SIAM J. Control and Optimization
Vol. 25, No. 3, May 1987

©1987 Society for Industrial and Applied Mathematics
018

LINEAR-QUADRATIC PROGRAMMING AND OPTIMAL CONTROL*

R.T. ROCKAFELLAR†

Abstract. A generalized approach is taken to linear and quadratic programming in which dual as well as primal variables may be subjected to bounds, and constraints may be represented through penalties. Corresponding problem models in optimal control related to continuous-time programming are then set up and theorems on duality and the existence of solutions are derived. Optimality conditions are obtained in the form of a global saddle point property which decomposes into an instantaneous saddle point condition on the primal and dual control vectors at each time, along with an endpoint condition.

Key words. Linear-quadratic programming, dual control problems, intertemporal programming, continuous-time programming, penalty representation of constraints

AMS(MOS) subject classifications. 49B10, 90C20, 90C05

1. Introduction. In finite-dimensional optimization a great importance is attached to problems of linear and quadratic programming. Such problems serve as mathematical models for a large number of applications. They are relatively easy to work with and possess duality properties that yield valuable insights and are the basis for many special algorithms. They are useful in methods of solving more general problems, for instance, in connection with sequential approximation or direction-finding subroutines. For such purposes they can be extended beyond traditional formulations to admit piecewise linear-quadratic objectives and penalty representations of constraints, although this possibility has not yet fully been utilized.

In optimal control there has not been a comparable emphasis on a "linear-quadratic" class of problems. The linear-quadratic regulator problem fits the picture to some degree but is virtually unconstrained. The continuous-time linear programming problems first introduced as "bottleneck" problems by Bellman [1] include certain types of control problems with constraints on states and controls (possibly mixed), but they carry no provision for quadratic terms in the objective and are very narrow in their treatment of initial and terminal conditions. Continuous-time linear programming problems do enjoy a strong duality theory, thanks to efforts of Tyndall [2], [3], Levinson [4], Grinold [5], [6], Schecter [7], Reiland [8], Meidan and Perold [9], and others. Continuous-time nonlinear programming has also been investigated, chiefly for duality; cf. Hanson [10], Hanson and Mond [11], Grinold [12], Farr and Hanson [13], Reiland and Hanson [14], Reiland [15]. This nonlinear literature covers certain classes of optimal control problems with quadratic terms, subject to the same limitations on the treatment of initial and terminal states. However, the quadratic case has not been worked out to take advantage of its special nature, and, in any case, the results are based on a Lagrange multiplier approach that does not yield even in finite dimensions a duality theory as broad and flexible as may currently be needed.

*Received by the editors December 16, 1985; accepted for publication (in revised form) June 24, 1986. This work was supported in part by a grant from the National Science Foundation at the University of Washington, Seattle.

† Department of Mathematics, University of Washington, Seattle, Washington 98195.

Our goal in this paper is to develop a theory of linear-quadratic programming-type problems specifically adapted to the optimal control setting and capable eventually of being used in new computational schemes, as well as directly.

Some of the motivation comes from mathematical modeling. Linear-quadratic models do not appear to have been used so far to their full potential. An obstacle may lie in the format in which finite-dimensional problems in linear programming and quadratic programming are ordinarily presented. In this format it is hard to deal with piecewise linear or piecewise quadratic functions, such as often are important in penalty representations, except by reformulations that disrupt the fundamental relationships, especially duality.

An alternative approach in finite dimensions, which we have followed recently in work on algorithms in stochastic programming [16], [17], [18], is to give primacy to an underlying saddle point problem (minimax problem). Thus we think of finite-dimensional linear-quadratic programming in a more general sense than usual as corresponding to finding a saddle point of a convex-concave quadratic (or linear) function on a product of polyhedral convex sets. Any such saddle point problem generates a primal problem of minimization and a dual problem of maximization. The classical case of linear and quadratic programming duality is the one where the polyhedral convex sets are orthants.

The problems in the general case could be reduced individually to the classical case, but by working directly in the broader format one gains several advantages. The most significant is the perception that bounds can reasonably be introduced for dual variables as well as primal variables, and moreover that this amounts to passing from exact representations of certain constraints to penalty representations.

We begin in §2 and §3 by explaining this unconventional approach to finite-dimensional linear-quadratic programming and the kinds of problem forms it handles. A particular aim is the elucidation of circumstances under which a model involving bounds on both primal and dual variables is appropriate, at least for computation. Then in §4 and §5 we introduce corresponding problems in optimal control, of a sort we call *intertemporal linear-quadratic programming*. The main results are obtained in §6. They consist of theorems on existence, duality, and the characterization of optimal controls. They are tied to an infinite-dimensional saddle point representation in terms of a convex-concave quadratic functional on a product of generalized polyhedral sets.

Our problems in optimal control have dynamics that are essentially linear, although "polyhedral differential inclusions" are also encompassed by the formulation. The expression of the objective and constraints involves, in general, terms that may be piecewise linear-quadratic. To clarify the nature of such terms in this introduction would take us too far. A brief description of one of the basic *linear* models covered by our theory is feasible, however, and may help to put the approach and results in perspective.

Over a fixed time interval $[t_0, t_1]$ we consider a dynamical system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + b(t), \qquad x(t_0) = B_e u_e + b_e, \tag{1.1}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^k$ is the instantaneous control and $u_e \in \mathbb{R}^{k_e}$ is an additional vector to be chosen, an "endpoint control." The incorporation of such a vector u_e may seem odd relative to the customary patterns in control theory, but it greatly aids in dualizing various conditions. Of course u_e could be trivialized by taking the dimension k_e to be 0 (then $x(t_0) = b_e$ in (1.1)). Another case to note is the one of a free initial point: $B_e = I$, $b_e = 0$ (then $x(t_0) = u_e$ in (1.1)). The subscript e will consistently be used in our notation for elements connected with endpoints.

For the basic linear case in question, the problem we associate with the system (1.1) takes the form

$$(\mathcal{P}_1)\qquad \begin{array}{lll} \text{minimize} & \multicolumn{2}{l}{\displaystyle\int_{t_0}^{t_1} [p(t)\cdot u(t) - c(t)\cdot x(t)]dt + [p_e\cdot u_e - c_e\cdot x(t_1)]} \\ \text{subject to} & (1.1) \text{ with} & C(t)x(t) + D(t)u(t) \geq q(t), \quad u(t) \geq 0, \\ & & C_e x(t_1) + D_e u_e \geq q_e, \quad u_e \geq 0. \end{array}$$

Discussion of the exact technical assumptions is postponed until §4. Observe, however, that the formulation allows for constraints only on the controls (rows of $C(t)$ consisting of 0's), constraints only on the states (rows of $D(t)$ consisting of 0's), and mixed constraints. The endpoint conditions allow for any system of finitely many linear equations or inequalities to be imposed on the pair $x(t_0), x(t_1)$ (as explained in detail in Examples 5.1 and 5.2 in §5).

In dualizing $(\mathcal{P}_1)$ we pass to the dynamical system

$$-\dot{y}(t) = A^*(t)y(t) + C^*(t)v(t) + c(t), \qquad y(t_1) = C_e^* v_e + c_e, \tag{1.2}$$

where $y(t) \in \mathbb{R}^n$ is the state, $v(t) \in \mathbb{R}^\ell$ is the instantaneous control, and $v_e \in \mathbb{R}^{\ell_e}$ is the endpoint control; the asterisk $*$ denotes the transpose of a matrix. The dual problem over the system (1.2) is

$$(\mathcal{Q}_1)\qquad \begin{array}{lll} \text{maximize} & \multicolumn{2}{l}{\displaystyle\int_{t_0}^{t_1} [q(t)\cdot v(t) - b(t)\cdot y(t)]dt + [q_e\cdot v_e - b_e\cdot y(t_0)]} \\ \text{subject to} & (1.2) \text{ with} & B_e^*(t)y(t) + D^*(t)v(t) \leq p(t), \quad v(t) \geq 0, \\ & & B_e y(t_0) + D_e^* v_e \leq p_e, \quad v_e \geq 0. \end{array}$$

Although $(\mathcal{P}_1)$ and $(\mathcal{Q}_1)$ have been written with inequality constraints only, there is no difficulty about extending the formulation to include equations in the manner familiar in linear programming. Thus, for example, the condition $C(t)x(t) + D(t)u(t) \geq q(t)$ in $(\mathcal{P}_1)$ can be converted to $C(t)x(t) + D(t)u(t) = q(t)$ by dropping the condition $v(t) \geq 0$ in $(\mathcal{Q}_1)$.

In contrast to $(\mathcal{P}_1)$ and $(\mathcal{Q}_1)$ the continuous-time linear programming problems mentioned earlier take the primal form

$$\begin{array}{ll} \text{minimize} & \displaystyle\int_{t_0}^{t_1} p(t)\cdot u(t)\,dt \\ \text{subject to} & \displaystyle\int_{t_0}^{t} K(t,\tau)u(\tau)d\tau + D(t)u(t) \geq q(t), \qquad u(t) \geq 0, \end{array}$$

and the dual form

$$\begin{array}{ll} \text{maximize} & \displaystyle\int_{t_0}^{t_1} q(t)\cdot v(t)\,dt \\ \text{subject to} & \displaystyle\int_{t}^{t_1} K^*(\tau,t)v(\tau)d\tau + D^*(t)v(t) \leq p(t), \qquad v(t) \geq 0, \end{array}$$

where the matrix $K(t,\tau)$ is some "kernel" with transpose $K^*(t,\tau)$. These are not necessarily problems of optimal control but become so in choosing

$$K(t,\tau) = C(t)\mathcal{A}(t)\mathcal{A}(\tau)^{-1}B(\tau)$$

with $A(t)$ the fundamental matrix corresponding to the differential equation (1.1) (i.e., $A(t)x_0$ is the unique solution to $\dot{x}(t) = A(t)x(t)$, $x(t_0) = x_0$), and setting

$$x(t) = A(t)\int_{t_0}^{t} A(\tau)^{-1}B(\tau)u(\tau)d\tau, \qquad y(t) = A^*(t)^{-1}\int_{t}^{t_0} A^*(\tau)C^*(\tau)v(\tau)d\tau.$$

Then one gets the case of $(\mathcal{P}_1)$ and $(\mathcal{Q}_1)$ where $b(t) = 0$, $c(t) = 0$, and all the e terms trivialize: the primal has $x(t_0) = 0$ but $x(t_1)$ free, whereas the dual has $y(t_1) = 0$ but $y(t_0)$ free.

In the work that has been done on special computational methods in continuous-time linear programming, e.g. Perold [19], [20], Anstreicher [21], attention has typically been limited further to the case where the kernel K is a *constant* matrix. In optimal control this corresponds not merely to having $A(t)$, $B(t)$ and $C(t)$ constant, but $A(t) \equiv 0$, a severe restriction.

Because of these distinctions and the desirability of being able to treat discrete-time analogues under the same heading, we shall refer to $(\mathcal{P}_1)$ and $(\mathcal{Q}_1)$ as problems of "intertemporal linear programming" (in continuous time) rather than "continuous-time linear programming."

The possiblity of mixed constraints on states and controls is important in accommodating many applications of an economic nature, involving planned activities with cumulative effects. But it also puts problems like $(\mathcal{P}_1)$ and $(\mathcal{Q}_1)$ beyond the range of the Pontryagin maximum principle. Mixed constraints can be readily handled, however, in the versions of optimal control and variational calculus that have been developed over the years in the conceptual framework of convex analysis and, more recently, nonsmooth analysis in the sense of Clarke [22].

The theory of convex problems of Bolza type, developed by the author in [23]–[29], is specifically applicable to problems $(\mathcal{P}_1)$, $(\mathcal{Q}_1)$ and their quadratic programming counterparts after a transformation which expresses everything through the trajectories x and y, as outlined in [30]. By this route it would be possible, with a degree of technical elaboration, to derive sharp duality theorems that characterize solutions and the circumstances in which they exist. Full justice to constraints involving states would, however, require us in the context of such duality to pass beyond the formulation of our primal and dual problems in terms of control *functions* u and v to one in which "impulse controls" may occur. An extension along those lines is indeed appropriate, and for the basic linear programming case in $(\mathcal{P}_1)$, $(\mathcal{Q}_1)$, it has been carried out by Murray [31] under a somewhat different choice of endpoint expressions.

For the present purpose we are able to postpone working with such an extension. We follow a different path and sidestep the difficulties posed by state constraints by appealing instead to alternative problem formulations where the constraints may be enforced by linear or piecewise linear-quadratic penalty expressions. We argue that as a practical matter of mathematical modeling and computation this is an often reasonable tactic which can be served by a much simpler theory where solutions always exist and strong duality always holds. The supporting results in finite-dimensional linear-quadratic programming provided in §§2 and 3 are critical in understanding this.

The saddle point representation furnished in §6 for the duality between our two infinite-dimensional problems of intertemporal linear-quadratic programming is of a kind not previously seen in optimal control. Moreover the representation has a separate decomposition property in each argument that may open the way to new saddle point techniques for computation such as extensions of the finite generation method

devised by R.J.-B. Wets and the author in a similar setting in stochastic programming [17]. Decomposition of the intertemporal saddle point condition leads to a characterization of optimality in terms of a "instantaneous" saddle point condition satisfied at each time t and an "endpoint" saddle point condition. This is a sort of "minimaximum principle" which has some precedent in continuous-time linear programming (Grinold [5, p. 46]) and the theory of Bolza problems (Rockafellar [23, Thm. 6]) but is new in this context of optimal control.

2. Linear-quadratic programming in finite dimensions. The infinite-dimensional control problems that are the subject of this paper, and our approach to them, will better be understood after a brief treatment of the formulation and duality properties of finite-dimensional linear-quadratic programming problems in the generalized sense. Such a treatment will also introduce facts and concepts that will be needed in later sections.

A simple foundation for almost all kinds of duality theory in optimization starts with a function $J(u,v)$ on a product set $U \times V$, where J is real-valued or possibly extended-real-valued. Regardless of the nature of J and the sets U and V (as long as the latter are nonempty), there is an associated *primal* problem

$$(\mathcal{P}_0) \qquad \text{minimize } f(u) \text{ over } U \quad \text{where } f(u) = \sup_{v\in V} J(u,v),$$

and a *dual* problem

$$(\mathcal{Q}_0) \qquad \text{maximize } g(v) \text{ over } V \quad \text{where } g(v) = \inf_{u\in U} J(u,v).$$

The relationship between these problems is tied to the *saddle point*, or *minimax* problem for J on $U \times V$, a saddle point being by definition a pair $(\bar{u},\bar{v}) \in U \times V$ such that

$$(2.1) \qquad J(u,\bar{v}) \geq J(\bar{u},\bar{v}) \geq J(\bar{u},v) \quad \text{for all } u \in U, v \in V.$$

The following facts are well known (cf. [32, Thm. 2], for example).

PROPOSITION 2.1. *It is always true that* $\inf(\mathcal{P}_0) \geq \sup(\mathcal{Q}_0)$. *Furthermore a pair* $(\bar{u},\bar{v})$ *is a saddlepoint of* J *on* $U \times V$ *if and only if* $\bar{u}$ *solves* $(\mathcal{P}_0)$, $\bar{v}$ *solves* $(\mathcal{Q}_0)$, *and* $\min(\mathcal{P}_0) = \max(\mathcal{Q}_0)$.

Here we use the notation that $\inf(\mathcal{P}_0)$ is the optimal value in $(\mathcal{P}_0)$, namely the infimum of f over U. We allow ourselves to write $\min(\mathcal{P}_0)$ in place of $\inf(\mathcal{P}_0)$ if the infimum is actually attained at some $\bar{u}$. Similarly for $\sup(\mathcal{Q}_0)$, $\max(\mathcal{Q}_0)$.

By finite-dimensional (*piecewise*) *linear-quadratic programming* in the general sense we shall mean the case of problems $(\mathcal{P}_0)$ and $(\mathcal{Q}_0)$ where U is a nonempty convex polyhedron in a space $\mathbf{R}^k$, V is a nonempty convex polyhedron in space $\mathbf{R}^\ell$, and J is a convex-concave function of the form

$$(2.2) \qquad J(u,v) = p\cdot u + v\cdot q + \tfrac{1}{2}u\cdot Pu - \tfrac{1}{2}v\cdot Qv - v\cdot Du,$$

where $p \in \mathbf{R}^k$, $q \in \mathbf{R}^\ell$, $P \in \mathbf{R}^{k\times k}$, $Q \in \mathbf{R}^{\ell\times\ell}$ and $D \in \mathbf{R}^{\ell\times k}$, with P and Q symmetric and positive *semidefinite*. When $P = 0$ and $Q = 0$, we speak of (*piecewise*) *linear programming* in the general sense. This includes classical linear programming, of course (cf. Example 3.1 below).

In the linear-quadratic programming case the objective functions in $(\mathcal{P}_0)$ and $(\mathcal{Q}_0)$ take the form

$$(2.3) \qquad f(u) = p\cdot u + \tfrac{1}{2}u\cdot Pu + \rho_{V,Q}(q - Du),$$

$$g(v) = q \cdot v - \tfrac{1}{2} v \cdot Qv - \rho_{U,P}(D^* v - p), \tag{2.4}$$

where

$$\rho_{V,Q}(s) = \sup_{v \in V} \{ s \cdot v - \tfrac{1}{2} v \cdot Qv \}, \tag{2.5}$$

$$\rho_{U,P}(r) = \sup_{u \in U} \{ r \cdot u - \tfrac{1}{2} u \cdot Pu \}. \tag{2.6}$$

When $P = 0$ and $Q = 0$, the functions $\rho_{V,Q}$ and $\rho_{U,P}$ reduce to the *support functions*

$$\sigma_V(s) = \sup_{v \in V} s \cdot v, \qquad \sigma_U(r) = \sup_{u \in U} r \cdot u. \tag{2.7}$$

The specific nature of these various expressions will be explored in the examples in §3. The central fact is that strong duality always holds for such problems.

THEOREM 2.2. *In the case where* $(\mathcal{P}_0)$ *and* $(\mathcal{Q}_0)$ *are finite-dimensional linear-quadratic programming problems in the general sense just described, one has*

$$\infty > \min(\mathcal{P}_0) = \max(\mathcal{Q}_0) > -\infty,$$

unless the optimal values $\inf(\mathcal{P}_0)$ *and* $\sup(\mathcal{Q}_0)$ *are both infinite. In particular, any finite-dimensional linear-quadratic programming problem with finite optimal value has an optimal solution.*

Theorem 2.2 can easily be derived from known results about quadratic programming in the standard sense, specifically the duality theorem of Dorn [33] and Cottle [34] and the existence criterion of Frank and Wolfe [35]. We have given the argument in full in [17, Thm. 2].

Incidentally, the suprema in (2.5) and (2.6) must be attained also, when finite. Indeed, these formulas give the optimal values in certain quadratic programming problems and are covered by the result just cited.

The sense in which the terminology "linear-quadratic programming in the general sense" is appropriate for the problems in Theorem 2.2 is elucidated by our next result.

PROPOSITION 2.3. *The function* $\rho_{V,Q}$ *is lower semicontinuous, convex, and piecewise linear-quadratic: its effective domain*

$$L = \{ s \in \mathbf{R}^\ell \mid \rho_{V,Q}(s) < \infty \} \tag{2.8}$$

is a nonempty convex polyhedron that can be decomposed into finitely many polyhedral convex sets, on each of which $\rho_{V,Q}$ *is quadratic (or linear).*

The same holds of course for $\rho_{U,P}$ *and its effective domain*

$$K = \{ r \in \mathbf{R}^k \mid \rho_{U,P}(r) < \infty \}. \tag{2.9}$$

Proof. Define

$$\varphi(v) = \begin{cases} \tfrac{1}{2} v \cdot Qv & \text{when } v \in V, \\ \infty & \text{when } v \notin V \end{cases} \tag{2.10}$$
$$= j_Q(v) + \delta_V(v),$$

where j_Q is the quadratic convex function corresponding to the positive definite form Q, and δ_V is the indicator of the convex polyhedron V:

$$\delta_V(v) = \begin{cases} 0 & \text{when } v \in V, \\ \infty & \text{when } v \notin V. \end{cases} \tag{2.11}$$

Clearly φ is convex, and its conjugate

$$\varphi^*(s) = \sup_{v \in \mathbf{R}^\ell} \{ s \cdot v - \varphi(s) \} \tag{2.12}$$

is given by

$$\varphi^*(s) = \rho_{V,Q}(s). \tag{2.13}$$

The latter is therefore lower semicontinuous and convex in s, and its effective domain L is a nonempty convex set (these properties being true for the conjugate of any proper convex function [36, §12]).

For each $s \in L$, the supremum in (2.12) (equivalently (2.5)) must actually be attained, as noted above. On the other hand we know from convex analysis [36, Thm. 23.5] that the supremum in (2.12) is attained at v if and only if $v \in \partial\varphi^*(s)$, which is equivalent to $s \in \partial\varphi(v)$. Thus L coincides with the effective domain of the subdifferential multifunction $\partial\varphi^*$, which is also the range of $\partial\varphi$. We shall use this fact to demonstrate that L is polyhedral and has the decomposition claimed.

Because $\varphi = j_Q + \delta_V$ and j_Q is finite everywhere on $\mathbf{R}^\ell$, we have by [36, Thm. 23.8] that

$$\partial\varphi(v) = \partial j_Q(v) + \partial\delta_V(v) = Qv + N_V(v), \tag{2.14}$$

where $N_V(v)$ is the normal cone to V at v [36, p. 215]. This normal cone is polyhedral, because V is polyhedral, and it depends only on the face of V to which v belongs. There are only finitely many faces of V, so it follows from (2.14) that $\partial\varphi$ is a polyhedral multifunction in the sense of Robinson [37], namely its graph in $\mathbf{R}^\ell \times \mathbf{R}^\ell$ is the union of finitely many polyhedral convex sets (one for each face of V). The same is then true for the multifunction $\partial\varphi^* = \partial\varphi^{-1}$, whose domain, already identified with L, must therefore be the projection of the union of finitely many polyhedral convex sets. We may conclude that the convex set L is actually polyhedral and can be decomposed into finitely many polyhedral convex sets L_i, over each of which the graph of $\partial\varphi^*$ is a polyhedral convex set. In the case of such a subset L_i having $\operatorname{int} L_i \neq \emptyset$, $\partial\varphi^*$ must by this reduce to a single-valued affine mapping on $\operatorname{int} L_i$, inasmuch as $\partial\varphi^*$ is single-valued almost everywhere on $\operatorname{int} L$ (a fact true of the subdifferential of any proper convex function on the interior of its effective domain [36, Thm. 24.5]). Therefore φ^* is quadratic (or linear) on $\operatorname{int} L_i$ by the lower semicontinuity of φ^*. For L_i with $\operatorname{int} L_i = \emptyset$, a slightly more general argument based on relative interiors of convex sets leads to the same conclusion. Thus the function $\varphi^* = \rho_{V,Q}$ is piecewise linear-quadratic as claimed. □

The terminology "linear programming in the general sense" in the case where $P = 0$ and $Q = 0$ is justified similarly. The functions $\rho_{V,Q}$ and $\rho_{U,P}$ reduce then to the support functions σ_V and σ_U in (2.7), which are polyhedral convex (piecewise linear) because U and V are polyhedral [36, Cor. 19.2.1].

Because $\rho_{V,Q}$ and $\rho_{U,P}$ can take ∞ as a value in some cases, the linear-quadratic programming problems $(\mathcal{P}_0)$ and $(\mathcal{Q}_0)$ may have implicit constraints. Thus in minimizing the function f given by (2.3) we are really interested only in the choices of u that satisfy

$$q - Du \in L \quad \text{as well as } u \in U. \tag{2.15}$$

Likewise in maximizing the function g in (2.4) we focus on v satisfying

$$D^* v - p \in K \quad \text{as well as } v \in V. \tag{2.16}$$

The polyhedral convexity of L and K in Proposition 2.3 together with that of U and V means that these constraint systems can be represented in principle by finitely many linear equations and inequalities.

A closer analysis of the sets L and K reveals additional structure that will be of use to us. Here we denote the *null space* of Q by

$$\operatorname{nl} Q = \{w \in \mathbf{R}^\ell \mid Qw = 0\}$$

and the *recession cone* [36, §8] of V by

$$\operatorname{rc} V = \{w \in \mathbf{R}^\ell \mid v + \lambda w \in V, \forall \lambda \geq 0\} \quad \text{for } v \in V.$$

The latter is the same regardless of the choice of $v \in V$. It is a polyhedral convex cone (always containing 0), because V is a polyhedral convex set [36, Thm. 19.5]. Indeed, if $V = \{v \mid Mv \leq m\}$, one has $\operatorname{rc} V = \{w \mid Mw \leq 0\}$. We denote the *polar* of a cone G as usual by

$$G^\circ = \{z \mid z \cdot w \leq 0, \forall w \in G\}. \tag{2.17}$$

PROPOSITION 2.4. *The effective domains L and K in Proposition 2.3 are the polar cones*

$$L = [\operatorname{rc} V \cap \operatorname{nl} Q]^\circ \quad \textit{and} \quad K = [\operatorname{rc} U \cap \operatorname{nl} P]^\circ. \tag{2.18}$$

Thus

$$L = \mathbf{R}^\ell \iff [\textit{the only } w \in \operatorname{rc} V \textit{ with } Qw = 0 \textit{ is } w = 0], \tag{2.19}$$

$$K = \mathbf{R}^k \iff [\textit{the only } z \in \operatorname{rc} U \textit{ with } Pz = 0 \textit{ is } z = 0]. \tag{2.20}$$

In particular $L = \mathbf{R}^\ell$ if V is bounded or if Q is positive definite, whereas $K = \mathbf{R}^k$ if U is bounded or if P is positive definite.

Proof. Let φ be given again by (2.10), so that $\varphi^* = \rho_{V,Q}$ as in (2.13). Since $L = \operatorname{dom} \varphi^*$ and L is closed, we have by [36, Thm. 13.3] that the indicator δ_L is conjugate to the recession function

$$(\operatorname{rc} \varphi)(w) = \lim_{\lambda \to \infty} \varphi(v + \lambda w)/\lambda,$$

where $v \in \operatorname{dom} \varphi = V$ (the limit being independent of the particular choice of v [36, Thm. 8.5]). The limit works out to

$$(\operatorname{rc} \varphi)(w) = \begin{cases} 0 & \text{if } w \in \operatorname{rc} V \text{ and } Qw = 0, \\ \infty & \text{otherwise}. \end{cases}$$

Thus $\operatorname{rc} \varphi = \delta_G$ for $G = \operatorname{rc} V \cap \operatorname{nl} Q$. The indicators δ_G and δ_L being conjugate to each other, we conclude that G and L are cones polar to each other [36, §14]. □

An important question of mathematical modeling and computation in applications both finite and infinite-dimensional is whether a problem $(\mathcal{P}_0)$, associated with a certain choice of J, U, and V, can reasonably be replaced by a more amenable problem $(\widehat{\mathcal{P}}_0)$ obtained in substituting for U and V a pair of smaller sets $\widehat{U}$ and $\widehat{V}$, e.g. bounded sets. The theorem we state next provides the answers for finite-dimensional linear-quadratic programming, although its full import will not be clear until the end of §3. It will be the basis for an infinite-dimensional generalization at the end of §6.

THEOREM 2.5. *Let $(\mathcal{P}_0)$ and $(\mathcal{Q}_0)$ be a pair of finite-dimensional linear-quadratic programming problems in the general sense. Consider also an auxiliary pair of such problems $(\widehat{\mathcal{P}}_0)$ and $(\widehat{\mathcal{Q}}_0)$ which corresponds to the same function J but subsets $\widehat{U} \subset U$ and $\widehat{V} \subset V$.*

(a) *If $\bar u$ and $\bar v$ are solutions to $(\mathcal{P}_0)$ and $(\mathcal{Q}_0)$ such that actually $\bar u \in \widehat{U}$ and $\bar v \in \widehat{V}$, then $\bar u$ and $\bar v$ are also solutions to $(\widehat{\mathcal{P}}_0)$ and $(\widehat{\mathcal{Q}}_0)$.*

(b) *Conversely, if $\bar u$ and $\bar v$ are solutions to $(\widehat{\mathcal{P}}_0)$ and $(\widehat{\mathcal{Q}}_0)$, and if U coincides with $\widehat{U}$ around $\bar u$ (i.e. $U \cap N = \widehat{U} \cap N$ for some neighborhood N of $\bar u$) and V coincides with $\widehat{V}$ around $\bar v$, then $\bar u$ and $\bar v$ are actually solutions to $(\mathcal{P}_0)$ and $(\mathcal{Q}_0)$.*

Proof. From Proposition 2.1 and Theorem 2.2 we know that $\bar u$ and $\bar v$ solve $(\mathcal{P}_0)$ and $(\mathcal{Q}_0)$ if and only if $(\bar u, \bar v)$ is a saddle point of J relative to $U \times V$. Likewise, $\bar u$ and $\bar v$ solve $(\widehat{\mathcal{P}}_0)$ and $(\widehat{\mathcal{Q}}_0)$ if and only if $(\bar u, \bar v)$ is a saddle point of J relative to $\widehat{U} \times \widehat{V}$. The former trivially implies the latter when $\widehat{U} = U$ and $\widehat{V} = V$, and this establishes (a). Under the assumptions in (b), $(\bar u, \bar v)$ is a saddle point relative to certain neighborhoods of $\bar u$ in U and $\bar v$ in V, i.e. it is a *local* saddle point relative to $U \times V$. But any local saddle point must be a global saddle point by the convexity-concavity of J. □

3. Basic models in linear-quadratic programming. The nature of the ρ functions appearing in the finite-dimensional linear-quadratic programming problems in §2 is revealed more clearly in the examples that follow. These examples illustrate various possibilities in formulation that one needs to appreciate in order to see the broad scope of the optimal control problems which will be introduced in §4.

Example 3.1. (*Classical linear programming.*) Let $P = 0$, $Q = 0$, $U = \mathbf{R}^k_+$, $V = \mathbf{R}^\ell_+$. Then

$$\rho_{V,Q}(s) = \sigma_{\mathbf{R}^\ell_+}(s) = \begin{cases} 0 & \text{if } s \leq 0, \\ \infty & \text{if } s \not\leq 0, \end{cases} \tag{3.1}$$

$$\rho_{U,P}(r) = \sigma_{\mathbf{R}^k_+}(r) = \begin{cases} 0 & \text{if } r \leq 0, \\ \infty & \text{if } r \not\leq 0. \end{cases} \tag{3.2}$$

It follows that in $(\mathcal{P}_0)$ we

$$\begin{array}{lll} \text{minimize} & p \cdot u & \\ \text{subject to} & Du \geq q, & u \geq 0, \end{array}$$

whereas in $(\mathcal{Q}_0)$ we

$$\begin{array}{lll} \text{maximize} & q \cdot v & \\ \text{subject to} & D^* v \leq p, & v \geq 0. \end{array}$$

Note the role of ∞ in (3.1) and (3.2) in representing constraints in these problems as discussed in connection with the sets L and K in Proposition 2.3.

Versions of linear programming that involve equality constraints or variables not restricted to be nonnegative correspond to other choices of U and V as polyhedral convex cones.

Example 3.2. (*Standard quadratic programming.*) Let $Q = 0$ (but $P \neq 0$) and take $U = \mathbf{R}^k$, $V = \mathbf{R}^\ell_+$. Then (2.8) holds, and in $(\mathcal{P}_0)$ we

$$\begin{array}{ll} \text{minimize} & p \cdot u + \frac{1}{2} u \cdot Pu \\ \text{subject to} & Du \geq q. \end{array}$$

This is quadratic programming in the traditional sense. To see what the dual is we must determine

$$\rho_{\mathbf{R}^k,P}(r) = \sup_{u \in \mathbf{R}^k} \{r \cdot u - \tfrac{1}{2} u \cdot Pu\}. \tag{3.3}$$

If P is positive definite, we easily calculate the supremum to be $\frac{1}{2}r \cdot P^{-1}r$, so that in $(\mathcal{Q}_0)$ we

$$\text{maximize} \quad q \cdot v - \tfrac{1}{2}[D^* v - p] \cdot P^{-1}[D^* v - p] \qquad \text{subject to} \quad v \geq 0.$$

If P is only positive semidefinite, the dualization is more subtle and is facilitated by an algebraic normalization. First we can decompose $U = \mathbb{R}^k$ into $U_1 \times U_2$, where $U_1 = \{u \mid Pu = 0\}$ and $U_2 = U_1^\perp$. Then by a change of coordinates if necessary we can actually suppose that $U_1 \times U_2 = \mathbb{R}^{k_1} \times \mathbb{R}^{k_2}$ for some $k_1 + k_2 = k$, so that $P = \text{diag}\,(P_1, 0)$ for an positive definite matrix $P_1 \in \mathbb{R}^{k_1 \times k_1}$. Writing $u = (u_1, u_2)$ with $u_1 \in \mathbb{R}^{k_1}, u_2 \in \mathbb{R}^{k_2}$, and correspondingly $r = (r_1, r_2)$ in (3.4), we calculate

$$\text{(3.4)} \qquad \begin{aligned} \rho_{U,P}(r_1, r_2) &= \sup_{u_1,u_2} \{r_1 \cdot u_1 + r_2 \cdot u_2 - \tfrac{1}{2}u_1 \cdot P_1 u_1\} \\ &= \begin{cases} \tfrac{1}{2}r_1 \cdot P_1^{-1} r_1 & \text{if } r_2 = 0, \\ \infty & \text{if } r_2 \neq 0. \end{cases} \end{aligned}$$

Also writing $p = (p_1, p_2)$ and $D = (D_1, D_2)$, we see that in $(\mathcal{P}_0)$ we

$$\text{minimize} \quad p_1 \cdot u_1 + p_2 \cdot u_2 + \tfrac{1}{2}u_1 \cdot P_1 u_1 \qquad \text{subject to} \quad D_1 u_1 + D_2 u_2 \geq q$$

whereas in $(\mathcal{Q}_0)$ we

$$\text{maximize} \quad q \cdot v - \tfrac{1}{2}[D_1^* v - p_1] \cdot P_1^{-1}[D_1^* v - p_1] \qquad \text{subject to} \quad D_2^* v = p_2, \quad v \geq 0.$$

Mixed systems of equality and inequality constraints can be handled by choosing $V = \mathbb{R}_+^{\ell_1} \times \mathbb{R}^{\ell_2}$ for some $\ell_1 + \ell_2 = \ell$.

With further algebra transformations it is possible actually to normalize the study of quadratic programming to the case where the matrix P is always *diagonal*. All one has to do is provide a factorization

$$\text{(3.5)} \qquad P = M^* M \quad \text{with} \quad M \in \mathbb{R}^{m \times k} \text{ for some dimension } m.$$

Then the problem $(\mathcal{P}_0)$ at the beginning of this example can be written as:

$$\text{minimize} \quad p \cdot u + 0 \cdot u' + \tfrac{1}{2}u' \cdot u' \quad \text{over all } (u, u') \in \mathbb{R}^k \times \mathbb{R}^m \qquad \text{satisfying} \quad Du + 0u' \geq q, \quad Mu - Iu' = 0.$$

This can be identified as a quadratic programming problem which can be written in terms of the enlarged vector (u, u') in the same format as the original $(\mathcal{P}_0)$, but with mixed equality and inequality constraints and a diagonalized quadratic form (actually with diagonal entries that are 0 for the components of u and 1 for the components of u').

Incidentally, some quadratic programming models can be set up more easily by taking advantage of the matrix Q instead of P. For example, the problem

$$\text{minimize} \quad \tfrac{1}{2}|Du - q|^2 \qquad \text{subject to} \quad u \geq 0,$$

where $|\cdot|$ is the Euclidean norm, can be regarded as the case of $(\mathcal{P}_0)$ where $p = 0$, $P = 0$, $Q = I$, $U = \mathbb{R}_+^k$, $V = \mathbb{R}^\ell$, inasmuch as

$$\text{(3.6)} \qquad \rho_{\mathbb{R}^\ell, I}(s) = \tfrac{1}{2}|s|^2.$$

The corresponding dual problem $(\mathcal{Q}_0)$ is:

$$\text{maximize} \quad q \cdot v - \tfrac{1}{2}|v|^2 \qquad \text{subject to} \quad D^* v \leq 0.$$

Example 3.3. (*Basic piecewise linear programming.*) Suppose $P = 0$, $Q = 0$. Let U be any convex polyhedron in $\mathbb{R}^k$ (expressible by some system of linear constraints which, for now, does not need to be specified), and let V be the *unit simplex* in $\mathbb{R}^\ell$:

$$\text{(3.7)} \qquad V = \{v \in \mathbb{R}_+^\ell \mid v \cdot \mathbf{1} = 1\} \quad \text{where } \mathbf{1} = (1, 1, \dots, 1).$$

Then

$$\text{(3.8)} \qquad \rho_{V,Q}(r) = \sigma_V(r) = \max_{i=1,\dots,\ell} r_i \quad \text{for } r = (r_1, \dots, r_\ell).$$

It follows that in $(\mathcal{P}_0)$ we

$$\text{minimize} \quad p \cdot u + \max_{i=1,\dots,\ell} \{q_i - d_i \cdot u\} \quad \text{over } u \in U,$$

where q_i is the ith component of q and d_i the ith row of D. The "max" expression in the objective in $(\mathcal{P}_0)$ is the pointwise maximum of a finite collection of affine functions of u and represents a general piecewise linear (i.e. polyhedral convex) function of u in the sense of [36, §19]. In the corresponding dual problem $(\mathcal{Q}_0)$ we

$$\text{maximize} \quad q \cdot v - \sigma_U(p - D^* v) \qquad \text{subject to} \quad v \geq 0, \quad v \cdot \mathbf{1} = 1,$$

where σ_U is the support function of U as in (2.7).

The constraint structure represented so far by the set U can be handled more directly under a different choice of notation. Still with $P = 0$ and $Q = 0$, simply take $U = \mathbb{R}^k$ but

$$V = \left\{ v \in \mathbb{R}_+^\ell \mid \sum_{i=1}^m v_i = 1 \right\} \quad \text{for an index } m \text{ satisfying } 1 < m < \ell,$$

where v_i is the ith component of v. This time

$$\text{(3.9)} \qquad \rho_{V,Q}(r) = \sigma_V(r) = \begin{cases} \max_{i=1,\dots,m} r_i & \text{if } r_{m+1} \geq 0, \dots, r_\ell \geq 0, \\ \infty & \text{otherwise.} \end{cases}$$

Then in $(\mathcal{P}_0)$ we

$$\text{minimize} \quad p \cdot u + \max_{i=1,\dots,m} \{q_i - d_i \cdot u\} \qquad \text{subject to} \quad d_i \cdot u \geq q_i \quad \text{for } i = m+1, \dots, \ell,$$

whereas in $(\mathcal{Q}_0)$ we

$$\text{maximize} \quad \sum_{i=1}^{\ell} q_i v_i \qquad \text{subject to} \quad v_i \geq 0 \quad \text{for } i = 1, \dots, \ell, \quad \sum_{i=1}^m v_i = 1, \quad \sum_{i=1}^{\ell} v_i d_i = p.$$

Example 3.4. (*Bounded linear programming.*) The linear programming problems in Example 3.1 are stated in terms of unbounded variables, but in practice this may not always be wise or convenient. Many linear programming codes ask the user to specify both upper and lower bounds for the vector u in the primal problem, say $\hat{u}^- \le u \le \hat{u}^+$. The effects on duality, however, are not widely appreciated. In fact there is reason impose upper and lower bounds on the dual variables too, say $\hat{v}^- \le v \le \hat{v}^+$. What this corresponds to is a representation of constraints in terms of *linear penalties*, like those in the currently popular ℓ_1 penalty function approach to nonlinear programming (cf. Fletcher [38]).

To be specific, suppose $P = 0$, $Q = 0$, and let U and V be vectorial intervals ("boxes") defined by upper and lower bounds:

$$U = [\hat{u}^-, \hat{u}^+], \qquad V = [\hat{v}^-, \hat{v}^+].$$

Adopting the notation

$$[s]_+ = \max\{0, s\}, \qquad [s]_- = \min\{0, s\} \tag{3.10}$$

in the vectorial sense, where the max is taken component by component (so that $s = [s]_+ + [s]_-$), we get

$$\rho_{V,Q}(s) = \sigma_V(s) = \max_{\hat{v}^- \le v \le \hat{v}^+} v \cdot s = \hat{v}^+ \cdot [s]_+ + \hat{v}^- \cdot [s]_-, \tag{3.11}$$

$$\rho_{U,P}(r) = \sigma_U(r) = \max_{\hat{u}^- \le u \le \hat{u}^+} u \cdot r = \hat{u}^+ \cdot [r]_+ + \hat{u}^- \cdot [r]_-. \tag{3.12}$$

It follows that in $(\mathcal{P}_0)$ we

$$\begin{aligned} &\text{minimize} && p \cdot u + \hat{v}^+ \cdot [q - Du]_+ + \hat{v}^- \cdot [q - Du]_- \\ &\text{subject to} && \hat{u}^- \le u \le \hat{u}^+, \end{aligned}$$

whereas in $(\mathcal{Q}_0)$ we

$$\begin{aligned} &\text{maximize} && q \cdot v + \hat{u}^+ \cdot [D^* v - p]_+ + \hat{u}^- \cdot [D^* v - p]_- \\ &\text{subject to} && \hat{v}^- \le v \le \hat{v}^+. \end{aligned}$$

Observe that these problems have piecewise linear objectives of a special kind. The optimal values are always finite, so optimal solutions always exist (Theorem 2.2).

Bounded linear programming in this sense may be a more natural vehicle in some applications than standard linear programming. Furthermore, problems in such a format can be solved directly, without reformulating them in the traditional way. Versions of the simplex method developed by Fourer [39] and the author [40, Chap. 11] can be used instead, for example.

Example 3.5. (*Bounded quadratic programming.*) This is a extension of the preceding example to allow for quadratic terms. Let

$$U = [\hat{u}^-, \hat{u}^+] \quad \text{and} \quad V = [\hat{v}^-, \hat{v}^+]$$

again, and take

$$P = \operatorname{diag}[\beta_1, \dots, \beta_k], \qquad Q = \operatorname{diag}[\gamma_1, \dots, \gamma_\ell],$$

where $\beta_j \ge 0$, $\gamma_i \ge 0$. (The assumption of a diagonal form for P and Q does not entail the loss of generality that might be imagined; cf. Example 3.2.) The calculation of the ρ functions (2.5) and (2.6) decomposes into one-dimensional calculations of the form

$$\max_{\alpha \in [\alpha^-, \alpha^+]} \{\tau\alpha - \tfrac{1}{2}\lambda\alpha^2\} \tag{3.13}$$

for various intervals $[\alpha^-, \alpha^+]$ and constants $\lambda \ge 0$. The maximum value in (3.13) is a function of $\tau \in \mathbf{R}$ that depends on the parameters α^-, α^+ and λ, and it is given by

$$\theta(\tau; \alpha^-, \alpha^+, \lambda) = \begin{cases} (2\tau - \lambda\alpha^+)(\alpha^+/2) & \text{when } \tau \ge \lambda\alpha^+, \\ (1/2\lambda)\tau^2 & \text{when } \lambda\alpha^- \le \tau \le \lambda\alpha^+, \\ (2\tau - \lambda\alpha^-)(\alpha^-/2) & \text{when } \tau \le \lambda\alpha^-. \end{cases} \tag{3.14}$$

Despite its formula, this function of τ has a simple form and a natural meaning. In the case where $\lambda = 0$, it vanishes at $\tau = 0$, is linear with slope α_+ for $\tau > 0$ and linear with slope α_- for $\tau < 0$. In the case where $\lambda > 0$, it has a similar structure but with a quadratic interpolation instead of a "corner." Indeed, it is the unique *smooth* function whose values are given by $\alpha^+\tau$ + const. for τ sufficiently high, by $\alpha^-\tau$ + const. for τ sufficiently low, and by $(1/2\lambda)\tau^2$ on the interval between.

With this notation, and denoting the components of p, q, and D by p_j, q_i, d_{ij}, and so forth, we can express the primal and dual problems as follows. In $(\mathcal{P}_0)$ we

$$\begin{aligned} &\text{minimize} && \sum_{j=1}^{k} [p_j u_j + \tfrac{1}{2}\beta_j u_j^2] + \sum_{i=1}^{\ell} \theta\left(q_i - \sum_{j=1}^{k} d_{ij} u_j; \hat{v}_i^-, \hat{v}_i^+, \gamma_i\right) \\ &\text{subject to} && \hat{u}_j^- \le u_j \le \hat{u}_j^+ \quad \text{for } j = 1, \dots, k, \end{aligned}$$

whereas in $(\mathcal{Q}_0)$ we

$$\begin{aligned} &\text{maximize} && \sum_{i=1}^{\ell} [q_i v_i - \tfrac{1}{2}\gamma_i v_i^2] - \sum_{j=1}^{k} \theta\left(\sum_{i=1}^{\ell} v_i d_{ij} - p_j; \hat{u}_j^-, \hat{u}_j^+, \beta_j\right) \\ &\text{subject to} && \hat{v}_i^- \le v_i \le \hat{v}_i^+ \quad \text{for } i = 1, \dots, \ell. \end{aligned}$$

When $\beta_j = 0$, $\gamma_i = 0$, these problems reduce to the bounded linear case in Example 3.4. They are useful in modeling situations where constraints are not necessarily sharp, as in stochastic programming (see Rockafellar and Wets [18] and King et al. [41]). Thus for instance if $[\hat{v}_i^-, \hat{v}_i^+] = [0, \alpha_i^+]$ the corresponding θ term in $(\mathcal{P}_0)$ imposes no penalty if the putative constraint $\sum_{j=1}^{k} d_{ij} u_j \ge q_i$ is satisfied, a slight penalty at a marginal cost that grows linearly (at the rate $1/\gamma_i$) from 0 as this constraint begins to be violated, and eventually for large violations a penalty with constant marginal cost α_i^+.

Of course it is also possible to get versions of these problems in which the penalty expressions do not eventually become linear but stay quadratic for arbitrarily large violations. These correspond to limiting cases of $\theta(\tau; \alpha^-, \alpha^+, \lambda)$ where $\alpha^- = -\infty$ or $\alpha^+ = \infty$, or both. They can be obtained by taking U and V not to be "boxes" but orthants or products of orthants and subspaces, as in Example 2.3. □

In understanding the relationship between penalty models such as Examples 3.4 and 3.5 and the more traditional models without penalties, such as Examples 3.1 and 3.2, the facts in Theorem 2.5 are essential. As an illustration of the way Theorem 2.5 can be employed, let us look again at the standard linear programming problems in Example 3.1. Suppose we know that an optimal solution $\bar{u}$ to $(\mathcal{P}_0)$ will exist within certain upper bounds, say $\bar{u} \le \hat{u}$, and also that a dual optimal solution $\hat{v}$ to $(\mathcal{Q}_0)$ will exist within certain upper bounds, say $\bar{v} \le \hat{v}$. Then according to Theorem 2.5(a), $\bar{u}$ and $\bar{v}$ can be found by solving, instead of the given problems, the bounded linear programming problems in Example 3.4 with

$$U = [\hat{u}^-, \hat{u}^+] = [0, \hat{u}], \qquad V = [\hat{v}^-, \hat{v}^+] = [0, \hat{v}]. \tag{3.15}$$

The idea here that dual bounds can be given along with primal bounds is not so far-fetched as it might seem. The components of a dual optimal solution $\bar{v}$ often have interpretation as marginal prices, or as rates of change with respect to certain perturbations of constraints. Economic limitations or experience may dictate appropriate bounds. Anyway, there is no great harm in going ahead with solving the bounded versions of the problems in terms of *estimated* bounds $\hat{u}$ and $\hat{v}$. If solutions $\bar{u}$ and $\bar{v}$ are obtained for which the upper bounds are not tight, then $\bar{u}$ and $\bar{v}$ actually solve the original problems, according to Theorem 2.5(b). If the upper bounds are tight in some components, they can be loosened and the procedure repeated.

4. Intertemporal linear-quadratic programming. The general problems of optimal control that are the main object of our study can now be formulated. The time interval $[t_0, t_1]$ is fixed. The primal problem is:

$(\mathcal{P})$ Over the dynamical system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + b(t) \quad \text{a.e.}, \qquad x(t_0) = B_e u_e + b_e,$$

with control space

$$\mathcal{U} = \{(u, u_e) \mid u \in \mathcal{L}^1, u(t) \in U(t) \text{ a.e.}, \quad u_e \in U_e\}$$

minimize the functional

$$\mathcal{F}(u, u_e) = \int_{t_0}^{t_1} [p(t) \cdot u(t) + \tfrac{1}{2} u(t) \cdot P(t)u(t) - c(t) \cdot x(t)]dt + [p_e \cdot u_e + \tfrac{1}{2} u_e \cdot P_e u_e - c_e \cdot x(t_1)]$$
$$+ \int_{t_0}^{t_1} \rho_{V(t),Q(t)}(q(t) - C(t)x(t) - D(t)u(t)) + \rho_{V_e,Q_e}(q_e - C_e x(t_1) - D_e u_e).$$

The dual problem is:

$(\mathcal{Q})$ Over the dynamical system

$$-\dot{y}(t) = A^*(t)y(t) + C^*(t)v(t) + c(t) \text{ a.e.}, \qquad y(t_1) = C_e^* v_e + c_e,$$

with control space

$$\mathcal{V} = \{(v, v_e) \mid v \in \mathcal{L}^1, v(t) \in V(t) \text{ a.e.}, \quad v_e \in V_e\},$$

maximize the functional

$$\mathcal{G}(v, v_e) = \int_{t_0}^{t_1} [q(t) \cdot v(t) - \tfrac{1}{2} v(t) \cdot Q(t)v(t) - b(t) \cdot y(t)]dt + [q_e \cdot v_e - \tfrac{1}{2} v_e \cdot Q_e v_e - b_e \cdot y(t_0)]$$
$$- \int_{t_0}^{t_1} \rho_{U(t),P(t)}(B^*(t)y(t) + D^*(t)v(t) - p(t))dt - \rho_{V_e,P_e}(B_e^* y(t_0) + D_e^* v_e - p_e).$$

Here

$$u(t) \in \mathbf{R}^k, \quad u_e \in \mathbf{R}^{k_e}, \quad x(t) \in \mathbf{R}^n, \quad v(t) \in \mathbf{R}^\ell, \quad v_e \in \mathbf{R}^{\ell_e}, \quad y(t) \in \mathbf{R}^n,$$

and dimensions of the other elements are determined accordingly. The matrices $P(t)$, P_e, $Q(t)$ and Q_e are assumed to be symmetric and positive *semidefinite* (possibly 0). The sets $U(t) = \mathbf{R}^k$, $U_e = \mathbf{R}^{k_e}$, $V(t) \subset \mathbf{R}^\ell$ and $V_e \subset \mathbf{R}^{\ell_e}$ are assumed to be polyhedral

convex. The ρ terms are defined by (2.5) and (2.6). In general they are piecewise linear-quadratic convex functions that may take on the value ∞; cf. Proposition 2.3. Various cases based in part on the finite-dimensional models in §3 will be viewed in §5. First we must clarify our technical foundations.

All the data elements in problems $(\mathcal{P})$ and $(\mathcal{Q})$, namely

$$A(t), B(t), C(t), D(t), b(t), c(t), P(t), Q(t), p(t), q(t), U(t), V(t),$$

are assumed to depend *continuously* on t. For the sets $U(t)$ and $V(t)$ this means continuity with respect to the usual notions of convergence of subsets of Euclidean space that are not necessarily bounded; see Salinetti and Wets [42] for an exposition of the convex case. Thus the multifunctions $t \mapsto U(t)$ and $t \mapsto V(t)$ should be lower semicontinuous and of closed graph. Lower semicontinuity of $t \mapsto U(t)$ implies that the multifunction $t \mapsto \text{int}\, U(t)$ is of open graph; indeed, by virtue of the convexity of $U(t)$, lower semicontinuity is equivalent to the latter property if $\text{int}\, U(t) \neq \emptyset$ for all $t \in [t_0, t_1]$ (Rockafellar [43, p. 458]). A special case of continuous dependence, of course, is the one where $U(t)$ and $V(t)$ are constant with respect to t.

Under these assumptions the dynamical systems in $(\mathcal{P})$ and $(\mathcal{Q})$ are well defined with respect to the control spaces $\mathcal{U}$ and $\mathcal{V}$. They determine unique absolutely continuous functions x and y from $[t_0, t_1]$ to $\mathbf{R}^n$.

In showing that the integrals in the objective functionals in $(\mathcal{P})$ and $(\mathcal{Q})$ are well defined too, we shall make use of the following.

PROPOSITION 4.1. *The expression $\rho_{V(t),Q(t)}(s)$ is lower semicontinuous jointly in t and s, in fact continuous relative to $\{(t,s) \mid s \in \text{int}\, L(t)\}$, where*

$$L(t) = \{s \in \mathbf{R}^\ell \mid \rho_{V(t),Q(t)}(s) < \infty\}. \tag{4.1}$$

Moreover $L(t)$ depends lower semicontinously on t.

The same holds for the expression $\rho_{U(t),P(t)}(r)$ and the effective domain

$$K(t) = \{r \in \mathbf{R}^k \mid \rho_{U(t),P(t)}(r) < \infty\}. \tag{4.2}$$

Proof. Our argument is based on showing that the function $\rho_{V(t),Q(t)}$ depends *epicontinuously* on t, i.e. its epigraph set

$$E(t) = \{(s, \alpha) \in \mathbf{R}^\ell \times \mathbf{R} \mid \rho_{V(t),Q(t)}(s) \le \alpha\}, \tag{4.3}$$

which is convex, depends continuously on t. Epicontinuity corresponds to a notion of function convergence first considered by Wijsman [44] and subsequently developed by others; see Wets [45]. It yields all the properties claimed. Indeed, if the multifunction $t \mapsto E(t)$ is continuous, then by definition it is lower semicontinuous and of closed graph. The closed graph property is equivalent to the lower semicontinuity of the function

$$(t, s) \mapsto \rho_{V(t),Q(t)}(s). \tag{4.4}$$

The lower semicontinuity of $t \mapsto E(t)$ implies from its definition the lower semicontinuity of the domain multifunction $t \mapsto L(t)$, since $L(t)$ is the projection in $\mathbf{R}^\ell$ of the epigraph (4.3). (Recall from Proposition 2.3 that $L(t)$ is a *closed* convex set, since it is polyhedral.) The multifunction $t \mapsto \text{int}\, L(t)$ is then of open graph, as cited above, i.e. the set $\{(t,s) \mid s \in \text{int}\, L(t)\}$ is open in the space $[t_0, t_1] \times \mathbf{R}^\ell$. The upper semicontinuity of $\rho_{V(t),Q(t)}(s)$ on this open set follows then from the lower semicontinuity of $t \mapsto E(t)$ again and the corresponding openness of $\{(t, s, \alpha) \mid (s, \alpha) \in \text{int}\, E(t)\}$, and with the lower semicontinuity noted earlier for (4.4) one gets continuity.

To prove that $\rho_{V(t),Q(t)}$ depends epicontinuously on t, we resort to the notation of Proposition 2.3, where now, however, everything depends on t. We identify $\rho_{V(t),Q(t)}$ with the conjugate φ_t^* of the convex function $\varphi_t = j_{Q(t)} + \delta_{V(t)}$, where $\delta_{V(t)}$ is the indicator of $V(t)$ and

$$j_{Q(t)}(v) = \tfrac{1}{2}v \cdot Q(t)v. \tag{4.5}$$

Trivially $\delta_{V(t)}$ depends epicontinuously on t, since its epigraph is just $V(t) \times \mathbf{R}_+$. Furthermore the convex function $j_{Q(t)}$ is finite everywhere on $\mathbf{R}^\ell$, and its values depend continuously on t because $Q(t)$ depends continuously on t. This implies by Wets [45, p. 392] that $j_{Q(t)}$ depends epicontinuously on t and by McLinden and Bergstrom [46, Thm. 6] that the sum $\varphi_t = j_{Q(t)} + \delta_{V(t)}$ depends epicontinuously on t. The operation of passing to the conjugate of a convex function is known to preserve epicontinuity (Wijsman [44]), so we may conclude that the function $\rho_{V(t),Q(t)} = \varphi_t^*$ does depend epicontinuously on t, as claimed. □

THEOREM 4.2. *In problem $(\mathcal{P})$ the control space $\mathcal{U}$ is a nonempty closed convex subset of $\mathcal{L}^1([t_0,t_1],\mathbf{R}^k) \times \mathbf{R}^{k_e}$, and the objective functional $\mathcal{F}$ is well defined, lower semicontinuous and convex, with values that are finite or ∞.*

Likewise, in problem $(\mathcal{Q})$ the control space $\mathcal{V}$ is a nonempty closed convex subset of $\mathcal{L}^1([t_0,t_1],\mathbf{R}^\ell) \times \mathbf{R}^{\ell_e}$, and the objective functional $\mathcal{G}$ is well defined, upper semicontinuous and concave, with values that are finite or $-\infty$.

Proof. Only the first half has to be argued; the second half is parallel. The convexity and closedness of $\mathcal{U}$ is obvious from the convexity and closedness of the sets $U(t)$ and U_e. The nonemptiness of $\mathcal{U}$ comes from the nonemptiness of $U(t)$ and U_e and the continuity of $t \mapsto U(t)$: the selection theorem of Michael [47] asserts that any lower semicontinuous multifunction from $[t_0,t_1]$ to $\mathbf{R}^k$ with nonempty closed convex values has a continuous selection. Thus there actually exist pairs (u,u_e) in $\mathcal{U}$ with u continuous rather than just $\mathcal{L}^1$.

The mapping $(u,u_e) \mapsto x$ from $\mathcal{L}^1([t_0,t_1],\mathbf{R}^k) \times \mathbf{R}^{k_e}$ into $\mathcal{C}([t_0,t_1],\mathbf{R}^n)$ is affine and continuous, even compact:

$$x(t) = M(t)\Big(B_e u_e + b_e + \int_{t_0}^{t} M(\tau)^{-1}[B(\tau)u(\tau) + b(\tau)]d\tau\Big), \tag{4.6}$$

where $M(t)$ is the matrix with the property that $\xi(t) = M(t)x_0$ is the solution to $\dot\xi(t) = A(t)\xi(t)$, $\xi(t_0) = x_0$. The terms

$$\int_{t_0}^{t_1} [p(t)\cdot u(t) - c(t)\cdot x(t)]dt + [p_e \cdot u_e - c_e \cdot x(t_1)]$$

in $\mathcal{F}(u,u_e)$ therefore give a continuous, affine functional of (u,u_e). The mapping that takes a pair (u,u_e) in $\mathcal{L}^1([t_0,t_1],\mathbf{R}^k) \times \mathbf{R}^{k_e}$ into the pair (s,s_e) in $\mathcal{L}^1([t_0,t_1],\mathbf{R}^\ell) \times \mathbf{R}^{\ell_e}$ given by

$$s(t) = q(t) - C(t)x(t) - D(t)u(t), \qquad s_e = q_e - C_e x(t_1) - D_e u_e, \tag{4.7}$$

is affine and continuous too.

It remains only to show that the expressions

$$I_1(u,u_e) = \int_{t_0}^{t_1} u(t)\cdot P(t)u(t)dt + u_e \cdot P_e u_e,$$

$$I_2(s,s_e) = \int_{t_0}^{t_1} \rho_{V(t),Q(t)}(s(t))dt + \rho_{V_e,Q_e}(s_e)$$

give well defined, lower semicontinuous, convex functionals on $\mathcal{L}^1([t_0,t_1],\mathbf{R}^k) \times \mathbf{R}^{k_e}$ and $\mathcal{L}^1([t_0,t_1],\mathbf{R}^\ell) \times \mathbf{R}^{\ell_e}$ respectively, with values that are finite or ∞. Certainly the continuity of $P(t)$ in t and the lower semicontinuity of $\rho_{V(t),Q(t)}(s)$ jointly in t and s proved in Proposition 4.1) ensure that the integrands for I_1 and I_2 are measurable in t.

All the terms in the formula for I_1 are nonnegative and convex, because $P(t)$ and P_e are positive semidefinite. Therefore I_1 is a well defined convex functional with values in $[0,\infty]$. Its lower semicontinuity follows from Fatou's lemma, since every norm-convergent sequence in $\mathcal{L}^1([t_0,t_1],\mathbf{R}^k)$ has a subsequence that converges pointwise almost everywhere.

The argument for I_2 is the same, after a normalization. We showed at the outset of this proof that $\mathcal{U}$ contains a pair (u,u_e) with u actually continuous. The same applies to $\mathcal{V}$. Taking (v,v_e) to be such a pair in $\mathcal{V}$ and observing from the definition of the ρ functions that then

$$\rho_{V(t),Q(t)}(s) \geq s(t)\cdot v(t), \qquad \rho_{V_e,Q_e}(s_e) \geq s_e \cdot v_e,$$

we can write

$$I_2(s,s_e) = I_3(s,s_e) + \int_{t_0}^{t_1} s(t)\cdot v(t)dt + s_e \cdot v_e,$$

where

$$I_3(s,s_e) = \int_{t_0}^{t_1} [\rho_{V(t),Q(t)}(s(t)) - s(t)\cdot v(t)]dt + [\rho_{V_e,Q_e}(s_e) - s_e \cdot v_e].$$

Thus I_2 differs by only a continuous linear functional from a functional I_3 whose terms are all convex and nonnegative. As with I_1 we can see that I_3 is well defined with values in $[0,\infty]$ and is convex and lower semicontinuous. Therefore I_2 has these required properties, except that its values will generally be in $(-\infty,\infty]$. □

It is evident that in the minimization in $(\mathcal{P})$ we are really interested only in the controls $(u,u_e) \in \mathcal{U}$ yielding $\mathcal{F}(u,u_e) < \infty$. Such controls have to satisfy

$$q(t) - C(t)x(t) - D(t)u(t) \in L(t) \text{ a.e.} \quad \text{and} \quad q_e - C_e x(t_1) - D_e u_e \in L_e, \tag{4.8}$$

where $L(t)$ and L_e are the effective domains of $\rho_{V(t),Q(t)}$ and ρ_{V_e,Q_e} (cf. Proposition 2.3). Similarly, in the maximization in $(\mathcal{Q})$ we are really interested only in the controls (v,v_e) yielding $\mathcal{G}(v,v_e) > -\infty$, and these have to satisfy

$$B^*(t)y(t) + D^*(t)v(t) - p(t) \in K(t) \text{ a.e.} \quad \text{and} \quad B_e^* y(t_0) + D_e^* v_e - p_e \in K_e, \tag{4.9}$$

where $K(t)$ and K_e are the effective domains of $\rho_{U(t),P(t)}$ and ρ_{U_e,P_e}. These implicit constraints can be regarded as "linear," incidentally, since the sets $L(t)$, L_e, $K(t)$ and K_e are polyhedral convex cones (Propositions 2.3 and 2.4).

As stated in §1, our approach in this paper to such implicit constraints involving the states $x(t)$ and $y(t)$ is to skirt them when convenient by adopting alternative problem formulations where they have no force, specifically because $L(t)$ and L_e are all of $\mathbf{R}^\ell$ and $\mathbf{R}^{\ell_e}$, or $K(t)$ and K_e are all of $\mathbf{R}^k$ and $\mathbf{R}^{k_e}$. Accordingly the following type of assumption will sometimes be of importance to us.

We shall say that the *primal finiteness condition* is satisfied if the functions $\rho_{V(t),Q(t)}$ and ρ_{V_e,Q_e} are finite everywhere (i.e. $L(t) = \mathbf{R}^\ell$ and $L_e = \mathbf{R}^{\ell_e}$). Likewise, the *dual finiteness condition* is satisfied if the functions $\rho_{U(t),P(t)}$ and ρ_{U_e,P_e} are finite everywhere (i.e. $K(t) = \mathbf{R}^k$ and $K_e = \mathbf{R}^{k_e}$). Criteria for this are furnished by Proposition 2.4.

PROPOSITION 4.3. *If the primal finiteness condition is satisfied, then* $\mathcal{F}(u, u_e)$ *in* $(\mathcal{P})$ *is finite for all* $(u, u_e) \in \mathcal{U}$ *with* $u \in \mathcal{L}^\infty$.

Likewise, if the dual finiteness condition is satisfied, then $\mathcal{G}(v, v_e)$ *in* $(\mathcal{Q})$ *is finite for all* $(v, v_e) \in \mathcal{V}$ *with* $v \in \mathcal{L}^\infty$.

Proof. Under the primal finiteness condition the convex functions $\rho_{V(t),Q(t)}$ and ρ_{V_e,Q_e} are finite on $\mathbf{R}^\ell$ and $\mathbf{R}^{\ell_e}$ and therefore continuous on these spaces, inasmuch as a convex function on a finite-dimensional space is continuous on any open set where it is finite [36, §10]. Moreover $\rho_{V(t),Q(t)}(s)$ is continuous jointly in t and s by Proposition 4.1 and consequently is bounded above and below on $[t_0, t_1] \times S$ for any bounded subset $S \subset \mathbf{R}^\ell$. For the function $s(t)$ in (4.7), then, the expression $\rho_{V(t),Q(t)}(s(t))$ is $\mathcal{L}^\infty$ in t when $u(t)$ is $\mathcal{L}^\infty$ in t, as is the expression $u(t) \cdot P(t)u(t)$. All the integrals in the formula for $\mathcal{F}(u, u_e)$ are therefore finite when $u \in \mathcal{L}^\infty$. The argument for $\mathcal{G}(v, v_e)$ under the dual finiteness condition runs the same way. □

The reader may wonder why we have formulated problems $(\mathcal{P})$ and $(\mathcal{Q})$ with control spaces involving $\mathcal{L}^1$ rather than $\mathcal{L}^\infty$. Matters would be simpler in some respects with $\mathcal{L}^\infty$, and for applications $\mathcal{L}^\infty$ is apparently more natural. The work done in continuous-time programming uses $\mathcal{L}^\infty$ too. Of course, our problems include the $\mathcal{L}^\infty$ case by simple restriction. The real reason for taking $\mathcal{L}^1$, however, is not extra generality but the need for allowing ample controls in order to close a possible duality gap between $(\mathcal{P})$ and $(\mathcal{Q})$. The payoff will come in our result on strong duality, Theorem 6.3.

5. Special cases of the optimal control models. Our task now is to illuminate the scope of the problems $(\mathcal{P})$ and $(\mathcal{Q})$ introduced in §4. We explain how they cover the linear programming models $(\mathcal{P}_1)$ and $(\mathcal{Q}_1)$ in §1 and much more.

The treatment of endpoints $x(t_0)$ and $x(t_1)$ in $(\mathcal{P})$ and $y(t_0)$ and $y(t_1)$ in $(\mathcal{Q})$ departs from the traditional patterns in the literature on optimal control. We therefore begin by considering various important cases embedded in our formulation and the way they come to be dualized.

Example 5.1. (*Problems with fixed endpoints.*) How can one represent in terms of the endpoint provisions in the structure of $(\mathcal{P})$ a problem in which an integral

(5.1)

$$\int_{t_0}^{t_1} [p(t) \cdot u(t) + \tfrac{1}{2}u(t) \cdot P(t)u(t) - c(t) \cdot x(t) + \rho_{V(t),Q(t)}(q(t) - C(t)x(t) - D(t)u(t))]dt$$

is minimized over all pairs x, u, satisfying $u(t) \in U(t)$ a.e., $u \in \mathcal{L}^1$,

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + b(t) \quad \text{a.e.}, \qquad x(t_0) = a_0, \quad x(t_1) = a_1, \tag{5.2}$$

where a_0 and a_1 are fixed points in $\mathbf{R}^n$? The requirement $x(t_0) = a_0$ can be handled by setting $b_e = a_0$ and trivializing the u_e vector by taking $\mathbf{R}^{k_e}$ to be *zero-dimensional* (so $U_e = \{0\}$, $B_e = 0$, $D_e = 0$, $p_e = 0$, $P_e = 0$). Only the term

$$\rho_{V_e,Q_e}(q_e - C_e x(t_1)) - c_e \cdot x(t_1) \tag{5.3}$$

remains then in the endpoint expression for $(\mathcal{P})$. This can be made to represent the requirement $x(t_1) = a_1$ as follows. First choose $V_e = \mathbf{R}^n$ and $Q_e = 0$, so that

$$\rho_{V_e,Q_e}(s_e) = \sigma_{\mathbf{R}^n}(s_e) = \begin{cases} 0 & \text{if } s_e = 0, \\ \infty & \text{if } s_e \neq 0. \end{cases} \tag{5.4}$$

Then

$$\rho_{V_e,Q_e}(q_e - C_e x(t_1)) = \begin{cases} 0 & \text{if } C_e x(t_1) = q_e, \\ \infty & \text{if } C_e x(t_1) \neq q_e. \end{cases} \tag{5.5}$$

Now all one has to do is take $C_e = I$, $q_e = a_1$, $c_e = 0$.

Note that the dual problem $(\mathcal{Q})$ in this case has as its endpoint term

$$q_e \cdot v_e - \tfrac{1}{2}v_e \cdot Q_e v_e - b_e \cdot y(t_0) - \rho_{V_e,P_e}(B_e^* y(t_0) + D_e^* v_e - p_e) = a_1 \cdot y(t_1) - a_0 \cdot y(t_0). \tag{5.6}$$

In $(\mathcal{Q})$, therefore, one maximizes

(5.7)

$$\int_{t_0}^{t_1} [q(t) \cdot v(t) - \tfrac{1}{2}v(t) \cdot Q(t)v(t) - b(t) \cdot y(t) - \rho_{U(t),P(t)}(B^*(t)y(t) + D^*(t)v(t) - p(t))]dt$$
$$+ a_1 \cdot y(t_1) - a_0 \cdot y(t_0)$$

over all pairs y, v, such that $v(t) \in V(t)$ a.e., $v \in \mathcal{L}^1$, and

$$-\dot{y}(t) = A^*(t)y(t) + C^*(t)v(t) + c(t) \quad \text{a.e.} \tag{5.8}$$

(with no restriction on the endpoints $y(t_0)$ and $y(t_1)$).

Of course one can stop with (5.4), (5.5), and have in place of $x(t_1) = a_1$ the more general constraint $C_e x(t_1) = q_e$ for some matrix C_e and vector q_e. In $(\mathcal{Q})$ this would correspond to replacing the term $a_1 \cdot y(t_1)$ in (5.7) by $q_e \cdot v_e$, where v_e is unrestricted but $y(t_1) = C_e^* v_e$ in (5.8) (if $c_e = 0$ still).

If we only want $x(t_0) = a_0$ in (5.2), so that $x(t_1)$ is a free endpoint in $(\mathcal{P})$, and correspondingly want to incorporate a term $-d_1 \cdot x(t_1)$ in the objective (5.1), we can represent this by trivializing the vector v_e too, i.e. by taking $\mathbf{R}^{\ell_e}$ to be zero-dimensional (so that $V_e = \{0\}$, $C_e = 0$, $q_e = 0$, $Q_e = 0$), and setting $c_e = d_1$. Then the term (5.3) reduces to $-d_1 \cdot x(t_1)$. In the corresponding version of $(\mathcal{Q})$ the term $a_1 \cdot y(t_1)$ drops from (5.7) but $y(t_1) = d_1$ is added to (5.8). Thus $(\mathcal{Q})$ is a problem of the same type but with $y(t_1)$ fixed and $y(t_0)$ free.

Example 5.2. (*General linear constraints on endpoints.*) Instead of fixed endpoints let us consider a much more general case where the functional (5.1) is to be minimized over all pairs x, u, satisfying $u(t) \in U(t)$ a.e., $u \in \mathcal{L}^1$,

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + b(t) \quad \text{a.e.}$$

and a constraint system of the form

$$A_0 x(t_0) + A_1 x(t_1) \geq a \tag{5.9}$$

on the endpoints, with $a \in \mathbf{R}^d$. This can be placed in the form of $(\mathcal{P})$ by choosing $B_e = I$ and $b_e = 0$ (so that $x(t_0) = u_e$ in $(\mathcal{P})$) and then setting $U_e = \mathbf{R}^n$, $D_e = A_0$, $C_e = A_1$, $q_e = a$, $V_e = \mathbf{R}^d_+$, $Q_e = 0$. Then

$$\rho_{V_e,Q_e}(q_e - C_e x(t_1) - D_e u_e) = \begin{cases} 0 & \text{if (5.9) holds}, \\ \infty & \text{otherwise}. \end{cases} \tag{5.10}$$

Taking $p_e = 0$, $P_e = 0$, $c_e = 0$, we get all the endpoint terms other than (5.9) to drop out, and $(\mathcal{P})$ then represents the problem as specified.

The corresponding dual problem $(\mathcal{Q})$ maximizes

$$\int_{t_0}^{t_1} [q(t) \cdot v(t) - \tfrac{1}{2}v(t) \cdot Q(t)v(t) - b(t) \cdot y(t) - \rho_{U(t),P(t)}(B^*(t)y(t) + D^*(t)v(t) - p(t))]dt + a \cdot v_e$$

over all y, v, v_e satisfying $v(t) \in V(t)$ a.e., $v \in \mathcal{L}^\infty$, $v_e \in \mathbb{R}^d_+$,

$$(5.11)\quad -\dot{y}(t) = A^*(t)y(t) + C^*(t)v(t) + c(t) \text{ a.e.,} \qquad y(t_0) = -A_0^* v_e, \quad y(t_1) = A_1^* v_e.$$

Obviously the inequality in (5.9) can be converted to an equation by taking $V_e = \mathbb{R}^d$ instead of $\mathbb{R}^d_+$. For a particularly interesting case of this, let $A_0 = -I$, $A_1 = I$, $a = 0$. Then (5.9) reduces to the requirement that $x(t_0) = x(t_1)$, and the endpoint conditions in (5.11) reduce correspondingly to $y(t_0) = y(t_1)$ ("periodic" boundary conditions).

Example 5.3. (*Basic intertemporal linear programming.*) Problems $(\mathcal{P})$ and $(\mathcal{Q})$ turn into the basic linear programming models $(\mathcal{P}_1)$ and $(\mathcal{Q}_1)$ described in §1 when $P(t), P_e, Q(t)$ and Q_e are zero matrices and

$$(5.12)\qquad U(t) = \mathbb{R}^k_+, \quad U_e = \mathbb{R}^{k_e}_+, \quad V(t) = \mathbb{R}^\ell_+, \quad V_e = \mathbb{R}^{\ell_e}_+$$

in the pattern of Example 3.1. By choosing products of orthants and subspaces in (5.12) instead of merely orthants, one obtains the versions of these problems having a mixture of equality and inequality constraints. Neither the primal nor the dual finiteness condition (as defined in the last section, before Proposition 4.3) is satisfied in any such formulation, however.

The endpoint conditions in Examples 5.1 and 5.2 all fit into the mold of this example, since only linear constraints are involved.

Example 5.4. (*Bounded intertemporal linear programming.*) With $P(t)$, P_e, $Q(t)$ and Q_e still taken to be zero matrices as in the preceding example, replace (4.8) by a choice of vectorial intervals giving upper and lower bounds on the various control vectors:

(5.13)

$$U(t) = [\hat{u}^-(t), \hat{u}^+(t)], \quad U_e = [\hat{u}_e^-, \hat{u}_e^+], \quad V(t) = [\hat{v}^-(t), \hat{v}^+(t)], \quad V_e = [\hat{v}_e^-, \hat{v}_e^+].$$

The assumption of continuous dependence of $U(t)$ and $V(t)$ on t is satisfied if the vectors $\hat{u}^-(t)$, $\hat{u}^+(t)$, $\hat{v}^-(t)$ and $\hat{v}^+(t)$ depend continuously on t. In this case *the primal and dual finiteness conditions are both satisfied.* In the notation introduced in Example 3.4 the objective in $(\mathcal{P})$ is to minimize

$$\begin{aligned}\mathcal{F}(u, u_e) = &\int_{t_0}^{t_1} [p(t)\cdot u(t) - c(t)\cdot x(t)]dt + [p_e\cdot u_e - c_e\cdot x(t_1)]\\ &+ \int_{t_0}^{t_1} \hat{v}^-(t)\cdot[q(t) - C(t)x(t) - D(t)u(t)]_- dt + \hat{v}_e^-\cdot[q_e - C_e x(t_1) - D_e u_e]_-\\ &+ \int_{t_0}^{t_1} \hat{v}^+(t)\cdot[q(t) - C(t)x(t) - D(t)u(t)]_+ dt + \hat{v}_e^+\cdot[q_e - C_e x(t_1) - D_e u_e]_+ ,\end{aligned}$$

while the objective in $(\mathcal{Q})$ is to maximize

$$\begin{aligned}\mathcal{G}(v, v_e) = &\int_{t_0}^{t_1} [q(t)\cdot v(t) - b(t)\cdot y(t)]dt + [q_e\cdot v_e - b_e\cdot y(t_0)]\\ &- \int_{t_0}^{t_1} \hat{u}^-(t)\cdot[B^*(t)y(t) + D^*(t)v(t) - p(t)]_- dt - \hat{u}_e^-\cdot[B_e^* y(t_0) + D_e^* v_e - p_e]_-\\ &- \int_{t_0}^{t_1} \hat{u}^+(t)\cdot[B^*(t)y(t) + D^*(t)v(t) - p(t)]_+ dt - \hat{u}_e^+\cdot[B_e^* y(t_0) + D_e^* v_e - p_e]_+ .\end{aligned}$$

For instance, by taking

$$(5.14)\qquad V(t) = [-\lambda\mathbf{1}, \lambda\mathbf{1}], \qquad V_e = [-\lambda_e\mathbf{1}, \lambda_e\mathbf{1}],$$

where $\mathbf{1}$ denotes a vector $(1, 1, \ldots, 1)$ of appropriate dimension, we obtain in $(\mathcal{P})$ the objective

(5.15)

$$\begin{aligned}\mathcal{F}(u, u_e) = &\int_{t_0}^{t_1} [p(t)\cdot u(t) - c(t)\cdot x(t)]dt + [p_e\cdot u_e - c_e\cdot x(t_1)]\\ &+ \lambda\int_{t_0}^{t_1} \|q(t) - C(t)x(t) - D(t)u(t)\|_1 dt + \lambda_e\|q_e - C_e x(t_1) - D_e u_e\|_1,\end{aligned}$$

where

$$(5.16)\qquad \|s\|_1 = \|(s_1, \ldots, s_\ell)\|_1 = |s_1| + \ldots + |s_\ell|.$$

This corresponds to a mathematical model in which constraints of the form

$$(5.17)\qquad C(t)x(t) + D(t)u(t) = q(t) \quad \text{a.e.,} \qquad C_e x(t_1) + D_e u_e = q_e,$$

are to be enforced by linear penalties with parameter values $\lambda > 0$ and $\lambda_e > 0$ sufficiently high.

These ideas are useful in particular in penalty representations of endpoint constraints like the ones discussed in Examples 5.1 and 5.2. Thus a condition $x(t_1) = a_1$ can be modeled by a term $\lambda\|(x(t_1) - a_1\|_1$ in the objective (the case of $C_e = I$, $D_e = 0$ and $q_e = a_1$ in (5.15) and (5.17)). A condition $x(t_0) = a_0$ corresponds of course to a trivial interval $U_e = [0, 0]$ and needs no penalty representation.

Example 5.5. (*Intertemporal piecewise linear programming.*) In the general case where $P(t) = 0$, $P_e = 0$, $Q(t) = 0$ and $Q_e = 0$, one minimizes in $(\mathcal{P})$ the objective

$$\begin{aligned}\mathcal{F}(u, u_e) = &\int_{t_0}^{t_1} [p(t)\cdot u(t) - c(t)\cdot x(t)]dt + [p_e\cdot u_e - c_e\cdot x(t_1)]\\ &+ \int_{t_0}^{t_1} \sigma_{V(t)}(q(t) - C(t)x(t) - D(t)u(t))dt + \sigma_{V_e}(q_e - C_e x(t_1) - D_e u_e)\end{aligned}$$

and one maximizes in $(\mathcal{Q})$ the objective

$$\begin{aligned}\mathcal{G}(v, v_e) = &\int_{t_0}^{t_1} [q(t)\cdot v(t) - b(t)\cdot y(t)]dt + [q_e\cdot v_e - b_e\cdot y(t_0)]\\ &- \int_{t_0}^{t_1} \sigma_{U(t)}(B^*(t)y(t) + D^*(t)v(t) - p(t))dt - \sigma_{U_e}(B_e^* y(t_0) + D_e^* v_e - p_e),\end{aligned}$$

where the σ terms are support functions defined by (2.7) and are polyhedral convex (piecewise linear).

There are two different ways of using this general piecewise linear model, beyond those already covered in Examples 5.3 and 5.4, that deserve emphasis here. The first is in problems where the objective directly involves piecewise linear terms expressed as the pointwise maximum of finite collections of affine functions. This case corresponds to the patterns in Example 3.3 and need not be written out in detail. One has

$$V(t) = [\text{simplex in } \mathbb{R}^{\ell_1}] \times [\text{orthant or interval in } \mathbb{R}^{\ell_2}],$$

and similarly for V_e. Note that in taking in an interval for the second term in each product one has a case where $V(t)$ and V_e are both bounded, so *the primal finiteness condition is satisfied.*

The other way of using this model is less obvious but important in reaching formulations of intertemporal linear programming problems that satisfy the primal and dual finiteness conditions. As already noted in Example 5.3, those conditions are never fulfilled in the basic case of $(\mathcal{P}_1)$ and $(\mathcal{Q}_1)$, but they can be brought to bear by passing to a bounded linear programming formulation as in Example 5.4. A more subtle approach is possible, however, in which only *some* of the constraints receive a linear penalty representation, namely those that definitely involve the state $x(t)$ (or $y(t)$). This might turn out to be a valuable consideration in the application of numerical methods for finding solutions.

For example, suppose we are dealing with a problem initially in the $(\mathcal{P}_1)$ format but with constraints partitioned to clarify the involvement of $x(t)$:

minimize

$$\int_{t_0}^{t_1} [p(t)\cdot u(t) - c(t)\cdot x(t)]dt + [p_e\cdot u_e - c_e\cdot x(t_1)]$$

subject to

$$\begin{aligned} &C_1(t)x(t) + D_1(t)u(t) \ge q_1(t),\\ &D_2(t)u(t) \ge q_2(t), \quad u(t)\ge 0,\\ &C_{e1}x(t_1) + D_{e1}u_e \ge q_{e1},\\ &D_{e2}u_e \ge q_{e2}, \quad u_e \ge 0,\end{aligned}$$

where $q_1(t)\in\mathbf{R}^{\ell_1}$, $q_2(t)\in\mathbf{R}^{\ell_2}$, $q_{e1}\in\mathbf{R}^{\ell_{e1}}$, $q_{e2}\in\mathbf{R}^{\ell_{e2}}$. The $(\mathcal{P}_1)$ format corresponds to choosing

$$C(t) = \begin{bmatrix} C_1(t)\\ 0\end{bmatrix},\quad D(t) = \begin{bmatrix} D_1(t)\\ D_2(t)\end{bmatrix},\quad C_e = \begin{bmatrix} C_{e1}\\ 0\end{bmatrix},\quad D_e = \begin{bmatrix} D_{e1}\\ D_{e2}\end{bmatrix},$$

$$U(t) = \mathbf{R}^k_+,\quad U_e = \mathbf{R}^{k_e}_+,\quad V(t) = \mathbf{R}^\ell_+,\quad V_e = \mathbf{R}^{\ell_e}_+$$

(with $\ell = \ell_1+\ell_2$ and $\ell_e = \ell_{e1}+\ell_{e2}$). An alternative formulation, however, is to take

$$\begin{aligned} &C(t) = C_1(t),\quad D(t) = D_1(t),\quad q(t) = q_1(t),\\ &U(t) = \{u\ge 0 \mid D_2(t)u\ge q_2\},\quad V(t) = \mathbf{R}^{\ell_1}_+,\\ &C_e = C_{e1},\quad D_e = D_{e1},\quad q_e = q_{e1},\\ &U_e = \{u_e\ge 0\mid D_{e2}u_e\ge q_{e2}\},\quad V_e = \mathbf{R}^{\ell_{e1}}_+.\end{aligned}$$

If $U(t)$ and U_e happen to be bounded sets, we *have the dual boundedness condition satisfied* in this formulation even though it was not satisfied in the formulation as $(\mathcal{P}_1)$.

What effect does this alternative have on the nature of the dual problem? One maximizes the expression

$$\int_{t_0}^{t_1} [q_1(t)\cdot v_1(t) - b(t)\cdot y(t)]dt + [q_{e1}\cdot v_{e1} - b_e\cdot y(t_0)]$$

$$-\int_{t_0}^{t_1} \sigma_{U(t)}(B^*(t)y(t) + D_1^*(t)v_1(t) - p(t))dt - \sigma_{U_e}(B_e^* y(t_0) + D_{e1}^* v_{e1} - p_e)$$

subject to $v_1(t)\in\mathbf{R}^{k_1}_+$ and $v_{e1}\in\mathbf{R}^{k_{e1}}_+$. One has

$$\begin{aligned} -\sigma_{U(t)}(r) &:= -\sup\{r\cdot u \mid u\ge 0,\ D_2(t)u\ge q_2(t)\}\\ &= \inf\{-r\cdot u\mid u\ge 0,\ D_2(t)u\ge q_2(t)\}\\ &= \sup\{q_2(t)\cdot v_2\mid v_2\ge 0,\ D_2^*(t)v_2\le -r\}\end{aligned}$$

by finite-dimensional linear programming duality, so that

$$\begin{aligned} &-\sigma_{U(t)}(B^*(t)y(t) + D_1^*(t)v_1(t) - p(t))\\ &\quad = \sup\{q_2(t)\cdot v_2\mid v_2\ge 0,\ B^*(t)y(t_0) + D_1^*(t)v_1(t) + D_2^*(t)v_2\le p(t)\}.\end{aligned}$$

Similarly

$$\begin{aligned} &-\sigma_{U_e}(B_e^* y(t_0) + D_{e1}^* v_{e1} - p_e)\\ &\quad = \sup\{q_{e2}\cdot v_{e2}\mid v_{e2}\ge 0,\ B_e^* y(t) + D_{e1}^* v_{e1} + D_{e2}^* v_{e2}\le p_e\}.\end{aligned}$$

The dual problem for the alternative approach is therefore essentially the same as $(\mathcal{Q}_1)$, except that the v_2 and v_{e2} components in $\mathbf{R}^{\ell_2}$ and $\mathbf{R}^{\ell_{e2}}$ have been "maximized out." These components can ultimately be recovered if necessary, but in the meantime we do not need to worry about them in connection with theorems about optimality conditions, existence and duality, in particular the $\mathcal{L}^\infty$ requirement on $v(t)$ in $(\mathcal{Q})$.

Of course, in order for this approach to work, we must also be able to verify the assumption of continuous dependence of $U(t)$ on t. When $U(t) = \{u\ge 0\mid D_2(t)u\ge q_2(t)\}$, this is satisfied for instance if $D_2(t)$ and $q_2(t)$ do not actually depend on t, or if there is a *continuous* function u such that $u(t)\ge 0$ and $D_2(t)u(t) > q_2(t)$ (strict inequality in every component). For the latter and also more general cases involving a possible mixture of equality and inequality constraints, see Rockafellar [48, Cor. 3.3].

Similar ideas can be applied to a partitioning of the constraints of a problem $(\mathcal{Q}_1)$ into those that affect $y(t)$ and those that do not. In $(\mathcal{P}_1)$ this would correspond to dynamics $\dot{x} = Ax + Bu + b$, $x(t_0) = B_e u_e + b_e$, where $B = [B_1, 0]$ and $B_e = [B_{e1}, 0]$, i.e. not all components of u and u_e are directly active in the dynamics.

Example 5.6. (*Linear-quadratic regulator problem and generalizations.*) Consider now a classical type of problem having the form

$$\text{(5.18)}\qquad \begin{aligned} \text{minimize}\quad &\frac{1}{2}\int_{t_0}^{t_1} [u(t)\cdot P(t)u(t) + (x(t)-\hat{x}(t))\cdot R(t)(x(t)-\hat{x}(t))]dt\\ &+\frac{1}{2}(x(t_1)-a_1)\cdot R_e(x(t_1)-a_1)\\ \text{subject to}\quad &u\in\mathcal{L}^\infty([t_0,t_1],\mathbf{R}^k),\\ &\dot{x}(t) = A(t)x(t) + B(t)u(t) + b(t)\ \text{a.e.},\qquad x(t_0) = a_0,\end{aligned}$$

where a_0 and a_1 are given points, $\hat{x}$ is a given function (continuous), $P(t)$ is positive *definite*, and $R(t)$ and R_e are positive *semidefinite*. This can be formulated as a problem $(\mathcal{P})$ by introducing factorizations

$$\text{(5.19)}\qquad R(t) = C^*(t)Q(t)^{-1}C(t)\quad\text{and}\quad R_e = C_e^* Q_e^{-1} C_e,$$

where $Q(t)$ and Q_e are positive *definite*. (If $R(t)$ and R_e themselves are positive definite, one can of course take $C(t) = I$, $Q(t) = R(t)^{-1}$, $C_e = I$, $Q_e = R_e^{-1}$, in (5.19).) Set

$$p(t) = 0,\quad c(t) = 0,\quad D(t) = 0,\quad q(t) = C(t)\hat{x}(t),\quad U(t) = \mathbf{R}^k,\quad V(t) = \mathbf{R}^\ell.$$

Then in the general format of $(\mathcal{P})$ the terms

$$p(t)\cdot u(t)\cdot P(t)u(t) - c(t)\cdot x(t) + \rho_{V(t),Q(t)}(q(t) - C(t)x(t) - D(t)u(t))$$

reduce to

$$\tfrac{1}{2}u(t)\cdot P(t)u(t) + \tfrac{1}{2}(x(t)-\hat{x}(t))\cdot R(t)(x(t)-\hat{x}(t)).$$

For the endpoints, trivialize u_e by taking $\mathbf{R}^{k_e}$ to be zero-dimensional (so $U_e = \{0\}$, $p_e = 0$, $P_e = 0$, $B_e = 0$, $D_e = 0$) and let $b_e = a_0$, $q_e = C_e a_1, c_e = 0$. The terms

$$p_e \cdot u_e + \tfrac{1}{2} u_e \cdot P_e u_e - c_e \cdot x(t_1) + \rho_{V_e,Q_e}(q_e - C_e x(t_1) - D_e u_e)$$

in $(\mathcal{P})$ then reduce to

$$\tfrac{1}{2}(x(t_1) - a_1) \cdot R_e(x(t_1) - a_1),$$

and we get the desired problem (5.18) as a special case of $(\mathcal{P})$. The corresponding dual $(\mathcal{Q})$ has the form

$$\begin{aligned}
&\text{minimize} && \int_{t_0}^{t_1} [\hat{x}(t) \cdot C^*(t)v(t) - b(t) \cdot y(t)]dt + [a_1 \cdot C_e^* v_e - a_0 \cdot y(t_0)] \\
(5.20)\quad & && - \tfrac{1}{2} \int_{t_0}^{t_1} [v(t) \cdot Q(t)v(t) + y(t) \cdot S(t)y(t)]dt - \tfrac{1}{2} v_e \cdot Q_e v_e \\
&\text{subject to} && v \in \mathcal{L}^\infty([t_0, t_1], \mathbf{R}^\ell), \quad v_e \in \mathbf{R}^{\ell_e}, \\
& && -\dot{y}(t) = A^*(t)y(t) + C^*(t)v(t) \text{ a.e.}, \quad y(t_1) = C_e^* v_e,
\end{aligned}$$

where

$$S(t) = B(t)P(t)^{-1}B^*(t). \tag{5.21}$$

Note that in this example *the primal and dual boundedness conditions are both satisfied.*

Generalizations of the linear-quadratic regulator problem can be made in several directions without going beyond the format of our problem $(\mathcal{P})$. For instance, instead of letting $u(t)$ be a free vector in $\mathbf{R}^k$ one can insist on bounds $\hat{u}^-(t) \le u(t) \le \hat{u}^+(t)$. Dually one can introduce bounds

$$-\lambda \mathbf{1} \le v(t) \le \lambda \mathbf{1} \quad \text{and} \quad -\lambda_e \mathbf{1} \le v_e \le \lambda_e \mathbf{1}$$

for parameter values $\lambda > 0$, $\lambda_e > 0$. The effect of this on the formulation of the original problem (5.18) is to replace the purely quadratic penalty expressions by terms that are quadratic near the origin but eventually grow at a linear rate. Thus for example if

$$R(t) = \mu I \quad \text{and} \quad R_e = \mu_e I \quad \text{for} \quad \mu > 0, \ \mu_e > 0$$

(corresponding in (5.19) to $C(t) = I$, $Q(t) = \mu^{-1} I$, $C_e = I$, $Q_e = \mu_e^{-1} I$) one has terms

$$(\mu/2) \int_{t_0}^{t_1} |x(t) - \hat{x}(t)|^2 dt + (\mu_e/2)|x(t_1) - a_1|^2$$

in (5.18) that are replaced by

$$\sum_{i=1}^{n} \left[\int_{t_0}^{t_1} \psi(|x_i(t) - \hat{x}_i(t)|)dt + \psi_e(|x_i(t_1) - a_{1i}|) \right],$$

where $x_i(t)$ and a_{1i} are the ith components of $x(t)$ and a_1 and ψ is the growth function defined by

$$\psi(\tau) = \begin{cases} (\mu/2)\tau^2 & \text{when} \quad 0 \le \tau \le \lambda/\mu, \\ \lambda(\tau - (\lambda/\mu)) + (\lambda^2/2\mu) & \text{when} \quad \tau \ge \lambda/\mu, \end{cases}$$

and similarly ψ_e in terms of λ_e and μ_e.

Still other generalizations of the linear-quadratic regulator problem are covered by the patterns in the next example.

Example 5.7. (*Bounded intertemporal quadratic programming.*) This corresponds to the finite-dimensional bounded quadratic programming models in Example 3.5 in the same way that Example 5.4 corresponds to the finite-dimensional bounded linear programming models in Example 3.4. Due to all the notation involved, we shall not write these problems out in full. The point is, however, that these are formulations of considerable versatility which allow for quadratic terms without damaging the explicit, symmetric nature of the dualization.

Example 5.8. (*Problems whose duals are essentially finite-dimensional.*) Suppose in problem $(\mathcal{P})$ that $C(t) \equiv 0$. Then in $(\mathcal{Q})$ the trajectory y is uniquely determined from v_e alone. Although $v(t)$ still appears in the objective in $(\mathcal{Q})$, it does so in a very simple way: the value chosen for $v(t)$ has no connection to past or future. At each time t one can just take $v(t)$ to maximize the expression

$$q(t) \cdot v(t) - \tfrac{1}{2} v(t) \cdot Q(t)v(t) - b(t) \cdot y(t) - \rho_{V(t),Q(t)}(B^*(t)y(t) + D^*(t)v(t) - p(t))$$

over $V(t)$, where $y(t)$ is already fixed. In this sense $(\mathcal{Q})$ is really a problem in v_e alone and is therefore finite-dimensional. (Of course $v(t)$ must ultimately be an $\mathcal{L}^1$ function of t.)

6. Saddle points and optimality. The duality between problems $(\mathcal{P})$ and $(\mathcal{Q})$ will be established by associating them with an infinite-dimensional saddle point problem. This will lead to the principal results of this paper, which concern the existence and optimality properties of solutions to $(\mathcal{P})$ and $(\mathcal{Q})$.

The saddle point representation we aim at follows the general guidelines at the beginning of §2. We take the control spaces $\mathcal{U}$ and $\mathcal{V}$ already introduced in §4 (which are nonempty by Theorem 4.2) and define on $\mathcal{U} \times \mathcal{V}$ a certain functional $\mathcal{J}$, namely

$$\mathcal{J}(u, u_e; v, v_e) = \int_{t_0}^{t_1} J(t, u(t), v(t))dt + J_e(u_e, v_e) - [(u, u_e), (v, v_e)] \tag{6.1}$$

under the convention $\infty - \infty = \infty$ (see below), where

$$J(t, u, v) = p(t) \cdot u + q(t) \cdot v + \tfrac{1}{2} u \cdot P(t)u - \tfrac{1}{2} v \cdot Q(t)v - v \cdot D(t)u, \tag{6.2}$$

$$J_e(u_e, v_e) = p_e \cdot u_e + q_e \cdot v_e + \tfrac{1}{2} u_e \cdot P_e u_e - \tfrac{1}{2} v_e \cdot Q_e v_e - v_e \cdot D_e u_e, \tag{6.3}$$

and

$$\begin{aligned}
[(u, u_e), (v, v_e)] &= \int_{t_0}^{t_1} y(t) \cdot [B(t)u(t) + b(t)]dt + y(t_0) \cdot [B_e u_e + b_e] \\
&= \int_{t_0}^{t_1} x(t) \cdot [C^*(t)v(t) + c(t)]dt + x(t_1) \cdot [C_e^* v_e + c_e].
\end{aligned} \tag{6.4}$$

The common value of the two expressions for $[(u, u_e), (v, v_e)]$ in (6.4) stems from the integration-by-parts formula

$$\int_{t_0}^{t_1} y(t) \cdot \dot{x}(t)dt + y(t_0) \cdot x(t_0) = -\int_{t_0}^{t_1} x(t) \cdot \dot{y}(t)dt + x(t_1) \cdot y(t_1).$$

The term $[(u, u_e), (v, v_e)]$, which is affine in (u, u_e) for fixed (v, v_e) and affine in (v, v_e) for fixed (u, u_e), as well as continuous with respect to all arguments, embodies the fundamental connection between the control systems in $(\mathcal{P})$ and $(\mathcal{Q})$.

The convention $\infty - \infty = \infty$ mentioned in the definition (6.1) of $\mathcal{J}$ refers to possible ambiguities in the value of the integral of $J(t, u(t), v(t))$. In general, since

$u(t)$ and $v(t)$ are only $\mathcal{L}^1$ in t, the integral of the term $u(t)\cdot P(t)u(t)$ might be ∞, the integral of $v(t)\cdot Q(t)v(t)$ might be $-\infty$, and the integral of $v(t)\cdot D(t)u(t)$ might be either. We use $\infty-\infty=\infty$ to resolve any dilemmas in extended arithmetic that might arise. This amounts to taking the integral term in (6.1) to be ∞ if $J(t,u(t),v(t))$ is not majorized by any $\mathcal{L}^1$ function of t. Of course if $J(t,u(t),v(t))\le\alpha(t)$ for an $\mathcal{L}^1$ function α, then the integral has an unambiguous value which is finite or $-\infty$, whereas if $J(t,u(t),v(t))\ge\beta(t)$ for an $\mathcal{L}^1$ function β, it is finite or ∞. Actually there is no difficulty at all if $u\in\mathcal{L}^\infty$ or $v\in\mathcal{L}^\infty$: one has

$$J(u,u_e;v,v_e)<\infty \quad\text{when}\quad u\in\mathcal{L}^\infty, \tag{6.5}$$

$$J(u,u_e;v,v_e)>-\infty \quad\text{when}\quad v\in\mathcal{L}^\infty, \tag{6.6}$$

and therefore $J(u,u_e;v,v_e)$ finite when both $u\in\mathcal{L}^\infty$ and $v\in\mathcal{L}^\infty$.

Anyway, under the specified convention J is a well-defined functional on $\mathcal{U}\times\mathcal{V}$ which is quadratic convex in (u,u_e) and quadratic concave in (v,v_e). The convention $\infty-\infty=-\infty$ could have been used instead and would have led to a functional $\tilde{J}$ that would serve our purposes in equivalent fashion; we shall occasionally make use of $\tilde{J}$ in our proofs. Obviously from (6.5) and (6.6), J and $\tilde{J}$ agree whenever $u\in\mathcal{L}^\infty$ or $v\in\mathcal{L}^\infty$.

THEOREM 6.1. *Problems $(\mathcal{P})$ and $(\mathcal{Q})$ are the primal and dual optimization problems associated with the saddle point problem for J on $\mathcal{U}\times\mathcal{V}$. Thus the functional $\mathcal{F}$ which in $(\mathcal{P})$ is minimized over $\mathcal{U}$ is given by*

$$\mathcal{F}(u,u_e)=\sup_{(v,v_e)\in\mathcal{V}} J(u,u_e;v,v_e), \tag{6.7}$$

whereas the functional $\mathcal{G}$ which in $(\mathcal{Q})$ is maximized over $\mathcal{V}$ is given by

$$\mathcal{G}(v,v_e)=\inf_{(u,u_e)\in\mathcal{U}} J(u,u_e;v,v_e). \tag{6.8}$$

Proof. In establishing (6.7) we take the second of the expressions in (6.4) for the term $[(u,u_e),(v,v_e)]$ in the definition (6.1) of J, so that

(6.9)
$$\begin{aligned}J(u,u_e;v,v_e)=&\int_{t_0}^{t_1}[p(t)\cdot u(t)+\tfrac12 u(t)\cdot P(t)u(t)-c(t)\cdot x(t)]dt\\&+\int_{t_0}^{t_1}(v(t)\cdot[q(t)-C(t)x(t)-D(t)u(t)]-\tfrac12 v(t)\cdot Q(t)v(t))dt\\&+[p_e\cdot u_e+\tfrac12 u_e\cdot P_eu_e-c_e\cdot x(t_1)]\\&+v_e\cdot[q_e-C_ex(t_1)-D_eu_e]-\tfrac12 v_e\cdot Q_ev_e.\end{aligned}$$

From the definition (2.5) of the functions $\rho_{V(t),Q(t)}$ and ρ_{V_e,Q_e} it is clear that

$$\mathcal{F}(u,u_e)\ge J(u,u_e;v,v_e)\quad\text{for all } (u,u_e)\in\mathcal{U},\ (v,v_e)\in\mathcal{V}, \tag{6.10}$$

and that the desired equation (6.7) can be verified by showing that the equation

$$\sup_{\substack{v\in\mathcal{L}^\infty\\ v(t)\in V(t)}}\int_{t_0}^{t_1}[v(t)\cdot s(t)-\tfrac12 v(t)\cdot Q(t)v(t)]dt=\int_{t_0}^{t_1}\rho_{V(t),Q(t)}(s(t))dt \tag{6.11}$$

holds for arbitrary $s\in\mathcal{L}^1$. This equation can be written as

$$\sup_{v\in\mathcal{L}^\infty}\int_{t_0}^{t_1}[v(t)\cdot s(t)-\varphi_t(v(t))]dt=\int_{t_0}^{t_1}\varphi_t^*(s(t))dt \tag{6.12}$$

for the convex function $\varphi_t(v)=j_{Q(t)}(v)+\delta_{V(t)}(v)$ utilized in the proofs of Proposition 2.3, 2.4, and 4.1. It holds by [49, Thm. 2] (or [50, Thm. 3C]) if $\varphi(t,v)=\varphi_t(v)$ is a so-called *normal integrand* and the left side of (6.12) is not $-\infty$. Actually $\varphi(t,v)$ is lower semicontinuous jointly in t and v, inasmuch as $Q(t)$ and $V(t)$ depend continuously on t, whereas normality merely requires $\varphi(t,v)$ to be lower semicontinuous in v for fixed t and measurable in (t,v) with respect to the σ-algebra in $[t_0,t_1]\times\mathbf{R}^\ell$ generated by the Lebesgue sets in $[t_0,t_1]$ and the Borel sets in $\mathbf{R}^\ell$ [50, Thm. 24]. Thus φ is normal. Furthermore the left side of (6.12), or equivalently of (6.11), cannot be $-\infty$, because the integral is finite when $v\in\mathcal{L}^\infty$, and we do know (from the proof of Theorem 4.2) that $\mathcal{V}$ contains at least one pair (v,v_e) with v actually continuous.

Our argument has not only verified (6.7) but shown that the same would be true if J were replaced by the alternative functional $\tilde{J}$ using $\infty-\infty=-\infty$ instead of $\infty-\infty=\infty$. Indeed, (6.10) still holds for $\tilde{J}$, since $J\ge\tilde{J}$. Everything else is unchanged, because we relied only on $v\in\mathcal{L}^\infty$, and for such v the values of J and $\tilde{J}$ agree. This symmetry is all we need to conclude that (6.8) is valid too. □

THEOREM 6.2 (Weak Duality). *For the optimal control problems $(\mathcal{P})$ and $(\mathcal{Q})$ it is always true that*

$$\inf(\mathcal{P})\ge\sup(\mathcal{Q}).$$

Furthermore a pair $((\bar u,\bar u_e),(\bar v,\bar v_e))$ is a saddle point of J on $\mathcal{U}\times\mathcal{V}$ if and only if $(\bar u,\bar u_e)$ solves $(\mathcal{P})$, $(\bar v,\bar v_e)$ solves $(\mathcal{Q})$, and $\min(\mathcal{P})=\max(\mathcal{Q})$ (finite).

Proof. This is just a repeat of the general facts in Proposition 2.1 for the specific case in Theorem 6.1. □

A stronger result is obtained by appealing to the *finiteness conditions* for $(\mathcal{P})$ and $(\mathcal{Q})$ that were introduced at the end of §4. We wish to emphasize again, as in §1, that this is by no means the most general result on strong duality. Rather, it is presented as a relatively simple result which is easy to work with and already capable of covering many important cases, especially in view of the modeling possibilities explained in §5.

THEOREM 6.3 (Strong Duality). *If the primal finiteness condition is satisfied, then*

$$\inf(\mathcal{P})=\max(\mathcal{Q})<\infty, \tag{6.13}$$

and moreover the dual objective $\mathcal{G}$ is weakly sup-compact relative to $\mathcal{V}$, i.e. all level sets of the form

$$\{(v,v_e)\in\mathcal{V}\mid \mathcal{G}(v,v_e)\ge\alpha\}\quad\textit{for } \alpha\in\mathbf{R} \tag{6.14}$$

are weakly compact in $\mathcal{L}^1([t_0,t_1],\mathbf{R}^\ell)\times\mathbf{R}^{\ell_e}$.

Likewise, if the dual finiteness condition is satisfied, then

$$\min(\mathcal{P})=\sup(\mathcal{Q})>-\infty, \tag{6.15}$$

and moreover the primal objective $(\mathcal{P})$ is weakly inf-compact relative to $\mathcal{U}$, i.e. all level sets of the form

$$\{(u,u_e)\in\mathcal{U}\mid \mathcal{F}(u,u_e)\le\alpha\}\quad\textit{for } \alpha\in\mathbf{R} \tag{6.16}$$

are weakly compact in $\mathcal{L}^1([t_0,t_1],\mathbf{R}^k)\times\mathbf{R}^{k_e}$.

Thus if both finiteness conditions are satisfied, solutions exist to both $(\mathcal{P})$ and $(\mathcal{Q})$, and

$$\min(\mathcal{P})=\max(\mathcal{Q})\quad(\textit{finite}). \tag{6.17}$$

Proof. Our proof of the formulas (6.7) and (6.8) in Theorem 6.1 gave something slightly stronger that will now be of use: if we denote by $\mathcal{U}^\infty$ and $\mathcal{V}^\infty$ the subsets of $\mathcal{U}$ and $\mathcal{V}$ having $u \in \mathcal{L}^\infty$ and $v \in \mathcal{L}^\infty$, then

$$\mathcal{F}(u,u_e) = \sup_{(v,v_e)\in\mathcal{V}^\infty} J(u,u_e;v,v_e) \quad \text{for all } (u,u_e), \tag{6.18}$$

$$\mathcal{G}(v,v_e) = \inf_{(u,u_e)\in\mathcal{U}^\infty} J(u,u_e;v,v_e) \quad \text{for all } (v,v_e). \tag{6.19}$$

In order to obtain (6.13) it will be enough by this to demonstrate

$$\inf_{\mathcal{U}^\infty}\sup_{\mathcal{V}} J = \max_{\mathcal{V}}\inf_{\mathcal{U}^\infty} J, \tag{6.20}$$

since the inequalities

$$\inf_{\mathcal{U}^\infty}\sup_{\mathcal{V}} J \ge \inf_{\mathcal{U}}\sup_{\mathcal{V}} J \ge \sup_{\mathcal{V}}\inf_{\mathcal{U}} J$$

hold trivially. The one-sided minimax theorem of Moreau [51] will justify (6.20) provided we can show that under the primal finiteness condition $J(u,u_e;v,v_e)$ is weakly sup-compact in (v,v_e) relative to $\mathcal{V}$ when $(u,u_e)\in\mathcal{U}^\infty$. The latter will also give us the claimed sup-compactness of $\mathcal{G}$ via (6.19).

Fix $(u,u_e)\in\mathcal{U}^\infty$. Taking J as expressed in (6.9) and introducing $s(t)$ and s_e as in (4.5), we have

$$J(u,u_e;v,v_e) = \int_{t_0}^{t_1} [v(t)\cdot s(t) - \tfrac{1}{2}v(t)\cdot Q(t)v(t)]dt + [v_e\cdot s_e - \tfrac{1}{2}v_e\cdot Q_e v_e] + \text{const.} \tag{6.21}$$

for all $(v,v_e)\in\mathcal{V}$, where $s(t)$ is $\mathcal{L}^\infty$ in t. The required sup-compactness property of J is the weak compactness of the level sets

$$\{(v,v_e)\in\mathcal{V} \mid J(u,u_e;v,v_e)\ge\alpha\} \quad \text{for } \alpha\in\mathrm{R}.$$

We recognize now that this is the same as the weak compactness of the level sets

$$\{(v,v_e)\in\mathcal{V} \mid \tfrac{1}{2}\int_{t_0}^{t_1} v(t)\cdot Q(t)v(t)dt + \tfrac{1}{2}v_e\cdot Q_e v_e - \langle\langle(v,v_e),(s,s_e)\rangle\rangle \le \beta\} \tag{6.22}$$

for $\beta\in\mathrm{R}$, where

$$\langle\langle(v,v_e),(s,s_e)\rangle\rangle = \int_{t_0}^{t_1} v(t)\cdot s(t)dt + v_e\cdot s_e. \tag{6.23}$$

Once again the convex function

$$\varphi_t(v) = j_{Q(t)}(v) + \delta_{V(t)}(v) = \begin{cases} \tfrac{1}{2}v\cdot Q(t)v & \text{if } v\in V(t),\\ \infty & \text{if } v\notin V(t)\end{cases}$$

will be useful, together with

$$\varphi_e(v_e) = j_{Q_e}(v_e) + \delta_{V_e}(v_e) = \begin{cases} \tfrac{1}{2}v_e\cdot Q_e v_e & \text{if } v_e\in V_e,\\ \infty & \text{if } v_e\notin V_e.\end{cases}$$

The convex functional

$$I(v,v_e) = \int_{t_0}^{t_1}\varphi_t(v(t))dt + \varphi_e(v_e)$$

is well defined on $\mathcal{L}^1([t_0,t_1],\mathrm{R}^\ell)\times\mathrm{R}^{\ell_e}$ with values in $[0,\infty)$, and in terms of it the set (6.22) can be written as

$$\{(v,v_e)\in\mathcal{L}^1([t_0,t_1],\mathrm{R}^\ell)\times\mathrm{R}^{\ell_e} \mid I(v,v_e) - \langle\langle(v,v_e),(s,s_e)\rangle\rangle \le \beta\}. \tag{6.24}$$

We shall be able to establish the weak compactness of this set for arbitrary $(s,s_e)\in\mathcal{L}^\infty([t_0,t_1],\mathrm{R}^\ell)\times\mathrm{R}^{\ell_e}$ and $\beta\in R$ by means of the theory of integral functional conjugate to each other [49], [50].

Let us think of the spaces $\mathcal{L}^1([t_0,t_1],\mathrm{R}^\ell)\times\mathrm{R}^{\ell_e}$ and $\mathcal{L}^\infty([t_0,t_1],\mathrm{R}^\ell)\times\mathrm{R}^{\ell_e}$ as dual to each other under the pairing (6.23). The pairing formula and the formula for I can actually be viewed as integrals over a measure space that is the union of $[t_0,t_1]$ and an atom $\{e\}$ of measure 1. In this sense I is an integral functional, pure and simple. The functional

$$I^*(s,s_e) = \int_{t_0}^{t_1}\varphi_t^*(s(t))dt + \varphi_e^*(s_e)$$

on $\mathcal{L}^\infty([t_0,t_1],\mathrm{R}^\ell)\times\mathrm{R}^{\ell_e}$, where φ_t^* and φ_e^* are conjugate to φ_t and φ_e, is an integral functional too, and I and I^* are conjugate to each other by [49, Thm. 2] (or [50, Thm. 3C]) with respect to the pairing (6.23). Indeed $\varphi_t^* = \rho_{V(t),Q(t)}$ and $\varphi_e^* = \rho_{V_e,Q_e}$, so φ_t^* and φ_e^* are finite convex functions on R^ℓ under the primal finiteness condition we are assuming. Furthermore $\varphi_t^*(s)$ is for each $s\in\mathrm{R}^\ell$ continuous in t by Proposition 4.1, hence integrable over $[t_0,t_1]$. These properties for I^* plug into the weak inf-compactness criterion of [49, p. 538] for integral functional on $\mathcal{L}^1$-type spaces and prove the required weak compactness of all level sets of the form (6.24) for the conjugate functional $I=(I^*)^*$.

The proof of (6.15) and the weak compactness of the sets (6.16) follows now by symmetry. □

COROLLARY 6.4. *Suppose the primal and dual finiteness conditions both hold. Then in order that $(\bar{u},\bar{u}_e)$ solve $(\mathcal{P})$ and $(\bar{v},\bar{v}_e)$ solve $(\mathcal{Q})$, it is both necessary and sufficient that $(\bar{u},\bar{u}_e),(\bar{v},\bar{v}_e))$ be a saddle point of J on $\mathcal{U}\times\mathcal{V}$.*

Proof. According to Theorem 6.2 the saddle point condition is always sufficient, and if $\min(\mathcal{P}) = \max(\mathcal{Q})$ it is also necessary. Necessity therefore follows from the primal and dual finiteness conditions by the result just proved in Theorem 6.3. □

The saddle point condition in Corollary 6.4 means that $(\bar{u},\bar{u}_e)\in\mathcal{U},(\bar{v},\bar{v}_e)\in\mathcal{V}$, and

$$J(\bar{u},\bar{u}_e;v,v_e)\le J(\bar{u},\bar{u}_e;\bar{v},\bar{v}_e)\le J(u,u_e;\bar{v},\bar{v}_e) \tag{6.25}$$

for all $(u,u_e)\in\mathcal{U}$ and $(v,v_e)\in\mathcal{V}$. This "global" condition actually decomposes, as we show next, into an "instantaneous" saddle point condition at each time t and an "endpoint" saddle point condition.

THEOREM 6.5 (Minimaximum Principle). *For $((\bar{u},\bar{u}_e),(\bar{v},\bar{v}_e))$ to be a saddle point of J on $\mathcal{U}\times\mathcal{V}$, it is necessary and sufficient that the following conditions hold (in addition to $\bar{u}(t)$ and $\bar{v}(t)$ being $\mathcal{L}^1$ in t). For almost every $t\in[t_0,t_1]$*

(6.26) *$(\bar{u}(t),\bar{v}(t))$ is a saddlepoint relative to $U(t)\times V(t)$ for*
$J(t,u,v) - u\cdot B^*(t)\bar{y}(t) - v\cdot C(t)\bar{x}(t)$,

and also

(6.27) *$(\bar{u}_e,\bar{v}_e)$ is a saddlepoint relative to $U_e\times V_e$ for*
$J_e(u_e,v_e) - u_e\cdot B_e^*\bar{y}(t_0) - v_e\cdot C_e\bar{x}(t_1)$,

where $\bar{x}$ and $\bar{y}$ are the primal and dual state functions corresponding to $(\bar{u},\bar{u}_e)$ and $(\bar{v},\bar{v}_e)$.

Proof. The saddle point condition (6.25) for J on $\mathcal{U}\times\mathcal{V}$ is equivalent by Theorem 6.1 to the condition

$$\mathcal{F}(\bar{u},\bar{u}_e) = J(\bar{u},\bar{u}_e;\bar{v},\bar{v}_e) = \mathcal{G}(\bar{v},\bar{v}_e). \tag{6.28}$$

Let us write this as

$$\mathcal{F}(\bar u, \bar u_e) + \bar\alpha = J(\bar u, \bar u_e; \bar v, \bar v_e) + \bar\alpha = \mathcal{G}(\bar v, \bar v_e) + \bar\alpha, \tag{6.29}$$

where

$$\bar\alpha = \int_{t_0}^{t_1} [c(t)\cdot \bar x(t) + b(t)\cdot \bar y(t)]dt + c_e \cdot \bar x(t_1) + b_e \cdot \bar y(t_0) - [(\bar u, \bar u_e), (\bar v, \bar v_e)].$$

The alternative expressions for $[(\bar u, \bar u_e), (\bar v, \bar v_e)]$ in (6.4) give

$$\bar\alpha = \int_{t_0}^{t_1} [c(t)\cdot \bar x(t) - \bar u(t)\cdot B^*(t)\bar y(t)]dt + [c_e \cdot \bar x(t_1) - \bar u_e \cdot B_e^* \bar y(t_0)]$$
$$= \int_{t_0}^{t_1} [b(t)\cdot \bar y(t) - \bar v(t)\cdot C(t)\bar x(t)]dt + [b_e \cdot \bar y(t_0) - \bar v_e \cdot C_e \bar x(t_1)].$$

Using these along with the formulas defining $\mathcal{F}$, $\mathcal{G}$, and J, we get expressions of the form

$$\mathcal{F}(\bar u, \bar u_e) = \int_{t_0}^{t_1} \bar f_t(\bar u(t))dt + \bar f_e(\bar u_e),$$
$$\mathcal{G}(\bar v, \bar v_e) + \bar\alpha = \int_{t_0}^{t_1} \bar g_t(\bar v(t))dt + \bar g_e(\bar v_e),$$
$$J(\bar u, \bar u_e; \bar v, \bar v_e) + \alpha = \int_{t_0}^{t_1} \bar J_t(\bar u(t), \bar v(t))dt + \bar J_e(u_e, v_e),$$

where

$$\bar f_t(u) = [p(t) - B^*(t)\bar y(t)]\cdot u + \tfrac{1}{2}u\cdot P(t)u + \rho_{V(t),Q(t)}(q(t) - C(t)\bar x(t) - D(t)u), \tag{6.30}$$

$$\bar f_e(u_e) = [p_e - B_e^*\bar y(t_0)]\cdot u + \tfrac{1}{2}u_e\cdot P_e u_e + \rho_{V_e,Q_e}(q_e - C_e\bar x(t_1) - D_e u_e), \tag{6.31}$$

$$\bar g_t(v) = [q(t) - C(t)\bar x(t)]\cdot v - \tfrac{1}{2}v\cdot Q(t)v - \rho_{V(t),P(t)}(B^*(t)\bar y(t) + D^*(t)v - p(t)), \tag{6.32}$$

$$\bar g_e(v_e) = [q_e - C_e\bar x(t_1)]\cdot v_e - \tfrac{1}{2}v_e\cdot Q_e v_e - \rho_{V_e,P_e}(B_e^*\bar y(t_0) + D_e^* v_e - p_e), \tag{6.33}$$

$$\bar J_t(u, v) = J(t, u, v) - u\cdot B^*(t)\bar y(t) - v\cdot C(t)\bar x(t), \tag{6.34}$$

$$\bar J_e(u_e, v_e) = J_e(u_e, v_e) - u_e\cdot B_e^*\bar y(t_0) - v_e\cdot C_e\bar x(t_1). \tag{6.35}$$

The saddle point condition on $((\bar u, \bar u_e), (\bar v, \bar v_e))$, written as (6.27), is equivalent under this formulation to

$$\int_{t_0}^{t_1} \bar f_t(\bar u(t))dt + \bar f_e(\bar u_e) = \int_{t_0}^{t_1} \bar J_t(\bar u(t), \bar v(t))dt + \bar J_e(\bar u_e, \bar v_e)$$
$$= \int_{t_0}^{t_1} \bar g_t(\bar v(t))dt + \bar g_e(\bar v_e). \tag{6.36}$$

But

$$\bar f_t(u) = \sup_{v\in V(t)} \bar J_t(u, v), \qquad \bar g_t(v) = \inf_{u\in U(t)} \bar J_t(u, v), \tag{6.37a}$$

$$\bar f_e(u_e) = \inf_{v\in V_e} \bar J_e(u_e, v_e), \qquad \bar g_e(v_e) = \sup_{u\in U_e} \bar J_e(u_e, v_e), \tag{6.37b}$$

by the definition of the ρ terms in (6.30) – (6.33), so

$$f_t(u) \ge J_t(u, v) \ge g_t(v) \quad \text{for all } u \in U(t), v \in V(t),$$
$$f_e(u_e) \ge J_e(u_e, v_e) \ge g_e(v_e) \quad \text{for all } u_e \in U_e, v_e \in V_e.$$

Since the left side of (6.36) cannot be $-\infty$, whereas the right side cannot be ∞ (from the corresponding facts about $\mathcal{F}(u, u_e)$ and $\mathcal{G}(v, v_e)$ in Theorem 6.1), condition (6.36) holds if and only if

$$\bar f_t(\bar u(t)) = \bar J_t(\bar u(t), \bar v(t)) = \bar g_t(\bar v(t)) \quad \text{a.e.}, \qquad \bar f_e(\bar u_e) = \bar J_e(\bar u_e, \bar v_e) = \bar g_e(\bar v_e).$$

In view of (6.37a) and (6.37b) these are precisely the "instantaneous" and "endpoint" saddle point conditions asserted in the theorem. □

Theorem 6.5 has an interesting interpretation in the context of the finite-dimensional linear-quadratic programming problems in §2, as revealed by its proof. We shall formulate this as a corollary.

Corresponding to the trajectories $\bar x$ and $\bar y$, consider the "instantaneous" primal and dual problems associated with the linear-quadratic form $\bar J_t(u, v)$ on $U(t)\times V(t)$, where $\bar J_t$ is given by (6.34), namely:

$(\mathcal{P}_t(\bar x, \bar y))$	minimize $\bar f_t(u)$ over $u \in V(t)$	where $\bar f_t$ is given by (6.30),
$(\mathcal{Q}_t(\bar x, \bar y))$	maximize $\bar g_t(v)$ over $v \in V(t)$	where $\bar g_t$ is given by (6.31).

Consider too the "endpoint" primal and dual problems associated with the linear-quadratic form $\bar J_e(u_e, v_e)$ on $U_e \times V_e$, where $\bar J_e$ is given by (6.35), namely:

$(\mathcal{P}_e(\bar x, \bar y))$	minimize $\bar f_e(u_e)$ over $u_e \in U_e$	where $\bar f_e$ is given by (6.32),
$(\mathcal{Q}_e(\bar x, \bar y))$	maximize $\bar g_e(v_e)$ over $v_e \in V_e$	where $\bar g_e$ is given by (6.33).

COROLLARY 6.6. *For $((\bar u, \bar u_e), (\bar v, \bar v_e))$ to be a saddle point of J on $\mathcal{U}\times\mathcal{V}$, it is necessary and sufficient that the following conditions hold (in addition to $u(t)$ and $v(t)$ being $\mathcal{L}^1$ in t). For almost every $t \in [t_0, t_1]$*

$$\text{(6.38)} \qquad \begin{array}{l} \bar u(t) \textit{ solves the instantaneous primal } (\mathcal{P}_t(\bar x, \bar y)), \textit{ and} \\ \bar v(t) \textit{ solves the instantaneous dual } (\mathcal{Q}_t(\bar x, \bar y)), \end{array}$$

and furthermore

$$\text{(6.39)} \qquad \begin{array}{l} \bar u_e \textit{ solves the endpoint primal } (\mathcal{P}_e(\bar x, \bar y)), \textit{ and} \\ \bar v_e \textit{ solves the endpoint dual } (\mathcal{Q}_e(\bar x, \bar y)). \end{array}$$

Proof. Because the instantaneous and endpoint problems fall in the category of finite-dimensional linear-quadratic programming, we can apply Theorem 2.2 to them and see that (6.38) entails

$$\min(\mathcal{P}_t(\bar x, \bar y)) = \max(\mathcal{Q}_t(\bar x, \bar y)),$$

and (6.39) entails

$$\min(\mathcal{P}_e(\bar x, \bar y)) = \max(\mathcal{Q}_e(\bar x, \bar y)).$$

It follows then from Proposition 2.1 that (6.38) is equivalent to (6.26), whereas (6.39) is equivalent to (6.27). □

Our final result extends Theorem 2.5 to the infinite-dimensional case. It provides a basis for the idea that in intertemporal linear-quadratic programming as well as in finite-dimensional linear-quadratic programming, a given pair of problems $(\mathcal{P})$ and $(\mathcal{Q})$ can often be remodeled, at least for computational purposes, by a more tractable pair $(\widehat{\mathcal{P}})$ and $(\widehat{\mathcal{Q}})$ in the pattern of *bounded* linear or quadratic programming as in Examples 5.4 and 5.7.

THEOREM 6.7. *Consider along with* $(\mathcal{P})$ *and* $(\mathcal{Q})$ *an auxiliary pair of problems* $(\widehat{\mathcal{P}})$ *and* $(\widehat{\mathcal{Q}})$ *under the same assumptions and defined by the same data, except with the control sets* $U(t), U_e, V(t)$, *and* V_e *replaced by sets*

$$\widehat{U}(t) \subset U(t), \quad \widehat{U}_e \subset U_e, \quad \widehat{V}(t) \subset V(t), \quad \widehat{V}_e \subset V_e. \tag{6.40}$$

Suppose $\min(\widehat{\mathcal{P}}) = \max(\widehat{\mathcal{Q}})$, *as would be true in particular by Theorem* 6.3 *if the sets* $\widehat{U}(t), \widehat{U}_e, \widehat{V}(t)$ *and* $\widehat{V}_e$ *are all bounded.*

(a) *If* $(\overline{u}, \overline{u}_e)$ *and* $(\overline{v}, \overline{v}_e)$ *satisfy the instantaneous and endpoint conditions in Theorem* 6.5 (*or Corollary* 6.6) *and also are such that*

$$\overline{u}(t) \in \widehat{U}(t) \text{ a.e.}, \quad \overline{u}_e \in \widehat{U}_e, \quad \overline{v}(t) \in \widehat{V}(t) \text{ a.e.}, \quad \overline{v}_e \in \widehat{V}_e,$$

then $(\overline{u}, \overline{u}_e)$ *solves not only* $(\mathcal{P})$ *but* $(\widehat{\mathcal{P}})$, *and* $(\overline{v}, \overline{v}_e)$ *solves not only* $(\mathcal{Q})$ *but* $(\widehat{\mathcal{Q}})$.

(b) *If* $(\overline{u}, \overline{u}_e)$ *solves* $(\widehat{\mathcal{P}})$ *and* $(\overline{v}, \overline{v}_e)$ *solves* $(\widehat{\mathcal{Q}})$, *and if* $U(t)$ *and* $V(t)$ *coincide with* $\widehat{U}(t)$ *and* $\widehat{V}(t)$ *around* $\overline{u}(t)$ *and* $\overline{v}(t)$ *for almost every* t, *while* $\widehat{U}_e$ *and* $\widehat{V}_e$ *coincide with* U_e *and* V_e *around* $\overline{u}_e$ *and* $\overline{v}_e$, *then actually* $(\overline{u}, \overline{u}_e)$ *solves* $(\mathcal{P})$ *and* $(\overline{v}, \overline{v}_e)$ *solves* $(\mathcal{Q})$.

(The terminology about "coinciding" is defined in the statement of Theorem 2.5.)

Proof. Under the assumptions in (a), $(\overline{u}, \overline{u}_e)$ and $(\overline{v}, \overline{v}_e)$ give a saddle point of J on $\mathcal{U} \times \mathcal{V}$ (by Theorem 6.5, or as the case may be, Corollary 6.6), and this saddle point happens to lie in $\widehat{\mathcal{U}} \times \widehat{\mathcal{V}}$ (where $\widehat{\mathcal{U}}$ and $\widehat{\mathcal{V}}$ are the control spaces corresponding to $(\widehat{\mathcal{P}})$ and $(\widehat{\mathcal{Q}})$). Then $((\overline{u}, \overline{u}_e), (\overline{v}, \overline{v}_e))$ is also a saddle point for J relative to $\widehat{\mathcal{U}} \times \widehat{\mathcal{V}}$. Theorem 6.2, applied to both pairs of problems, yields the conclusions.

Under the assumptions in (b) we know by Theorem 6.2, as applied to $(\widehat{\mathcal{P}})$ and $(\widehat{\mathcal{Q}})$, that $((\overline{u}, \overline{u}_e), (\overline{v}, \overline{v}_e))$ is a saddle point of J relative to $\widehat{\mathcal{U}} \times \widehat{\mathcal{V}}$. The instantaneous conditions and endpoint conditions in Theorem 6.5 must therefore be satisfied relative to $\widehat{U}(t) \times \widehat{V}(t)$ and $\widehat{U}_e \times \widehat{V}_e$. But by Theorem 2.5 and our hypothesis about the sets coinciding locally, the same conditions are then satisfied relative to $U(t) \times V(t)$ and $U_e \times V_e$. Theorem 6.5 tells us now that $((\overline{u}, \overline{u}_e), (\overline{v}, \overline{v}_e))$ is a saddle point also for J relative to $\mathcal{U} \times \mathcal{V}$. Then $(\overline{u}, \overline{u}_e)$ and $(\overline{v}, \overline{v}_e)$ are optimal for $(\mathcal{P})$ and $(\mathcal{Q})$ by Theorem 6.2. □

To make the best use of Theorem 6.7 in the manner outlined at the end of §3 for the finite-dimensional case, it would be helpful to have criteria under which $(\mathcal{P})$ and $(\mathcal{Q})$ have solutions $(\overline{u}, \overline{u}_e)$ and $(\overline{v}, \overline{v}_e)$ with $\overline{u}$ and $\overline{v}$ actually in $\mathcal{L}^\infty$. Then, for example, Theorem 6.7 can be applied with the subsets (6.40) taken to be intervals adequately large. Such criteria can be developed, but we shall not address the issue here. Results of this nature for the cases covered by continuous-time programming may be gleaned from Grinold [5], [6] and Reiland [8], [15].

REFERENCES

[1] R.E. BELLMAN, *Dynamic Programming*, Princeton Univ. Press, 1957.

[2] W.F. TYNDALL, *A duality theorem for a class of continuous linear programming problems*, SIAM J. Appl. Math., 13 (1965), pp. 644–666.

[3] ——, *An extended duality theorem for continuous linear programming problems*, SIAM J. Appl. Math., 15 (1967), pp. 1294–1298.

[4] N. LEVINSON, *A class of continuous linear programming problems*, J. Math. Anal. Appl., 16 (1966), pp. 73–83.

[5] R. GRINOLD, *Symmetric duality for continuous linear programs*, SIAM J. Appl. Math., 18 (1970), pp. 84–97.

[6] ——, *Continuous programming part one: linear objectives*, J. Math. Anal. Appl., 28 (1969), pp. 32–51.

[7] M. SCHECTER, *Duality in continuous linear programming*, J. Math. Anal. Appl., 37 (1972), pp. 130–141.

[8] T.W. REILAND, *Optimality conditions and duality in continuous programming*, II: *The linear problem revisited*, J. Math. Anal. Appl., 77 (1980), pp. 329–343.

[9] R. MEIDAN AND A.F. PEROLD, *Optimality conditions and strong duality in abstract and continuous time linear programming*, J. Optim. Theory Appl., 40 (1983), pp. 61–76.

[10] M.A. HANSON, *Duality for a class of infinite programming problems*, SIAM J. Appl. Math., 16 (1968), pp. 318–323.

[11] M.A. HANSON AND B. MOND, *A class of continuous convex programming problems*, J. Math. Anal. Appl., 22 (1968), pp. 427–437.

[12] R. GRINOLD, *Continuous programming part two: nonlinear objectives*, J. Math. Anal. Appl., 27 (1969), pp. 639–655.

[13] W.H. FARR AND M.A. HANSON, *Continuous time programming with nonlinear constraints*, J. Math. Anal. Appl., 45 (1974), pp. 96–115.

[14] T.W. REILAND AND M.A. HANSON, *Generalized Kuhn–Tucker conditions and duality for continuous nonlinear programming problems*, J. Math. Anal. Appl., 74 (1980), pp. 578–598.

[15] T.W. REILAND, *Optimality conditions and duality in continuous programming*, I: *Convex programs and a theorem of the alternative*, J. Math. Anal. Appl., 77 (1980), pp. 297–325.

[16] R.T. ROCKAFELLAR AND R.J.-B. WETS, *A dual solution procedure for quadratic stochastic programs with simple recourse*, in Numerical Methods, V. Pereyra and A. Reinoza, eds., Lecture Notes in Math. 1005, Springer-Verlag, Berlin, New York, 1983, pp. 252–265.

[17] ——, *A Lagrangian finite generation technique for solving linear quadratic problems in stochastic programming*, Math. Programming Stud., 28 (1986), pp. 63–93.

[18] ——, *Linear quadratic programming problems with stochastic penalties: the finite generation algorithm*, in Numerical Techniques for Stochastic Optimization Problems, Y. Ermoliev and R.J.-B. Wets, eds., Springer-Verlag, Berlin, New York, 1986.

[19] A.F. PEROLD, *Fundamentals of a continuous-time simplex method*, Technical Report SOL 78–26 (1978), Dept. of Operations Research, Stanford Univ., Stanford, CA.

[20] ——, *Extreme points and basic feasible solutions in continuous-time linear programming*, this Journal, 19 (1981), pp. 52–63.

[21] K.M. ANSTREICHER, *Generation of feasible descent directions in continuous time linear programming*, Technical Report SOL 83–18 (1983), Dept. of Operations Research, Stanford Univ., Stanford, CA.

[22] F.H. CLARKE, *Optimization and Nonsmooth Analysis*, Wiley–Interscience, New York, 1983.

[23] R.T. ROCKAFELLAR, *Conjugate convex functions in optimal control and the calculus of variations*, J. Math. Anal. Appl., 32 (1970), pp. 174–222.

[24] ——, *Generalized Hamiltonian equations for convex problems of Lagrange*, Pacific J. Math., 33 (1970), pp. 411–428.

[25] ——, *Existence and duality theorems for convex problems of Bolza*, Trans. Amer. Math. Soc., 159 (1971), pp. 1–40.

[26] ——, *State constraints in convex problems of Bolza*, SIAM J. Control, 10 (1972), pp. 691–715.

[27] ——, *Semigroups of convex bifunctions generated by Lagrange problems in the calculus of variations*, Math. Scand., 36 (1975), pp. 137–158.

[28] ——, *Dual problems of Lagrange for arcs of bounded variation*, in Calculus of Variations and Control Theory, D.L. Russell, ed., Academic Press, New York, 1976, pp. 155–192.

[29] ——, *Optimality conditions for convex control problems with nonnegative states and the possibility of jumps*, in Game Theory and Mathematical Economics, O. Moeschlin, ed., North-Holland, Amsterdam, New York, 1981, pp. 339–349.

[30] ——, *Duality in optimal control*, in Mathematical Control Theory, W.A. Coppel, ed., Lecture Notes in Math., Springer-Verlag, Berlin, New York, 680 (1978), pp. 219–257.

[31] J.M. MURRAY, *Some existence and regularity results for dual linear control problems*, J. Math. Anal. Appl., 112 (1985), pp. 190–209.

[32] R.T. ROCKAFELLAR, *Conjugate Duality and Optimization*, Regional Conference Series in Appl. Math., 61, Society for Industrial and Applied Mathematics, Philadelphia, 1974.

[33] W.S. DORN, *Duality in quadratic programming*, Quart. Appl. Math., 18 (1960), pp. 155–162.

[34] R.W. COTTLE, *Symmetric dual quadratic programs*, Quart. Appl. Math., 21 (1963), pp. 237–243.

[35] M. FRANK AND P. WOLFE, *An algorithm for quadratic programming*, Naval Res. Logist. Quart., 3 (1956), pp. 95–110.

[36] R.T. ROCKAFELLAR, *Convex Analysis*, Princeton Univ. Press, Princeton, NJ, 1970.

[37] S.M. ROBINSON, *Some continuity properties of polyhedral multifunctions*, Math. Programming Stud., 14 (1981), pp. 206–214.

[38] R. FLETCHER, *Penalty functions*, in Mathematical Programming: The State of the Art, A. Bachem et al., eds, Springer–Verlag, 1983, pp. 87–114.

[39] R. FOURER, *A simplex algorithm for piecewise linear programming*, I: *Derivation and proof*, Math. Programming, 33 (1985), pp. 204–233.

[40] R.T. ROCKAFELLAR, *Network Flows and Monotropic Optimization*, Wiley–Interscience, New York, 1984.

[41] A. KING, R.T. ROCKAFELLAR, L. SOMYODY AND R. J.-B. WETS, *Lake eutrophication management: the Lake Balaton project*, in Numerical Techniques for Stochastic Optimization Problems, Y. Ermoliev and R. J.-B. Wets, eds., Springer–Verlag, Berlin, New York, 1986.

[42] G. SALINETTI AND R.J.-B. WETS, *On the convergence of sequences of convex sets in finite dimensions*, SIAM Rev., 21 (1979), pp. 18–33.

[43] R.T. ROCKAFELLAR, *Integrals which are convex functionals*, II, Pacific J. Math., 39 (1971), pp. 439–469.

[44] R.A. WIJSMAN, *Convergence of sequences of convex sets, cones and functions*, II, Trans. Amer. Math. Soc., 123 (1966), pp. 32–45.

[45] R.J.-B. WETS, *Convergence of convex functions, variational inequalities and convex optimization problems*, in Variational Inequalities and Complementarity Problems, R.W. Cottle et al., eds., John Wiley, New York, 1980, pp. 375–403.

[46] L. MCLINDEN AND R.C. BERGSTROM, *Preservation of convergence of convex sets and functions*, Trans. Amer. Math. Soc., 268 (1981), pp. 127–142.

[47] E. MICHAEL, *Continuous selections*, I, Ann. of Math., 63 (1956), pp. 361–382.

[48] R.T. ROCKAFELLAR, *Lipschitzian properties of multifunctions*, J. Nonlin. Anal. Th. Meth. Appl., 9 (1985), pp. 867–885.

[49] R.T. ROCKAFELLAR, *Integrals which are convex functionals*, Pacific J. Math., 24 (1968), pp. 525–539.

[50] R.T. ROCKAFELLAR, *Integral functionals, normal integrands and measurable selections*, in Nonlinear Operators and the Calculus of Variations, L. Waelbroeck, ed., Lecture Notes in Math., Springer-Verlag, Berlin, New York, 543, 1976, pp. 157–207.

[51] J.-J. MOREAU, *Théorèmes* 'inf-sup,' C.R. Acad. Sci. Paris, 258 (1964), pp. 2720–2722.