SIAM J. CONTROL AND OPTIMIZATION
Vol. 31, No. 3, pp. 569-603, May 1993

002

LIPSCHITZIAN STABILITY IN NONLINEAR CONTROL AND OPTIMIZATION*

ASEN L. DONTCHEV† AND WILLIAM W. HAGER‡

Abstract. This paper studies Lipschitz properties, relative to the parameter p, of the set of solutions to problems of the form

$$\text{Find } z \in \Omega_p \text{ such that } T_p(z) \in F_p(z).$$

As applications, various problems in control and optimization are examined, focusing in particular on the stability of the feasible set of a control problem, and the stability of solutions of infinite-dimensional mathematical programs and optimal control problems. In another application, an estimate is obtained for the error in the Euler approximation to an optimal control problem.

Key words. stability, sensitivity, feasible set, optimal set, discrete approximations, Euler approximation, sufficient optimality conditions

AMS(MOS) subject classifications. 49K40, 49M25, 90C31, 93B05

1. Introduction. This paper presents a general framework for analyzing Lipschitz stability in control and optimization. As applications of the theory, we study the dependence on a parameter of the set of controls and states that satisfies given inequality constraints. We also study the dependence on a parameter of the optimal solutions of various problems in nonlinear control and optimization.

The paper begins by studying (in § 2) the following problem:

$$\text{Find } z \in \Omega_p \text{ such that } T_p(z) \in F_p(z), \tag{1}$$

where p is a parameter, T_p maps Ω_p to Y_p, Y_p is a normed linear space, and $F_p(z)$ is a subset of Y_p for each $z \in \Omega_p$. Loosely speaking, we proceed in the following way: Along with (1), we consider an *auxiliary problem*

$$\text{Find } z \in \Omega_p \text{ such that } L_p(z) + y \in F_p(z), \tag{2}$$

where $y \in Y$ is treated as a new parameter. It turns out that if L_p approximates T_p in a suitable sense, and if the set of solutions of (2) possesses certain Lipschitz properties with respect to y, uniformly in p, then the set of solutions of (1) will have analogous properties with respect to p. In particular, if T_p is smooth, then L_p can be its linearization. For a nonsmooth T_p, we should choose a nonsmooth L_p.

Our abstract approach is based on a refinement of the set-valued contracting mapping principle (Lemma 1). An existence result given in Theorem 1, leads to various stability results. In particular, Corollary 1 obtains an estimate for the distance from a reference point to the set of solutions of (1). In Corollary 2, we assume that Ω, F, and L are independent of p, obtaining an implicit function theorem: If the solution set of (2) is pseudo-Lipschitz with respect to y around some given point, and if L strongly approximates T_p, then the set of solutions of (1) is pseudo-Lipschitz as well. Corollary 3 obtains a result related to metric regularity of the map $T - F$.

* Received by the editors July 9, 1990; accepted for publication (in revised form) December 4, 1991. This research was supported by the United States Army Research Office contract DAAL03-89-G-0082, by National Science Foundation grant DMS 9022899, and by the Bulgarian Ministry of Science contract 127. The work was performed while the first author was a visitor at the University of Florida.

† Mathematical Reviews, 416 Fourth Street, Ann Arbor, Michigan 48107 (on leave from the Institute of Mathematics, Bulgarian Academy of Sciences, Sofia, Bulgaria).

‡ Department of Mathematics, University of Florida, Gainesville, Florida 32611.

Generalized equations of the form (1) have been considered by Robinson in a series of papers [35]-[38] with Ω and F independent of p. Our analysis contains some of his results. While the theory of [35]-[38] is applied to finite-dimensional mathematical programming problems, our focus here is infinite-dimensional optimization, primarily optimal control. Our analysis is more in the spirit of [19] and [20].

Recently, a different approach to sensitivity based on nonsmooth analysis and the differentiability properties of set-valued maps was developed by Aubin [4], Aubin and Frankowska [6], Rockafellar [39] and [41], King and Rockafellar [24], Mordukhovich [32], and others. In [16] this approach is applied to various control problems. A motivation for the nonsmooth approach to Lipschitz stability is given by Rockafellar in [40].

An outline of our paper follows, while detailed comments connecting specific results in our paper to related literature appear throughout the paper. Section 3 examines the feasible set for a nonlinear control system with inequality state and control constraints that depend on a parameter. We show that if the functions defining the constraints are sufficiently smooth, and if an interior point condition holds for a linearized system, then the feasible set is pseudo-Lipschitz. Moreover, the interior point condition holds if the gradients of the active constraints satisfy an independence condition, the same condition that appeared in Hager's analysis [18] of Lipschitz continuity in time for an optimal control. At the end of § 3, we present an example of a nonsmooth control system with state and control constraints, and we demonstrate a method for proving local controllability.

Section 4 considers a quadratic minimum problem in a reflexive Banach space with linear cone constraints. We show that a coercivity condition together with surjectivity of the gradients of the (active) constraints guarantee local Lipschitz continuity of the solution relative to the data. In § 5 we apply this result to a nonlinear optimization problem, and a quadratic program plays the role of an auxiliary problem.

In § 6 we consider a nonlinear control problem with convex control constraints. The treatment of the control problem requires special care due to the discrepancy between the function spaces needed for coercivity and for differentiability. An example shows that the method of analysis can still be applied, even when the coercivity condition is violated.

Finally, in § 7 we obtain error estimates for Euler's approximation to a nonlinear optimal control problem with convex control constraints. In this case, the parameter p in (1) corresponds to the mesh spacing. The key step in the analysis is to show that the solution of a perturbed discrete linear-quadratic problem, related to the auxiliary problem (2), depends Lipschitz continuously on a parameter, uniformly in the mesh spacing. Our method makes use of the so-called averaged modulus of smoothness, introduced by Sendov and Popov [42]. When the optimal contol has bounded variation, the error in the discrete control is on the order of the mesh spacing.

2. Abstract theory. Let Z be a Banach space, let Ω be a closed subset of Z, let Y be a normed vector space, and let 2^Y denote the collection of subsets of Y. Given a map $T:\Omega \to Y$ and a map $F:\Omega \to 2^Y$, we consider the following problem:

$$\text{Find } z \in \Omega \text{ such that } T(z) \in F(z). \tag{3}$$

Of course, for appropriate choices of Ω, T, and F, (3) may represent an equation, an inclusion, or a variational inequality. In this section, conditions are formulated that guarantee a solution to (3). This existence theorem applied to perturbations of (3) yields stability results. Throughout this paper, $\|\cdot\|$ denotes a norm in the appropriate space. Given subsets P and $Q \subset Z$, the one-sided distance from P to Q (or excess function), denoted $\|P-Q|$, is defined by

$$\|P-Q| = \sup_{p\in P} \inf_{q\in Q} \|p-q\|.$$

If Q is empty, we set $\|P-Q|=\infty$. Given $z \in Z$, let $B_r(z)$ denote the closed ball with center z and radius r.

We will use a contraction mapping principle for set-valued maps to obtain an existence result for (3). The proof that follows is similar to the usual proofs for the existence of a fixed point (see [22, p. 31] or [34]); however, since our Lipschitz assumption (b) below is weaker than the usual Lipschitz assumption, we include a proof for completeness. Although this fixed point result is stated for a Banach space, it holds in any complete metric space.

LEMMA 1. *Let $\Phi:\Omega\to 2^{\Omega}$ with $\Phi(z)$ closed for every $z\in\Omega$. Suppose that there exist real numbers r and λ, and $z_0\in\Omega$ with the following properties*:

$$\text{(a)}\quad 0\leqq\lambda<1 \quad \textit{and} \quad \frac{\|z_0-\Phi(z_0)|}{1-\lambda}<r,$$

$$\text{(b)}\quad \|\Phi(y)\cap B_r(z_0)-\Phi(z)|\leqq\lambda\|y-z\| \textit{ for every } y \textit{ and } z\in B_r(z_0)\cap\Omega.$$

Then there exists $z\in B_r(z_0)\cap\Omega$ such that $z\in\Phi(z)$. If Φ is single-valued, then assumption (a) *can be replaced by*

$$\text{(a}'\text{)}\quad 0\leqq\lambda<1 \quad \textit{and} \quad \frac{\|z_0-\Phi(z_0)\|}{1-\lambda}\leqq r,$$

and there exists a unique $z\in B_r(z_0)\cap\Omega$ with $z=\Phi(z)$.

Proof. By assumption (a), there exists $z_1\in\Phi(z_0)$ such that $\|z_1-z_0\|<r(1-\lambda)$. Proceedings by induction, suppose that there exists $z_{k+1}\in\Phi(z_k)\cap B_r(z_0)$ for $k=1,2,\dots,n-1$ with $\|z_{k+1}-z_k\|<r(1-\lambda)\lambda^k$. By assumption (b) and the induction hypothesis, we have

$$\|z_n-\Phi(z_n)|\leqq\|\Phi(z_{n-1})\cap B_r(z_0)-\Phi(z_n)|\leqq\lambda\|z_n-z_{n-1}\|<r(1-\lambda)\lambda^n.$$

Hence, there exists $z_{n+1}\in\Phi(z_n)$ such that $\|z_{n+1}-z_n\|<r(1-\lambda)\lambda^n$. By the triangle inequality,

$$\|z_{n+1}-z_0\|\leqq\sum_{k=0}^{n}\|z_{k+1}-z_k\|<r(1-\lambda)\sum_{k=0}^{n}\lambda^k<r,$$

so that $z_{n+1}\in B_r(z_0)$. This completes the induction step.

By the triangle inequality and for $n>m$, we have

$$\|z_n-z_m\|\leqq\sum_{k=m}^{n-1}\|z_{k+1}-z_k\|\leqq r(1-\lambda)\sum_{k=m}^{n-1}\lambda^k<r\lambda^m.$$

Thus the z_k form a Cauchy sequence that converges to some limit $z\in B_r(z_0)\cap\Omega$. By assumption (b),

$$\|z_k-\Phi(z)|\leqq\|\Phi(z_{k-1})\cap B_r(z_0)-\Phi(z)|\leqq\lambda\|z_{k-1}-z\|.$$

Again, by the triangle inequality,

$$\|z-\Phi(z)|\leqq\|z-z_k\|+\|z_k-\Phi(z)|\leqq\|z-z_k\|+\lambda\|z_{k-1}-z\|,$$

which approaches zero as k increases. Since $\Phi(z)$ is closed, it follows that $z \in \Phi(z)$. If Φ is single-valued, then by (b), z is the unique element in $B_r(z_0) \cap \Omega$ for which $z = \Phi(z)$. □

Given $y \in Y$ and a mapping L from Z to Y, consider the *auxiliary problem*

$$\text{Find } z \in \Omega \text{ such that } L(z) + y \in F(z). \tag{4}$$

Note that the set of solutions to (4) is closed if L is continuous and the graph of F is closed. Given $z_0 \in Z$ and $y_0 \in Y$, define the parameters D_r and δ by

$$D_r = \sup_{\substack{y,z \in B_r(z_0) \cap \Omega \\ y \neq z}} \frac{\|T(z) - T(y) - L(z) + L(y)\|}{\|z - y\|} \quad \text{and} \quad \delta = \|T(z_0) - L(z_0) - y_0\|.$$

If L is a bounded, linear operator, and z_0 lies in the interior of Ω, then $D_r \to 0$ when $r \to 0$ if and only if T is strictly (Fréchet) differentiable at z_0 (see [5, p. 16]). We have the following generalization of [20, Thm. 1].

THEOREM 1. *Let γ and r be real numbers that satisfy the relations*

$$0 \leq \gamma D_r < 1 \quad \text{and} \quad r > \frac{\gamma\delta}{1 - \gamma D_r}. \tag{5}$$

Defining the set

$$\Delta_r = \bigcup_{z \in B_r(z_0) \cap \Omega} \{T(z) - L(z)\},$$

let Ψ denote a map from Δ_r to 2^Ω with the following properties: $z_0 \in \Psi(y_0)$, $\Psi(y)$ is a closed, nonempty subset of the solutions to (4) for each $y \in \Delta_r$, and

$$\|\Psi(y_1) \cap B_r(z_0) - \Psi(y_2)| \leqq \gamma \|y_1 - y_2\| \text{ for every } y_1 \text{ and } y_2 \in \Delta_r \cup \{y_0\}. \tag{6}$$

Then (3) has a solution $z \in B_r(z_0)$. If in addition there is only one solution of (4) for every $y \in \Delta_r$, then z is the unique solution of (3) in $B_r(z_0)$, and the second condition in (5) can be weakened to $r \geqq \gamma\delta/(1 - \gamma D_r)$.

Proof. We apply Lemma 1 with $\Phi(z) = \Psi(T(z) - L(z))$. Thus z is a solution to (3) if $z \in \Phi(z)$. By (6), we have

$$\|\Phi(z) \cap B_r(z_0) - \Phi(y)| \leqq \gamma \|T(z) - T(y) - L(z) + L(y)\| \leqq \gamma D_r \|z - y\|$$

whenever y and $z \in B_r(z_0) \cap \Omega$. Hence, Φ satisfies (b) of Lemma 1 with constant $\lambda = \gamma D_r$. Since $z_0 \in \Psi(y_0)$, it follows that

$$\begin{aligned} \|z_0 - \Phi(z_0)| &\leqq \|\Psi(y_0) \cap B_r(z_0) - \Psi(T(z_0) - L(z_0))| \\ &\leqq \gamma \|y_0 + L(z_0) - T(z_0)\| = \gamma\delta. \end{aligned}$$

Dividing this inequality by $1 - \lambda$, we see that $r > \|z_0 - \Phi(z_0)|/(1 - \lambda)$. Since the contraction property of Lemma 1 holds on $B_r(z_0) \cap \Omega$, there exists $z \in B_r(z_0)$ with $z \in \Phi(z)$. If the solution of (4) is unique for every $y \in \Delta_r$, then Φ is single-valued on $B_r(z_0) \cap \Omega$. By Lemma 1, there exists a unique $z \in B_r(z_0) \cap \Omega$ with $z = \Phi(z)$. Hence, there is a unique solution to (3) in $B_r(z_0)$. □

Remark 1. Note that if $\sigma \geqq rD_r + \delta$, then $\Delta_r \subset B_\sigma(y_0)$. To prove this, we take the norm of the identity

$$T(z) - L(z) - y_0 = [T(z) - T(z_0) - L(z) + L(z_0)] + [T(z_0) - L(z_0) - y_0],$$

where $z \in B_r(z_0) \cap \Omega$, and we apply the triangle inequality to obtain the relation

$$\|T(z) - L(z) - y_0\| \leqq rD_r + \delta \leqq \sigma.$$

Now we consider a family of equations, each equation depending on a parameter p contained in a metric space P. Associated with each $p \in P$, there is a closed subset Ω_p of a Banach space Z_p, a normed vector space Y_p, and a pair of maps $T_p : \Omega_p \to Y_p$ and $F_p : \Omega_p \to 2^{Y_p}$. Analogous to (3), we study the following problem:

$$\text{(7)} \qquad \text{Find } z \in \Omega_p \text{ such that } T_p(z) \in F_p(z).$$

Let 0 be a fixed element of P. Using Theorem 1, we will study the continuity of the map $p \to \Sigma(p)$, where $\Sigma(p)$ is the set of solutions of (7), making use of the following auxiliary problem:

$$\text{(8)} \qquad \text{Find } z \in \Omega_p \text{ such that } L_p(z) + y \in F_p(z),$$

where $L_p : Z_p \to Y_p$, and $y \in Y_p$. We give three specific results assuming the maps appearing in (7) and (8) satisfy certain conditions near a reference point. The parameters D_r, Δ_r, and δ of Theorem 1 now depend on p as follows:

$$D_r(p) = \sup_{\substack{y,z \in B_r(z_p) \cap \Omega_p \\ y \neq z}} \frac{\| T_p(z) - T_p(y) - L_p(z) + L_p(y) \|}{\| z - y \|},$$

$$\Delta_r(p) = \Delta_r(p, z_p), \quad \text{where } \Delta_r(p, x) = \bigcup_{z \in B_r(x) \cap \Omega_p} \{T_p(z) - L_p(z)\}, \quad \text{and}$$

$$\delta(p) = \| T_p(z_p) - L_p(z_p) - y_p \|.$$

(Although the norms above may depend on p, this dependence is not indicated explicitly. In § 5 we consider a finite-dimensional discretization of an optimal control problem, in which case the norms depend on the mesh spacing.)

COROLLARY 1. *Let Ψ_p denote a map from a neighborhood of y_p to 2^{Ω_p} with the following properties: $z_p \in \Psi_p(y_p)$, $\Psi_p(y)$ is a closed, nonempty subset of the solutions to (8) for each $y \in \Delta_\sigma(p)$, where $\sigma > 0$, and for some γ and $\alpha > 0$, we have*

$$\text{(9)} \qquad \|\Psi_p(y_1) \cap B_\alpha(z_p) - \Psi_p(y_2)| \leqq \gamma \| y_1 - y_2 \| \text{ for every } y_1 \text{ and } y_2 \in \Delta_\sigma(p) \cup \{y_p\}.$$

If $D_r(p)$ and $\delta(p)$ tend to zero as r and p tend to zero, then for each $\gamma^+ > \gamma$ and for each p sufficiently close to zero, (7) has a solution z such that

$$\text{(10)} \qquad \| z_p - z \| \leqq \gamma^+ \| T_p(z_p) - L_p(z_p) - y_p \|.$$

If there is only one solution to (8) for every $y \in \Delta_\sigma(p)$, then z, satisfying (10), is the unique solution of (7) in a neighborhood of z_p.

Proof. Apply Theorem 1 with $\delta = \delta(p)$ and $r = \gamma^+ \delta(p)$. If $\delta = 0$, then z_p is a solution of (7), and (10) holds with $z = z_p$. If $\delta > 0$, then choose p sufficiently close to 0 that

$$\sigma \geqq r, \quad \gamma D_r(p) < 1, \quad \alpha \geqq r, \quad \text{and} \quad \gamma^+ > \frac{\gamma}{1 - \gamma D_r(p)}.$$

Hence, Theorem 1 yields (10). □

Now we wish to start with a given solution z_0 of (7) associated with $p = 0$ and show that for small perturbations in the parameter, we can solve the equation, and in some sense, the solution is "well behaved." In this analysis, we allow T to depend on p, while Ω and F are independent of p. That is, the following problem is considered:

$$\text{(11)} \qquad \text{Find } z \in \Omega \text{ such that } T_p(z) \in F(z).$$

To study the continuity of the solution map, we work with the fixed auxiliary problem (4) (L is independent of p).

We define an analogue $E_r(p)$ of D_r in which T is replaced by T_p, below:

$$E_r(p) = \sup_{\substack{y,z \in B_r(z_0)\cap\Omega \\ y \neq z}} \frac{\|T_p(z) - T_p(y) - L(z) + L(y)\|}{\|z-y\|}. \tag{12}$$

Following the terminology of Robinson [38], we say that $L(z)$ strongly approximates $T_p(z)$ at $z = z_0$ and $p = 0$ if and only if $E_r(p) \to 0$ as p and r tend to zero. Note that L does not need to be smooth. For example, if $T_p(z) = f_p(g(z))$, where f_p is Fréchet differentiable and g is Lipschitz, but not necessarily differentiable, then $L(z) = f_0'[g(z_0)]g(z)$ strongly approximates $T_p(z)$ at $z = z_0$ and $p = 0$ under appropriate continuity assumptions (see [38]).

In the following corollary, we take $z_p = z_0$ and $y_p = y_0$, and we replace assumption (9) by pseudo-Lipschitz continuity. Recall (see [4]) that the map Ψ is pseudo-Lipschitz with modulus γ, around a point (y_0, z_0) in the graph of Ψ, if there exist neighborhoods V of y_0 and U of z_0 such that

$$\|\Psi(y_1) \cap U - \Psi(y_2)| \leqq \gamma \|y_1 - y_2\|$$

whenever y_1 and $y_2 \in V$. Letting $\Sigma_r(p) = \Sigma(p) \cap B_r(z_0)$ denote the restriction of $\Sigma(p)$ to $B_r(z_0)$, we have the following corollary.

COROLLARY 2. *Given $z_0 \in \Sigma(0)$, define $y_0 = T_0(z_0) - L(z_0)$, and let $\Psi(y)$ denote the set of solutions to* (4). *We assume that Ψ is closed and nonempty-valued near y_0, that Ψ is pseudo-Lipschitz with modulus γ around (y_0, z_0), and that L strongly approximates $T_p(z)$ at $p = 0$ and $z = z_0$. If $T_p(z)$ is continuous in p at $p = 0$, uniformly in a neighborhood of $z = z_0$, then for each $\gamma^+ > \gamma$ and for r sufficiently small, there exists $s > 0$ such that $\Sigma_r(p)$ is nonempty for every $p \in B_s(0)$; moreover, for each p and $q \in B_s(0)$ and for each $z_p \in \Sigma_r(p)$, there exists $z_q \in \Sigma(q)$ such that*

$$\|z_p - z_q\| \leqq \gamma^+ \|T_q(z_p) - T_p(z_p)\|. \tag{13}$$

If there is only one solution to (4) *for every y near y_0, then the solution of* (11) *is unique in $B_r(z_0)$ for every $p \in B_s(0)$. Moreover, the $z_q \in \Sigma(q)$ satisfying* (13) *also lies in $B_r(z_0)$.*

Proof. Define the parameters

$$d(a,s) = \sup_{\substack{p,q \in B_s(0) \\ z \in B_a(z_0)}} \|T_p(z) - T_q(z)\| \quad \text{and} \quad \bar{D}_\rho = \sup_{p \in B_\rho(0)} E_\rho(p).$$

Let U and V be the neighborhoods of z_0 and y_0 appearing in the definition of pseudo-Lipschitz continuity. Choose σ sufficiently small that $B_\sigma(y_0) \subset V$. Choose ρ sufficiently small that $B_\rho(z_0) \subset U$

$$\sigma > \rho \bar{D}_\rho \quad \text{and} \quad \gamma < \gamma^+(1 - \gamma \bar{D}_\rho). \tag{14}$$

Choose a and s sufficiently small that

$$\rho \geqq 2a, \quad \rho \geqq s, \quad a > \gamma^+ d(a,s), \quad \text{and} \quad \sigma \geqq \rho \bar{D}_\rho + d(a,s). \tag{15}$$

Let $p \in B_s(0)$. Referring to Remark 1, we see that $\Delta_a(p, z_0) \subset B_\sigma(y_0)$. Theorem 1 with T replaced by T_p and with $r = a$ implies that $\Sigma_a(p)$ is nonempty.

Given p and $q \in B_s(0)$ and $z_p \in \Sigma_a(p)$, let us apply Theorem 1 with z_0, y_0, and T in the theorem replaced by z_p, $y_p = T_p(z_p) - L(z_p)$, and T_q, respectively. In Theorem 1, we take

$$r = \gamma^+ \delta, \quad \text{where } \delta = \|T_q(z_p) - T_p(z_p)\| \leqq d(a,s).$$

If $\delta=0$, then (13) holds trivially by taking $z_q=z_p$. If $\delta>0$, then by (15) we have

$$r>\frac{\gamma\delta}{1-\gamma D_\rho},$$

where $D_\rho=E_\rho(q)\leqq\bar{D}_\rho$. Since $B_r(z_p)\subset B_\rho(z_0)$, we have $\Delta_r(p,z_p)\subset\Delta_\rho(p,z_0)\subset B_\sigma(y_0)$. Hence, condition (6) of Theorem 1 holds, and (13) is established.

If the map Ψ is single-valued, then there exists a unique solution of (11) in $B_r(z_0)$ for every $p\in B_s(0)$. Similar to the proof of Theorem 1, the map $\Phi_p(z)=\Psi(T_p(z)-L(z))$ is a contraction on the ball $B_a(z_0)$ with contraction constant $\lambda=\gamma\bar{D}_\rho$ for each $p\in B_s(0)$. The distance from z_p to z_q is estimated by the following sequence of inequalities:

$$\begin{aligned}\|z_p-z_q\|=\|\Phi_p(z_p)-\Phi_q(z_q)\|&\leqq\|\Phi_p(z_p)-\Phi_q(z_p)\|+\|\Phi_q(z_p)-\Phi_q(z_q)\|\\&\leqq\lambda\|z_p-z_q\|+\gamma\|T_p(z_p)-T_q(z_p)\|,\end{aligned}$$

which yields (13). □

Observe that if there exists a constant κ such that T_p satisfies the Lipschitz condition

$$\|T_p(z)-T_q(z)\|\leqq\kappa\text{ distance }\{p,q\}$$

for every z in a neighborhood of z_0, then (13) implies that for every $z_p\in\Sigma_a(p)$, there exists $z_q\in\Sigma(q)$ such that

$$\|z_p-z_q\|\leqq\gamma^+\kappa\text{ distance }\{p,q\};$$

that is, the map Σ is pseudo-Lipschitz around $p=0$ and $z=z_0$. Thus we conclude that if the auxiliary problem strongly approximates the original problem, and if the solution map of the auxiliary problem is pseudo-Lipschitz, then the solution map of the original problem is pseudo-Lipschitz as well.

Remark 2. Corollary 2 is a generalization of Theorem 2.1 in [37] and of Theorem 3.2 in [38]. In [37] Ω is a closed convex set, $F(z)$ is the normal cone to Ω at z, $T_p(z)$ is Fréchet differentiable with respect to z around $z=z_0$ and $p=0$, and both $T_p(z)$ and its derivative $T'_p(z)$ are continuous with respect to z and p at $z=z_0$ and $p=0$. Furthermore, it is assumed that (4) with

$$L(z)=T_0(z_0)+T'_0(z_0)(z-z_0)$$

has a unique solution that is Lipschitz near $y_0=0$. In [38] $F(z)=0$, $L(z)$ strongly approximates $T_p(z)$ at $z=z_0$ and $p=0$, and the assumptions for $L(z)$ are equivalent to the condition that L^{-1} is single-valued and Lipschitz near 0.

Corollary 2 is an implicit function theorem in which we avoid the surjectivity (interiority) condition, for a suitably defined derivative, that is usually present in a Graves-type theorem (see [7, p. 95]). For example, given a closed-valued map $F:Z\to 2^Y$ and given (z_0,y_0) in the graph of F, let $f:Z\to Y$ be a continuous function that is strictly Fréchet differentiable at z_0. Let us apply Corollary 2 with $L(z)=-f'(z_0)(z-z_0)$, $p=y\in Y$, $T_p(z)=p-f(z)$, and Ψ defined by

$$\Psi(y)=\{z\in Z: y-f'(z_0)(z-z_0)\in F(z)\}.$$

By Corollary 2, Ψ is pseudo-Lipschitz around (y_0,z_0) if and only if the map $[F+f]^{-1}$ is pseudo-Lipschitz around $(y_0+f(z_0),z_0)$.

In the remainder of the paper, we also make use of the following result.

COROLLARY 3. *If the assumptions of Corollary* 2 *hold, then for each* $\gamma^+>\gamma$, *there exist positive constants* a, ρ, *and* ε *with the following properties*: *For every* $z\in B_a(z_0)$, $p\in B_\rho(0)$, *and* $w_p\in F(z)$ *with* $\|T_p(z)-w_p\|\leqq\varepsilon$, *there exists* $z_p\in\Sigma(p)$ *such that*

$$\|z-z_p\|\leqq\gamma^+\|T_p(z)-w_p\|. \tag{16}$$

Proof. Choose σ and ρ as in the proof of Corollary 2 to satisfy (14). Let a and $\varepsilon > 0$ be small enough that

$$\frac{\rho}{2} \geqq a \geqq \gamma^+ \varepsilon \quad \text{and} \quad \sigma \geqq \rho \bar{D}_\rho + \varepsilon. \tag{17}$$

Given $p \in B_\rho(0)$, $z \in B_a(z_0)$, and $w_p \in F(z)$ with $\|T_p(z) - w_p\| \leqq \varepsilon$, let us define $y_p = w_p - L(z)$. We apply Theorem 1 with y_0, z_0, and T replaced by y_p, z, and T_p, and with $\delta = \|T_p(z) - w_p\|$. If $\delta = 0$, then $z \in \Sigma(p)$, and (16) holds. If $\delta > 0$, we take $r = \gamma^+\delta$. Since $\delta \leqq \varepsilon$, it follows from (14) and (17) that

$$\frac{\gamma\delta}{1 - \gamma E_\rho(p)} < r = \gamma^+\delta \leqq \gamma^+\varepsilon \leqq a \leqq \frac{\rho}{2} < \rho.$$

Hence, $B_r(z) \subset B_\rho(z_0)$, which implies that $\Delta_r(p, z) \subset \Delta_\rho(p, z_0)$. By Remark 1 and (17), we have $\Delta_\rho(p, z_0) \subset B_\sigma(y_0)$ so that assumption (6) of Theorem 1 holds. By Theorem 1, there exists a solution $z_p \in B_r(z)$ to (11), where $r = \gamma^+\delta$, which establishes (16). □

Remark 3. Relation (16) implies that

$$\|z - \Sigma(p)| \leqq \gamma^+ \|T_p(z) - w_p\|.$$

Since the left-hand side of this inequality does not depend on w_p, it follows from Corollary 3 that when $\|T_p(z) - F(z)|$ is sufficiently small, we have

$$\|z - \Sigma(p)| \leqq \gamma^+ \|T_p(z) - F(z)|. \tag{18}$$

In particular, if $T_p(z) = T(z) + p$, we conclude that the map $T - F$ is *metrically regular* around $(z_0, 0)$. It turns out that metric regularity is equivalent to the pseudo-Lipschitz property (see Penot [33]). For a discussion of related results, see Cominetti [10], and the references therein.

Corollary 3 is a generalization of Theorem 1 in [36] where the estimate (18) is obtained under the following conditions: F is a closed, convex cone, independent of z; $T_p(z)$ is continuously Fréchet differentiable; and interior point regularity holds. This regularity condition implies, via the celebrated Robinson–Ursescu theorem (see [36] and [43]), that the solution map of the linearized (auxiliary) problem is pseudo-Lipschitz.

3. Feasibility and controllability. As a first appliction of the abstract theory, we study the continuity of the map "parameter → feasible set" of a nonlinear control system with constraints. The following model problem is analyzed: Given an interval $I = [0, 1]$, the state x is a map from I to R^n, while the control u is a map from I to R^m. Given θ between 1 and ∞, let $L^\theta(R^m)$ denote the space of functions $u: I \to R^m$ with $|u(t)|^\theta$ integrable where $|.|$ is the Euclidean norm. Let $W^{1,\theta}(R^n)$ denote the space of functions $x: I \to R^n$ with both x and its derivative in $L^\theta(R^n)$. We often omit the argument R^n or R^m when the context is clear. Of course, when θ is ∞, these spaces are modified in the standard fashion: L^∞ is the space of essentially bounded functions, and $W^{1,\infty}$ is the space of Lipschitz continuous functions (or, equivalently, the space of essentially bounded functions with essentially bounded derivatives). Given functions

$$f_p: R^{n+m} \times I \to R^n, \quad K_p: R^m \times I \to R^\mu, \quad \text{and} \quad S_p: R^n \times I \to R^\nu,$$

where p is a parameter, and, given a starting condition $a \in R^n$, the feasible set $\Sigma(p)$ consists of the set of $u \in L^\infty$ and $x \in W^{1,\theta}$ that satisfy the relations

$$\begin{aligned} &\dot{x}(t) = f_p(x(t), u(t), t) \quad \text{and} \quad K_p(u(t), t) \leqq 0 \quad \text{a.e. } t \in I, \\ &x(0) = a, \qquad S_p(x(t), t) \leqq 0 \qquad \text{for every } t \in I. \end{aligned} \tag{19}$$

Using the notation of (11), the feasible set in the control problem consists of those $z \in \Omega$ such that $T_p(z) \in F(z)$, where

$$\Omega = \{z = (x, u): x \in W^{1,\theta}, u \in L^\infty, x(0) = a\},$$

$$T_p(x, u) = \begin{bmatrix} f_p(x, u) - \dot{x} \\ K_p(u) \\ S_p(x) \end{bmatrix}, \quad \text{and} \quad F(x, u) = \begin{bmatrix} 0 \\ L_-^\infty \\ L_-^\infty \end{bmatrix}.$$

Here L_-^∞ denotes the nonpositive functions in L^∞.

Given a pair $z_0 = (x_0, u_0)$ that is feasible for (19) when $p = 0$, we wish to study the behavior of $\Sigma(p)$ for p near zero. Throughout this section, we make the following assumption: There exists a closed set $\Delta \subset R^n \times R^m \times I$ and a $\delta > 0$ such that $(x_0(t), u_0(t), t)$ lies in Δ for almost every $t \in I$, the distance from $(x_0(t), u_0(t), t)$ to the boundary of Δ in the hyperplane $R^n \times R^m \times \{t\}$ is at least δ for almost every $t \in I$, the derivatives of $f_p(x, u, t)$, $K_p(u, t)$, and $S_p(x, t)$ with respect to x and u exist on Δ, and these derivatives along with the function values are continuous with respect to $(x, u, t) \in \Delta$ and p near zero. From the development in § 2, Lipschitz properties of the solution map for the nonlinear problem are related to Lipschitz properties of the solution map for an auxiliary problem (4) when y is in a neighborhood of $T_0(z_0)$. We consider the following linearization of (19):

$$\begin{aligned} &\dot{x}(t) - \dot{x}_0(t) = A(t)(x(t) - x_0(t)) + B(t)(u(t) - u_0(t)) + y_1(t), \\ (20) \qquad &K(t)(u(t) - u_0(t)) + y_2(t) \leqq 0, \\ &S(t)(x(t) - x_0(t)) + y_3(t) \leqq 0, \end{aligned}$$

where $y_1 \in L^\theta$, y_2 and $y_3 \in L^\infty$, and

$$A(t) = \nabla_x f_0(x_0(t), u_0(t), t),$$

$$B(t) = \nabla_u f_0(x_0(t), u_0(t), t),$$

$$K(t) = \nabla_u K_0(u_0(t), t),$$

$$S(t) = \nabla_x S_0(x_0(t), t).$$

Above any equality or inequality involving measurable functions is interpreted in the sense "almost everywhere."

From the development of § 2, we see that pseudo-Lipschitz continuity of the feasible map can be deduced from the following three conditions:

(i) Lipschitz continuity of $T_p(z)$ with respect to p,

(ii) $\bar{D}_\rho$ is sufficiently small,

(iii) The solution map associated with the linearized system is pseudo-Lipschitz.

With regard to condition (i), Lipschitz continuity of $T_p(z)$ with respect to p is equivalent to Lipschitz continuity of $f_p(z)$, $K_p(z)$, and $S_p(z)$ with respect to p. Also, $\bar{D}_\rho$ tends to zero as ρ tends to zero under our smoothness assumptions. In the following two lemmas, we study the Lipschitz continuity of the solution map for the linearized system.

LEMMA 2. *Let $z_0 = (x_0, u_0) \in \Omega$ be feasible in* (19) *when $p = 0$, let $\Lambda(y)$ denote the set of solutions $(x, u) \in W^{1,\theta} \times L^\infty$ to* (20), *and define $y_0 = T_0(z_0)$. If there exist $\alpha > 0$,*

$w\in W^{1,\theta}$, and $v\in L^{\infty}$ such that

$$\begin{aligned}(21)\qquad &\dot{w}(t)=A(t)w(t)+B(t)v(t), \qquad w(0)=0,\\ &(K(t)v(t)+K_0(u_0(t),t))_i\leqq -\alpha, \qquad i=1,2,\ldots,\mu,\\ &(S(t)w(t)+S_0(x_0(t),t))_i\leqq -\alpha, \qquad i=1,2,\ldots,\nu,\end{aligned}$$

then Λ is pseudo-Lipschitz around (y_0, z_0).

Proof. This result follows from the Robinson-Ursescu theorem (see [36] and [43]) as stated (for example) by Aubin and Ekeland in [5, p. 132] or Clarke [9, p. 236]. That is, if there exists $r>0$ such that for each y in a neighborhood of y_0, we can find $z\in B_r(z_0)$ with $z\in\Lambda(y)$, then the map $y\to\Lambda(u)\cap B_r(z_0)$ is Lipschitz continuous, from which it follows that Λ is pseudo-Lipschitz around (y_0, z_0). The proof of the lemma proceeds as follows: Given $y\in Y$ and a pair (w, v) satisfying (21), let x denote the solution to the differential equation in (20) corresponding to the control $u=v+u_0$ and the starting condition $x(0)=a$. Observe that x can be expressed as

$$x=w+x_0+My_1,$$

where M is a bounded linear map from L^{θ} to $W^{1,\theta}$. Hence, we have

$$S(x-x_0)+y_3=S(w+My_1)+y_3\leqq -\alpha+SMy_1+y_3-S_0(x_0).$$

Similarly, putting $u=v+u_0$ into the control constraint of (20) gives

$$K(u-u_0)+y_2\leqq -\alpha+y_2-K_0(u_0).$$

Thus there exists $\sigma>0$ such that x and u satisfy the constraints in (20) whenever $y\in B_\sigma(y_0)$, where

$$y_0=T_0(z_0)=\begin{bmatrix}0\\ K_0(u_0)\\ S_0(x_0)\end{bmatrix}.$$

By the triangle inequality, we have

$$\|u-u_0\|+\|x-x_0\|\leqq\|v\|+\|w\|+\|My_1\|\leqq\|v\|+\|w\|+\sigma\|M\|$$

for every $y\in B_\sigma(y_0)$. Setting $r=\|v\|+\|w\|+\sigma\|M\|$, it follows that $B_r(z_0)\cap\Lambda(y)$ is nonempty whenever $y\in B_\sigma(y_0)$. This completes the proof. □

The proof of Lemma 2 provides a way to construct a single-valued Ψ: For each $y\in Y$, x is the solution of the differential equation (20) corresponding to $u=v+u_0$. Now we present a condition that yields the existence of a state and control satisfying the interior point condition of Lemma 2. This condition is the same one that appeared in the study [18] of Lipschitz continuous solutions in optimal control. First, we provide some terminology. We say that a function g is piecewise continuous if there exists a finite sequence $\{t_i\}$ with

$$0=t_0<t_1<t_2<\cdots<t_N=1,$$

such that g is continuous on the open interval (t_i, t_{i+1}) for each i, and one-sided limits exist at each t_i. A function is piecewise continuously differentiable if it is continuous and its derivative is piecewise continuous. If $K(t)$ is the coefficient matrix for u in (20), then $K^B(t)$ and $K^N(t)$ denote the submatrices of $K(t)$ consisting of rows associated with those indices i for which either

$$K_0(u_0(t),t)_i=0 \quad\text{or}\quad K_0(u_0(t),t)_i<0, \quad\text{respectively.}$$

In other words, $K^B(t)$ and $K^N(t)$ are the submatrices corresponding to the binding and the nonbinding constraints at time t. The submatrices S^B and S^N of S are defined in a similar fashion. Basically, we will show that if the columns of $K^B(t)^T$ and

$B(t)^T S^B(t)^T$ are uniformly linearly independent, then the interior point condition of Lemma 2 is satisfied.

LEMMA 3. *Let* $z_0 = (x_0, u_0) \in \Omega$ *be a point feasible for* (19) *when* $p = 0$, *and suppose that* $S_0(a, 0) < 0$, $S_0(x_0)$ *is piecewise continuously differentiable, both* $K_0(u_0)$ *and the matrices* K *and* B *are piecewise continuous, and at each* t *where these piecewise continuous functions are continuous, the following independence condition holds: There exists* $\beta > 0$ *such that*

$$|K^B(t)^T b + B(t)^T S^B(t)^T c| \geqq \beta(|b| + |c|)$$

for every b *and* c. *Moreover, at a time* t *of discontinuity, this independence condition also holds, but with* t *replaced by both* t^+ *and* t^-, *and with the binding constraint set replaced by those of* $K_0(u_0(t^+))$ *and* $K_0(u_0(t^-))$, *respectively. Then there exist* $w \in W^{1,\infty}$ *and* $v \in L^\infty$ *that satisfy hypothesis* (21) *of Lemma* 2 *for some* $\alpha > 0$.

Proof. Before considering the state constraints, we give a proof in the case of no state constraints. We first show that there exist a scalar $\delta > 0$ and sequences $\{t_i\}$ and $\{\tau_i\}$ such that

$$\begin{aligned} &t_i \leqq \tau_i \leqq t_{i+1} \quad \text{for } 0 \leqq i \leqq N, \quad \tau_0 = t_0 = 0, \quad \tau_N = t_{N+1} = 1, \\ &K_0^{N_i}(u_0(t), t) \leqq -\delta \quad \text{for each } t_i \leqq t \leqq t_{i+1}, \end{aligned} \tag{22}$$

$$|K^{B_i}(t)^T b| \geqq \delta |b| \quad \text{for each } t_i \leqq t \leqq t_{i+1}, \quad \text{for every } b, \tag{23}$$

where the B_i superscript means those rows (or components) associated with the binding constraints at τ_i, while the N_i superscript means those rows (or components) associated with nonbinding constraints at τ_i. The right-hand side of the inequality $K_0^{N_i}(u_0(t), t) \leqq -\delta$ is interpreted as a vector with every component equal to $-\delta$.

To prove (22) and (23), we verify that they are satisfied on the closure of each open interval J where K and $K_0(u_0)$ are continuous. Let us define the parameter $\varepsilon(t)$ by

$$\varepsilon(t) = \underset{1 \leqq i \leqq \mu}{\text{minimum}} \{-K_0(u_0(t), t)_i : K_0(u_0(t), t)_i \neq 0\}.$$

If all the constraints are binding at t, then we set $\varepsilon(t) = +\infty$. The value of $K_0(u_0)$ at an endpoint of J is taken to be its limit at that point. By the continuity assumptions, it follows that for each $t \in J$, there exists an open ball O_t, containing t, such that

$$K_0^{N_t}(u_0(s)) \leqq -\frac{\varepsilon(t)}{2} \quad \text{for every } s \in O_t$$

and

$$|K^{B_t}(s)^T b| \geqq \frac{\beta}{2} |b| \quad \text{for every } s \in O_t,$$

where the superscript B_t stands for rows binding at t, while the superscript N_t stands for components nonbinding at t. If t is an endpoint of J, then the open ball is replaced by a half-open ball. By compactness, this cover of J has a finite subcover. In (22) and (23), the τ_i are the centers of the balls in the subcover, while the t_i are arbitrary points in the overlap region between adjacent balls. The parameter δ is given by

$$\delta = \tfrac{1}{2} \text{minimum} \{\varepsilon(\tau_0), \varepsilon(\tau_1), \ldots, \varepsilon(\tau_N), \beta\}.$$

The control v that satisfies the interior point condition of Lemma 2 can be constructed in the following way: Given $\sigma > 0$ and t between t_i and t_{i+1}, $v(t)$ is the minimum norm solution of the equation $K^{B_i}(t)v(t) = -\sigma$ (if there are no binding constraints, set $v(t) = 0$). Since $K_0(u_0(t), t)$ is nonpositive, we have

$$K^{B_i}(t)v(t) + K_0^{B_i}(u_0(t), t) \leqq -\sigma. \tag{24}$$

By (23), the smallest singular value of $K^{B_i}(t)$ is bounded from below by δ. It follows that $v(t)$ has the following bound in the Euclidean norm:

$$|v(t)| \leqq \frac{\sigma\sqrt{\mu}}{\delta},$$

where μ is the number of component of K_0. Hence, by (22) v also satisfies the inequality

$$K^{N_i}(t)v(t) + K_0^{N_i}(u_0(t), t) \leqq \frac{\sigma\sqrt{\mu}}{\delta}|K^{N_i}(t)| - \delta. \tag{25}$$

Relations (24) and (25) imply that v satisfies the interior point condition (21) for σ sufficiently small.

When state constraints are present, this proof must be modified in several ways. By incorporating the state constraints in the definition of $\varepsilon(t)$, we can choose δ to satisfy the additional relation

$$S_0^{N_i}(x_0(t), t) \leqq -\delta \quad \text{for each } t_i \leqq t \leqq t_{i+1}.$$

In addition, the independence condition (23) generalizes to the form

$$|K^{B_i}(t)^T b + B(t)^T S^{B_i}(t)^T c| \geqq \delta(|b| + |c|). \tag{26}$$

Similar to the control constrained case, we wish to construct a control v_σ and a state w_σ such that $K^{B_i}v_\sigma(t) = -\sigma$ and $S^{B_i}w_\sigma(t) = -\sigma$ for t between t_i and t_{i+1}, and both v_σ and w_σ are bounded pointwise by a constant times σ. This construction is complicated by the fact that v_σ and w_σ must satisfy the linear differential equation in (21).

The proof proceeds by induction, interval by interval, from left to right. Suppose that on the interval $[t_0, t_k]$ we can construct a control v_σ and a corresponding state w_σ such that for each σ sufficiently small, we have

$$|w_\sigma(t_k)| \leqq c\sigma, \qquad S^{B_{k-1}}w_\sigma(t_k) = -\sigma,$$

where c is independent of σ, and with $\alpha = \sigma$, the control $v = v_\sigma$ and the state $w = w_\sigma$ satisfy the relations (21) on the interval $[0, t_k]$. We now show that this construction can be continued on the interval $[t_k, t_{k+1}]$. The control v and the state w on the new interval are chosen to satisfy the relations

$$\begin{aligned} &K^{B_k}(t)v(t) = -\sigma, \quad \text{for } t_k < t \leqq t_{k+1}, \\ &(S^{B_k}(t)w(t))_i = (S^{B_k}(t_k)w_\sigma(t_k))_i + \gamma_i(t - t_k) \quad \text{for } t_k < t \leqq t_k + \varepsilon, \\ &S^{B_k}(t)w(t) = -\sigma \quad \text{for } t_k + \varepsilon < t \leqq t_{k-1}. \end{aligned} \tag{27}$$

The parameters γ_i and ε are selected so that $(S^{B_k}(t_k+\varepsilon)w(t_k+\varepsilon))_i = -\sigma$, or equivalently, so that

$$\gamma_i = -\frac{(S^{B_k}(t_k)w_\sigma(t_k))_i + \sigma}{\varepsilon}.$$

Since $w_\sigma(t_k)$ tends to zero as σ tends to zero, it follows that for any ε, γ_i tends to zero as σ tends to zero.

To obtain a control that satisfies (27), we differentiate the second and third equations in (27) and we substitute from the state equation $\dot{w} = Aw + Bv$ to obtain

$$\begin{aligned} &S^{B_k}(t)B(t)v(t) = \gamma_i - \frac{dS^{B_k}(t)}{dt}w(t) - S^{B_k}(t)A(t)w(t) \quad \text{for } t_k < t \leqq t_k + \varepsilon, \\ &S^{B_k}(t)B(t)v(t) = -\frac{dS^{B_k}(t)}{dt}w(t) - S^{B_k}(t)A(t)w(t) \quad \text{for } t_k + \varepsilon < t \leqq t_{k+1}. \end{aligned} \tag{28}$$

By the independence condition (26), the minimum norm control $v(t)$ that satisfies the equation $K^{B_k}(t)v(t) = -\sigma$ along with (28), where w is the solution to

$$\dot{w}(t) = A(t)w(t) + B(t)v(t), \; w(t_k) = w_\sigma(t_k),$$

is bounded pointwise by a constant times σ. If the ith state constraint is binding at both τ_{k-1} and τ_k, then by the construction of w, we have

$$(S(t)w(t))_i = -\sigma \quad \text{for } t_k \leqq t \leqq t_{k+1},$$

which implies that

$$(S(t)w(t) + S_0(x_0(t), t))_i \leqq -\sigma \quad \text{for } t_k \leqq t \leqq t_{k+1}.$$

If the ith state constraint is nonbinding at either τ_{k-1} or τ_k, then $S_0(x_0(t_k), t_k)_i \leqq -\delta$. Hence, for ε sufficiently small, $S_0(x_0(t), t)_i \leqq -\delta/2$ for t between t_k and $t_k + \varepsilon$. Taking σ sufficiently small yields

$$(S(t)w(t) + S_0(x_0(t), t))_i \leqq -\delta/4 \quad \text{for } t_k \leqq t \leqq t_k + \varepsilon.$$

Now consider t in the interval $[t_k + \varepsilon, t_{k+1}]$. If the ith constraint is binding at τ_k, then

$$(S(t)w(t) + S_0(x_0(t), t))_i \leqq -\sigma.$$

If the ith constraint is nonbinding at τ_k, then

$$(S(t)w(t) + S_0(x_0(t), t))_i \leqq |S(t)w(t)| - \delta.$$

Since w is bounded by a constant times σ, the induction step is complete. □

To conclude, we state a specific sensitivity result for the feasibility problem (19) based on Corollary 3.

THEOREM 2. *If (x_0, u_0) is feasible in (19) when $p = 0$, and there exist $\alpha > 0$, $w \in W^{1,\theta}$, and $v \in L^\infty$ satisfying (21), then for each p in a neighborhood of 0 and for each (x, u) in a neighborhood of $z_0 = (x_0, u_0)$, there exists x_p and u_p that are feasible in (19), and we have*

$$(29) \quad \|x_p - x\|_{W^{1,\theta}} + \|u_p - u\|_{L^\infty} \leqq c(\|f_p(x, u) - \dot{x}\|_{L^\theta} + \|K_p(u)_+\|_{L^\infty} + \|S_p(x)_+\|_{L^\infty}),$$

where the "+" subscript stands for the positive part and c is independent of p.

Proof. Relation (29) follows from Corollary 3 and Lemma 2, where we identify the z of Corollary 3 with the pair (x, u), while w_p is identified with the triple $(0, K_p(u)_-, S_p(x)_-)$. Here the subscript "$-$" stands for the negative part. □

Remark 4. Using generalized derivatives of set-valued maps, a result related to Theorem 2 is established in [16, Thm. 10.1] for a problem with final state constraints. A linear system with convex state and control constraints is studied in [14].

We can use the same approach to study local controllability. The following simple example illustrates the basic ideas. Let us consider the nonsmooth control system

$$(30) \qquad \dot{x}(t) = |x(t)| + x(t)^5 u(t) + u(t) \quad \text{a.e. } t \in I,$$

with the constraints

$$x(0) = 0, \quad x(t) \geqq t - 1 \quad \text{for every } t \in I, \quad -1 \leqq u(t) \leqq 1 \quad \text{a.e. } t \in I, \quad x \in W^{1,\infty}, \quad u \in L^\infty.$$

A control system is *locally controllable* around 0 at $t = 1$ if for each a near zero that satisfies the state constraints at $t = 1$, there exists a feasible trajectory with $x(1) = a$. We will apply Corollary 1 with the following identifications:

$$\Omega = \{(x, u) \in W^{1,\infty} \times L^\infty : x(0) = 0, x(t) \geqq t - 1 \text{ for every } t \in I, -1 \leqq u(t) \leqq 1 \text{ a.e. } t \in I\},$$

$$T_p(x, u) = \begin{bmatrix} \dot{x} - |x| - x^5 u - u \\ x(1) - p \end{bmatrix}, \quad \text{and} \quad F_p(x, u) = 0.$$

Local controllability is equivalent to existence of a solution to (7) for every $p \geqq 0$, p sufficiently small.

In applying Corollary 1, we take $P = R_+$, the nonnegative real numbers, $z_p = z_0 = (x_0, u_0) = 0$, $y_0 = 0$, and

$$L_p(z) = \begin{bmatrix} \dot{x} - |x| - u \\ x(1) \end{bmatrix}.$$

With this definition, the auxiliary problem becomes the following:

$$\text{Find } (x, u) \in \Omega \text{ such that} \tag{31}$$
$$\dot{x}(t) - |x(t)| - u(t) + y_1(t) = 0, \quad x(1) + y_2 = 0.$$

Hypothesis (9) of Corollary 1 is satisfied if there exists a single-valued map Ψ from $L^\infty \times R$ to $W^{1,\infty} \times L^\infty$, with $\Psi(0) = 0$, with $(x, u) = \Psi(y_1, y_2)$ a solution of (31), and with Ψ Lipschitz continuous on the set $\Delta_\sigma(p)$ of Corollary 1. It can be verified that the following choice for Ψ has the desired properties:

$$\Psi(y) = \begin{bmatrix} \dfrac{1 - e^t}{e - 1} y_2 \\ \dfrac{y_2}{1 - e} - y_1(t) \end{bmatrix}.$$

(Note that if $(y_1, y_2) \in \Delta_\sigma(p)$, then $y_2 \leqq 0$ since $p \geqq 0$.) Hence, the control system (30) is locally controllable around 0 at $t = 1$.

4. Quadratic programs. In applying the results of § 2 to problems in optimal control and mathematical programming, we must derive Lipschitz results for the auxiliary problem. This section collects properties of quadratic programs that are relevant to the analysis.

LEMMA 4. *Let X denote a reflexive Banach space, let $\Lambda \subset X$ be a nonempty closed, convex subset, and consider the problem*

$$\text{minimize } \tfrac{1}{2}\langle Ax, x\rangle + \langle \phi, x\rangle \quad \text{over } x \in \Lambda, \tag{32}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between X and the dual space X^*, $\phi \in X^*$, $A: X \to X^*$ is a continuous linear operator, and $\langle Ax, y\rangle = \langle Ay, x\rangle$ for every x and $y \in X$. If there exists a constant $\alpha > 0$ such that

$$\langle A(x_1 - x_2), x_1 - x_2\rangle \geqq \alpha \|x_1 - x_2\|^2 \quad \text{for every } x_1 \text{ and } x_2 \in \Lambda, \tag{33}$$

then there is a unique solution $\bar{x}$ to (32), *and $\bar{x}$ is the unique solution to the following variational inequality*:

$$\text{Find } \bar{x} \in \Lambda \text{ such that } \langle A\bar{x} + \phi, x - \bar{x}\rangle \geqq 0 \quad \text{for every } x \in \Lambda. \tag{34}$$

If x_1 and x_2 denote the solutions of (32) *corresponding to $\phi = \phi_1$ and $\phi = \phi_2$, then we have*

$$\|x_1 - x_2\| \leqq \|\phi_1 - \phi_2\| / \alpha. \tag{35}$$

Proof. In a Hilbert space, the existence of a solution $\bar{x}$ to (32) and the correspondence between a solution to (32) and a solution to (34) is found for example, in [26, Chap. 1]. The Lipschitz result (35) is found in [20, Lemma 1] for a Hilbert space. These Hilbert space proofs are also valid in a reflexive Banach space. □

The usual second-order sufficient condition for (32) has the form

$$\langle A(x - \bar{x}), x - \bar{x}\rangle \geqq \alpha \|x - \bar{x}\|^2 \quad \text{for every } x \in \Lambda, \tag{36}$$

where $\alpha > 0$. Hence, condition (33) is stronger than the second-order sufficient condition. An important difference between (33) and (36) is that after small perturbations in A, (33) still holds for some $\alpha > 0$; after small perturbations in A and $\bar{x}$, (36) may not hold for any $\alpha > 0$.

Under the hypotheses of Lemma 4, let us consider the constraint set

$$\Lambda = \{x \in X \colon Bx + \psi \in K\}, \tag{37}$$

where $B: X \to W$ is a continuous, linear operator; W is a Banach space; $\psi \in W$; and $K \subset W$ is a closed, convex cone with vertex at the origin. In this case, (32) takes the form

$$\text{minimize } \tfrac{1}{2}\langle Ax, x\rangle + \langle \phi, x\rangle \quad \text{subject to } Bx + \psi \in K. \tag{38}$$

Given x_1 and $x_2 \in \Lambda$, observe that $v = x_1 - x_2$ has the property that $Bv \in K - K$. Hence, when Λ is given by (37), (33) holds if

$$\langle Av, v\rangle \geqq \alpha \|v\|^2 \quad \text{whenever } Bv \in K - K. \tag{39}$$

Conversely, if B is surjective and (33) holds, then (39) holds. Hence, (39) and (33) are equivalent when B is surjective.

Letting K^+ denote the polar cone defined by

$$K^+ = \{\lambda \in W^* \colon \langle \lambda, k\rangle \geqq 0 \text{ for every } k \in K\},$$

suppose that there exists $\lambda \in K^+$ and $\bar{x} \in X$ satisfying

$$A\bar{x} - B^*\lambda + \phi = 0 \quad \text{and} \quad \langle \lambda, B\bar{x} + \psi\rangle = 0, \quad \text{where } B\bar{x} + \psi \in K. \tag{40}$$

It follows that

$$0 = \langle A\bar{x} + \phi, x - \bar{x}\rangle - \langle \lambda, Bx + \psi\rangle \leqq \langle A\bar{x} + \phi, x - \bar{x}\rangle$$

for every $x \in X$ with $Bx + \psi \in K$. By Lemma 4, $\bar{x}$ is the unique solution to (32). Note that the conditions $\langle \lambda, B\bar{x} + \psi\rangle = 0$ and $B\bar{x} + \psi \in K$ of (40) are often written in the compact form

$$B\bar{x} + \psi \in \partial K^+(\lambda),$$

where $\partial K^+(\lambda) = \{w \in W^{**} \colon \langle w, \mu - \lambda\rangle \geqq 0 \text{ for each } \mu \in K^+\}$ is the normal cone at λ to the set K^+.

If $\bar{x}$ is a solution to (38) and B is surjective, it is known (see Kurcyusz [25]) that there exists a unique Lagrange multiplier $\lambda \in K^+$ satisfying (40). Assuming that B is surjective and (33) holds, let us study the dependence of the solution and the multiplier associated with (38) on the parameters ϕ and ψ. Given $\phi = \phi_i \in X^*$ and $\psi = \psi_i \in W$ for $i = 1$ and 2, let $\bar{x}_i$ be the corresponding solutions to (38), and let λ_i be the associated multipliers satisfying (40). If $x = \tilde{x}_1$ is any solution to $Bx = -\psi_1$, then by the open mapping principle (see [7, p. 57]), B^{-1} is Lipschitz continuous, and there exists a solution $x = \tilde{x}_2$ to $Bx = -\psi_2$ such that

$$\|\tilde{x}_1 - \tilde{x}_2\| \leqq c\|\psi_1 - \psi_2\|,$$

where c is independent of ψ_1 and ψ_2. Making the change of variables $\bar{x}_i = w_i + \tilde{x}_i$ in (40), we obtain

$$Aw_i - B^*\lambda_i + A\tilde{x}_i + \phi_i = 0, \quad Bw_i \in K, \quad \text{and} \quad \langle \lambda_i, Bw_i\rangle = 0. \tag{41}$$

Hence, $w = w_i$ is the solution to the problem

$$\begin{aligned} &\text{minimize} && \tfrac{1}{2}\langle Aw, w\rangle + \langle A\tilde{x}_i + \phi_i, w\rangle \\ &\text{subject to} && Bw \in K. \end{aligned}$$

By Lemma 4, inequality (35), we have

$$\alpha \|w_1 - w_2\| \leqq \|\phi_1 - \phi_2\| + \|A\| \, \|\tilde{x}_1 - \tilde{x}_2\| \leqq \|\phi_1 - \phi_2\| + c\|A\| \, \|\psi_1 - \psi_2\|.$$

Taking the norm of the identity $x_1 - x_2 = w_1 - w_2 + \tilde{x}_1 - \tilde{x}_2$, and applying the triangle inequality, we obtain

$$\|x_1 - x_2\| \leqq c\|\phi_1 - \phi_2\| + c\|\psi_1 - \psi_2\|,$$

where c is independent of the ϕ_i and the ψ_i. Moreover, since B^* is one-to-one and $(B^*)^{-1}$ is a continuous linear operator, (40) implies that $\|\lambda_1 - \lambda_2\|$ has a similar bound. These observations are summarized in the following lemma.

LEMMA 5. *Suppose that B is surjective and* (33) *holds. If x_i and λ_i are the solutions to* (40) *corresponding to $\phi = \phi_i$ and $\psi = \psi_i$, $i = 1$ and 2, then there exists a constant c, depending only on A and B, such that*

$$\|x_1 - x_2\| + \|\lambda_1 - \lambda_2\| \leqq c\|\phi_1 - \phi_2\| + c\|\psi_1 - \psi_2\|.$$

Clearly, the coercivity condition (33) is preserved after small perturbations in A. In the context of (38) with B surjective, we now show that the coercivity condition (33) is preserved after small perturbations in B, and after arbitrary perturbations in ψ. Since B is surjective, we observed earlier that (33) is equivalent to (39). Since ψ does not appear in (39), coercivity is preserved after any perturbation in ψ. Now let us consider the effect of changes in B. Given a bounded linear operator $\bar{B}: X \to W$, the open mapping principle implies that there exists a constant c, depending only on B, with the following property: If $\bar{B}\bar{v} \in K - K$, we can find $v \in X$ with $Bv = \bar{B}\bar{v} \in K - K$ and

$$\|v - \bar{v}\| \leqq c\|B - \bar{B}\| \, \|\bar{v}\|. \tag{42}$$

By the triangle inequality, we have

$$(1 - c\|B - \bar{B}\|)\|\bar{v}\| \leqq \|v\| \leqq (1 + c\|B - \bar{B}\|)\|\bar{v}\|. \tag{43}$$

Defining $\delta v = v - \bar{v}$ and applying (33) yields

$$\langle A\bar{v}, \bar{v}\rangle \geqq \alpha\|v\|^2 - 2\langle Av, \delta v\rangle + \langle A\delta v, \delta v\rangle. \tag{44}$$

Utilizing the inequality $2ab \leqq \rho a^2 + b^2/\rho$, where ρ is an arbitrary scalar, we have

$$|\langle Av, \delta v\rangle| \leqq \|A\| \, \|v\| \, \|\delta v\| \leqq \frac{\alpha}{4}\|v\|^2 + \frac{\|A\|^2}{\alpha}\|\delta v\|^2.$$

This inequality, coupled with relations (42)-(44), yields the following result.

LEMMA 6. *If the coercivity condition* (39) *holds and B is surjective, then for $\bar{A}$ in a neighborhood of A and for $\bar{B}$ in a neighborhood of B, there exists $\alpha > 0$ such that*

$$\langle \bar{A}v, v\rangle \geqq \alpha\|v\|^2 \text{ whenever } \bar{B}v \in K - K.$$

We observe that in certain cases, the coercivity condition (33) holds for the set Λ if it holds on a subset $\Lambda \cap \Gamma$.

LEMMA 7. *Let Γ and Λ be convex subsets of X, and suppose that there exists $\alpha > 0$ such that*

$$\langle A(x_1 - x_2), x_1 - x_2\rangle \geqq \alpha\|x_1 - x_2\|^2 \quad \text{for every } x_1 \text{ and } x_2 \in \Lambda \cap \Gamma. \tag{45}$$

If int Γ, *the interior of* Γ, *intersects* Λ, *then* (33) *holds.*

Proof. Given $v_i \in \Lambda$, $i=1$ and 2, and $v \in \operatorname{int} \Gamma \cap \Lambda$, define $v_i(\beta)$ by

$$v_i(\beta) = \beta v_i + (1-\beta)v.$$

Since Λ is convex, $v_i(\beta) \in \Lambda$ for $i=1$ and 2, whenever $\beta \in [0,1]$. Since $v_i(\beta) \to v$ as $\beta \to 0$, we can choose $\beta > 0$ small enough that $v_i(\beta) \in \Gamma$ for $i=1$ and 2. Applying (45) with $x_i = v_i(\beta)$, we obtain (33). □

Remark 5. Suppose that $B=(B_1, B_2)$, $\psi = (\psi_1, \psi_2)$, and $K = K_1 \times K_2$, where B_1 is surjective, and where there exists $v \in X$ with $B_1 v + \psi_1 \in K_1$ and $B_2 v + \psi_2 \in \operatorname{int} K_2$. Combining Lemmas 6 and 7, we see that if the coercivity condition (33) holds, then for $\bar{A}$ and $\bar{B}_1$ near A and B_1 respectively, there exists $\alpha > 0$ such that

$$\langle \bar{A}v, v\rangle \geqq \alpha \|v\|^2 \quad \text{whenever } \bar{B}_1 v \in K_1 - K_1.$$

5. Optimal solutions. In this section, we use Corollaries 1 and 2 to study an optimal solution of the problem

$$\text{minimize } C_p(x) \quad \text{over } x \in \Omega_p, \tag{46}$$

where p is a parameter in a metric space P, Ω_p is a closed, convex nonempty subset of a reflexive Banach space X, and $C_p : X \to R$. Given a local minimizer x_0 of (46) corresponding to $p=0$, we assume that for each $p \in P$, the functional $C_p(x)$ is twice Fréchet differentiable with respect to x in a neighborhood of x_0, the derivatives $C'_p(x)$ and $C''_p(x)$ with respect to x are continuous in p and x in a neighborhood of $p=0$ and $x = x_0$, and there exists $\alpha > 0$ such that

$$\langle C''_0(x_0)(x_1-x_2), x_1-x_2\rangle \geqq \alpha \|x_1-x_2\|^2 \quad \text{for every } x_1, x_2 \in \bigcup_{p\in P} \Omega_p. \tag{47}$$

In addition, we assume that $\lim_{p\to 0} |\Omega_p - \Omega_0| = 0$, where $|A-B| = \|A-B| + \|B-A|$ is equivalent to the Hausdorff distance between the sets A and B.

THEOREM 3. *For each $\beta > 1/\alpha$ and $\gamma > \|C'_0(x_0)\|$, there exists $s > 0$ with the following property: For each $p \in B_s(0)$, we can find a strict local minimizer x_p of* (46) *such that*

$$\|x_p - x_0\| \leqq \beta \|C'_p(x_0) - C'_0(x_0)\| + \sqrt{\gamma/\alpha}\, |\Omega_p - \Omega_0|^{1/2}. \tag{48}$$

If $C'_0(x_0) = 0$, then we require $\gamma > \|C''_0(x_0)\|$, and we replace the exponent $\frac{1}{2}$ in (48) *by* 1.

Proof. Since x_0 is a local minimizer of (46) and Ω_0 is convex, we have

$$\langle C'_0(x_0), x - x_0\rangle \geqq 0 \quad \text{for every } x \in \Omega_0. \tag{49}$$

Given $p \in P$, (47) and Lemma 4 imply that there exists a unique $\xi_p \in \Omega_p$ satisfying the relation

$$\langle C'_0(x_0) + L(\xi_p - x_0), x - \xi_p\rangle \geqq 0 \quad \text{for every } x \in \Omega_p, \tag{50}$$

where $L = C''_0(x_0)$. Adding (49) with $x = z_0$ and (50) with $x = z_p$ to inequality (47) with $x_1 = \xi_p$ and $x_2 = x_0$, we get

$$\alpha \|\xi_p - x_0\|^2 \leqq \langle L(\xi_p - x_0), z_p - x_0\rangle + \langle C'_0(x_0), z_0 - \xi_p + z_p - x_0\rangle$$

for every $z_0 \in \Omega_0$ and $z_p \in \Omega_p$. From this, it follows that

$$\alpha \|\xi_p - x_0\|^2 \leqq \|L\| \, \|\xi_p - x_0\| \, \|x_0 - \Omega_p| + \|C'_0(x_0)\| (\|x_0 - \Omega_p| + \|\xi_p - \Omega_0|).$$

Consequently, $\xi_p \to x_0$ as $p \to 0$, and we have

$$\alpha \|\xi_p - x_0\|^2 \leqq (\|L\| \, \|\xi_p - x_0\| + \|C'_0(x_0)\|) |\Omega_p - \Omega_0|. \tag{51}$$

Now let us consider the following problem:

$$\text{Find } x_p \in \Omega_p \text{ such that } \langle C'_p(x_p), x - x_p\rangle \geqq 0 \quad \text{for every } x \in \Omega_p. \tag{52}$$

We apply Corollary 1 to this problem, making the following identifications:

$$z=x, \quad T_p=C'_p, \quad y_p=C'_0(x_0)-L(x_0), \quad \text{and} \quad F_p(x)=\partial\Omega_p(x).$$

The auxiliary problem is the following:

(53) $\qquad$ Find $x\in\Omega_p$ such that $\langle L(x)+y, w-x\rangle \geqq 0$ for every $w\in\Omega_p$.

By Lemma 4, there exists a unique solution of (53) for each $y\in X^*$, and this solution is a Lipschitz continuous function of y with Lipschitz constant $1/\alpha$, independent of p. By Corollary 1 and for any constant $\beta>1/\alpha$, there exists a solution x_p to (52) for p near zero, and we have

$$\|x_p-\xi_p\| \leqq \beta\|C'_p(\xi_p)-L(\xi_p)-y_p\|. \tag{54}$$

By the definition of y_p, it follows that

$$\|C'_p(\xi_p)-L(\xi_p)-y_p\| = \|C'_p(\xi_p)-C'_0(x_0)-L(\xi_p-x_0)\|.$$

By the continuous differentiability assumptions, it follows that for any $\varepsilon>0$, there exists an $s>0$ such that

$$\|C'_0(\xi_p)-C'_0(x_0)-L(\xi_p-x_0)\| \leqq \varepsilon\|\xi_p-x_0\| \quad \text{for every } p\in B_s(0)$$

and

$$\|C'_p(\xi_p)-C'_0(\xi_p)\| \leqq \|C'_p(x_0)-C'_0(x_0)\|+\varepsilon\|\xi_p-x_0\| \quad \text{for every } p\in B_s(0).$$

These inequalities, combined with (51), (54), and the triangle inequality, yield (48). By (47), x_p is a strict local minimizer of (46) for p near 0. □

Remark 6. In general, the exponent $\frac{1}{2}$ in (48) is sharp (see [13, p. 13]). Theorem 3 is a generalization of Proposition 1.2 in [13].

Typically, Theorem 3 yields a Hölder-type estimate for x_p-x_0. However, when the constraint set is described by equalities and inequalities that possess certain regularity properties, a Lipschitz estimate can be established. We consider the following problem:

$$\text{minimize } C_p(x) \quad \text{subject to } G_p(x)\in K, \tag{55}$$

where $G_p: X\to W$, W is a Banach space, and $K\subset W$ is a closed, convex cone with vertex at the origin. Letting x_0 denote a local minimizer of (55) corresponding to $p=0$, we assume henceforth in this section that C_p and G_p possess the following smoothness properties: $C_p(x)$ and $G_p(x)$ are twice Fréchet differentiable in x in a neighborhood of $p=0$ and $x=x_0$, and these derivatives are continuous in p and x at $p=0$ and $x=x_0$. The functions $G_p(x)$, $C'_p(x)$, and $G'_p(x)$ are Lipschitz in $p\in P$, uniformly in x near x_0. Let H_p denote the Lagrangian defined by

$$H_p(x,\lambda)=C_p(x)-\langle\lambda, G_p(x)\rangle,$$

where $\lambda\in W^*$. The first-order necessary conditions associated with a solution to (55) can be expressed in the following way:

$$\nabla_x H_p(x_p,\lambda_p)=0 \quad \text{and} \quad G_p(x_p)\in\partial K^+(\lambda_p), \quad \text{where } x_p\in X \quad \text{and} \quad \lambda_p\in K^+. \tag{56}$$

It is well known (see [25]) that if $G'_0(x_0)$ is surjective, then there exists λ_0 satisfying (56) for $p=0$. Our Lipschitz result is based on the following coercivity condition: There exists $\alpha>0$ such that

$$\begin{aligned}&\langle\nabla^2_{xx}H_0(x_0,\lambda_0)(x_2-x_1), x_2-x_1\rangle \geqq \alpha\|x_2-x_1\|^2\\ &\text{whenever } G_0(x_0)+G'_0(x_0)(x_i-x_0)\in K, \quad \text{for } i=1,2.\end{aligned} \tag{57}$$

THEOREM 4. *If $G_0'(x_0)$ is surjective and the coercivity condition* (57) *holds, then there exists $s>0$ such that* (55) *has a strict local minimizer x_p for each $p \in B_s(0)$, and both x_p, and the associated* (*unique*) *multiplier $\lambda_p \in K^+$ satisfying the first-order necessary condition* (56), *are Lipschitz continuous functions of $p \in B_s(0)$.*

Proof. We apply Corollary 2 with the following identifications: $z=(x, \lambda)$, $Z= X \times X^*$, $\Omega = X \times K^+$,

$$T_p(x, \lambda) = \begin{bmatrix} \nabla_x H_p(x, \lambda) \\ G_p(x) \end{bmatrix}, \quad \text{and} \quad F(x, \lambda) = \begin{bmatrix} 0 \\ \partial K^+(\lambda) \end{bmatrix}.$$

Hence, the problem "Find $z \in \Omega$ such that $T_p(z) \in F(z)$" is the same as finding x and λ satisfying the first-order necessary condition for (55). In the auxiliary problem, we take $L = T_0'(z_0)$. Hence, the auxiliary problem is equivalent to the following: Given $\phi \in X^*$ and $\psi \in W$, find $x \in X$ and $\lambda \in K^+$ such that

$$\begin{aligned} &Ax - B^*\lambda + \phi = 0 \quad \text{and} \quad Bx + \psi \in \partial K^+(\lambda), \\ &\text{where } A = \nabla^2_{xx} H_0(x_0, \lambda_0) \quad \text{and} \quad B = G_0'(x_0). \end{aligned} \tag{58}$$

By Lemma 5, the solution to (58) is a Lipschitz continuous function of ϕ and ψ. By Corollary 2, there exists a solution $z_p = (x_p, \lambda_p)$ to (56), which is a Lipschitz continuous function of p for p near 0. By Lemma 6, the coercivity condition (57) holds when the zeros are replaced by p near 0. Hence, the second-order sufficiency condition holds (see Maurer and Zowe [30]), and x_p is a strict local minimizer of (55) for p near 0. □

Theorem 4 yields Lipschitz continuity without assuming strict complementary slackness. For an illustration, suppose that $G_p = (g_p, h_p)$ and $K = K_g \times K_h$, where K_g and K_h are closed convex cones with vertices at the origin of the associated Banach spaces. In this cases, the optimization problem (55) takes the form

$$\begin{aligned} &\text{minimize} \quad C_p(x) \\ &\text{subject to} \quad g_p(x) \in K_g, \quad h_p(x) \in K_h. \end{aligned} \tag{59}$$

The Lagrangian H_p is given by

$$H_p(x, \mu, \nu) = C_p(x) - \langle \mu, g_p(x) \rangle - \langle \nu, h_p(x) \rangle,$$

where $\lambda = (\mu, \nu)$ is the multiplier in the dual space. Again, if x_0 is a local minimizer of (59) and $G_0'(x_0)$ is surjective, then there exists a multiplier $\lambda_0 = (\mu_0, \nu_0)$ satisfying (56) for $p=0$. Let us assume that $h_0(x_0)=0$ and $\nu_0 \in \text{int } K_h^+$. In finite dimensions, h_p corresponds to the part of the inequality constraints that are active at $p=0$ with associated multipliers that are strictly positive. We make the following coercivity assumption:

$$\begin{aligned} &\langle \nabla^2_{xx} H_0(x_0, \mu_0, \nu_0)(x_2 - x_1), x_2 - x_1 \rangle \geqq \alpha \|x_2 - x_1\|^2 \\ &\text{whenever } g_0(x_0) + g_0'(x_0)(x_i - x_0) \in K_g, \quad h_0'(x_0)(x_i - x_0) = 0, \quad \text{for } i = 1, 2. \end{aligned} \tag{60}$$

By Theorem 4, the following optimization problem has a local minimizer for p near 0 that depends Lipschitz continuously on p:

$$\begin{aligned} &\text{minimize } C_p(x) \\ &\text{subject to} \quad g_p(x) \in K_g, \quad h_p(x) = 0. \end{aligned} \tag{61}$$

Observe that this problem differs from (59) since the constraint $h_p(x) \in K_h$ associated with (59) has been replaced by $h_p(x) = 0$. Exploiting the assumption that ν_0 lies in the interior of K_h^+, we show that this local minimizer for (61) is also a local minimizer of (59).

COROLLARY 4. *If* $(g_0'(x_0), h_0'(x_0))$ *is surjective, the coercivity condition* (60) *holds, and* ν_0 *lies in the interior of* K_h^+, *then there exists* $s > 0$ *such that* (59) *has a strict local minimizer* x_p *for each* $p \in B_s(0)$, *and both* x_p, *and the associated multipliers* $\mu_p \in K_g^+$ *and* $\nu_p \in K_h^+$ *satisfying the first-order necessary condition, are Lipschitz continuous functions of* $p \in B_s(0)$.

Proof. We apply the proof given for Theorem 4 to problem (61) replacing K by $K_g \times \{0\}$ and replacing the coercivity condition (57) by (60). It follows that there exists a solution x_p of (61) and associated Lagrange multipliers μ_p and ν_p that are Lipschitz continuous functions of p near 0 and that satisfy the first-order necessary conditions for (61). Since $\nu_p \in \text{int } K_h^+$ for p near zero, the first-order necessary conditions for (59) hold. By Lemma 8 of Appendix 1, x_p is a strict local minimizer of (61). □

Finally, let us suppose that $G_p = (f_p, g_p, h_p)$ and $K = K_f \times K_g \times K_h$, where K_f, K_g, and K_h are closed convex cones with vertices at the origin of the associated Banach spaces. Hence, the optimization problem (55) takes the form

$$\begin{aligned} &\text{minimize} \quad C_p(x) \\ &\text{subject to} \quad f_p(x) \in K_f, \quad g_p(x) \in K_g, \quad h_p(x) \in K_h. \end{aligned} \tag{62}$$

If $f_0(x_0) \in \text{int } K_f$, then under the hypotheses of Corollary 4, the solution x_p of (59) is a Lipschitz continuous function of p and $f_p(x_p) \in K_f$ for p near 0. Hence, the local minimizer x_p of (59) is a local minimizer of (62).

Remark 7. Theorem 4 and Corollary 4 yield Lipschitz continuity of a local minimizer in a neighborhood of a reference point, without assuming strict complementary slackness. In finite dimensions, this problem was studied by Hager in [18] and by Robinson in [37]. Note that the coercivity assumption (60) is slightly weaker, in the infinite-dimensional context, than the coercivity assumption used in earlier work (see [18], [23], [37]) since (60) only requires coercivity relative to those x_i satisfying the constraint $g_0(x_0) + g_0'(x_0)(x_i - x_0) \in K_g$.

In comparing Corollary 4 to the recent paper [23] of Ito and Kunisch, note that in [23] the infinite-dimensional constraints are linear inequalities, the problem is formulated in a Hilbert space, and the nonlinear constraints for which the associated dual multiplier can vanish are finite-dimensional. Alt presents in [1] and [3] a stability analysis that is related to ours, but different. In [1] he considers a cone constrained problem, under the assumption that any neighborhood of the reference point contains a solution of the perturbed problem. In [3] he studies Lipschitz continuity of the solution of a problem with nonlinear cone constraints and with equality constraints under a weaker constraint qualification (Robinson's constraint regularity condition), but a stronger coercivity—(60) is required to hold on the kernel of the gradients of the equality constraints; moreover, he assumes that the variation of the Lagrange multipliers for the perturbed problem can be estimated in terms of the variation in the solution and the variation in the parameter (under our surjectivity condition, this hypothesis is satisfied). Recently, Malanowski [28] has obtained a Lipschitz continuity result in a Hilbert space setting that parallels the analysis of Ito and Kunisch [23], using a regularity condition for the constraints that is weaker than surjectivity.

6. Optimal control. Let us consider a nonlinear optimal control problem with control constraints

$$\begin{aligned} &\text{minimize} \quad \int_I g_p(x(t), u(t))\, dt \\ (63)\qquad &\text{subject to} \quad \dot{x}(t) = f_p(x(t), u(t)) \quad \text{and} \quad u(t) \in U \quad \text{a.e. } t \in I, \\ &\qquad\qquad x(0) = a, \quad x \in W^{1,\theta}, \quad u \in L^\infty, \end{aligned}$$

where $f_p : R^{n+m} \to R^n$, $g_p : R^{n+m} \to R$, $U \subset R^m$ is nonempty, closed, and convex, a is the given starting condition, and $\theta \in [1, \infty]$. We assume that there exists a solution (x_0, u_0) to (63) corresponding to $p = 0$, and we wish to show that there exists a nearby solution for p in a neighborhood of zero. To this end, suppose that there exists a closed set $\Delta \subset R^n \times R^m$ and a $\delta > 0$ such that $(x_0(t), u_0(t))$ lies in Δ for almost every $t \in I$, the distance from $(x_0(t), u_0(t))$ to the boundary of Δ is at least δ for almost every $t \in I$, and the first two derivatives of $f_p(x, u)$ and $g_p(x, u)$ with respect to x and u exist on Δ, and these derivatives, along with the function value $f_p(x, u)$, are continuous with respect to $(x, u) \in \Delta$ and p near zero.

Let H_p denote the Hamiltonian defined by

$$H_p(x, u, \lambda) = g_p(x, u) + \lambda^T f_p(x, u).$$

If (x_p, u_p) is a solution of (63), then the minimum principle [22, p. 134] implies the following:

$$\nabla_u H_p(x_p(t), u_p(t), \lambda_p(t))^T (v - u_p(t)) \geqq 0 \quad \text{a.e. } t \in I \text{ and for every } v \in U,$$

where $\lambda = \lambda_p$ is the solution of the adjoint equation

$$\dot{\lambda}(t) = -\nabla_x H_p(x(t), u(t), \lambda(t)) \quad \text{a.e. } t \in I, \quad \lambda(1) = 0,$$

associated with $x = x_p$ and $u = u_p$. Let $f_0^*(t)$ and $H_0^*(t)$ stand for $f_0(x_0(t), u_0(t))$ and $H_0(x_0(t), u_0(t), \lambda_0(t))$, and define the matrices

$$A(t) = \nabla_x f_0^*(t), \quad B(t) = \nabla_u f_0^*(t), \quad Q(t) = \nabla_{xx}^2 H_0^*(t),$$
$$R(t) = \nabla_{uu}^2 H_0^*(t), \quad S(t) = \nabla_{xu}^2 H_0^*(t).$$

The following coercivity assumption will be utilized: There exists $\alpha > 0$ such that

$$(64) \qquad \int_I (x(t)^T Q(t) x(t) + u(t)^T R(t) u(t) + 2x(t)^T S(t) u(t))\, dt \geqq \alpha \int_I |u(t)|^2\, dt$$

whenever $x \in W^{1,2}$, $x(0) = 0$, $u \in L^2$, $\dot{x} = Ax + Bu$, $u = v - w$ for some v and $w \in L^2$ with $v(t)$ and $w(t) \in U$ for almost every $t \in I$. By taking $v = w$ except on a small interval, it can be shown, below, that a pointwise coercivity condition holds (see the recent paper [15]):

$$(65) \qquad u^T R(t) u \geqq \alpha |u|^2 \quad \text{a.e. } t \in I \text{ whenever } u = v - w \text{ with } v \text{ and } w \in U.$$

THEOREM 5. *If the coercivity condition* (64) *holds, then there exists positive constants* κ, r, *and* s *such that for each* $p \in B_s(0)$, (63) *has a strict local minimizer* $(x_p, u_p) \in B_r(x_0, u_0)$ *and the relation*

$$\begin{aligned} &\|x_p - x_q\|_{W^{1,\theta}} + \|u_p - u_q\|_{L^\infty} + \|\lambda_p - \lambda_q\|_{W^{1,\theta}} \\ &\quad \leqq \kappa(\|f_q(x_p, u_p) - f_p(x_p, u_p)\|_{L^\theta} \\ &\qquad + \|\nabla_u H_q(x_p, u_p, \lambda_p) - \nabla_u H_p(x_p, u_p, \lambda_p)\|_{L^\infty} \\ &\qquad + \|\nabla_x H_q(x_p, u_p, \lambda_p) - \nabla_x H_p(x_p, u_p, \lambda_p)\|_{L^\theta}) \end{aligned}$$

holds whenever p *and* $q \in B_s(0)$.

Proof. We apply Corollary 2 where the z of Corollary 2 is identified with (x, u, λ), while Ω, T_p, and F are defined in the following way:

$$\Omega = \{(x, u, \lambda): x \in W^{1,\theta}, u \in L^\infty, \lambda \in W^{1,\theta}, x(0) = a,$$
$$\lambda(1) = 0, u(t) \in U \text{ a.e. } t \in I\},$$

$$T_p(x, u, \lambda) = \begin{bmatrix} \nabla_x H_p(x, u, \lambda) + \dot{\lambda} \\ \nabla_u H_p(x, u, \lambda) \\ f_p(x, u) - \dot{x} \end{bmatrix}, \quad \text{and} \quad F(x, u, \lambda) = \begin{bmatrix} 0 \\ \partial U(u) \\ 0 \end{bmatrix}.$$

The space Y containing the range of T_p is $L^\theta \times L^\infty \times L^\theta$. We make the following choice for the operator L of Corollary 2: $L(z) = M(z - z_0)$, where

$$M(x, u, \lambda) = \begin{bmatrix} A^T\lambda + Qx + Su + \dot{\lambda} \\ Ru + S^T x + B^T \lambda \\ Ax + Bu - \dot{x} \end{bmatrix}.$$

It can be verified that under the smoothness assumptions, $E_r(p) \to 0$ as p and r tend to zero, where $E_r(p)$ is defined in (12).

Given q_i and s_i in L^θ and r_i in L^∞ for $i = 1$ and 2, let us consider the following problem:

$$\text{(66)} \qquad \text{Find } (x, u, \lambda) \in \Omega \text{ such that } L(x, u, \lambda) + \begin{bmatrix} q_i \\ r_i \\ s_i \end{bmatrix} \in F(x, u, \lambda).$$

In [20, Lemma 3], we show that when the coercivity assumptions (64) and (65) hold, (66) has a unique solution (x_i, u_i, λ_i), and the following Lipschitz property holds:

$$\|x_2 - x_1\|_{W^{1,\theta}} + \|u_2 - u_1\|_{L^\infty} + \|\lambda_2 - \lambda_1\|_{W^{1,\theta}}$$
$$\leqq \gamma[\|q_2 - q_1\|_{L^\theta} + \|r_2 - r_1\|_{L^\infty} + \|s_2 - s_1\|_{L^\theta}].$$

Hence, Ψ is Lipschitz, and by Corollary 2, problem (11) has a locally unique solution that satisfies the conclusion of Theorem 5. Since the estimate of Theorem 5 yields an L^∞ perturbation in both state and the control, and since the coercivity assumption (64) is preserved after small perturbations in Q, R, S, A, and B, it follows from Corollary 5 in Appendix 1 that the solution of (11) provided by Corollary 2 is a strict local minimizer for the optimal control problem (63) when p is near zero. □

We show by an example that Lipschitz continuity can be obtained without the coercivity condition (64). Consider the following problem:

$$\begin{aligned} &\text{minimize} && x_1(1) + x_2(1) \\ \text{(67)} \qquad &\text{subject to} && \dot{x}_1 = p x_1 \sin x_2 + u_1, \quad \dot{x}_2 = u_2, \, x_1(0) = 1, \quad x_2(0) = 1, \\ & && u_1^2 + u_2^2 \leqq 2, \end{aligned}$$

where p is a real parameter. For $p = 0$, the optimal solution is $u_0 = (-1, -1)$ and $x_0 = (1 - t, 1 - t)$, and the corresponding adjoint variable is $\lambda_0 = (-1, -1)$. The auxiliary problem has the form

$$\dot{x}_i = u_i + s_i, \quad x_i(0) = 1, \quad \dot{\lambda}_i = q_i, \quad \lambda_i(1) = -1, \qquad \text{for } i = 1, 2,$$
$$(\lambda + r)^T (v - u) \geqq 0 \quad \text{whenever } v_1^2 + v_2^2 \leqq 2.$$

The control solving the auxiliary problem is

$$u = \sqrt{2}\, w/|w|, \quad \text{where } w_i(t) = r_i(t) - 1 + \int_1^t q(s)\, ds.$$

Hence, the solution of the auxiliary problem is unique and Lipschitz continuous, with respect to $y=(s, r, q)$ around 0, as a function from L^∞ to $W^{1,\infty}\times L^\infty$. By Corollary 2 and for p near zero, there exists a solution (x_p, u_p, λ_p) of the first-order necessary conditions associated with (67), which is unique in a neighborhood of (x_0, u_0, λ_0) and which is a Lipschitz continuous function of p near 0. Using the uniqueness of (x_0, u_0), it can be shown that (x_p, u_p) is the unique solution of (67). In this example, we use the strong convexity of the constraining set instead of the coercivity condition to ensure Lipschitz continuity.

Remark 8. Lipschitz results for problems with convex cost, linear dynamics, and linear inequality state and control constraints are obtained by Dontchev [13, Chap. 2] using duality theory and the regularity of the optimal control established by Hager [18]. Later, Malanowski [27] studied a problem with a quadratic cost functional, linear inequality state and control constraints, and system dynamics that are linear with respect to the control. A similar problem without state constraints, but with convex control constraints, is considered by Alt in [3] using Robinson's strong regularity condition. In [2] Alt considered a nonlinear problem with inequality control constraints. He obtains an estimate for the optimal control, assuming existence of a solution to the perturbed problem in a neighborhood of the reference point (see Remark 7).

7. Euler's method. Again, let us consider a nonlinear control problem with control constraints, below:

$$\begin{aligned} &\text{minimize} \quad \int_I g(x(t), u(t))\, dt \\ &\text{subject to} \quad \dot{x}(t)=f(x(t), u(t)) \quad \text{and} \quad u(t)\in U \quad \text{a.e. } t\in I, \\ &\qquad\qquad x(0)=a, \quad x\in W^{1,\infty}, \quad u\in L^\infty, \end{aligned} \tag{68}$$

where $f: R^{n+m}\to R^n$, $g: R^{n+m}\to R$, $U\subset R^m$ is nonempty, closed, and convex, and a is the given starting condition. We assume that there exists a solution (x^*, u^*) to (68) with u^* Riemann integrable, that there exists a closed set $\Delta\subset R^{n+m}$ where both f and g are twice continuously differentiable, and that there exists $\delta>0$ such that $(x^*(t), u^*(t))\in\Delta$ and the distance from $(x^*(t), u^*(t))$ to the boundary of Δ is at least δ for every $t\in I$. When we write $\dot{x}^*$, we mean a function whose values on I coincide with those of $f(x^*, u^*)$.

Let H denote the Hamiltonian defined by

$$H(x, u, \lambda)=g(x, u)+\lambda^T f(x, u),$$

and let $\lambda=\lambda^*$ be the solution of the adjoint equation

$$\dot{\lambda}(t)=-\nabla_x H(x(t), u(t), \lambda(t)) \quad \text{a.e. } t\in I, \lambda(1)=0, \tag{69}$$

associated with $x=x^*$ and $u=u^*$. By the minimum principle [22, p. 134], we have

$$\nabla_u H(x^*(t), u^*(t), \lambda^*(t))^T(v-u^*(t))\geqq 0 \quad \text{a.e. } t\in I \text{ and for every } v\in U. \tag{70}$$

Given a natural number N, let $h=1/N$ be the mesh spacing, and let x_i and u_i denote approximations to $x(t)$ and $u(t)$ at $t=t_i=ih$. We consider the Euler discretization of (68) given by

$$\begin{aligned} &\text{minimize} \quad \sum_{i=0}^{N-1} hg(x_i, u_i) \\ &\text{subject to} \quad x_{i+1}=x_i+hf(x_i, u_i) \quad \text{and} \\ &\qquad\qquad u_i\in U, i=0, 1, \dots, N-1, x_0=a. \end{aligned} \tag{71}$$

If (x^h, u^h) denotes a solution to (71), let $\lambda = \lambda^h$ denote the solution of the discrete adjoint equation

$$\lambda_i = \lambda_{i+1} + h\nabla_x H(x_i, u_i, \lambda_{i+1}), \qquad i = N-1, N-2, \ldots, 0, \lambda_N = 0, \tag{72}$$

associated with $x = x^h$ and $u = u^h$. By the discrete minimum principle [22, p. 280], we have

$$\nabla_u H(x_i^h, u_i^h, \lambda_{i+1}^h)^T(v - u_i^h) \geqq 0 \quad \text{for all } v \in U, \qquad i = 0, 1, \ldots, N-1. \tag{73}$$

To estimate the distance between (x^*, u^*) and (x^h, u^h), we need a coercivity-type assumption for the discrete problem. Define the following matrices:

$$A(t) = \nabla_x f^*(t), \quad B(t) = \nabla_u f^*(t), \quad Q(t) = \nabla_{xx}^2 H^*(t),$$
$$R(t) = \nabla_{uu}^2 H^*(t), \quad S(t) = \nabla_{xu}^2 H^*(t).$$

Here $f^*(t)$ and $H^*(t)$ stand for $f(x^*(t), u^*(t))$ and $H(x^*(t), u^*(t), \lambda^*(t))$, respectively. Letting A_i, B_i, Q_i, S_i, and R_i denote the corresponding time-varying matrices evaluated at $t = t_i$, we assume that there exists a scalar $\alpha > 0$, α independent of N, such that

$$u^T R_i u \geqq \alpha |u|^2, 0 \leqq i \leqq N-1 \quad \text{whenever } u = v - w \text{ with } v \text{ and } w \in U, \tag{74}$$

$$\sum_{i=0}^{N-1} x_i^T Q_i x_i + u_i^T R_i u_i + 2x_i^T S_i u_i \geqq \alpha \sum_{i=0}^{N-1} |u_i|^2 \tag{75}$$

whenever $u_i = v_i - w_i$ for some v_i and $w_i \in U$, and

$$x_{i+1} = x_i + hA_i x_i + hB_i u_i, \qquad i = 0, 1, \ldots, N-1, \quad x_0 = 0. \tag{76}$$

Obviously, the discrete condition (74) holds if there exists $\alpha > 0$ such that

$$u^T R(t) u \geqq \alpha |u|^2 \quad \text{for every } t \in I \text{ and for each } u = v - w \text{ with } v \text{ and } w \in U.$$

In Appendix 2, we show that assumption (75) for the discrete problem can be deduced from an analogous assumption for the continuous problem. In analyzing the discrete problem (71), we utilize a discrete L^p norm defined by

$$(\|u\|_{L^p})^p = \sum_{i=0}^{N-1} h|u_i|^p, 1 \leqq p < \infty, \quad \text{and} \quad \|u\|_{L^\infty} = \text{maximum } \{|u_i|: 0 \leqq i < N\}.$$

If ϕ and v satisfy the finite difference system

$$\phi_{i+1} = \phi_i + hA_i\phi_i + hv_i, \qquad i = 0, 1, \ldots, N-1, \quad \phi_0 = 0,$$

then there exists a constant c, independent of h, such that

$$|\phi_j| \leqq c\|v\|_{L^1} \leqq c\|v\|_{L^2} \quad \text{for each } j = 0, 1, \ldots, N. \tag{77}$$

Squaring this inequality, multiplying by h, and summing over j yields

$$\|\phi\|_{L^2} \leqq c\|v\|_{L^2}.$$

Hence, if the coercivity condition (75) holds relative to the control, then the following joint state-control coercivity condition holds: There exists $\alpha > 0$ such that

$$h \sum_{i=0}^{N-1} x_i^T Q_i x_i + u_i^T R_i u_i + 2x_i^T S_i u_i \geqq \alpha(\|x\|_{L^2}^2 + \|u\|_{L^2}^2)$$

whenever $u_i = v_i - w_i$ for some v_i and $w_i \in U$, and

$$x_{i+1} = x_i + hA_i x_i + hB_i u_i, \qquad i = 0, 1, \ldots, N-1, \quad x_0 = 0.$$

Our convergence result for the discrete problem is expressed in terms of a modulus of smoothness introduced by Sendov and Popov [42]. The local modulus of continuity $\omega(u; t, h)$ of the function u is defined by

$$\omega(u; t, h) = \sup \{|u(a) - u(b)|: a, b \in [t - h/2, t + h/2] \cap I\},$$

while the average modulus of smoothness τ is given by

$$\tau(u; h) = \int_I \omega(u; t, h)\, dt.$$

In [42, pp. 8–11] it is shown that $\tau(u; h) \to 0$ as $h \to 0$ if and only if the bounded function u is Riemann integrable on I; moreover, $\tau(u; h) = O(h)$ if and only if u has bounded variation on I. The main result in this section is the following theorem.

THEOREM 6. *If u^* is Riemann integrable and the coercivity assumptions* (74) *and* (75) *hold, then for all N sufficiently large, there exists a local minimizer (x^h, u^h) of* (71) *such that*

$$\underset{0 \leqq i \leqq N-1}{\text{maximum}} |u^*(t_i) - u_i^h| = O(h + \tau(u^*; h)),$$

$$\underset{0 \leqq i \leqq N}{\text{maximum}} |x^*(t_i) - x_i^h| = O(h + \tau(u^*; h)),$$

$$\underset{0 \leqq i \leqq N}{\text{maximum}} |\lambda^*(t_i) - \lambda_i^h| = O(h + \tau(u^*; h)),$$

$$\underset{0 \leqq i \leqq N-1}{\text{maximum}} \left| \dot{x}^*(t_i) - \frac{x_{i+1}^h - x_i^h}{h} \right| = O(h + \tau(u^*; h)).$$

Hence, if u^* has bounded variation, then each of these error estimates is of order h.

Proof. We apply Corollary 1 to the necessary conditions associated with the discrete problem (71). The parameter p of Corollary 1 is identified with the mesh spacing h the set Ω_p consists of discrete triples (x, u, λ), where $u_i \in U$ for each i. Component i, $0 \leqq i \leqq N-1$, of the operators T_p and F_p, denoted T_i^h and F_i^h, respectively, is the following:

$$T_i^h(x, u, \lambda) = \begin{bmatrix} \nabla_x H(x_i, u_i, \lambda_{i+1}) + (\lambda_{i+1} - \lambda_i)/h \\ \nabla_u H(x_i, u_i, \lambda_{i+1}) \\ f(x_i, u_i) - (x_{i+1} - x_i)/h \end{bmatrix} \quad \text{and} \quad F_i^h(x, u, \lambda) = \begin{bmatrix} 0 \\ \partial U(u_i) \\ 0 \end{bmatrix}.$$

Given $z = (x, u, \lambda)$ in the discrete space Z_p associated with Ω_p, we use the L^∞ norm for each of the three components x, u, and λ of z. In the discrete space Y_p associated with the range of T_p, we use the L^1 norm for the first and last component, and the L^∞ norm for the middle component. That is, if $y = (a, b, c) \in Y_p$, then

$$\|y\|_p = \|a\|_{L^1} + \|b\|_{L^\infty} + \|c\|_{L^1}.$$

The point z_p of Corollary 1 is given by $z_p = (x^I, u^I, \lambda^I)$, where

$$x_i^I = x^*(t_i), \quad u_i^I = u^*(t_i), \quad \lambda_i^I = \lambda^*(t_i).$$

Also, in Corollary 1, component i of the point y_p, denoted y_i^h, is the triple

$$y_i^h = \begin{bmatrix} 0 \\ \nabla_u H(x_i^I, u_i^I, \lambda_i^I) \\ 0 \end{bmatrix}.$$

Observe that with this choice for y_p, we have $y_p \in F(z_p)$. We define a linear operator M^h that acts on a discrete triple (x, u, λ) to produce a vector whose ith component is

$$M_i^h(x, u, \lambda) = \begin{bmatrix} A_i^T \lambda_{i+1} + Q_i x_i + S_i u_i + (\lambda_{i+1} - \lambda_i)/h \\ R_i u_i + S_i^T x_i + B_i^T \lambda_{i+1} \\ A_i x_i + B_i u_i - (x_{i+1} - x_i)/h \end{bmatrix}.$$

Taking $L_p(z) = M^h(x, u, \lambda) - M^h(x^I, u^I, \lambda^I)$, observe that $L_p(z_p) = 0$.

It can be verified that under the smoothness assumptions and for ρ smaller than δ, we have $D_\rho(h) \to 0$ as $h \to 0$. Now consider the term

$$(T_p(z_p) - y_p)_i = \begin{bmatrix} \nabla_x H(x_i^I, u_i^I, \lambda_{i+1}^I) + (\lambda_{i+1}^I - \lambda_i^I)/h \\ \nabla_u H(x_i^I, u_i^I, \lambda_{i+1}^I) - \nabla_u H(x_i^I, u_i^I, \lambda_i^I) \\ f(x_i^I, u_i^I) - (x_{i+1}^I - x_i^I)/h \end{bmatrix}. \tag{78}$$

The middle component of this vector is $O(h)$ since λ^* is Lipschitz continuous. Since the analysis of the first and last component in (78) is similar, we only focus on the last component

$$\begin{aligned} \left| f(x_i^I, u_i^I) - \frac{x_{i+1}^I - x_i^I}{h} \right| &\leqq \frac{1}{h} \int_{t_i}^{t_{i+1}} |f(x_i^I, u_i^I) - \dot{x}^*(t)| \, dt \\ \leqq \frac{1}{h} \int_{t_i}^{t_{i+1}} |f(x_i^I, u_i^I) - f(x^*(t), u^*(t))| \, dt &\leqq \frac{c}{h} \int_{t_i}^{t_{i+1}} [h + \omega(u^*; t, 2h)] \, dt, \end{aligned} \tag{79}$$

where c denotes a generic constant, independent of h. Multiplying (79) by h, summing over i, and exploiting the inequality $\tau(u; kh) \leqq k\tau(u; h)$ for each natural number k (see [42, p. 11]), it follows that

$$\| T_p(z_p) - y_p \|_p = O(h + \tau(u^*; h)).$$

Next, we must analyze the auxiliary problem and establish the existence of a constant γ satisfying (9). The analysis essentially parallels that of [20] except that continuous norms are replaced by their discrete analogues. We must examine how the solution to the following system depends on the perturbations q_i, r_i, and s_i:

$$\begin{gathered} A_i^T \lambda_{i+1} + Q_i x_i + S_i u_i + \frac{\lambda_{i+1} - \lambda_i}{h} + q_i = 0, \qquad \lambda_N = 0, \\ (R_i u_i + S_i^T x_i + B_i^T \lambda_{i+1} + r_i)(v - u_i) \geqq 0 \quad \text{for every } v \in U, \\ A_i x_i + B_i u_i - \frac{x_{i+1} - x_i}{h} + s_i = 0, \qquad x_0 = a, \end{gathered} \tag{80}$$

$i = 0, 1, \ldots, N-1$. Note that system (80) constitutes the first-order necessary conditions (see [22, p. 280]) associated with the following quadratic program:

$$\begin{aligned} &\text{minimize} \quad h \sum_{i=0}^{N-1} \frac{1}{2} x_i^T Q_i x_i + \frac{1}{2} u_i^T R_i u_i + x_i^T S_i u_i + q_i^T x_i + r_i^T u_i \\ &\text{subject to} \quad x_{i+1} = x_i + hA_i x_i + hB_i u_i + hs_i \quad \text{and} \\ &\qquad\qquad u_i \in U, \quad 0 \leqq i \leqq N-1, \quad x_0 = a. \end{aligned} \tag{81}$$

By Lemma 4 and the discussion that follows it, there is a one-to-one correspondence between a solution to (81) and a solution to (80) when (75) holds.

Now consider the perturbations (q^i, r^i, s^i) for $i=1$ and 2. Let (x^i, u^i, λ^i) denote the associated solutions to (80). Referring to Lemma 4 (see [20, § 2] for more details), we have

$$\|u^1-u^2\|_{L^2} \leqq c(\|q^1-q^2\|_{L^1}+\|r^1-r^2\|_{L^2}+\|s^1-s^2\|_{L^1}),$$

where c is a constant independent of h. Utilizing (77), we also conclude that

$$\|x^1-x^2\|_{L^\infty}+\|\lambda^1-\lambda^2\|_{L^\infty} \leqq c(\|q^1-q^2\|_{L^1}+\|r^1-r^2\|_{L^2}+\|s^1-s^2\|_{L^1}). \tag{82}$$

Finally, by (74) and Lemma 4, we have

$$\|u^1-u^2\|_{L^\infty} \leqq c(\|r^1-r^2\|_{L^\infty}+\|x^1-x^2\|_{L^\infty}+\|\lambda^1-\lambda^2\|_{L^\infty}).$$

Combining this with (82) yields

$$\begin{aligned}&\|x^1-x^2\|_{L^\infty}+\|u^1-u^2\|_{L^\infty}+\|\lambda^1-\lambda^2\|_{L^\infty}\\ &\leqq c(\|q^1-q^2\|_{L^1}+\|r^1-r^2\|_{L^\infty}+\|s^1-s^2\|_{L^1}).\end{aligned}$$

Hence, there exists a constant γ such that (9) holds with $\sigma=\infty$.

By Corollary 1, there exists a solution to the discrete necessary conditions (72) and (73) associated with (71) that satisfies the first three estimates of Theorem 6. The discrete and continuous state equations, along with the previously established error estimates, imply that

$$\left|\dot{x}^*(t_i)-\frac{x_{i+1}^h-x_i^h}{h}\right| = |f(x^*(t_i), u^*(t_i))-f(x_i^h, u_i^h)| = \mathrm{O}(h+\tau(u^*; h)),$$

which gives the last estimate of Theorem 6. The fact that x^h and u^h are local minimizers for (71) follows from Corollary 6 in Appendix 1, the coercivity condition (75), and the fact that coercivity condition (75) is preserved after small perturbations in Q_i, R_i, S_i, A_i, and B_i. □

Remark 9. Note that the coercivity assumptions (74) and (75) do not necessarily imply that an optimal control is either unique or continuous. For example, if $g(x, u)=(u^2-1)^2$, $f=0$, and $U=R^1$, then for each measurable set $M\subset 1$, the function defined by

$$u(t)=1 \quad \text{for } t\in M \quad \text{and} \quad u(t)=-1 \quad \text{for } t\notin M$$

is an optimal control that satisfies (74) and (75).

Remark 10. Results most closely related to Theorem 6 include the papers of Budak, Berkovich, and Solov'eva [8] and Cullum [11] in which convergence of the optimal value associated with discrete approximations to state and control constrained problems is established. Mordukhovich [31] shows that the discrete optimal cost converges to the true optimal cost if and only if a relaxation of the control problem is stable. Estimates for the error in the optimal control associated with higher-order discretizations of unconstrained nonlinear problems are derived by Hager [17]. Dontchev [12] obtains an error estimate for Euler's approximation applied to an optimal control problem with convex cost, linear system dynamics, and linear inequality state and control constraints.

Appendix 1: Sufficient optimality conditions. We begin by establishing the sufficient optimality result needed for Corollary 4. Let us consider the following optimization problem:

$$\begin{aligned}&\text{minimize} \quad C(z)\\ &\text{subject to} \quad g(z)\in K_g, \quad h(z)\in K_h,\end{aligned} \tag{83}$$

where $g: Z \to W_g$ and $h: Z \to W_h$, W_g and W_h are Banach spaces, and K_g and K_h are closed, convex cones with vertices at the origin of their respective spaces. The Lagrangian H associated with (83) is given by

$$H(z, \mu, \nu) = C(z) - \langle \mu, g(z) \rangle - \langle \nu, h(z) \rangle,$$

where $\mu \in W_g^*$ and $\nu \in W_h^*$. Letting z^* be a point that is feasible in (83), we assume that C, g, and h are twice Fréchet differentiable at z^*. The first-order necessary conditions associated with (83) have the following form: There exists $\mu \in K_g^+$ and $\nu \in K_h^+$ such that

$$\nabla_z H(z^*, \mu, \nu) = 0, \quad \langle \mu, g(z^*) \rangle = 0, \quad \langle \nu, h(z^*) \rangle = 0. \tag{84}$$

Although the following lemma makes the same surjectivity assumption that appears in Corollary 4, this assumption can be replaced by any condition that ensures regularity of the linearized system (see Robinson [35] or Maurer and Zowe [30]).

LEMMA 8. *Suppose that z^* is feasible in* (83), $\mu \in K_g^+$, $\nu \in \operatorname{int} K_h^+$, *the first-order necessary conditions* (84) *hold, the operator*

$$\begin{bmatrix} h'(z^*) \\ g'(z^*) \end{bmatrix}$$

is surjective, and there exists $\alpha > 0$ such that

$$\langle \nabla_{zz}^2 H(z^*, \mu, \nu)(z - z^*), z - z^* \rangle \geqq \alpha \|z - z^*\|^2$$

$$\textit{whenever } g(z^*) + g'(z^*)(z - z^*) \in K_g \quad \textit{and} \quad h'(z^*)(z - z^*) = 0.$$

Then z^* is a strict local minimizer for (83).

In comparing this result to Maurer and Zowe's classic sufficient optimality result [30], observe that the constraint $h'(z^*)(z - z^*) = 0$ in the coercivity condition above corresponds to a constraint of the form $h'(z^*)(z - z^*) \in K_h$ in [30]. In this respect, the coercivity condition of Lemma 8 is weaker than that of [30]. On the other hand, Lemma 8 assumes that $\nu \in \operatorname{int} K_h^+$, while [30] only assumes that $\nu \in K_h^+$.

Proof. Throughout this proof, we let ε denote a generic positive constant that can be made arbitrarily small for z sufficiently close to z^*, we let α denote a generic positive constant that is uniformly bounded away from zero for z near z^*, and we let β denote a generic constant that is uniformly bounded from above for z near z^*. Expanding $H(z, \mu, \nu)$ in a Taylor series about $z = z^*$, we have

$$H(z, \mu, \nu) = H(z^*, \mu, \nu) + \nabla_z H(z^*, \mu, \nu)(z - z^*) + \tfrac{1}{2} \nabla_{zz}^2 H(z^*, \mu, \nu)(z - z^*, z - z^*) + R(z),$$

where $R(z) \leqq \varepsilon \|z - z^*\|^2$. By the first-order necessary conditions, $H(z^*, \mu, \nu) = C(z^*)$ and $\nabla_z H(z^*, \mu, \nu) = 0$. Hence, it follows that

$$C(z) = C(z^*) + M(z) + R(z),$$

$$\text{where } M(z) = \langle \mu, g(z) \rangle + \langle \nu, h(z) \rangle + \tfrac{1}{2} \nabla_{zz}^2 H(z^*, \mu, \nu)(z - z^*, z - z^*).$$

If z is feasible in (83), then since $\nu \in \operatorname{int} K_h^+$, we have $\langle \nu, h(z) \rangle \geqq \alpha \|h(z)\|$. Thus we have

$$\langle \mu, g(z) \rangle + \langle \nu, h(z) \rangle \geqq \alpha \|h(z)\|.$$

By the complementary slackness condition, $h(z^*) = 0$, and by the differentiability assumption,

$$h(z) = h'(z^*)(z - z^*) + o(\|z - z^*\|) \quad \text{and} \quad g(z) = g(z^*) + g'(z^*)(z - z^*) + o(\|z - z^*\|).$$

Referring to [35, Thm. 1], the surjectivity assumption implies that for each z near z^*, there exists an associated $y \in Z$ such that $h'(z^*)(y-z^*)=0$, $g(z^*)+g'(z^*)(y-z^*) \in K_g$, and

$$\|y-z\| \leqq \beta\|h'(z^*)(z-z^*)\| + \beta \inf\{\|g(z^*)+g'(z^*)(z-z^*)-k\| : k \in K_g\}.$$

This bound, combined with the Taylor expansions of g and h, implies that for each z near z^*, with z feasible in (83), there exists an associated $y \in Z$ such that $h'(z^*)(y-z^*)=0$, $g(z^*)+g'(z^*)(y-z^*) \in K_g$, and $\|y-z\| \leqq \beta\|h(z)\| + \varepsilon\|z-z^*\|$. Applying the triangle inequality yields

$$\|y-z\| \leqq \beta\|h(z)\| + \varepsilon\|y-z^*\|.$$

By the coercivity assumption,

$$\nabla^2_{zz} H(z^*, \mu, \nu)(z-z^*, z-z^*) \geqq \alpha\|y-z^*\|^2 - \beta\|y-z^*\|\,\|y-z\| - \beta\|y-z\|^2.$$

These inequalities, along with the relation $h(z^*)=0$, imply that for z near z^* with z feasible in (83), we have

$$M(z) \geqq \alpha\|h(z)\| + \alpha\|y-z^*\|^2.$$

Recall that the remainder term R has the bound $R(z) \leqq \varepsilon\|z-z^*\|^2$. By the triangle inequality,

$$\|z-z^*\| \leqq \|z-y\| + \|y-z^*\| \leqq (1+\varepsilon)\|y-z^*\| + \beta\|h(z)\|,$$

from which it follows that

$$\|z-z^*\|^2 \leqq \beta\|y-z^*\|^2 + \beta\|h(z)\|^2.$$

Hence, for z near z^* with z feasible in (83), we have

$$\begin{aligned} C(z)-C(z^*) = M(z)+R(z) &\geqq \alpha\|h(z)\| + \alpha\|y-z^*\|^2 - \varepsilon\|z-z^*\|^2 \\ &\geqq \alpha\|h(z)\| + \alpha\|z-z^*\|^2 - \beta\|h(z)\|^2 \geqq \alpha\|z-z^*\|^2, \end{aligned}$$

which completes the proof. □

Next, we obtain sufficient optimality conditions that are applicable to optimal control problems. A number of relevant sufficient optimality conditions have appeared in the literature; for example, see Ioffe [21] and, in particular, the results of Maurer [29]. Although the basic strategy for obtaining sufficient optimality results in the optimal control setting is developed nicely by Maurer in [29], the precise results that we need in §§ 6 and 7 are not stated in his paper. For completenesss, we give a brief, self-contained treatment of the results needed in our paper. We begin with the abstract problem

$$\text{(85)} \qquad \begin{array}{ll} \text{minimize} & C(z) \\ \text{subject to} & z \in \Lambda, \end{array}$$

where Λ is a subset of a normed vector space Z, and C is a real-valued function. As in [29], we assume that there are two different norms, denoted by $\|\cdot\|$ and $|||\cdot|||$, associated with Z.

LEMMA 9. *Suppose that z^* satisfies the constraints of* (85) *and that there exists a functional M and a scalar $\alpha > 0$ with the following property*:

(86) $\qquad M(z) \geqq \alpha|||z-z^*|||^2$ *for each $z \in \Lambda$ with $\|z-z^*\|$ sufficiently small*

and

$$\frac{C(z)-C(z^*)-M(z)}{|||z-z^*|||^2}\to 0 \quad as \ \|z-z^*\|\to 0 \quad with \ z\in\Lambda. \tag{87}$$

Then z^* is a strict local minimizer for (85).

Proof. By the hypotheses above, we have

$$C(z)-C(z^*)\geqq \alpha|||z-z^*|||^2+o(|||z-z^*|||^2)$$

as $\|z-z^*\|\to 0$ with $z\in\Lambda$, which implies that z^* is a strict local minimizer for (85). □

In the application of Lemma 9, the following observation is helpful.

LEMMA 10. *Suppose that there exists a scalar $\alpha>0$, a bilinear form b that is bounded relative to the norm $|||\cdot|||$, and a set T such that*

$$b(z-z^*, z-z^*)\geqq \alpha|||z-z^*|||^2 \quad for \ every \ z\in T.$$

If for each $z\in\Lambda$, there exists $y\in T$ such that $|||z-y|||=o(|||z-z^*|||)$, then for each $\beta<\alpha$, we have

$$b(z-z^*, z-z^*)\geqq \beta|||z-z^*|||^2 \quad for \ all \ z\in\Lambda \ with \ \|z-z^*\| \ sufficiently \ small.$$

Proof. Given $z\in\Lambda$, let $y\in T$ be the hypothesized point for which $|||z-y|||=o(|||z-z^*|||)$. Since the bilinear form b is bounded, there exists a constant c such that

$$b(z-z^*, z-z^*)\geqq \alpha|||y-z^*|||^2-c|||z-y|||^2-c|||z-y|||\,|||y-z^*|||.$$

The inequality

$$|||y-z^*|||\geqq|||z-z^*|||-|||z-y|||=|||z-z^*|||-o(|||z-z^*|||)$$

completes the proof. □

We now apply Lemmas 9 and 10 to optimal control problems. Note that in the following result, we neither assume an interior point nor controllability.

COROLLARY 5. *Suppose that x^* and u^* are feasible for the optimal control problem* (68), *that f and g satisfy the differentiability conditions given below* (68), *that $\lambda=\lambda^*$ is the solution to the adjoint equation* (69) *associated with $x=x^*$ and $u=u^*$, and that the minimum principle* (70) *holds. If there exists $\alpha>0$ such that*

$$\int_I (x(t)^TQ(t)x(t)+u(t)^TR(t)u(t)+2x(t)^TS(t)u(t))\,dt\geqq\alpha\int_I|u(t)|^2\,dt \tag{88}$$

whenever $x\in W^{1,2}$, $x(0)=0$, $u\in L^2$, $\dot{x}=Ax+Bu$, $u=v-u^*$ for some $v\in L^2$ with $v(t)\in U$ for almost every $t\in I$, then u^* is a strict local minimizer for (68).

Proof. We apply Lemmas 9 and 10 with the following identifications: The z of Lemma 9 is the pair (x, u), the space Z is $W^{1,\infty}\times L^\infty$, the norm $|||\cdot|||$ associated with Z is the L^2 inner product norm, $C(z)$ is the integral cost function in (68), and Λ consists of those (x, u) in a convex neighborhood of (x^*, u^*) that satisfy the constraints

$$F(z)=0, \quad \text{where } F(z)=f(x,u)-\dot{x}, \qquad u(t)\in U \quad \text{a.e. } t\in I, \quad \text{and} \quad x(0)=a.$$

The functional M is defined by

$$M(z)=\langle\nabla_u H(x^*, u^*, \lambda^*), u-u^*\rangle+b(z-z^*, z-z^*),\ z=(x,u),$$

where

$$b(\delta z, \delta z)=\int_I(\delta x(t)^TQ(t)\delta x(t)+\delta u(t)^TR(t)\delta u(t)+2\delta x(t)^TS(t)\delta u(t))\,dt,$$

$$\delta z=(\delta x, \delta u).$$

The set T of Lemma 10 is given by

$$T=\{z=(x, u)\in Z\colon F'(z^*)(z-z^*)=0,\ x(0)=a,\ u(t)\in U \text{ a.e. } t\in I\}.$$

To verify identity (87), first note that for $z\in\Lambda$, we have

$$C(z)=\int_I g(z)+F(z)^T\lambda^*\,dt.$$

Hence, after an integration by parts and a Taylor expansion, we obtain (87). By the minimum principle (70), $M(z)\geqq b(z-z^*, z-z^*)$. By the coercivity condition (88) and the fact that coercivity with respect to the control implies coercivity with respect to the state (see [20]), we have

$$b(z-z^*, z-z^*)\geqq \alpha|||z-z^*|||^2 \quad \text{for some } \alpha>0.$$

Thus (86) follows from Lemma 10 if, for each $z=(x, u)\in\Lambda$, we can establish the existence of $y\in T$ with $|||z-y|||=o(|||z-z^*|||)$. We construct y in the following way: Let w be the solution to

$$\dot{w}=\nabla_x f(x^*, u^*)w-F'(z^*)(z-z^*), \qquad w(0)=0, \tag{89}$$

and define $y=(w+x, u)$. Observe that $y\in T$. Also, by (89) we have

$$\|w\|_{L^2}\leqq c\|F'(z^*)(z-z^*)\|_{L^2},$$

where c is a generic constant. From the relation

$$\|F'(z^*)(z-z^*)\|_{L^2}=\|F(z)-F(z^*)-F'(z^*)(z-z^*)\|_{L^2}=o(\|z-z^*\|_{L^2}),$$

we conclude that $\|z-y\|_{L^2}=\|w\|_{L^2}=o(\|z-z^*\|_{L^2})$, which completes the proof. □

Now let us consider the finite-dimensional optimization problem

$$\begin{aligned}&\text{minimize} \quad C(x, u)\\ &\text{subject to} \quad F(x, u)=0, \qquad u\in\Omega\subset R^m, \quad x\in R^n,\end{aligned} \tag{90}$$

where Ω is convex and F maps R^{m+n} to R^n. Let z denote the pair (x, u), and for $\lambda\in R^n$, let H be the Lagrangian defined by

$$H(z, \lambda)=C(z)+\lambda^T F(z).$$

COROLLARY 6. *Suppose that x^* and u^* are feasible for* (90), *that F and C are twice differentiable at $z^*=(x^*, u^*)$, and that $\nabla_x F(x^*, u^*)$ is nonsingular. If there exists a multiplier $\lambda^*\in R^n$ such that*

$$\nabla_x H(x^*, u^*, \lambda^*)=0 \quad \textit{and} \quad \nabla_u H(x^*, u^*, \lambda^*)^T(u-u^*)\geqq 0 \quad \textit{for every } u\in\Omega$$

and

$$(z-z^*)^T\nabla^2_{zz}H(z^*, \lambda^*)(z-z^*)\geqq \alpha|z-z^*|^2$$

whenever $F'(z^*)(z-z^*)=0$ for some $z=(x, u)$ with $u\in\Omega$, then z^* is a local minimizer for (90).

Proof. We apply Lemma 9 with

$$M(z)=\nabla_z H(z^*, \lambda)(z-z^*)+\tfrac{1}{2}(z-z^*)^T\nabla^2_{zz}H(z^*, \lambda^*)(z-z^*).$$

The set T of Lemma 10 is given by

$$T=\{z=(x, u)\in R^{m+n}\colon F'(z^*)(z-z^*)=0,\ u\in\Omega\}.$$

The y of Lemma 10 is constructed in the following way: Given $z=(x, u)$ with $u\in\Omega$, $y=(x+w, u)$, where w is the solution to

$$\nabla_x F(x^*, u^*)w = -F'(z^*)(z-z^*).$$

Observe that $y\in T$. If $F(z)=0$, then

$$|F'(z^*)(z-z^*)| = |F(z)-F(z^*)-F'(z^*)(z-z^*)| = \mathrm{o}(|z-z^*|).$$

Since $\nabla_x F(x^*, u^*)$ is nonsingular, we have

$$|z-y| = |w| \leqq c|F'(z^*)(z-z^*)| = \mathrm{o}(|z-z^*|). \qquad \square$$

Appendix 2: The coercivity condition. Here we show that the discrete coercivity condition (75) of § 7 can be deduced from an analogous continuous condition.

LEMMA 11. *Suppose that the matrices A, B, Q, R, and S of § 7 are continuous and that there exists $\beta>0$ such that*

$$\int_I (x(t)^T Q(t)x(t)+u(t)^T R(t)u(t)+2x(t)^T S(t)u(t))\,dt \geqq \beta\int_I |u(t)|^2\,dt \tag{91}$$

whenever $x\in W^{1,2}$, $x(0)=0$, $u\in L^2$, $\dot{x}=Ax+Bu$, $u=v-w$ for some v and $w\in L^2$ with $v(t)$ and $w(t)\in U$ for almost every $t\in I$. Then there exists $\alpha>0$ satisfying the discrete coercivity condition (75).

Proof. Given sequences $\{x_i\}$ and $\{u_i\}$ that satisfy the linear equation (76) where $u_i=v_i-w_i$ for some v_i and $w_i\in U$, let u^h denote the piecewise constant extension of the u_i defined by

$$u^h(t)=u_i, \qquad t_i\leqq t<t_{i+1}, \qquad i=0, 1, \dots, N-1,$$

and let x^h be the solution of

$$\dot{x}^h = Ax^h+Bu^h, \qquad x^h(0)=0.$$

Define $y_i=x^h(t_i)$ and let y^h be the piecewise constant extension of the y_i. Since u^h is piecewise constant,

$$\int_I |u^h(t)|^2\,dt = h\sum_{i=0}^{N-1}|u_i|^2.$$

We will show that, for $x=x^h$ and $u=u^h$,

$$|\text{left side of (91)} - \text{left side of (75)}| \leqq h\varepsilon^h \sum_{i=0}^{N-1}|u_i|^2, \tag{92}$$

where ε^h denotes a generic constant that tends to zero as h tends to zero. Hence, (75) follows from (91) when h is sufficiently small.

Let us begin with the quadratic control terms in (75) and (91). Since u^h is equal to u_i on the interval $[t_i, t_{i+1}]$, it follows that

$$h\sum_{i=0}^{N-1}u_i^T R_i u_i - \int_I u^h(t)^T R(t)u^h(t)\,dt = h\sum_{i=0}^{N-1}u_i^T\delta R_i u_i,$$

where

$$\delta R_i = R_i - \frac{1}{h}\int_{t_i}^{t_{i+1}} R(t)\,dt. \tag{93}$$

Since $R(t)$ is continuous in t, (93) approaches zero, uniformly in i, as $N \to \infty$. Hence, we have

$$\left| h \sum_{i=0}^{N-1} u_i^T R_i u_i - \int_I u^h(t)^T R(t) u^h(t)\, dt \right| \leqq h\varepsilon^h \sum_{i=0}^{N-1} |u_i|^2.$$

Now let us consider the quadratic state terms in (75) and (91). As with the quadratic control term, we have

$$\left| h \sum_{i=0}^{N-1} y_i^T Q_i y_i - \int_I y^h(t)^T Q(t) y^h(t)\, dt \right| \leqq h\varepsilon^h \sum_{i=0}^{N-1} |y_i|^2. \tag{94}$$

From the differential equation satisfied by x^h, we have

$$\|y^h\|_{L^2}^2 \leqq \|y^h\|_{L^\infty}^2 \leqq \|x^h\|_{L^\infty}^2 \leqq c\|u^h\|_{L^2}^2 = ch \sum_{i=0}^{N-1} |u_i|^2, \tag{95}$$

where c denotes a generic constant that is independent of h for h sufficiently small. Combining (94) and (95) yields

$$\left| h \sum_{i=0}^{N-1} y_i^T Q_i y_i - \int_I y^h(t)^T Q(t) y^h(t)\, dt \right| \leqq h\varepsilon^h \sum_{i=0}^{N-1} |u_i|^2. \tag{96}$$

Since y^h is the piecewise constant extension of x^h, it follows from the equation for x^h that

$$\|y^h - x^h\|_{L^2} \leqq h\|\dot{x}^h\|_{L^2} \leqq ch\|u^h\|_{L^2}.$$

This estimate, along with (95), implies that

$$\left| \int_I y^h(t)^T Q(t) y^h(t) - x^h(t)^T Q(t) x^h(t)\, dt \right| \leqq ch\|u^h\|_{L^2}^2 = h\varepsilon^h \sum_{i=0}^{N-1} |u_i|^2. \tag{97}$$

Finally, let us consider the difference

$$\sum_{i=0}^{N-1} y_i^T Q_i y_i - x_i^T Q_i x_i.$$

Integrating the differential equation for x^h over the interval $[t_i, t_{i+1}]$ gives

$$y_{i+1} = y_i + hA_i y_i + hB_i u_i + e_i, \tag{98}$$

where

$$e_i = -h\delta B_i u_i - h\delta A_i y_i + \int_{t_i}^{t_{i+1}} A(t)(x^h(t) - y_i)\, dt.$$

The factors δA_i and δB_i are defined by

$$\delta A_i = A_i - \frac{1}{h}\int_{t_i}^{t_{i+1}} A(t)\, dt \quad \text{and} \quad \delta B_i = B_i - \frac{1}{h}\int_{t_i}^{t_{i+1}} B(t)\, dt.$$

Subtracting the finite difference equation (76) from (98) gives

$$|y_j - x_j| \leqq c \sum_{i=0}^{N-1} |e_i|.$$

From the definition of e_i, we have

$$|e_i| \leqq h\varepsilon^h(|u_i| + |y_i|).$$

It follows that

$$|y_j - x_j| \leqq h\varepsilon^h \sum_{i=0}^{N-1} |u_i| + |y_i| \leqq \varepsilon^h \|u^h\|_{L^2}.$$

Summing over j yields

$$h \sum_{j=0}^{N-1} |y_j - x_j|^2 \leqq \varepsilon^h \|u^h\|_{L^2}^2.$$

Hence, we have

$$\left| h \sum_{i=0}^{N-1} y_i^T Q_i y_i - x_i^T Q_i x_i \right| \leqq \varepsilon^h \|u^h\|_{L^2}^2. \tag{99}$$

The triangle inequality, along with (96), (97), and (99), gives

$$\left| h \sum_{i=0}^{N-1} x_i^T Q_i x_i - \int_I x^h(t)^T Q(t) x^h(t)\, dt \right| \leqq h\varepsilon^h \sum_{i=0}^{N-1} |u_i|^2.$$

Since the cross-product term $x_i^T S_i u_i$ can be analyzed in a similar manner, the proof of (92) is complete. □

REFERENCES

[1] W. ALT, *Stability of solutions for a class of nonlinear cone constrained optimization problems, Part* 1: *Basic theory*, Numer. Funct. Anal. Optim., 10 (1989), pp. 1053-1064.
[2] ———, *Stability of solutions to control constrained nonlinear optimal control problems*, Appl. Math. Optim., 21 (1990), pp. 53-68.
[3] ———, *Parametric optimization with applications to optimal control and sequential quadratic programming*, Bayreuth. Math. Schr., 35 (1991), pp. 1-37.
[4] J.-P. AUBIN, *Lipschitz behavior of solutions to convex minimization problems*, Math. Oper. Res., 9 (1984), pp. 87-111.
[5] J.-P. AUBIN AND I. EKELAND, *Applied Nonlinear Analysis*, Wiley-Interscience, New York, 1984.
[6] J.-P. AUBIN AND H. FRANKOWSKA, *On inverse function theorems for set-valued maps*, J. Math. Pures Appl., 66 (1987), pp. 71-89.
[7] ———, *Set-Valued Analysis*, Birkhäuser, Boston, 1990.
[8] B. M. BUDAK, E. M. BERKOVICH, AND E. N. SOLO'EVA, *The convergence of difference approximations in optimal control problems*, Zh. Vychisl. Mat. i Mat. Fiz., 9 (1969), pp. 522-547.
[9] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, John Wiley, New York, 1983.
[10] R. COMINETTI, *Metric regularity, tangent sets, and second-order optimality conditions*, Appl. Math. Optim., 21 (1990), pp. 265-287.
[11] J. CULLUM, *Finite-dimensional approximations of state-constrained continuous optimal control problems*, SIAM J. Control, 10 (1972), pp. 649-670.
[12] A. L. DONTCHEV, *Error estimates for a discrete approximation to constrained control problems*, SIAM J. Numer. Anal., 13 (1981), pp. 500-514.
[13] ———, *Perturbations, Approximations and Sensitivity Analysis of Optimal Control Systems*, Lecture Notes in Control and Inform. Sci., Vol. 52, Springer-Verlag, New York, 1983.
[14] ———, *On the admissible controls of constrained linear systems*, C. R. Acad. Bulgare Sci., 42 (1989), pp. 33-36.
[15] J. C. DUNN AND T. TIAN, *Variants of the Kuhn–Tucker sufficient conditions in cones of nonnegative functions*, SIAM J. Control Optim., 30 (1992), pp. 1361-1384.
[16] H. FRANKOWSKA, *Some inverse mapping theorems*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 7 (1990), pp. 183-234.
[17] W. W. HAGER, *Rate of convergence for discrete approximations to unconstrained control problems*, SIAM J. Numer. Anal., 13 (1976), pp. 449-471.
[18] ———, *Lipschitz continuity for constrained processes*, SIAM J. Control Optim., 17 (1979), pp. 321-338.
[19] ———, *Approximations to the multiplier method*, SIAM J. Numer. Anal., 22 (1985), pp. 16-46.
[20] ———, *Multiplier methods for nonlinear optimal control*, SIAM J. Numer. Anal., 27 (1990), pp. 1061-1080.

[21] A. D. IOFFE, *Necessary and sufficient conditions for a local minimum. 3: Second order conditions and augmented duality*, SIAM J. Control Optim., 17 (1979), pp. 266–288.

[22] A. D. IOFFE AND V. M. TIHOMIROV, *Theory of Extremal Problems*, North-Holland, Amsterdam, 1979.

[23] K. ITO AND K. KUNISCH, *Sensitivity analysis of solutions to optimization problems in Hilbert spaces with applications to optimal control and estimation*, J. Differential Equations, 99 (1992), pp. 1–40.

[24] A. J. KING AND R. T. ROCKAFELLAR, *Sensitivity analysis for nonsmooth generalized equations*, Math. Programming, 55 (1992), pp. 192–212.

[25] S. KURCYUSZ, *On the existence and nonexistence of Lagrange multipliers in Banach spaces*, J. Optim. Theory Appl., 20 (1976), pp. 81–110.

[26] J. L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*, S. K. Mitter, trans., Springer-Verlag, Berlin, New York, 1971.

[27] K. MALANOWSKI, *On stability of solutions to constrained optimal control problems for systems with control appearing linearly*, Arch. Automat. Telemech., 33 (1988), pp. 483–497.

[28] ———, *Second order conditions and constraint qualifications in stability and sensitivity analysis of solutions to optimizations problems in Hilbert spaces*, Appl. Math. Optim., 25 (1992), pp. 51–79.

[29] H. MAURER, *First and second-order sufficient optimality conditions in mathematical programming and optimal control*, Math. Programming Study, 14 (1981), pp. 163–177.

[30] H. MAURER AND J. ZOWE, *First- and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, Math. Programming, 16 (1979), pp. 98–110.

[31] B. S. MORDUKHOVICH, *On difference approximations of optimal control systems*, Appl. Math. Mech., 42 (1978), pp. 452–461.

[32] ———, *Sensitivity analysis in nonsmooth optimization*, in Theoretical Aspects of Industrial Design, SIAM, Philadelphia, PA, 1992, pp. 32–46.

[33] J.-P. PENOT, *Metric regularity, openness, and Lipschitz multifunctions*, Nonlinear Anal. Theory Methods Appl., 13 (1989), pp. 629–643.

[34] S. M. ROBINSON, *An inverse-function theorem for a class of multivalued functions*, Proc. Amer. Math. Soc., 41 (1973), pp. 211–218.

[35] ———, *Stability theory for systems of inequalities, Part I: Linear systems*, SIAM J. Numer. Anal., 12 (1976), pp. 754–772.

[36] ———, *Stability theory for systems of inequalities, Part II: Differentiable nonlinear systems*, SIAM J. Numer. Anal., 13 (1976), pp. 497–513.

[37] ———, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp, 43–62.

[38] ———, *An implicit-function theorem for a class of nonsmooth functions*, Math. Oper. Res., 16 (1991), 292–309.

[39] R. T. ROCKAFELLAR, *Lipschitzian properties of multifunctions*, Nonlinear Anal., 9 (1985), pp. 867–885.

[40] ———, *Lipschitzian stability in optimization: The role of nonsmooth analysis*, in Nondifferentiable Optimization: Motivation and Applications, V. F. Demyanov and D. Pallaschke, eds., Lecture Notes in Economics and Mathematical Systems, Vol. 255, Springer-Verlag, New York, 1985, pp. 55–73.

[41] ———, *Proto-differentiability of set-valued mappings and its application in optimization*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 6 (1989), pp. 449–482.

[42] B. SENDOV AND V. A. POPOV, *The Averaged Moduli of Smoothness*, John Wiley, New York, 1988.

[43] C. URSESCU, *Multifunctions with closed convex graph*, Czechoslovak Math. J., 25 (1975), pp. 438–441.