\dot{\psi}_z = -H_z[\lambda] = -\psi_x \sum \hat{u}_i r_i(\hat{x}), \tag{3.2}$$

the transversality conditions

$$\psi_x(0) = l'_{x_0}[\lambda] = \beta_0, \qquad \psi_x(T) = -l'_{x_T}[\lambda] = -\beta_T, \tag{3.3}$$

$$\psi_z(0) = l'_{z_0}[\lambda] = \alpha_0, \qquad \psi_z(T) = l'_{z_T}[\lambda] = 0, \tag{3.4}$$

and the maximality condition

$$H[\lambda](\hat{z}, \hat{x}, \hat{u}) = \max_{|u| \leq 1} H[\lambda](\hat{z}, \hat{x}, u) = \text{const } \geq 0. \tag{3.5}$$

From (3.5) it follows that $\psi_x \cdot \sum \hat{u}_i r_i(\hat{x}) = \text{const } \geq 0$; then from (3.2) we have $\dot{\psi}_z = -\text{const}$, and (3.4) implies

$$\psi_z(0) - \psi_z(T) = \alpha_0 = \int_0^T \text{const}\, dt;$$

hence,

$$\text{const } = \psi_x \cdot \sum \hat{u}_i r_i(\hat{x}) = \frac{\alpha_0}{T} \geq 0. \tag{3.6}$$

Now we introduce the following conventions with respect to the terms. By a *trajectory* in Problem (Z) we mean an admissible triple $w = (z, x, u)$. Taking into account that u is uniquely determined by z and x from system (2.1), and the choice of z exerts influence only on the parametrization of one and the same curve in the space x, sometimes the function $x(t)$ itself will be called a trajectory. Following A. A. Milyutin, by an *extremal* we call a pair (λ, w), where w is a trajectory and λ is a tuple of corresponding Lagrange multipliers satisfying all the above-noted relations, except for the nontriviality. (In [13], such pair is called a biextremal.) Thus, different extremals can correspond to one and the same trajectory. An extremal (λ, w) is called *trivial* if $\lambda = 0$; otherwise, the extremal is called *nontrivial.* (In works of differential geometers (see [3, 13, 17]) the passage from a trajectory to a nontrivial extremal is called the Hamiltonian lift.)

The set of all λ satisfying relations (3.1)–(3.5) for the given trajectory $\hat{w}$ of Problem (Z) and the normalization

$$\alpha_0 + |\beta_0| + |\beta_T| + \max_t |\psi(t)| \;=\; 1,$$

we denote by $\Lambda(Z, \hat{w})$, omitting Z and $\hat{w}$ if it does not lead to confusion. It is easy to see that $\Lambda(Z, \hat{w})$ is a finite-dimensional compact set, since each λ is defined by its own α and β; hence, the choice of normalization in the definition of this compact set would be, in essence, of no importance for us.

A trajectory $\hat{w}$ is called a *stationary* trajectory in Problem (Z) if $\Lambda(Z, \hat{w})$ is not empty, i.e., if this trajectory can be completed up to a nontrivial extremal $(\lambda, \hat{w})$. (Obviously, any admissible trajectory can be completed up to a trivial extremal.)

For every $\lambda \in \Lambda(Z, \hat{w})$, two essentially distinct cases are possible:

$$(1) \quad \alpha_0 > 0, \qquad \text{and} \qquad (2) \quad \alpha_0 = 0.$$

If $\alpha_0 > 0$, then the extremal $(\lambda, \hat{w})$ is said to be *normal.* In this case (3.5) and (3.6) imply

$$\sum \hat{u}_i(\psi_x, r_i(\hat{x})) = \max_{|u|\le 1} \sum u_i(\psi_x, r_i(\hat{x})) = \frac{\alpha_0}{T} > 0, \tag{3.7}$$

and since the unit ball in the Euclidean metric is a strictly convex set, the control $\hat{u}(t)$ is uniquely expressed through $\psi_x(t)$, namely, $\hat{u}_i(t) = c\,(\psi_x(t), r_i(\hat{x}(t)))$, $i = 0, 1, \ldots, k-1$, where c is some constant.

If $\alpha_0 = 0$, then the extremal $(\lambda, \hat{w})$ is said to be *abnormal.* This term has appeared already in the CCV [2]; it denotes the degenerate case in the sense that the objective functional of the problem does not enter the first order necessary conditions. For Problem (Z), in this case, (3.5) implies that

$$\psi_x(t)\, r_i(\hat{x}(t)) = 0 \qquad \forall i = 0, 1, \ldots, k-1, \qquad \text{i.e. } \psi_x(t) \perp \Gamma(\hat{x}(t)); \tag{3.8}$$

therefore, the condition (3.5) does not select $\hat{u}$ from the whole unit ball. In optimal control, one says in such a situation that the control $\hat{u}$ is *singular* with respect to the set $|u| \le 1$ (and, taking into account the fact that this takes place at every time instant, one sometimes even says that the control $\hat{u}$ is *totally singular*; but we do not use this "strengthened" term). Conversely, it is easy to see that (3.8) implies, by virtue of (3.7), that $\alpha_0 = 0$. Thus, the equality $\alpha_0 = 0$ is equivalent to the fulfilment of (3.8), i.e., in Problem (Z), the abnormality of an extremal $(\lambda, \hat{w})$ is equivalent to its singularity.

Thus, a nontrivial extremal $(\lambda, \hat{w})$ can be either normal $(\alpha_0 > 0)$, or abnormal, the latter in Problem (Z) is the same as singular $(\alpha_0 = 0)$.

A stationary trajectory $\hat{w}$ is said to be *normal* if we have $\alpha_0 > 0 \quad \forall \lambda \in \Lambda(\hat{w})$; it is said to be *abnormal* if $\exists \lambda \in \Lambda(\hat{w})$ for which $\alpha_0 = 0$; and it is said to be *strictly abnormal* or *singular* if we have $\alpha_0 = 0 \quad \forall \lambda \in \Lambda(\hat{w})$.

(This notion of singularity corresponds to the general notion of singular constraint in the abstract extremum problem; see [9]. In our case, namely the objective functional of the problem is singular, since the multiplier α_0 corresponds exactly to this functional.)

Note that if a trajectory $\hat{w}$ is abnormal, then the set $\Lambda(\hat{w})$ always consists of more than one element, since in this case, along with any $\lambda \in \Lambda(\hat{w})$, we have also $-\lambda \in \Lambda(\hat{w})$.

The abnormality of $\hat{w}$ means that the equality-type constraints (2.1) and (2.2) in Problem (Z) are degenerate in the first order, i.e., they do not satisfy the known conditions of the implicit function theorem, or, as is conventional now, the Lyusternik condition: the derivative of the operator that defines these equality-type constraints should be "onto"(surjective). If we consider the operator

$$g: \ \mathbb{R} \times AC \times L_\infty[0,T] \ \longrightarrow \ L_1[0,T] \times \mathbb{R}^n \times \mathbb{R}^n,$$

$$(z, x, u) \ \longmapsto \ (\dot{x} - z \sum u_i r_i(x), \ x(0), x(T)),$$

then the Lyusternik condition means that $g'(\hat{w})$ is "onto", and the abnormality of $\hat{w}$ means that $\operatorname{Im} g'(\hat{w})$ is not the whole image space.
Sometimes, instead of this "initial" equality operator, one considers the operator

$$p: \ L_\infty[0,T] \longrightarrow \mathbb{R}^n, \qquad u \mapsto x(T),$$

where x is the solution to the equation $\dot{x} = \hat{z} \sum u_i r_i(x)$ with the initial condition $x(0) = a$. It is not difficult to verify that the derivatives $g'(\hat{w})$ and $p'(\hat{u})$ are surjective or not surjective simultaneously. (This is due to the following general fact: if linear operators $A: X \to Y$ and $B: X \xrightarrow{\text{onto}} Z$ are given, then the joint operator $(A, B):$ $X \to Y \times Z$ is surjective if and only if the operator $A: \ker B \to Y$ is surjective.) Thus, in Problem (Z), the abnormality of $\hat{w}$, i.e., the existence of $\lambda \in \Lambda(\hat{w})$ with $\alpha_0 = 0$, is equivalent to the degeneracy of the operator g or p at the point $\hat{w}$. It is worth noting that in the general case, the degeneracy of the equality-type constraints is a more strong condition than the existence of $\lambda \in \Lambda(\hat{w})$ with $\alpha_0 = 0$ (the equality-type constraints can be nondegenerate, but there can exist $\lambda \in \Lambda(\hat{w})$ with $\alpha_0 = 0$).

Note that in Problem (Z), the whole tuple $\lambda = (\alpha_0, \beta_0, \beta_T, \psi_x, \psi_z)$ is determined by the function $\psi_x(t)$, which in the sequel is denoted by $\psi(t)$ (the vectors β_0 and β_T are determined from the transversality conditions (3.3), the numbers α_0 and ψ_z are obtained from relation (3.6), from the adjoint equation (3.2), and from the transversality conditions (3.4).) Hence, instead of the set $\Lambda(\hat{w})$, one can consider its projection on the component ψ_x, which is denoted by $\Psi(\hat{w})$.

For the abnormal extremal, $\alpha_0 = 0$ and $\psi_z = 0$, the function $\psi = \psi_x$ satisfies the adjoint equation (3.1)

$$\dot{\psi} = -\psi \cdot \sum \hat{u}_i r_i'(\hat{x}) \tag{3.9}$$

and relations (3.8):

$$\psi(t)\, r_i(\hat{x}(t)) = 0, \qquad i = 0, 1, \ldots, k-1. \tag{3.10}$$

The set of all nonzero Lipschitzian n-dimensional functions $\psi(t)$ satisfying (3.9) and (3.10) with the normalization $|\psi(0)| = 1$ we denote by $\Psi_0(Z, \hat{w})$, sometimes omitting Z and $\hat{w}$. These are those $\psi \in \Psi$ for which $\alpha_0 = 0$. Thus, a trajectory $\hat{w}$ is abnormal if and only if there exists a function $\psi(t)$ satisfying the above conditions, i.e., if the set $\Psi_0(\hat{w})$ is not empty. The trajectories with nonempty $\Lambda(\hat{w})$, but with empty $\Psi_0(\hat{w})$, are normal stationary trajectories. Obviously, for the singular trajectories $\Psi(\hat{w}) = \Psi_0(\hat{w})$.

Let us now write the MP for Problem (Z) in the case of an arbitrary submetric, which generates the mixed constraint

$$\varphi(x, u) \leq 1. \tag{3.11}$$

Recall that, due to the assumptions made, the function φ is continuous in (x, u) and is sublinear in u. We assume for the time being that φ is smooth in some tube around the trajectory $(\hat{x}(t), \hat{u}(t))$ and that its derivatives are continuous at the points of this trajectory. This property is necessary for us only for writing out the MP in Problem (Z). In the sequel, starting from Sec. 4, we will not use the MP for this problem with the initial constraint (3.11), since we will pass to Problem (Z) with a more wide constraint.

The MP for problems with mixed constraints, which is a generalization of the Pontryagin Maximum Principle, was obtained in the late 1960-s by A. Ya. Dubovitskii and A. A. Milyutin (see [12, 14]) and by L. Neustadt and K. Makowsky [18]. Being applied to Problem (Z) with constraint (3.11), it will differ from the above-presented MP (for the constraint $|u| \leq 1$) only by the fact that to constraint (3.11), there corresponds its own multiplier $\mu(\cdot) \in L_1[0, T]$, $\mu(t) \geq 0$ a.e., satisfying the complementarity slackness condition

$$\mu(t)\Big(\varphi(\hat{x}(t), \hat{u}(t)) - 1\Big) = 0 \qquad \text{a.e. on } [0, T]; \tag{3.12}$$

this multiplier also enters the adjont equation:

$$\dot{\psi} = -\hat{z}\,\psi \sum u_i\, r_i'(\hat{x}) \;+\; \mu(t)\frac{\partial \varphi}{\partial x}(\hat{x},\hat{u}), \tag{3.13}$$

and appears in an additional condition, condition of stationarity w.r.t. u :

$$\hat{z}(\psi, r_i(\hat{x})) - \mu(t)\frac{\partial \varphi}{\partial u}(\hat{x},\hat{u}) = 0 \qquad \forall i = 0,1,\dots,k-1. \tag{3.14}$$

The maximality condition becomes

$$\max_{u\in U(\hat{x}(t))} H[\lambda](\hat{z},\hat{x}(t),u) \;=\; H[\lambda](\hat{z},\hat{x}(t),\hat{u}(t)) \;= \text{const} \,\geq 0. \tag{3.15}$$

Other conditions remain the same. As before, all multipliers are expressed through $\psi = \psi_x$. The following still holds:

$$\max_{u} H(\hat{z},\hat{x},u) = \hat{H} = H(\hat{z},\hat{x},\hat{u}) = \hat{z}\sum \hat{u}_i(\psi_i, r_i(\hat{x})) = \text{const} \;= \frac{\alpha_0}{T};$$

from here, by (3.14) and (3.12), we have

$$\mu(t) = \mu(t)\,\varphi(\hat{x},\hat{u}) = \mu(t)\sum\big(\frac{\partial \varphi}{\partial u_i},\hat{u}_i\big) = \hat{z}\sum \hat{u}_i\,(\psi, r_i(\hat{x})) = \hat{H} = \frac{\alpha_0}{T}, \tag{3.16}$$

and, therefore, $\mu(t) = \text{const} \,= \hat{H}$. If the extremal $(\lambda,\hat{w})$ is abnormal, i.e., if $\alpha_0 = 0$, then, due to (3.16), also $\mu(t) = 0$ almost everywhere, i.e., the new element in the MP in fact disappears.

This implies that if $(\lambda,\hat{w})$ is an abnormal extremal in Problem (Z) with constraint (3.11), then it will also be an abnormal extremal in Problem (Z) with any other (smooth) mixed constraint $p(x,u) \leq 0$ (not necessarily related to some submetric), in particular, with the constraint $p(x,u) = (l(x), u - v(x)) \leq 0$, where l and v are from the definition of the support hyperplane, and also for free u, i.e., with the constraint $u \in \mathbb{R}^k$.

Note that if an extremal is normal, then, by virtue of (3.16), $\mu(t) = \frac{\alpha_0}{T} > 0$, and then the maximality condition (3.15) or the condition of complementary slackness (3.12) implies that $\varphi(\hat{x}(t),\hat{u}(t)) \equiv 1$, i.e., the motion is performed with the maximal possible speed. If, on the other hand, $\alpha_0 = 0$, i.e., if the extremal is abnormal, then this relation does not follow from the MP. Nevertheless, we can always assume that this relation holds by virtue of the following argument. If on some time interval $\varphi(\hat{x}(t),\hat{u}(t)) < 1$, then one can consider a new control $u'(t) = c(t)\hat{u}(t)$, where $c(t) > 1$ on this interval, and $c(t) = 1$ outside this interval. The state component $x'(t)$ of the corresponding trajectory $(\hat{z}, x'(t), u'(t))$ will move along the same curve $\hat{\chi}$, but will pass the way from a to b in a smaller time T'. Then one can pass the same way in time $T > T'$ with a smaller bound on the speed $z' < z$. Therefore, the given trajectory $\hat{w}$ will not be even a weak minimum point in Problem (Z). Thus, we will assume that $\varphi(\hat{x}(t),\hat{u}(t)) \equiv 1$ for the examined trajectory.

Let us now return to the question on the expansion of the admissible control set. Let the functions $l(x)$ and $v(x)$ from the definition of the support hyperplane be such that in a neghborhood of the set $\hat{\chi}$,

$$\varphi(x, v(x)) = 1, \qquad (l(x), v(x)) = 1, \tag{3.17}$$

i.e., the set $U(x)$ is contained in the halfspace $(l(x), u) \le 1$. (This can always be obtained by the corresponding normalization of these functions.) Problem (Z) with the constraint $\varphi_*(x,u) = (l(x),u) \le 1$ will be called Problem (Z_*). It is an extension of Problem (Z) with the initial constraint $\varphi(x,u) \le 1$.

Lemma 3.1. *If $(\lambda, \hat{w})$ is an extremal in Problem (Z) with the constraint $\varphi(x,u) \le 1$, then it will also be an extremal in Problem (Z_*) with the constraint $\varphi_*(x,u) \le 1$.*

Proof. As was already said, this statement is obvious for the abnormal extremals. Hence, it is sufficient to consider the case where $\alpha_0 > 0$ for the given extremal. By definition, for all x from some neighborhood of $\hat{\chi}$, the inequality $\varphi(x,u) \le 1$ implies $\varphi_*(x,u) \le 1$, and both these inequalities turn into equalities for $u = v(x)$. Then in a neighborhood of any point $(x_0, u_0 = v(x_0))$, the inequality $\varphi_*(x,u) \le \varphi(x,u)$ holds. (In fact, since $\varphi = \varphi_* = 1$ at this point itself, both these functions are positive in a neighborhood of this point; if $\varphi(x,u) = c > 0$, then $\varphi(x, u/c) = 1$; therefore, $\varphi_*(x, u/c) \le 1$, i.e., $\varphi_*(x,u) \le c$.) This implies that the gradients of functions φ and φ_* are collinear at any such point. Since both these functions are sublinear in u, are equal and positive at the point $u_0 = v(x_0)$, we have that the gradients of these functions in u are simply equal, and, moreover, are not equal to zero. Therefore, the gradients of these functions in x are also equal, i.e., finally, $\operatorname{grad}\varphi = \operatorname{grad}\varphi_*$ at any point of the form $(x_0, u_0 = v(x_0))$. And then, for a given λ, relations (3.13) and (3.14) hold not only with the function φ, but also with φ_*, i.e., $\lambda \in \Lambda(Z_*, \hat{w})$, or, which is the same, $(\lambda, \hat{w})$ is an extremal in Problem (Z_*). The lemma is proved.

Corollary 1. *The set $\Lambda(\hat{w})$ in Problems (Z) and (Z_*) is one and the same.*

Indeed, since in the passage from the constraint $\varphi \le 1$ to the constraint $\varphi_* \le 1$ the admissible set expands, the set $\Lambda(\hat{w})$ can only narrow. But, according to Lemma 3.1, no narrowing takes place, therefore, these sets coincide.

This fact readily implies

Corollary 2. *The trajectory $\hat{w}$ is singular in Problem (Z) if and only if it is singular in Problem (Z_*) .*

In addition, the following is valid.

Corollary 3. *If the trajectory* $\hat{w}$ *is singular in Problem* (Z) *with the constraint* $\varphi(x,u) \le 1$, *then it will be stationary and, therefore, automatically singular also for free* u. *The set* $\Lambda(\hat{w})$ *will not change under this procedure.*

Indeed, in the passage from the constraint $\varphi(x,u) \le 1$ to $u \in \mathbb{R}^k$, the set $\Lambda(\hat{w})$ can only narrow. However, since for the singular trajectory, any $\lambda \in \Lambda(\hat{w})$ has $\mu(t) = 0$, relations (3.13) and (3.14) (the only relations in the MP that contain φ), as was already said, will hold also for the problem with free u. Thus, in this passage the trajectory $\hat{w}$ will remain stationary, and the set $\Lambda(\hat{w})$ will not change.

Note that if the extremal is normal, then the corresponding Lagrange function $L[\lambda](w) = -H[\lambda](z,x,u) + \mu\varphi(x,u)$ has the coefficient $L_{uu}[\lambda] = \mu\varphi_{uu}$ (here we take into account the fact that $H_{uu}[\lambda] = 0$), and assuming that the matrix φ_{uu} is positive definite on the subspace $\varphi_u \bar{u} = 0$ along the trajectory $\hat{w}(t)$ (which always holds in the case of sub-Riemannian metrics), we obtain that the strengthened Legendre condition is satisfied for this extremal. Therefore, the case of a normal extremal can be considered within the framework of the CCV (see, e.g., [13]). But, for the abnormal extremals, $L_{uu}[\lambda] = 0$, i.e., the principal assumption of the CCV, the strengthened Legendre condition, is not satisfied, hence, this case requires a special consideration. It is precisely this case that is the subject of the present paper.

4 Passage to an Associated Basis

Up to now all our considerations were in an arbitrary basis. Now we will consider some special bases.

Definition 4.1. Following A. A. Milyutin [15], a basis in $\Gamma(x)$ is said to be *associated* for the trajectory $\hat{x}(t)$ if $r_0(\hat{x}(t)) = \frac{d}{dt}\hat{x}(t)$ on $[0,T]$.

In any such basis the examined control is $\hat{u}(t) = (1,0,\ldots,0)$. The state components $\hat{z}$ and $\hat{x}(t)$ remain unchanged, since they do not depend on the choice of basis in $\Gamma(x)$.

Definition 4.2. An associated basis in $\Gamma(x)$ is said to be *support (strictly support)* for a submetric q if, in addition to the above property, in some neighborhood of $\hat{\chi}$, the subspace $\Gamma_0(x) = \mathrm{Lin}\{r_1(x),\ldots,r_{k-1}(x)\}$ is a support (strictly support) hyperplane to the hodograph $F(x)$ of the submetric q in the sense of Definitions 2.1 and 2.2.

The existence of a twice smooth associated basis that is support (strictly support) for a given metrics, is ensured by Assumption A2 (Assumption A3, respectively). Problem (Z) in the associated support basis has the same form as in Sec. 2, but now the term with $i = 0$ will be singled out in the notation of the control system:

Problem (Z) **.**

$$\dot{x} = z\left(u_0 r_0(x) + \sum_{i=1}^{k-1} u_i r_i(x)\right), \tag{4.1}$$

$$\dot{z} = 0, \qquad x(0) = a, \quad x(T) = b, \tag{4.2}$$

$$u = (u_0, u_1, \dots, u_{k-1}) \in U(x),$$

$$J = z(0) \longrightarrow \min .$$

In addition, we set $\hat{z} = 1$ for convenience (having made a linear change of time, if necessary); then T is the time of motion.

Just for this very Problem (Z) in an associated support basis, we will apply the quadratic sufficient conditions for the Pontryagin and strong minima, obtained for the problems of general form in [9]. This will be our "starting"position.

As was already said, in the case of a sub-Riemannian metrics, one can assume that $U(x)$ is the unit ball which does not depend on x. As for the case of an arbitrary submetric, the admissible control set depends on x, i.e., there is a mixed constraint on x, u, and this is rather inconvenient, since the quadratic conditions obtained in [9] (and they are the most general of all known ones) do not allow the presence of mixed constraints.

However, this difficulty can easily be overcome when obtaining sufficient conditions. As in [9], we replace the family of sets $U(x)$ by the ambient constant set $\tilde{U}$ that includes this family. It is clear that any sufficient condition for a minimum of some type (Pontryagin, strong) in Problem (Z) with the set $\tilde{U}$ will automatically be a sufficient condition for the minimum of the same type also in Problem (Z) with the set $U(x)$.

Moreover, for obtaining sufficient conditions for the Π-minimum, we adopt Assumption A2 and simply take the half-space $U_* = \{u_0 \le 1\}$ as $\tilde{U}$. From Definition 4.2, it follows that this set indeed contains $U(x)$ for all x from a neighborhood of the set $\hat{\chi}$. If the trajectory was singular in Problem (Z) with the set $U(x)$, then, according to Corollary 2 of Lemma 3.1, this trajectory will remain singular under the passage to the set U_* .

To obtain sufficient conditions for the strong minimum, we adopt Assumption A3. Then, from Definition 4.2, it follows that in a corresponding associated basis, which is strictly support one for the given submetrics, for any x from some neighborhood of the set $\hat{\chi}$, the hyperplane $u_0 = 1$ is strictly support to the set $U(x)$ at the point $\hat{u} = (1, 0, \ldots, 0)$, i.e., $\hat{u}$ is a unique maximum point of the linear functional $p(u) = u_0$ on the set $U(x) \subset \mathbb{R}^k$. The sublinearity and positivity of $\varphi(x, \cdot)$ imply that the set $U(x)$ is compact, and from the continuity of φ it follows that this set is Hausdorff continuously depending on x. Then one can state that there exists a convex compact set $\tilde{U} \subset \mathbb{R}^k$, which also has $\hat{u}$ as a unique maximum point of the functional $p(u) = u_0$ and such that $U(x) \subset \tilde{U}$ for any x from some neighborhood of $\hat{\chi}$. (The proof of this fact is given in Appendix B.)

We consider first the case of the strong minimum. (The case of the Pontryagin minimum is considered below in Part III.)

We study Problem (Z) with the above control set $\tilde{U}$, which does not depend on x. Let us call it Problem $(\tilde{Z})$. Note that the particular form of the set $\tilde{U}$ is of no importance for us; we are interested only in the existence of this set; further, we will pass from this set to the halfspace U_* that contains this set, and this halfspace does not depend on the choice of $\tilde{U}$.

Thus, in the case of an arbitrary submetric (satisfying Assumption A3), we arrive at Problem $(\tilde{Z})$ in an associated basis with a convex compact set $\tilde{U}$, not depending on x, for which $\hat{u} = (1, 0, \ldots, 0)$ is a unique maximum point of the linear functional $p(u) = u_0$. (In the case of a sub-Riemannian metrics, $U(x)$ is the unit ball, therefore, no passage to $\tilde{U}$ is required.) The examined trajectory is still $\hat{w} = (\hat{z} = 1, \hat{x}(t)$, $\hat{u} = (1, 0, \ldots, 0))$. Corollary 1 of Lemma 3.1 implies the following statement.

Lemma 4.1. *The set* $\Lambda(\hat{w})$ *in Problems* (Z) *and* $(\tilde{Z})$ *is one and the same.*

Indeed, since $U(x) \subset \tilde{U} \subset U_*$, the set $\Lambda(\hat{w})$ for Problem $(\tilde{Z})$ occupies an intermediate place between $\Lambda(Z, \hat{w})$ and $\Lambda(Z_*, \hat{w})$, and since these two sets coincide, the set $\Lambda(\tilde{Z}, \hat{w})$ also coincides with them.

This implies that the singularity of the trajectory $\hat{w}$ in Problem (Z) is equivalent to its singularity in Problem $(\tilde{Z})$ and in Problem (Z_*). We assume that $\hat{w}$ is singular, and from now on our goal is to obtain sufficient conditions for the presence of the strong minimum in Problem $(\tilde{Z})$ at the given trajectory. Now we can apply the results of [9] to this situation.

5 Application of General Sufficient Conditions to Problem $(\tilde{Z})$

According to [9, Sec. 8], for obtaining sufficient conditions in Problem $(\tilde{Z})$, we have to consider the following Problem (Z_*) with the extended control set U_* in the form of a halfspace:

Problem (Z_*) **.**

$$\dot{x} = z\left(u_0 r_0(x) + \sum_{i=1}^{k-1} u_i r_i(x)\right), \tag{5.1}$$

$$\dot{z} = 0, \qquad x(0) = a, \quad x(T) = b, \tag{5.2}$$

$$u_0 \le 1, \quad \text{components } u_1, \dots, u_{k-1} \text{ are free},$$

$$J = z(0) \longrightarrow \min\,,$$

in which the same singular trajectory $\hat{w}$ is studied. Obviously, this extension does not depend on a particular realization of the set $\tilde{U}$: the halfspace will be one and the same $(u_0 \le 1)$. It turns out that the quadratic sufficient conditions for the weak minimum in Problem (Z_*) yield the strong minimum in Problem $(\tilde{Z})$. More precisely, Theorem 8.1 from [9], being applied to the given situation, yields the following result.

Theorem 5.1. *Let the singular trajectory $\hat{w}$ in Problem (Z_*) satisfy the γ-sufficient condition for the weak minimum. Then $\hat{w}$ is a point of the strict strong minimum in Problem $(\tilde{Z})$. Moreover, there exist $\varepsilon, C > 0$ such that, on the set $|z-1| + \|x - \hat{x}\|_\infty < \varepsilon$, for any $w = (z, x, u)$ satisfying equations (5.1), $\dot{z} = 0$ and constraints $x(0) = a$, $u \in \tilde{U}$, the following estimate holds:*

$$(z-1)^+ + |x(T) - b| \;\ge\; C\gamma(w - \hat{w}). \tag{5.3}$$

Now it is possible, using the formulas from [9, Secs. 4, 5] and [10], to write the γ-sufficient condition for Problem (Z_*), and thereby to obtain the sufficient condition for a strong minimum in Problem $(\tilde{Z})$. But we will proceed in another way. We will show that one can pass from the γ-sufficient condition in Problem (Z_*) to the γ-sufficient condition in another, simpler problem (which differs mainly by the fact that the inequality-type constraint $u_0 \le 1$ is replaced by the equality-type constraint $u_0 = 1$), and then to some system that corresponds to the study of the trajectory $\hat{w}$ for rigidity. In so doing, the weak γ-sufficiency in Problem (Z_*) will turn out to be exactly equivalent to the γ-sufficient condition for the rigidity. The efficient instrument for carrying out these passages will be Theorem 5.2 from [9] on a "nonvariational" equivalent of the γ-sufficient conditions.

First, let us perform a rather simple passage from the γ-conditions in Problem (Z_*) to the γ-conditions in a new Problem (S), which differs from Problem (Z_*) only by the fact that the multiplier z stands only before the very first term in the control system:

$$\dot{x} = z\, u_0\, r_0(x) + \sum_{i=1}^{k-1} u_i\, r_i(x)\,. \tag{5.4}$$

Theorem 5.2. *γ-sufficient condition for the weak minimum in Problem (Z_*) is equivalent to the γ-sufficient condition for the weak minimum in Problem (S).*

In view of Theorem 5.2 (a) from [9], this statement stems from the following property that is intuitively obvious. Recall that for Problem (Z_*) the *violation function* is defined as

$$\sigma(w) = \int_0^T \left| \dot{x} - z\left(u_0 r_0(x) + \sum_{i=1}^{k-1} u_i r_i(x)\right)\right| dt \;+\; \int_0^T |\dot{z}|\, dt \;+ \tag{5.5}$$

$$+\, |x(0)-a| + |x(T)-b| \;+\; \operatorname{vrai\,max}_t\,(u_0-1)^+ \;+\; (z-1)^+\,,$$

and the violation function for Problem (S) is defined similarly.

Lemma 5.1. *The fulfillment of the inequality $\sigma(w) \ge C\gamma(w)$ for some $C>0$ in some neighborhood of the point $\hat{w}$ for Problem (Z_*) is equivalent to its fulfillment for some $C>0$ in some neighborhood of the point $\hat{w}$ for Problem (S).*

Proof. This follows from the simple fact that, for z close to $\hat{z}=1$, the mapping $A:\ W \to W$, defined as

$$w = (z,\, x,\, u_0,\, u_i,\; i=1,\dots,k-1) \longmapsto \tilde{w} = (z,\, x,\, u_0,\, \tilde{u}_i = \frac{u_i}{z},\; i=1,\dots,k-1),$$

is a homeomorphism, and the violation functions in these two problems are connected by the relation $\sigma_*(w) = \sigma_S(Aw)$. For both problems,

$$\gamma(w) = |z-1|^2 + \bar{y}_0^2(T) + \sum_{i=1}^{k-1} y_i^2(T) + \int_0^T \left(\bar{y}_0^2 + \sum_{i=1}^{k-1} y_i^2\right) dt, \tag{5.6}$$

where $\bar{y}_0(t) = y_0(t) - t$,

$$\dot{y}_i = u_i, \qquad y_i(0) = 0 \qquad \forall i = 0,1,\dots,k-1, \tag{5.7}$$

and under the mapping A, only the components y_i, $i \ge 1$, will change; therefore, for $|z-1| \le \frac{1}{2}$, i.e., for $\frac{1}{2} \le z \le \frac{3}{2}$, we have

$$\frac{4}{9}\gamma(w) \;\le\; \gamma(Aw) \;\le\; 4\gamma(w).$$

Thus, if $\sigma_*(w) \ge C\gamma(w)$ in some neighborhood of $\hat{w}$, then on that very set $\sigma_S(Aw) \ge \frac{C}{4}\gamma(Aw)$, and, therefore, in some neighborhood of $\hat{w}$ there will be $\sigma_S(\tilde{w}) \ge \frac{C}{4}\gamma(\tilde{w})$. The reverse passage is valid by similar arguments. The lemma is proved.

6 Passage to $u_0 = 1$

Thus we passed to the γ-sufficiency in Problem (S). Let us show now that it is possible to pass from the halfspace $u_0 \le 1$ to the subspace $u_0 = 1$, i.e., to the following

Problem (S_1).

$$\dot{x} = z\, r_0(x) + \sum_{i=1}^{k-1} u_i\, r_i(x), \tag{6.1}$$

$$\dot{z} = 0, \qquad x(0) = a, \quad x(T) = b,$$

$$J = z(0) \longrightarrow \min.$$

Here Eq. (6.1) is obtained from Eq. (5.4) for $u_0 = 1$. Since u_0 disappeared, now we have not k, but $k-1$ controls, and all of them are free. The examined trajectory (we will denote it by the same letter) is $\hat{w} = (\hat{z} = 1,\, \hat{x}(t), \;\; \hat{u}_i = 0,\, i = 1, \dots, k-1)$, and the control $\hat{u} = (0, \dots, 0) \in \mathbb{R}^{k-1}$. Note that in systems where the component u_0 is present, the control $\hat{u} = (1, 0, \dots, 0) \in \mathbb{R}^k$. We hope that it will always be clear from the context which $\hat{u}$ is meant and that no confusion will arise here.

In the passage from Problem (S) to Problem (S_1) the set of admissible trajectories has narrowed (due to the narrowing of the set of admissible controls); therefore, the set $\Lambda(\hat{w})$ of Lagrange tuples can expand, and a priori there is a danger that a "normal" tuple λ with $\alpha_0 > 0$ can appear. But in the given case this does not happen.

Lemma 6.1. *In Problems (S) and (S_1), the set $\Lambda(\hat{w})$ is one and the same.*

Proof. As was already said, the inclusion $\Lambda(\hat{w}) \subset \Lambda_1(\hat{w})$ is obvious. Conversely, if $\lambda \in \Lambda_1(\hat{w})$, i.e., if a given λ ensures that the MP holds in Problem (S_1), then

$$\psi(t)\, r_i(\hat{x}(t)) = 0 \qquad \forall\, i = 1, \dots, k-1, \tag{6.2}$$

and, as in Sec. 3, we easily obtain

$$\psi(t)\, r_0(\hat{x}(t)) = \text{const} = \frac{\alpha_0}{T} \ge 0. \tag{6.3}$$

But then the given λ ensures that the MP holds also in Problem (S), i.e., $\lambda \in \Lambda(\hat{w})$. (Recall that for an arbitrary stationary trajectory in Problems (S) and (S_1), it is not required that $\alpha_0 = 0$.) The lemma is proved.

Theorem 6.1. *The γ-sufficient condition for the weak minimum in Problem (S) is equivalent to the γ-sufficient condition for the weak minimum in Problem (S_1).*

In the proof, we again use the "nonvariational" interpretation of these conditions. Let us note that both system (5.4), $\dot{z}=0$ with the initial condition $x(0)=a$, and the system 6.1, $\dot{z}=0$ with the initial condition $x(0)=a$ obviously satisfy the Lyusternik condition at $\hat{w}$, and let us introduce the following two sets:

$$\mathcal{D}(S)=\{w \mid \text{ Eq. (5.4) is satisfied, } \dot{z}=0,\ x(0)=a,\ u_0 \le 1\},$$

$$\mathcal{D}(S_1)=\{w \mid \text{ Eq. (5.4) is satisfied, } \dot{z}=0,\ x(0)=a,\ u_0 = 1\}.$$

Taking into account Theorem 5.2 (b) from [9], the proof of Theorem 6.1 is a result of the following

Lemma 6.2. *The validity of the inequality $\sigma(w)\ge C\gamma(w)$ for some $C>0$ on the set $\mathcal{D}(S)$ in some L_∞-neighborhood of the point $\hat{w}$ for Problem (S) is equivalent to the validity of this inequality for some $C>0$ on the set $\mathcal{D}(S_1)$ in some L_∞-neighborhood of the point $\hat{w}$ for Problem (S_1).*

(Let us explain here that for a point w satisfying Eq. (5.1), its closeness to $\hat{w}$ in the norm of the space W is equivalent to its closeness to $\hat{w}$ in the uniform $(L_\infty-)$ norm $|z-\hat{z}|+\|x-\hat{x}\|_\infty+\|u-\hat{u}\|_\infty$. Hence we formulate Lemma 6.2 in terms of L_∞-norm.)

Proof. Recall that the order γ for Problem (S) is given by formulas (5.6), (5.7), and then, for Problem (S_1), it is given by these formulas for $\bar{y}_0(t)\equiv 0$. Let us introduce the following notations: $\bar{z}=z-1$,

$$\eta(y)=\sum_{i=1}^{k-1} y_i^2(T)+\int_0^T \sum_{i=1}^{k-1} y_i^2\,dt\,. \tag{6.4}$$

Note that on $\mathcal{D}(S)$, $\dot{\bar{y}}_0=u_0-1\le 0$, hence $\bar{y}_0(t)$ is monotone, and since $\bar{y}_0(0)=0$, we have $\max|\bar{y}_0(t)|\le|\bar{y}_0(T)|$. Therefore, one may not include $\int_0^T \bar{y}_0^2\,dt$ into γ, since it is estimated through $\bar{y}_0^2(T)$. Thus, on the set $\mathcal{D}(S)$, one can consider

$$\gamma(w)\ =|\bar{z}|^2+|\bar{y}_0(T)|^2+\eta(y), \tag{6.5}$$

and on the set $\mathcal{D}(S_1)$ this value turns into

$$\gamma_1(w)\ =|\bar{z}|^2+\eta(y). \tag{6.6}$$

Since $\mathcal{D}(S)\supset\mathcal{D}(S_1)$, it is sufficient to prove the implication $(\Longleftarrow)$ in the statement of the lemma.

Suppose that, for some $\varepsilon>0$, the following estimate holds on the set $B_\varepsilon(\hat{w})\cap \mathcal{D}(S_1)$:

$$\sigma(w)=(z-1)^+ +|x(T)-b|\ge C\gamma(w-\hat{w}). \tag{6.7}$$

Let us show that then the same estimate holds also on the set $B_\delta(\hat{w}) \cap \mathcal{D}(S)$ for some $\delta > 0$. Let $\delta > 0$ be arbitrary for now. Take any point $w \in B_\delta(\hat{w}) \cap \mathcal{D}(S)$. By definition, this point satisfies Eq. (5.1) with $u_0 \le 1$, and we have to pass to Eq. (6.1), i.e., to equation (5.1) with $u_0 = 1$, in order to obtain the estimate required.

To this end, bearing in mind the remark of Sec. 3, we reparametrize the trajectory w in such a way that $u_0 = 1$. Namely, assuming that $\delta \le \frac{1}{2}$, we have $1 - \delta \le u_0(t) \le 1$, and let us consider the Lipschitzian mapping $s: [0,T] \to [0,T]$ defined by

$$s(0) = 0, \qquad \dot{s}(t) = \frac{u_0(t)}{\theta}, \tag{6.8}$$

where $\theta = (1/T)\int_0^T u_0(t)\,dt$ is the mean value of the function $u_0(t)$ on the closed interval $[0,T]$. Then $s(T) = (1/\theta)\int_0^T u_0(t)\,dt = T$, and

$$1 - \delta \le \theta \le 1, \qquad 1 - \delta \le \dot{s}(t) \le \frac{1}{1-\delta}, \tag{6.9}$$

where $1 - \delta \ge \frac{1}{2}$; therefore, $s(t)$ is a bi-Lipschitzian mapping of the interval $[0,T]$ onto itself. Denote by $t(s)$ the inverse mapping and introduce the functions

$$x'(s) = x(t(s)), \qquad u_i'(s) = \theta \frac{u_i(t(s))}{u_0(t(s))}, \qquad i = 1, \ldots, k-1. \tag{6.10}$$

It is easy to verify that

$$\frac{dx'(s)}{ds} = \frac{dx}{dt} \Big/ \frac{ds}{dt} = (\theta z)\, r_0(x'(s)) + \sum_{i=1}^{k-1} u_i'(s)\, r_i(x'(s)). \tag{6.11}$$

Taking $z' = \theta z$, we obtain the trajectory $w' = (z', x', u_i', i = 1, \ldots, k-1)$ satisfying Eq. (6.1), and thereby $w' \in \mathcal{D}(S_1)$. Note that $x'(T) = x(T)$, i.e., the endpoints of the new trajectory and those of the old one coincide.

First let us show that the trajectory w' is close to $\hat{w}$. The trajectory $\hat{w}$ satisfies system (6.1) with $\hat{z} = 1$, $\hat{u}_i = 0$, $i \ge 1$ and with the initial condition $\hat{x}(0) = a$, the trajectory w' satisfies system (6.11) with u_i' and with the same initial condition $x'(0) = a$, therefore, the difference $x' - \hat{x}$ satisfies the equation

$$\frac{d}{dt}(x' - \hat{x}) = z'(r_0(x') - r_0(\hat{x})) + (z' - 1)\, r_0(x) + \sum_{i=1}^{k-1} u_i'\, r_i(x'). \tag{6.12}$$

From (6.9), it follows that $|z' - z| = (1-\theta)|z| \le \delta|z| \le \delta(1+\delta)$, and since $|z - \hat{z}| \le \delta$, we have $|z' - \hat{z}| \le \delta(2+\delta) < 3\delta$.

By definition $\|x - \hat{x}\|_\infty \le \delta$; therefore, the modules of the vectors $r_i(x'(t))$, $i = 0, 1, \ldots, k-1$, are bounded by a common constant (which does not depend on the specific values of $x'(t)$); moreover,

$$|\, r_0(x'(t)) - r_0(\hat{x}(t))\,| \le \text{const}\ |x'(t) - \hat{x}(t)|.$$

Since $\|u_i\|_\infty \le \delta$, it follows from (6.10) that $\|u_i'\|_\infty \le \frac{\delta}{1-\delta}$, and then from (6.12) and from the known Grownwall-type estimates, it follows that $\|x' - \hat{x}\|_C \le f(\delta)$, where $f(\delta) \to 0$ as $\delta \to 0$. Thus, $\|w' - \hat{w}\|_\infty \le 3\delta + (k-1)\frac{\delta}{1-\delta} + f(\delta) = f_1(\delta) \to 0$ as $\delta \to 0$, i.e., the closeness of w' to $\hat{w}$ is proved.

For a sufficiently small $\delta > 0$ we have $\|w' - \hat{w}\|_\infty < \varepsilon$ and $w' \in \mathcal{D}(S_1)$; then, by the premise of the lemma, estimate (6.7) holds for w', i.e., the following is true:

$$(z'-1)^+ + |x(T) - b| \;\ge\; C\gamma_1(w') \;= C\left(|\bar{z}'|^2 + \eta(y')\right), \tag{6.13}$$

where

$$\dot{y}_i' = u_i', \qquad y_i'(0) = 0, \qquad i = 1, \ldots, k-1. \tag{6.14}$$

We have to show that estimate (6.7) is valid for the trajectory w, i.e., that for some $\delta > 0$ and for some $C > 0$ not depending on $w \in B_\delta(\hat{w}) \cap \mathcal{D}(S)$, the following holds:

$$(z-1)^+ + |x(T) - b| \;\ge\; C\gamma(w) = C\left(|\bar{z}|^2 + |\bar{y}_0(T)|^2 + \eta(y)\right). \tag{6.15}$$

We prove this by contradiction. Assume that for all δ and $C > 0$, there is no such estimate on $B_\delta(\hat{w}) \cap \mathcal{D}(S)$, i.e., that there exists a sequence $w_n = (z_n, x_n, u_n) \in \mathcal{D}(S)$ such that $w_n \ne \hat{w}$, $\|w_n - \hat{w}\| \to 0$, and (omitting the subscript n)

$$(z-1)^+ + |x(T) - b| \;\le\; o\left(|\bar{z}|^2 + |\bar{y}_0(T)|^2 + \eta(y)\right). \tag{6.16}$$

Let us then show, that the corresponding sequence $w_n' \in \mathcal{D}(S_1)$, constructed as in above, will also violate estimate (6.13), i.e., that the following inequality will be held (we again omit n) :

$$(z'-1)^+ + |x(T) - b| \le o\left(|\bar{z}'|^2 + \eta(y')\right). \tag{6.17}$$

Recall that η is given by formula (6.4). Let us first prove that $\eta(y) \simeq \eta(y')$ (have the same order, i.e., estimate one by another with some constants). Indeed, according to (6.14), (6.10), and (6.8), for each $i \ge 1$ and for any $\tau \in [0, T]$, the following holds:

$$y_i'(\tau) = \int_0^\tau u_i'(s)\, ds = \int_0^\tau \theta \frac{u_i(t(s))}{u_0(t(s))}\, ds = \int_0^{t(\tau)} u_i(t)\, dt = y_i(t(\tau)),$$

(in particular, $y_i'(T) = y_i(T)$); therefore,

$$\int_0^T |y_i'(s)|^2 ds \;=\; \int_0^T |y_i(t(s))|^2 ds \;=\; \int_0^T |y_i(t)|^2\, \dot{s}(t)\, dt,$$

from which, due to (6.9), we obtain the estimate

$$(1-\delta)\int_0^T |y_i(t)|^2 dt \;\le\; \int_0^T |y_i'(s)|^2 ds \;\le\; \frac{1}{1-\delta}\int_0^T |y_i(t)|^2 dt$$

and hence, the same estimate is valid for η :

$$(1-\delta)\,\eta(y) \le \;\eta(y') = \int_0^T \sum_{i=1}^{k-1} |y_i'(s)|^2 ds \,+ \sum_{i=1}^{k-1} |y_i'(T)|^2 \;\le \frac{1}{1-\delta}\,\eta(y), \tag{6.18}$$

i.e., $\eta(y) \simeq \eta(y')$ in realty.

Further, we put $p = (1/T)\int_0^T (1-u_0)\,dt = -(1/T)\bar{y}_0(T)$. Then $p = 1-\theta$, and, according to (6.9), $0 \le p \le \delta$. Moreover,

$$\bar{z}' = z' - 1 = \theta z - 1 = \theta(1+\bar{z}) - 1 = \theta\bar{z} + (\theta - 1) = \theta\bar{z} - p, \tag{6.19}$$

from which, taking into account that $\theta \le 1$, we obtain the estimate

$$|\bar{z}'|^2 \le \; |\bar{z}|^2 \,+\, \frac{1}{T^2}\,|\bar{y}_0(T)|^2.$$

On the other hand, (6.19) implies $\theta\bar{z} = \bar{z}' + p$, and, taking into account that $\theta \ge \frac{1}{2}$, we obtain

$$|\bar{z}|^2 \le 4\left(|\bar{z}'|^2 \,+\, \frac{1}{T^2}\,|\bar{y}_0(T)|^2\right).$$

Due to this estimate and (6.18), from (6.16) it follows that

$$(z-1)^+ + |x(T) - b| \;\le\; o\left(|\bar{z}'|^2 + p^2 + \eta(y')\right). \tag{6.20}$$

If, at the same time, $p^2 \le \mathcal{O}(|\bar{z}'|^2 + \eta(y'))$, then from here and from the inequality $z' = \theta z \le z$, (6.17) obviously follows, and we arrive at the desired contradiction.

It remains to consider the case where, on the contrary,

$$p^2 \gg |\bar{z}'|^2 + \eta(y'), \qquad \text{i.e.,} \qquad |\bar{z}'|^2 + \eta(y') \;\le\; o(p^2). \tag{6.21}$$

Then the leading term in the right-hand side of (6.20) is p^2, i.e., the following holds:

$$(\bar{z})^+ + |x(T) - b| \;\le\; o(p^2). \tag{6.22}$$

Let us show that in this case we also come to a contradiction, but now without using of (6.17). From (6.19), we have $\theta\bar{z} = \bar{z}' + p$, and then from (6.22), it follows that $(\bar{z}' + p)^+ \le o(p^2)$. But, by virtue of (6.21), $\bar{z}' = o(p)$, and since $p \ge 0$, this implies (we write the subscript n again) that $p_n = 0$ for a sufficiently large n, and also that $\bar{z}_n' = 0$. But then, by virtue of (6.19), $\bar{z}_n = 0$, i.e., $z_n = 1$, and, by virtue of (6.21), $\eta(y_n') = 0$, i.e., $u_n'(t) \equiv 0$, and hence, $u_n(t) \equiv 0$. In addition, from $p_n = 0$ it follows that $\theta_n = 1$, $u_{0,n}(t) \equiv 1$, and then $x_n = \hat{x}$ as well. Thus, for sufficient large n the trajectory w_n coincides with $\hat{w}$, which contradicts our assumption.

Lemma 6.2 is proved, and therefore Theorem 6.1 is also proved.

7 Quadratic Conditions in Problem (S_1) and in System (R)

We have thus passed to the γ-sufficient conditions for the weak minimum in Problem (S_1). Let us now write these conditions.

Since our $\hat{w}$ is singular, i.e., $\forall \lambda \in \Lambda(S_1, \hat{w})$, the multiplier by the functional $\alpha_0 = 0$, the functional J does not enter the Lagrange function and hence, into its second variation.

Further, since there are no inequality-type constraints in Problem (S_1), in the notation of the cone of critical variations, one can ignore the single inequality that corresponds to the functional J (since the nonnegativity of the even function — which is homogeneous of second degree in our case — on the halfspace is equivalent to its nonnegativity on the whole space), therefore, the functional J will not enter the formulation of the γ-conditions at all.

Here the set $\Lambda(\hat{w})$ consists of all collections $\lambda = (\psi = \psi_x, \;\; \psi_z = 0, \;\; \alpha_0 = 0, \;\; \beta_0 = \psi(0), \;\; \beta_T = -\psi(T))$ for which the following is satisfied:

(a) the adjoint equation

$$\dot{\psi} = -\psi\, r_0'(\hat{x}), \tag{7.1}$$

(b) the relations

$$\psi(t)\, r_i(\hat{x}(t)) = 0 \qquad \forall\, i = 0, 1, \ldots, k-1, \tag{7.2}$$

(c) the normalization

$$|\psi(0)| = 1. \tag{7.3}$$

The choice of a particular normalization is not essential here, since $\Lambda(\hat{w})$ is finite-dimensional compact set. Since the whole tuple here is determined by the function $\psi(t)$, one can consider, instead of $\Lambda(\hat{w})$, the set $\Psi(\hat{w})$, consisting of the corresponding functions $\psi(t)$. Since the trajectory $\hat{w}$ is singular, as was already noted, $\Psi(\hat{w}) = \Psi_0(\hat{w})$.

The cone of critical variations $\mathcal{K}$ (here it is a subspace) is given by the following equalities:

$$\dot{\bar{x}} = \bar{z} r_0(\hat{x}) + r_0'(\hat{x})\bar{x} + \sum_{i=1}^{k-1} \bar{u}_i r_i(\hat{x}), \tag{7.4}$$

$$\dot{\bar{z}} = 0, \qquad \bar{x}(0) = \bar{x}(T) = 0. \tag{7.5}$$

Passing to the Goh variables, i.e., setting

$$\dot{\bar{y}}_i = \bar{u}_i, \qquad \bar{y}_i(0) = 0, \qquad i = 1, \ldots, k-1, \tag{7.6}$$

$$\bar{x} = \bar{\xi} + \sum_{j=1}^{k-1} \bar{y}_j r_j(\bar{x}), \tag{7.7}$$

we obtain that $\mathcal{K}$ is given by Eqs. (7.6) and by

$$\dot{\bar{\xi}} = \bar{z}r_0 + r_0'\bar{\xi} + \sum_{j=1}^{k-1} \bar{y}_j[r_0, r_j], \qquad \bar{\xi}(0) = 0, \tag{7.8}$$

$$\bar{\xi}(T) + \sum \bar{y}_j(T) r_j(\hat{x}(T)) = 0. \tag{7.9}$$

(Here $[f(x), g(x)] = f'(x)g(x) - g'(x)f(x)$ is the Lie bracket of two vector fields.)

For each ψ we define the Lagrange function

$$\Phi[\psi](w) = \psi_0(x_0 - a) - \psi_T(x_T - b) + \int_0^T \psi\left(\dot{x} - zr_0(x) - \sum u_i r_i(x)\right) dt, \tag{7.10}$$

and consider its second variation at the trajectory $\hat{w}$; this variation can be brought to the following form (see Appendix C):

$$\begin{aligned}\Omega\,[\psi](\bar{z}, \bar{\xi}, \bar{y}, \bar{u}) \;=\; & \frac{1}{2}\left.\left\{ \sum_{i,j=1}^{k-1} \bar{y}_i\,\bar{y}_j\,\psi\,(r_i'(\hat{x})\,r_j(\hat{x})) \right\}\right|_T + \\ + \int_0^T & \left(-\frac{1}{2}\,\psi\,(r_0''\,\bar{\xi}, \bar{\xi}) \;-\; \bar{z}\,\psi\,(r_0'\,\bar{\xi}) \;+\; \sum_{i=1}^{k-1} \bar{y}_i\,\psi\,[r_i, r_0]'\,\bar{\xi} + \right. \\ & \left. + \frac{1}{2}\sum_{i,j=1}^{k-1} \bar{y}_i\,\bar{y}_j\,\psi\,[\,[r_i, r_0], r_j\,] \;+\; \frac{1}{2}\sum_{i,j=1}^{k-1} \bar{y}_i\,\bar{u}_j\,\psi\,[r_i, r_j] \right) dt.\end{aligned} \tag{7.11}$$

Note that in the last term here, the matrix of coefficients is skew-symmetrical. These formulas has been already presented in [8, 15, 10].

For an arbitrary set $M = \{\psi\}$ define the functional

$$\Omega\,[M](\bar{w}) \;=\; \sup_{\psi \in M} \Omega\,[\psi](\bar{w}).$$

According to the general theory [9], we must isolate from the set Ψ subsets $G_a(\Psi)$ consisting of those $\psi \in \Psi$ for which the quadratic form (7.11) satisfies the Goh conditions. To this end, let us find the corresponding coefficients of the quadratic form.

Recall that the quadratic form of general form with the zero Legendre term is represented in the Goh variables as follows:

$$\begin{aligned}\Omega[\lambda](\bar{\xi}, \bar{y}, \bar{u}) = \; & g[\lambda](\bar{\xi}_0, \bar{\xi}_T, \bar{y}_T) + \\ + \int_0^T & \Big((D[\lambda]\bar{\xi}, \bar{\xi}) + (P[\lambda]\bar{\xi}, \bar{y}) + (Q[\lambda]\bar{y}, \bar{y}) + (V[\lambda]\bar{y}, \bar{u}) \Big)\, dt,\end{aligned} \tag{7.12}$$

$$\dot{\bar{\xi}} = A(t)\bar{\xi} + B_1(t)\bar{y}, \qquad \dot{\bar{y}} = \bar{u}, \quad \bar{y}(0) = 0, \tag{7.13}$$

where $g[\lambda]$ is a quadratic form of the endpoint values, all the matrices in this expression are measurable and bounded, among them $V[\lambda](t)$ and $B_1(t)$ are Lipschitzian, and moreover, $V[\lambda](t)$ is skew-symmetric and $Q[\lambda](t)$ is symmetric (see [5, 9]).

For an arbitrary set $M = \{\lambda\}$ and for any number $a \in \mathbb{R}$, the set $G_a(M)$ by definition consists of all $\lambda \in M$, for which

$$(1) \quad V[\lambda](t) \equiv 0, \qquad (2) \quad Q[\lambda](t) \geq a \quad \text{almost everywhere.} \tag{7.14}$$

(The inequality $Q \geq a$ for a symmetrical $m \times m$ - matrix Q means that $\forall h \in \mathbb{R}^m \quad (Qh, h) \geq a(h, h)$.)

Let us return to the quadratic form (7.11). Here the variable $\bar{z}$ plays the role of an additional component $\bar{\xi}_0$, and the last two terms in (7.12) have the following form:

$$(V[\lambda](t)\,\bar{y}, \bar{u}) = \frac{1}{2} \sum_{i,j=1}^{k-1} \bar{y}_i\, \bar{u}_j\, \psi\,[r_i, r_j]\,(\hat{x}(t))\,, \tag{7.15}$$

$$(Q[\lambda](t)\,\bar{y}, \bar{y}) = \frac{1}{2} \sum_{i,j=1}^{k-1} \bar{y}_i\, \bar{y}_j\, \psi\,[\,[r_i, r_0], r_j\,]\,(\hat{x}(t))\,. \tag{7.16}$$

Therefore, the set $G_a(\Psi)$ consists of all $\psi \in \Psi$ for which the following conditions hold on the interval $[0, T]$:

$$\psi(t)\,[r_i, r_j](\hat{x}(t)) \;=\; 0 \qquad \forall\, i, j = 1, \ldots, k-1, \tag{7.17}$$

$$\frac{1}{2} \sum_{i,j=1}^{k-1} \bar{y}_i\, \bar{y}_j\, \psi[\,[r_i, r_0], r_j\,]\,(\hat{x}(t)) \;\geq\; a\,|\,\bar{y}\,|^2 \qquad \forall\, \bar{y} = (\bar{y}_1, \ldots, \bar{y}_{k-1}) \in \mathbb{R}^{k-1}. \tag{7.18}$$

Note that, by virtue of (7.17), the last term in (7.11) vanishes and then the control $\bar{u}$ does not explicitely enter Ω. Then, one can take $\bar{y}(t)$ as a new control, and (7.18) is nothing else but the classical Legendre condition with respect to this new control. In the term outside the integral (7.11), the vector $\bar{y}(T)$ can be replaced by an arbitrary vector $h \in \mathbb{R}^{k-1}$ (see [5, 7, 9]), but we will not use this fact here.

Now let us write the quadratic order γ. According to the general theory [5, 7, 9], we must a priori take

$$\gamma(\bar{w}) = \; |\bar{z}|^2 + |\bar{x}(0)|^2 + |\bar{y}(T)|^2 + \int_0^T |\bar{y}|^2\, dt.$$

But since we have to consider not arbitrary $\bar{w}$, but only $\bar{w} \in \mathcal{K}$, and since on this subspace, $\bar{x}(0) = 0$ and, as is shown in [9, Sec. 6.2], there is the estimate

$$|\bar{z}| + |\bar{y}(T)| \,\leq\, \text{const}\ \|\bar{y}\|_1\,,$$

one can leave only the integral term in γ, i.e., the following holds:

$$\gamma(\bar{w}) \simeq \gamma'(\bar{w}) = \int_0^T |\bar{y}|^2\, dt. \tag{7.19}$$

Thus, we have found the set $\Psi = \Psi(\hat{w})$ (in our case, it coincides with $\Psi_0(\hat{w})$), its subsets $G_a(\Psi)$, the family of the quadratic forms $\Omega[\psi](\bar{w})$, and the comparison functional $\gamma' \simeq \gamma$. The quadratic γ-sufficient condition of the weak minimum for the point $\hat{w}$ in Problem (S_1) consists in the fact that, for some $a > 0$,

$$\Omega[\Psi](\bar{w}) \geq a \int_0^T |\bar{y}|^2\, dt \qquad \forall\, \bar{w} \in \mathcal{K}. \tag{7.20}$$

According to [4, 5, 7], this condition is equivalent to the fact that, for the same $a > 0$,

$$\Omega\,[G_a(\Psi)](\bar{w}) \;\geq\; a \int_0^T |\bar{y}|^2\, dt \qquad \forall\, \bar{w} \in \mathcal{K}. \tag{7.21}$$

(The latter inequality automatically implies that $G_a(\Psi)$ is not empty, otherwise, by definition, $\Omega = -\infty$ as supremum over $\emptyset$.)

From here, taking into account Theorems 6.1, 5.2, and 5.1, we obtain the final form of the quadratic sufficient conditions for the strong minimum for a singular trajectory $\hat{w}$ in Problem (Z).

Theorem 7.1. *Let $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = (1, 0, \ldots, 0))$ be a singular trajectory of Problem (Z), and the inequality (7.20) or (7.21) hold for this trajectory for some $a > 0$. Then $\hat{w}$ is a point of the strict strong minimum in Problem $(\tilde{Z})$, and, therefore, it is a point of strict strong minimum in Problem (Z).*

Recall that by virtue of Corollaries 1, 2 of Lemma 3.1, the trajectory $\hat{w}$ is singular in Problem (Z) if and only if it is singular in Problem (Z_*) (with the constraint $u_0 \leq 1$.) The set $\Lambda(\hat{w})$ for these problems is one and the same. In this sense, problems (Z) are the same for all submetrics $\varphi(x,u)$ having one and the same support hyperplane in a neighborhood of $\hat{x}(t)$. One can introduce the following definition.

Definition 7.1. We say that the submetrics q_1 and q_2 satisfying Assumption A2 are *equivalent* in a neighborhood of the curve $\hat{x}(t)$, if in a neighborhood of the set $\hat{\chi}$ the vector $r_0(x)$ and the subspace $\Gamma_0(x)$, taking part in the definition of the support hyperplane, are common for these submetrics.

In other words, the submetrics q_1 and q_2 are equivalent in a neighborhood of the curve $\hat{x}(t)$, if in a neighborhood of the set $\hat{\chi}$ there exists an associated support basis that is common for these submetrics. It is easy to see that this property holds irrespective of the parametrization of the curve $\hat{x}$.

For all submetrics equivalent to the given one, in any associated support basis, the tangent halfspace to the hodograph at the point $r_0(x)$, for x from some neighborhood of the set $\hat{\chi}$, is given by the inequality $u_0 \le 1$. Let us introduce also the following definition.

Definition 7.2. An equivalence class, i.e., the set of all submetrics, every two of which are equivalent to each other in a neighborhood of the curve $\hat{x}(t)$, will be called a *sheaf of submetrics* equivalent in a neighborhood of $\hat{x}(t)$.

A sheaf of submetrics that are equivalent in a neighborhood of $\hat{x}(t)$ is completely determined by specifying an associated basis.

Definition 7.3. The set of all submetrics from a given sheaf that have the subspace $\Gamma_0(x)$ as a strict support hyperplane will be called a *strict sheaf of submetrics,* equivalent in a neighborhood of $\hat{x}(t)$.

In these terms, we obtain the following strengthening of Theorem 7.1.

Theorem 7.2. *Let $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = (1, 0, \ldots, 0))$ be a singular trajectory of Problem (Z_*), written in some associated basis, and let the inequality (7.20) or (7.21) be satisfied for this trajectory for some $a > 0$. Then $\hat{w}$ is a point of the strict strong minimum in Problem (Z) with any submetric from the strict sheaf corresponding to this basis, i.e., a submetric for which this basis is an associated one ($r_0(\hat{x}(t)) = \dot{\hat{x}}(t)$), and for all x from some neighborhood of $\hat{\chi}$, the hodograph $U(x)$ is contained in the halfspace $u_0 \le 1$ intersecting with the subspace $u_0 = 1$ at the single point $\hat{u}$.*

This is precisely the result of the "direct" application of the general quadratic conditions for the strong minimum from [9] to Problem (Z). A shortcoming of this result is the presence of the assumption on the singularity of $\hat{w}$ in it. Below, in Secs. 8–10, we will show that one can omit this assumption.

Now note that conditions (7.20) and (7.21) exactly coincide with the quadratic sufficient conditions obtained recently by A. A. Milyutin [15] for the so-called rigidity of the trajectory $\hat{w}$ for the following system.

System (R) **.**

$$\dot{x} = z\, r_0(x) + \sum_{i=1}^{k-1} u_i\, r_i(x),$$

$$\dot{z} = 0, \qquad x(0) = a, \quad x(T) = b.$$

Recall that smooth Γ-admissible curve $\hat{x}(t)$, connecting two points a and b, is called *rigid* [3], if in some its neighborhood with respect to the norm $\|x\|_C + \|\dot{x}\|_\infty$, any other Γ-admissible curve, connecting the same points is a reparametrization of the curve $\hat{x}(t)$.

This property is equivalent to the fact that (see [15]) in any associated basis, the trajectory $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = 0)$ is isolated among all trajectories of System (R) on the fixed interval $[0, T]$ with respect to the norm $\|w\|_\infty = |z| + \|x\|_C + \|u\|_\infty$.

Definition 7.4. Following [15], the trajectory $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = 0)$ of System (R) is called quadratically rigid if, for some $a > 0$, inequality (7.20) or (7.21) holds, where $\Psi = \Psi_0 = \Psi_0(\hat{w})$ is specified, as for Problem (S_1), by relations (7.1)–(7.3).

In [15] it is proved that any quadratically rigid trajectory is rigid. (The latter definition is motivated by this very fact.) In the same paper it is also shown that property of the *quadratic rigidity does not depend on the choice of associated basis* (in particular, conditions (7.17), (7.18) hold independently on this choice), so one can speak about the quadratic rigidity of the curve $\hat{x}(t)$. (A Γ-admissible curve $\hat{x}(t)$ is quadratically rigid, if the corresponding trajectory $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = 0)$ of System (R) is quadratically rigid in some, and therefore, in any associated basis.) In addition, it is shown that this property does not depend on the choice of parametrization of the curve $\hat{x}$.

Using the above terminology, we obtain the following reformulation of Theorem 7.2, which relates the concepts of rigidity and strong minimality.

Theorem 7.3. *Let a singular trajectory $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = (1, 0, \ldots, 0))$ of Problem (Z_*) written in an associated basis be quadratically rigid. Then it is a point of the strict strong minimum in Problem (Z) with any submetric from the strict sheaf that is determined by this basis.*

To prove this theorem, we have covered the following path:

$$\text{Problem } (Z) \to \text{Problem } (\tilde{Z}) \to \text{Problem } (Z_*) \to$$

$$\to \text{Problem } (S) \to \text{Problem } (S_1) \to \text{System } (R).$$

At this point we finish the procedure of the direct application of the general quadratic sufficient conditions for singular trajectories to the problem on geodesics, and pass to the elimination of the assumption on the singularity of $\hat{w}$.

Part II

Sufficient Conditions for Strong Minimum for Arbitrary Quadratically Rigid Trajectories

8 Description of the Situation

Our goal in this part of the paper is to eliminate the assumption of the singularity in Theorem 7.4, i.e., to prove the Main Theorem 1 that was announced in Introduction. Now let us formulate this theorem in the above notions and notation.

Theorem 8.1 (Main Theorem 1). *Let a Γ-admissible curve $\hat{x}(t)$ connecting the points a and b be quadratically rigid, i.e., let the inequality (7.20) or (7.21) for some $a>0$ hold for the trajectory $\hat{w}=(\hat{z}=1, \hat{x}(t), \hat{u}=0)$ of System (R) in some associated basis. Then, in any submetric, not necessarily related to this basis, that has a strict support hyperplane in a neighborhood of $\hat{x}(t)$, the curve $\hat{x}(t)$ yields the strict strong minimum of distance between the points a and b, i.e., in Problem (Z) written in any associated basis for $\hat{x}(t)$, that is strictly support one for the given submetric, the trajectory $\hat{w}=(\hat{z}=1, \hat{x}(t), \hat{u}=(1,0,\ldots,0))$ is a point of the strict strong minimum.*

In order to prove this theorem, we try to cover the same path as in Part I, but in the reverse order: from System (R) to Problem (Z). But there is a rather serious trouble waiting for us on this path: in the passage from System (R) to Problem (S_1), there can appear $\lambda \in \Lambda(S_1, \hat{w})$ for which

$$\psi(t)\, r_0(\hat{x}(t)) = \frac{\alpha_0}{T} > 0\,, \tag{8.1}$$

and then, in the further passage to Problem (S), the trajectory $\hat{w}$ becomes nonsingular, since it has λ with $\alpha_0 > 0$. (In Problem (S_1) itself, the singularity is still retained, since the control $u=(u_1,\ldots,u_{k-1})$ is free in this problem; therefore, any stationary trajectory is singular; but, in Problem (S) the control $u_0 \le 1$ is added, and it is just with respect to this control inequality (8.1) violates the singularity: $H_{u_0} = \psi r_0(\hat{x}) > 0$, i.e., the maximum of H over $u_0 \le 1$ is strictly attained.) Previously, in Part I, this was not the case, for we have just assumed that the trajectory $\hat{w}$ was singular in Problem (Z), which is equivalent to its singularity in Problem (S).

For nonsingular trajectories, the fulfilment of quadratic sufficient conditions with our γ is a too weak condition to guarantee in any somewhat general situation the presence of even the weak minimum and, the more so, the Pontryagin or the strong minima. Thus, in the direct passage from System (R) to Problem (S_1), we immediately leave the framework of the theory of quadratic conditions for singular regimes, on which our considerations are based, and, having only estimates (7.20) and (7.21), we cannot say anything more.

In order to overcome this difficulty, we use a method proposed by A. A. Milyutin. First, let us describe the situation more precisely.

Thus, let the curve $\hat{x}(t)$, $t \in [0,T]$, of the distribution $\Gamma(x)$ be such that in an associated basis $r_0(x), \ldots, r_{k-1}(x)$, the trajectory $\hat{w} = (\hat{z} = 1,\ \hat{x}(t),\ \hat{u}_i = 0,$ $i = 1, \ldots, k-1)$ of System (R) corresponding to this curve satisfies the quadratic sufficient conditions (7.20) or (7.21). As was already said, these conditions will then be satisfied in any other associated basis.

Let now in $\Gamma(x)$ an arbitrary submetric $q(x,\dot{x})$ be given, not necessarily related to the above basis, and having a strict support hyperplane $\Gamma_0(x)$ in a neighborhood of the curve $\hat{x}(t)$. Replacing, if necessary, the parametrization of the curve $\hat{x}$, one can assume that $q(\hat{x}(t), \dot{\hat{x}}(t)) = 1$. The time T and the base vector field $r_0(x)$ will correspondingly change under this procedure, but the fulfilment of conditions (7.20), (7.21) with some $a > 0$ for the trajectory $\hat{w}$ in a new associated basis will remain unchanged. Replacing the vector fields $r_1(x), \ldots, r_{k-1}(x)$ by base fields for the hyperplane $\Gamma_0(x)$ (in some neighborhood of the set $\hat{\chi}$), we will have an associated basis in $\Gamma(x)$ that will be strictly support for the given submetric and in which the trajectory $\hat{w}$ will satisfy System (R), and, according to [15], it will still satisfy conditions (7.20) and (7.21) with some $a > 0$. (It is just due to this fact that an *arbitrary* submetrics appears in the statement of Theorem 8.1.)

Let us fix the obtained basis, the closed interval $[0,T]$, and the trajectory $\hat{w}$; all our further considerations will be carried out namely for these objects. In the chosen basis, in some neighborhood of $\hat{\chi}$, the unit ball $U(x)$ of the given submetric is a convex compact set in the space $\mathbb{R}^k$, which is contained in the halfspace $u_0 \le 1$ and intersects the hyperplane $u_0 = 1$ at the unique point $\hat{u} = (1, 0, \ldots, 0)$.

Let us now choose any convex compact set $\tilde{U} \subset \mathbb{R}^k$ containing $U(x)$ for all x from some neighborhood of the set $\hat{\chi}$ and also intersecting with the halfspace $u_0 = 1$ at the unique point $\hat{u} = (1, 0, \ldots, 0)$. (Such compact set has already been used in Sec. 4; its existence is proved in Appendix B.) As before, if we show that $\hat{w}$ is a point of the strict strong minimum in Problem (Z) with the set $\tilde{U}$, then the more so it will be such in Problem (Z) with the set $U(x)$, and thereby Theorem 8.1 will be proved.

Thus, our starting position is as follows: the trajectory $\hat{w}$ is quadratically rigid in System (R), and we want to prove that it is a point of the strict strong minimum in Problem (Z) with the set $\tilde{U}$.

In essence, if one removes all the preliminaries that are not necessary now, we have the following situation. In an open set in $\mathbb{R}^n$, there given an arbitrary basis of a distribution $\Gamma(x)$, i.e., there given k linear independent vector fields $r_0(x), \ldots, r_{k-1}(x)$. In this basis, System (R) is considered, and a trajectory $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = 0)$ of this system on a closed interval $[0, T]$ is given; this trajectory is quadratically rigid, i.e., it satisfies inequalities (7.20) and (7.21) for some $a > 0$. There also given an arbitrary convex compact set $\tilde{U} \in \mathbb{R}^k$ containing in the halfspace $u_0 \leq 1$ and intersecting with the hyperplane $u_0 = 1$ at the unique point $\hat{u} = (1, 0, \ldots, 0)$; Problem (Z) is considered for this compact set on the same interval $[0, T]$. It is required to prove that then the trajectory $\hat{w} = (\hat{z} = 1,\ \ \hat{x}(t),\ \ \hat{u}_0 = 1,\ \ \hat{u}_1 = \cdots = \hat{u}_{k-1} = 0)$ yields the strict strong minimum in this problem.

Let us pass to the proof. As was already said, if we add now the functional $J = z(0) \to \min$ to System (R), then in the obtained Problem (S_1) there can appear nonsingular ψ, i.e., those satisfying (8.1), and we want to avoid this circumstance. The method proposed by A. A. Milyutin consists in the following: first, one should pass from System (R) to some "relaxed" System (R'), and just from this system then pass to the corresponding problems (S) and (S_1) (more precisely, to analogs of these problems, which will be denoted by other letters.)

Consider the first step of this chain of passages.

9 Passage to System (R')

As before, let $\Psi_0 = \Psi_0(\hat{w})$ be the set of all Lipschitzian n-dimensional functions on $[0, T]$, satisfying the following relations:

$$\dot{\psi}(t) = -\psi(t)\, r_0'(\hat{x}(t)), \tag{9.1}$$

$$\psi(t)\, r_i'(\hat{x}(t)) = 0 \qquad \forall\, i = 0, 1, \ldots, k-1. \tag{9.2}$$

In $\mathbb{R}^n$, we consider the linear subspace $M = \{\psi(T) \,|\, \psi \in \Psi_0\}$ and choose (and fix for the sequel) some orthogonal complement L to it. Then $\mathbb{R}^n = M \oplus L$, and any vector $x \in \mathbb{R}^n$ can be represented in the form $x = x_M + x_L$, where $x_M = \pi_M x$, $x_L = \pi_L x$, and π_M, π_L are the corresponding projections on M and L. Assuming that M and L are coordinate planes in $\mathbb{R}^n$, one can write $x = (x_M, x_L)$.

Note a simple property that we will need below.

Proposition 9.1. *If a function $\psi(t)$ satisfies Eq. (9.1) and $\psi(T) \in M$, then $\psi \in \Psi_0$ and relations (9.2) automatically hold for this function.*

Indeed, in this case, $\psi(T) = \tilde{\psi}(T)$ for some $\tilde{\psi} \in \Psi_0$ (by the definition of the set M), and since both functions ψ and $\tilde{\psi}$ satisfy one and the same differential equation (9.1), we have $\psi(t) = \tilde{\psi}(t)$ on the whole closed interval $[0,T]$, i.e., $\psi \in \Psi_0$, and, hence, (9.2) holds.

Denote $\Delta x(T) = x(T) - b$ and consider the following system.

System (R').

$$\dot{x} = z r_0(x) + \sum_{i=1}^{k-1} u_i r_i(x),$$

$$\dot{z} = 0, \quad x(0) = a,$$

$$\Delta x_M(T) = \pi_M(x(T) - b) = 0,$$

$$|\Delta x_L(T)|^2 \le 0.$$

It differs from System (R) by the fact that, instead of the endpoint equality $x(T)-b = 0$, i.e., instead of two equalities $\Delta x_M(T) = 0$ and $\Delta x_L(T) = 0$, there is now only one equality, and the other is written in the form of inequality $|\Delta x_L(T)|^2 \le 0$.

It is clear that the set of admissible trajectories in both systems is one and the same, so nothing was changed at the first glance. But the cone of critical variations $\mathcal{K}(R')$ for System (R') at the trajectory $\hat{w}$ is wider then that cone for System (R). In both cases, it is a subspace, but, if previously it was given by the relations

$$\dot{\bar{x}} = \bar{z}\, r_0(\hat{x}) \;+\; r_0'(\hat{x})\, \bar{x} \;+\; \sum_{i=1}^{k-1} \bar{u}_i\, r_i(\hat{x}),$$

$$\dot{\bar{z}} = 0, \quad \bar{x}(0) = 0, \quad \bar{x}_M(T) = 0, \quad \bar{x}_L(T) = 0,$$

now the last of these relations is absent: $\bar{x}_L(T)$ is free. In addition, as we will show below, the set $\Lambda(\hat{w})$ of Lagrange collections for System (R') is wider than that for System (R). (The set $\Lambda(\hat{w})$ for any system is defined in the same way as for any minimization problem, with the only difference that here there is no the objective functional, and hence there will be no corresponding multiplier α_0. In other words, the set $\Lambda(\hat{w})$ for any system is the set of all those Lagrange collections for the minimization problem with the given system and with an arbitrary functional, whose coefficient by the functional $\alpha_0 = 0$. Actually, we have used this property in Sec. 7.)

To each $\lambda \in \Lambda(R', \hat{w})$, there correspond the Lagrange function and its second variation $\Omega[\lambda](\bar{w})$.

Definition 9.1. We say that the trajectory $\hat{w}$ of System (R') satisfies the weak γ-sufficiency if, for some $a > 0$,

$$\max_{\Lambda(R',\hat{w})} \Omega[\lambda](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall\, \bar{w} \in \mathcal{K}(R'). \tag{9.3}$$

For System (R), the weak γ-sufficiency was called the quadratic rigidity (see Sec. 7).

Theorem 9.1. *Let the weak γ-sufficiency (i.e., the quadratic rigidity) be satisfied for the trajectory $\hat{w}$ in System (R). Then it is satisfied also in System (R').*

Before giving a proof, let us make more precise the structure of the set $\Lambda(R')$. The previous set $\Lambda(R)$ consisted of the normalized collections $\lambda = (\psi = \psi_x(t), \psi_z, \beta_0, \beta_M, \beta_L)$ such that for the corresponding functions

$$H[\lambda](z, x, u) \;=\; z(\psi, r_0(x)) \;+\; \textstyle\sum u_i(\psi, r_i(x)),$$

$$l[\lambda](x(0), x_M(T), x_L(T)) \;=\; \beta_0\,(x(0) - a) \;+\; \beta_M \Delta x_M(T) \;+\; \beta_L \Delta x_L(T)$$

the following relations hold:

$$\dot{\psi}_x = -H_x[\lambda] = \; -\hat{z}\,\psi_x r_0'(\hat{x}),$$

$$\dot{\psi}_z = -H_z[\lambda] = \; -\psi_x r_0(\hat{x}),$$

$$\psi_x(0) = l'_{x(0)}[\lambda] = \beta_0, \qquad \psi_x(T) = -l'_{x(T)}[\lambda] = -(\beta_M, \beta_L),$$

$$\psi_z(0) = \; \psi_z(T) = \; 0,$$

$$H_{u_i}[\lambda] = \; (\psi, r_i(\hat{x})) = 0, \qquad i = 1, \ldots, k-1,$$

$$H[\lambda](\hat{z}, \hat{x}, \hat{u}) = \; \hat{z}\,(\psi, r_0(\hat{x})) = \text{const}\,.$$

These conditions, as was already shown in Sec. 3, are reduced to the fact that the function $\psi(t) = \psi_x(t)$ satisfies relations (9.1) and (9.2). The transversality conditions are not essential here, and the nontriviality of the collection λ is equivalent to the nontriviality of the function ψ (here the function $\psi_z \equiv 0$). Note that here, $\forall\, \lambda$ we have $\beta_L = 0$, since $\psi(T) \in M$ by the definition of the subspace M.

For the new system, the set $\Lambda(R')$ consists of collections $\lambda = (\psi = \psi_x(t), \psi_z, \beta_0, \beta_M, \alpha_L)$, where α_L is now a scalar, $\alpha_L \geq 0$, satisfying the same conditions, in which only the endpoint function $l[\lambda]$ is changed; it is now equal to

$$l[\lambda](x(0), x_M(T), x_L(T)) = \; \beta_0\,(x(0) - a) + \beta_M \Delta x_M(T) + \alpha_L |\Delta x_L(T)|^2,$$

therefore, the transversality conditions will now be as follows:

$$\psi(0) = \beta_0, \qquad \psi(T) = (-\beta_M, 0). \tag{9.4}$$

The last relation simply means that $\psi(T) \in M$. Here still $\psi = \psi_x$ satisfies (9.1) and (9.2), and $\psi_z \equiv 0$ (hence ψ_z is not considered in the sequel).

Thus, if $\lambda \in \Lambda(R')$, then ψ satisfies relations (9.1), (9.2), and $\psi(T) \in M$, i.e., no new ψ as compared with $\Lambda(R)$ appear.

On the other hand, if $\lambda = (\psi, \beta_0, \beta_M, \beta_L) \in \Lambda(R)$, then ψ satisfies relations (9.1), (9.2), and $\psi(T) \in M$. But then ψ satisfies (9.4), and, therefore, for any $\alpha \geq 0$, the collection $\lambda = (\psi, \beta_0, \beta_M, \alpha) \in \Lambda(R')$.

Thus, the store of functions $\psi = \psi_x$ in the sets $\Lambda(R)$ and $\Lambda(R')$ is one and the same. However, one more element appears in the set $\Lambda(R')$; namely,

$$\lambda_0 = (\psi(t) \equiv 0, \quad \beta_0 = 0, \quad \beta_M = 0, \quad \alpha_L = 1),$$

and, as can be easily shown, any element of $\Lambda(R')$ is a convex combination of some element having $\psi \neq 0$ and $\alpha_L = 0$, and of this "labelled" element λ_0. The above considerations imply the following statement (we have actually proved it).

Lemma 9.1. *There exists an injection $\pi : \Lambda(R) \to \Lambda(R')$, for which the Lagrange function $\Phi[\lambda](z, x, u)$ does not change, and $\Lambda(R')$ is the convex hull of $\pi(\Lambda(R))$ and the point λ_0 up to normalization.*

Indeed, consider the mapping

$$\lambda = (\psi, \beta_0, \beta_M, \beta_L = 0) \longmapsto \lambda' = (\psi, \beta_0, \beta_M, \alpha_L = 0).$$

as π. Obviously, it satisfies the required properties. The Lagrange function for such λ and λ' in both systems is one and the same. More precisely, $\Phi[\lambda, R] = \Phi[\lambda', R']$. Then the second variations of these functions (at the point $\hat{w}$) are related exactly in the same way: $\Omega[\lambda, R](\bar{w}) = \Omega[\lambda', R'](\bar{w})$.

Proof of Theorem 9.1. From Lemma 9.1, it follows that the left-hand side of (9.3) is the maximum of the following two values: $\Omega[\Lambda(R)](\bar{w})$ and $\Omega[\lambda_0](\bar{w})$. For the point λ_0, we have $\Phi[\lambda_0](w) = |\Delta x_L(T)|^2$; therefore, $\Omega[\lambda_0](\bar{w}) = |\bar{x}_L(T)|^2$. Thus, the statement of the theorem is reduced to the following. For some $a > 0$, let

$$\Omega[\Lambda(R)](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall\, \bar{w} \in \mathcal{K}(R). \tag{9.5}$$

Then, for some $a' > 0$,

$$\max\{\, \Omega[\Lambda()](\bar{w}),\, |\bar{x}_L(T)|^2 \,\} \;\geq\; a'\gamma(\bar{w}) \qquad \forall\, \bar{w} \in \mathcal{K}('). \tag{9.6}$$

Let us prove this statement by contradiction. Suppose that (9.6) is not satisfied, i.e., there exists a sequence $\bar w_n \in \mathcal{K}(R')$ for which the following simultaneously hold:

$$\Omega[\Lambda()](\bar w_n) \;\le\; o\,(\gamma(\bar w_n)), \tag{9.7}$$

$$|\bar x_{L,n}(T)|^2 \;\le\; o\,(\gamma(\bar w_n)).$$

Denote $\gamma(\bar w_n) = \gamma_n$ for brevity. The last inequality means that the variation $\bar w_n$ violates the constraint $\bar x_L(T) = 0$ from amongst those ones that specify the subspace $\mathcal{K}(R)$ (and only this constraint, since $\bar w_n \in \mathcal{K}(R')$), by a value of order $o\,(\sqrt{\gamma_n})$. Then from the Banach open mapping theorem (the distance from a point to the kernel of a linear surjective operator is estimated by the norm of the image of this point), it follows that there exists a correction $\tilde w_n \in \mathcal{K}(R')$, such that

$$||\tilde w_n|| = \;|\tilde z_n| + ||\tilde x_n||_C + ||\tilde u_n||_\infty \le o\,(\sqrt{\gamma_n}) \tag{9.8}$$

$$\text{and} \qquad w'_n = \bar w_n + \tilde w_n \in \mathcal{K}(R), \tag{9.9}$$

i.e., $x'_n(T) \;=\; \bar x_n(T) + \tilde x_n(T) \;=\; 0.$ It is easy to see that for the new sequence $\gamma(w'_n) = \gamma(\bar w_n + \tilde w_n) = \gamma_n + o(\gamma_n)$, i.e., the new infinitesimal value $\gamma'_n = \gamma(w'_n)$ is equivalent to the old γ_n.

Now let us estimate the value of $\Omega[\Lambda(R)]$ on the sequence w'_n, taking (9.7) into account. To this end, let us first estimate the change of the individual quadratic form for each λ. It is convenient to carry out such estimation for the following general quadratic form that has the zero Legendre term (the bars over the variables are not written here for convenience):

$$\Omega(x,u) = (Sp,p) + \int_0^T ((D(t)x,x) + 2(x,C(t)u))\;dt,$$

and which is considered on linear equation

$$\dot x = A(t)x + B(t)u, \tag{9.10}$$

where $p = (x_0, x_T)$, the matrix S is of dimension $2n \times 2n$, the matrices D and A of corresponding dimensions are measurable and bounded, and the matrices B and C are Lipschitzian. (The quadratic forms of our family $\Omega[\lambda]$ obviously belong to this general class. It should be only noted that as the phase variable we have the pair (z,x), not x.) For system (9.10) we have, by definition,

$$\gamma(w) = \;|x(0)|^2 + |y(T)|^2 + \int_0^T |y|^2\,dt, \qquad \text{where } \dot y = u,\;\; y(0) = 0,$$

and one can show (see Appendix D) that the following estimate is valid:

$$|x(0)|^2 + |x(T)|^2 + \int_0^T |x|^2\,dt \;\le\; \text{const}\,\cdot\gamma(w). \tag{9.11}$$

Lemma 9.2. *Let sequences* $w_n = (x_n, u_n)$ *and* $\tilde{w}_n = (\tilde{x}_n, \tilde{u}_n)$ *satisfy the linear system (9.10), and let*

$$||\tilde{w}_n||_\infty = ||\tilde{x}_n||_\infty + ||\tilde{u}_n||_\infty \le o(\sqrt{\gamma_n}), \tag{9.12}$$

where $\gamma_n = \gamma(w_n)$. *Then, for* $w_n' = w_n + \tilde{w}_n$, *we have* $\gamma(w_n') \sim \gamma_n$ *and*

$$\Omega(w_n') = \Omega(w_n) + o(\gamma_n).$$

The proof is also given in Appendix D.

In our situation, this lemma implies that $\forall \lambda$,

$$\Omega[\lambda](w_n') = \Omega[\lambda](\bar{w}_n) + a_n[\lambda] \cdot \gamma_n,$$

where $a_n[\lambda] \to 0$. Since the dependence on λ here is linear, for any bounded set $M = \{\lambda\}$ the following holds:

$$\sup_{\lambda \in M} \Omega[\lambda](w_n') = \sup_{\lambda \in M} \Omega[\lambda](\bar{w}_n) + b_n \gamma_n,$$

where $b_n = b_n(M) \to 0$. In particular, for $M = \Lambda(R)$ we have

$$\Omega[\Lambda()](w_n') = \Omega[\Lambda()](\bar{w}_n) + o(\gamma_n).$$

But then, from (9.7), we obtain $\Omega[\Lambda()](w_n') \le o(\gamma_n)$, which, in view of (9.9), contradicts (9.5). Theorem 9.1 is proved.

Thus, we are now in System (R'), and the weak γ-sufficiency is satisfied for the trajectory $\hat{w}$ in this system.

10 Passage to Problems (P_1), (P), and (Y_*)

Consider now Problem (P_1), which is obtained by adding the functional $J = z(0) \to \min$ to System (R'), and let us see what the set $\Lambda(P_1, \hat{w})$ will look like here. This set consists of all normalized collections $\lambda = (\psi = \psi_x(t), \psi_z, \beta_0, \beta_M, \ \alpha_0, \alpha_L)$, where $\alpha_0 \ge 0, \ \ \alpha_L \ge 0$, such that, for the corresponding functions

$$H[\lambda](z, x, u) = z(\psi, r_0(x)) + \sum u_i(\psi, r_i(x)) + \psi_z \cdot 0,$$

$$l[\lambda](z_0, x_0, x_T) = \alpha_0 z(0) + \beta_0(x(0) - a) + \beta_M \Delta x_M(T) + \alpha_L |\Delta x_L(T)|^2,$$

the following conditions hold (we take into account that $\hat{z} = 1$) :

$$H[\lambda](\hat{z}, \hat{x}, \hat{u}) = \psi\, r_0(\hat{x}) = \text{const}\,,$$

$$\dot{\psi}_z = -\psi\, r_0(\hat{x}), \qquad \psi_z(0) = \alpha_0, \quad \psi_z(T) = 0$$

(this implies $\psi r_0(\hat{x}) = \alpha_0/T \geq 0$),

$$\dot{\psi} = -\psi\, r_0'(\hat{x}), \qquad \psi(0) = \beta_0, \quad \psi(T) = (-\beta_M, 0),$$

and still $\psi\, r_i(\hat{x}) = 0, \quad i = 1, \ldots, k-1.$

Here, a priori is possible that $\alpha_0/T = \psi r_0(\hat{x}) > 0$, i.e., that a "nonsingular" $\psi = \psi_x$ can appear. But, from the transversality condition, $\psi(T) = (-\beta_M, 0)$, i.e., $\psi(T) \in M$, and from Proposition 9.1 it follows that $\psi r_0(\hat{x}) = 0$, i.e., $\alpha_0 = 0$ (and also $\psi_z = 0$).

Thus, in the passage from System (R') to Problem (P_1) no new $\psi = \psi_x$ appear, and hence the Lagrange sets for them coincide: $\Lambda(P_1, \hat{w}) = \Lambda(R', \hat{w})$. (More precisely, the set $\Lambda(P_1, \hat{w})$ consists of all elements of the form $(\alpha_0 = 0,\, \lambda)$, where $\lambda \in \Lambda(R', \hat{w})$, and only of such elements.)

Further, the cone of critical variations in Problem (P_1) is a halfspace in the subspace $\mathcal{K}(R')$, namely,

$$\mathcal{K}(P_1) = \{\bar{w} \in \mathcal{K}(R') \mid \bar{z}(0) \leq 0\}.$$

(The additional inequality corresponds to the functional J of Problem (P_1), which was not present in Problem (R').) But from the point of view of the nonnegativity of any homogeneous function of the second degree, a linear space and its halfspace are equivalent; therefore, one can take $\mathcal{K}(P_1) = \mathcal{K}(R')$.

These two facts imply

Theorem 10.1. *Let the weak γ-sufficiency be satisfied for the trajectory $\hat{w}$ in System (R'). Then it is satisfied also in Problem (P_1), i.e., for some $a > 0$, we have*

$$\max_{\Lambda(P_1)} \Omega[\lambda](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall\, \bar{w} \in \mathcal{K}(P_1). \tag{10.1}$$

Thus, we are now in Problem (P_1), and the weak γ-sufficiency (10.1) is satisfied for $\hat{w}$. Note that Problem (P_1) almost coincides with Problem (S_1); the only difference is that instead of the constraint $\Delta x(T) = x(T) - b = 0$ we now have the following two constraints: $\Delta x_M(T) = 0, \quad |\Delta x_L(T)|^2 \leq 0.$

Let us proceed further. Previously, we passed to Problem (S_1) from Problem (S). Let us now make the corresponding reverse passage, namely, from Problem (P_1) to the following problem.

Problem (P).

$$\dot{x} = zu_0 r_0(x) + \sum_{i=1}^{k-1} u_i r_i(x), \tag{10.2}$$

$$\dot{z} = 0, \quad u_0 \le 1,$$

$$x(0) = a, \quad \Delta x_M(T) = 0, \quad |\Delta x_L(T)|^2 \le 0.$$

$$J = z(0) \to \min,$$

i.e., we have "released" u_0. (We had $u_0 = 1$, and now we have $u_0 \le 1$.)

Since the admisssible control set has expanded, the corresponding Lagrange set Λ can only narrow. But, in this case, no narrowing will take place.

Lemma 10.1. *Problems (P) and (P_1) have one and the same set $\Lambda(\hat{w})$.*

Proof. The introduction of $u_0 \le 1$ leads to the appearance of the additional condition

$$H_{u_0}[\lambda] = \psi\, r_0(\hat{x}) \ge 0 \tag{10.3}$$

in the MP. But, for any $\lambda \in \Lambda(P_1)$, the following simply holds: $\psi r_0(\hat{x}) = 0$; hence, $\lambda \in \Lambda(P)$. The inverse inclusion $\Lambda(P) \subset \Lambda(P_1)$, as was already said, is obvious.

Remark. This lemma would be valid also for the usual boundary condition $x(T) = 0$, i.e., for the passage from Problem (S_1) to Problem (S), since, $\forall\, \lambda \in \Lambda(S_1)$, the following holds: $\psi\, r_0(\hat{x}) = \alpha_0/T \ge 0$, and this is just the additional condition (10.3) in the MP for Problem (S).

Now it is important for us that $\psi\, r_0(\hat{x}) = 0$ in Problem (P_1), from which, by Lemma 10.1, it follows that the trajectory $\hat{w}$ is singular in Problem (P).

Theorem 10.2. *The weak γ-sufficiency for the trajectory $\hat{w}$ in Problem (P_1) is equivalent to the weak γ-sufficiency for this trajectory in Problem (P).*

The proof completely repeats the proof of Theorem 6.1 with the only difference that in the violation function, instead of $|x(T)-b|$, one should everywhere take $|\Delta x_M(T)|+|\Delta x_L(T)|^2$. Since the term $|x(T) - b|$ actually took only "passive" participation in all above considerations, nothing will change under such replacement. The detailed verification of this fact is left to the reader.

Let us now pass from Problem (P) to Problem (Y_*), which is obtained from Problem (P) by the replacement of the differential equation (10.2) with the "initial"equation (5.1).

Problem (Y_*).

$$\dot{x} \;=\; z\,\left(u_0 r_0(x) \,+\, \sum_{i=1}^{k-1} u_i r_i(x)\right),$$

$$\dot{z} = 0, \quad u_0 \le 1,$$

$$x(0) = a, \quad \Delta x_M(T) = 0, \quad |\Delta x_L(T)|^2 \le 0,$$

$$J = z(0) \to \min\,.$$

Theorem 10.3. *The weak γ-sufficiency for the trajectory $\hat{w}$ in Problem (P) is equivalent to the weak γ-sufficiency for this trajectory in Problem (Y_*).*

The proof is completely similar to the proof of Theorem 5.2.

At last, using Theorem 8.1 from [9], we can pass to the final point of our chain of passages, namely, to

Problem $(\tilde{Y})$.

$$\dot{x} \;=\; z\,\left(u_0 r_0(x) \,+\, \sum_{i=1}^{k-1} u_i r_i(x)\right), \tag{10.4}$$

$$\dot{z} = 0, \quad u \in \tilde{U}, \quad x(0) = a, \tag{10.5}$$

$$\Delta x_M(T) = 0, \quad |\Delta x_L(T)|^2 \le 0, \tag{10.6}$$

$$J = z(0) \to \min,$$

which differs from Problem (Y_*) by the fact that here the admissible control set is not the halfspace $u_0 \le 1$, but is the convex compact set $\tilde{U}$ described in the end of Sec. 8. Namely, similar to Theorem 5.1, the following holds:

Theorem 10.4. *Let the trajectory $\hat{w}$ satisfy the γ-sufficient condition for the weak minimum in Problem (Y_*). Then $\hat{w}$ is a point of the strict strong minimum in Problem $(\tilde{Y})$ with the set $\tilde{U}$. Moreover, there exist $\varepsilon,\, C > 0$, such that for all $w = (z, x, u)$ from the set $|z-1| + \|x-\hat{x}\|_C < \varepsilon$, satisfying (10.4) and (10.5), the following estimate holds:*

$$(z-1)^+ \;+ |\Delta x_M(T)| + |\Delta x_L(T)|^2 \;\ge\; C\,\gamma(w-\hat{w}). \tag{10.7}$$

Now it remains to note, that if one neglects the γ-estimate (10.7), i.e., if the second statement of Theorem 10.4 is ignored, then the strict strong minimum (at the point $\hat{w}$) in Problem $(\tilde{Y})$ coincides with the strict strong minimum in Problem $(\tilde{Z})$, which (see Sec. 4) differs from Problem $(\tilde{Y})$ only by the fact that the boundary conditions in it are of the initial form: $\Delta x(T) = x(T) - b = 0$, since, outside the γ-estimates, the constraints (10.6) have the same sense as $\Delta x(T) = 0$.

Thus, we have covered the path from System (R) to Problem $(\tilde{Z})$:

$$\text{System } (R) \to \text{System } (R') \to \text{Problem } (P_1) \to \text{Problem } (P) \to$$
$$\to \text{Problem } (Y_*) \to \text{Problem } (\tilde{Y}) \to \text{Problem } (\tilde{Z}),$$

but this time not assuming the singularity of the trajectory $\hat{w}$ in Problem $(\tilde{Z})$.

Theorems 9.1, 10.1–10.4 imply

Theorem 10.5. *Let the trajectory $\hat{w}$ in System (R) satisfy the weak γ-sufficiency, i.e., let $\hat{w}$ be quadratically rigid. Then, it is a point of the strict strong minimum in Problem $(\tilde{Z})$.*

Thus, Theorem 8.1, the Main Theorem 1 of this paper, is proved.

This theorem is stronger than Theorem 5.2 from [1], since in the latter (a) only two-dimensional distributions and only sub-Riemmanian metrics were admitted; (b) it was assumed that inequality (7.21) was satisfied for an individual quadratic form $\Omega[\psi](\bar{w})$ for some $\psi \in G_a(\Psi_0)$, which is a more restrictive requirement; (c) not the strong minimum was guaranteed, but only the minimum with respect to $\|w\|_1$.

Part III
Sufficient Conditions for Pontryagin Minimality

Consider again the initial Problem (Z) on the curve of minimum length, but now we will study the presence of not the strong minimum, but of the Pontryagin minimum (briefly, Π-minimum) at the given trajectory $\hat{w}$. Here it will be sufficient for us to require from the submetric to satisfy Assumption A2 on the existence of a twice smooth support (not necessarily strictly support) hyperplane in a neighborhood of the trajectory $\hat{x}(t)$.

11 Passage to Problems (Z_*) and (S)

Let us again write Problem (Z) in an associated basis that is support for a given submetric, and let us pass directly to Problem (Z_*) with the subspace of controls.

Problem (Z_*).

$$\dot{x} = z\left(u_0 r_0(x) + \sum_{i=1}^{k-1} u_i r_i(x)\right), \tag{11.1}$$

$$\dot{z} = 0, \qquad x(0) = a, \quad x(T) = b, \tag{11.2}$$

$$u_0 \le 1, \qquad u_1, \ldots, u_{k-1} \text{ are free,}$$

$$J = z(0) \longrightarrow \min,$$

It is obvious that any sufficient condition for the Π-minimum in Problem (Z_*) will automatically be a sufficient condition for the Π-minimum in the initial Problem (Z); therefore, we can apply the general γ-sufficient conditions for Π-minimum from [9] to Problem (Z_*). However, as in Sec. 5, let us first carry out a number of passages that will enable us to simplify these conditions. Introduce the following notions.

Definition 11.1. We will say that a sequence $w_n = (z_n, x_n, u_n)$ converges in the Pontryagin sense to $\hat{w} = (\hat{z}, \hat{x}, \hat{u})$ and we will write $w_n \xrightarrow{\Pi} \hat{w}$ if

$$|z_n - \hat{z}| + \|x_n - \hat{x}\|_\infty + \|u_n - \hat{u}\|_1 \to 0, \qquad \|u_n - \hat{u}\|_\infty \le \mathcal{O}(1).$$

Let $\mathcal{D}$ be a subset of W containing $\hat{w}$.

Definition 11.2. We will say that the property $\mathcal{F}$ holds on the set $\mathcal{D}$ if there exists $C > 0$ such that for any sequence $w_n \in \mathcal{D}$, $w_n \xrightarrow{\Pi} \hat{w}$, the estimate

$$(z_n - 1)^+ + |x_n(T) - b| \ge C\gamma(w_n - \hat{w}) \tag{11.3}$$

holds for all sufficiently large n.

Now let us introduce the set

$$\mathcal{D}(Z_*) = \{w \mid \text{Eq. (11.1) holds}, \ \dot{z} = 0, \ x(0) = a, \ u_0 \le 1\}.$$

Theorems 7.1 and 7.2 in [9] imply

Theorem 11.1. *The γ-sufficient condition for the Π-minimum for the trajectory $\hat{w}$ in Problem (Z_*) is equivalent to the fulfilment of the property $\mathcal{F}$ on the set $\mathcal{D}(Z_*)$.*

Further, as in Sec. 5, let us pass to Problem (S), which differs from Problem (Z_*) by the fact that Eq. (11.1) is replaced by the following equation:

$$\dot{x} = zu_0 r_0(x) + \sum_{i=1}^{k-1} u_i r_i(x). \tag{11.4}$$

By analogy with Theorem 5.2, the following holds.

Theorem 11.2. *The γ-sufficient condition for the Π-minimum in Problem (Z_*) is equivalent to the γ-sufficient condition for the Π-minimum in Problem (S).*

The proof, in view of Theorem 7.2 (ii) from [9], is a consequence of the following lemma which is similar to Lemma 5.1.

Lemma 11.1. *The fulfilment of the property $\mathcal{F}$ on the set $\mathcal{D}(Z_*)$ is equivalent to its fulfilment on the set $\mathcal{D}(S)$.*

(The set $\mathcal{D}(S)$ was introduced in Sec. 6.) The proof almost literally repeats the proof of Lemma 5.1.

Now let us write the γ-sufficient condition for the Π-minimum in Problem (S). To this end, we must define the corresponding objects. First, it is the cone of critical variations for Problem (S) at the point $\hat{w}$, which, according to [9, Sec. 4], is $\mathcal{K} \cap \mathcal{N}$, where $\mathcal{K}$ is given by the linearization of all constraints of the problem, except for the constraint on the control, and $\mathcal{N}$ consists of variations that are, for every t, tangent to the admissible control set at the point $\hat{u}(t)$. In our case, $\forall\, t$, the tangent cone to the set $U_* = \{u \in \mathbb{R}^k \,|\, u_0 \le 1\}$ at the point $\hat{u}(t)$ is one and the same: $N = \{\bar{u} \in \mathbb{R}^k \mid \bar{u}_0 \le 0\}$.

Thus, $\mathcal{K}$ consists of all $\bar{w} = (\bar{z}, \bar{x}, \bar{u})$ satisfying the equations

$$\dot{\bar{x}} = \ \bar{z}\, r_0(\hat{x}) + r_0'(\hat{x})\bar{x} + \bar{u}_0 r_0(\hat{x}) + \sum_{i=1}^{k-1} \bar{u}_i r_i(\hat{x}), \tag{11.5}$$

$$\dot{\bar{z}} = 0, \quad \bar{x}(0) = \bar{x}(T) = 0, \tag{11.6}$$

and the inequality $\bar{z} \le 0$ (the linearization of the functional), while $\mathcal{N} = \{\bar{w} \mid \bar{u}_0(t) \le 0 \text{ a.e.}\}$.

In addition, for the cone N, we determine the maximal linear subspace $N_0 \subset N$; in our case it is the hyperplane $\bar{u}_0 = 0$.

As was already proved, $\Lambda(S, \hat{w}) = \Lambda(Z, \hat{w})$. For each λ we must consider the corresponding Lagrange function $\Phi[\lambda](w)$ and take its second variation $\Omega[\lambda](\bar{w})$ at the trajectory $\hat{w}$. In the above, we wrote the quadratic form (7.11) for Problem (S_1).

Since Problem (S), as compared with Problem (S_1), involves an additional control u_0, now two more terms will appear in the quadratic form, namely,

$$- \int_0^T \Big(\bar{z} \bar{u}_0 \, \psi r_0(\hat{x}) \; + \; \bar{u}_0 \, \psi(r_0'(\hat{x}) \, \bar{x}) \Big) \, dt, \tag{11.7}$$

but it would not affect the selection of λ into sets $G_a(\Lambda)$: one can show that, like for Problem (S_1), this selection is specified by conditions (7.17) and (7.18); no other conditions would appear (see Appendix C). However, we will not use this fact, as well as a particular form of the functional $\Omega[\lambda](\bar{w})$; we only endow it with the subscript S.

Further, according to [9, Sec. 4] (see also [5, 8]), for our control system (11.4), (11.2) and for any λ, we define a cubic functional $\rho[\lambda](\bar{w})$; this functional is considered under the linear constraints (11.5) and (11.6) for $\bar{u}(t) \in N_0$ almost everywhere (i.e., for $\bar{u}_0(t) = 0$ a.e. in this case). Using this functional, we define subsets $E_a(\Lambda)$ of the sets $G_a(\Lambda)$. Precisely, $\forall \, a$ the set $E_a(\Lambda)$ consists of all those $\lambda \in G_a(\Lambda)$, for which the functional $\rho[\lambda](\bar{w})$ satisfies an additional equality-type condition on the abovenoted subspace of $\bar{w}$. In our case, this condition means (see Appendix E) that

$$\psi(t) \, [\, r_i, [r_j, r_s] \,] \, (\hat{x}(t)) = 0 \qquad \forall \, i, j, s = 1, \ldots, k-1. \tag{11.8}$$

Note that, in contrast to conditions (7.17) and (7.18), this condition now *depends on the choice of associated basis.* Moreover, as was shown by A. A. Milyutin, if it is satisfied for some basis, then it is not satisfied for all other sufficiently close bases.

Now we can formulate a sufficient condition for the Π-minimum in Problem (S). Theorems 7.1, 7.2, and 7.4 from [9] imply

Theorem 11.3. *Suppose that, for some* $a > 0$,

$$\Omega_S \, [E_0(\Lambda(S))](\bar{w}) \geq a \gamma(\bar{w}) \qquad \forall \, \bar{w} \in \mathcal{K} \cap \mathcal{N}. \tag{11.9}$$

Then $\hat{w}$ *is a point of the strict* Π *-minimum in Problem* (S).
Inequality (11.9) is equivalent to the inequality

$$\Omega_S \, [E_a(\Lambda(S))](\bar{w}) \geq a \gamma(\bar{w}) \qquad \forall \, \bar{w} \in \mathcal{K} \cap \mathcal{N}, \tag{11.10}$$

(here $E_a(\Lambda(S))$ *is automatically nonempty), and is also equivalent to the fulfilment of the property* $\mathcal{F}$ *on the set* $\mathcal{D}(S)$.

12 Passage to $u_0 = 1$

Thus, we have the γ-sufficient condition for the Π-minimum for the trajectory $\hat{w}$ in Problem (S). Let us now pass from the halfspace $u_0 \le 1$ to the subspace $u_0 = 1$, i.e., to Problem (S_1). However, here it is not possible to make this passage directly, as it was done in Sec. 6. We will make this passage in two steps: first, we will pass to an intermediate Problem $(S_{\frac{1}{2}})$, in which the control set $U_{\frac{1}{2}}$ is the stripe $\frac{1}{2} \le u_0 \le 1$, and only then to Problem (S_1) with the control set $u_0 = 1$. The reasons for using the intermediate problem will be given below in the proof of Lemma 12.1.

Consider the first step of this passage. Since in a uniform neighborhood of $\hat{u}(t)$ the set U_* coincides with the set $U_{1/2}$, in the passage from Problem (S) to Problem $(S_{\frac{1}{2}})$ the set $\Lambda(\hat{w})$ will not change; hence, the family of Lagrange functions and the corresponding families of quadratic forms $\Omega[\lambda](\bar{w})$ and cubic functionals $\rho[\lambda](\bar{w})$ will not change. In addition, the tangent cone N to the sets $u_0 \le 1$ and $\frac{1}{2} \le u_0 \le 1$ at the point $\hat{u}$ (for which $\hat{u}_0 = 1$) will be one and the same. Therefore, the sets $G_a(\Lambda)$ and $E_a(\Lambda)$ for Problem $(S_{\frac{1}{2}})$ will also be the same as for Problem (S). Thus, all the objects taking part in the statement of the γ-sufficient conditions for the Π-minimum, i.e., in the inequalities (11.9) and (11.10), will be the same for both problems. Hence, the following is valid.

Theorem 12.1. *The γ-sufficient condition for the Π-minimum in Problem (S) coincides with the γ-sufficient condition for the Π-minimum in Problem $(S_{\frac{1}{2}})$.*

Now let us consider the passage from Problem $(S_{\frac{1}{2}})$ to Problem (S_1) with the subspace $u_0 = 1$ as the control set. As was established in Sec. 6, the set $\Lambda(\hat{w})$ is not changed in this process.

Theorem 12.2. *The γ-sufficient condition for the Π-minimum in Problem $(S_{\frac{1}{2}})$ is equivalent to the γ-sufficient condition for the Π-minimum in Problem (S_1).*

The proof, in view of Theorem 7.2 (ii) from [9], is a consequence of the following lemma.

Lemma 12.1. *The validity of property $\mathcal{F}$ on the set $\mathcal{D}(S_{\frac{1}{2}})$ is equivalent to its validity on the set $\mathcal{D}(S_1)$.*

The proof is similar to the proof of Lemma 6.2. Since $\mathcal{D}(S_{\frac{1}{2}}) \supset \mathcal{D}(S_1)$, it is sufficient, as before, to prove the implication $(\Leftarrow)$. We will prove the contrapositive implication. To this end, it is sufficient to prove that there exists $\varepsilon_0 > 0$ such that if $0 < \varepsilon \le \varepsilon_0$ and if a sequence $w_n \in \mathcal{D}(S_{\frac{1}{2}})$, $w_n \ne \hat{w}$, $w_n \xrightarrow{\Pi} \hat{w}$, satisfies the estimate

$$\sigma(w_n) = (z_n - 1)^+ + |x_n(T) - b| \le \varepsilon\, \gamma(w_n - \hat{w}), \tag{12.1}$$

then there exists a sequence $w'_n \in \mathcal{D}(S_1)$, $w'_n \neq \hat{w}$, $w'_n \xrightarrow{\Pi} \hat{w}$, for which

$$\sigma(w'_n) \;\leq\; Q\varepsilon\, \gamma(w'_n - \hat{w}), \tag{12.2}$$

where Q is a constant that does not depend on the sequence and on ε.

Let us follow the proof of Lemma 6.2. Let be given a sequence w_n satisfying (12.1) with some $\varepsilon > 0$. As before, reparametrize each its term in such a way that for a new sequence w'_n we will have $u'_{0,n}(t) = 1$. Since we have $\frac{1}{2} \leq u_{0,n}(t) \leq 1$ (i.e., $\delta = 1/2$), the function $s_n(t)$ defined by formula (6.8) will still define a bi-Lipschitzian mapping of the interval $[0,T]$ onto itself ($\frac{1}{2} \leq \dot{s}_n(t) \leq 2$); therefore, the required parametrization is available. (It is just at this point we use the fact that $u_{0,n}(t)$ is bounded away from zero. If, as before, we had only the inequality $u_{0,n}(t) \leq 1$, then, since now u_n converges to $\hat{u}$ not uniformly, but only in the Pontryagin sense, for all n the inequality $u_{0,n}(t) < 0$ could hold on a set of positive measure, and then the above reparametrization of the trajectory w_n would not be available.)

The trajectory w'_n constructed by the formulas (6.10) and (6.11) belongs to $\mathcal{D}(S_1)$. Since $\int_0^T |1 - u_{0,n}|\, dt \to 0$, we have $\theta_n \to 1 - 0$, and therefore, also $z'_n \to 1 - 0$, and since $\forall\, i = 1, \ldots, k-1$,

$$\int_0^T |u'_{i,n}(s)|\, ds = \int_0^T |u_{i,n}(t)|\, dt \to 0,$$

from (6.12) and Gronwall-type estimates we easily obtain $\|x'_n - \hat{x}\|_C \to 0$. Therefore, $w'_n \xrightarrow{\Pi} \hat{w}$.

The order γ for Problems $(S_{\frac{1}{2}})$ and (S_1) is determined by the same formulas (6.5) and (6.6); therefore, inequality (12.1) for w_n means, similar to (6.16), that the following holds (we omit the subscript n):

$$(z-1)^+ + |x(T) - b| \;\leq\; \varepsilon \left(|\bar{z}|^2 + |\bar{y}_0(T)|^2 + \eta(y)\right). \tag{12.3}$$

And we want to show that then the following inequality holds:

$$(z'-1)^+ + |x(T) - b| \leq Q\varepsilon \left(|\bar{z}'|^2 + \eta(y')\right), \tag{12.4}$$

where Q does not depend on the sequence and on ε.

Since inequalities (6.9) now hold for $\delta = \frac{1}{2}$, then (6.18) also holds for this δ, i.e.,

$$\frac{1}{2}\eta(y) \;\leq\; \eta(y') \;\leq\; 2\,\eta(y). \tag{12.5}$$

Further, we set $p = -(1/T)\bar{y}_0(T)$ and by analogy with (6.20), we obtain the estimate

$$(z-1)^+ + |x(T) - b| \;\leq\; \varepsilon \cdot C(T) \left(|\bar{z}'|^2 + p^2 + \eta(y')\right), \tag{12.6}$$

where $C(T)$ is a constant that depends only on T.

Denote by $\mathcal{M}_\varepsilon$ the class of all sequences $w_n \in \mathcal{D}(S_{\frac{1}{2}})$, $w_n \xrightarrow{\Pi} \hat{w}$, $w_n \neq \hat{w}$ (for large n), satisfying the estimate (12.3), or, which is the same, satisfying the estimate (12.1). Assume that there exist $\varepsilon_0 > 0$ and a constant K such that for any sequence from the class $\mathcal{M}_{\varepsilon_0}$, the following estimate holds:

$$p^2 \le K\left(|\bar{z}'|^2 + \eta(y')\right). \tag{12.7}$$

Since under the decrease of ε, the class $\mathcal{M}_\varepsilon$ narrows, this constant K fits for any class $\mathcal{M}_\varepsilon$ with $\varepsilon < \varepsilon_0$. And since it is just sufficient for us to consider $\varepsilon \le \varepsilon_0$, from (12.6), (12.7), and from the inequality $z' \le z$, we obtain the required estimate (12.4).

Now let us show that such a constant K exists for the class $\mathcal{M}_\varepsilon$ even for any $\varepsilon > 0$. Assume that for some $\varepsilon > 0$ there is no such constant, i.e., for any $\alpha > 0$, there exists a sequence w_n from the class $\mathcal{M}_\varepsilon$ for which

$$|\bar{z}'|^2 + \eta(y') \le \alpha\, p^2. \tag{12.8}$$

Then, from (12.6), we obtain

$$(\bar{z})^+ + |x(T) - b| \;\le\; \varepsilon \cdot C(T)\,(1+\alpha)\,p^2. \tag{12.9}$$

Here, $\theta \bar{z} = \bar{z}' + p$, and $\theta \to 1$; therefore,

$$(\bar{z}' + p)^+ \;\le\; 2\,C\varepsilon\,(1+\alpha)\,p^2. \tag{12.10}$$

From (12.8), it follows that $|\bar{z}'| \le \sqrt{\alpha}p$. Let us assume that $\alpha \le 1/4$. Then $p \ge 0$ implies $(\bar{z}' + p)^+ \ge (1 - \sqrt{\alpha})p \ge \frac{1}{2}p$, and, by virtue of (12.10), $\frac{1}{2}p \le 3C\varepsilon\, p^2$, and from here, since $p_n \to 0+$ (we again write the subscript n), it obviously follows that $p_n = 0$ for large n.

But then also $\bar{z}'_n = 0$, and $\bar{z}_n = 0$, i.e., $z_n = 1$, and, by virtue of (12.8), we have $\eta(y'_n) = 0$, hence, $\eta(y_n) = 0$, i.e., $u_{i,n}(t) \equiv 0$, $i = 1, \dots, k-1$. In addition, $p_n = 0$ implies $\theta_n = 1$, $u_0(t) \equiv 1$, and then $x_n = \hat{x}_n$. Thus, for large n, the trajectory w_n coincides with $\hat{w}$, which contradicts the definition of the class $\mathcal{M}_\varepsilon$. Therefore, the required constant K does exist. (It is clear from the above arguments that $K = 4$ obviously fits.) Lemma 12.1 is proved, and thus, Theorem 12.2 is also proved along with this lemma.

13 Quadratic Conditions for Π-Minimum in Problem (S_1)

Thus, we have passed to the γ-sufficient conditions for the Π-minimum in Problem (S_1). In this problem, the control $u \in \mathbb{R}^{k-1}$ is free, therefore, according to Sec. 11 (see also [5, 9]), the γ-conditions for the Π-minimum differ from the γ-conditions for the weak minimum by the additional condition (11.8), to which $\lambda \in \Lambda(\hat{w})$ or, which is the same, $\psi \in \Psi(\hat{w})$, should satisfy. The conditions for weak minimum in Problem (S_1) were written out in Sec. 7. The cone of critical variations here is the subspace given by relations (7.6), (7.8), and (7.9); the quadratic form $\Omega[\psi](\bar{w})$ has the form (7.11), the order γ is given by formula (7.19). The set $G_a(\Psi)$ consists of all $\psi \in \Psi$ satisfying conditions (7.17), and (7.18), and the set $E_a(\Psi)$ consists of those ψ that satisfy both these conditions and relations (11.8).

According to Theorem 7.1(ii) from [9], the γ-sufficient condition for the Π-minimum for the point $\hat{w}$ in Problem (S_1) consists in the fact that, for some $a > 0$, the following holds:

$$\Omega\,[E_0(\Psi)](\bar{w}) \geq a \int_0^T |\bar{y}|^2\,dt \qquad \forall\, \bar{w} \in \mathcal{K}. \tag{13.1}$$

(This inequality automatically implies that $E_0(\Psi)$ is nonempty.)
According to [4, 5, 7], this condition is equivalent to the fact that, for the same $a > 0$, the following holds:

$$\Omega\,[E_a(\Psi)](\bar{w}) \geq a \int_0^T |\bar{y}|^2\,dt \qquad \forall\, \bar{w} \in \mathcal{K}. \tag{13.2}$$

From here, in view of Theorems 12.2, 12.1, and 11.2, we obtain the final form of quadratic sufficient conditions for the Π-minimum for the singular trajectory $\hat{w}$ in Problem (Z_*).

Theorem 13.1. *Let $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = (1, 0, \ldots, 0))$ be a singular trajectory of Problem (Z_*), and, for some $a > 0$, let inequality (13.1) or (13.2) be satisfied for this trajectory. Then $\hat{w}$ is a point of the strict Π-minimum in Problem (Z_*).*

Recall that Π-minimum in Problem (Z_*) with the subspace of controls $u_0 \leq 1$ means that for any bounded set $B \subset \mathbb{R}^k$, in Problem (Z_*) with the control set $\{u_0 \leq 1\} \cap B$, the point $\hat{w}$ gives the minimum with respect to the norm $\|w\|_1 = |z| + \|x\|_C + \|u\|_1$.

Now, as in Sec. 7, let us use Corollary 2 of Lemma 3.1, which asserts that the singularity of the trajectory $\hat{w}$ in Problem (Z_*) is equivalent to its singularity in the initial Problem (Z), and use the notion of the sheaf of submetrics corresponding to a given basis. Then we obtain the following strengthening of Theorem 13.1.

Theorem 13.2. *Let* $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = (1, 0, \dots, 0))$ *be a singular trajectory of Problem* (Z_*) *written in an associated basis, and let this trajectory satisfy inequality (13.1) or (13.2) for some* $a > 0$. *Then in Problem* (Z) *with any submetric from the sheaf that corresponds to this basis,* $\hat{w}$ *is a point of the strict minimum with respect to the norm* $\|w\|_1$.

The above is exactly the result of the "direct" application of the general quadratic conditions for the Π-minimum from [9] to Problem (Z). As in Sec. 7, a shortcoming of this result is the assumption of the singularity of $\hat{w}$ that is present in it. In Sec. 14, we will show that here again one can get rid of this assumption.

For the time being, as in Sec. 7, note that the functional J in no way enters conditions (13.1) and (13.2), and these conditions themselves coincide with the γ-sufficient conditions for the so-called Pontryagin rigidity of the trajectory $\hat{w}$ for System (R), that were obtained in [15].

Definition 13.1 [15]. A smooth Γ-admissible curve $\hat{x}(t)$, $t \in [0, T]$, connecting the points a and b, is called *Pontryagin-type rigid* (briefly, Π-rigid) with respect to a given associated basis, if the corresponding trajectory $(\hat{x}(t), \hat{u}(t) = (1, 0, \dots, 0))$ of the system $\dot{x} = u_0 r_0(x) + \sum u_i r_i(x)$ has the following property: for any K, $h > 0$, there exists an $\varepsilon > 0$ such that any trajectory $(x(t), u(t))$, $t \in [0, T]$, of this system connecting the same points and satisfying the inequalities $\|x - \hat{x}\|_C + \|u - \hat{u}\|_1 < \varepsilon$, $\|u\|_\infty \leq K$, and $u_0(t) \geq h$ a.e., is a reparametrization of the trajectory $(\hat{x}(t), \hat{u}(t))$.

It is easy to show that this property is equivalent to the fact that, for any K, the trajectory $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = 0)$ of System (R) is isolated with respect to the norm $\|w\|_1$ in the set of all trajectories of this system on the given interval $[0, T]$, satisfying the constraint $\|u\|_\infty \leq K$.

Definition 13.2 [15]. The trajectory $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = 0)$ of System (R) is called *quadratically* Π*-rigid* in a given associated basis, if it satisfies inequality (13.1) or (13.2) for some $a > 0$.

In [15], it is proved that any quadratically Π-rigid trajectory is Π-rigid in the given basis. The property of quadratic Π-rigidity, in contrast to that of "ordinary" rigidity, depends on the choice of associated basis (since the fulfilment of condition (11.8) depends on the choice of basis); more precisely, this property depends on the choice of subspace $\Gamma_0(x)$ and does not depend on the choice of basis in $\Gamma_0(x)$ and also on the parametrization of the curve $\hat{x}$.

Using the introduced notions, we get the following reformulation of Theorem 13.2, which relates the concepts of Π-rigidity and Π-minimum.

Theorem 13.3. *Let a singular trajectory* $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = (1, 0, \dots, 0))$ *of Problem* (Z_*), *written in an associated basis, be quadratically* Π*-rigid. Then it is a point of the strict* Π*-minimum in Problem* (Z_*), *or, in other words, it is a point of the strict minimum with respect to the norm* $\|w\|_1$ *in Problem* (Z) *with any submetric from the sheaf determined by the given basis.*

In order to prove this theorem, we have covered the following path:

$$\text{Problem } (Z) \to \text{Problem } (Z_*) \to \text{Problem } (S) \to$$

$$\to \text{Problem } (S_{\frac{1}{2}}) \to \text{Problem } (S_1) \to \text{System } (R)\,.$$

Here we complete the procedure of direct application of the general quadratic sufficient conditions for the Π-minimum for singular trajectories to the problem on geodesics, and pass to the elimination of the assumption of singularity of $\hat{w}$.

14 Conditions for Π-Minimality of Quadratically Π-Rigid Trajectories

In this section, our goal is to get rid of the assumption of singularity in Theorems 13.1–13.3, i.e., to prove the following theorem.

Theorem 14.1 (Main Theorem 2). *Let the trajectory* $\hat{w} = (\hat{z} = 1, \hat{x}(t), \quad \hat{u} = (0, \dots, 0))$ *of System* (R), *written in an associated basis be quadratically* Π*-rigid, i.e., let it satisfy inequality (13.1) or (13.2) for some* $a > 0$. *Then it is a point of the strict* Π*-minimum in Problem* (Z_*) *for this basis, and, therefore, it is a point of the strict minimum with respect to* $\|w\|_1$ *in Problem* (Z) *with any submetric from the sheaf determined by this basis.*

To prove this theorem, we will cover the same path as in Secs. 11–13, but in the reverse order: from System (R) to Problem (Z_*), as was done in Part II. In contrast to the situation of Sec. 8, now the basis is fixed: it is one and the same in System (R) and in Problem (Z_*). In addition, since the second statement of Theorem 14.1 (about Problem (Z)) is a trivial consequence of the first one (about Problem (Z_*)), and Problem (Z_*) does not involve a submetric, the submetric in fact does not take part in the proof of this theorem at all.

Thus, let the trajectory $\hat{w}$ of System (R) satisfy conditions (13.1) or (13.2) for $a > 0$. As in Sec. 9, let us first pass to System (R'). As was shown, in this passage the subspace of critical variations expands, and the set $\Lambda(\hat{w})$ expands too. Lemma 9.1, which describes the set $\Lambda(R', \hat{w})$, is still valid. Moreover, this lemma implies the following property that will be useful for us.

Lemma 14.1. *Under the injection from Lemma* 9.1, *the set* $E_a(\Lambda(R'))$ *is, up to normalization, the convex hull of the set* $\pi((E_a(\Lambda(R))))$ *and of the point* λ_0. *(Obviously,* $\lambda_0 \in E_a(\Lambda(R'))$, *for it has* $\psi \equiv 0$.*)*

The passage to System (R') is ensured by the following theorem.

Theorem 14.2. *Let the quadratic* Π *-rigidity holds for the trajectory* $\hat{w}$ *in System* (R). *Then it holds also in System* (R').

The proof is similar to the proof of Theorem 9.1 with the only difference that the quadratic condition of Π-rigidity involves the maximum not over all $\lambda \in \Lambda$, but only over $\lambda \in E_0(\Lambda)$ (or over $E_a(\Lambda)$). However, this fact does not affect the validity of all considerations in the proof of Theorem 9.1.

Further, we pass to Problem (P_1), which is obtained by adding the functional $J = z(0) \to \min$ to System (R'). As was proved in Sec. 10, the subspace of critical variations and the set $\Lambda(\hat{w})$ do not change under this procedure; therefore, by analogy with Theorem 10.1, the following theorem holds.

Theorem 14.3. *Let the Pontryagin* γ *-sufficiency hold for the trajectory* $\hat{w}$ *in System* (R'). *Then it holds also in Problem* (P_1), *i.e., for some* $a > 0$, *we have*

$$\max_{E_a(\Lambda(P_1))} \Omega[\lambda](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall\, \bar{w} \in \mathcal{K}(P_1). \tag{14.1}$$

The next step consists in the passage from Problem (P_1) to Problem $(P_{\frac{1}{2}})$ in which the equation

$$\dot{x} = z\, r_0(x) + \sum_{i=1}^{k-1} u_i r_i(x) \tag{14.2}$$

is replaced by

$$\dot{x} = z\, u_0 r_0(x) + \sum_{i=1}^{k-1} u_i r_i(x), \tag{14.3}$$

and the control u_0 that has appeared is bounded by the inequalities $\frac{1}{2} \leq u_0 \leq 1$. By analogy with Theorem 12.2, the following theorem holds.

Theorem 14.4. *The* γ *-sufficient condition for the* Π *-minimum in Problem* (P_1) *is equivalent to the* γ *-sufficient condition for the* Π *-minimum in Problem* $(P_{\frac{1}{2}})$.

With account for Theorem 7.2 (ii) from [9], the proof is a consequence of the following lemma.

Lemma 14.2. *The fulfilment of the property* $\mathcal{F}$ *on the set* $\mathcal{D}(P_1)$ *is equivalent to its fulfilment on the set* $\mathcal{D}(P_{\frac{1}{2}})$.

The proof repeats the proof of Lemma 12.1 with the only difference that now one has to take $|\Delta x_M(T)| + |\Delta x_L(T)|^2$ instead of $|x(T) - b|$ in the violation function. Notice again that since the term $|x(T) - b|$ took only "passive" part in all considerations, nothing will change under such replacement. The detailed verification of this fact is left to the reader.

Then we pass from Problem $(P_{\frac{1}{2}})$ to Problem (P) with the control set $u_0 \leq 1$. Since this set and the set $\frac{1}{2} \leq u_0 \leq 1$ coincide in a neighborhood of the point $\hat{u}_0 = 1$, the tangent cone N and the set $\Lambda(\hat{w})$ will not change under such passage. Therefore (as in Sec. 12), the families of Lagrange functions, the families of second variations $\Omega[\lambda](\bar{w})$ of these functions, and the families of corresponding cubic functionals $\rho[\lambda](\bar{w})$ will not change as well. Thus, all the objects taking part in the formulation of the γ-sufficient conditions for the Π-minimum in both problems will be the same. Hence, by analogy with Theorem 12.1, the following theorem holds.

Theorem 14.5. *The* γ*-sufficient condition for the* Π*-minimum in Problem* $(P_{\frac{1}{2}})$ *is equivalent to the* γ*-sufficient condition for the* Π*-minimum in Problem* (P).

Finally, let us pass from Problem (P) to Problem (Y_*) (see Sec. 10), which is obtained as a result of replacement of the equation (14.3) by the initial equation (11.1) (in which the whole right-hand side, not only its first term, is multiplied by Z).

Lemma 14.3. *The validity of property* $\mathcal{F}$ *on the set* $\mathcal{D}(P)$ *is equivalent to its validity on the set* $\mathcal{D}(Y_*)$.

The proof is similar to the proof of Lemma 11.1, which, in turn, almost literally repeats the proof of Lemma 5.1.

This lemma and Theorem 7.2 (ii) from [9] yield the following theorem, which is similar to Theorem 11.2.

Theorem 14.6. *The* γ*-sufficient condition for the* Π*-minimum in Problem* (P) *is equivalent to the* γ*-sufficient condition for the* Π*-minimum in Problem* (Y_*).

Summing up the passages made, from Theorems 14.2–14.6, we obtain the following theorem.

Theorem 14.7. *Let the trajectory $\hat{w}$ in System (R) satisfy the quadratic Π-rigidity, i.e., let inequality (13.1) or (13.2) hold for some $a > 0$. Then $\hat{w}$ satisfies the γ-sufficient condition for the Π-minimum in Problem (Y_*), and hence is a point of the strict Π-minimum in Problem (Y_*).*

Now, as in Sec. 10, it remains to note that the strict Π-minimum in Problem (Y_*) coincides with the strict Π-minimum in Problem (Z_*), which differs from Problem (Y_*) only by the fact that the right boundary condition in it has the initial form $x(T) - b = 0$. Therefore, Theorem 14.7 implies Theorem 14.1; the proof of the latter one has been just the aim of this section.

Thus, here we have covered the following path:

$$\text{System } (R) \to \text{System } (R') \to \text{Problem } (P_1) \to \text{Problem } (P_{\frac{1}{2}}) \to$$

$$\to \text{Problem } (P) \to \text{Problem } (Y_*) \to \text{Problem } (Z_*)\,,$$

but now without assumption on the singularity of the trajectory $\hat{w}$ in Problem (Z_*).

Part IV
Special Cases, Examples and Proofs of Auxiliary Statements

15 Special Cases

It will be shown here into what transform the Main Theorems in two special cases, namely, when the distribution is two-dimensional and when the segment of the curve is sufficiently small.

15.1 Two-dimensional case.

Let $k = \dim \Gamma(x) = 2$. Then, in Problem (S_1) and System (R), the control is one-dimensional: $u = u_1 \in \mathbb{R}$. As is well known, the problems with one-dimensional control are essentially simpler than those with multidimensional control. In the case of one-dimensional control, condition (11.8) (and also condition (7.17)) is automatically satisfied; hence, $E_a(\Lambda) = G_a(\Lambda)$, and the γ-sufficient condition for the Π-minimum in Problems (Z_*), (S), and (S_1) coincides with the γ-sufficient condition for the weak minimum, while the quadratic rigidity in System (R) coincides with the quadratic Π-rigidity. But, since the quadratic rigidity of a trajectory is preserved in any associated basis, Theorem 14.1 implies the following theorem for the two-dimensional distribution:

Theorem 15.1. *Let a Γ-admissible curve $\hat{x}(t)$ connecting points a, b be quadratically rigid, i.e., let in some associated basis the trajectory $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = 0)$ of System (R) satisfy inequality (7.20) or (7.21) for some $a > 0$. Then, in any associated basis, the trajectory $\hat{w}$ is a point of the strict Π-minimum in Problem (Z_*), and hence, for any submetric that has a support (not necessarily strictly support) hyperplane in a neighborhood of $\hat{x}(t)$, the trajectory $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = (1, 0, \dots, 0))$ is a point of the strict minimum in Problem (Z) with respect to $\|w\|_1$.*

Let us compare this theorem with Theorem 8.1. In both these theorems, the hypothesis is one and the same. The conclusion of Theorem 15.1 differs from that of Theorem 8.1 by the fact that, on the one hand, no strictness of the support hyperplane is now required, but, on the other hand, only the minimality with respect to $\|w\|_1$ is guaranteed for the given trajectory, instead of the strong minimality.

Theorem 15.1 is still stronger than Theorem 5.2 from [1]: in the latter only the sub-Riemannian metric is admitted, and inequality (7.21) should be satisfied for an individual quadratic form for some $\psi \in G_a(\Psi_0))$.

15.2 Sufficient conditions for small segments of curve.

Let $\hat{x}(t)$, $t \in [0, T]$, be a Γ-admissible curve, and let $\hat{w} = (\hat{z} = 1, \hat{x}(t), \hat{u} = (1, 0, \dots, 0))$ be the corresponding trajectory of Problem (Z) in some associated basis. We will consider this problem and also the corresponding System (R) on any segment $\Delta = [t_1, t_2] \subset [0, T]$ with the endpoint conditions $x(t_1) = \hat{x}(t_1)$ and $x(t_2) = \hat{x}(t_2)$, i.e., we will study the question of minimality and of rigidity of the given segment of the curve in the set of all Γ-admissible curves connecting the initial and terminal points of this segment. The corresponding problem, system, and trajectory will all be marked by the subscript Δ.

For the trajectory $\hat{w}_\Delta$, it will be convenient for us to introduce the set

$$G_+(\Psi_0(\hat{w}_\Delta)) = \bigcup_{a>0} G_a(\Psi_0(\hat{w}_\Delta)). \tag{15.1}$$

First, let us give the following definition, which do not involve the submetric.

Definition 15.1. Following [15], we call the curve $\hat{x}(t)$ *quadratically rigid on small segments* if there exists $\delta > 0$ such that for any segment $\Delta \subset [0,T]$ of length $|\Delta| \le \delta$, the set $G_+(\Psi_0(\hat{w}_\Delta))$ is not empty, i.e, there exists an n-dimensional Lipschitz function $\psi(t)$ that satisfies the adjoint equation (7.1), relations (7.2) and (7.17), and inequality (7.18) for some $a = a(\Delta) > 0$ on Δ.

As is shown in [15, Sec. 6], the fulfilment of this "local" property does not imply even the stationarity of the trajectory $\hat{w}$ on the whole interval $[0,T]$, i.e., it does not imply the nonemptiness of $\Psi_0(\hat{w})$, not to mention the nonemptiness of $G_+(\Psi_0(\hat{w}))$. On the other hand, the following theorem is proved in [15, Sec. 5] and [9, Sec. 6].

Theorem 15.2. *Let the curve $\hat{x}(t)$ be quadratically rigid on small segments. Then there exists $\delta > 0$ such that for any $\Delta \subset [0,T]$, $|\Delta| \le \delta$, inequality (7.21) holds for some $a = a(\Delta) > 0$, i.e., the curve $\hat{x}_\Delta(t)$ is quadratically rigid.*

Thus, if the curve is quadratically rigid on small segments, then all its sufficiently small segments are indeed quadratically rigid, i.e., the sense of the above-introduced notion corresponds to its wording.

Now let some submetric be given. Let us write Problem (Z) in an associated basis for the given curve $\hat{x}(t)$, which is a support one for this submetric.

Definition 15.2. We say that a curve $\hat{x}(t)$ yields the *Pontryagin (strong) minimum in Problem (Z) on small segments* if there exists $\delta > 0$ such that, for any $\Delta \subset [0,T]$, $|\Delta| \le \delta$, the trajectory $\hat{w}_\Delta = (\hat{z} = 1, \hat{x}(t), \hat{u} = (1,0,\ldots,0))$ yields the Pontryagin (respectively, strong) minimum in Problem (Z_Δ).

Definition 15.3. We say that a curve $\hat{x}(t)$ yields the *global minimum of distance on small segments* if there exists $\delta > 0$ such that, for any $\Delta \subset [0,T]$, $|\Delta| \le \delta$, the curve $\hat{x}_\Delta$ is the shortest one among all Γ-admissible curves connecting the same initial and terminal points, i.e., the corresponding trajectory $\hat{w}_\Delta$ is a point of the global minimum in Problem (Z_Δ) with the given submetric.

Lemmas 8.1–8.3 from [9] imply the following lemma.

Lemma 15.1. *If the curve $\hat{x}(t)$ yields the strong (the strict strong) minimum of distance on small segments in Problem (Z), then it yields the global (the strict global) minimum of distance on small segments as well.*

From here, by Theorems 15.2 and 8.1, we obtain the following theorem.

Theorem 15.3 (sufficient condition for the global minimum on small segments). *Let a curve $\hat{x}(t)$ be quadratically rigid on small segments. Then, for any submetric on*

$\Gamma(x)$ having a strict support hyperplane in a neighborhood of $\hat{x}(t)$, this curve yields the strict global minimum of distance on small segments.

This theorem is stronger than the results of [13, Theorem 5] and [1, Corollary 5.2], since in both these papers: (a) only two-dimensional distributions and only sub-Riemannian metrics are admitted; (b) it is assumed that $G_+(\Psi_0(\hat{w}[0,T]))$ is nonempty, whereas in Theorem 15.3 the trajectory $\hat{w}$ can be even nonstationary on the whole interval $[0,T]$.

Let us pass to the consideration of Theorem 14.1 on the Pontryagin minimum. Recall that the Pontryagin minimum in Problem (Z) with any submetric is equivalent to the minimum with respect to $\|w\|_1$. First, we prove the following simple fact.

Lemma 15.2. *If the trajectory $\hat{w}$ yields the Π-minimum on small segments in Problem (Z), then it yields the strong minimum on small segments as well, and hence, by Lemma 15.1, it also yields the global minimum on small segments. The same is true for the strict minima.*

Proof. Let there exist $\delta > 0$ such that for any closed interval $\Delta \subset [0,T]$ of length $|\Delta| \le \delta$, there exists $\varepsilon = \varepsilon(\Delta) > 0$ such that for any admissible trajectory $w_\Delta = (z,x,u)$ of Problem (Z_Δ) satisfying the inequality

$$|z-1| + \|x-\hat{x}\|_C + \|u-\hat{u}\|_1 < 2\varepsilon, \tag{15.2}$$

the inequality $J(w_\Delta) \ge J(\hat{w}_\Delta)$ holds. It follows from the compactness arguments that ε can be assumed to be common for all Δ, $|\Delta| \le \delta$.

Thus, there exist $\delta, \varepsilon > 0$ such that for any closed interval $\Delta \subset [0,T]$ of length $|\Delta| \le \delta$, for any admissible trajectory $w_\Delta = (z,x,u)$ of Problem (Z_Δ), satisfying inequality (15.2), we have $J(w_\Delta) \ge J(\hat{w}_\Delta)$.

Let a number K be such that $|U(x)| \le K$ for all x from the ε-neighborhood of the set $\hat{\chi}$. Then, for any admissible w_Δ satisfying the inequality $|z-1| + \|x-\hat{x}\|_C < \varepsilon$, we have $u(t) \in U(x(t))$, and hence, $\int_\Delta |u-\hat{u}|\,dt \le 2K\delta$; therefore, assuming $\delta < \varepsilon/(2K)$, we obtain that, for all such w_Δ, inequality (15.2) automatically holds, hence $J(w_\Delta) \ge J(\hat{w}_\Delta)$, and thus the strong minimality of w_Δ on small segments is established. The lemma is proved.

Similar to (15.1), let us now introduce the following set:

$$E_+(\Psi_0(\hat{w}_\Delta)) = \bigcup_{a>0} E_a(\Psi_0(\hat{w}_\Delta)), \tag{15.3}$$

whose nonemptiness, as was already noted, depends on the choice of associated basis, and let us give the following definition.

Definition 15.4. The curve $\hat{x}(t)$ is called *quadratically* Π*-rigid on small segments in a given associated basis* if there exists $\delta > 0$ such that for any segment $\Delta \subset [0, T]$ of length $|\Delta| \le \delta$, the set $E_+(\Psi_0(\hat{w}_\Delta))$ is nonempty, i.e., there exists an n-dimensional Lipschitzian function $\psi(t)$ satisfying the adjoint equation (7.1), relations (7.2), (7.17), and (11.8), and inequality (7.18) on Δ for some $a = a(\Delta) > 0$.

Lemma 15.2 and Theorem 14.1 imply the following theorem.

Theorem 15.4. *Let the trajectory $\hat{w}$ of System (R) written in some associated basis be quadratically Π-rigid on small segments. Then, for any submetric from the sheaf determined by this basis, this trajectory yields the strict global minimum in Problem (Z_Δ) for small Δ, i.e., it is strictly shortest between its endpoints on small segments.*

For the two-dimensional distribution, this theorem is stated as follows. (We again take into account the fact that for the two-dimensional $\Gamma(x)$, the quadratic Π-rigidity is equivalent to the "ordinary" quadratic rigidity, and the latter does not depend on the choice of associated basis.)

Theorem 15.5. *Let the trajectory $\hat{w}$ of System (R) be quadratically rigid on small segments. Then this trajectory yields the strict global minimum of distance on small segments for any submetric on $\Gamma(x)$, having a support hyperplane in a neighborhood of $\hat{x}(t)$.*

This theorem is somewhat stronger than Theorem 15.3 (but only for two-dimensional distributions), since the strictness of the support hyperplane is not required here. It follows also from Theorem 15.1 and Lemma 15.2.

16 Examples

Here we consider several examples.

Example 1. Let the distribution in $\mathbb{R}^3$ be given by the following two vector fields:

$$r_0(x) = b(x_1)\frac{\partial}{\partial x_2} + c(x_1)\frac{\partial}{\partial x_3} \qquad \text{and} \qquad r_1(x) = \frac{\partial}{\partial x_1},$$

where b and c are twice smooth functions of the scalar argument, satisfying the conditions

$$b(0) = 1, \qquad c(0) = c'(0) = 0, \qquad c''(0) \neq 0. \tag{16.1}$$

The corresponding control system has the form $\quad \dot{x} = u_0\, r_0(x) + u_1\, r_1(x),$ i.e.,

$$\dot{x}_1 = u_1, \qquad \dot{x}_2 = b(x_1)\, u_0, \qquad \dot{x}_3 = c(x_1)\, u_0.$$

We will study the trajectory $\hat{x}(t) = (0,t,0),\;\; \hat{u}(t) = (1,0),\;\; t \in [0,T]$, with an arbitrary fixed $T > 0$. This trajectory connects the points $(0,0,0)$ and $(0,T,0)$.

(This example is a generalization of examples from [16, 13, 19]. The example from [16] corresponds to $b \equiv 1,\;\; c = x_1^2$; the example from [13] to $b = 1 - x_1,\;\; c = x_1^2$, and the example from [19] to $c(x_1) = b(x_1) x_1^2$ with some specific function b. In all these papers, the sub-Riemannian metric, in which the given basis is orthonormal, is considered.)

The basis $r_0,\, r_1$ is associated for $\hat{x}(t)$, and the corresponding System (R) has the following form:

$$\dot{x}_1 = u_1, \qquad \dot{x}_2 = z\, b(x_1), \qquad \dot{x}_3 = z\, c(x_1), \qquad \dot{z} = 0,$$

$$x(0) = (0,0,0), \qquad x(T) = (0,T,0).$$

The control in this system is one-dimensional; $\hat{u}_1 = 0,\;\; \hat{z} = 1$. Here $H = \psi_1 u_1 + \psi_2 z b(x_1) + \psi_3 z c(x_1)$, and, by definition, the set Ψ_0 consists of all 3-dimensional functions $\psi(t)$ normalized by $|\psi(0)| = 1$, satisfying the adjoint system (7.1), i.e.,

$$\dot{\psi}_1 = -\psi_2\, b'(0) - \psi_3\, c'(0), \qquad \dot{\psi}_2 = 0, \qquad \dot{\psi}_3 = 0,$$

and relations (7.2), i.e., $\psi_1 = 0,\quad \psi_2 b(0) + \psi_3 c(0) = 0$.

In view of (16.1), this implies that Ψ_0 consists of two constant vectors $\psi = \pm(0,0,1)$. For each of these vectors, the Lagrange function (7.10) is of the form

$$\Phi[\psi](z,x,u) = \psi_3 \left(x_3(0) - x_3(T) + \int_0^T (\dot{x}_3 - z c(x_1))\, dt \right),$$

and the second variation of this function is

$$\Omega\,[\psi](\bar{z},\bar{x},\bar{u}) \;=\; -\psi_3 \int_0^T c''(0)\, \bar{x}_1^2\, dt.$$

The subspace of critical variations is given by the linearization of all constraints of System (R). Among the obtained relations we will have the equation $\dot{\bar{x}}_1 = \bar{u}_1$, $\bar{x}_1(0) = 0$. Since the Goh variable $\bar{y} = \bar{y}_1 \in \mathbb{R}^1$ satisfies the same equation $\dot{\bar{y}}_1 = \bar{u}_1$, $\bar{y}_1(0) = 0$, we have $\bar{x}_1 = \bar{y}_1$, and, choosing $\psi_3 = \operatorname{sign} c''(0)$, we obtain

$$\Omega\,[\psi](\bar{z},\bar{x},\bar{u}) \;\geq\; |c''(0)| \cdot \int_0^T |\bar{y}|^2\, dt,$$

whence inequality (7.20) is satisfied.

Then, from Theorem 8.1 we conclude that, for any T and any submetric having a strict support hyperplane in a neighborhood of $\hat{x}(t)$ (in particular, for any sub-Riemannian metric), the trajectory $\hat{x}(t)$ yields the strict strong minimum, and any sufficiently small segment of this trajectory yields the strict global minimum of distance between the endpoints of this segment. If the submetric has only a nonstrict support hyperplane, then, by Theorem 15.1, the trajectory $\hat{w}$ yields the strict minimum with respect to $\|w\|_1$, and any sufficiently small segment of this trajectory still yields the strict global minimum by Theorem 15.5. (In [16, 13, 19], for the corresponding examples, the strict global minimality is established only for small segments and only for sub-Riemmanian metrics.)

Notice that if $b'(0) = 0$, then the trajectory $\hat{x}$ can be nonsingular for some submetrics in the corresponding Problems (Z) and (S_1). For example, if one takes the standard Euclidean metric $\varphi(x,u) = |u|$ in the above basis (i.e., if one assumes that this basis is orthonormal), then this basis will be a support one for this metric (the hodograph $|u| \le 1$ is contained in the halfspace $u_0 \le 1$), and in the corresponding Problem (Z), or, what is equivalent, in Problem (S_1), which is obtained by adding the functional $J = z(0) \to \min$ to System (R), the maximum principle for the given trajectory will be fulfilled also with the vector $\psi = (0,1,0)$, which does not belong to $\Psi_0(\hat{w})$, since it is not orthogonal to $\Gamma(\hat{x})$. (For this vector, $H(\hat{w}) = \psi_2 \hat{z}\, b(0) = 1 > 0$.) However, by Theorems 8.1, 15.1, and 15.5, this fact does not affect the validity of the statements on the minimality of $\hat{w}$.

Example 2. In [13], the proof of minimality of a certain class of abnormal trajectories is reduced (with the help of some special transformations) to the study of the following system, which the authors of that paper call the "normal form":

$$\dot{x}_1 = u\, a(x) + v\, x_1 b_1(x),$$

$$\dot{x}_2 = v\,(1 + x_1 b_2(x)),$$

$$\dot{x}_i = v\, x_1 b_i(x), \qquad i = 3, \ldots, n.$$

The examined trajectory is $\hat{u} \equiv 0$, $\hat{v} \equiv 1$, $x(0) = (0, x_2^0, 0, \ldots, 0)$, i.e., $\hat{x}(t) = (0,\, x_2^0 + t,\, 0, \ldots, 0)$.

It is assumed that $b_3(x) = x_1\eta_1(x) + x_3\eta_3(x) + \ldots + x_n\eta_n(x)$, all functions a, b_i, and η_i are smooth, and $\eta_1(\hat{x}(t)) \neq 0 \quad \forall\, t$.

The assumptions on the coefficient b_3 imply that

$$b_3(\hat{x}(t)) = 0, \qquad \frac{\partial b_3}{\partial x_1}(\hat{x}(t)) \neq 0 \qquad \forall\, t. \tag{16.2}$$

(In essence, those assumptions are equivalent to these conditions.)

In [13], the sub-Riemannian metric with the unit ball $u^2+v^2 \le 1$ was considered, and it was proved that the sufficiently small segments of the given trajectory yield the strict global minimum with respect to this submetric.

Here $u = u_1$ and $v = u_0$ in our notations, and the basis in which the system is written is associated for the trajectory $\hat{x}(t)$. We put $a(x) \equiv 1$ (or, which is the same, we introduce a new control $u_1 = a(x)u$), fix any interval $[0,T]$, and show that the set $G_+(\Psi_0)$ is not empty for the given trajectory; hence, this trajectory is quadratically rigid on small segments. From here, by Theorem 15.5, it follows that $\hat{x}$ yields the strict global minimum on small segments with respect to any submetric having twice smooth support hyperplane in a neighborhood of $\hat{x}(t)$, in particular, with respect to any sub-Riemannian metric, not necessarily having the given basis as orthonormal.

System (R) here has the following form (we make the change of variables $v \mapsto z$) :

$$\dot{x}_1 = u + z\, x_1 b_1(x), \qquad \dot{x}_2 = z\,(1 + x_1 b_2(x)), \qquad \dot{x}_i = z\, x_1 b_i(x), \qquad i = 3, \ldots, n.$$

The Pontryagin function is $H = u\,\psi_1 + z\,\psi_2 + z\,x_1 \sum_{i=1}^n \psi_i b_i(x)$. The set Ψ_0 consists of all n-dimensional functions $\psi(t)$ normalized by $|\psi(0)| = 1$ satisfying the adjoint system (7.1):

$$\dot{\psi}_1 = \; -\sum_{i=1}^{n} \psi_i b_i(\hat{x}), \tag{16.3}$$

$$\dot{\psi}_2 = 0, \qquad \dot{\psi}_3 = 0, \quad i = 3, \ldots, n,$$

and the orthogonality conditions (7.2):

$$u\,\psi_1 + v\,\psi_2 + v\,\hat{x}_1 \sum_{i=1}^{n} \psi_i b_i(\hat{x}) = 0 \qquad \forall\, u,\, v;$$

from which $\psi_1 = 0, \quad \psi_2 = 0$.

We will not describe the whole set Ψ_0; indicate only the following two constant vectors contained in this set: $\psi = \theta\,(0,0,1,0,\ldots,0)$, with $\theta = \pm 1$ (here the fulfilment of (16.3) is ensured by the relation $b_3(\hat{x}) = 0$). For each of these ψ the Lagrange function is

$$\Phi[\psi](z,x,u) = \; \theta \left(x_3(0) - x_3(T) + \int_0^T (\dot{x}_3 - z\,x_1 b_3(x))\, dt \right),$$

and the second variation of this function at the trajectory $\hat{x}(t)$ is

$$\Omega\,[\psi](\bar{z},\bar{x},\bar{u}) \; = \; -\theta \int_0^T \hat{z}\,\bar{x}_1 \Big(\bar{x}_1 \eta_1(\hat{x}) + \bar{x}_3 \eta_3(\hat{x}) + \ldots + \bar{x}_n \eta_n(\hat{x}) \Big)\, dt.$$

The subspace of critical variations is given by the relations

$$\dot{\bar{x}}_1 = \, \bar{u} \, + \, \bar{x}_1 b_1(\hat{x}), \qquad \dot{\bar{x}}_2 = \, \bar{z} \, + \, \bar{x}_1 b_2(\hat{x}),$$

$$\dot{\bar{x}}_i = \bar{x}_1 b_i(\hat{x}), \qquad i = 3, \ldots, n;$$

$$\bar{x}(0) = \bar{x}(T) = 0.$$

Passing to the Goh variables, i.e., taking

$$\bar{x}_1 = \bar{\xi}_1 + \bar{y}, \qquad \dot{\bar{y}} = \bar{u}, \quad \bar{y}_0 = 0,$$

$$\bar{x}_2 = \bar{\xi}_2, \qquad \bar{x}_i = \bar{\xi}_i, \quad i = 3, \ldots, n,$$

in the new variables we obtain

$$\Omega\,[\psi](\bar{z}, \bar{\xi}, \bar{y}) \;=\; -\theta \int_0^T \left((\bar{\xi}_1 + \bar{y})^2 \eta_1(\hat{x}) + (\bar{\xi}_1 + \bar{y}) \sum_{i=3}^n \bar{\xi}_i\, \eta_i(\hat{x}) \right) dt, \tag{16.4}$$

$$\text{where} \quad \dot{\bar{\xi}}_i = (\bar{\xi}_1 + \bar{y})\, b_i(\hat{x}), \qquad \bar{\xi}_i(0) = \bar{\xi}_i(T) = 0, \quad i = 1; \quad i = 3, \ldots, n. \tag{16.5}$$

One can see from the above that the coefficient $Q[\lambda] \;=\; \psi\,[[r_1, r_0], r_1]$ in (7.12) and (7.16), which stands by $\bar{y}^2$, is equal to $-\theta \eta_1(\hat{x}(t))$. Since, by the assumption, $\eta_1(\hat{x}(t)) \neq 0$ (i.e., $\frac{\partial b_3}{\partial x_1}(\hat{x}(t)) \neq 0$), then choosing $\theta = -\operatorname{sign} \eta_1(\hat{x})$, we obtain that inequality (7.18) holds with some $a > 0$, i.e., $\psi \in G_+(\Psi_0)$. (But, obviously, $-\psi \notin G_+(\Psi_0)$.) Thus, the set $G_+(\Psi_0)$ for this trajectory is indeed nonempty, which is the required result.

Moreover, one can find, for the functional (16.4) and for the above θ, the exact bound of those T for which this functional will be positive definite. To this end, taking $\bar{y}$ as a new control and taking $\bar{\xi}_i$, $i = 3, \ldots, n$, as new state variables, we obtain an ordinary quadratic functional of CCV, and we must find the conjugate point of this functional. However, in this case, one can take $\bar{x}_1$ as a new control, since $\bar{u}$ enters the functional (16.4) and system (16.5) only through $\bar{x}_1$. One can show that $\int_0^T |\bar{x}_1|^2\, dt \simeq \int_0^T |\bar{y}|^2\, dt$ (see estimate (9.11)); therefore, the positive definiteness of Ω with respect to $\|\bar{y}\|_2^2$ coincides with its positive definiteness with respect to $\|\bar{x}_1\|_2^2$.

Denote $\bar{x}_1 = \bar{v}$ and set, for brevity, $\eta_i(\hat{x}(t)) = \eta_i(t)$, $b_i(\hat{x}(t)) = b_i(t)$. Assuming, without loss of generality, that $\eta_1(t) > 0$, we have

$$\Omega = \int_0^T \left(\bar{v}^2 \eta_1(t) + \bar{v} \sum_{i=3}^n \bar{\xi}_i\, \eta_i(t) \right) dt,$$

$$\text{where} \qquad \dot{\bar{\xi}}_i = b_i(t)\, \bar{v}, \qquad \bar{\xi}_i(0) = \bar{\xi}_i(T) = 0, \quad i = 3, \ldots, n.$$

The conjugate point T_0 is the minimal value of T for which there exists a nonzero stationary trajectory of this functional on the interval $[0, T]$, i.e., a nontrival solution to the Euler–Lagrange–Jacobi equation for this functional. Let us write this equation. Here we have

$$H = \sum_{i=3}^n \bar{\psi}_i b_i(t) \bar{v} - \eta_1(t) \bar{v}^2 - \bar{v} \sum_{i=3}^n \bar{\xi}_i\, \eta_i(t),$$

$$\dot{\bar{\psi}}_i \;= -H_{\bar{\xi}_i} = \; \bar{v}\eta_i(t),$$

$$H_{\bar{u}} = \; \sum \bar{\psi}_i b_i(t) - 2\eta_1(t)\bar{v} - \sum \bar{\xi}_i\, \eta_i(t) = 0,$$

from which, $\bar{v}$ is expressed through $\bar{\psi}_i$ and $\bar{\xi}_i$:

$$\bar{v} = \; \frac{1}{2\,\eta_1(t)}\left(\sum \bar{\psi}_i\beta_i(t) - \sum \bar{\xi}_i\,\eta_i(t)\right).$$

Taking into account this expression, we obtain the following closed system of differential equations for $\bar{\xi}$ and $\bar{\psi}$:

$$\dot{\bar{\xi}}_i = b_i(t)\bar{v}, \qquad \dot{\bar{\psi}}_i = \eta_i(t)\bar{v}, \qquad i = 3,\ldots,n. \tag{16.6}$$

We are interested in the solutions to this system with the following initial conditions:

$$\bar{\xi}_i(0) = 0, \qquad i = 3,\ldots,n. \tag{16.7}$$

Let $X(t)$ and $P(t)$ be $(n-2)\times(n-2)$ -matrices whose columns are a fundamental system of solutions $(\bar{\xi},\bar{\psi})$ to system (16.6), (16.7). Then T_0 is the first value of $t>0$ for which $\det X(t) = 0$.

If $T<T_0$, then Ω (i.e., functional (16.4) for $\theta=-1$) is positive definite; hence, condition (7.21) is satisfied; therefore, this trajectory is quadratically rigid, and, by Theorem 8.1, it yields the strict strong minimum for any submetric having a strict support hyperplane in a neighborhood of $\hat{x}(t)$, while, by Theorem 15.1, this trajectory yields the strict Π -minimum for any submetric having simply a support hyperplane.

If $T\ge T_0$, then functional (16.4) is not positive definite. Nevertheless, condition (7.21) can be satisfied, since it involves the maximum of $\Omega[\psi]$ over all $\psi\in G_+(\Psi_0)$. For example, if not only the function $b_3(x)$, but also the function $b_4(x)$ satisfies condition (16.2), then, by analogy with preceding, the set Ψ_0 contains also two constant vectors $\psi=\theta\,(0,0,0,1,0,\ldots,0)$, $\theta=\pm1$; hence, for $\theta=-\operatorname{sign}\frac{\partial b_4}{\partial x_1}(\hat{x}(t))$, we will again obtain $\psi\in G_+(\Psi_0)$, and it is possible to find the conjugate point T_0' for the corresponding quadratic form. Then condition (7.21) is certainly satisfied for all $T<\max(T_0,T_0')$. However, this value can still not be the exact bound of those T for which this condition is satisfied. In order to determine this bound, one has to describe the whole set $G_+(\Psi_0)$ and find the "conjugate point" for the functional (7.21). The Jacobi theory for functionals of such kind is presented in [4].

Example 3 is due to A. A. Milyutin. In the space $\mathbb{R}^5$ with the coordinates $x=(x_0,x_1,x_2,x_3,x_4)$, let us consider a three-dimensional distribution $\Gamma(x)$ generated by the vector fields $r_0(x) = e_0$, $r_1(x) = A_1 x$, $r_2(x) = A_2 x$, where $e_0 = (1,0,0,0,0)$ is a base vector, and the matrices A_1, A_2 act as follows:

$$A_1 x = (0,\ x_0 + x_3,\ 0,\ x_1 + x_3,\ x_4)', \qquad A_2 x = (0,\ 0,\ x_0,\ x_3,\ x_2 + x_4)'$$

(here the prime denotes the passage from a row vector to the column vector), or, using the notation of the vector fields in the form of differential operators,

$$\begin{aligned} r_0(x) &= \tfrac{\partial}{\partial x_0} \quad \text{(a constant field)}\ , \\ r_1(x) &= (x_0 + x_3)\tfrac{\partial}{\partial x_1} + (x_1 + x_3)\tfrac{\partial}{\partial x_3} + x_4 \tfrac{\partial}{\partial x_4}, \\ r_2(x) &= x_0 \tfrac{\partial}{\partial x_2} + x_3 \tfrac{\partial}{\partial x_3} + (x_2 + x_4)\tfrac{\partial}{\partial x_4}. \end{aligned}$$

(It is not difficult to verify that $\forall\, x \quad \Gamma^3(x) = \mathbb{R}^5$, hence, Γ is bracket generating, although the dimension of $\Gamma(x)$ and $\Gamma^2(x)$ can be different at different points.) The following control system corresponds to this distribution:
$\dot x = u_0 e_0 + u_1 A_1 x + u_2 A_2 x$.

Consider the trajectory $\hat x_0 = t$, $\hat x_1 = \hat x_2 = 0$, $\hat x_3 = 1$, $\hat x_4 = -1$ on some interval $[t_0, T]$. It connects the points $x(t_0) = (t_0, 0, 0, 1, -1)$ and $x(T) = (T, 0, 0, 1, -1)$. Along this trajectory, $\forall\, t \quad \dim \Gamma(\hat x(t)) = 3$, $\dim \Gamma^2(\hat x(t)) = 4$. The given basis is, obviously, an associated one, and the corresponding System (R) has the form $\dot x = z\, e_0 + u_1 A x_1 + u_2 A_2 x$, $\dot z = 0$, i.e.,

$$\begin{gathered} \dot x_0 = z, \qquad \dot z = 0, \\ \dot x_1 = u_1(x_0 + x_3), \qquad \dot x_2 = u_2\, x_0, \\ \dot x_3 = u_1(x_1 + x_3) + u_2 x_3, \\ \dot x_4 = u_1 x_4 + u_2(x_2 + x_4). \end{gathered}$$

The control in this system is now two-dimensional. Let us find the set $\Psi_0(\hat w)$. Here the Pontryagin function is $\qquad H = z\psi_0 + u_1(\psi, A_1 x) + u_2(\psi, A_2 x) =$

$$= \psi_0 z + \psi_1 u_1(x_0 + x_3) + \psi_2 u_2 x_0 + \psi_3 u_1(x_1 + x_3) + \psi_3 u_2 x_3 + \psi_4 u_1 x_4 + \psi_4 u_2(x_2 + x_4).$$

The adjoint system has the form $\dot\psi = -\hat u_1 \psi A_1 - \hat u_2 \psi A_2 = 0$, i.e., $\psi = \text{const}$, and the condition $\psi \perp \Gamma(\hat x(t))$ means that $\psi_0 = 0$,

$$\begin{aligned} H_{u_1} &= \psi A_1 \hat x = \psi_1(\hat x_0 + \hat x_3) + \psi_3(\hat x_1 + \hat x_3) + \psi_4 \hat x_4 = 0, \\ H_{u_2} &= \psi A_2 \hat x = \psi_2 \hat x_0 + \psi_3 \hat x_3 + \psi_4(\hat x_2 + \hat x_4) = 0. \end{aligned}$$

For the given trajectory, the two last relations mean that $\forall\, t$

$$\psi_1(t+1)+\psi_3-\psi_4=0, \qquad \psi_2\, t+\psi_3-\psi_4=0,$$

whence $\psi_1 = \psi_2 = 0$, $\psi_3 = \psi_4$. Thus, up to normalization, the set Ψ_0 consists of the two constant vectors $\psi = \theta\,(0,0,0,1,1)$, $\theta = \pm 1$. For each of these vectors, the Lagrange function has the form

$$\Phi[\psi](z,x,u) = \theta\Big(x_3(t_0)-x_3(T)+x_4(t_0)-x_4(T)+$$
$$+\int_{t_0}^{T}(\dot x_3+\dot x_4)\,dt-\int_{t_0}^{T}\big(u_1(x_1+x_3)+u_2x_3+u_1x_4+u_2(x_2+x_4)\big)\,dt\Big),$$

and its second variation at the trajectory $\hat w(t)$ is

$$\Omega\,[\psi](\bar z,\bar x,\bar u) \;=\; -\theta\int_{t_0}^{T}\Big(\bar u_1(\bar x_1+\bar x_3+\bar x_4)+\bar u_2(\bar x_2+\bar x_3+\bar x_4)\Big)\,dt.$$

The subspace of critical variations is given by

$$\begin{gathered}\dot{\bar x}_0 = \bar z, \qquad \dot{\bar z}=0,\\ \dot{\bar x}_1 = \bar u_1\,(t+1), \qquad \dot{\bar x}_2 = \bar u_2\, t,\\ \dot{\bar x}_3 = \bar u_1+\bar u_2, \qquad \dot{\bar x}_4 = -\bar u_1-\bar u_2,\\ \bar x(t_0) = \bar x(T) = 0.\end{gathered}$$

Let us make the Goh transformation, i.e., let us set

$$\begin{gathered}\dot{\bar y}_1=\bar u_1, \quad \dot{\bar y}_2=\bar u_2, \quad \bar y_1(t_0)=\bar y_2(t_0)=0,\\ \bar z=\bar\xi_z, \qquad \bar x_0=\bar\xi_0,\\ \bar x_1=\bar\xi_1+\bar y_1(t+1), \qquad \bar x_2=\bar\xi_2+\bar y_2 t,\\ \bar x_3=\bar\xi_3+\bar y_1+\bar y_2, \qquad \bar x_4=\bar\xi_4-\bar y_1-\bar y_2.\end{gathered}$$

Then $\bar\xi$ is subjected to the equation

$$\dot{\bar\xi}=\bar z\,e_0-\bar y\,e_1-\bar y_2\,e_2, \qquad \bar\xi(t_0)=0.$$

Note that here $\bar\xi_3=\bar\xi_4=0$ and $\bar x_3+\bar x_4=0$.

In the new variables, we obtain

$$\Omega=-\theta\int_{t_0}^{T}\Big(\bar u_1(\bar\xi_1+(t+1)\bar y_1)\;+\;\bar u_2(\bar\xi_2+t\bar y_2)\Big)\,dt,$$

$$\left.\begin{aligned}\dot{\bar\xi}_1&=-\bar y_1, \qquad \bar\xi_1(t_0)=0,\\ \dot{\bar\xi}_2&=-\bar y_2, \qquad \bar\xi_2(t_0)=0.\end{aligned}\right\} \tag{16.8}$$

The following relations should hold at the right end:

$$\bar{x}_1(T) = \bar{\xi}_1(T) + (T+1)\,\bar{y}_1(T) = 0,$$
$$\bar{x}_2(T) = \bar{\xi}_2(T) + T\,\bar{y}_2(T) = 0,$$
$$\bar{x}_3(T) = \bar{y}_1(T) + \bar{y}_2(T) = 0.$$

Setting $\bar{y}_1(T) = \beta$, we obtain from the last relation that $\bar{y}_2(T) = -\beta$, while two preceding relations mean that

$$\bar{\xi}_1(T) = -(T+1)\beta, \qquad \bar{\xi}_2(T) = T\beta. \tag{16.9}$$

Integrating by parts, we obtain

$$-\theta\,\Omega = \left(\bar{y}_1\bar{\xi}_1 + \bar{y}_2\bar{\xi}_2\right)\Big|_T + \left(\tfrac{t+1}{2}\,\bar{y}_1^2 + \tfrac{t}{2}\,\bar{y}_2^2\right)\Big|_T - \int_{t_0}^{T}(-\bar{y}_1^2 + \frac{\bar{y}_1^2}{2} - \bar{y}_2^2 + \frac{\bar{y}_2^2}{2})\,dt =$$
$$= -\tfrac{T+1}{2}\,\bar{y}_1^2(T) - \tfrac{T}{2}\,\bar{y}_2^2(T) + \tfrac{1}{2}\int_{t_0}^{T}(\bar{y}_1^2 + \bar{y}_2^2)\,dt.$$

Thus, for $\theta = -1$, we have

$$2\,\Omega = -(2T+1)\,\beta^2 + \int_{t_0}^{T}(\bar{y}_1^2 + \bar{y}_1^2)\,dt, \tag{16.10}$$

moreover, according to (16.9), we also have

$$\begin{aligned} \int_{t_0}^{T}\bar{y}_1\,dt &= -\bar{\xi}_1(T) = -(T+1)\,\beta, \\ \int_{t_0}^{T}\bar{y}_2\,dt &= -\bar{\xi}_2(T) = T\beta. \end{aligned} \tag{16.11}$$

Here one can assume already that $\bar{y}_i \in L_2$ and consider them as new controls, while $\bar{\xi}_i$ can be considered as new state variables that are connected by relations (16.11) at the time instant T. From here one can see that the coefficient by $\bar{y}^2$ in the given quadratic form is $Q = I$ (the identity 2×2-matrix); therefore, $\psi \in G_{\frac{1}{2}}(\Psi_0)$, whereas for $\theta = +1$, the matrix $Q = -I$, and so ψ does not satisfy condition (7.18) even with $a = 0$.

Thus, the set $G_+(\Psi_0)$ consists of a unique constant vector $\psi = -(0,0,0,1,1)$, to which there corresponds the quadratic form (16.10) with relations (16.8) and (16.11). Therefore, the fulfillment of inequality (7.21) is equivalent to its fulfillment for the quadratic form (16.10).

Let us establish what are those t_0 and T for which this quadratic form satisfies (7.21) for some $a > 0$, i.e., for which it is positive definite. To this end, as is known from the CCV, it is necessary to find the conjugate point of this quadratic form. Having fixed the time instant T, one has to move t_0 (namely t_0, and not T, for $\bar{\xi}(t_0) = 0$), and find the maximum value of t_0 for which $\Omega(\bar{y}) \le 0$ for some nonzero function $\bar{y}$. (Since the strengthened Legendre condition with respect to the new control is fulfilled in our case, namely, $Q > 0$, we will always have that $\Omega(\bar{y}) > 0$ for t_0 close to T.) This can be done, for example, in the following way: having noticed that, for any fixed β, the functional Ω splits into the sum of functionals of $\bar{y}_1$ and $\bar{y}_2$, then find the minimum of Ω independently over $\bar{y}_1$ and over $\bar{y}_2$. Thereby, we will find $\min \Omega$ for the given β. It is easy to see that here we have

$$\bar{y}_1 = \text{const} \; = \frac{T+1}{\Delta}\beta, \qquad \bar{y}_2 = \text{const} \; = -\frac{T}{\Delta}\beta,$$

where $\Delta = T - t_0$, and then, for the given β,

$$\min \Omega = \; \Big(-(2T+1) + \frac{(T+1)^2}{\Delta} + \frac{T^2}{\Delta}\Big)\,\beta^2.$$

(Naturally, the obtained expression is quadratic in β.)

It is seen from here that Ω will be positive definite if and only if $(2T+1) < (2T^2+2T+1)/\Delta$. Therefore, if $2T+1 \le 0$, i.e., if $T \le -\frac{1}{2}$, then any $\Delta > 0$ fits, i.e., the given segment of the trajectory is quadratically rigid for any $t_0 < T$, and hence, by Theorem 8.1, this segment yields the strict strong minimum for any submetric that has a strict support hyperplane in a neighborhood of $\hat{x}(t)$. If, on the other hand, $2T+1 > 0$, i.e., if $T > -\frac{1}{2}$, then the given segment of the trajectory will be quadratically rigid and will give the strict strong minimum, respectively, only for $\Delta < \frac{2T^2+2T+1}{2T+1}$. For $\Delta > \frac{2T^2+2T+1}{2T+1}$, i.e., for sufficiently distant negative t_0, we obtain that $\Omega < 0$ (for some $\bar{y}$), and thus, according to [15], the quadratic *necessary* condition for rigidity is violated; therefore, this segment of the trajectory is not rigid. As for minimality of this segment, here we cannot state anything, having for the time being only the necessary conditions at our disposal.

It is interesting to note that the point $T_* = -\frac{1}{2}$, which plays some “critical” role in this example, is not a priori specified by anything on the examined trajectory.

17 Appendices

Appendix A. Proof of Lemma 2.1.

Intuitively, the validity of this lemma is absolutely obvious, but the formal proof calls for somewhat nontrivial constructions. Let us consider the case of nonstrict minima here.

It is clear that it is sufficient to prove the direct implication, i.e., to obtain the minimum in the "geometric sense" from the presence of the strong minimum in Problem (Z). (The inverse implication is obvious.)

From the very beginning, we restrict ourselves to the consideration of a sufficiently small neighborhood $\mathcal{O}(\hat{\chi})$ of the set $\hat{\chi}$. Consider the hodograph of the submetric $F(x) = \{\, \bar{x} \in \Gamma(x) \,|\, q(x,\bar{x}) \le 1\}$ and the mapping $p(x):\, \mathbb{R}^k \to \mathbb{R}^n, \quad u \mapsto \sum u_i r_i(x)$, for which $p(x)\,\mathbb{R}^k = \Gamma(x)$, and $p(x)\,U(x) = F(x)$.

By Assumption A2, we have a smooth hyperplane $\Gamma_0(x)$ in $\Gamma(x)$ that is support to the hodograph $F(x)$ at the point $r_0(x)$ on the set $\mathcal{O}(\hat{\chi})$. Let us extend it up to the smooth hyperplane $H(x)$ in the whole space $\mathbb{R}^n$ in such a way that $H(x) \bigcap \Gamma(x) = \Gamma_0(x)$ and $r_0(x) \notin H(x)$. Making a smooth change of coordinates in $\mathcal{O}(\hat{\chi})$, one can assume that the vector $r_0(x) = (1,0,\dots,0)$, the hyperplane $H(x)$ is constant and is given by the relation $x_1 = 0$, the set $F(x)$ is contained in the halfspace $x_1 \le 1$, and the trajectory $\hat{w}(t)$ has the form $\hat{z} = 1, \quad \hat{x}(t) = (t,0,\dots,0), \quad \hat{u}(t) = (1,0,\dots,0)$.

Assume that this trajectory yields the strong minimum in Problem (Z). This means, by definition, that the following property holds.

Property A. If w_m is a sequence of admissible trajectories, and if $|z_m - 1| + \|x_m - \hat{x}\|_C \to 0$, then $z_m \ge 1$ for all m sufficiently large.

By virtue of the specific character of Problem (Z), this implies the following (formally, more strong) property.

Property B. If admissible trajectories w_m are such that $\|x_m - \hat{x}\|_C \to 0$, then $z_m \ge 1$ for all m sufficiently large.

Indeed, assume that $z_m < 1$ for some subsequence. If $z_m \to 1$, then, by Property A we have $z_m \ge 1$, which contradicts the condition $z_m < 1$. If, on the other hand, $z_m \le \text{const} < 1$ for some subsequence, then there exists such a subsequence $\beta_m \ge 1$ that $z'_m = \beta_m\, z_m \to 1-0$. We set $u'_m = u_m/\beta_m$. Since $\beta_m \ge 1$, we still have $u'_m(t) \in U(x_m(t))$. Then the trajectory $w'_m = (z'_m, x_m, u'_m)$ with the same x_m will be admissible, but, as was just established, we have a contradiction for this trajectory: $z'_m < 1$ and at the same time, $z'_m \ge 1$. Thus, Property B is proved.

For any curve $x(t), \quad t \in [0,T]$, we set $\delta(x) = \max \operatorname{dist}(x(t), \hat{\chi})$. We have to prove that if $\delta(x_m) \to 0$, then $z_m \ge 1$ for large m. (All the time we speak about the admissible trajectories of Problem (Z).) Assume the contrary: there exists a sequence approaching to $\hat{\chi}$, i.e., $\delta(x_m) \to 0$, but $z_m < 1$. Let us show that in this case $\|x_m - \hat{x}\|_C \to 0$, and thereby, we obtain a contradiction with Property B. To this end, it is sufficient to prove the following lemma.

Lemma A.1. *Let $\delta(x_m) \to 0$ and $z_m \le 1 + o(1)$. Then $\|x_m - \hat{x}\|_C \to 0$.*

Proof. Let π be the projection of $\mathbb{R}^n$ onto the coordinate x^1 : $\pi(x) = x^1$ (the number of the coordinate here is written as the superscript, since the subscript indicates the number of the term of the sequence). Set $s_m(t) = \pi(x_m(t)) = x_m^1(t)$. Since $\delta(x_m) \to 0$, we have that $x_m(t)$ is uniformly close to $\hat{x}(s_m(t)) = (s_m(t), 0, \ldots, 0) = (x_m^1(t), 0, \ldots, 0)$. Since $\dot{x}_m \in z_m\, F(x_m(t))$ and $z_m \le 1 + o(1)$, and the set $F(x_m(t))$ is contained in the halfspace $x_1 \le 1$, we have

$$\dot{s}_m(t) = \dot{x}_m^1(t) \le 1 + \alpha_m, \qquad \text{where } \alpha_m \to 0.$$

Moreover, by virtue of the endpoint conditions $x_m(0) = \hat{x}(0)$ and $x_m(T) = \hat{x}(T)$, we have $s_m(0) = 0$ and $s_m(T) = T$. Let us show that $\|s_m(t) - t\|_C \to 0$. Then $\hat{x}(s_m(t))$ will be uniformly close to $\hat{x}(t)$ (due to its Lipschitz continuity) and, since $x_m(t)$ is uniformly close to $\hat{x}(s_m(t))$, we will have that $x_m(t)$ is uniformly close to $\hat{x}(t)$, and thus Lemma A.1 will be proved. (Notice that generally the curve $\hat{x}(s_m(t))$ is not admissible; it is only an intermediate point for the estimate of the distance between $x_m(t)$ and $\hat{x}(t)$.)

Thus, the whole matter is reduced to the proof of the following property of functions of one variable.

Lemma A.2. *Let absolutely continuous (Lipschitz continuous in our case) functions* $s_m : [0,T] \to \mathbb{R}$ *be such that* $s_m(0) = 0$, $s_m(T) = T$, *and* $\dot{s}_m(t) \le 1 + \alpha_m$ *almost everywhere, where* $\alpha_m \to 0$. *Then* $\|s_m(t) - t\|_C \to 0$.

Proof. Making change of variables $v_m(t) = s_m(t) - t$, we have $v_m(0) = 0$, $v_m(T) = 0$, and $\dot{v}_m(t) \le \alpha_m \to 0$ almost everywhere, and it is required to prove that $\|v_m\|_C \to 0$.

From the presence of the upper bound for $\dot{v}_m(t)$, it obviously follows that it is sufficient to prove the convergence $v_m(t_*) \to 0$ for each t_*. Since always

$$v_m(t_*) = \int_0^{t_*} \dot{v}_m(t)\, dt \le \alpha_m\, t_* \to 0,$$

then, in the case where there is no convergence to zero, we have $\liminf v_m(t_*) < 0$, or, passing to a subsequence, $v_m(t_*) \le -h < 0$. But since the inequality

$$v_m(T) - v_m(t_*) \le \int_{t_*}^{T} \alpha_m\, dt \to 0,$$

holds on the interval $[t_*, T]$ as well, summing it with the preceding inequality, we obtain that $v_m(T) \le -h + o(1) < 0$, which contradicts the condition $v_m(T) = 0$. Lemma A.2 is proved, and thus, Lemma A.1 is proved, and, hence, Lemma 2.1 for nonstrict minima is also proved.

Now it is easy to prove this lemma also for the strict minima. We leave this as an exercise to the reader.

Appendix B

Let $\mathcal{X}$ be an arbitrary metric compact set, and $\varphi : \mathcal{X} \times \mathbb{R}^k \to \mathbb{R}$ be a continuous function, which, $\forall\, x \in \mathcal{X}$, is a positive sublinear function in $u \in \mathbb{R}^k$.

Lemma B.1. *For each $x \in \mathcal{X}$, the set $U(x) = \{u \,|\, \varphi(x,u) \le 1\}$ is a solid compact set, which continuously depends on x in the Hausdorff metric.*

Proof. The closedness of $U(x)$ follows from the continuity of φ in u, while the boundedness of this set follows from the sublinearity and from the positivity of φ in u; therefore the compactness of $U(x)$ is proved. Since $\varphi(x,0) = 0$ and φ is continuous in u, we have $0 \in \operatorname{int} U(x)$. (All these are well-known facts of convex analysis.)

Now let $x_n \to x_0$. We have to prove that $U(x_n) \to U(x_0)$ in the Hausdorff metric. For the compact sets, this is equivalent to the fulfillment of the following two properties: (a) if $u_n \in U(x_n)$, $u_n \to u_0$, then $u_0 \in U(x_0)$, and (b) if $u_0 \in U(x_0)$, then there exist $u_n \in U(x_n)$ such that $u_n \to u_0$.

In the case (a), $u_n \in U(x_n)$ means that $\varphi(x_n, u_n) \le 1$, and then, by virtue of the continuity of φ, we obtain in the limit, $\varphi(x_0, u_0) \le 1$, i.e., indeed $u_0 \in U(x_0)$.

In the case (b), we have $\varphi(x_0, u_0) \le 1$; therefore, $\varphi(x_n, u_0) \le \alpha_n \to 1$, whence, due to the sublinearity of φ, we obtain $\varphi(x_n, \frac{u_0}{\alpha_n}) \le 1$, i.e., $u_n = \frac{u_0}{\alpha_n} \in U(x_n)$, and, $u_n \to u_0$. Lemma is proved.

Lemma B.2. *Let $U(x)$ be a family of compact sets in $\mathbb{R}^k$, continuously depending, in the Hausdorff metric, on a parameter x from a metric compact set $\mathcal{X}$. Then the union of these compact sets $U' = \bigcup_x U(x)$ is also a compact set.*

Proof. The continuity of $U(x)$ obviously implies that the function $|U(x)| = \max\{\,|u| \,:\, u \in U(x)\,\}$ is continuous, then, on the compact set $\mathcal{X}$, it is bounded from above by a number ρ. This means that each $U(x)$ is contained in the ball of radius ρ; therefore, the set U' is also contained in this ball, i.e., is bounded. Let us prove its closedness. Let $u_n \in U'$, $u_n \to u_0$. Then $u_n \in U(x_n)$ for some $x_n \in \mathcal{X}$. By compactness of $\mathcal{X}$, we have $x_{n_m} \to x_0 \in \mathcal{X}$ for some subsequence. But then, due to the continuity of $U(x)$, we obtain $u_0 = \lim u_{n_m} \in U(x_0)$; therefore $u_0 \in U'$. Lemma is proved.

Lemma B.3. *Let, the under conditions of the preceeding lemma, all compact sets $U(x)$ be convex, and let the point $\hat{u} \in \mathbb{R}^k$ be such that $\forall\, x \in \mathcal{X}$, it is the unique maximum point of the linear function $l(u) = u_0$ on $U(x)$. Then there exists a convex solid compact set $\tilde{U}$ containing all $U(x)$, $x \in \mathcal{X}$, for which $\hat{u}$ is also the unique maximum point of the function $l(u)$ on $\tilde{U}$.*

Proof. One can always assume that at least one of the compact sets $U(x)$ is solid. (If not, then we add the convex hull of any solid compact set lying in the halfspace

$l(u) < l(\hat u)$ and of the point $\hat u$ to this family.) We set $U' = \bigcup_x U(x)$. By Lemma B2, it is a compact set; moreover, it is solid. It is clear that $\hat u$ is still a unique maximum point of the function $l(u)$ on this set. Let us take its convex hull $\tilde U = \mathrm{co}\, U'$. As is known, this hull is also compact. Let us show that $\hat u$ is a unique maximum point of $l(u)$ on it. Indeed, assume that there exists $u \in \tilde U$, $u \neq \hat u$, for which $l(u) \geq l(\hat u)$. Then, by definition, $u = \sum_{i=1}^{m} \alpha_i u_i'$, where all $u_i' \in U'$, $\alpha_i > 0$, and $\sum \alpha_i = 1$. Since here $l(u_i') \leq l(\hat u)$ $\forall\, i$, the inequality $l(u) \geq l(\hat u)$ is fulfilled only in the case when $\forall\, i$ $l(u_i') = l(\hat u)$, i.e., when all $u_i' = \hat u$. But then their convex combination $u = \hat u$, which contradicts the assumption that $u \neq \hat u$. Lemma is proved.

(Obviously, all these three lemmas remain valid for an arbitrary compact set $\mathcal{X}$; one should just replace the corresponding sequences by generalized sequences.)

In the context of Sec. 4, one can take as the compact set $\mathcal{X}$ the closure of any sufficiently small neighborhood of the set $\hat\chi$.

Appendix C. Proof of formula (7.11)

From (7.10), the expression for Ω in the initial variables directly follows:

$$\Omega[\psi](\bar z, \bar x, \bar u) = -\int_0^T \Big(\bar z\, \psi(r_0'\, \bar x) + \frac{1}{2}\, \psi(r_0''\, \bar x, \bar x) + \sum_{i=1}^{k-1} \bar u_i\, \psi(r_i'\, \bar x)\Big)\, dt. \tag{17.1}$$

Substituting $\bar x = \bar\xi + \sum \bar y_j\, r_j(\hat x)$ into it and removing the parentheses, we obtain in the new variables:

$$\Omega[\psi](\bar z, \bar\xi, \bar y, \bar u) = -\int_0^T \Big(\bar z\, \psi(r_0'\bar\xi) + \bar z \sum \bar y_j\, \psi(r_0'\, r_j) + \frac{1}{2}\, \psi(r_0''\, \bar\xi, \bar\xi) + \tag{17.2}$$

$$+ \sum \bar y_j\, \psi(r_0''\, \bar\xi, r_j) + \frac{1}{2} \sum_{ij} \bar y_i \bar y_j\, \psi(r_0''\, r_i, r_j) + \sum \bar u_i\, \psi(r_i'\, \bar\xi) + \sum_{ij} \bar u_i \bar y_j\, \psi(r_i'\, r_j)\Big)\, dt.$$

Let us try to transform the last two terms in such a way that $\bar u_i$ will be excluded, if possible. To this end, let us integrate these terms by parts taking into account that $\bar u_i dt = d\bar y_i$. To begin with, notice that for any symmetric absolutely continuous matrix $S(t)$, we have $(S\bar y, \bar y)^{\cdot} = (\dot S \bar y, \bar y) + 2(S\bar y, \bar u)$; therefore,

$$\int_0^T (S(t)\bar y, \bar u)\, dt = \frac{1}{2}\, (S\bar y, \bar y)\Big|_T - \frac{1}{2} \int_0^T (\dot S(t)\bar y, \bar y)\, dt.$$

For an arbitrary absolutely continuous matrix $C(t)$ we have $C = S + V$, where $S = \frac{1}{2}(C + C^*)$ is the symmetrical part of this matrix and $V = \frac{1}{2}(C - C^*)$ is its skew-symmetrical part. Then, $(S\bar y, \bar y) = (C\bar y, \bar y)$, and, therefore,

$$\int_0^T (C(t)\bar y, \bar u)\, dt = \frac{1}{2}\, (C\bar y, \bar y)\Big|_T - \frac{1}{2} \int_0^T (\dot C \bar y, \bar y)\, dt + \frac{1}{2} \int_0^T ((C - C^*)\bar y, \bar u)\, dt.$$

Then the last term in (17.2) can be written in the form

$$-\int_0^T \sum_{ij} \bar u_i \bar y_j\, \psi(r_i' r_j)\, dt = -\frac{1}{2}\, \psi(r_i' r_j)\, \bar y_i \bar y_j \Big|_T +$$

$$+\int_0^T \frac{1}{2} \sum \big(\psi\,(r_i' r_j)\big)^{\cdot}\, \bar y_i \bar y_j\, dt - \int_0^T \frac{1}{2} \sum \psi(r_i' r_j - r_j' r_i)\, \bar u_i \bar y_j\, dt.$$

Opening the total derivative with respect to t in the middle term by the formula

$$\frac{d}{dt}\,\varphi(x)\Big|_{\hat x(t)} = \varphi'(\hat x)\,\frac{d\hat x}{dt} = \varphi'(\hat x)\, r_0(\hat x),$$

and using the Lie brackets $[f, g] = f'g - g'f$, we obtain the following expression for the last term of (17.2):

$$-\frac{1}{2}\,\psi\,(r_i' r_j)\, \bar y_i \bar y_j\Big|_T + \int_0^T \frac{1}{2} \sum \psi\,(-r_0' r_i' r_j + r_i'' r_0 r_j + r_i' r_j' r_0)\, \bar y_i \bar y_j\ dt- \tag{17.3}$$

$$-\int_0^T \frac{1}{2} \sum \psi\,[r_i, r_j]\, \bar u_i \bar y_j\, dt$$

Now let us integrate by parts the next to the last term of (17.2), taking into account Eq. (7.8) for $\bar\xi$:

$$\int_0^T - \sum \bar u_i\ \psi(r_i'\,\bar\xi)\, dt = -\sum \bar y_i\, \psi(r_i'\,\bar\xi)\Big|_T -$$

$$-\int \sum \bar y_i\, \psi(r_0' r_i'\,\bar\xi)\, dt\ +\ \int \sum \bar y_i\, \psi(r_i''\, r_0, \bar\xi)\, dt\ +$$

$$+\int \sum_i \bar y_i\, \psi r_i'\, (\bar z\, r_0 + r_0'\,\bar\xi + \sum_j \bar y_j\, [r_0, r_j])\, dt\ . \tag{17.4}$$

Now let us collect the similar terms in (17.2)–(17.4).

The terms of the type $\bar z\,\bar y$ are present in (17.2) and in (17.4):

$$\int \sum_i \big(-\bar z\, \bar y_i\, \psi(r_0' r_i) + \bar z\, \bar y_i\, \psi(r_i' r_0)\big)\, dt = \int \sum_i \bar z\, \bar y_i\, \psi\,[r_i, r_0]\, dt. \tag{17.5}$$

But it is easy to see that the following holds along $\hat x(t)$:

$$\psi\,[r_i, r_0] = -\frac{d}{dt}\ (\psi, r_i) = 0, \qquad \text{since, by virtue of the MP,} \quad \psi r_i(\hat x) = 0; \tag{17.6}$$

therefore, the term (17.5) disappears.

The terms of the type $\bar z\,\bar\xi$ and $\bar\xi\bar\xi$ are present only in (17.2); they are

$$\int \Big(-\bar z\, \psi(r_0'\,\bar\xi) - \frac{1}{2}\,\psi\,(r_0''\,\bar\xi, \bar\xi)\Big)\, dt. \tag{17.7}$$

The terms of the type $\bar\xi\,\bar y$ are present in (17.2) and in (17.4); they add up to

$$\int \sum_i \Big(-\bar y_i\,\psi(r_0''\,\bar\xi, r_i) - \bar y_i\,\psi(r_0' r_i'\,\bar\xi) + \bar y_i\,\psi(r_i''\,r_0, \bar\xi) + \bar y_i\,\psi(r_i' r_0'\,\bar\xi)\Big)\,dt. \tag{17.8}$$

Since the matrix of second derivatives of every component of every vector field $r_i(x)$ is symmetrical, then $\forall\,\bar\xi,\bar\eta \in \mathbb{R}^n \quad (r_i''\,\bar\xi, \bar\eta) = (r_i''\,\bar\eta, \bar\xi)$, therefore, we can interchange r_0 and $\bar\xi$ in the third term of the obtained relation, and then, having noticed that

$$\frac{d}{dx}\,[r_i, r_0]\,\bar\xi = \frac{d}{dx}\Big((r_i' r_0 - r_0' r_i)\,\bar\xi\Big) = r_i''\bar\xi\,r_0 + r_i' r_0'\,\bar\xi - r_0''\bar\xi\,r_i - r_0' r_i'\,\bar\xi,$$

we can write the whole expression (17.8) in the form

$$\int \sum_i \bar y_i\,\psi\,[r_i, r_0]'\bar\xi\,dt\ . \tag{17.9}$$

The terms of the type $\bar y\,\bar y$ from (17.2)–(17.4) add up to

$$\int \frac{1}{2}\,\sum \bar y_i \bar y_j\,\psi(-r_0'' r_i r_j - r_0' r_i' r_j + r_i'' r_0 r_j + r_i' r_j' r_0 + 2 r_i'\,[r_0, r_j])\,dt.$$

It is not difficult to verify that this value coincides with the following one:

$$\int \frac{1}{2}\,\sum_i \bar y_i \bar y_j\,\psi\,[r_i, [r_0, r_j]\,]\,dt. \tag{17.10}$$

To this end, one should open all Lie brackets and interchange subscripts i and j in some terms. This can be done, since the relation $\sum a_{ij}\bar y_i \bar y_j = \sum a_{ji}\bar y_i \bar y_j$ holds for any quadratic form.

Further, the term of the type $\bar y\,\bar u$ is present only in (17.3); it is equal to

$$\int \frac{1}{2}\,\sum_i \psi\,[r_i, r_j]\,\bar y_i \bar u_j\,dt \tag{17.11}$$

(here we interchanged some subscripts as well as the sign).

Finally, consider the terms outside the integral that have appeared due to the integration by parts. They are present in (17.3) and (17.4):

$$-\sum \frac{1}{2}\,\psi(r_i' r_j)\,\bar y_i \bar y_j\Big|_T - \sum \bar y_i\,\psi(r_i'\,\bar\xi)\Big|_T\ .$$

But, by virtue of the endpoint relation (7.9), $\bar\xi(T) = -\sum \bar y_j(T)\,r_j(\hat x(T))$; therefore, the term outside the integral is equal to

$$\frac{1}{2}\,\sum_i \psi\,(r_i' r_j)\,\bar y_i \bar y_j\Big|_T\ . \tag{17.12}$$

Summing up (17.7)–(17.12), we obtain formula (7.11).

Now let us consider the quadratic form $\Omega_S[\psi](\bar w)$ for Problem (S) corresponding to Eq. (11.4). The linearization (11.5) of this equation differs from the linearization (7.4) of Eq. (6.1) of Problem (S_1) only by the fact that now the sum $\sum \bar u_i r_i(\hat x)$ is taken over i not from 1, but from 0 to $k-1$. As was already said, two additional terms (11.7) will appear in the second variation of Lagrange function for Eq. (11.4) as compared with (7.11). The first of them is indeed a new one: $-\int \bar z \bar u_0\,(\psi\, r_0)\,dt$, and, since, by virtue of the MP, we have $\psi\, r_0(\hat x) = \text{const}$ $(=0$ for the singular trajectory), this term is equal to

$$-\,\bar z \bar y_0\,(\psi\, r_0)\Big|_T\,. \tag{17.13}$$

As for the second term, it corresponds to the fact that the sum in the last term of (17.1) should again be taken over $i = 0, 1, \ldots, k-1$. Then $\Omega_S[\psi]$ will have the form (7.11), in which all sums are taken over i, j beginning from 0, plus the additional term (17.13) outside the integral. Thus, finally, Ω_S will differ from (7.11) by the presence of the following additional terms (they correspond to $i = 0,\;\; j = 0)$:

$$\Big(-\bar z \bar y_0\,(\psi\, r_0) + \frac{1}{2}\,\bar y_0^2\,\psi(r_0' r_0) + \frac{1}{2}\sum_{i=1}^{k-1} \bar y_i \bar y_0\,\psi(r_i' r_0 + r_0' r_i)\Big)\Big|_T +$$

$$+\int_0^T \frac{1}{2}\sum_{i=1}^{k-1} \bar y_i \bar y_0\,\psi\,[\,[r_i, r_0], r_0]\,dt.$$

(Here we have used (17.6) and the fact that $[r_0(x), r_0(x)] = 0$ at any point x.)

Now let us define the set $G_a(\Lambda(S))$ taking into account that Problem (S) involves the pointwise constraint $u_0(t) \le 1$, and therefore, that Ω_S should be considered on the pointwise cone $N = \{\bar u \in \mathbb{R}^k \mid \bar u_0 \le 0\}$ (tangent to this constraint). For a quadratic form of the general form (7.12), under the presence of the constraint $\bar u(t) \in N$, this set, according to [7, 8, 9], consists of all $\lambda \in \Lambda(S)$ satisfying conditions (7.14) on the maximal subspace N_0 containing in N, and, moreover, satisfying the following condition:

$$(V[\lambda](t)\,\bar y, \bar u) = 0 \qquad \forall\, \bar y \in N_0,\;\; \bar u \in N. \tag{17.14}$$

For Problem (S), the subspace $N_0 = \{\bar u \mid \bar u_0 = 0\}$; therefore, conditions (7.14) mean that conditions (7.17) and (7.18) are satisfied, while the additional condition (17.14) means that condition (7.17) is satisfied also for $j = 0:$ $\;\psi(t)[r_i, r_0](\hat x(t)) = 0$. But, according to (17.6), this condition is automatically satisfied by virtue of the MP. Thus, for Problem (S), the set $G_a(\Lambda(S))$ is defined by the same conditions (7.17) and (7.18).

Appendix D

Proof of the estimate (9.11). We set $x = \xi + By$. From Eq. (9.10), we have $\dot{\xi} = A\xi + (AB - \dot{B})y$, and since $y(0) = 0$, we obtain $\xi(0) = x(0)$. Then $\|\xi\|_\infty \le \text{const}\,(|x(0)| + \|y\|_2) \le \text{const}\sqrt{\gamma(w)}$, and, therefore, $\|x\|_2 \le \|\xi\|_2 + \|By\|_2 \le \text{const}\sqrt{\gamma(w)}$. In addition, $|x(T)| \le |\xi(T)| + |By(T)| \le \text{const}\sqrt{\gamma(w)}$. Summing up these estimates, we obtain the estimate (9.11).

Proof of Lemma 9.2. This lemma practically repeats Lemma 6.4 from [5]; however, we will give its prove here for completeness of the presentation. Since (x_n, u_n) satisfy Eq. (9.10), then by virtue of estimate (9.11),

$$|x_n(0)|^2 + |x_n(T)|^2 + \int_0^T |x_n|^2\, dt \le \mathcal{O}(\gamma_n). \tag{17.15}$$

Then we have

$$\begin{aligned} \Omega(w_n') = {} & \Omega(w_n) + \Omega(\tilde{w}_n) + 2(Sp_n, \tilde{p}_n) + \\ & + 2\int_0^T \Big((Dx_n, \tilde{x}_n) + (x_n, C\tilde{u}_n) + (\tilde{x}_n, Cu_n)\Big)\, dt. \end{aligned} \tag{17.16}$$

Estimate (9.12) implies $\Omega(\tilde{w}_n) = o(\gamma_n)$. It follows from (17.15) that $|p_n| = |(x_n(0), x_n(T))| \le \mathcal{O}(\sqrt{\gamma_n})$, and (9.12) implies that $|\tilde{p}_n| \le o(\sqrt{\gamma_n})$; therefore, the mixed term outside the integral in (17.16) is also $o(\gamma_n)$. The first two integral mixed terms satisfy the following estimate (we write it out only for the second term):

$$\int \Big|(x_n, C\tilde{u}_n)\Big|\, dt \;\le \text{const}\; ||x_n||_2\, ||\tilde{u}_n||_2 \;=\; o\,(\gamma_n) \tag{17.17}$$

which is valid by virtue of the same estimates (17.15) and (9.12).

It remains to estimate the last integral term in (17.16), for which the direct estimate, similar to (17.17), is not valid. We will estimate this term using the principal method that is applied in the study of problems which are linear in control; namely, we will intergrate it by parts, bearing in mind that $u_n = \dot{y}_n$. We then obtain

$$\int_0^T (\tilde{x}_n, Cu_n)\, dt \;=\; (\tilde{x}_n, Cy_n)\Big|_0^T - \int_0^T \Big((\tilde{x}_n, \dot{C}y_n) + (\dot{\tilde{x}}_n, Cy_n)\Big)\; dt.$$

Substituting $\dot{\tilde{x}}_n = A\tilde{x}_n + B\tilde{u}_n$ in this relation and taking into account (9.12), we again obtain that the whole this value is $o(\gamma_n)$. Thus, Lemma 9.2 is proved.

Appendix E

Here we write out the cubic functional $\rho[\lambda](\bar w)$ and present the definition of the set $E_a(\Lambda)$ for Problem (S) from Sec. 11.

Assume that for the general system that is linear in control: $\dot x = f_0(x,t)+F(x,t)u$, where $u \in \mathbb{R}^k$ and F is a $n \times k-$ matrix, the following problem is considered: $g(p)=0,\ \ \varphi_i(p) \le 0,\ \ i=1,\dots,\nu,\ \ J=\varphi_0(p) \to \min,\ \ u \in U,$ where $p=(x_0,x_T)$, and U is a polyhedral set in $\mathbb{R}^k$, and let an examined trajectory $\hat w = (\hat x, \hat u)$ be given. We will assume that $\hat u(t)$ is continuous and $\forall\, t$ lies in the relative interior of one and the same face of the set U.

According to [5, 8, 9], we must consider the following expansion of the system equation at the trajectory $\hat w$ up to the quadratic term of the type $\bar x\, \bar u$:

$$\dot{\bar x} = A(t)\bar x + B(t)\bar u + (R(t)\bar x, \bar u) + \dots\,,$$

this expansion is determined by the matrices $A(t) = f_0'(\hat x(t),t) + F'(\hat x(t),t)\hat u(t)$, $B(t) = F(\hat x(t),t)$, and by the tensor $R(t) = F'(\hat x(t),t)$ (the prime denotes the derivative with respect to x), and for each $\lambda \in \Lambda$ we must determine the cubic functional

$$\rho[\lambda](\bar w) = \int_0^T \Big[-\Big(\frac{1}{2}\, H_{uxx}[\lambda]\bar x, \bar x, \bar u) + (H_{xu}[\lambda]\bar y, (R(t)\bar x, \bar u)\Big)\Big]\, dt.$$

Further, we must put here $\bar x = B(t)\bar y$, where, as always in this paper, $\dot{\bar y} = \bar u$, $\bar y(0)=0$; as a result, we obtain the functional

$$\eta[\lambda](\bar y) = \int_0^T \Big(\mathcal{E}[\lambda](t)\, \bar y, \bar y, \bar u\Big)\, dt, \tag{17.18}$$

where $\mathcal{E}[\lambda](t)$ is a third-rank tensor in the space $\mathbb{R}^k$. For each fixed t_*, this tensor determines the following differential 1-form:

$$\omega[\lambda](t_*) = \Big(\mathcal{E}[\lambda](t_*)\, \bar y, \bar y, d\bar y\Big).$$

Let N be the tangent cone to the set U at the point $\hat u(t)$. Due to the assumptions on $\hat u(t)$, this cone does not depend on t. Let N_0 be the maximal linear subspace that is contained in N. Then, according to [5], the set $E_a(\Lambda)$ consists of all $\lambda \in G_a(\Lambda)$ for which the 1-form $\omega[\lambda](t_*)$ is closed (or, equivalently, is exact, since everything takes place in $\mathbb{R}^k$) for every t_*: $d\omega[\lambda](t_*) = 0$. Here the differential is taken with respect to $\bar y$, while λ and t_* play the role of parameters.

Let us reveal what this condition means for Problem (S) in Sec. 11. The state variables here are z, x; the Pontryagin function is $H[\lambda](z,x,u) = zu_0(\psi, r_0(x)) + \sum u_j(\psi, r_j(x))$. The expansion of system (11.4) at the trajectory $\hat w$ has the form

$$\dot{\bar x} = \bar z r_0(\hat x) + r_0'(\hat x)\bar x + \sum_{i=0}^{k-1} \bar u_i r_i(\hat x) + \sum_{i=0}^{k-1} \bar u_i r_i'(\hat x)\bar x + \dots\,, \tag{17.19}$$

the above values $B(t)\bar{u}$ and $(R(t)\bar{x},\bar{u})$ are, respectively, the next to the last and the last written terms in this expansion. Then we have the following on the subspace $N_0=\{\bar{u}_0=0\}$:

$$\rho[\lambda](\bar{w}) = \int_0^T \left(-\frac{1}{2}\sum_{i=1}^{k-1}\bar{u}_i\,\psi\Big(r_i''(\hat{x})\bar{x},\bar{x}\Big)+\sum_{j=1}^{k-1}\bar{y}_j\,\psi\Big(r_j'(\hat{x})(\sum_{i=1}^{k-1}\bar{u}_i r_i'(\hat{x})\bar{x})\Big)\right)dt. \tag{17.20}$$

(The variables $\bar{z}$, $\bar{u}_0$ do not enter this expression; $(r_i''\bar{x},\bar{x})$ is a vector-valued quadratic form; we then take the inner product of this form and the covector ψ.) Substituting $\bar{x}=\sum\bar{y}_s r_s(\bar{x})$ into the above expression, we obtain

$$\eta[\lambda](\bar{y}) = \int_0^T\Big(-\frac{1}{2}\sum_{ijs}\bar{u}_i\bar{y}_j\bar{y}_s\,\psi(r_i''r_jr_s)+\sum_{ijs}\bar{u}_i\bar{y}_j\bar{y}_s\,\psi(r_j'r_i'r_s)\Big)\,dt,$$

and hence

$$\omega[\lambda](t_*) = -\frac{1}{2}\sum_{ijs}\psi(r_i''r_jr_s)\,\bar{y}_j\bar{y}_s d\bar{y}_i\;+\;\sum_{ijs}\psi(r_j'r_i'r_s)\,\bar{y}_j\bar{y}_s d\bar{y}_i.$$

All the coefficients here are constant (frozen at the point t_*); therefore, it is not difficult to verify that the closedness of this 1-form is equivalent to the fulfilment of the following relations:

$$\psi\,[r_s\,[r_i,r_j]\,](\hat{x}(t_*))=0,\qquad\forall\, i,j,s\,=1,\ldots,k-1. \tag{17.21}$$

Indeed, we have

$$d\,\omega[\lambda](t_*) = -\frac{1}{2}\sum\psi(r_i''r_jr_s)\Big(\bar{y}_s\,d\bar{y}_j\wedge d\bar{y}_i+\bar{y}_j\,d\bar{y}_s\wedge d\bar{y}_i\Big)\;+$$
$$+\sum\psi(r_j'r_i'r_s)\Big(\bar{y}_s\,d\bar{y}_j\wedge d\bar{y}_i+\bar{y}_j\,d\bar{y}_s\wedge d\bar{y}_i\Big)=0.$$

The relation $d\omega=0$ (we omit λ,t_*) means that, for any triple of subscripts $i,j,s=1,\ldots,k-1$, the coefficients by the monomials $\bar{y}_s\,(d\bar{y}_j\wedge d\bar{y}_i)$ and $\bar{y}_s\,(d\bar{y}_i\wedge d\bar{y}_j)$ add up to zero, i.e.,

$$-\frac{1}{2}\,\psi(r_i''r_jr_s)+\frac{1}{2}\,\psi(r_j''r_ir_s)-\frac{1}{2}\,\psi(r_i''r_sr_j)+\frac{1}{2}\,\psi(r_j''r_sr_i)\;+$$
$$+\psi(r_j'r_i'r_s)-\psi(r_i'r_j'r_s)+\psi(r_s'r_i'r_j)-\psi(r_s'r_j'r_i)\;=\;0.$$

One can change the order of arguments in the first two terms of this relation (r_j, r_s and r_i, r_s, respectively) and then, collecting similar terms, we obtain

$$-\,\psi(r_i''r_sr_j)\;+\;\psi(r_j''r_sr_i)\;+ \tag{17.22}$$
$$+\,\psi(r_j'r_i'r_s)-\psi(r_i'r_j'r_s)\;+\;\psi(r_s'r_i'r_j)-\psi(r_s'r_j'r_i)\;=\;0.$$

(The last four terms are not changed.)

On the other hand, if one opens the double Lie brackets in relation (17.21), then one obtains exactly relation (17.22). Thus, the equivalence of the condition $d\omega = 0$ to relations (17.21) is proved.

References

[1] A. A. Agrachev and A. V. Sarychev, "Strong minimality of abnormal geodesics for 2-distributions," *J. of Dynamical and Control Systems,* **1**, No. 2, 139–176 (1995).

[2] G. A. Bliss, *Lectures on the calculus of variations*, Univ. of Chicago Press (1946); Russian transl., IL, Moscow (1950).

[3] R. Bryant and L. Hsu, "Rigidity of integral curves of rank 2 distributions," *Invent. Math.,* **114**, 435–461 (1993).

[4] A. V. Dmitruk, "Jacobi-type conditions for the problem of Bolza with inequalities," *Math. Notes of the Acad. Sci. USSR,* **35**, No. 6, 427–435 (1984).

[5] A. V. Dmitruk, "Quadratic order conditions of a Pontryagin minimum in an optimal control problem, linear in the control. I. Theorem on deciphering," *Mathematics of the USSR, Izvestija,* **28**, No. 2, 275–303 (1987).

[6] A. V. Dmitruk, "Second order necessary and sufficient conditions of a Pontryagin minimum for singular regimes," *Lect. Notes in Control and Inf. Sci.,* **180**, 334–343 (1992).

[7] A. V. Dmitruk, *Quadratic Order Conditions of the Pontryagin Minimum for Singular Extremals in Optimal Control Problems* [in Russian], *Dr. Sci. Dissertation,* St. Petersburg State University (1994).

[8] A.V.Dmitruk, "Second order necessary and sufficient conditions of a Pontryagin minimum for singular boundary extremals," In: *Proc. Third Intern. Congr. Industr. Appl. Math., Hamburg, 1995*, Zeitschrift für Angewandte Mathematik und Mechanik, Issue 3, pp. 411–412 (1995).

[9] A. V. Dmitruk, "Quadratic order conditions of a local minimum for singular extremals in a general optimal control prooblem," In: *Proc. of Symposia in Pure*

Math., Vol. 64, *Differential Geometry and Control* (G. Ferreyra et al., eds.), Amer. Math. Society, pp. 163–198 (1998).

[10] A. V. Dmitruk, "Quadratic sufficient conditions for a strong minimality of abnormal sub-Riemannian geodesics," *Russian Journal of Math. Physics,* **6**, No. 3, 363–372 (1999).

[11] A. V. Dmitruk, A. A. Milyutin, and N. P. Osmolovskii, "Lyusternik's theorem and the theory of extremum," *Russian Math. Surveys,* **35**, No. 6, 11-51 (1980).

[12] A. Ya. Dubovitskii and A. A. Milyutin, "Theory of maximum principle," In: *Methods of Extremal Problems Theory in Economics* [in Russian], Nauka, TsEMI, (1981), pp. 6–47.

[13] W. Liu and H. J. Sussmann, "Shortest paths for sub-Riemannian metrics on rank 2 distributions," *Mem. Amer. Math. Soc.,* **118**, No. 564 (1995).

[14] A. A. Milyutin, "Maximum principle in regular optimal control problem," In: *Necessary Condition in Optimal Control* [in Russian], Nauka, Moscow (1990).

[15] A. A. Milyutin, "The notion of rigidity and optimal control theory," *Tr. Mosk. Mat. Obshch. (Trans. of Moscow Math. Soc.),* **60**, 98–154 (1998).

[16] R. Montgomery, "Abnormal minimizers," *SIAM J. Control and Optimiz.,* **32**, No. 4, 1605–1620 (1994).

[17] R. Montgomery, "Survey of singular curves in sub-Riemannian geometry," *J. of Dynamical and Control Systems,* **1**, No. 1, 49–90 (1995).

[18] K. Makowski and L. W. Neustadt, "Optimal control problems with mixed control-phase variable equality and inequality constraints," *SIAM J. Control and Optimiz.,* **12**, No. 2, 184–228 (1974).

[19] N. N. Petrov, "On shortest sub-Riemannian geodesics," *Differents. uravn. (Diff. Equations),* **30**, No. 5 (1994).