Nonlinear Analysis, 1997, v. 30, No. 4, 2439–2448.

QUADRATIC ORDER CONDITIONS OF A LOCAL MINIMUM FOR ABNORMAL EXTREMALS

A. V. DMITRUK
Central Economic-Mathematical Institute, Russian Academy of Sciences,
Russia 117418, Moscow, ul. Krasikova, 32

It is well known in the theory of extremal problems that the abnormal case, i.e. the case when equality constraints are degenerate at the examined point, is a difficult subject to obtain higher order conditions of a local minimum. Especially it is true for necessary conditions. The matter is that "standard" necessary conditions, relevant to the general case, are always trivially fulfilled in the abnormal case and do not provide any information about the presence or absence of a local minimum at the given point. Here we present a method of treatment extremal problems with degenerate equality constraints, originally proposed by A.A.Milyutin. It consists of the passing from the given problem to another one, in which the equality constraints are nondegenerate. Application of this method and of its refinement allows one to obtain informative quadratic order necessary conditions for local minima in some classes of problems.

Key words and phrases: equality constraints, Lyusternik condition, Lagrange multipliers, second and third variations of Lagrange function, weak and Pontryagin minimum, quadratic order conditions, finite codimension, Legendre type conditions.

1. GENERAL NECESSARY CONDITIONS

Consider the following Problem I:

$$J = \varphi_0(w) \to \min, \qquad \varphi_i(w) \le 0, \quad i = 1, \dots, \nu, \qquad g(w) = 0,$$

where w belongs to a Banach space W, the image of g lies in a finite-dimensional space R^q, and all the functions possess expansions up to quadratic terms in a neighborhood of a point w^0, which is examined to be a point of local minimum.

To be precise, we assume that for the mapping g there exists a bilinear mapping $Q_g : W \times W \to R^q$, such that

$$g(w + \bar{w}) = g(w) + g'(w^0)\,\bar{w} + Q_g(\bar{w}, \bar{w}) + \zeta_g(w, \bar{w})\ \|\ \bar{w}\ \|^2,$$

where $\zeta_g(w_n, \bar{w}_n) \to 0$ for any sequences $w_n \to w^0$ and $\bar{w}_n \to 0$. (This assumption can be weakened, see [3].) We denote $2Q_g(\bar{w}, \bar{w}) = g''[w^0]\,(\bar{w})$. The same assumption we take for all φ_i, $i = 1, \dots, \nu$.

We assume that w^0 is a stationary point, i.e. there exists a collection of Lagrange multipliers $\lambda = (\alpha_0, \dots, \alpha_\nu, \beta)$, where $\alpha_i \ge 0$, $i = 0, \dots, \nu$, and $\beta \in R^q$, such that $\sum_{i=0}^{\nu} |\alpha_i| + |\beta| = 1$, $\alpha_i \varphi_i(w^0) = 0\ \forall\, i = 1, \dots, \nu$, and Lagrange function $\Phi[\lambda](w) = \alpha_0\varphi_0(w) + \dots + \alpha_\nu\varphi_\nu(w) + \beta g(w)$ is stationary at w^0: $\Phi'[\lambda](w^0) = 0$. Denote by Λ the set of all such collections for w^0. It is obviously a finite–dimensional compact set.

For any λ let $\Omega[\lambda](\bar{w})$ be the second variation of $\Phi[\lambda](w)$ at w^0. Since λ, generally, is not unique, the question arises, which functional of the family $\Omega[\lambda](\bar{w})$ one has to consider. The answer is [4] that a'priori one has to take all this family, and to consider its maximum over $\lambda \in \Lambda$. This somewhat unusual functional appears in the finite-dimensional problem already. Generally, for any set $M = \{\lambda\}$ define the functional $\Omega\,[M](\bar{w}) \;:=\; \sup_{\lambda\in M}\,\Omega[\lambda](\bar{w})$.

Denote by $\mathcal{K}$ the so-called cone of critical variations, consisting of all $\bar{w} \in W$ such that $\varphi_i'(w^0)\bar{w} \le 0$ for all $i \in I$, and $g'(w^0)\bar{w} = 0$, where $I = \{i:\ \varphi_i(w^0) = 0\} \cup \{0\}$ is the set of active indices.

In [4] the following necessary condition was obtained, which we regard as "the standard" (or "canonical") one.

THEOREM 1. Let w^0 be a local minimum point in Problem I. Then for all $\bar{w} \in \mathcal{K}$

$$\Omega\,[\Lambda](\bar{w}) \;=\; \sup_{\lambda\in\Lambda}\,\Omega[\lambda](\bar{w}) \;\ge\; 0. \tag{1}$$

Note that since $\Omega[\lambda](\bar{w})$ is linear with respect to λ, the set Λ can be replaced in (1) by its convex hull $\operatorname{co}\Lambda$. However, it is desirable not to expand but to narrow the set of λ, over which the supremum is taken, because the more narrow set of λ, the more strong is necessary condition (1). In [3] A.A.Milyutin proved (among other results) the following strengthening of this condition, which we also regard as a "standard" one.

For any set $M = \{\lambda\}$ denote by M^+ the set of all $\lambda \in M$ such that the quadratic functional $\Omega[\lambda](\bar{w})$ is nonnegative on a subspace in W of a finite codimension (which may depend on λ).

THEOREM 2 (on finite codimensions). If condition (1) holds on $\mathcal{K}$, then $(\operatorname{co}\Lambda)^+$ is nonempty, and for all $\bar{w} \in \mathcal{K}$

$$\Omega\,[(\operatorname{co}\Lambda)^+](\bar{w}) \;\ge\; 0. \tag{2}$$

COROLLARY. If w^0 is a local minimum point in Problem I, then $(\operatorname{co}\Lambda)^+$ is nonempty, and for all $\bar{w} \in \mathcal{K}$ inequality (2) holds.

2. FINITE CODIMENSIONS AND ITERATIONS OF GOH CONDITIONS

In order to show what means $\lambda \in (\operatorname{co}\Lambda)^+$, let us consider briefly the case, when Ω is given in the classical integral form (here we omit the argument λ, and the bar over w): $w = (x,u) \in W = AC^m\,[0,T] \times L^r_\infty\,[0,T]$, where AC^m is the space of $m-$dimensional absolutely continuous functions,

$$\dot{x} = A(t)x + B(t)u, \quad x(0) = 0, \tag{3}$$

$$\Omega(w) = (Sx_T, x_T) + \int_0^T [(Q(t)x,x) + (P(t)x,u) + (R_0(t)u,u)]\,dt, \tag{4}$$

the matrices A, Q, R_0 are measurable and essentially bounded, and P, B are smooth enough. Recall the known Goh transformation: $(x,u) \to (\xi, y, u)$, where $\xi = x - By$, and

$$\dot{y} = u, \qquad y(0) = 0, \tag{5}$$

so that

$$\dot{\xi} = A\xi + B_1 y, \qquad \xi(0) = 0, \tag{6}$$

$$\Omega(x,u) = \Omega_1(\xi, y, u) = (S_1(\xi_T, y_T), (\xi_T, y_T)) +$$

$$+\int_0^T [(Q_1(t)\xi,\xi)+(P_1(t)\xi,y)+(R_1(t)y,y)+(V_1(t)y,u)+(R_0(t)u,u)]\,dt, \tag{7}$$

where S_1 is an $(m+r)\times(m+r)-$ matrix, $B_1 = AB - \dot{B}$, and the matrix V_1 is smooth skew-symmetric. The term $(C(t)\xi,u)$, and the symmetric part of $(V_1(t)y,u)$ are reduced to the presented ones after the integration by parts in view of (5), (6).

This transformation can be repeated (if P_1 and B_1 are again smooth) in such a way that now $(\xi,y)\to(\eta,z)$, $\eta=\xi-B_1z$, $\dot{z}=y$, $z(0)=0$, $\dot{\eta}=A\eta+B_2z$ $(B_2=AB_1-\dot{B}_1)$, $\eta(0)=0$, and so on, until the smoothness of the initial P, B allows one to do it.

The classical Legendre condition for Ω means that $R_0(t)\ge 0$ a.e. on $[0,T]$. It may happen to exist nontrivial intervals (a,b), on which $R_0(t)=0$ a.e. We will call them intervals of 0-degeneracy. We will say that Ω satisfies Goh conditions of the 1-st degree, or, simply, 1-Goh conditions, if $R_0(t)\ge 0$, and on any interval of 0-degeneracy $V_1(t)=0$ and $R_1(t)\ge 0$ a.e. If there are no intervals of 0-degeneracy, then by definition Ω satisfies 1-Goh conditions. An interval of 0-degeneracy, on which $R_1(t)=0$, we will call interval of 1-degeneracy. Ω satisfies 2-Goh conditions, if it satisfies 1-Goh conditions and on any interval of 1-degeneracy $V_2(t)=0$ and $R_2(t)\ge 0$ a.e., and so on. We then say that Ω satisfies k-Goh conditions, if it satisfies $(k-1)$-Goh conditions and on any interval of $(k-1)$-degeneracy $V_k(t)=0$ and $R_k(t)\ge 0$ a.e. Thus, if Ω satisfies k-Goh conditions for some k, and on any interval of $(k-1)$-degeneracy $R_k(t)>0$ a.e., then there are no intervals of k-degeneracy, and automatically Ω satisfies m-Goh conditions for all $m>k$.

Lemma 1. Suppose that Ω of the form (3), (4) is nonnegative on a subspace in W of a finite codimension. Then Ω satisfies k-Goh conditions for all $k\ge 1$.

The proof follows, essentially, from that for the case $k=1$, which we leave as an exercise to the reader.

Observe however, that conditions (1) and (2) are informative only if the mapping g is nondegenerate at w^0, or, as is also said, satisfies Lyusternik condition: $g'(w^0)$ maps onto. Indeed, in this case $\mathrm{co}\,\Lambda$ does not contain 0, hence it can be shown that $\mathrm{co}\,\Lambda\subset[\xi,\eta]\cdot\Lambda$, where $\eta\ge\xi>0$, which readily implies that conditions (1) and (2) remain valid (in fact, equivalent to themselves) if one substitute $\mathrm{co}\,\Lambda$ by Λ. But in the opposite case, when $g'(w^0)$ is not onto, there exists $\beta\ne 0$ such that $\beta g'(w^0)=0$, and hence the collection $\lambda_1=(0,\ldots,0,\beta)$ together with $-\lambda_1$ belongs to Λ, which implies that $0\in\mathrm{co}\,\Lambda$ and then obviously $0\in(\mathrm{co}\,\Lambda)^+$, which yields $\Omega[\mathrm{co}\,\Lambda](\bar{w})\ge\Omega[(\mathrm{co}\,\Lambda)^+](\bar{w})\ge\Omega[0](\bar{w})=0.$

Thus, in the degenerate case conditions (1) and (2) are trivially fulfilled regardless of the presence or absence of a local minimum at w^0. So, it is highly desirable to have necessary conditions, that remain informative in the degenerate case. To our knowledge, the first such conditions were obtained, for optimal control problems, by A.J.Krener [1], A.A.Agrachev and R.V.Gamkrelidze [2] and then by others. The method of [1, 2] consists of choosing appropriate families of control variations of the so-called needle type, concentrated near an arbitrary point t_* and parameterized by the widths of the needles, and then of analyzing expansions of the corresponding state variations.

3. MILYUTIN's APPROACH: WEAKENING THE EQUALITY CONSTRAINTS

An original method to overcome the difficulty of degeneracy was proposed by A.A.Milyutin [3]. Its key idea is to replace the initial problem with degenerate equality constraints by another one with nondegenerate equalities, and to apply "standard" quadratic necessary conditions to this new problem. To be more precise, suppose that g is degenerate at w^0. Then g can be taken as $g=(g_1,g_2)$, where g_1 is nondegenerate at w^0 and $g_2'(w^0)=0$, and thereby Problem I can be presented in the form (Problem Ia):

$$J=\varphi_0(w)\to\min,\quad \varphi_i(w)\le 0,\quad i=1,\ldots,\nu,$$

$$g_1(w) = 0, \qquad g_2(w) = 0.$$

Denote $q_2 = \dim g_2$ and suppose that the quadratic mapping $g_2''[w^0](\cdot): W \to R^{q_2}$ (that is the quadratic part of the expansion of g at w^0) satisfies the following nondegeneracy condition, proposed by A.A.Milyutin [3]:

Definition. We say that $g_2''[w^0](\cdot)$ is nondegenerate in Milyutin's sense, if there exists a closed convex solid (c.c.s.) cone $K_2 \subset R^{q_2}$ such that for any natural number m there exists a number $Q(m)$ such that for any $z \in K_2$, and any subspace $\Gamma \subset W$ with codim $\Gamma \le m$ there exists $\bar{w} \in \Gamma$ such that

$$g_2''[w^0](\bar{w}) = z \qquad \text{and} \qquad \| \bar{w} \| \le Q(m)\sqrt{|z|}\ .$$

This condition is in a sense a quadratic analogue of the (first order) Lyusternik nondegeneracy condition. It was also shown in [3] that the failure of this condition (i.e. the degeneracy of $g_2''[w^0]$) means that there exists $\beta_2 \ne 0$ such that $\beta_2 g_2''[w^0](\bar{w}) = 0$ on a subspace $\Gamma \subset W$ of a finite codimension. K_2 is called *the super-Legendre cone* for quadratic mapping $g_2''[w^0](\cdot)$. The following basic result was established.

Theorem 3 (A.A.Milyutin). Let w^0 be a local minimum point in Problem Ia, and suppose that $g_2''[w^0]$ is nondegenerate in the above sense. Then for any c.c.s. cone $K \subset -\mathrm{int}\, K_2 \cup \{0\}$ there exists a number $C > 0$ such that w^0 remains a local minimum point in the following Problem II:

$$J \;=\; \max_{i\in I} \varphi_i(w) \;+\; C|g_2(w)| \;\to \min,$$

$$g_1(w) = 0, \qquad g_2(w) \in K.$$

Observe that in passing from Problem Ia to Problem II the nondegenerate equality constraints $(g_1(w) = 0)$ remain unaffected, whereas the degenerate ones are replaced by an inequality. Thus, the equality constraints in Problem II are nondegenerate, and the "standard" quadratic necessary conditions can be efficient. To apply them we have to present the "nonfunctional" inequality $g_2(w) \in K$ in a "functional" form. It can easily be done if we choose the cone K to be finite–faced, i.e. if K is given by a finite number of inequalities $(b_s, z) \le 0,\ s = 1, \dots, \sigma$. Then the inequality constraint in Problem II takes the form:

$$(b_s, g_2(w)) \le 0, \quad s = 1, \dots, \sigma. \tag{8}$$

The nonsmoothness of the functional J does not matter much, because J is the maximum of a finite number of smooth functionals, whereas the whole higher–order theory [4] is valid for problems with such functionals.

Theorem 3 can be regarded as "a theorem on weakening the equality constraints", on the quadratic level. It is proved in [3] for a wide abstract class of problems, and is well suitable for obtaining conditions of a local minimum that are in terms of second variations of Lagrange functions. Note again a rather unexpected trick in Theorem 3: the passage from equality to inequality constraints. This trick is in a sense reverse to the known trick of Valentine (1939), in which one passes from inequality to equality constraints by adding the squares of new variables. Perhaps, in that time it was believed that equality constraints were more convenient to deal with than inequality constraints (note, that the latter were not considered in the classics). However, today, after several decades that inequality constraints have been included into investigation, and have become an ordinary object, it turned out that, on the contrary, they are in a sense even more convenient than equality constraints. We mention e.g. the definiteness of the sign of Lagrange multipliers for inequality constraints, and essentially more weak, rough assumptions for them, as compared with always rather strong, fine assumptions for equality constraints.

Before applying conditions (1) or (2) to Problem II, let us establish a relation between the sets of Lagrange multipliers Λ_{I} and Λ_{II}, and Lagrange functions Φ_{I} and Φ_{II} for Problems Ia and II, resp., the inequalities in the latter being in the form (8).

Lemma 2. There exists numbers $\eta \geq \xi > 0$ and a mapping $\pi: \Lambda_{\mathrm{II}} \to [\xi, \eta] \cdot \Lambda_{\mathrm{I}}$ such that for any $\lambda'' \in \Lambda_{\mathrm{II}}$ and $\lambda' = \pi\lambda''$ the corresponding Lagrange functions coincide:

$$\Phi_{\mathrm{II}}[\lambda''](w) = \Phi_{\mathrm{I}}[\lambda'](w) \quad \forall\, w \in W, \quad \text{hence their second variations coincide:}$$

$$\Omega_{\mathrm{II}}[\lambda''](\bar{w}) = \Omega_{\mathrm{I}}[\lambda'](\bar{w}) \quad \forall\, \bar{w} \in W, \quad \text{which implies that} \quad \pi(\Lambda_{\mathrm{II}}^{+}) \subset [\xi, \eta] \cdot \Lambda_{\mathrm{I}}^{+}.$$

The cones of critical variations in Problems Ia and II obviously coincide (because $g_2'(w^0) = 0$), so we denote them by the same letter $\mathcal{K}$.

Theorem 3 together with Theorem 1, Theorem on finite codimensions and Lemma 2 yield the following result for the initial problem.

Theorem 4 (A.A.Milyutin). Let w^0 be a local minimum point in Problem I. Then Λ^+ is nonempty, and for all $\bar{w} \in \mathcal{K}$

$$\Omega\,[\Lambda^+](\bar{w}) \geq 0. \tag{9}$$

Proof. Take Problem I in the form of Problem Ia, and suppose first that $g_2''[w^0]$ is nondegenerate in the sense of Milyutin. Take the c.c.s. cone K in the above finite–faced form. By theorem 3 w^0 is a local minimum point in Problem II. Applying Theorems 1 and 2 to Problem II, and taking into account the nondegeneracy of g_1, we get that $\Lambda_{\mathrm{II}}^{+}$ is nonempty, and $\quad \Omega_{\mathrm{II}}[\Lambda_{\mathrm{II}}^{+}](\bar{w}) \geq 0 \quad \forall\, \bar{w} \in \mathcal{K}$. From here and Lemma 2 it follows immediately that Λ_{I}^{+} is nonempty, and $\Omega_{\mathrm{I}}[\Lambda_{\mathrm{I}}^{+}](\bar{w}) \geq 0 \quad \forall\, \bar{w} \in \mathcal{K}$, q.e.d.

Suppose now that $g_2''[w^0]$ is degenerate in Milyutin's sense. In this case, as it was mentioned above, there is a $\beta_2 \in R^{q_2}$, $|\beta_2| = 1$, such that $\beta_2 g_2''[w^0](\bar{w}) = 0$ on a subspace in W of a finite codimension. Then $\lambda_2 = (\alpha = 0, \beta_1 = 0, \beta_2)$ together with $-\lambda_2$, obviously belonging to Λ, belong also to Λ^+. From here we get for all w : $\quad \Omega\,[\Lambda^+](w) \geq |\Omega\,[\lambda_2](w)| \geq 0$, q.e.d.

Condition (9) differs from condition (2) only in the absence of "co". This, however, makes condition (9) informative in the abnormal case.

Theorem 4 covers, in particular, the case of a weak minimum in optimal control problems, and, in view of Lemma 1 and other statements of this kind, strengthens the necessary conditions from [1, 2] and some others.

Remark 1. Theorem 4 remains valid in a more general setting: actually one need not require the equality constraint in Problem I to be finite-dimensional; it is sufficient to require that only the degenerate part of it is finite-dimensional. Thus, if $g = (g_1, g_2)$, where $g_1'(w^0)$ is onto, and if $\dim g_2 < \infty$, then Theorem 4 still holds.

Remark 2. Recently A.V.Arutyunov [10] obtained a refinement of condition (9), in which the set Λ^+ is narrowed to the set of all λ such that $\Omega\,[\lambda](\bar{w}) \geq 0$ on a subspace of bounded codimension, the upper bound of codimension being dependent on the number of constraints in the problem.

Let us now confine our consideration to optimal control problems, linear in the control, and pass from a weak to a more strong type of minimum. On a fixed time interval $[t_0, t_1]$ we consider the control system

$$\dot{x} = f(x,t) + F(x,t)u, \tag{10}$$

the terminal constraints

$$\varphi_i(x_0, x_1) \le 0, \quad i = 1, \ldots, \nu, \qquad K(x_0, x_1) = 0 \quad (\dim K = q), \tag{11}$$

where $x_0 = x(t_0),\ x_1 = x(t_1),$ and the cost functional $\quad J = \varphi_0(x_0, x_1) \to \min.$

Here $x \in AC^m[t_0, t_1],\ u \in L^r_\infty[t_0, t_1],\quad x$ is the state variable and u is the control (of dimensions m and r resp.), the reference point is (x^0, u^0).

In order to put this problem in the form of Problem I, one can take as independent variables $w = (x_0, u) \in R^m \times L^r_\infty$, and consider $x(t)$ to be the corresponding solution of the control system, whence $x_1 = x(t_1)$ depends on x_0 and u. By this simple trick the control system is no longer a constraint in the problem, and the only equality constraint $(K = 0)$ is finite-dimensional. However, in view of Remark 1, it is not necessary to do this passage, so we still denote the above problem by Problem I.

There may also be the nonfunctional constraint on the control $u \in U(t)$ (where $U(t)$ is a convex body, Hausdorff continuous in t), but formally we will not include it in the statement of the problem (because this would bring an essential nonsmoothness into it), instead we will use it for the description of the class of admissible variations, with respect to which the notion of minimum is considered.

Recall the notion of *Pontryagin minimum.* Let Π be the set of all sequences $w_n = (x_n, u_n)$ such that $\|x_n - x^0\|_C \to 0,\quad \|u_n - u^0\|_1 \to 0,$ and $\|u_n\|_\infty \le \mathcal{O}(1)$. We call them *Pontryagin sequences.*

Definition (see [4–6]. The point $w^0 = (x^0, u^0)$ is called a Pontryagin minimum point (or, briefly, a $\Pi-$ minimum point) in some problem, if there are no sequence $\{w_n\} \in \Pi$ such that for all n the point w_n satisfies all the constraints, and gives a lesser value to the cost functional: $J(w_n) < J(w^0)$.

In other words, a $\Pi-$ minimum is an L_1-minimum with respect to the control on any uniformly bounded control set.

Note that the set Π include, in particular, the so-called "needle-type" variations (sequences), so, strictly speaking, the Π-minimum is not a "local" minimum w.r.t. the control. Obviously, it occupies an intermediate position between the classic weak and strong minima.

This notion, however, is not very convenient to operate with, because if we include the constraint $u \in U$ in the problem, it would be too hard to keep it precisely satisfied, while making variations of the control. So, we also take the following notion, more convenient to work with. Denote by $\Pi(U)$ the set of all sequences from Π, such that $u_n(t) \in U(t) + v_n(t),$ where $\|v_n\|_\infty \to 0.$ (This set obviously possess an important property: it admits the addition of the uniformly small variations of the control. In fact, one can consider any subset $\Pi' \subset \Pi$ possessing this property.)

Definition. The point w^0 is called a $\Pi(U)-$minimum point in the above Problem I, if there are no sequence $\{w_n\} \in \Pi(U)$ such that for all n the point w_n satisfies constraints (10), (11), and $J(w_n) < J(w^0)$.

When analysing Π-minimum points for problems, linear in the control, it turns out that one has to take into account not only second variations of Lagrange functions, but also their third variations (see details in [5–9]). Theorems 3, 4 are not sufficient to treat fully this case, because if the mapping $g_2''[w^0]$ is degenerate, we cannot pass to Problem II, and, unlike in Theorem 4, cannot choose multipliers with required properties of the third variation of Lagrange function, since we have no information about the third derivative of g_2. In this case the following theorem is suggested [6, 7].

Let the initial control problem be of the form (Problem Ib):

$$J = \varphi_0(w) \to \min, \qquad \varphi_i(w) \le 0, \quad i = 1, \ldots, \nu,$$

$$g_k(w) = 0, \quad k = 1,\ 2,\ 3,$$

where the mapping g_1 is nondegenerate at $w^0 = (x^0, u^0)$, $g_2'(w^0) = 0$, $g_3'(w^0) = 0$, the mapping $g_2''[w^0]$ satisfies Milyutin's nondegeneracy condition, $g_3''[w^0]$ is "totally degenerate" in Milyutin's sense (i.e. vanishes on a subspace of finite codimension), and the cubic mapping $g_3'''[w^0]$ is nondegenerate in a cubic sense, described below. Then the following theorem holds.

THEOREM 5. Let w^0 be a $\Pi(U)$-minimum point in Problem Ib. Then there exists c.c.s. cone $K \subset -\mathrm{int}\, K_2 \cup \{0\}$ and numbers $C, \delta > 0$ such that w^0 remains a $\Pi(U)$-minimum point in the following Problem III:

$$J \;=\; \max_{i \in I} \varphi_i(w) \;+\; C|g_2(w)| \;\to \min,$$

$$g_1(w) = 0, \qquad g_2(w)) \in K, \qquad |g_3(w)| \;\le\; \delta\, |g_2(w)|.$$

This is a theorem on weakening the equality constraints "on the cubic level". It is proved in [6, 7] not for a general abstract problem (as Theorem 3 is), but only for optimal control problems, linear in the control, and we don't know, whether a similar theorem is valid for another class of problems. Theorem 5 allows one to obtain informative necessary conditions of a quadratic order for a Pontryagin minimum in the cases when *a*) the control is free, or *b*) there is a control constraint $u \in U(t)$ and the reference control $u^0(t)$ goes strictly inside $\mathrm{int}\, U(t)$. (To apply Theorem 5 for the "exact" $\Pi-$minimum, we consider a more narrow set $U_\mu(t) = u^0(t) + \mu(U(t) - u^0(t))$, where $0 < \mu < 1$, so that for any sequence from $\Pi(U_\mu)$ we have $u_n(t) \in U(t)$ for n large enough, and hence $\Pi-$minimum with $u \in U$ implies $\Pi(U_\mu)-$minimum. We apply Theorem 5 to $\Pi(U_\mu)-$minimum, and then consider the limit as $\mu \to \infty$.)

For any set $M = \{\lambda\}$ denote by $E(M)$ the set of all $\lambda \in M$ such that the second variation $\Omega[\lambda](\bar w)$ of Lagrange function $\Phi[\lambda](w)$ satisfies Goh conditions (of the first degree), and the sum of the second and third variations of Lagrange function satisfies the pointwise condition (15) below with $a = 0$ (a new condition of Legendre type), and put $E^+(M) = E(M) \cap M^+$. Next one can prove Lemma 3 that, like Lemma 2, establish a relation between the sets Λ_{I}, Λ_{III}, and Lagrange functions Φ_{I}, Φ_{III} for Problems Ib and III, resp., with the additional property

$$\pi\ (E^+(\Lambda_{\mathrm{III}})) \;\subset\; [\xi, \eta] \cdot E^+(\Lambda_{\mathrm{I}}). \tag{12}$$

This property is also valid under conditions of Lemma 2. The cones of critical variations in Problems Ib and III again obviously coincide. The general theory [4] and Theorem on finite codimensions allows one to obtain the following result [7, 8].

THEOREM 6. Let $w^0 = (x^0, u^0)$ be a Π-minimum point in Problem I with the constraint $u \in U(t)$. Then $E^+(\mathrm{co}\,\Lambda)$ is nonempty, and for all $\bar w \in \mathcal{K}$

$$\Omega\ [E^+(\mathrm{co}\,\Lambda)](\bar w) \;\ge\; 0. \tag{13}$$

Like before, this condition is informative only in the normal case, when g satisfies Lyusternik condition. However, due to Theorems 3 and 5 one can remove the "co" from Theorem 6, so that the following theorem holds.

THEOREM 7. Let $w^0 = (x^0, u^0)$ be a Π-minimum point in Problem I with the constraint $u \in U(t)$. Then $E^+(\Lambda)$ is nonempty, and for all $\bar{w} \in \mathcal{K}$

$$\Omega\,[E^+(\Lambda)]\,(\bar{w}) \;\geq\; 0. \tag{14}$$

PROOF. We give it here for $\Pi(U_\mu)-$minimum. Suppose first, that we can take Problem I in the form of Problem Ia, where $g_2''[w^0]$ satisfies Milyutin's nondegeneracy condition. In this case, with account of property (12), the proof repeats that of Theorem 4.

Now suppose that $g_2''[w^0]$ is degenerate, and consider Problem I as Problem Ib, when $g_3'''[w^0]$ is nondegenerate in the cubic sense. Then by Theorem 5 w^0 is a $\Pi(U_\mu)$-minimum point in Problem III. Applying Theorem 6 to Problem III, we get, due to the nondegeneracy of g_1, that $E^+(\Lambda_{\mathrm{III}})$ is nonempty, and $\Omega_{\mathrm{III}}\,[E^+(\Lambda_{\mathrm{III}})]\,(\bar{w}) \geq 0 \quad \forall\, \bar{w} \in \mathcal{K}.$ From here and Lemma 3 immediately follows that $E^+(\Lambda_{\mathrm{I}})$ is nonempty, and $\Omega_{\mathrm{I}}\,[E^+(\Lambda_{\mathrm{I}})]\,(\bar{w}) \geq 0 \quad \forall\, \bar{w} \in \mathcal{K},$ q.e.d.

Finally, suppose that $g_3'''\,[w^0]$ is degenerate in the cubic sense. Then it can be shown (see the end of next chapter), that there is a $\beta_3 \neq 0$, such that $\lambda_3 = (\alpha = 0, \beta_1 = 0, \beta_2 = 0, \beta_3)$ together with $-\lambda_3$ belong to $E^+(\Lambda_{\mathrm{I}})$, from which inequality (14) follows trivially.

5. THE NEW LEGENDRE TYPE CONDITION

Now adduce in brief the above-mentioned pointwise condition, which, in our opinion, may present an intrinsic interest. Fix any λ and define the cubic functional (we again omit the λ):

$$\rho(\bar{w}) = \int_{t_0}^{t_1} [-(H_{uxx}\bar{x}, \bar{x}, \bar{u}) + 2((F_x'\bar{x}, \bar{u}), H_{xu}\bar{y})]\,dt,$$

where $H(x,u,t) = \psi(t)(f(x,t) + F(x,t)u)$, and $\dot{\bar{y}} = \bar{u}$, $\bar{y}(t_0) = 0$. It is shown in [6] that $\rho(\bar{w})$ is in a sense the principal part of the third variation of Lagrange function $\Phi[\lambda](w)$ at w^0, the latter being considered on any Pontryagin sequence. Replacing here $\bar{x}$ by $F(x^0(t), t)\bar{y}$ (which, in fact, corresponds to the use of Goh transformation, see [5–9]), we come to the following cubic functional:

$$e(\bar{w}) = \int (\mathcal{E}(t)\bar{y}, \bar{y}, \bar{u}))dt,$$

where $\mathcal{E}$ is a tensor with smooth coefficients. Next define the functional

$$L(\bar{y}) \;=\; \int\, [(R(t)\bar{y}, \bar{y}) \;+\; (\mathcal{E}(t)\bar{y}, \bar{y}, \bar{u})]\,dt,$$

where R is the matrix (R_1) from the second variation (7) of Lagrange function. We call it Legendre part of Lagrange function.

Now fix an arbitrary t_*, and consider $L(\bar{y})$ with the coefficients, frozen at t_* (omitting the bars):

$$L\,[t_*](y) \;= \int [(R(t_*)y, y) \;+\; (\mathcal{E}(t_*)y, y, u)]\,dt.$$

The above-mentioned condition, that makes the selection λ to $E(M)$, is: for any t_* and any absolutely continuous function $y(t)$, such that $\dot{y} = u \in U(t_*)$, and $y(t_0) = y(t_1) = 0$, the following inequality holds:

$$L[t_*](y) \ \geq \ a \int (y, y) \, dt, \tag{15}$$

where $a = 0$ for necessary conditions and $a > 0$ for sufficient ones. Observe, that this condition involves not only the second variation of Lagrange function (as usual), but also its third variation and the admissible control set $U(t)$, frozen at t_*. Condition (15) is to be verified at each point t_* separately, and because of this we call it a condition of Legendre type. However, for each t_* it leads to an auxiliary optimal control problem: to find the maximal a for which inequality (15) holds true for all $y(t)$. This auxiliary problem is rather nontrivial, and has intrinsic interest; see more about it in [8, 9].

THE CUBIC NONDEGENERACY CONDITION. The mapping $g: \ R^m \times L_\infty^r \to R^q, \ (x_0, u) \to K(x_o, x_1)$, where $x_1 = x(t_1)$ is the endpoint of the solution to (11), can be shown to possess the following expansion at $w^0 = (x_0^0, u^0)$ [6, 7]:

$$g(w^0 + \bar{w}) = g(w^0) + g'(w^0)\bar{w} + \frac{1}{2} g''[w^0](\bar{w}) + \int (\mathcal{E}_g(t)\bar{y}, \bar{y}, \bar{u}) \, dt + \zeta(\bar{w}),$$

where $\mathcal{E}_g$ is a $q-$dimensional $r \times r \times r$-tensor, and $\zeta(\bar{w}_n) \to 0$ for any Pontryagin sequence $\bar{w}_n$.

Define the $q-$dimensional differential 1-form:

$$\omega(t) = (\mathcal{E}_g(t)y, y, dy),$$

where t is regarded as a parameter.

We say that g is *cubic degenerate at* w^0, if there exists a nonzero $b \in R^q$, such that the scalar differential 1-form $(b, \omega(t)) = (b, \mathcal{E}_g(t)y, y, dy))$ for all t is closed: $d(b, \omega(t)) = 0$, the differential being taken w.r.t. y. If there is no such b, we say that g is *cubic nondegenerate at* w^0. In the degenerate case for any t_* and for any cycle $y(t)$ obviously $\int b(\mathcal{E}(t_*)y(t), y(t), u(t)) \, dt = 0$, hence only the first term remains in $L[t_*]$. If in addition, $g''[w^0]$ is "totally degenerate" in Milyutin's sense (i.e. vanishes on a subspace of finite codimension), hence $b \, g''[w^0](\bar{w}) = 0$ on this subspace too, and then for $\lambda = (\alpha = 0, \beta = b)$ and $-\lambda$ the second variation $\Omega[\lambda](\bar{w}) = \pm b \, g''[w^0](\bar{w})$ satisfies Goh conditions, whence $R(t) = 0$. (Recall that in our problem $\Omega[\lambda]$ has the form (4) with $R_0(t) = 0$, and Goh conditions mean that $V_1(t) = 0$ and $R_1(t) \geq 0$.) Thus, for this λ, for any t_* we get $L[t_*](y) \equiv 0$, so condition (15) holds trivially with $a = 0$. We apply these considerations to the mapping g_3 in the proof of Theorem 7.

6. SUFFICIENT CONDITIONS

In conclusion, we briefly touch on the question of sufficient conditions. This is a matter of a quite different character. The main difficulty here is not the degeneracy of the equality constraints, but the search for a proper "order of estimation", i.e. a positive functional $\gamma(\bar{w})$, regarding to which all the functionals in the problem can be taken into account. A proper order of estimation (usually quadratic one, but this is not necessary) allows one to obtain sufficient conditions for a local minimum that are close to necessary ones. This means that, replacing the nonstrict inequalities in Theorems 1, 2, 4, 6, 7 ($\sup \Omega[\lambda](\bar{w}) \geq 0$) by the estimates: $\sup \Omega[\lambda](\bar{w}) \geq a\gamma(\bar{w})$, with $a > 0$, one gets the corresponding sufficient conditions. (However, the proof of this is not trivial, see e.g. [4, 6]). These sufficient conditions, as it can easily be understood, are informative regardless of the degeneracy or nondegeneracy of the equality constraints.

A general theory of such "adjoint pairs" of necessary and sufficient conditions was developed by A.A.Milyutin and his co-workers in [4]. The order $\gamma(\bar{w})$ turns out to be specific for every specific class of problems. For the general nonlinear control problem it contains the integral of the square of the control variation (see [4]):

$$\gamma_0(\bar{w}) \;=\; |\bar{x}(t_0))|^2 + \int |\bar{u}(t)|^2\,dt,$$

but for problems linear in the control this order is obviously too rough. For this class the proper order is:

$$\gamma(\bar{w}) = \;|\bar{x}(t_0)|^2 \;+\; |\bar{y}(t_1)|^2 \;+\; \int |\bar{y}(t)|^2\,dt, \qquad (16)$$

$$\dot{\bar{y}} = \bar{u}, \qquad \bar{y}(t_0) = 0,$$

which, as one can observe, contains (in addition to endpoint terms) only the integral of the squares of state variables ([5–9]).

The described approach allows one to obtain, in particular, sufficient conditions of the order (16) for a minimality of abnormal geodesics in the sub-Riemannian geometry. This will be presented in the nearest papers of the author.

ACKNOWLEDGEMENTS

This work was supported by the Russian Basic Research Foundation, grant 94-01-00337, and by the International Science Foundation, grant MGG300. The author thanks Prof. A.A.Milyutin for numerous discussions and valuable suggestions.

REFERENCES

1. KRENER A.J., The high order maximal principle and its application to singular extremals, *SIAM J. on Control* **15**, 256–293 (1977).

2. AGRACHEV A.A. & GAMKRELIDZE R.V., An optimality principle of the second order for a time-optimality problem, *Mathematics of the USSR, Sbornik* **100**, No. 4, (1976).

3. MILYUTIN A.A., Quadratic conditions of an extremum in smooth problems with a finite-dimensional image, in *Metody teorii ekstremal'nyh zadach v ekonomike (Methods of the theory of extremal problems in economics)* (Edited by V.L.Levin), pp. 138–177, Nauka, Moscow, (1981) (in Russian).

4. LEVITIN E.S., MILYUTIN A.A. & OSMOLOVSKII N.P., Conditions of higher order for a local minimum in problems with constraints, *Russian Math. Surveys* **33**, 97–168 (1978).

5. DMITRUK A.V., Quadratic order conditions of a Pontryagin minimum in an optimal control problem, linear in the control, with a constraint on the control, *Soviet Math. Doklady* **28**, No. 2 (1983).

6. DMITRUK A.V., Quadratic order conditions of a Pontryagin minimum in an optimal control problem, linear in the control, *Mathematics of the USSR, Izvestija* **28**, No. 2 (1987) and **31**, No. 1 (1988).

7. DMITRUK A.V., Fine theorems on weakening the equality constraints in optimal control problems, linear in the control, *Siberian Math. Journal* **31**, No. 2 (1990).

8. DMITRUK A.V., Second order necessary and sufficient conditions of a Pontryagin minimum for singular regimes, *Lecture Notes in Control and Information Sciences* **180**, 334–343 (1990).

9. DMITRUK A.V., Second order optimality conditions for singular extremals, in *Computational Optimal Control* (Edited by R.Bulirsch and D.Kraft), ISNM **115**, pp. 71–81, Birkhäuser (1994).

10. ARUTYUNOV A.V., On the theory of abnormal extremal problems with finite-dimensional image, *Doklady of the Russian Acad. Sci.* **348**, 295–298 (1996).