SIAM J. CONTROL OPTIM.
Vol. 36, No. 2, pp. 698–718, March 1998

011

LIPSCHITZIAN STABILITY FOR STATE CONSTRAINED NONLINEAR OPTIMAL CONTROL*

A. L. DONTCHEV† AND W. W. HAGER‡

Abstract. For a nonlinear optimal control problem with state constraints, we give conditions under which the optimal control depends Lipschitz continuously in the L^2 norm on a parameter. These conditions involve smoothness of the problem data, uniform independence of active constraint gradients, and a coercivity condition for the integral functional. Under these same conditions, we obtain a new nonoptimal stability result for the optimal control in the L^∞ norm. And under an additional assumption concerning the regularity of the state constraints, a new tight L^∞ estimate is obtained. Our approach is based on an abstract implicit function theorem in nonlinear spaces.

Key words. optimal control, state constraints, Lipschitzian stability, implicit function theorem

AMS subject classifications. 49K40, 90C31

PII. S0363012996299314

1. Introduction.

We consider the following optimal control problem involving a parameter:

$$\text{minimize} \int_0^1 h_p(x(t), u(t))dt \tag{1}$$

subject to

$$\dot{x}(t) = f_p(x(t), u(t)) \text{ a.e. } t \in [0, 1], \quad x(0) = x^0,$$
$$g_p(x(t)) \le 0 \text{ for all } t \in [0, 1], \; u \in L^\infty, \; x \in W^{1,\infty},$$

where the state $x(t) \in \mathbf{R}^n$, $\dot{x} \equiv \frac{d}{dt}x$, the control $u(t) \in \mathbf{R}^m$, the parameter p lies in a metric space, the functions $h_p : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$, $f_p : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^n$, and $g_p : \mathbf{R}^n \to \mathbf{R}^k$. Throughout the paper, $L^\alpha(J; \mathbf{R}^m)$ denotes the usual Lebesgue space of functions $u : J \to \mathbf{R}^m$ with $|u(\cdot)|^\alpha$ integrable, equipped with its standard norm

$$\| u \|_{L^\alpha} = \left(\int_J |u(t)|^\alpha dt \right)^{1/\alpha},$$

where $| \cdot |$ is the Euclidean norm. Of course, $\alpha = \infty$ corresponds to the space of essentially bounded functions. Let $W^{m,\alpha}(J; \mathbf{R}^n)$ be the usual Sobolev space consisting of vector-valued functions whose jth derivative lies in L^α for all $0 \le j \le m$; its norm is

$$\| u \|_{W^{m,\alpha}} = \sum_{j=0}^{m} \| u^{(j)} \|_{L^\alpha} .$$

When either the domain J or the range $\mathbf{R}^n$ is clear from context, it is omitted. We let H^m denote the space $W^{m,2}$, and Lip denote $W^{1,\infty}$, the space of Lipschitz continuous

*Received by the editors February 26, 1996; accepted for publication (in revised form) January 21, 1997. This research was supported by National Science Foundation grant DMS 9404431.
http://www.siam.org/journals/sicon/36-2/29931.html

†Mathematical Reviews, University of Michigan, Ann Arbor, MI 48107 (ald@math.ams.org).
‡Department of Mathematics, University of Florida, Gainesville, FL 32611-8105 (hager@math.ufl.edu).

functions. Subscripts on spaces are used to indicate bounds on norms; in particular, $W^{m,\alpha}_\kappa$ denotes the set of functions in $W^{m,\alpha}$ with the property that the L^α norm of the mth derivative is bounded by κ, and Lip_κ denotes the space of Lipschitz continuous functions with Lipschitz constant κ. Throughout, c is a generic constant, independent of the parameter p and time t, and $B_a(x)$ is the closed ball centered at x with radius a. The L^2 inner product is denoted $\langle \cdot, \cdot \rangle$, the complement of a set $\mathcal{A}$ is $\mathcal{A}^c$, and the transpose of a matrix B is B^T. Given a vector $y \in \mathbf{R}^m$ and a set $\mathcal{A} \subset \{1, 2, \ldots, m\}$, $y_\mathcal{A}$ denotes the subvector consisting of components associated with indices in $\mathcal{A}$. And if $Y \in \mathbf{R}^{m\times n}$, then $Y_\mathcal{A}$ is the submatrix consisting of rows associated with indices in $\mathcal{A}$.

We wish to study how a solution to either (1) or the associated variational system representing the first-order necessary condition depends on the parameter p. We assume that the problem (1) has a local minimizer $(x, u) = (x_*, u_*)$ corresponding to a reference value $p = p_*$ of the parameter, and the following smoothness condition holds.

Smoothness. The local minimizer (x_*, u_*) of (1) lies in $W^{2,\infty} \times$ Lip. There exists a closed set $\Delta \subset \mathbf{R}^n \times \mathbf{R}^m$ and a $\delta > 0$ such that $B_\delta(x_*(t), u_*(t)) \subset \Delta$ for every $t \in [0, 1]$. The function values and first two derivatives of $f_p(x, u)$, $g_p(x, u)$, and $h_p(x, u)$, and the third derivatives of $g_p(x)$, with respect to x and u, are uniformly continuous relative to p near p_* and $(x, u) \in \Delta$. And when either the first two derivatives of $f_p(x, u)$ and $h_p(x, u)$ or the first three derivatives of $g_p(x)$, with respect to x and u, are evaluated at (x_*, u_*), the resulting expression is differentiable in t, and the L^∞ norm of the time derivative is uniformly bounded relative to p near p_*.

Let A, B, and K be the matrices defined by

$$A = \nabla_x f_*(x_*, u_*), \quad B = \nabla_u f_*(x_*, u_*), \quad \text{and } K = \nabla_x g_*(x_*).$$

Here and elsewhere the * subscript is always associated with p_*. Let $\mathcal{A}(t)$ be the set of indices of the active constraints at $(x_*(t), p_*)$; that is,

$$\mathcal{A}(t) = \{i \in \{1, 2, \cdots, k\} : g_*(x_*(t))_i = 0\}.$$

We introduce the following assumption.

Uniform independence at $\mathcal{A}$. The set $\mathcal{A}(0)$ is empty and there exists a scalar $\alpha > 0$ such that

$$\left| \sum_{i \in \mathcal{A}(t)} v_i K_i(t) B(t) \right| \geq \alpha |v_{\mathcal{A}(t)}|$$

for each $t \in [0, 1]$ where $\mathcal{A}(t) \neq \emptyset$ and for each choice of v.

Uniform independence implies that the state constraints are first-order (see [12] for the definition of the order of a state constraint). This condition can be generalized to higher order state constraints (see Maurer [17]), however, the generalization of the stability results in this paper to higher order state constraints is not immediate.

It is known (see, for instance, Theorem 7.1 of the recent survey [12] and the regularity analysis in [8]) that under appropriate assumptions, the first-order necessary conditions (Pontryagin's minimum principle) associated with a solution (x_*, u_*) of (1) can be written in the following way. There exist $\psi_* \in W^{2,\infty}$ and $\nu_* \in$ Lip such that $x = x_*$, $\psi = \psi_*$, $u = u_*$, and $\nu = \nu_*$ are a solution at $p = p_*$ of the variational system:

$$\dot{x} = f_p(x, u), \; x(0) = x^0, \tag{2}$$

$$(3)\qquad \dot{\psi} = -\nabla_x H_p(x, \psi, u, \nu),\ \ \psi(1) = 0,$$

$$(4)\qquad 0 = \nabla_u H_p(x, \psi, u, \nu),$$

$$(5)\qquad g_p(x) \in N(\nu),\ \ \nu(1) \le 0,\ \ \dot{\nu} \ge 0 \ \text{ a.e.}$$

Here H_p is the Hamiltonian defined by

$$H_p(x, \psi, u, \nu) = h_p(x, u) + \psi^{\mathsf{T}}\ f_p(x, u) - \nu^{\mathsf{T}} \nabla g_p(x) f_p(x, u),$$

and the set-valued map N is defined in the following way: given a nondecreasing Lipschitz continuous function ν, a continuous function y lies in $N(\nu)$ if and only if

$$y(t) \le 0,\ \dot{\nu}(t)^{\mathsf{T}} y(t) = 0 \text{ for a.e. } t \in [0, 1],\ \text{ and } \ \nu(1)^{\mathsf{T}} y(1) = 0.$$

Defining

$$Q = \nabla_{xx} H_*(w_*),\ \ M = \nabla_{xu} H_*(w_*),\ \text{and } R = \nabla_{uu} H_*(w_*),$$

where $w_* = (x_*, \psi_*, u_*, \nu_*)$, let $\mathcal{B}$ be the quadratic form

$$\mathcal{B}(x, u) = \frac{1}{2} \int_0^1 x(t)^{\mathsf{T}} Q(t) x(t) + u(t)^{\mathsf{T}} R(t) u(t) + 2x(t)^{\mathsf{T}} M(t) u(t) dt,$$

and let L be the linear and continuous operator from $H^1 \times L^2$ to L^2 defined by $L(x, u) = \dot{x} - Ax - Bu$. We introduce the following growth assumption for the quadratic form.

Coercivity. There exists a constant $\alpha > 0$ such that

$$\mathcal{B}(x, u) \ \ge\ \alpha \langle u, u \rangle \quad \text{for all } \ (x, u) \in \mathcal{M},$$

where

$$(6)\qquad \mathcal{M} = \{(x, u) : x \in H^1,\ u \in L^2, L(x, u) = 0,\ x(0) = 0\}.$$

In the terminology of [12], the form of the minimum principle we employ is the "indirect adjoining approach with continuous adjoint function." A different approach, found in [13], for example, involves a different choice for the multipliers and for the Hamiltonian. The multipliers in these two approaches are related in a linear fashion as shown in [11]. Normally, the multiplier ν, associated with the state constraint, and the derivative of ψ have bounded variation. In our statement of the minimum principle above, we are implicitly assuming some additional regularity so that ν and $\dot{\psi}$ are not only of bounded variation, but Lipschitz continuous. This regularity can be proved under the uniform independence and coercivity conditions (see [8]).

In section 3 we establish the following result.

THEOREM 1.1. *Suppose that the problem* (1) *with* $p = p_*$ *has a local minimizer* (x_*, u_*) *and that the smoothness and the uniform independence conditions hold. Let* ψ_* *and* ν_* *be the associated multipliers satisfying the variational system* (2)–(5) *with* $\psi_* \in W^{2,\infty}$ *and* $\nu_* \in Lip$*. If the coercivity condition holds, then there exist a constant* μ *and neighborhoods* V *of* p_* *and* U *of* $w_* = (x_*, \psi_*, u_*, \nu_*)$ *in* $W^{1,\infty} \times W^{1,\infty} \times L^\infty \times L^\infty$*, such that for every* $p \in V$*, there is a unique solution* $w = (x, \psi, u, \nu) \in U$ *to the first-order necessary conditions* (2)–(5) *with the property that* $(\dot{x}, \dot{\psi}, u, \nu) \in \mathrm{Lip}_\mu$ *and* (x, u) *is a local minimizer of the problem* (1) *associated with* p*. Moreover, for every* $p_i \in V, i = 1, 2,$ *if* $w_i = (x_i, \psi_i, u_i, \nu_i)$ *is the corresponding solution of* (2)–(5)*, the following estimate holds:*

$$(7)\qquad \|x_1 - x_2\|_{H^1} + \|\psi_1 - \psi_2\|_{H^1} + \|u_1 - u_2\|_{L^2} + \|\nu_1 - \nu_2\|_{L^2} \le cE_2,$$

where

$$E_\alpha = \|f_{p_1}(x_1, u_1) - f_{p_2}(x_1, u_1)\|_{L^\alpha} + \|\nabla_x H_{p_1}(w_1) - \nabla_x H_{p_2}(w_1)\|_{L^\alpha} + \|\nabla_u H_{p_1}(w_1) - \nabla_u H_{p_2}(w_1)\|_{L^\alpha} + \|g_{p_1}(x_1) - g_{p_2}(x_1)\|_{W^{1,\alpha}}.$$

In addition, we have

$$\|x_1 - x_2\|_{W^{1,\infty}} + \|\psi_1 - \psi_2\|_{W^{1,\infty}} + \|u_1 - u_2\|_{L^\infty} + \|\nu_1 - \nu_2\|_{L^\infty} \le cE_2^{2/3}.$$

The proof of Theorem 1.1 is based on an abstract implicit function theorem appearing in section 2. In section 4 we show that the L^∞ estimate of Theorem 1.1 can be sharpened if the points where the state constraints change between active and inactive are separated. In section 5 we comment briefly on related work.

2. An implicit function theorem in nonlinear spaces.

The following lemma provides a generalization of the implicit function theorem that can be applied to nonlinear spaces. To simplify the notation, we let $\|x - y\|_X$ denote the distance between the elements x and y of the metric space X.

LEMMA 2.1. *Let X and Π be metric spaces with X complete, let Y be a subset of Π, and let P be a set. Given $w_* \in X$ and $r > 0$, let W denote the ball $B_r(w_*)$ in X and suppose that $T : W \times P \to Y$ and $F : X \to 2^\Pi$ (the subsets of Π) have the following properties.*

(P1) *$T(w_*, p_*) \in F(w_*)$ for some $p_* \in P$.*
(P2) *For some $\beta > 0$, $\|T(w_*, p_*) - T(w_*, p)\|_\Pi \le \beta$ for all $p \in P$.*
(P3) *For some $\epsilon > 0$, $\|T(w_1, p) - T(w_2, p)\|_\Pi \le \epsilon\|w_1 - w_2\|_X$ for all $w_1, w_2 \in W$ and $p \in P$.*
(P4) *F^{-1} restricted to Y is single-valued and Lipschitz continuous, with Lipschitz constant λ.*

If $\epsilon\lambda < 1$ and $r \ge \lambda\beta/(1 - \epsilon\lambda)$, then for each $p \in P$, there exists a unique $w \in W$ such that $T(w, p) \in F(w)$. Moreover, for every $p_i \in P, i = 1, 2$, if w_i denotes the w associated with p_i, then we have

$$\|w_1 - w_2\|_X \le \frac{\lambda}{1 - \lambda\epsilon}\|T(w_1, p_1) - T(w_1, p_2)\|_\Pi. \tag{8}$$

Proof. Fix $p \in P$ and define $\Phi(w) = F^{-1}(T(w, p))$ for $w \in W$. Observe that

$$\begin{aligned}\|\Phi(w_1) - \Phi(w_2)\|_X &= \|F^{-1}(T(w_1, p)) - F^{-1}(T(w_2, p))\|_X \\ &\le \lambda\|T(w_1, p) - T(w_2, p)\|_\Pi \le \lambda\epsilon\|w_1 - w_2\|_X\end{aligned}$$

for each $w_1, w_2 \in W$. Since $\lambda\epsilon < 1$, Φ is a contraction on W with contraction constant $\lambda\epsilon$. Let $w \in W$. Since $w_* = F^{-1}(T(w_*, p_*))$ and $r \ge \lambda\beta/(1 - \epsilon\lambda)$, we have

$$\begin{aligned}\|w_* - \Phi(w)\|_X &= \|F^{-1}(T(w_*, p_*)) - F^{-1}(T(w, p))\|_X \\ &\le \lambda(\|T(w, p) - T(w_*, p)\|_\Pi + \|T(w_*, p) - T(w_*, p_*)\|_\Pi) \\ &\le \lambda(\epsilon r + \beta) \le r.\end{aligned}$$

Thus Φ maps W into itself. By the Banach contraction mapping principle, there exists a unique $w \in W$ such that $w = \Phi(w)$. Since $w = \Phi(w)$ is equivalent to $T(w, p) \in F(w)$ for $w \in W$, we conclude that for each $p \in P$, there is a unique $w(p) \in W$ such that

$T(w(p), p) \in F(w(p))$. Defining $w_1 = w(p_1)$ and $w_2 = w(p_2)$, we have

$$\begin{aligned} \|w_1 - w_2\|_X &= \|F^{-1}(T(w_1, p_1)) - F^{-1}(T(w_2, p_2))\|_X \\ &\le \lambda \|T(w_1, p_1) - T(w_2, p_2)\|_\Pi \\ &\le \lambda \|T(w_1, p_1) - T(w_1, p_2)\|_\Pi + \lambda \|T(w_1, p_2) - T(w_2, p_2)\|_\Pi \\ &\le \lambda \|T(w_1, p_1) - T(w_1, p_2)\|_\Pi + \lambda\epsilon \|w_1 - w_2\|_X. \end{aligned}$$

Rearranging this inequality, the proof is complete. □

Let X, Y, and $\mathcal{P}$ be metric spaces and let $w_* \in X$. Using the terminology of [3], $f : X \times \mathcal{P} \to Y$ is strictly stationary at $w = w_*$, uniformly in p near p_*, if for each $\epsilon > 0$, there exists $\delta > 0$ with the property that

$$\|f(w_1, p) - f(w_2, p)\|_Y \le \epsilon \|w_1 - w_2\|_X$$

for all $w_1, w_2 \in B_\delta(w_*)$ and $p \in B_\delta(p_*)$.

THEOREM 2.2. *Let X be a complete metric space, let Π be a linear metric space, let Y be a subset of Π, and let $\mathcal{P}$ be a metric space. Suppose that $\mathcal{F} : X \to 2^\Pi$, that $\mathcal{T} : X \times \mathcal{P} \to \Pi$, that $\mathcal{L} : X \to \Pi$ is continuous, and that for some $w_* \in X$ and $p_* \in \mathcal{P}$ we have:*

(Q1) *$\mathcal{T}(w_*, p_*) \in \mathcal{F}(w_*)$;*
(Q2) *$\mathcal{T}(w_*, \cdot\,)$ is continuous at p_*;*
(Q3) *$\mathcal{T}(w, p) - \mathcal{L}(w)$ is strictly stationary at $w = w_*$, uniformly in p near p_*;*
(Q4) *$(\mathcal{F} - \mathcal{L})^{-1}$ restricted to Y is single-valued and Lipschitz continuous, with Lipschitz constant λ;*
(Q5) *$\mathcal{T} - \mathcal{L}$ maps a neighborhood of (w_*, p_*) into Y.*

Then for each $\lambda_+ > \lambda$, there exist neighborhoods W of w_* and P of p_* such that for each $p \in P$, a unique $w \in W$ exists satisfying $\mathcal{T}(w, p) \in \mathcal{F}(w)$; moreover, for every $p_i \in P, i = 1, 2$, if w_i denotes the $w \in W$ associated with p_i, then we have

$$\|w_1 - w_2\|_X \le \lambda_+ \|\mathcal{T}(w_1, p_1) - \mathcal{T}(w_1, p_2)\|_\Pi.$$

Proof. By (Q5) there exist neighborhoods U' of w_* and P' of p_* such that $\mathcal{T}(w, p) - \mathcal{L}(w) \in Y$ for each $w \in U'$ and $p \in P'$. We apply Lemma 2.1 with the following identifications: X, Y, and Π are as defined in the statement of the theorem, $F(w) = \mathcal{F}(w) - \mathcal{L}(w)$, and $T(w, p) = \mathcal{T}(w, p) - \mathcal{L}(w)$. (P1) and (P4) follow immediately from (Q1) and (Q4), respectively. Choose $\epsilon > 0$ such that $\epsilon < (\lambda_+ - \lambda)/(\lambda_+ \lambda)$. Since $\lambda_+ > \lambda$, it follows that for this choice of ϵ, we have $\epsilon\lambda < 1$ and $\lambda/(1 - \lambda\epsilon) < \lambda_+$. By (Q3) and the identity $T(w_1, p_1) - T(w_1, p_2) = \mathcal{T}(w_1, p_1) - \mathcal{T}(w_1, p_2)$, there exist neighborhoods $P = B_r(p_*) \subset P'$ of p_* and $W = B_r(w_*) \subset U'$ of w_* such that (P3) of Lemma 2.1 holds. Let β satisfy $\lambda\beta/(1 - \epsilon\lambda) \le r$, and by (Q2), choose P smaller if necessary so that (P2) holds. By Lemma 2.1, for each $p \in P$, there exists a unique $w \in W$ such that $T(w, p) \in F(w)$, and the estimate (8) holds. Since $T(w, p) \in F(w)$ if and only if $\mathcal{T}(w, p) \in \mathcal{F}(w)$, the proof is complete. □

A particular case of Theorem 2.2 corresponds to the well-known Robinson implicit function theorem [20] in which X is a Banach space, Π is its dual X^*, $\mathcal{F}(w) = N_\Omega(w)$, Ω is a closed, convex set in X, $N_\Omega(w)$ is the normal cone to the set Ω at the point w, $\mathcal{T}$ is differentiable with respect to w, both $\mathcal{T}$ and its derivative $\nabla_w \mathcal{T}$ are continuous in a neighborhood of (w_*, p_*), and $\mathcal{L}(w) = \mathcal{T}(w_*, p_*) + \nabla_w \mathcal{T}(w_*, p_*)(w - w_*)$ is the linearization of $\mathcal{T}$. The Robinson framework is applicable to control problems with control constraints after the range space X^* is replaced by a general Banach space Y (see the discussion in section 5). However, for problems with state constraints, there are difficulties in applying Robinson's theory since stability results for state constrained quadratic problems, analogous to the results for control constrained problems, have not been established. In our previous paper [3], we extend Robinson's work in several different directions. For the solution map of a generalized equation in a linear metric space, we showed that Aubin's pseudo-Lipschitz property, that the existence of a Lipschitzian selection, and that local Lipschitzian invertibility are "robust" under nonlinear perturbations that are strictly stationary at the reference point. In Theorem 2.2, we focus on the latter property, giving an extension of our earlier result to nonlinear spaces. In this nonlinear setting, we are able to analyze the state constrained problem, obtaining a Lipschitzian stability result for the solution.

3. Lipschitzian stability in L^2. To prove Theorem 1.1, we apply Theorem 2.2 using the following identifications. First, we define

$$w = (x, \psi, u, \nu), \tag{9}$$

where

$$x,\ \psi \in W^{2,\infty}_{\mu} \text{ (with the } H^1 \text{ norm)},\ x(0) = x^0,\ \psi(1) = 0, \tag{10}$$

$$u,\ \nu \in \mathrm{Lip}_{\mu} \text{ (with the } L^2 \text{ norm)},\ \nu(1) \le 0 \text{ and } \dot{\nu} \ge 0 \text{ a.e.} \tag{11}$$

An appropriate value for μ is chosen later in the analysis. The space X consists of the collection of functions x, ψ, u, and ν satisfying (10) and (11) with the norm defined in (10) and (11). Observe that the norms we use are not the natural norms. For example, the u and ν components of elements in X lie in $W^{1,\infty}$, but we use the L^2 norm to measure distance. Despite the apparent mismatch of space and norm, X is complete by Lemma 3.2 below.

The functions $\mathcal{T}$ and $\mathcal{F}$ of Theorem 2.2 are selected in the following way:

$$\mathcal{T}(w,p) = \begin{pmatrix} \dot{x} - f_p(x,u) \\ \dot{\psi} + \nabla_x H_p(x,u,\psi,\nu) \\ \nabla_u H_p(x,u,\psi,\nu) \\ g_p(x) \end{pmatrix} \text{ and } \mathcal{F}(w) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ N(\nu) \end{pmatrix}. \tag{12}$$

The continuous operator $\mathcal{L}$ is obtained by linearizing the map $\mathcal{T}(\cdot, p_*)$ in L^∞ at the reference point $w_* = (x_*, \psi_*, u_*, \nu_*)$. In particular,

$$\mathcal{L}(w) = \begin{pmatrix} \dot{x} - Ax - Bu \\ \dot{\psi} + A^{\mathsf{T}}\psi + Qx + Mu - (\dot{K}^{\mathsf{T}} + A^{\mathsf{T}}K^{\mathsf{T}})\nu \\ Ru + M^{\mathsf{T}}x + B^{\mathsf{T}}\psi - B^{\mathsf{T}}K^{\mathsf{T}}\nu \\ Kx \end{pmatrix}. \tag{13}$$

Defining $\pi_* = \mathcal{T}(w_*, p_*) - \mathcal{L}(w_*)$, let a_*, s_*, r_*, and b_* denote the components of π_*:

$$\begin{aligned} a_* &= -f_*(x_*, u_*) + Ax_* + Bu_*, \\ s_* &= \nabla_x H_*(w_*) - A^{\mathsf{T}}\psi_* - Qx_* - Mu_* + (\dot{K}^{\mathsf{T}} + A^{\mathsf{T}}K^{\mathsf{T}})\nu_*, \\ r_* &= \nabla_u H_*(w_*) - Ru_* - M^{\mathsf{T}}x_* - B^{\mathsf{T}}\psi_* + B^{\mathsf{T}}K^{\mathsf{T}}\nu_*, \\ b_* &= g_*(x_*) - Kx_*. \end{aligned}$$

The space Π is the product $L^2 \times L^2 \times L^2 \times H^1$, while the elements π in Y have the form $\pi = (a, s, r, b)$, where

$$\begin{aligned} &a,\ s,\ r\ \in \mathrm{Lip} \text{ (with the } L^2 \text{ norm)},\ \ b \in W^{2,\infty} \text{ (with the } H^1 \text{ norm)}, \\ &\|a - a_*\|_{W^{1,\infty}} + \|r - r_*\|_{W^{1,\infty}} + \|s - s_*\|_{W^{1,\infty}} + \|b - b_*\|_{W^{2,\infty}} \le \kappa, \end{aligned} \tag{14}$$

where κ is a small positive constant chosen so that two related quadratic programs, (37) and (41), introduced later have the same solution. As we will see, the constant μ associated with the space X must be chosen sufficiently large relative to κ. Note that the inverse $(\mathcal{F}-\mathcal{L})^{-1}\pi$ is the solution (x,ψ,u,ν) of the linear variational system:

$$\dot{x} = Ax + Bu - a,\ x(0) = x^0, \tag{15}$$

$$\dot{\psi} = -A^{\mathsf{T}}\psi - Qx - Mu + (\dot{K} + A^{\mathsf{T}}K^{\mathsf{T}})\nu - s,\ \psi(1) = 0, \tag{16}$$

$$0 = Ru + M^{\mathsf{T}}x + B^{\mathsf{T}}\psi - B^{\mathsf{T}}K^{\mathsf{T}}\nu + r, \tag{17}$$

$$Kx + b \in N(\nu),\ \nu(1) \le 0,\ \dot{\nu} \ge 0\ \text{ a.e.} \tag{18}$$

Referring to the assumptions of Theorem 2.2, (Q1) holds by the definition of X, and by the minimum principle, (Q2) follows immediately from the smoothness condition. In Lemma 3.3, we deduce (Q3) from the smoothness condition and a Taylor expansion. In Lemma 3.6, (Q5) is obtained by showing that for w near w_* and p near p_*, $\mathcal{T}(w,p) - \mathcal{L}(w)$ and its associated derivatives are near those of $\pi_* = \mathcal{T}(w_*,p_*) - \mathcal{L}(w_*)$. Finally, in a series of lemmas, (Q4) is established through manipulations of quadratic programs associated with (15)–(18).

To start the analysis, we show that X is complete using the following lemma.

LEMMA 3.1. *If $u \in \mathrm{Lip}_\mu([0,1];\mathbf{R}^1)$, then we have*

$$\|u\|_{L^\infty} \le \max\{\ \sqrt{3}\|u\|_{L^2},\ \sqrt[3]{3\mu}\|u\|_{L^2}^{2/3}\ \}\ .$$

Proof. Since u is continuous, its maximum absolute value is achieved at some time t_m on the interval $[0,1]$. Let $u_m = u(t_m)$ denote the associated value of u. We consider two cases.

Case 1. $u_m > \mu$. Let us examine the maximum ratio between the ∞-norm and the 2-norm:

$$\text{maximize } \{\|u\|_{L^\infty}/\|u\|_{L^2} : \|u\|_{L^\infty} = u_m, u \in \ \mathrm{Lip}_\mu\}.$$

Since $u_m > \mu$, the maximum is attained by the linear function v satisfying $v(0) = u_m$ and $\dot{v} = -\mu$. The 2-norm of this function is readily evaluated:

$$\|v\|_{L^2}^2 = u_m^2(3 - 3\alpha + \alpha^2)/3,\ \text{ where } \alpha = \mu/u_m.$$

Since $\alpha \in [0,1]$ and since $3 - 3\alpha + \alpha^2 \ge 1$ on this interval, we have $\|v\|_{L^2}^2 \ge u_m^2/3$. Taking square roots gives

$$\|v\|_{L^\infty}/\|v\|_{L^2} \le \sqrt{3},$$

which establishes the lemma in Case 1.

Case 2. $u_m \le \mu$. In this case, let us examine the maximum ratio between the ∞-norm and the 2-norm to the 2/3-power:

$$\text{maximize } \{\|u\|_{L^\infty}/\|u\|_{L^2}^{2/3} : \|u\|_{L^\infty} = u_m, u \in \ \mathrm{Lip}_\mu\}.$$

The maximum is attained by the piecewise linear function v satisfying $v(0) = u_m$, $\dot{v} = -\mu$ on $[0, u_m/\mu]$, and $v = 0$ elsewhere. Since

$$\|v\|_{L^2}^2 = \frac{u_m^3}{3\mu},$$

it follows that

$$\|v\|_{L^\infty}/\|v\|_{L^2}^{2/3} \le \sqrt[3]{3\mu},$$

which completes the proof of Case 2. □

LEMMA 3.2. *The space X of functions w satisfying* (9), (10), *and* (11) *is complete.*

Proof. Suppose that $w_k = (x_k, u_k, \psi_k, \nu_k)$ is a Cauchy sequence in X. We analyze the ν-component of w_k. The sequence ν_k is a Cauchy sequence in L^∞ by Lemma 3.1. Since L^∞ is complete, there exists a limit point $\bar{\nu} \in L^\infty$. Since the ν_k converge pointwise to $\bar{\nu}$ and since each of the ν_k is Lipschitz continuous with Lipschitz constant μ, $\bar{\nu}$ is Lipschitz continuous with Lipschitz constant μ. Since each of the ν_k is nondecreasing, it follows from the pointwise convergence that $\bar{\nu}$ is nondecreasing; hence, $\dot{\bar{\nu}} \ge 0$. Since $\nu_k(1) \le 0$ for each k, the pointwise convergence implies that $\bar{\nu}(1) \le 0$. This shows that the ν-component of X is complete. The other components can be analyzed in a similar fashion. □

LEMMA 3.3. *If the smoothness condition holds, then for $\mathcal{T}$ and $\mathcal{L}$ defined in* (12) *and* (13)*, respectively, $\mathcal{T} - \mathcal{L}$ is strictly stationary at w_*, uniformly in p near p_*.*

Proof. Only the first component of $\mathcal{T}(w,p) - \mathcal{L}(w)$ is analyzed, since the other components are treated in a similar manner. To establish strict stationarity for the first component, we need to show that for any given $\epsilon > 0$,

$$\|(f_p(x,u) - f_p(y,v)) - A(x-y) - B(u-v)\|_{L^2} \le \epsilon\|x-y\|_{H^1} + \epsilon\|u-v\|_{L^2}, \tag{19}$$

for p near p_* and for (x,u) and $(y,v) \in W_\mu^{2,\infty} \times \mathrm{Lip}_\mu$ near (x_*, u_*) in the norm of $H^1 \times L^2$, where $A = \nabla_x f_*(x_*, u_*)$ and $B = \nabla_u f_*(x_*, u_*)$. By Lemma 3.1, (x,u) and (y,v) are also near (x_*, u_*) in L^∞. After writing the difference $f_p(x,u) - f_p(y,v)$ as an integral over the line segment connecting (x,u) and (y,v), we have

$$(f_p(x,u) - f_p(y,v)) - A(x-y) - B(u-v) = (A_p - A)(x-y) + (B_p - B)(u-v),$$

where (A_p, B_p) is the average of the gradient of f_p along the line segment connecting (x,u) and (y,v). By the smoothness condition, $\|A_p - A\|_{L^\infty} \to 0$ and $\|B_p - B\|_{L^\infty} \to 0$ as p approaches p_* and as both (x,u) and (y,v) approach (x_*, u_*) in L^∞. This completes the proof. □

LEMMA 3.4. *If the smoothness condition holds, then for $\mathcal{T}$ and $\mathcal{L}$ defined in* (12) *and* (13)*, respectively, and for any choice of the parameter $\kappa > 0$ in* (14)*, there exists $\delta > 0$ such that $\mathcal{T}(w,p) - \mathcal{L}(w) \in Y$ for all $p \in B_\delta(p_*)$ and $w \in B_\delta(w_*) \cap X$.*

Proof. Again, we focus on the first component of $\mathcal{T} - \mathcal{L}$, since the other components are treated in a similar manner. Referring to the definition of Y, we should show that

$$\|(f_p(x,u) - f_*(x_*, u_*)) - A(x-x_*) - B(u-u_*)\|_{W^{1,\infty}} \le \kappa/4 \tag{20}$$

for p near p_* and for $(x,u) \in W_\mu^{2,\infty} \times \mathrm{Lip}_\mu$ near (x_*, u_*) in the norm of $H^1 \times L^2$. The $W^{1,\infty}$ norm in (20) is composed of two norms, the L^∞ norm of the function values, and the L^∞ norm of the time derivative. By the same expansion used in Lemma 3.3, we obtain the bound

$$\|(f_p(x,u) - f_*(x_*, u_*)) - A(x-x_*) - B(u-u_*)\|_{L^\infty} \le \kappa/8$$

for p near p_* and for (x,u) near (x_*, u_*). Differentiating the expression within the norm of (20) gives

$$\begin{aligned} &\tfrac{d}{dt}\big(f_p(x,u) - f_*(x_*, u_*) - A(x-x_*) - B(u-u_*)\big) \\ &= \big(\nabla_x f_p(x,u) - A\big)\dot{x} + \big(\nabla_u f_p(x,u) - B\big)\dot{u} - \dot{A}(x-x_*) - \dot{B}(u-u_*). \end{aligned}$$

By the smoothness condition, $\dot{A}$ and $\dot{B}$ lie in L^∞, and by the definition of X, we have $\|\dot{u}\|_{L^\infty} \le \mu$. By the triangle inequality and by Lemma 3.1,

$$\|\dot{x}\|_{L^\infty} \le \|\dot{x}_*\|_{L^\infty} + \|\dot{x} - \dot{x}_*\|_{L^\infty} \le \|\dot{x}_*\|_{L^\infty} + \sqrt[3]{3\mu}\|x - x_*\|_{H^1}^{2/3}$$

for x near x_*. Moreover, by Lemma 3.1 and by the smoothness condition, $\nabla_x f_p(x,u)$ approaches A and $\nabla_u f_p(x,u)$ approaches B in L^∞ as p approaches p_* and (x,u) approaches (x_*,u_*). Hence, for p near p_* and (x,u) near (x_*,u_*), we have

$$\left\| \frac{d}{dt}\big(f_p(x,u) - f_*(x_*,u_*) - A(x - x_*) - B(u - u_*)\big)\right\|_{L^\infty} \le \kappa/8.$$

Analyzing each of the components of $\mathcal{T} - \mathcal{L}$ in this same way, the proof is complete. □

We now begin a series of lemmas aimed at verifying (Q4). After a technical result (Lemma 3.5) related to the constraints, a surjectivity property (Lemma 3.6) is established for the linearized constraint mapping. Then we study a quadratic program corresponding to the linear variational system (15)–(18). We show that the solution (Lemma 3.9) and the multipliers (Lemma 3.10) depend Lipschitz continuously on the parameters. And utilizing the solution regularity derived in [8], the solution and the multipliers lie in X for μ sufficiently large.

To begin, let I be any map from $[0,1]$ to the subsets of $\{1,2,\ldots,k\}$ with the property that the following sets I_i are closed for every i:

$$I_i = I^{-1}(i) = \{t \in [0,1] : i \in I(t)\}\ .$$

We establish the following decomposition property for the interval $[0,1]$.

LEMMA 3.5. *If uniform independence at I holds, then for every $\alpha', 0 < \alpha' < \alpha$, there exist sets $J_1, J_2, \ldots, J_l$, corresponding points $0 = \tau_1 < \tau_2 < \cdots < \tau_{l+1} = 1$, and a positive constant $\rho < \min_i(\tau_{i+1} - \tau_i)$ such that for each $t \in [\tau_i - \rho, \tau_{i+1} + \rho] \cap [0,1]$, we have $I(t) \subset J_i$, and if J_i is nonempty, then*

$$(21)\qquad \left|\sum_{j\in J_i} v_j K_j(t) B(t)\right| \ge \alpha' |v_{J_i}|$$

for every choice of v. The set J_1 can always be chosen empty.

Proof. For each $t \in (0,1)$ with $I(t)^c \ne \emptyset$, there exists an open interval O centered at t with $O \subset \cap_{i\in I(t)^c} I_i^c$. If $t = 0$ or 1, then we can choose a half-open interval O, with t the closed end of the interval, such that $O \subset \cap_{i\in I(t)^c} I_i^c$. If $I(t)^c$ is empty, take $O = [0,1]$. For any fixed $t \in [0,1]$ with $I(t) \ne \emptyset$, choose O smaller if necessary so that

$$(22)\qquad \left|\sum_{i\in I(t)} v_i K_i(s) B(s)\right| \ge \alpha' |v_{I(t)}|$$

for each $s \in O$ and for each choice of v. Since B and K are continuous, it is possible to choose O in this way. Observe that by the construction of O, we have $I(s) \subset I(t)$ for each $s \in O$ and (22) holds if $I(t)$ is nonempty. Given any interval O on $(0,1)$, let $O_{1/2}$ denote the open interval with the same center but with half the length; for the open intervals associated with $t = 0$ or 1, let $O_{1/2}$ denote the half-open interval with the same endpoint, 0 or 1, but with half the length. The sets $O_{1/2}$ form an open cover of $[0,1]$. Let $O_1, O_2, \ldots, O_l$ be a finite subcover of $[0,1]$ and let $t_1, t_2, \ldots, t_l$ denote the associated centers of interior intervals, and the closed endpoint of the intervals associated with $t = 0$ or 1. It can be arranged so that no O_i is contained in the union of other elements of the subcover (by discarding these extra sets if necessary). Arrange the indices of the O_i so that the left side of O_i is to the left of the left side of O_{i+1} for each i. Let $\tau_1, \tau_2, \ldots, \tau_{l-1}$ denote the successive left sides of the O_i, and let ρ be 1/4 of the length of the smallest O_i. Defining $J_i = I(t_i)$ for $i \geq 1$, it follows from the construction of the O_i that $I(t) \subset J_i$ and (22) holds for each t in an interval associated with t_i and with length twice that of O_i. Since $(\tau_i, \tau_{i+1}) \subset O_i$, we have (21). By taking ρ smaller if necessary, we can enforce the condition $\rho < \min_i(\tau_{i+1} - \tau_i)$. ∎

LEMMA 3.6. *If uniform independence at I holds, then for each $a \in L^\infty$ and $b \in W^{1,\infty}$, there exist $x \in W^{1,\infty}$ and $u \in L^\infty$ such that $L(x,u) + a = 0$, $x(0) = x^0$, and*

$$K_j(t)x(t) + b_j(t) = 0 \quad \text{for each} \quad j \in I(t), \quad t \in [0,1]. \tag{23}$$

This (x,u) pair is an affine function of (a,b), and for each $\alpha \geq 1$, there exists a constant $c > 0$ such that

$$\|x_1 - x_2\|_{W^{1,\alpha}} + \|u_1 - u_2\|_{L^\alpha} \leq c(\|a_1 - a_2\|_{L^\alpha} + \|b_1 - b_2\|_{W^{1,\alpha}}) \tag{24}$$

for every $(a_i, b_i) \in L^\infty \times W^{1,\infty}$, $i = 1, 2$, where (x_i, u_i) is the pair associated with (a_i, b_i).

Proof. We use the decomposition provided by Lemma 3.5 to enforce the equations

$$\dot{x}(t) - A(t)x(t) - B(t)u(t) + a(t) = 0, \quad x(0) = x^0, \tag{25}$$

$$K_j(t)x(t) + b_j(t) = 0 \quad \text{for each} \quad j \in J_i \setminus J_{i-1}, \quad t \in [\tau_i + \rho, \tau_{i+1}], \tag{26}$$

$$K_j(t)x(t) + b_j(t) = 0 \quad \text{for each} \quad j \in J_i \cap J_{i-1}, \quad t \in [\tau_i, \tau_{i+1}], \tag{27}$$

$i = 2, 3, \ldots, l$. Since J_1 is empty, (23) holds trivially on $[\tau_1, \tau_2] = [0, \tau_2]$. Suppose that $i > 1$, and let us consider (23) on the interval $[\tau_i, \tau_{i+1}]$. Since $I(t) \subset J_i$ for $t \in [\tau_i, \tau_{i+1}]$, we conclude that any $j \in I(t)$ is contained in either $J_i \cap J_{i-1}$ or $J_i \setminus J_{i-1}$. If $j \in J_i \cap J_{i-1}$, then by (27), (23) holds. If $j \in J_i \setminus J_{i-1}$, then by the construction of the J_i, $j \notin I(t)$ for $t \in [\tau_i, \tau_i + \rho]$. Hence, (26) implies that (23) holds.

Suppose that $j \in J_i$ and let σ_j be any given Lipschitz continuous function. Observe that if

$$K_j(\tau_i)x(\tau_i) + \sigma_j(\tau_i) = 0 \quad \text{and} \quad \frac{d}{dt}\big(K_j(t)x(t) + \sigma_j(t)\big) = 0 \quad \text{a.e.} \quad t \in [\tau_i, \tau_{i+1}], \tag{28}$$

then $K_j(t)x(t) + \sigma_j(t) = 0$ for all $t \in [\tau_i, \tau_{i+1}]$. Carrying out the differentiation in the second relation of (28) and substituting for $\dot{x}$ using the state equation (25), we obtain a linear equation for u. By Lemma 3.5, this equation has a solution, and for fixed t and x, the minimum norm solution can be written:

$$u(t,x) = M_i(t)[-\dot{\sigma}_{J_i}(t) + K_{J_i}(t)a(t) - \dot{K}_{J_i}(t)x - K_{J_i}(t)A(t)x], \tag{29}$$

where

$$M_i(t) = (K_{J_i}(t)B(t))^{\mathsf{T}}[K_{J_i}(t)B(t)(K_{J_i}(t)B(t))^{\mathsf{T}}]^{-1}. \tag{30}$$

In the special case where J_i is empty, we simply set $u(t,x) = 0$.

These observations show how to construct x and u in order to satisfy (26) and (27). On the initial interval $[0, \tau_2]$, u is simply 0 and x is obtained from (25). Assuming x and u have been determined on the interval $[0, \tau_i]$, their values on $[\tau_i, \tau_{i+1}]$ are obtained in the following way: the control is given in feedback form by (29), where for $j \in J_i \cap J_{i-1}$,

$$\sigma_j(t) = b_j(t) \;\text{ for }\; t \in [\tau_i, \tau_{i+1}]. \tag{31}$$

For $j \in J_i \setminus J_{i-1}$, $\sigma_j(t) = b_j(t)$ for $t \in [\tau_i + \rho, \tau_{i+1}]$, while σ_j is linear on $[\tau_i, \tau_i + \rho]$ with

$$\sigma_j(\tau_i) = -K_j(\tau_i)x(\tau_i) \;\text{ and }\; \sigma_j(\tau_i + \rho) = b_j(\tau_i + \rho). \tag{32}$$

With this choice for σ, the first equation in (28) is satisfied, and with x and u given by (25) and (29), respectively, the second equation in (28) is satisfied. Also, by the choice of σ,

$$K_j(t)x(t) + \sigma_j(t) = K_j(t)x(t) + b_j(t) = 0$$

for each $j \in J_i \cap J_{i-1}$ and $t \in [\tau_i, \tau_{i+1}]$, and for each $j \in J_i \setminus J_{i-1}$ and $t \in [\tau_i + \rho, \tau_{i+1}]$. Hence, (26) and (27) hold, which yields (23).

For $j \in J_i$, it follows from the definition of σ that

$$|\dot{\sigma}_j(t)| \le c\big(|x(\tau_i)| + \|b\|_{W^{1,\infty}}\big) \;\text{ a.e. }\; t \in [\tau_i, \tau_{i+1}].$$

When u in (29) is inserted in (25) and this bound on $|\dot{\sigma}_j(t)|$ is taken into account, we obtain by induction that $x \in W^{1,\infty}$ and $u \in L^\infty$. By the equations (25) for the state, (29) for the control, and (31)–(32) for σ, (x, u) is an affine function of (a, b). Moreover, the change $(\delta x, \delta u)$ in the state and control associated with the change $(\delta a, \delta b)$ in the parameters satisfies

$$\|\delta x\|_{W^{1,\alpha}([0,\tau_i])} + \|\delta u\|_{L^\alpha([0,\tau_i])} \le c\big(\|\delta a\|_{L^\alpha([0,\tau_i])} + \|\delta \dot{\sigma}\|_{L^\alpha([0,\tau_i])}\big), \tag{33}$$

for each i where σ is specified in (31)–(32).

To complete the proof, we need to relate the σ term of (33) to the b term of (24). For $j \in J_i$, $\delta\sigma_j(t) = \delta b_j(t)$ if $t \in [\tau_i + \rho, \tau_{i+1}]$ or if $j \in J_{i-1}$ and $t \in [\tau_i, \tau_i + \rho]$. For $j \in J_i \setminus J_{i-1}$ and $t \in [\tau_i, \tau_i + \rho]$, we have

$$\begin{aligned}|\delta\dot{\sigma}_j(t)| &\le \big(|\delta b_j(\tau_i + \rho)| + |K_j(\tau_i)\delta x(\tau_i)|\big)/\rho \le c\big(\|\delta b\|_{L^\infty} + |\delta x(\tau_i)|\big)\\ &\le c\big(\|\delta b\|_{W^{1,\alpha}} + |\delta x(\tau_i)|\big).\end{aligned}$$

Consequently, for almost every $t \in [\tau_i, \tau_{i+1}]$,

$$|\delta\dot{\sigma}(t)| \le c\big(\|\delta b\|_{W^{1,\alpha}} + |\delta\dot{b}(t)| + |\delta x(\tau_i)|\big). \tag{34}$$

Since $\delta x(0) = 0$, let us proceed by induction and assume that

$$|\delta x(\tau_i)| \le c(\|\delta a\|_{L^\alpha} + \|\delta b\|_{W^{1,\alpha}}) \;\text{ for }\; i = 1,\ 2, \ldots, j.$$

Combining this with (34) and (33) for $i = j + 1$ gives

$$\|\delta x\|_{W^{1,\alpha}([0,\tau_{j+1}])} + \|\delta u\|_{L^\alpha([0,\tau_{j+1}])} \le c\ \big(\ \|\delta a\|_{L^\alpha} + \|\delta b\|_{W^{1,\alpha}}\ \big)\ .$$

Since $|\delta x(\tau_{j+1})| \le \|\delta x\|_{W^{1,\alpha}([0,\tau_{j+1}])}$, the induction step is complete. □

In the following lemma, we prove a pointwise coercivity result for the quadratic form $\mathcal{B}$. See [4] and [7] for more general results of this nature.

LEMMA 3.7. *If coercivity holds, then there exists a scalar $\alpha > 0$ such that*

$$\mathcal{B}(x,u) \geq \alpha[\langle x,x\rangle + \langle u,u\rangle + \langle \dot{x},\dot{x}\rangle] \quad \text{for all} \quad (x,u) \in \mathcal{M} \tag{35}$$

and

$$v^{\mathsf{T}} R(t) v \geq \alpha v^{\mathsf{T}} v \text{ for every } t \in [0,1] \text{ and } v \in \mathbf{R}^m. \tag{36}$$

Proof. If $(x,u) \in \mathcal{M}$, then $L(x,u) = 0$ and $x(0) = 0$. Hence, the L^2 norm of x and $\dot{x}$ are bounded in terms of the L^2 norm of u, and (35) follows directly from the coercivity condition. To establish (36), we consider the control u_ϵ defined by

$$u_\epsilon(s) = \begin{cases} v & \text{for } t - \epsilon/2 \leq s \leq t + \epsilon/2, \\ 0 & \text{otherwise.} \end{cases}$$

Let the state x_ϵ be the solution to $L(x_\epsilon, u_\epsilon) = 0$, $x_\epsilon(0) = 0$. For any $t \in (0,1)$, we have

$$\lim_{\epsilon \to 0} \frac{\mathcal{B}(x_\epsilon, u_\epsilon)}{\epsilon} = v^{\mathsf{T}} R(t) v \quad \text{and} \quad \lim_{\epsilon \to 0} \frac{\langle u_\epsilon, u_\epsilon \rangle}{\epsilon} = v^{\mathsf{T}} v.$$

Combining this with the coercivity condition gives (36). ◻

Consider the following linear-quadratic problem involving the parameters a, s, $r \in L^\infty$ and $b \in W^{1,\infty}$:

$$\begin{aligned} &\text{minimize } \mathcal{B}(x,u) + \langle s,x\rangle + \langle r,u\rangle \\ &\text{subject to} \\ &\quad L(x,u) + a = 0, \quad x(0) = x^0, \\ &\quad K_{I(t)}(t)x(t) + b_{I(t)}(t) \leq 0 \text{ for all } t \in [0,1], \\ &\quad x \in W^{1,\infty}([0,1];\mathbf{R}^n), \quad u \in L^\infty([0,1];\mathbf{R}^m). \end{aligned} \tag{37}$$

If the feasible set for (37) is nonempty, then coercivity implies the existence of a unique minimizer over $H^1 \times L^2$. Using the following lemma, we show that this minimizer lies in $W^{1,\infty} \times L^\infty$, and that it exhibits stability relative to the L^2 norm.

LEMMA 3.8. *If coercivity and uniform independence at I hold, then* (37) *has a unique solution for every $a, r, s \in L^\infty$ and $b \in W^{1,\infty}$. Moreover, the change $(\delta x, \delta u)$ in the solution to* (37) *corresponding to a change $(\delta a, \delta b, \delta s, \delta r)$ in the parameters satisfies the estimate*

$$\|\delta x\|_{H^1} + \|\delta u\|_{L^2} \leq c(\|\delta a\|_{L^2} + \|\delta b\|_{H^1} + \|\delta s\|_{L^2} + \|\delta r\|_{L^2}). \tag{38}$$

Proof. By Lemma 3.6, uniform independence at I implies that the feasible set for (37) is nonempty, while the coercivity condition implies the existence of a unique solution (x_*, u_*) in $H^1 \times L^2$. From duality theory (for example, see [10]), there exists $\lambda \in L^\infty$ with the property that $u = u_*$ is the minimum with respect to u of the expression

$$\mathcal{B}(x,u) + \langle s,x\rangle + \langle r,u\rangle + \langle \lambda, \dot{x} - Ax - Bu + a\rangle$$

over all $u \in L^\infty$. It follows that

$$R(t)u_*(t) + M(t)^{\mathsf{T}} x_*(t) + r(t) - B(t)^{\mathsf{T}} \lambda(t) = 0, \tag{39}$$

and by (36), $u_*(t)$ is uniformly bounded in t. From the equations $L(x_*, u_*) = 0$ and $x_*(0) = x^0$, $x_* \in W^{1,\infty}$.

The estimate (38) can be obtained, as in Lemma 5 in [2], by eliminating the perturbation in the constraints. Let Λ be the affine map in Lemma 3.6 relating the feasible pair (x, u) to the parameters (a, b). By making the substitution $(x, u) = (y, v) + \Lambda(a, b)$, we transform (37) to an equivalent problem of the form

$$
\begin{aligned}
(40) \qquad & \text{minimize } \mathcal{B}(y, v) + \langle \sigma, y \rangle + \langle \rho, v \rangle \\
& \text{subject to} \\
& \quad L(y, v) = 0, \quad y(0) = 0, \\
& \quad K_{I(t)}(t) y(t) \le 0 \text{ for all } t \in [0, 1], \\
& \quad y \in W^{1,\infty}([0, 1]; \mathbf{R}^n), \quad v \in L^\infty([0, 1]; \mathbf{R}^m).
\end{aligned}
$$

Here σ and ρ are affine functions of $a, b, s,$ and r. Utilizing the coercivity condition and the analysis of [9, section 2], we obtain the following estimate for the change $(\delta y, \delta v)$ corresponding to the change $(\delta\sigma, \delta\rho)$:

$$
\begin{aligned}
\alpha(\|\delta y\|_{H^1}^2 + \|\delta v\|_{L^2}^2) &\le \|\delta\sigma\|_{L^1} \|\delta y\|_{L^\infty} + \|\delta\rho\|_{L^2} \|\delta v\|_{L^2} \\
&\le \|\delta\sigma\|_{L^1} \|\delta y\|_{H^1} + \|\delta\rho\|_{L^2} \|\delta v\|_{L^2}.
\end{aligned}
$$

Hence,

$$
\|\delta y\|_{H^1} + \|\delta v\|_{L^2} \le c(\|\delta\sigma\|_{L^1} + \|\delta\rho\|_{L^2}).
$$

Taking into account the relations between (x, u), (y, v), (σ, ρ), and (a, b, s, r), the proof is complete. ☐

Now let us consider the full linear-quadratic problem where the subscript I on the state constraint has been removed:

$$
\begin{aligned}
(41) \qquad & \text{minimize } \mathcal{B}(x, u) + \langle s, x \rangle + \langle r, u \rangle \\
& \text{subject to} \\
& \quad L(x, u) + a = 0, \quad x(0) = x^0, \\
& \quad K(t) x(t) + b(t) \le 0 \text{ for all } t \in [0, 1], \\
& \quad x \in W^{1,\infty}([0, 1]; \mathbf{R}^n), \quad u \in L^\infty([0, 1]; \mathbf{R}^m).
\end{aligned}
$$

The first-order necessary conditions for this problem are precisely (15)–(18). Observe that x_*, u_*, ψ_*, and ν_* satisfy (15)–(18) when $\pi = \pi_*$. Since the first-order necessary conditions are sufficient for optimality when coercivity holds, (x_*, u_*) is the unique solution to (41) at $\pi = \pi_*$. In addition, if uniform independence holds, we now show that the multipliers ψ and ν satisfying (16)–(18) are unique; hence, x_*, u_*, ψ_*, and ν_* are the unique solution to (15)–(18) for $\pi = \pi_*$.

To establish this uniqueness property for the multipliers, we apply Lemma 3.5 to the active constraint map $\mathcal{A}$ of section 1. Let J_i be the index sets associated with $I = \mathcal{A}$ in Lemma 3.5. Since $\mathcal{A}(t) \subset J_i$ for each $t \in [\tau_i, \tau_{i+1}]$, the complementary slackness condition $\nu_*(1)^{\mathsf{T}} g_*(1) = 0$, associated with the condition (5) of the minimum principle, implies that $(\nu_*)_{J_l^c} = 0$ on $[\tau_l, 1]$, while (21) along with (16) and (17) imply that $(\nu_*)_{J_l}$ and ψ_* are uniquely determined on $[\tau_l, 1]$. Proceeding by induction, suppose that ψ_* and ν_* are uniquely determined on the interval $[\tau_{i+1}, 1]$. Since $(\nu_*)_{J_i^c}$ is constant on $[\tau_i, \tau_{i+1}]$, it is uniquely determined by the continuity of ν_*, while $(\nu_*)_{J_i}$

and ψ_* on $[\tau_i, \tau_{i+1}]$ are uniquely determined by (21), (16), and (17). This completes the induction step.

We now use Lemma 3.8 to show that the solution to (41) depends Lipschitz continuously on the parameters when coercivity and uniform independence at $\mathcal{A}$ hold. We do this by making a special choice for the map I. Again, let J_i be the index sets associated with $I = \mathcal{A}$ by Lemma 3.5. Since $\mathcal{A}(t) \subset J_i$ for each $t \in [\tau_i, \tau_{i+1}]$, the parameter

$$\epsilon_i = -\sup\{(g_*)_j(t) : t \in [\tau_i, \tau_{i+1}], j \in J_i^c\} \tag{42}$$

is strictly positive for each i. Setting $\epsilon = .5 \min \epsilon_i$, we consider (37) in the case $I = \mathcal{A}_\epsilon$ where $\mathcal{A}_\epsilon(t)$ is the index set associated with the ϵ-active constraints for the linearized problem:

$$\mathcal{A}_\epsilon(t) = \{i : K_i(t)x_*(t) + (b_*)_i(t) \geq -\epsilon\} = \{i : (g_*)_i(t) \geq -\epsilon\}. \tag{43}$$

Since $\mathcal{A}_\epsilon(t) \subset J_i$ for each $t \in [\tau_i, \tau_{i+1}]$, Lemma 3.5 implies that uniform independence at $\mathcal{A}_\epsilon$ holds.

We now observe that the solution (x_*, u_*) of (41) at $\pi = \pi_*$ is the solution of (37) for $I = \mathcal{A}_\epsilon$ and $\pi = \pi_*$. First, (x_*, u_*) is feasible in (37) since there are fewer constraints than in (41). By the choice $I = \mathcal{A}_\epsilon$, all feasible pairs for (37) near (x_*, u_*) are also feasible in (41). Since (x_*, u_*) is optimal in (41), it is locally optimal in (37) as well, and by the coercivity condition and Lemma 3.7, (x_*, u_*) is the unique minimizer of (37) for $\pi = \pi_*$. By Lemma 3.8, we have an estimate for the change in the solution to (37) corresponding to a change in the parameters. Since $\|\delta x\|_{L^\infty} \leq \|\delta x\|_{H^1}$, it follows that for small perturbations in the data, the solution to (37) is feasible, and hence optimal, for (41). Hence, our previous stability analysis for (37) provides us with a local stability analysis for (41). We summarize this result in the following way.

LEMMA 3.9. *If coercivity and uniform independence at $\mathcal{A}$ hold, then for s, r, and a in an L^∞ neighborhood of s_*, r_*, and a_*, respectively, and for b in a $W^{1,\infty}$ neighborhood of b_*, there exists a unique minimizer of* (41)*, and the estimate* (38) *holds. Moreover, taking $I = \mathcal{A}_\epsilon$ with $\epsilon = .5 \min \epsilon_i$, where ϵ_i is defined in* (42)*, the solutions to* (37) *and* (41) *are identical in these neighborhoods.*

Now let us consider the multipliers associated with (41).

LEMMA 3.10. *If coercivity and uniform independence at $\mathcal{A}$ hold, then for s, r, and a in an L^∞ neighborhood of s_*, r_*, and a_*, respectively, and for b in a $W^{1,\infty}$ neighborhood of b_*, there exists a unique minimizer of* (41) *and associated unique multipliers satisfying the estimate:*

$$\|\delta\psi\|_{H^1} + \|\delta\nu\|_{L^2} \leq c(\|\delta a\|_{L^2} + \|\delta b\|_{H^1} + \|\delta s\|_{L^2} + \|\delta r\|_{L^2}). \tag{44}$$

Proof. Let $\mathcal{A}_\epsilon$ be the ϵ-active constraints defined by (43), where $\epsilon = .5 \min \epsilon_i$. Let J_i be the index sets and let ρ be the positive number associated with $I = \mathcal{A}$ by Lemma 3.5. Consider $\pi = \pi_* + \delta\pi$ where $\delta\pi$ is small enough that the active constraint set for (41) is a subset of $\mathcal{A}_\epsilon(t)$ for each t. By the same analysis used to establish uniqueness of (ψ_*, ν_*), there exist unique Lagrange multipliers $(\psi, \nu) = (\psi_*, \nu_*) + (\delta\psi, \delta\nu)$ corresponding to $\pi = \pi_* + \delta\pi$. We will show that

$$\|\delta\psi\|_{H^1} + \|\delta\nu\|_{L^2} \leq c(\|\delta x\|_{L^2} + \|\delta u\|_{L^2} + \|\delta s\|_{L^2} + \|\delta r\|_{L^2}). \tag{45}$$

Combining this with Lemma 3.9 yields Lemma 3.10.

We prove (45) by induction. Let us start with the interval $[\tau_l - \rho, 1]$. If $i \in J_l^c$, then $\nu_i(t) = 0$ for each $t \in [\tau_l - \rho, 1]$. Hence, $\delta\nu_{J_l^c} = 0$ on $[\tau_l - \rho, 1]$. Multiplying (17) by KB, we can solve for $\delta\nu_{J_l}$ and substitute in (16) to eliminate ν. Since $\psi(1) = 0$, it follows that

$$\|\delta\psi\|_{H^1([\sigma-\rho,1])} + \|\delta\nu\|_{L^2([\sigma-\rho,1])} \le c(\|\delta x\|_{L^2} + \|\delta u\|_{L^2} + \|\delta s\|_{L^2} + \|\delta r\|_{L^2}) \tag{46}$$

for $\sigma = \tau_l$.

Proceeding by induction, suppose that (46) holds for $\sigma = \tau_{j+1}$; we wish to show that it holds for $\sigma = \tau_j$. If $i \in J_j^c$, then $\nu_i(t)$ is constant on $[\tau_j - \rho, \tau_{j+1}]$, and we have

$$\int_{\tau_j-\rho}^{\tau_{j+1}} \delta\nu_i(t)^2 dt = \frac{\tau_{j+1} - \tau_j + \rho}{\rho} \int_{\tau_{j+1}-\rho}^{\tau_{j+1}} \delta\nu_i(t)^2 dt.$$

Combining this with (46) for $\sigma = \tau_{j+1}$, it follows that

$$\|\delta\nu_i\|_{L^2([\sigma-\rho,1])} \le c(\|\delta x\|_{L^2} + \|\delta u\|_{L^2} + \|\delta s\|_{L^2} + \|\delta r\|_{L^2})$$

for $\sigma = \tau_j$. Again, multiplying (17) by KB, we solve for $\delta\nu_{J_j}$ and substitute in (16). Since $|\delta\psi(\tau_j)| \le \|\delta\psi\|_{H^1([\tau_j,1])}$, the induction bound (46) for $\sigma = \tau_{j+1}$ coupled with the bound already established for $\delta\nu_i$, $i \in J_j^c$, gives (46) for $\sigma = \tau_j$. This completes the induction. ∎

LEMMA 3.11. *Suppose that smoothness, coercivity, and uniform independence at* $\mathcal{A}$ *hold and let* κ *be small enough that* Y *is contained in the neighborhoods defined in Lemmas* 3.9 *and* 3.10. *Then for some* $\mu > 0$ *and for each* $\pi \in Y$, *there exists a unique solution* (x, u) *to* (41) *and associated multipliers* (ψ, ν) *satisfying the estimates* (38) *and* (44), $(x, \psi, u, \nu) = (\mathcal{F} - \mathcal{L})^{-1}\pi$, *and we have* $\dot{x}$, $\dot{\psi}$, u, $\nu \in \mathrm{Lip}_\mu$.

Proof. If $w = (x, \psi, u, \nu)$ denotes $(\mathcal{F} - \mathcal{L})^{-1}\pi$, then w satisfies the first-order necessary conditions (15)–(18) associated with (41). Lemmas 3.9 and 3.10 tell us that the unique solution and multipliers for (41) satisfy the estimates (38) and (44) for π near π_*. Since the first-order necessary conditions are sufficient for optimality when coercivity holds, the variational system (15)–(18) has a unique solution, for π near π_*, that is identical to the solution and multipliers for (41), and the estimates (38) and (44) are satisfied.

To complete the proof, we need to show that $\dot{x}$, $\dot{\psi}$, u, $\nu \in \mathrm{Lip}_\mu$ for some constant $\mu > 0$. This follows from the regularity results of [8], where it is shown that the solution to a constant coefficient, linear-quadratic problem satisfying the uniform independence condition and with R positive definite, Q positive semidefinite, and $M = 0$ has the property that the optimal u and associated ν are Lipschitz continuous in time, while the derivatives of x and ψ are Lipschitz continuous in time. Moreover, the Lipschitz constant in time is bounded in terms of the constant α in the uniform independence condition and the smallest eigenvalue of R. Exactly the same analysis applies to a linear-quadratic problem with time-varying coefficients; however, the bound for the Lipschitz constant of the solution depends on the Lipschitz constants of the matrices of the problem and of the parameters a, r, s, and $\dot{b}$, as well as on a uniform bound for the smallest eigenvalue of $R(t)$ on $[0, 1]$ and for the parameter α in the uniform independence condition. By Lemma 3.9 and with the choice for I given in the statement of the lemma, the quadratic programs (37) and (41) have the same solution for s, r, and a in an L^∞ neighborhood of s_*, r_*, and a_* and for b in a $W^{1,\infty}$ neighborhood of b_*. Hence, for parameters in this neighborhood of π_*, the indices of the active constraints are contained in $I(t)$ for each t, and the independence condition (21) holds. Lemma 3.7 provides a lower bound for the eigenvalues of

$R(t)$. If $(a, s, r, b) \in Y$, then the Lipschitz constants for a, s, r, and $\dot{b}$ are bounded by those for a_*, s_*, r_*, and $\dot{b}_*$ plus κ. Hence, taking μ sufficiently large, the proof is complete. □

Proof of Theorem 1.1. We apply Theorem 2.2 with the identifications given at the beginning of this section and with μ chosen sufficiently large in accordance with Lemma 3.11. The completeness of X is established in Lemma 3.2, (Q1) is immediate, (Q2) follows from smoothness, (Q3) is proved in Lemma 3.3, (Q4) follows from Lemma 3.11, and (Q5) is established in Lemma 3.4. Applying Theorem 2.2, the estimate (7) is established. Under the uniform independence condition, coercivity is a second-order sufficient condition for local optimality (see [4, Theorem 1]) which is stable under small changes in either the parameters or the solution of the first-order optimality conditions. Finally, we apply Lemma 3.1 to obtain the L^∞ estimate of Theorem 1.1. □

We note that the coercivity condition we use here is a strong form of a second-order sufficient optimality condition; it not only provides optimality, but also guarantees Lipschitz continuity of the optimal solution and multipliers when uniform independence holds. As recently proved in [6] for finite-dimensional optimization problems, Lipschitzian stability of the solution and multipliers necessarily requires a coercivity condition stronger than the usual second-order condition. For the treatment of second-order sufficient optimality under conditions equivalent to coercivity, see [18] and [21]. These sufficient conditions can be applied to state constraints of arbitrary order. For recent work concerning the treatment of second-order sufficient optimality in state constrained optimal control, see [16], [19], and [22].

4. Lipschitzian stability in L^∞.

One way to sharpen the L^∞ estimate of Theorem 1.1 involves an assumption concerning the regularity of the solution to the linear-quadratic problem (41). The time t is a contact point for the ith constraint of $Kx + b \le 0$ if $(K(t)x(t) + b(t))_i = 0$ and there exists a sequence $\{t_k\}$ converging to t with $(K(t_k)x(t_k) + b(t_k))_i < 0$ for each k.

Contact separation. There exists a finite set $I_1, I_2, \ldots, I_N$ of disjoint, closed intervals contained in $(0, 1)$ and neighborhoods of (a_*, r_*, s_*) in $W^{1,\infty}$ and of b_* in $W^{2,\infty}$ with the property that for each a, r, s, and b in these neighborhoods, and for each solution to (41), all contact points are contained in the union of the intervals I_i with exactly one contact point in each interval and with exactly one constraint changing between active and inactive at this point.

Observe that if for (1) with $p = p_*$, there are a finite number of contact points, at each contact point exactly one constraint changes between active and inactive, and each contact point in the linear-quadratic problem (41) depends continuously on the parameters, then contact separation holds. The finiteness of the contact set is a natural condition in optimal control; for example, in [5] it is proved that for a linear-quadratic problem with time invariant matrices and one state constraint, the contact set is finite when uniform independence and coercivity hold.

THEOREM 4.1. *Suppose that the problem* (1) *with $p = p_*$ has a local minimizer (x_*, u_*) and that smoothness, contact separation, and uniform independence at $\mathcal{A}$ hold. Let ψ_* and ν_* be the associated multipliers satisfying the first-order necessary conditions* (2)–(5). *If the coercivity condition holds, then there exist neighborhoods V of p_* and U of $w_* = (x_*, \psi_*, u_*, \nu_*)$ in $W^{1,\infty} \times W^{1,\infty} \times L^\infty \times L^\infty$, such that for every $p \in V$, there exists a unique solution $w = (x, \psi, u, \nu) \in U$ to the first-order necessary conditions* (2)–(5) *and (x, u) is a local minimizer of the problem* (1) *associated with p. Moreover, for every $p_i \in V, i = 1, 2$, if $w_i = (x_i, \psi_i, u_i, \nu_i)$ is the corresponding* *solution of* (2)–(5)*, the following estimate holds:*

$$\|x_1 - x_2\|_{W^{1,\infty}} + \|\psi_1 - \psi_2\|_{W^{1,\infty}} + \|u_1 - u_2\|_{L^\infty} + \|\nu_1 - \nu_2\|_{L^\infty} \le cE_\infty.$$

To prove this result, we need to supplement the 2-norm perturbation estimates provided by Lemmas 3.9 and 3.10 with analogous ∞-norm estimates.

LEMMA 4.2. *If coercivity, uniform independence at $\mathcal{A}$, and contact separation hold, then there exist neighborhoods of (a_*, r_*, s_*) in $W^{1,\infty}$ and of b_* in $W^{2,\infty}$ such that for each a_i, r_i, s_i, and b_i, $i = 1, 2$, in these neighborhoods, the associated solutions (x_i, u_i) of* (41) *satisfy*

$$\begin{aligned} &\|\delta x\|_{W^{1,\infty}} + \|\delta\psi\|_{W^{1,\infty}} + \|\delta u\|_{L^\infty} + \|\delta\nu\|_{L^\infty} \\ &\le c(\|\delta a\|_{L^\infty} + \|\delta b\|_{W^{1,\infty}} + \|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}). \end{aligned} \tag{47}$$

Proof. Letting $\mathcal{A}_\epsilon$ denote the ϵ-active set defined in (43), we again choose $\epsilon = .5 \min \epsilon_i$, where ϵ_i is defined in (42). We consider parameters a, r, s, and b chosen within the neighborhoods of the contact separation condition, and sufficiently close to a_*, r_*, s_*, and b_* that the active constraint set for the solution of the perturbed linear-quadratic problem (41) is contained in $\mathcal{A}_\epsilon(t)$ for each t. By eliminating the perturbations in the constraints, as we did in the proof of Lemma 3.8, there is no loss of generality in assuming that $a = b = 0$. We refer to the quadratic programs corresponding to the parameters (r_1, s_1) and (r_2, s_2) as Problems 1 and 2.

Let (x, u) be either (x_1, u_1) or (x_2, u_2). If $t \in (0, 1)$ is a time for which $K_i(t)x(t) = 0$ for some i, then $\frac{d}{dt}(K_i x) = \dot{K}_i x + K_i \dot{x} = 0$. Substituting for $\dot{x}$ using the state equation $\dot{x} = Ax + Bu$ and for u using the necessary condition (17) yields

$$K_i BR^{-1}(KB)^{\mathsf{T}}\nu = -\dot{K}_i x - K_i Ax + K_i BR^{-1}(B^{\mathsf{T}}\psi + M^{\mathsf{T}}x + r).$$

This equation has the form

$$N_i \nu = S_i x + T_i \psi + U_i r \tag{48}$$

for suitable choices of the row vectors N_i, S_i, T_i, and U_i. Hence, at any time t where $K_i(t)x_1(t) = K_i(t)x_2(t) = 0$, the change in solution and multipliers corresponding to a change in parameters satisfies the equation

$$N_i(t)\delta\nu(t) = S_i(t)\delta x(t) + T_i(t)\delta\psi(t) + U_i(t)\delta r(t). \tag{49}$$

By the contact separation condition, Problems 1 and 2 have the same active set near $t = 1$. Since the components of ν corresponding to inactive constraints are constant and since $\nu_i(1) = 0$ if $K_i(1)x(1) < 0$, it follows that $\delta\nu_i(t) = 0$ for t near 1 when $K_i x_1(1) < 0 > K_i x_2(1)$. The relation (49) combined with uniform independence, with the L^2 estimates provided in Lemmas 3.9 and 3.10, and with a bound for the L^∞ norm in terms of the H^1 norm, gives

$$\|\delta\nu\|_{L^\infty[t,1]} \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}). \tag{50}$$

Using the bound (36) of Lemma 3.7 in (17) and applying Gronwall's lemma to (16), we have

$$\begin{aligned} &\|\delta x\|_{W^{1,\infty}[t,1]} + \|\delta\psi\|_{W^{1,\infty}[t,1]} + \|\delta u\|_{L^\infty[t,1]} + \|\delta\nu\|_{L^\infty[t,1]} \\ &\le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}) \end{aligned} \tag{51}$$

for all $t < 1$ in some neighborhood of $t = 1$. As t decreases, this estimate is valid until the first contact point is reached for either Problem 1 or Problem 2. Proceeding by induction, suppose that we have established (51) up to some contact point; we now wish to show that (51) holds up to the next contact point.

Again, by the contact separation condition, there is precisely one constraint, say constraint j, that makes a transition between active and inactive at the current contact point. Suppose that on the interval (α, β), the active sets for Problems 1 and 2 differ by the element j, and let τ be the first contact point to the left of α for either Problem 1 or Problem 2. If there is no such point, we take $\tau = 0$. By the contact separation condition, the difference $\alpha - \tau$ is uniformly bounded away from zero for all choices of the parameters s and r near s_* and r_*. There are essentially two cases to consider.

Case 1. Constraint j is active in Problem 2 to the left of $t = \beta$, and constraint j is active in Problem 1 to the left of $t = \alpha$.

Case 2. Constraint j is active in Problem 2 to the right of $t = \alpha$, and constraint j is active in Problem 1 to the right of $t = \beta$.

Case 1. Since constraint j is active in both Problems 1 and 2 at $t = \alpha$, it follows from (49) and from the uniform independence condition that

$$|\delta\nu_\Gamma(\alpha)| \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}) + c|\delta\nu_{\Gamma^c}(\alpha)|,$$

where Γ is the set of indices of active constraints at $t = \alpha$. Since $\delta\nu_i$ is constant for $i \in \Gamma^c$ on (α, β), the induction hypothesis yields

$$|\delta\nu_{\Gamma^c}(\alpha)| = |\delta\nu_{\Gamma^c}(\beta)|_{L^\infty} \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}). \tag{52}$$

Hence, we have

$$|\delta\nu(\alpha)| \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}). \tag{53}$$

Since ν_j is constant in Problem 1 on (α, β), and since it is monotone in Problem 2, the bound (53) coupled with the bound (51) at $t = \beta$ implies that

$$\|\delta\nu_j\|_{L^\infty([\alpha,\beta])} \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}). \tag{54}$$

Since $\delta\nu_i$ is constant on (α, β) for $i \in \Gamma^c$, it follows from (51) that

$$\|\delta\nu_{\Gamma^c}\|_{L^\infty([\alpha,\beta])} \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}). \tag{55}$$

Relation (49), for $i \in \Gamma^- = \Gamma \setminus \{j\}$, along with (54) and (55) yield

$$\|\delta\nu_{\Gamma^-}\|_{L^\infty([\alpha,\beta])} \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}). \tag{56}$$

Combining (54)–(56) gives

$$\|\delta\nu\|_{L^\infty([\alpha,\beta])} \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}). \tag{57}$$

On the interval from $t = \alpha$ down to the next contact point τ, precisely the same constraints are active in both Problems 1 and 2. Again, the relation (49) combined with uniform independence, with the L^2 estimates provided in Lemmas 3.9 and 3.10, and with a bound for the L^∞ norm in terms of the H^1 norm gives

$$\|\delta\nu\|_{L^\infty([\tau,\alpha])} \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}). \tag{58}$$

Relation (50) for $t = \beta$, along with (57) and (58), gives

$$\|\delta\nu\|_{L^\infty([\tau,1])} \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}).$$

And combining this with (15)–(17) gives (51) for $t = \tau$. This completes the induction step in Case 1.

Case 2. The mean value theorem implies that for some $\gamma \in (\tau, \alpha)$, we have

$$\begin{aligned}(\alpha - \tau)\frac{d}{dt}K_j(t)\delta x(t)\Big|_{t=\gamma} &= K_j(\alpha)\delta x(\alpha) - K_j(\tau)\delta x(\tau) \\ &\le 2\|K_j\|_{L^\infty}\|\delta x\|_{L^\infty} \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}).\end{aligned}$$

Hence, even though the derivative of $K_j x_i$ may not vanish on (τ, α), the derivative of the change $K_j\delta x$ is still bounded by the perturbation in the parameters at some $\gamma \in (\tau, \alpha)$:

$$\left|\frac{d}{dt}(K_i(t)\delta x(t))\right|_{t=\gamma} \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty})/(\alpha - \tau). \tag{59}$$

Since α and τ lie in disjoint closed sets I_k associated with the contact separation condition, $\alpha - \tau$ is bounded away from zero by the distance between the closest pair of sets. Focusing on the left side of (59), we substitute $\delta\dot{x} = A\delta x + B\delta u$, and we substitute for δu using (17) to obtain the relation

$$N_j(\gamma)\delta\nu(\gamma) = S_j(\gamma)\delta x(\gamma) + T_j(\gamma)\delta\psi(\gamma) + U_j(\gamma)\delta r(\gamma) + \Delta_j, \tag{60}$$

where $|\Delta_j| \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty})/(\alpha - \tau)$. Let Γ denote the set of indices of the active constraints at $t = \beta$. Combining (60) with (49) for $i \in \Gamma^- = \Gamma \setminus \{j\}$ gives

$$|\delta\nu_\Gamma(\gamma)| \le c(\|\delta r\|_{L^\infty} + \|\delta s\|_{L^\infty}) + c|\delta\nu_{\Gamma^c}(\gamma)|.$$

The analysis for Case 1 can now be applied, starting with (52) but with α replaced by γ. ∎

Remark 4.3. In the proof of Lemma 4.2, we needed to ensure that the difference $\alpha - \tau$, appearing in Case 2, was bounded away from zero. The contact separation condition ensures that this difference is bounded away from zero, since α and τ lie in disjoint closed intervals I_k. On the other hand, any condition that ensures a positive separation for the contact points α and τ in Case 2 can be used in place of the contact separation assumption of Theorem 4.1 and Lemma 4.2.

Proof of Theorem 4.1. The functions $\mathcal{T}$, $\mathcal{F}$, and $\mathcal{L}$ and the sets X, Π, and Y are the same as in the proof of Theorem 1.1 except that L^2 is replaced by L^∞ and H^1 is replaced by $W^{1,\infty}$ everywhere. Except for this change in norms, and the replacement of the L^2 estimates (38) and (44) referred to in Lemma 3.11 by the corresponding L^∞ estimate (47) of Lemma 4.2, the same proof used for Theorem 1.1 can be used to establish Theorem 4.1. ∎

5. Remarks. As mentioned in section 2, Theorem 2.2 is a generalization of Robinson's implicit function theorem [20] to nonlinear spaces. His theorem assumes that the nonlinear term is strictly differentiable and that the inverse of the linearized map is Lipschitz continuous. In optimal control, the latter condition amounts to Lipschitz continuity in L^∞ of the solution-multiplier vector associated with the linear-quadratic approximation. For problems with control constraints, this property for the solution is obtained, for example, in [1] or [4].

In this paper, we obtain Lipschitzian stability results for state constrained problems utilizing a new form of the implicit function theorem applicable to nonlinear spaces. We obtain optimal Lipschitzian stability results in L^2 and nonoptimal stability results in L^∞ under the uniform independence and the coercivity conditions. And with an additional contact separation condition, we obtain a tight L^∞ stability result. These are the first L^∞ stability results that have been established for state constrained control problems.

The uniform independence condition was introduced in [8], where it was shown that this condition together with the coercivity condition yield Lipschitz continuity in time of the solution and the Lagrange multipliers of a convex state and control constrained optimal control problem. Using Hager's regularity result, Dontchev [1] proved that the solution of this problem has a Lipschitz-type property with respect to perturbations. Various extensions of these results have been proposed by several authors. A survey of earlier results is given in [2].

In a series of papers (see [14], [15], and the references therein), Malanowski studied the stability of optimal control problems with constraints. In [15] he considers an optimal control problem with state and control constraints. His approach differs from ours in the following ways: he uses an implicit function theorem in linear spaces and a compactness argument, and the second-order sufficient condition he uses is different from our coercivity condition. Although there are some similar steps in the analysis of L^2 stability, the two approaches mainly differ in their abstract framework.

A prototype of Lemma 3.5 is given in [1, Lemma 2.5]. Lemma 3.6 is related to Lemma 3 in [2], although the analysis in Lemma 3.6 is much simpler since we ignore indices outside of $\mathcal{A}(t)$. In the analysis of the linear-quadratic problem (37), we follow the approach in [4].

Acknowledgment. The authors wish to thank both Kazimierz Malanowski for his comments on an earlier version of this paper and the reviewers for their constructive suggestions.

REFERENCES

[1] A. L. Dontchev, *Perturbations, Approximations and Sensitivity Analysis of Optimal Control Systems*, Lecture Notes in Control and Inform. Sci. 52, Springer-Verlag, New York, 1983.

[2] A. L. Dontchev and W. W. Hager, *Lipschitzian stability in nonlinear control and optimization*, SIAM J. Control Optim., 31 (1993), pp. 569–603.

[3] A. L. Dontchev and W. W. Hager, *An inverse function theorem for set-valued maps*, Proc. Amer. Math. Soc., 121 (1994), pp, 481–489.

[4] A. L. Dontchev, W. W. Hager, A. B. Poore, and B. Yang, *Optimality, stability and convergence in nonlinear control*, Appl. Math. Optim., 31 (1995), pp. 297–326.

[5] A. L. Dontchev and I. Kolmanovsky, *On regularity of optimal control*, in Proc. French-German Conf. on Optimization, Lecture Notes in Economics and Mathematical Systems 429, R. Durier and C. Michelot, eds., Springer-Verlag, New York, 1995, pp. 125–135.

[6] A. L. Dontchev and R. T. Rockafellar, *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*, SIAM J. Optim., 6 (1996), pp. 1087–1105.

[7] 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. 1361–1348.

[8] W. W. Hager, *Lipschitz continuity for constrained processes*, SIAM J. Control Optim., 17 (1979), pp. 321–337.

[9] W. W. Hager, *Multiplier methods for nonlinear optimal control*, SIAM J. Numer. Anal., 27 (1990), pp. 1061–1080.

[10] W. W. Hager and G. D. Ianculescu, *Dual approximations in optimal control*, SIAM J. Control Optim., 22 (1984), pp. 423–465.

[11] W. W. Hager and S. K. Mitter, *Lagrange duality theory for convex control problems*, SIAM J. Control Optim., 14 (1976), pp. 843–856.

[12] R. F. Hartl, S. P. Sethi, and R. G. Vickson, *A survey of the maximum principles for optimal control problems with state constraints*, SIAM Rev., 37 (1995), pp. 181–218.
[13] A. D. Ioffe and V. M. Tihomirov, *Theory of Extremal Problems*, North-Holland, Amsterdam, 1979.
[14] K. Malanowski, *Two-norm approach in stability and sensitivity analysis of optimization and optimal control problems*, Adv. Math. Sci. Appl., 2 (1993), pp. 397–443.
[15] K. Malanowski, *Stability and sensitivity of solutions to nonlinear optimal control problems*, Appl. Math. Optim., 32 (1995), pp. 111–141.
[16] K. Malanowski, *Sufficient optimality conditions for optimal control subject to state constraints*, SIAM J. Control Optim., 35 (1997), pp. 205–227.
[17] H. Maurer, *On the Minimum Principle for Optimal Control Problems with State Constraints*, Schriftenreihe des Rechenzentrums der Univ. Münster Nr. 41, Münster, 1979.
[18] H. Maurer, *First and second order sufficient optimality conditions in mathematical programming and optimal control*, Math. Programming Stud., 14 (1981), pp. 163–177.
[19] H. Maurer and S. Pickenhain, *Second order sufficient conditions for optimal control problems with control-state constraints*, J. Optim. Theory Appl., 86 (1995), pp. 649–667.
[20] S. M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.
[21] G. Sorger, *Sufficient conditions for nonconvex control problems with state constraints*, J. Optim. Theory Appl., 62 (1989), pp. 289–310.
[22] V. Zeidan, *The Riccati equation for optimal control problems with mixed state-control constraints; necessity and sufficiency*, SIAM J. Control Optim., 32 (1994), pp. 1297–1321.