

On the optimal value function of optimal control problems

B. GOLLAN

Diese Arbeit untersucht die Optimalwertfunktion von parametrisierten deterministischen Optimal Control Problemen mit verschiedenen Zustandsbeschränkungen. Die Hauptergebnisse enthalten Abschätzungen über den Einfluß von kleinen Variationen des Parameters p auf den Optimalwert $V = V(p)$. Im einzelnen werden obere Schranken für verschiedene Richtungsableitungen von V und Abschätzungen für einen verallgemeinerten Subgradienten $\partial V(\cdot)$ gegeben. Diese Abschätzungen sind definiert durch Multiplikatoren, welche notwendige Optimalitätsbedingungen erster oder höherer Ordnung entlang einer Optimallösung des ungestörten Ausgangsproblems erfüllen. Alle Ausführungen basieren auf der Trennbarkeit oder Nicht-Trennbarkeit von geeigneten approximierenden Kegeln und setzen keine oder nur sehr schwache Regularitätsbedingungen voraus.

В работе исследуется функция оптимального значения параметризованных детерминистических задач оптимального управления с разными ограничениями на состояния. Главные результаты содержат оценки о влиянии маленьких вариаций параметра p на оптимальное значение $V = V(p)$. В частности, даются верхние границы для разных производных по направлению функции V и оценки для некоторого обобщенного субградиента $\partial V(\cdot)$. Эти оценки определены множителями, выполняющими необходимые условия оптимальности первого и высшего порядка вдоль оптимального решения невозмущенной исходной задачи. Все рассуждения основаны на отделимости или неотделимости подходящих аппроксимирующих конусов и предполагаются либо никакие или лишь очень слабые условия регулярности.

The optimal value of parameterized deterministic optimal control problems with state and mixed state-control constraints is studied. The main results are estimates for the effect of small variations of the parameter p on the optimal value $V = V(p)$. In particular, upper bounds for several directional derivatives of V as well as estimates for a generalized subgradient $\partial V(\cdot)$ are given. These estimates are given in terms of multipliers which satisfy first or higher order necessary optimality conditions along an optimal solution of the unperturbed problem. The theory is based on the separability or nonseparability of suitable approximating cones and requires no or only very weak regularity assumptions.

1. Introduction

This paper is concerned with a particular topic in the field called sensitivity of optimal control problems. We consider several classes of deterministic optimal control problems which contain a (vector-valued) parameter p . This parameter describes perturbations of a reference problem which is given by $p = 0$. We give criteria which ensure that perturbed problems have admissible solutions, together with an estimate of the respective value of the objective function. Thus we obtain estimates for the local behaviour of the optimal value function $V(p)$ which is defined as the optimal value of the problem with parameter p . Note that the results of this paper do not concern the form of optimal solutions of perturbed problems. This topic and the related question of continuity of V form a quite different subject. It requires

much stronger assumptions and also different arguments. Such results can be found in the work of DONTCHEV [4, 5, 6] and others. Preliminary results concerning our topic have been obtained in [7, 21] for standard optimal control problems and in [1] for more general problems with constraints $\theta(t, u, x) \leq 0$. However, there the assumptions involve a more or less detailed a-priori knowledge about perturbed problems, such as the following: For any small perturbation the problem has a (unique) optimal solution which is close to the (unique) solution of the unperturbed problem.

More advanced are the results obtained by MAURER in [18–20]. He applies perturbation results from nonlinear programming to optimal control problems. His results require only information about some optimal reference solution $(\bar{x}(t), \bar{u}(t))$ of the unperturbed problem, yet no assumptions about neighbourhoods of $p = 0$. Furthermore, the case of state constraints $\psi(p, x) \leq 0$ is carefully treated. However, MAURER needs some other regularity conditions, such as the complete controllability of the linearized dynamical equation $\dot{x} = f_x x + f_u u$ along $(\bar{x}(t), \bar{u}(t))$.

Our results in Chapter 3 are closely related to those in [18–20], but they are obtained without any of the regularity assumptions mentioned above. This is possible by using general results from [12, 13], which are based on the separability or non-separability of suitable convex sets. Via the notion of a derived set or cone, which was introduced by Hestenes and is closely related to Pontryagin's cone of attainability, these results apply directly to different types of optimal control problems. Therefore, whenever necessary optimality conditions (expressed in terms of adjoint variables and other multipliers) are proved via derived sets, they can be used to obtain sensitivity results similar to that of Theorem 3.1. This holds in particular for the necessary conditions in [3, 14–17, 25–29], i.e., also for particular classes of control problems involving time delays.

In addition to Maurer's results we consider also problems with mixed constraints $\theta(t, x, u, p) \leq 0$. In Chapter 4 we show that sometimes sharper results than in Chapter 3 can be obtained by using multipliers which satisfy higher order necessary conditions as in [17, 28]. In Chapter 5 we extend the results of Chapter 3 and give estimates for a generalized gradient $\partial V(0)$. The notion of $\partial V(\cdot)$ used there requires only lower semicontinuity of V at $p = 0$ and thus covers a large class of applications. We conclude with an economic model which involves several control variables as well as several mixed constraints.

The general problem we deal with is the following:

Find an arc $\alpha_0 = (x(\cdot), u(\cdot), b)$, $u(\cdot)$ piecewise continuous, such that

$$I^0(\alpha) = g^0(b) + \int_{t_0}^{t_1} L^0(t, x(t), u(t)) dt \quad (1.1)$$

is minimized on the set \mathfrak{A} of admissible arcs α on $[t_0, t_1]$ given by the following constraints:

$$\begin{aligned} \dot{x}^i &= f^i(t, x, u), & 1 \leq i \leq n, & & u(t) \in U, \\ \psi^\alpha(t, x) &\leq 0, & 1 \leq \alpha \leq A', & & \psi^\alpha(t, x) = 0, & A' < \alpha \leq A, \\ \theta^\beta(t, x, u) &\leq 0, & 1 \leq \beta \leq B', & & \theta^\beta(t, x, u) = 0, & B' < \beta < B, \\ I^\gamma(\alpha) &\leq 0, & 1 \leq \gamma \leq G', & & I^\gamma(\alpha) = 0, & G' < \gamma \leq G, \\ x^i(t_j) &= X_j^i(b), & t_j &= T^j(b), & j &= 0, 1, \end{aligned} \quad (1.2)$$

where $I^\gamma(\alpha) = g^\gamma(b) + \int_{t_0}^{t_1} L^\gamma(t, x(t), u(t)) dt$.

U is an open subset of \mathbf{R}^m , and b ranges in some open subset of \mathbf{R}^k . The functions $f^i, \theta^\beta, L^\gamma, g^\nu, X_j^i, T^i$ are once and φ^α twice continuously differentiable with respect to their arguments on suitable domains of definition.

From now on we always use the following conventions:

If no other specification is made, then the index i runs from 1 to n , α from 1 to A , β from 1 to B , γ from 1 to G . If in a product of one or more terms some index appears at least twice, then one has to sum up with respect to this index. For partial derivatives of a function, say, of f with respect to x , we use both the notation $\frac{\partial f}{\partial x}$ and f_x . If a function of the form $f(t, x, u)$ is evaluated at $(t, x(t), u(t))$, we sometimes write shortly $f(t) := f(t, x(t), u(t))$.

Perturbations of a general form of such problems are considered in Chapter 3. To do this, we need a characterization of an optimal arc α_0 of (1.1), (1.2) by means of necessary conditions. Such conditions have been developed in [25, 26], e.g., and the subsequent Theorem 1.1 is a special form of [26, Theorem 9.1].

We assume that the $(A + B) \times (A + B + m)$ -matrix

$$\begin{pmatrix} \frac{\partial \theta^\beta}{\partial u^i} & \delta_{\beta e} \theta^e & 0 \\ \frac{\partial \varphi^\alpha}{\partial u^i} & 0 & -\delta_{\alpha \sigma} \varphi^\sigma \end{pmatrix} \text{ has rank } A + B \tag{1.3}$$

along an optimal arc $\alpha_0 = (x(t), u(t), b)$ on a set R_ε , where j runs from 1 to m , ρ from 1 to B , σ from 1 to A , and where we denote:

$$\varphi^\alpha(t, x, u) = \psi_t^\alpha(t, x) + \psi_x^\alpha f^i(t, x, u),$$

$$R_\varepsilon = \{z = (t, x(t), u) \mid \varphi^\alpha(z) \geq 0 \text{ for all } \alpha \text{ with } \psi^\alpha(t, x(t)) = 0,$$

$$\text{or } \varphi^\alpha(z) \leq 0 \text{ for all } \alpha \text{ with } \psi^\alpha(t, x(t)) = 0, \theta^\beta(z) \leq 0, 1 \leq \beta \leq B'$$

$$\theta^\beta(z) = 0, B' < \beta \leq B, |u - u(t)| \leq \varepsilon, t_0 \leq t \leq t_1\}.$$

In the sequel we use the Hamiltonian

$$H(t, x, u, y, \mu, \nu) = y^0 L^0 + y^i f^i - \lambda^\gamma L^\gamma - \mu^\alpha \varphi^\alpha - \nu^\beta \theta^\beta.$$

Theorem 1.1: *Let $\alpha_0 = (x(t), u(t), b)$ be an optimal solution of problem (1.1), (1.2) and assumption (1.3) hold on a set R_ε . Then there are multipliers $y^0, x^\alpha, \lambda^\gamma, \mu^\alpha(t), \nu^\beta(t), y^i(t)$, defined on $[t_0, t_1]$, with the following properties:*

- (i) $y^0 \leq 0, \lambda^\gamma \geq 0, 1 \leq \gamma \leq G', x^\alpha \geq 0, 1 \leq \alpha \leq A', (y^0, y(t), \lambda, x) \neq 0$ for any t , and $\lambda^\gamma = 0$, if $I^\gamma(\alpha_0) < 0$.
- (ii) $H(t, x(t), u(t), y(t), \mu(t), 0) \geq H(t, x(t), u, y(t), \mu(t), 0)$ for all $u \in U$ with $(t, x(t), u) \in R_\varepsilon$.
- (iii) The functions $y^i(\cdot)$ are continuous on $[t_0, t_1]$, and continuously differentiable where $u(\cdot)$ is continuous. Together with $\mu(\cdot), \nu(\cdot)$ they satisfy $y^i = -H_{x^i}, H_{u^i} = 0$, along α_0 on those intervals where $u(\cdot)$ is continuous.
- (iv) $H(t) := H(t, x(t), u(t), y(t), \mu(t), \nu(t))$ is continuous on $[t_0, t_1]$, and continuously differentiable where $u(\cdot)$ is continuous. There $\frac{dH}{dt} = H_t$.

(v) The transversality conditions hold at a_0 (with $1 \leq s \leq k$):

$$0 = -y^0 \frac{\partial g^0}{\partial b^s} + \lambda^r \frac{\partial g^r}{\partial b^s} + (\kappa^a \psi_{z^i}^a(t_0) - y^i(t_0)) \frac{\partial X_0^i}{\partial b^s} + y^i(t_1) \frac{\partial X_1^i}{\partial b^s} \\ + (\kappa^a \psi_{z^i}^a(t_0) + H(t_0)) \frac{\partial T^0}{\partial b^s} - H(t_1) \frac{\partial T^1}{\partial b^s}.$$

(vi) $\mu^a(\cdot), v^\beta(\cdot)$ are continuous on intervals of continuity of $u(\cdot)$, and $\mu^a(\cdot)$ is further continuous at those t , where $\varphi^a(t)$ is discontinuous.

(vii) For $1 \leq \alpha \leq A'$, $\mu^a(\cdot)$ is a nonincreasing function which is constant on intervals upon which $\varphi^a(t) < 0$.

For $1 \leq \beta \leq B'$, $v^\beta(\cdot)$ is nonnegative. Moreover, $v^\beta(t) \theta^\beta(t) = 0$ (β not summed), on $[t_0, t_1]$.

(viii) If $\varphi^a(t_1) < 0$, then $\mu^a(t_1) = 0$, $1 \leq \alpha \leq A'$. If also $\varphi^a(t_0) < 0$, then $\mu^a(t_0) = \kappa^a$.

Except for the second part of statement (viii), this result has been proved in [25, 26]. The last statement can be proved as in [11] (where it is proved for a somewhat simpler problem) by investigating the problem (1.1), (1.2) reversed in time.

2. Some standard perturbations

Starting with (1.1), (1.2) as unperturbed problem, we consider standard perturbations of the terminal and the isoperimetric constraints. That means that for some $p = (p^1, \dots, p^G, \dots, p^{G+n})$ we replace the respective constraints in (1.2) by

$$I^r(a) \leq p^r, \quad 1 \leq \gamma \leq G', \quad I^r(a) = p^r, \quad G' < \gamma \leq G, \\ x^i(t_1) = X_1^i(b) + p^{G+i}, \quad 1 \leq i \leq n. \quad (2.1)$$

Problem (1.1), (1.2), with the respective constraints in (1.2) modified according to (2.1), will be denoted by $(P1)_p$. Our aim is to characterize $V = V(p)$, the optimal value of problem $(P1)_p$ depending on p . To be more precise, we agree about the following throughout this paper: If a problem with a perturbation p has no admissible solution, we set $V(p) = +\infty$. If admissible solutions exist, we set $V(p) = \inf \{I^0(a) \mid a \in \mathfrak{A}(p)\}$, where $\mathfrak{A}(p)$ denotes the set of all admissible solutions of the respective perturbed problem. Of course, this infimum may be $-\infty$.

We assume that an optimal arc a_0 of the unperturbed problem $(P1)_0$ is given which satisfies (1.3). Then by Theorem 1.1 it makes sense to define

$$\Omega = \{(y^0, \omega) = (y^0, \kappa, \lambda, \mu(\cdot), v(\cdot), y(\cdot)) \mid (y^0, \omega) \text{ satisfies the} \\ \text{statements of Theorem 1.1 along } a_0\}, \quad (2.2)$$

$$\Omega_0 = \{\omega \mid (0, \omega) \in \Omega\} \cup \{0\},$$

$$\Omega_1 = \{\omega \mid (-1, \omega) \in \Omega\}.$$

The crucial elements of our sensitivity analysis are the two terms

$$s_j(p) = \sup_{\Omega_j} \{y^i(t_1) p^{G+i} - \lambda^r p^r\}, \quad j = 0, 1. \quad (2.3)$$

By the definition of Ω_0 we have always $s_0(p) \geq 0$. It is easily checked that Ω forms a convex cone, and that $\Omega_1 + \Omega_0 \subseteq \Omega_1$, if $\Omega_1 \neq \emptyset$.

A first sensitivity result is the following.

Theorem 2.1: Let a_0 be an optimal arc of problem $(P1)_0$ along which assumption (1.3) holds.

A) If $\Omega_1 \neq \emptyset$ and $p \in \text{int} \{p \mid s_0(p) = 0\}$, then $s_1(p) < \infty$, then for any $s > s_1(p)$ there is a function $r(\varepsilon): \mathbb{R} \rightarrow \mathbb{R}^n$ of type $o(\varepsilon)$ and some $\varepsilon_0 > 0$ with the following property: For any $\varepsilon \in (0, \varepsilon_0]$ the perturbed problem $(P1)_{\varepsilon p + r(\varepsilon)}$ has an admissible solution and

$$V(\varepsilon p + r(\varepsilon)) - V(0) \leq \varepsilon s.$$

B) If $\Omega_1 = \emptyset$ and $s_0(p) = 0$, then for any $s \in \mathbb{R}$ there is a function $r(\cdot)$ with all properties as in A).

At the time being we regard this result as a technical tool for proving the similar, but more general Theorem 3.1. After that we will clarify its implications by some corollaries and remarks.

Proof of Theorem 2.1:

Step 1: We prove the first statement of A). Define

$$Z = \{(y^0, y(t_1), -\lambda) \mid (y^0, x, \lambda, \mu(\cdot), v(\cdot), y(\cdot)) \in \Omega\},$$

$$Z_0 = \{(y(t_1), -\lambda) \mid (0, y(t_1), -\lambda) \in Z\} \cup \{0\},$$

$$Z_1 = \{(y(t_1), -\lambda) \mid (-1, y(t_1), -\lambda) \in Z\}.$$

Z resp. Z_0 are convex cones in \mathbb{R}^{1+n+G} resp. \mathbb{R}^{n+G} , Z_1 is a convex set. By arguments made below (preceding (2.6)), Z, Z_0, Z_1 are also closed. Furthermore, $\Omega_1 \neq \emptyset$ implies $Z_1 \neq \emptyset$, and this implies $Z_1 + Z_0 \subseteq Z_1$. Therefore $Z_0 \subseteq 0^+Z_1$, where 0^+Z_1 denotes the recession cone of Z_1 (cf. [22]). Conversely let $z \in 0^+Z_1$, i.e., there exist sequences $\eta_j \downarrow 0$, $z_j \in Z_1$ with $\eta_j z_j \rightarrow z$. Then $\eta_j(-1, z_j) \in Z$ for all j and $\eta_j(-1, z_j) \rightarrow (0, z)$. Since Z is closed, $(0, z) \in Z$ and hence $z \in Z_0$. Therefore $Z_0 = 0^+Z_1$ and, since $s_j(p)$, $j = 0, 1$, is the support function of Z_j , $p \in \text{int} \{p \mid s_0(p) = 0\}$ implies $s_1(p) < \infty$ (cf. [22, Theorem 14.2]).

Step 2: Following the approach taken in [25, 26], we first consider a special case of the given problem, where the constraints (1.2), (2.1) are of a simpler form, namely as follows:

$$\begin{aligned} \dot{x}^i &= f^i(t, x, u), \quad u(t) \in U, \quad \psi^\alpha(t, x) \leq 0, \quad 1 \leq \alpha \leq A', \\ I^\gamma(a) &\leq p^\gamma, \quad 1 \leq \gamma \leq G', \quad I^\gamma(a) = p^\gamma, \quad G' < \gamma \leq G, \\ x^i(t_0) &= X_0^i(b), \quad x^i(t_1) = X_1^i(b) + p^{G+i}, \quad t_0, t_1 \text{ fixed.} \end{aligned} \quad (2.4)$$

For the problem of minimizing (1.1) subject to constraints (2.4), necessary optimality conditions are given in [25]. For our purposes we have to sketch the crucial ideas of the proof. On a set \mathcal{C} of arcs a which satisfy (2.4) except for the terminal constraints at t_0 and t_1 , real valued functionals $J^\varrho(a)$, $0 \leq \varrho \leq \varrho_0$, $\varrho_0 > G + n$, are defined with the following properties:

1. If $J^\varrho(a) = 0$ for all $\varrho > G + n$, then

$$J^\varrho(a) = I^\varrho(a), \quad 0 \leq \varrho \leq G, \quad J^{G+i}(a) = x^i(t_1) - X_1^i(b), \quad 1 \leq i \leq n,$$

and, as a consequence thereof,

2. a is an admissible solution of (1.1), (2.4), if and only if

$$\begin{aligned} J^\varrho(a) &\leq p^\varrho, \quad 1 \leq \varrho \leq G', \\ J^\varrho(a) &= p^\varrho, \quad G' < \varrho \leq G + n, \\ J^\varrho(a) &= 0, \quad \varrho > G + n. \end{aligned} \quad (2.5)$$

Further, a_0 is an optimal solution if and only if it minimizes $J^0(a)$ subject to (2.5).

Now let a_0 be an optimal arc of the unperturbed problem (1.1), (2.4), and let (1.3) hold. Then a derived cone K (this notion is defined in [15]) for the functionals J^e at a_0 on $\{a \in \mathcal{C} \mid J^e(a) = 0, \rho > G'\}$ is constructed. Then Hestenes' general multiplier rule, [15], is applicable and results in the existence of multipliers $(\lambda^0, \lambda) = (\lambda^0, \dots, \lambda^{e_0})$, with $\lambda^0 \geq 0, 0 \leq \gamma \leq G'$, which are related to the multipliers of Theorem 1.1 as follows:

$$\lambda^0 = -y^0, \quad \lambda^1, \dots, \lambda^G \text{ coincide, and } \lambda^{G+1} = -y^1(t_1). \quad (2.6)$$

For such problems of minimizing a function $J^0(a)$ subject to constraints of the form (2.5), where multipliers of the unperturbed problem are obtained via a derived cone, general perturbation results have been given in [12] and been refined in [13]. In our case here they say the following:

Let Λ be the set of multipliers (λ^0, λ) satisfying the above multiplier rule, and let A_0, A_1 be defined via Λ similarly to (2.2). If $\lambda p \geq 0$ for all $\lambda \in A_0$, then statements as in Theorem 2.1 hold with $s_j(p) = \sup \{-\lambda p \mid \lambda \in A_j\}, j = 0, 1$.

By relations (2.6), these terms $s_j(p)$ coincide with those defined in (2.3). Therefore it follows from the relations between (2.4) and (2.5), that Theorem 2.1 holds at least in case of constraints of type (2.4).

Step 3: In case of a problem with general constraints (1.2), respectively (2.1), the proof of Theorem 1.1 is carried out as follows (cf. [26]): Additional variables are introduced such that the modified problem is equivalent to the original one and has constraints of the form (2.4). This means, that again a derived cone is used to obtain the multipliers. It follows that perturbations p in the general problem $(P1)_p$ can be characterized exactly in the way shown above. This proves Theorem 2.1 ■

3. General perturbations

In this chapter we consider the following optimal control problem

$$\begin{aligned} (P2)_p: \quad & V(p) = \inf I^0(a; p) \quad \text{subject to} \\ & \dot{x}^i = f^i(t, x, u, p), \quad u(t) \in U, \\ & \psi^\alpha(t, x, p) \leq 0, \quad 1 \leq \alpha \leq A', \quad \psi^\alpha(t, x, p) = 0, \quad A' < \alpha \leq A, \\ & \theta^\beta(t, x, u, p) \leq 0, \quad 1 \leq \beta \leq B', \quad \theta^\beta(t, x, u, p) = 0, \quad B' < \beta \leq B, \\ & I^\gamma(a, p) \leq 0; \quad 1 \leq \gamma \leq G', \quad I^\gamma(a, p) = 0, \quad G' < \gamma \leq G, \\ & x^i(t_j) = X_j^i(b, p), \quad t_j = T^j(b, p), \quad j = 0, 1, \end{aligned}$$

where

$$I^\gamma(a, p) = g^\gamma(b, p) + \int_{t_0}^{t_1} L^\gamma(t, x(t), u(t), p) dt, \quad 0 \leq \gamma \leq G.$$

U is an open subset of \mathbf{R}^m , b ranges in some open subset of \mathbf{R}^k , and p is some fixed parameter in \mathbf{R}^S . All functions which involve p are assumed to be differentiable with respect to p .

For fixed p , this problem is of the form (1.1), (1.2). Without loss of generality we consider $(P2)_0$ as unperturbed problem. We assume that we know an optimal solution $a_0 = (x(t), u(t), \bar{b})$ of $(P2)_0$, along which assumption (1.3) holds, hence $\Omega \neq \emptyset$. Let Ω_0, Ω_1 be defined as in (2.2). We consider problems perturbed by p and characterize

the local behaviour of the optimal value function $V(p)$ of problem $(P2)_p$. To this end we define for $\omega \in \Omega$, and $1 \leq \sigma \leq S$:

$$\begin{aligned} \bar{F}_\sigma(t) = F_\sigma(a_0, \omega(t)) := & \lambda^r \frac{\partial L^r}{\partial p^\sigma}(t) + \mu^a(t) \frac{\partial \varphi^a}{\partial p^\sigma}(t) + \nu^b(t) \frac{\partial \theta^b}{\partial p^\sigma}(t) \\ & - y^i(t) \frac{\partial f^i}{\partial p^\sigma}(t); \end{aligned} \tag{3.1}$$

$$\begin{aligned} \bar{F}_\sigma = \bar{F}_\sigma(a_0, \omega(\cdot)) := & -y^0 \frac{\partial g^0}{\partial p^\sigma}(\bar{b}, 0) + \lambda^r \frac{\partial g^r}{\partial p^\sigma}(\bar{b}, 0) + \kappa^a \frac{\partial \psi^a}{\partial p^\sigma}(t_0) \\ & + (\kappa^a \psi_z^a(t_0) - y^i(t_0)) \frac{\partial X_0^i}{\partial p^\sigma}(\bar{b}, 0) + y^i(t_1) \frac{\partial X_1^i}{\partial p^\sigma}(\bar{b}, 0) \\ & + (\kappa^a \psi_t^a(t_0) + H(t_0)) \frac{\partial T^0}{\partial p^\sigma}(\bar{b}, 0) - H(t_1) \frac{\partial T^1}{\partial p^\sigma}(\bar{b}, 0); \end{aligned} \tag{3.2}$$

$$s_j(p) = \sup \left\{ \left(\bar{F}_\sigma + \int_{t_0}^{t_1} F_\sigma(t) dt \right) p^\sigma \mid \omega \in \Omega_j \right\}, \quad j = 0, 1. \tag{3.3}$$

Theorem 3.1: *Let a_0 be an optimal solution of problem $(P2)_0$ which satisfies assumption (1.3). Then concerning perturbed problems $(P2)_p$ all the statements of Theorem 2.1 hold analogously with $s_1(p)$ and $s_0(p)$ given by (3.3). In particular, these imply for*

$$V_+(0, p) := \liminf_{\substack{\varepsilon \rightarrow 0 \\ r(\varepsilon) = o(\varepsilon)}} \frac{V(\varepsilon p + r(\varepsilon)) - V(0)}{\varepsilon} \text{ and any } p:$$

- A) If $\Omega_1 \neq \emptyset$, then $V_+(0; p) \leq s_1(p)$.
- B) If $\Omega_1 = \emptyset$ and $s_0(p) < \infty$, then $V_+(0; p) = -\infty$.

Remark 3.2: Assuming several regularity properties of a_0 (which imply $\Omega_0 = \{0\}$), MAURER, [18–20], obtains the estimate

$$\limsup_{\varepsilon \rightarrow 0} \frac{V(\varepsilon p) - V(0)}{\varepsilon} \leq s_1(p).$$

Without those regularity assumptions this inequality does not hold any more and has to be replaced by the somewhat weaker one in the above theorem.

Remark 3.3: A novelty of the above result is the fact that sensitivity results are given also in the case of abnormality, i.e., when $\Omega_0 \neq \{0\}$. Below we indicate, what abnormality means with respect to sensitivity of the optimal value.

A) If $\Omega_1 \neq \emptyset$ and $\Omega_0 \neq \{0\}$, then typically $s_1(p)$ will be finite for some p and infinite for others. The case $s_1(p) = \infty$ may be related to one of the following situations (the list is not claimed to be complete):

a) Perturbed problems $(P2)_{\varepsilon p + r(\varepsilon)}$ have feasible solutions and V is ‘nice’ (say, V has a directional derivative in direction p), yet the first order necessary conditions of Theorem 1.1 give too little information, i.e., they allow for too many multipliers such that $s_1(p) = \infty$.

b) The perturbed problems have feasible solutions, but V has a vertical tangent in direction p . For example, such a situation may occur along switching curves of bang-bang solutions. In BOLTJANSKI, [2, pp. 31–39], a very simple example shows this effect.

c) Perturbed problems $(P2)_{p+o(\varepsilon)}$ have no feasible solutions. Actually, Theorem 3.1 implies also results of the following type: If for some p statement b) or c) from above is true, then each optimal arc a_0 of $(P2)_0$ is abnormal. Precise formulations of this fact would be rather lengthy and we omit them. For problems in mathematical programming they are given in [12, Corollary 3.4].

B) If $\Omega_1 = \emptyset$ (i.e., $y^0 = 0$ always) and $s_0(p) < \infty$, then for suitable sequences $\varepsilon_j \downarrow 0$, $r(\varepsilon_j)$, we have by Theorem 3.1 that $\lim_{j \rightarrow \infty} (V(\varepsilon_j p + r(\varepsilon_j)) - V(0))/\varepsilon_j = -\infty$. This

means, roughly speaking, that in such a direction p the optimal value becomes drastically better. However, in detail the situation may be rather complicated. Using a very simple example without state constraints, Example 6.2 shows the following. We have $\Omega_1 = \emptyset$, for a certain direction p we have $s_0(p) = 0$, the directional derivative $\lim_{\varepsilon \downarrow 0} (V(\varepsilon p) - V(0))/\varepsilon$ exists and is finite, and only for suitable functions $r(\varepsilon)$ of type $o(\varepsilon)$ we have $\lim_{\varepsilon \downarrow 0} (V(\varepsilon p + r(\varepsilon)) - V(0))/\varepsilon = -\infty$.

Proof of Theorem 3.1: We take $\hat{x} = (x^1, \dots, x^{n+S}) := (x, p)$ as new state variable. Then problem $(P2)_p$ becomes the following:

$$\begin{aligned} V(p) &= \inf I^0(\hat{a}) \text{ subject to} \\ \dot{x}^i &= f^i(t, \hat{x}, u), \quad 1 \leq i \leq n, \quad \dot{x}^{n+\sigma} = 0, \quad 1 \leq \sigma \leq S, \\ \psi^\alpha(t, \hat{x}) &\leq 0, \quad 1 \leq \alpha \leq A', \quad \psi^\alpha(t, \hat{x}) = 0, \quad A' < \alpha \leq A, \\ \theta^\beta(t, \hat{x}, u) &\leq 0, \quad 1 \leq \beta \leq B', \quad \theta^\beta(t, \hat{x}, u) = 0, \quad B' < \beta \leq B, \\ I^\gamma(\hat{a}) &\leq 0, \quad 1 \leq \gamma \leq G', \quad I^\gamma(\hat{a}) = 0, \quad G' < \gamma \leq G, \\ x^i(t_j) &= X_j^i(b, c), \quad 1 \leq i \leq n, \quad j = 0, 1, \quad x^{n+\sigma}(t_0) = c^\sigma, \\ x^{n+\sigma}(t_1) &= p^\sigma, \quad 1 \leq \sigma \leq S, \quad t_j = T^j(b, c), \quad u(t) \in U, \end{aligned} \quad (3.4)$$

where $\hat{a} = (\hat{x}(\cdot), u(\cdot), b, c)$.

It is easy to see that there is a one-to-one correspondence between admissible (resp. optimal) arcs of $(P2)_p$ and (3.4). The advantage of (3.4), however, is that it contains the parameter p only in form of a standard perturbation as treated in Chapter 2. Therefore p -perturbed problems are characterized through Theorem 2.1 by $s_j(p) = \sup \{y^{n+\sigma}(t_1) p^\sigma \mid \hat{\omega}(\cdot) \in \hat{\Omega}_j\}$, where $\hat{\Omega}$ denotes the set of all multipliers related to an optimal solution \hat{a}_0 of (3.4) with $p = 0$. Checking the respective necessary conditions point by point, one sees the following:

$$\begin{aligned} (y^0, \kappa, \lambda, \mu(\cdot), \nu(\cdot), y(\cdot)) &\text{ is a multiplier for } a_0 \text{ if and only if} \\ (y^0, \kappa, \lambda, \mu(\cdot), \nu(\cdot), y(\cdot), y^{n+1}(\cdot), \dots, y^{n+S}(\cdot)) &\end{aligned} \quad (3.5)$$

is a multiplier for \hat{a}_0 .

Let $1 \leq \sigma \leq S$. Then the transversality conditions for the arc \hat{a}_0 yield

$$y^{n+\sigma}(t_0) = \bar{F}_\sigma, \quad \bar{F}_\sigma \text{ as in (3.2)}. \quad (3.6, i)$$

Part (iii) of Theorem 1.1 yields

$$\dot{y}^{n+\sigma}(t) = F_\sigma(t), \quad F_\sigma \text{ as in (3.1)}. \quad (3.6, ii)$$

Together with (3.5) it follows that $s_j(p)$ above coincides with (3.3) (the additional multipliers $y^{n+\sigma}$ are eliminated by (3.6), and because of (3.5) the multipliers on the right-hand sides of (3.6) are actually elements of Ω_j). This proves the first part of

Theorem 3.1. The results about $V_+(0; p)$ are immediate consequences thereof. Note that if $s_0(p) = \infty$ in case A), then also $s_1(p) = \infty$, and hence the claim is trivial. ■

Theorem 3.1 gives very general results, but usually one is rather interested in particular perturbations, where a parameter p^o enters only one type of the constraints or only a single constraint. Then the results become intuitively more appealing. In the following corollary we carry this out for some cases of particular interest.

Corollary 3.4: In the case 1. — 4. listed below, let q be a (scalar) component of a perturbation parameter p and let q enter only constraints of form (i) described below. For convenience we assume that $\Omega_1 \neq \emptyset$ and that V has directional Hadamard derivatives $V'(0, d) = \lim_{\epsilon \rightarrow 0, d' \rightarrow d} \frac{V(\epsilon d') - V(0)}{\epsilon}$ in all directions d .

I. Then in particular the right- respectively left hand partial derivatives of V with respect to q at $p = 0$ (denoted by $\frac{\partial V}{\partial q}(0+)$ respectively $\frac{\partial V}{\partial q}(0-)$) exist and together with the terms $s(\omega)$ given by (ii) below they satisfy

$$\frac{\partial V}{\partial q}(0+) \leq \sup \{s(\omega) \mid \omega \in \Omega_1\}, \quad \frac{\partial V}{\partial q}(0-) \geq \inf \{s(\omega) \mid \omega \in \Omega_1\}.$$

II. If further the partial derivative $\frac{\partial V}{\partial q}(0)$ exists and if Ω_1 contains only one element $\bar{\omega}$, then $\frac{\partial V}{\partial q}(0) = s(\bar{\omega})$.

We consider the following perturbed constraints (i), whose effect on V is described by the respective term (ii):

1. Perturbation of the terminal constraint:

$$(i) \quad x(t_1) = X_1(b, q) \quad (ii) \quad s(\omega) = y^i(t_1) \frac{\delta X_1^i}{\delta q}(\bar{b}, 0)$$

2. Perturbation of the initial constraint:

$$(i) \quad x(t_0) = X_0(b, q) \quad (ii) \quad s(\omega) = (x^\alpha / \psi_x^\alpha(t_0) - y^i(t_0)) \frac{\partial X_0^i}{\partial q}(\bar{b}, 0)$$

3. Perturbation of the state constraints:

$$(i) \quad \psi^\alpha(t, x, q) \begin{cases} \leq 0, & 1 \leq \alpha \leq A' \\ = 0, & A' < \alpha \leq A \end{cases} \quad (ii) \quad s(\omega) = x^\alpha \frac{\partial \psi^\alpha}{\partial q}(t_0) + \int_{t_0}^{t_1} \mu^\alpha(t) \frac{\partial \psi^\alpha}{\partial q}(t) dt.$$

As a special case thereof, let $\alpha, 1 \leq \alpha \leq A'$, be fixed and $\psi^\alpha(t_0) < 0$ and $\psi^\alpha(t_1) < 0$. Then:

$$(i) \quad \psi^\alpha(t, x) \leq q \quad (ii) \quad s(\omega) = \mu^\alpha(t_1) - \mu^\alpha(t_0) = -\mu^\alpha(t_0) (\leq 0).$$

4. Perturbation of the mixed constraints:

$$(i) \quad \theta^\beta(t, x, u, q) \begin{cases} \leq 0, & 1 \leq \beta \leq B' \\ = 0, & B' < \beta \leq B \end{cases} \quad (ii) \quad s(\omega) = \int_{t_0}^{t_1} v^\beta(t) \frac{\partial \theta^\beta}{\partial q}(t) dt.$$

As a special case thereof, let $\beta, 1 \leq \beta \leq B'$, be fixed. Then:

$$(i) \quad \theta^\beta(t, x, u) \leq q \quad (ii) \quad s(\omega) = - \int_{t_0}^{t_1} v^\beta(t) dt (\leq 0).$$

Of course, similar results can be stated also for perturbations of the dynamics f^t , of the end times t_0, t_1 , or of the isoperimetric constraints.

Remark 3.5: In [18–20] MAURER considers also perturbations p in a Banach space, which is out of reach in this approach. However, the treatment of mixed constraints and perturbations thereof seems to be difficult in the function space approach. Further, the preceding results are restricted to so-called first order state constraints $\psi^a(t, x) \leq 0$ (cf. [19]) which is implied by condition (1.3). For a scalar control u , [18–20] give similar results also in case of higher constraints (where $\varphi_u^a(t) = 0$ along boundary arcs, but for a higher derivative of ψ^a the partial derivative with respect to u is nonzero). At least for state constraints of order 2 such results can be obtained along the same lines as here. The suitable necessary conditions are provided by [29] and proved in the same way as sketched in Chapter 2. However, a detailed account would lead too far here.

Remark 3.6: If in a given problem U is not an open set (as required in (1.2)), the often this can be handled by adding suitable constraints. However, sometimes this procedure may be too complicated or violate condition (1.3). Then the following considerations may be useful: In [16], KNOBLOCH develops necessary optimality conditions for an optimal control problem with arbitrary control set U and one state constraint $\psi(t, x) \leq 0$. Instead of introducing a multiplier μ attached to ψ or φ , he derives a maximum principle with the classical Hamiltonian. Then the adjoint variables $y^i(\cdot)$ may be discontinuous at those times t , where an optimal arc α_0 hits the boundary $\psi(t, x) = 0$. Roughly speaking, it is shown in [8] that the sensitivity with respect to perturbations $\psi(t, x) \leq p$ is described by the sum of all jumps occurring in the adjoint variables. By suitable reformulations this result can be shown to be equivalent to part 3 of Corollary 3.4.

Remark 3.7: If the optimal solution of the unperturbed problem is not unique, then obviously all results in this chapter can be improved as follows. If M is a set of optimal solutions satisfying (1.3), then the results of Theorem 3.1 hold with each $\alpha \in M$ and hence yield upper bounds for $V_+(0, p)$ of the type $\inf_{\alpha \in M} s_1(p, \alpha)$.

4. Higher order conditions

Theorem 3.1 gives upper bounds of the lower Hadamard derivative $V_+(0; p)$ of the optimal value function V at 0. It may well occur that sometimes for some p one has only $V_+(0; p) < s_1(p)$, with a significant gap in between. One possible reason for such a gap may be the fact that first order necessary conditions as in Theorem 1.1 are not precise enough, i.e., they allow for too many multipliers. As a consequence, the supremum in (3.3) will be too large. In such cases, multiplier sets satisfying higher order necessary conditions may lead to upper bounds which are lower than those of Theorem 3.1.

For standard optimal control problems without state constraints, KNOBLOCH [17] proved a set of higher order necessary conditions which have the following properties:

1. In addition to the usual (first order) maximum principle, higher order conditions such as the generalized Legendre-Clebsch-conditions hold. Thus these conditions are satisfied by a multiplier set $\tilde{\Omega}$ which is contained in Ω .
2. This is achieved via the construction of a derived cone which contains the cone defined by first order approximations (the cone of attainability of Pontryagin).

Therefore, if $\tilde{s}_j(p)$ is defined similarly as $s_j(p)$ in (3.3), but with Ω replaced by $\tilde{\Omega}$, then all the conclusions of Theorem 3.1 hold also with $\tilde{s}_j(p)$ instead of $s_j(p)$, $j = 0, 1$. Since $\tilde{\Omega} \subseteq \Omega$ implies $\tilde{s}_j(p) \leq s_j(p)$, these higher order conditions may give better sensitivity results than those of Theorem 3.1, if the inclusion $\tilde{\Omega} \subseteq \Omega$ is proper. [10] contains such an example involving a simple system with three state and two control variables. There the generalized Legendre-Clebsch-condition results in a multiplier set $\tilde{\Omega}$ such that $V_+(0; p) = V'(0, p) = \tilde{s}_1(p) < s_1(p)$ for some directions p .

In a similar way one can obtain strengthened sensitivity results by means of the second order conditions in [28].

5. Generalized subgradients of the optimal value function

In Chapter 3 we proved that for any optimal solution a of $(P2)_0$ which satisfies (1.3), we have for all p that

$$\begin{aligned} V_+(0; p) &\leq s_1(p, a), & \text{if } \Omega_1(a) \neq \emptyset, \\ V_+(0; p) &= -\infty, & \text{if } \Omega_1(a) = \emptyset \text{ and } s_0(p, a) = 0. \end{aligned} \tag{5.1}$$

In order to draw further conclusions thereof, we denote for $1 \leq \sigma \leq S$, $j = 0, 1$:

$$\begin{aligned} \pi^\sigma(\omega) &= \bar{F}_\sigma(\omega) + \int_{t_0}^{t_1} F_\sigma(\omega(t)) dt, \\ \Pi &= \{(\pi^1, \dots, \pi^S) \mid \pi^\sigma = \pi^\sigma(\omega); (y^0, \omega(\cdot)) \in \Omega\}, \\ \Pi_j &= \{(\pi^1, \dots, \pi^S) \mid \pi^\sigma = \pi^\sigma(\omega); \omega \in \Omega_j\} \quad (\subseteq \Pi). \end{aligned} \tag{5.2}$$

Since Ω is a convex cone and Π is the image of Ω under a linear transformation, Π is a convex cone in \mathbb{R}^S . For similar reasons Π_0 is a convex cone and Π_1 is a convex set. We have $0 \in \Pi_0$, since $0 \in \Omega_0$, and $\Pi_1 + \Pi_0 \subseteq \Pi_1$, if $\Pi_1 \neq \emptyset$. If we have a set M of optimal solutions as in Remark 3.7, the sets in (5.2) depend on the respective solution $a \in M$. We denote this by $\Pi(a)$, $\Pi_j(a)$.

Let $\delta^*(p \mid C) := \sup \{cp \mid c \in C\}$ denote the support function of a convex set C . Then (5.1) implies that for any $a \in M$ and for all p

$$V_+(0; p) \leq \delta^*(p \mid \Pi_1(a)), \quad \text{if } \Pi_1(a) \neq \emptyset. \tag{5.3}$$

Relation (5.3) seems to be close to a statement about subgradients of V . However, since even 'nice' data of the problem imply not even continuity of V , a first difficulty arises in a suitable definition of ∂V . In order to cover a large class of cases we use a general notion introduced by Rockafellar, which assumes only (strict) lower semicontinuity of V at 0. There V is called strict lower semicontinuous at $p = 0$, if there exists some $\delta > V(0)$ such that $\min(\delta, V(p))$ is lower semicontinuous on a neighbourhood of $p = 0$ (for more details see [24]). For such a function V the generalized directional derivative at 0 in direction p is given by

$$V'(0; p) = \lim_{\epsilon \downarrow 0} \limsup_{\eta \rightarrow 0, |\eta - p| \leq \epsilon} \inf_{\tau} \frac{V(\tau p' + \eta) - V(\eta)}{\tau}, \tag{5.4}$$

where $\eta \rightarrow 0$ if and only if $\eta \rightarrow 0$ and $V(\eta) \rightarrow V(0)$. $V'(0; p)$ is a function convex and positively homogeneous in p , (cf. [23]). Hence it makes sense to define

$$\partial V(0) = \{y \in \mathbb{R}^S \mid yp \leq V'(0; p) \text{ for all } p\}.$$

This is a convex and closed set, possibly empty. It is compact and nonempty if and only if V is locally Lipschitz around $p = 0$. In a special case, (5.3) leads directly to estimates of $\partial V(0)$, namely when V is subdifferentially regular at $p = 0$ (cf. [23]), which implies that

$$V_+(0; p) = V'(0; p) \quad \text{for all } p. \quad (5.5)$$

Examples of subdifferentially regular functions are convex or continuously differentiable functions.

Theorem 5.1: Let M be defined as in Remark 3.7 and V be subdifferentially regular at $p = 0$.

(i) If $\Pi_1(a) \neq \emptyset$ for all $a \in M$, then $\partial V(0) \subseteq \bigcap_{a \in M} \text{cl } \Pi_1(a)$.

ii) If for some $a \in M$, $\Pi_1(a) = \emptyset$ and $\Pi_0(a)$ is not the whole space, then $\partial V(0) = \emptyset$.

Remark: Concerning (ii) above, note that $\Pi_0(a) = \mathbf{R}^s$ might be possible only in very pathological cases which are practically out of interest.

Proof: (i) From (5.3) and (5.5) it follows that for any $a \in M$, $\partial V(0) \subseteq \text{cl } \Pi_1(a)$.

(ii) By the assumption there is some p with $s_0(p, a) = 0$ (since Π_0 is a cone). Then by (5.1) $V'(0; p) = -\infty$ and this implies $\partial V(0) = \emptyset$ ■

Without assumption (5.5), results as in Theorem 5.1 are much harder to prove. For finite dimensional problems in nonlinear programming a weak regularity condition introduced by Rockafellar and called *tameness* is needed (cf. [24], also [13]). In the following we do something similar. However, in order not to blow up the technical parts of the proof too much, we limit ourselves to standard control problems without state constraints. That means that as in [15, chapter 6.2] we consider the following problem:

minimize $I^0(a)$ subject to

$$\dot{x}^i = f^i(t, x, u), \quad 1 \leq i \leq n, \quad u(t) \subseteq U,$$

$$I^r(a) \leq p^r, \quad 1 \leq r \leq G', \quad I^r(a) = p^r, \quad G' < \gamma \leq G, \quad (P3)_p$$

$$x^i(t_0) = x_0^i, \quad x^i(t_1) = X_1^i + p^{\sigma+i}, \quad 1 \leq i \leq n, \quad t_0, t_1 \text{ fixed.}$$

There U is some fixed subset of \mathbf{R}^m , and I^r are defined as in (1.2), except that the p^r are now constants. Note that in this case $\Pi = Z$, $\Pi_j = Z_j$, $j = 0, 1$, with Z, Z_0, Z_1 as in Step 1 of the proof of Theorem 2.1, and hence Π, Π_j are closed subsets of $\mathbf{R}^s = \mathbf{R}^{G+n}$. Further, $s_j(p, a) = \delta^*(p | \Pi_j(a))$, $j = 0, 1$. As analogue to the tameness assumption in [24] we use the following regularity condition:

There exists some $\delta > V(0)$, such that for any sequence (η_j) in the p -space with $\eta_j \rightarrow 0$ and $V(\eta_j) < \delta$ there exist a subsequence (again denoted by (η_j)), optimal solutions $a_j = (x_j, u_j)$ of $(P3)_{\eta_j}$ and an optimal solution $\bar{a} = (\bar{x}, \bar{u}) = (x, u)$ of $(P3)_0$, such that

$$\lim_{j \rightarrow \infty} \int_{t_0}^{t_1} |u_j(t) - \bar{u}(t)| dt = 0. \quad \text{There } |\cdot| \text{ denotes some norm in } \mathbf{R}^m.$$

Note that (5.6) requires neither uniqueness of the respective optimal solutions nor the existence of solutions for a whole neighbourhood of $p = 0$. In particular, (5.6) does not require continuity of V .

Theorem 5.2: *Let M be the set of all optimal solutions of $(P3)_0$ and let assumption (5.6) hold. Then*

- (i) V is strictly lower semicontinuous at $p = 0$.
- (ii) $V^\uparrow(0; p) \leq \delta^* \left(p \mid \bigcup_{a \in M} \Pi_1(a) + \bigcup_{a \in M} \Pi_0(a) \right)$ for all p , if $\Pi_1(a) \neq \emptyset$ for some $a \in M$.
 Moreover, if $\Pi_1(a) = \emptyset$ and $s_0(\bar{p}, a) < \infty$ for all $a \in M$, then $V^\uparrow(0; \bar{p}) = -\infty$ and $\partial V(0) = \emptyset$.
- (iii) $\partial V(0) \subseteq \text{cl co} \left[\bigcup_{a \in M} \Pi_1(a) + \bigcup_{a \in M} \Pi_0(a) \right]$.

Proof: Fix some p and some sequence (η_j) for which (5.6) applies. Let a_j, \bar{a} be as in (5.6).

Step 1: To prove (i), let $\bar{K}(\bar{a})$ resp. $\bar{K}(a_j)$ be the derived sets related to \bar{a} resp. a_j as defined in [15, p. 268]. $\bar{K}(a_j)$ has the form

$$\{k = (k^0, \dots, k^S) \mid k^\sigma = F_j^\sigma(t, x_j(t), u) - F_j^\sigma(t, x_j(t), u_j(t)), 0 \leq \sigma \leq S, \\ t \in (t_0, t_1), u \in U, u_j(\cdot) \text{ continuous at } t\},$$

with certain functions F_j^σ defined via solutions of the adjoint equations related to a_j (cf. [15: pp. 266–270]). $\bar{K}(\bar{a})$ is defined analogously.

Claim: $k \in \bar{K}(\bar{a}) \Rightarrow$ there exist $k_j \in \bar{K}(a_j)$ with $k_j \rightarrow k$. (5.7)

Proof: The components of k are of the form

$$k^\sigma = F^\sigma(\bar{t}, \bar{x}(\bar{t}), u) - F^\sigma(\bar{t}, \bar{x}(\bar{t}), \bar{u}(\bar{t})), \quad 0 \leq \sigma \leq S,$$

with some $\bar{t}, u, (t_0 < \bar{t} < t_1)$. (5.6) implies that $u_j(t) \rightarrow \bar{u}(t)$ for almost all t . Therefore one can find $t_j \rightarrow \bar{t}$ such that $u_j(\cdot)$ is continuous at t_j and $u_j(t_j) \rightarrow \bar{u}(\bar{t})$. Now define k_j by

$$k_j^\sigma = F_j^\sigma(t_j, x_j(t_j), u) - F_j^\sigma(t_j, x_j(t_j), u_j(t_j)), \quad (0 \leq \sigma \leq S),$$

with the same u as above in k . By [15, Appendix, Theorem 4.1], $x_j(t) \rightarrow \bar{x}(t)$ for every t . For the same reasons, the solutions of the adjoint equations related to \bar{a} , converge with $j \rightarrow \infty$ to the solutions of the adjoint equation related to \bar{a} . But these solutions define F_j^σ and F^σ . Therefore $k_j \rightarrow k$, which proves the claim (5.7).

Furthermore, these facts imply the following: If $V(\eta_j) < \delta$ for some sequence $\eta_j \rightarrow 0$, then $\liminf_{j \rightarrow \infty} V(\eta_j) = \liminf_{j \rightarrow \infty} I^0(a_j) = I^0(\bar{a}) = V(0)$. This proves (i).

Step 2: Proof of (ii). If $D := \delta^*(p \mid \Pi_1(\bar{a}) + \Pi_0(\bar{a})) = \infty$, then (ii) is trivial. If D is finite, choose $s > D$. In case that $\Pi_1(\bar{a}) = \emptyset$ and $s_0(p, \bar{a}) = 0, s \in \mathbb{R}$ can be chosen arbitrarily. Let $K(\bar{a})$ resp. $K(a_j)$ be the convex cones generated by $\bar{K}(\bar{a})$ resp. $\bar{K}(a_j)$. Then (5.7) holds analogously also with $K(\bar{a})$ and $K(a_j)$. These cones are derived cones for functionals J^σ at \bar{a} (resp. for J_j^σ at a_j) as in (2.5).

Claim: For any $\varepsilon > 0$ there exists some j_0 such that

$$\text{dist}((s, p), K(a_j) - \text{cl } L_{s,p}) \leq \varepsilon \quad \text{for all } j \geq j_0, \tag{5.8}$$

where

$$L_{s,p} = \{t = (t^0, \dots, t^S) \mid t^0 < xs, t^\sigma < xp^\sigma, 1 \leq \sigma \leq G', t^\sigma = xp^\sigma, \sigma > G'; \\ x > 0\}.$$

Proof: By the proof of Theorem 2.1 and [13: Proposition 3.1] we have $\text{dist}((s, p), K(\bar{a}) - \text{cl } L_{s,p}) = 0$, i.e., there exist $\bar{k} \in K(\bar{a}), \bar{l} \in \text{cl } L_{s,p}$, with $(s, p) = \bar{k} - \bar{l}$. Then (5.7), applied to $K(\bar{a}), K(a_j)$, yields (5.8).

Now fix $\varepsilon > 0$ and some sequence $\tau_j \downarrow 0$. By (5.8) there is some j_0 such that for any $j \geq j_0$ there exist $k_j \in K(a_j)$, $l_j \in L_{s,p}$, Δ_j with $|\Delta_j| \leq \varepsilon/2$ such that $(s, p) = k_j - l_j + \Delta_j$. Since $K(a_j)$ and $L_{s,p}$ are cones, this implies that there are $k_j \in K(a_j)$, Δ_j with $|\Delta_j| \leq \varepsilon/2$ such that

$$\begin{aligned} k_j^\sigma &< s + \Delta_j^\sigma, & k_j^\sigma &< p^\sigma + \Delta_j^\sigma, & 1 &\leq \sigma \leq G', & k_j^\sigma &= p^\sigma + \Delta_j^\sigma, \\ \sigma &> G'. \end{aligned} \quad (5.9)$$

By the definition of a derived cone, this implies that for any $j \geq j_0$ there exist $\delta_j > 0$ and a family of arcs $a_{j,\tau}$ satisfying the initial conditions of problem (P3) at t_0 , $\tau \in [0, \delta_j]$, such that

$$J_j^\sigma(a_{j,\tau}) = \eta_j^\sigma + \tau k_j^\sigma + r_j^\sigma(\tau), \quad \text{with } r_j^\sigma(\tau) = o(\tau). \quad (5.10)$$

By (5.6), [15, Appendix, Theorem 4.1] and arguments as in [15: p. 269], made separately for every j , it is possible to choose $\delta_j = \delta' > 0$, independent of j , and δ' small enough such that $|r_j^\sigma(\tau)/\tau| \leq \varepsilon/2$ for all $\tau \in (0, \delta')$ and $j \geq j_0$. For some $j_1 \geq j_0$ we have $\tau_j \leq \delta'$ for all $j \geq j_1$, and hence (5.9) and (5.10) imply: For any $j \geq j_1$ there is some p_j with $|p - p_j| \leq \varepsilon$ and an admissible solution \bar{a}_j of problem (P3) $_{\eta_j + \tau_j p_j}$ with $I^0(\bar{a}_j) \leq \tau_j(s + \varepsilon) + I^0(a_j)$. In particular, $V(\eta_j + \tau_j p_j) - V(\eta_j) \leq \tau_j(s + \varepsilon)$. Since ε , η_j , τ_j are freely chosen and $|p_j - p| \leq \varepsilon$, statement (ii) follows from the choice of s .

Step 3: Part (iii) follows from (ii), as it has been demonstrated in [13, Theorem 3.6]. ■

Remark 5.3: In Chapter 4 we showed how one can sharpen the results of Chapter 3 by using multiplier sets \bar{Q} which satisfy higher order necessary conditions. In a similar way one can refine also the previous results in this chapter. To this end let \bar{I}_j , \bar{H}_j be defined by \bar{Q} similarly as in (5.2). Then the results of Chapter 4 imply the analogue to (5.3), namely $V_+(0; p) \leq \delta^*(p | \bar{H}_1(a))$, if $\bar{H}_1(a) \neq \emptyset$. Hence it is clear that Theorem 5.1 holds also with the sets $\bar{H}_0(a)$, $\bar{H}_1(a)$. The proof of a refined version of Theorem 5.2, however, would exceed reasonable length.

6. Two examples

In the literature one can find a number of papers which apply the necessary-optimality conditions of Theorem 1.1 or a similar form to optimal control problems of practical interest. However, sometimes they lack mathematical rigour and lead to wrong results. In [30] for example, a nuclear reactor model involving mixed constraints is treated. The problem is stated over a time interval with free, but finite terminal time t_1 . The author presents an 'optimal' solution with $t_1 = \infty$, without realizing that consequently the problem (as it is stated) would not have an optimal solution at all. In [19, 20], on the contrary, two control problems in engineering with state inequality constraints are treated and a careful sensitivity analysis is given. With respect to mixed constraints [31] presents a nice example. The authors deal with an economic problem which involves three state and three control variables and five mixed constraints, and they use the conditions of Theorem 1.1 to characterize an optimal solution. Since sensitivity analysis of such constraints is a novelty of our approach, we state this problem in more detail.

Example 6.1 ([31]): Find $(b, u(\cdot), v(\cdot), w(\cdot))$ such that

$$bc\varphi - \int_{t_0}^{t_1} (gux^3 - du - rx^2 - cv) \varphi dt \text{ is minimized subject to}$$

$$\dot{x}^1 = v - ax^1, \quad \dot{x}^2 = cv + rx^2 + du - gux^3, \quad \dot{x}^3 = w,$$

$$x(t_0) = (x_0^1, x_0^2, x_0^3), \quad x_0^1 > 0; \quad x^1(t_1) = -b;$$

$$gu - \alpha + \beta x^3 - \gamma w \leq 0, \quad 0 \leq v \leq M, \quad 0 \leq u(t) \leq x(t).$$

$a, c, d, r, g, \alpha, \beta, \gamma, \varphi$ are functions of the time t and continuously differentiable on $[t_0, t_1]$. γ is negative, all others are positive on $[t_0, t_1]$.

In [31] the necessary conditions of Theorem 1.1 are applied and result in explicit formulas for the multipliers. Furthermore, the adjoint variables y^1, y^2, y^3 are interpreted according to [1] as marginal values, however without checking the set of assumptions made in [1]. The similar statements of Theorem 3.1 hold in any case. Here we limit ourselves to a short discussion of a scalar perturbation p of one of the mixed constraints, namely

$$0^1(t) = g(t)u(t) - \alpha(t) + \beta(t)x^3(t) - \gamma(t)w(t) \leq p.$$

Let $v^1(\cdot)$ be the related multiplier. Then, by [31], $v^1(t) = -y^3(t)/\gamma(t)$, which is nonnegative and continuous on $[t_0, t_1]$. $y^3(\cdot)$ is uniquely defined by the adjoint differential equation and $y^3(t_1) = 0$. If we assume that $V(p)$ is differentiable at $p = 0$ (which is a reasonable assumption, since here V is a monotone decreasing function), then by Corollary 3.4. II we have

$$\frac{\partial V}{\partial p}(0) = \int_{t_0}^{t_1} \frac{y^3(t)}{\gamma(t)} dt \quad (\leq 0).$$

The second example is to illustrate Remark 3.3.B and is of purely theoretical interest. Nevertheless in the following example one may think of a boat with maximal speed 1 moving on a river with a homogeneous current faster than 1.

Example 6.2 ([9]): Minimize t_1 subject to

$$\dot{x}^1 = c + u$$

$$\dot{x}^2 = v$$

$$U = \{(u, v) \mid u^2 + v^2 \leq 1\}, \quad c > 1.$$

$$x(0) = (0, 0), \quad x(t_1) = x_1 = (a, b), \quad \text{with } a, b > 0, \quad a^2 = (c^2 - 1)b^2.$$

This problem is not quite of form (P1), but rather than reformulating it we use the following supplement to Remark 3.6: Applying the usual maximum principle (cf. [2, 15]) for a problem with free final time t_1 , the adjoint variables $y^1(\cdot), y^2(\cdot)$ describe perturbations of (a, b) in the same way as those in Theorem 3.1. Without proof we state the following properties of the above problem:

1. Since $c > 1 \geq u^2 + v^2$, there is a set $P \subseteq \mathbb{R}^2$ such that the problem has a (admissible or optimal) solution if and only if $(a, b) \in P$. Actually P is a convex cone, and $a^2 = (c^2 - 1)b^2, a > 0$, describes the boundary points of P . Thus above we have $x_1 \in \partial P$.

2. The (unique) optimal control is $(\bar{u}(t), \bar{v}(t)) = (-1/c, \sqrt{c^2 - 1}/c)$, with $t_1 = ac/(c^2 - 1)$. The adjoints y^1, y^2 are constant on $[0, t_1]$ and have the values $(y^1, y^2) = \alpha(\bar{u}, \bar{v})$, with any $\alpha > 0$.

3. Since the Hamiltonian $H = y^0 + y^1(c + u) + y^2v$ vanishes identically along $(u(t), v(t))$, we get $y^0 = 0$. Thus $\Pi_1 = \emptyset, \Pi_0 = \{\beta(-1, \sqrt{c^2 - 1}) \mid \beta > 0\}$.

4. Let $p = (p^1, p^2)$ be a perturbation of $x_1 = (a, b)$. If $x_1 + p \notin P$, then $s_0(p) = \infty$, and Theorem 3.1 makes no statement. Otherwise $s_0(p) = 0$, and according to Theorem 3.1 we get $V_+(0; p) = -\infty$. Now choose in particular $\bar{p} = (\sqrt{c^2 - 1}, 1)$. Then $x_1 + \bar{p} \in \partial P$, and by note 2 we get $V(\varepsilon\bar{p}) = (a + \varepsilon\sqrt{c^2 - 1})c/(c^2 - 1)$. Therefore $\lim_{\varepsilon \downarrow 0} \frac{V(\varepsilon\bar{p}) - V(0)}{\varepsilon} = c\sqrt{c^2 - 1}$, whereas $V_+(0; \bar{p}) = -\infty$. This means that only for suitable functions $r(\varepsilon)$ of type $o(\varepsilon)$ with $x_1 + \varepsilon\bar{p} + r(\varepsilon) \in \text{int } P$ we have $\lim_{\varepsilon \downarrow 0} (V(\varepsilon\bar{p} + r(\varepsilon)) - V(0))/\varepsilon = -\infty$.

5. If $p(\bar{u}, \bar{v}) \leq 0$, then Theorem 3.1 implies that $V_+(0; p) = -\infty$. One sees directly from (5.4) that for all those p also $V'(0; p) = -\infty$. If $p(\bar{u}, \bar{v}) > 0$, then for $\varepsilon > 0$, small enough, and any $\tau > 0$ and any p' with $|p' - p| \leq \varepsilon$ we have $x_1 + \tau p' \in P$, hence $V(\tau p') = +\infty$. Therefore $V_+(0; p) = V'(0; p) = +\infty$, and V is subdifferentially regular at $p = 0$. Thus Theorem 5.1, (ii) applies and yields $\partial V(0) = \emptyset$.

REFERENCES

- [1] BERKOVITZ, L. D.: Variational Methods in Problems of Control and Programming. *J. Math. Anal. Appl.* **3** (1961), 145–169.
- [2] BOLTJANSKI, W. G.: Mathematische Methoden der optimalen Steuerung. Hanser: München 1972.
- [3] BUNCE, G. R.: A Maximum Principle for Time-Lag Control Problems with Bounded State. *J. Optimization Theory Appl.* **22** (1977), 563–606.
- [4] DONTCHEV, A. L.: Sensitivity Analysis of Linear Infinite-Dimensional Optimal Control Systems under Changes of System Order. *Control Cybernetics* **3** (1974), 21–35.
- [5] —: Sensitivity Analysis of Optimal Control Systems with small time delay. *Control Cybernetics* **4** (1975), 91–104.
- [6] DONTCHEV, A. L., and T. R. GICEV: Convex Singularly Perturbed Optimal Control Problem with Fixed Final State: Controllability and Convergence. *Math. Operationsforsch. Statist., Ser. Optimization* **10** (1979), 345–355.
- [7] EVERS, A. H.: Sensitivity Analysis in Dynamic Optimization. *J. Optimization Theory Appl.* **32** (1980), 17–37.
- [8] GOLLAN, B.: Optimal Control Probleme mit parameterabhängigen Zustandsbeschränkungen. Diplomarbeit: Universität Würzburg 1975.
- [9] —: Störungstheorie für abstrakte Optimierungsprobleme mit Anwendungen auf die Theorie optimaler Steuerungen. Dissertation: Universität Würzburg 1979.
- [10] —: Sensitivity Analysis in Optimization with Applications to Optimal Control Theory. In: *Lecture Notes in Mathematics* Vol. **44**. Marcel Dekker: New York 1979, 153–170.
- [11] —: On Optimal Control Problems with State Constraints. *J. Optimization Theory Appl.* **32** (1980), 75–80.
- [12] —: Perturbation Theory for Abstract Optimization Problems. *J. Optimization Theory Appl.* **35** (1981), 317–341.
- [13] —: On the Marginal Function in Nonlinear Programming. *Math. Oper. Res.* (to appear)
- [14] HALANAY, A.: Optimal Controls for Systems with Time Lags. *SIAM J. on Control* **6** (1968), 215–234.
- [15] HESTENES, M. R.: Calculus of Variations and Optimal Control Theory. Wiley: New York 1966.

- [16] KNOBLOCH, H. W.: Das Pontryaginsche Maximumprinzip für Probleme mit Zustandsbeschränkung. I, II. ZAMM 55 (1975), 545–556 und 621–634.
- [17] —: Higher Order Necessary Conditions in Optimal Control Theory. Lecture Notes in Control and Information Sciences, Vol. 34. Springer: Heidelberg—Berlin—New York 1981.
- [18] MAURER, H.: Optimale Steuerprozesse mit Zustandsbeschränkungen. Habilitationsschrift: Universität Würzburg 1976.
- [19] —: Differential Stability in Optimal Control Problems. Appl. Math. Optim. 5 (1979), 283–295.
- [20] —: First Order Sensitivity of the Optimal Value Function in Mathematical Programming and Optimal Control. In: Mathematical Programming with Data Perturbations. Lecture Notes in Pure and Applied Mathematics, Marcel Dekker (to appear).
- [21] PETERSON, D. W.: On Sensitivity in Optimal Control Problems. J. of Optimization Theory and Appl. 13 (1974), 56–73.
- [22] ROCKAFELLAR, R. T.: Convex Analysis. Princeton University Press: Princeton 1970.
- [23] —: Generalized Directional Derivatives and Subgradients of Nonconvex Functions. Can. J. Math. 32 (1980), 257–280.
- [24] —: Lagrange Multipliers and Subderivatives of Optimal Value Functions in Nonlinear Programming. In: Math. Programming Studies (to appear).
- [25] RUSSAK, I. B.: On Problems with Bounded State Variables. J. Optimization Theory and Appl. 5 (1970), 114–157.
- [26] —: On General Problems with Bounded State Variables. J. Optimization Theory and Appl. 6 (1970), 424–452.
- [27] RUSSAK, I. B.: Second Order Necessary Conditions for Problems with State Inequality Constraints. SIAM-J. Control 13 (1975), 372–388.
- [28] —: Second-Order Necessary Conditions for General Problems with State Inequality Constraints. J. Optimization Theory and Appl. 17 (1975), 43–92.
- [29] RUSSAK, I. B., and S. T. TAN: Necessary Conditions for Problems with Higher Derivative Bounded State Variables. J. Optimization Theory and Appl. 26 (1978), 601–636.
- [30] STRUGAR, P. V.: Maximum Neutron Flux in Thermal Research Reactors. J. Optimization Theory and Appl. 5 (1970), 301–312.
- [31] THOMPSON, R. G., GEORGE, M. D., BROWN, P. L., and M. S. PROCTOR: Optimal Production, Investment, and Output Price Controls for a Monopoly Firm of the Evans' Type. Econometrica 39 (1971), 119–129.

Manuskripteingang: 08. 05. 1981

VERFASSER:

Dr. BERNHARD GOLLAN
 Mathematisches Institut der Universität
 D-8700 Würzburg, Am Hubland, Federal Republic of Germany