HAMILTONIAN TRAJECTORIES AND DUALITY IN THE OPTIMAL CONTROL OF LINEAR SYSTEMS WITH CONVEX COSTS

*R.T. Rockafellar**

Abstract. A duality theorem is proved for problems of optimal control of linear dynamical systems in continuous time subject to linear constraints and convex costs, such as penalties. Optimality conditions are stated in terms of a "mini-maximum principle" in which the primal and dual control vectors satisfy a saddle point condition at almost every instant of time. This principle is shown to be equivalent to a generalized Hamiltonian differential equation in the primal and dual state variables, along with a transversality condition which likewise is in Hamiltonian form.

Keywords: convex optimal control, duality, Hamiltonian trajectories, generalized problems of Bolza, calculus of variations, continuous convex programming, intertemporal convex programming

* This work was supported in part by grants from the National Science Foundation and the Air Force Office of Scientific Research at the University of Washington, Seattle.

1. Introduction

This paper focuses on optimal control problems of convex type and the special properties they enjoy, in particular properties of duality. A fundamental problem form, intended for approximations of more complicated control situations as well as direct use in mathematical modeling, is introduced in terms of linear dynamics and linear constraints that may be represented by penalties, either finite or infinite. A duality theorem is proved and made the basis for deriving necessary and sufficient conditions for the optimality of control functions and state trajectories. The work extends the author's recent results on continuous time problems with piecewise linear-quadratic costs [**1**], [**2**]. It ties in more generally with the theory of dual problems of Bolza in the calculus of variations, as developed earlier by methods of convex analysis in Rockafellar [**3**], [**4**]. A bridge is thereby provided to a conceptual framework dominated by a Hamiltonian function and its gradients or subgradients in the expression of optimality condition.

The chief aim, besides setting up the duality, is to demonstrate that solutions to problems in the chosen class can be characterized in two quite different, yet equivalent ways. First there is a "minimaximum principle" which expresses the primal and dual optimal control vectors at any time as giving a saddle point of a certain convex-concave function. Second there is a generalized Hamiltonian differential equation in terms of primal and dual states but no direct mention of controls.

The minimaximum principle is suggestive of computational approaches that depend on generating sequences of control functions as in various algorithms of convex programming. The Hamiltonian system, on the other hand, is of interest in that it can be solved like an ordinary differential equation from any choice of initial primal and dual states. While this may or may not be a practical tool in calculating optimal trajectories, it reveals important information about such trajectories, for example that under our assumptions they can be realized by optimal control functions that are essentially bounded. Knowledge of the Hamiltonian function is crucial also to the prospects of applying Hamilton-Jacobi theory in its latest forms to convex problems of optimal control.

The model problem we start from is not the broadest possible that would fit under the heading of convex optimal control. It is selected rather to yield strong results while still encompassing a wide spectrum of applications. The details of structure are designed to facilitate dualization.

To help keep formulas compact and readable, we write x_t and u_t as the state and control vectors at time t instead of $x(t)$ and $u(t)$. These vectors belong to $\mathbb{R}^n$ and $\mathbb{R}^k$, respectively. We also make use of an auxiliary control vector $u_e \in \mathbb{R}^{k_e}$ which affects endpoint costs and constraints; the subscript $_e$ is utilized also to designate data elements connected with endpoints. See [**1**] for a discussion of the modeling possibilities with end-point controls. Inner products of vectors in $\mathbb{R}^n$ and $\mathbb{R}^k$ will be expressed in the notation $\langle \cdot, \cdot \rangle$ and the Euclidean norm by $|\cdot|$.

We denote by $\mathcal{U}$ the space of all control elements u consisting of a choice of vector u_e and an *essentially bounded*, measurable function $t \mapsto u_t$ defined over the interval $[t_0, t_1]$, which is fixed throughout the paper. We handle $\mathcal{U}$ as a Banach space in the norm $\|u\| =$

$\max\{|u_e|, \operatorname{ess\,sup}_t |u_t|\}$. Each $u \in \mathcal{U}$ will determine a state trajectory $x : t \mapsto x_t \in \mathbb{R}^n$, which will be Lipschitz continuous over $[t_0, t_1]$. The time derivative of x_t, which exists for almost every t, will be denoted by $\dot{x}_t$. The space of all such Lipschitz continuous arcs x in $\mathbb{R}^n$ will be denoted by $\mathcal{A}^\infty = \mathcal{A}_n^\infty[t_0, t_1]$. This is a Banach space in the norm $\|x\|_\infty = \max\{|x_{t_0}|, \operatorname{ess\,sup}_t \dot{x}_t\}$. (The superscript ∞ is a reminder that the derivative function $t \mapsto \dot{x}_t$ belongs to $\mathcal{L}_n^\infty[t_0, t_1]$.)

The control problem we address takes the form

$$(\mathcal{P}) \qquad \begin{aligned} &\text{minimize the functional } F(u) = \\ &\quad \int_{t_0}^{t_1} [\langle p_t, u_t\rangle + \varphi_t(u_t) + \psi_t(q_t - C_t x_t - D_t u_t) - \langle c_t, x_t\rangle]\, dt \\ &\quad + [\langle p_e, u_e\rangle + \varphi_e(u_e) + \psi_e(q_e - C_e x_{t_1} - D_e u_e) - \langle c_e, x_{t_1}\rangle] \\ &\text{over } u \in \mathcal{U}, \text{ where } x \text{ is determined from } u \text{ by} \\ &\quad \dot{x}_t = A_t x_t + B_t u_t + b_t \ \text{ a.e.}, \quad x_{t_0} = B_e u_e + b_e. \end{aligned}$$

Here φ_t and φ_e are functions on $\mathbb{R}^k$ and $\mathbb{R}^{ke}$, while ψ_t and ψ_e are functions on certain spaces $\mathbb{R}^l$ and $\mathbb{R}^{le}$. The dimensions of the various vectors and matrices in $(\mathcal{P})$ are of course completely determined by the dimensions of these spaces. In general we assume:

(A1) $\varphi_t, \varphi_e, \psi_t, \psi_e$, are lower semicontinuous, proper, convex functions.
(A2) φ_t and ψ_t depend epi-continuously on $t \in [t_0, t_1]$.
(A3) $A_t, B_t, b_t, C_t, c_t, D_t, p_t, q_t$, depend continuously on $t \in [t_0, t_1]$.

By (A3) we are assured in particular that each choice of $u \in \mathcal{U}$ gives rise to a unique trajectory x, which belongs to the space $\mathcal{A}^\infty$ because the function $t \mapsto A_t x_t + B_t u_t + b_t$ is essentially bounded. The mapping $u \mapsto x$ is continuous. The properness in (A1) asserts that the functions $\varphi_t, \varphi_e, \psi_t, \psi_e$, do not take on the value $-\infty$, although they might in some cases take on ∞ as long as they do not have this value everywhere. The role of ∞ is to provide an infinite penalty for certain constraint violations; more about this in a moment.

Assumption (A2) means that the epigraphs sets $\operatorname{epi} \varphi_t$ and $\operatorname{epi} \psi_t$, which are closed convex subsets of $\mathbb{R}^{k+1}$ and $\mathbb{R}^{l+1}$, depend continuously on t in the sense of set convergence. This form of continuity has been studied by many authors in recent years; see Salinetti and Wets [**5**], Wets [**6**], for properties and references. As a special case, of course, epi-continuity is present when φ_t and ψ_t do not actually vary with t.

Proposition 1.1. *Under (A1)–(A3), the functional F in problem $(\mathcal{P})$ is well defined on the Banach space $\mathcal{U}$ with values in $(-\infty, \infty]$. Furthermore, F is convex and lower semicontinuous.*

Proof. The epi-continuity of $t \mapsto \operatorname{epi} \varphi_t$ in (A2) entails that the function $(t, w) \mapsto \varphi_t(w)$ is lower semicontinuous on $[t_0, t_1] \times \mathbb{R}^k$. This function is definitely therefore a normal integrand in the sense of [**7**] and is bounded below on $[t_0, t_1] \times W$ for every bounded set $W \subset \mathbb{R}^k$. It follows that $\varphi_t(u_t)$ is measurable in t when u_t is measurable in t, and it is essentially bounded from below when u_t is essentially bounded in t. For any $u \in \mathcal{U}$, then, the integral of $\varphi_t(u_t)$ has a well defined value in $(-\infty, \infty]$. Similar properties hold for ψ_t. Since (A3) implies $q_t - C_t x_t - D_t u_t$ is a bounded measurable function of t when u_t is such a function of t (here we note that x_t, as determined by the dynamics, is continuous in t), we conclude that the integral of $\psi_t(q_t - c_t x_t - D_t u_t)$ likewise has a well defined value in $(-\infty, \infty]$ for any $u \in \mathcal{U}$. Thus $F(u)$ is well defined on $\mathcal{U}$ with values in $(-\infty, \infty]$. The convexity of F follows obviously from the convexity in (A1) and the fact that the dynamical mapping $u \mapsto x$ is affine. Lower semicontinuity in the norm topology of $\mathcal{U}$ follows from the lower semicontinuity in (A1) and continuity in (A3), as well as the continuity of $u \mapsto x$, by Fatou's lemma, cf. [**7**]. $\square$

Problem $(\mathcal{P})$ may involve *implicit constraints* beyond the ones already mentioned, due to the possiblity of ∞ values. Recall that the effective domain $\operatorname{dom} F$ consists of the elements $u \in \mathcal{U}$ such that $F(u) < \infty$; similarly for $\operatorname{dom} \varphi_t$ and $\operatorname{dom} \varphi_e$. Minimizing F over $\mathcal{U}$ is the same as minimizing F over $\operatorname{dom} F$. Obviously the condition $u \in \operatorname{dom} F$ requires u to belong to the set

$$U := \{u \in \mathcal{U} \mid \int_{t_0}^{t_1} \varphi_t(u_t)\, dt < \infty \text{ and } \varphi_e(u_e) < \infty\} \tag{1.1}$$

and satisfy

$$u_t \in U_t \text{ a.e. and } u_e \in U_e, \text{ where } U_t := \operatorname{dom} \varphi_t \text{ and } U_e := \operatorname{dom} \varphi_e, \tag{1.2}$$

$$\begin{gathered} q_t - C_t x_t - D_t u_t \in R_t \text{ a.e. and } q_e - C_e x_{t_1} - D_e u_e \in R_e, \\ \text{where } R_t := \operatorname{dom} \psi_t \text{ and } R_e := \operatorname{dom} \psi_e. \end{gathered} \tag{1.3}$$

The control problem *dual* to $(\mathcal{P})$ involves dual states $y_t \in \mathbb{R}^n$ and dual controls $v_t \in \mathbb{R}^l$ and $v_e \in \mathbb{R}^{l_e}$. Let us denote by $\mathcal{V}$ the set of control elements v consisting of a choice of v_e and an essentially bounded, measurable function $t \mapsto v_t$. This is a Banach space in the same way as described above for $\mathcal{U}$. The dual problem takes the form

$$(\mathcal{Q}) \quad \begin{aligned} &\text{maximize the functional } G(v) = \\ &\int_{t_0}^{t_1} [\langle q_t, v_t\rangle - \psi_t^*(v_t) - \varphi_t^*(B_t^* y_t + D_t^* v_t - p_t) - \langle b_t, y_t\rangle]\, dt \\ &\quad + [\langle q_e, v_e\rangle - \psi_e^*(v_e) - \varphi_e^*(B_e^* y_{t_0} + D_e^* v_e - p_e) - \langle b_e, y_{t_0}\rangle] \\ &\text{over } v \in \mathcal{V}, \text{ where } y \text{ is determined from } v \text{ by} \\ &\quad -\dot{y}_t = A_t^* y_t + C_t^* v_t + c_t \text{ a.e.}, \quad y_{t_1} = C_e^* v_e + c_e. \end{aligned}$$

The asterisk on a matrix denotes transpose, but on a convex function it indicates the conjugate function (Legendre-Fenchel transform) in the sense of convex analysis [**8**]. Note that the dual dynamical system goes backward in time and uniquely determines a Lipschitz continuous trajectory $y \in \mathcal{A}^\infty$ for each $v \in \mathcal{V}$.

Just as (A3) is preserved in passing to transposes, assumptions (A1) and (A2) imply the corresponding properties for the conjugate functions:

(A1*) $\varphi_t^*, \varphi_e^*, \psi_t^*, \psi_e^*$, are lower semicontinuous, proper, convex functions.
(A2*) φ_t^* and ψ_t^* depend epi-continuously on $t \in [t_0, t_1]$.

The equivalence between (A2) and (A2*) follows from Wijsman's theorem [**9**] on the continuity of the Legendre-Fenchel transform with respect to epi-convergence. We immediately therefore get the version of Proposition 1.1 that applies to the dual problem.

Proposition 1.2. *Under (A1)–(A3), the functional G in problem $(\mathcal{Q})$ is well defined on the Banach space $\mathcal{V}$ with values in $[-\infty, \infty)$. Furthermore, G is concave and upper semicontinuous.*

Implicit in $(\mathcal{Q})$ are the constraints that v should belong to the set

$$V := \{v \in \mathcal{V} \mid \int_{t_0}^{t_1} \psi_t^*(v_t)\,dt < \infty \text{ and } \psi_e^*(v_e) < \infty\} \tag{1.4}$$

and satisfy

$$v_t \in V_t \text{ a.e. and } v_e \in V_e, \text{ where } V_t := \operatorname{dom} \psi_t^* \text{ and } V_e := \operatorname{dom} \psi_e^*, \tag{1.5}$$

$$\begin{gathered} B_t^* y_t + D_t^* v_t - p_t \in S_t \text{ a.e. and } B_e^* y_{t_0} + D_e^* v_e - p_e \in S_e, \\ \text{where } S_t := \operatorname{dom} \varphi_t^* \text{ and } S_e := \operatorname{dom} \varphi_e^*. \end{gathered} \tag{1.6}$$

The special case of these primal and dual problems that was treated in [**1**] and [**2**] as *extended linear-quadratic* optimal control is obtained by taking

$$\begin{aligned} \varphi_t(u_t) &= \tfrac{1}{2}\langle u_t, P_t u_t\rangle \text{ for } u_t \in U_t, \quad \varphi_t(u_t) = \infty \text{ for } u_t \notin U_t, \\ \varphi_e(u_e) &= \tfrac{1}{2}\langle u_e, P_e u_e\rangle \text{ for } u_e \in U_e, \quad \varphi_e(u_e) = \infty \text{ for } u_e \notin U_e, \\ \psi_t^*(v_t) &= \tfrac{1}{2}\langle v_t, Q_t v_t\rangle \text{ for } v_t \in V_t, \quad \psi_t^*(v_t) = \infty \text{ for } v_t \notin V_t, \\ \psi_e^*(v_e) &= \tfrac{1}{2}\langle v_e, Q_e v_e\rangle \text{ for } v_e \in V_e, \quad \psi_e^*(v_e) = \infty \text{ for } v_e \notin V_e, \end{aligned} \tag{1.7}$$

for *polyhedral* sets U_t, U_e, V_t, V_e, and *positive semidefinite* symmetric matrices P_t, P_e, Q_t, Q_e. The philosophy behind this is fully explained in [**1**] and will not be repeated here, except to say that the functions $\psi_t, \psi_e, \varphi_t^*, \varphi_e^*$, are then piecewise linear-quadratic and yield a version of linear-quadratic optimal control in which piecewise linear-quadratic penalty terms may be present and are readily dualized.

In general, the terms involving ψ_t and ψ_e in $(\mathcal{P})$ may be viewed as *monitoring* the vectors $s_t = q_t - C_t x_t - D_t u_t$ and $s_e = q_e - C_e x_{t_1} - D_e u_e$. A simple example would be the one where ψ_t vanishes on a certain set K but has the value ∞ outside of K. Then the ψ_t term expresses through infinite penalties the condition that $s_t \in K$ a.e. This condition might represent a system of equations or inequalities. Instead ψ_t could have finite, positive values outside of K, and then we would have a finite penalty representation of such a constraint system. Similarly, ψ_e could play this role for constraints on the endpoint x_{t_1}, while φ_t^*

and φ_e^* could have such interpretations in the dual problem. Many examples are worked out in [**1**].

Our strongest results will eventually call for a further assumption:

(A4) φ_t and φ_e are coercive, while ψ_t and ψ_e are everywhere finite.

Coercivity of φ_t means that $\lim_{|w|\to\infty} \varphi_t(w)/|w| = \infty$, which is true in particular when the control set U_t in (1.2) is bounded; similarly for φ_e and U_e. It is known from convex analysis [**8**, §13] that φ_t and φ_e are coercive if and only if the conjugate functions φ_t^* and φ_e^* are finite everywhere. Likewise, ψ_t and ψ_e are finite everywhere if and only if ψ_t^* and ψ_e^* are coercive. Thus (A4), like the earlier assumptions, has an equivalent dual form:

(A4*) ψ_t^* and ψ_e^* are coercive, while φ_t^* and φ_e^* are everywhere finite.

The interpretation of (A4), then, is that *there are effectively no exact implicit constraints of type* (1.3) *and* (1.6) *in the primal and dual problems.* In other words, this additional assumption corresponds to the situation where all the monitoring of $q_t - C_t x_t - D_t u_t$ and $q_e - C_e x_{t_1} - D_e u_e$ in the primal problem and of $B_t^* y_t + D_t^* v_t - p_t$ and $B_e^* y_{t_0} + D_e^* v_e - p_e$ in the dual problem proceeds with *finite* values: no infinite penalties. Such a property may naturally be present in a given application, or it may be achieved as a mode of approximation for a problems one is really interested in. Anyway, one may argue that it is vital for the development of computational methods for problems like $(\mathcal{P})$ and $(\mathcal{Q})$. Conditions on x or on x and u jointly that are modeled as exact constraints can lead to serious numerical complications, whereas such conditions on u alone, as in (1.2), present relatively little difficulty. See [**1**] for more on this issue.

Proposition 1.3. *Under (A4), the epi-continuity assumption (A2) is equivalent to having $\varphi_t^*(r)$ and $\psi_t(s)$ be continuous in $t \in [t_0, t_1]$ for each $r \in \mathbb{R}^k$ and $s \in \mathbb{R}^l$. Then in fact $\varphi_t^*(r)$ is continuous with respect to (t, r), and $\psi_t(s)$ is continuous with respect to (t, s).*

Proof. For finite convex functions, epi-continuity with respect to t is equivalent to pointwise continuity with respect to t; see Wets [**5**, Corollaries 4 and 5]. Further, finite convex functions whose values depend continuously on t are jointly continuous in t and their other variables [**8**, Theorem 10.7]. □

Proposition 1.4. *Under assumptions (A1)–(A3), the sets U and V in (1.1) and (1.4) are convex and nonempty. When (A4) holds too, U is identical to the set of feasible controls for $(\mathcal{P})$, i.e. the elements $u \in \mathcal{U}$ for which $F(u)$ is finite, and likewise V is the set of feasible controls for $(\mathcal{Q})$. In particular, feasible controls do exist then for both problems.*

Proof. The convexity of U and V is obvious from their definitions by the convexity in (A1). Clearly $F(u) = \infty$ when $u \notin U$, and $G(v) = -\infty$ for $v \notin V$. According to (A2), the multifunction $t \mapsto \operatorname{epi}\varphi_t$, whose values are nonempty closed convex sets by (A1), is continuous. For such a multifunction the continuous selection theorem of Michael [**10**] applies: it is possible to choose $(u_t, \alpha_t) \in \operatorname{epi}\varphi_t$ continuously with respect to $t \in [t_0, t_1]$. Then $\varphi_t(u_t) \leq \alpha_t$, so the integral of $\varphi_t(u_t)$ cannot be ∞ and therefore must be finite. Taking any u_e in U_e, a set which is nonempty by the properness of φ_e in (A1), we obtain a control element $u \in U$. Thus $U \neq \emptyset$. Any $u \in U$, on the other hand, makes all the terms in the formula for $F(u)$ in $(\mathcal{P})$ be finite except perhaps for the integral of $\psi_t(s_t)$, where $s_t = q_t - C_t x_t + D_t u_t$. The function $t \mapsto s_t$ is essentially bounded in t by (A3). The continuity of $(t,s) \mapsto \psi_t(s)$, which is asserted by Proposition 1.3, implies that the latter function is bounded on $[t_0, t_1] \times W$ for any bounded set $W \subset \mathbb{R}^l$. We thereby obtain the essential boundedness of $\psi_t(s_t)$ in t and hence the finiteness of its integral. This yields the desired conclusion in the case of $(\mathcal{P})$. The corresponding result for $(\mathcal{Q})$ follows by duality. $\square$

2. Minimax representation.

The close relationship between problems $(\mathcal{P})$ and $(\mathcal{Q})$ that leads to their being called dual to each other stems from a joint representation in terms of a minimax problem in $\mathcal{U} \times \mathcal{V}$. To give this, we introduce the functional

$$J(u,v) := \int_{t_0}^{t_1} J_t(u_t, v_t)\, dt + J_e(u_e, v_e) - j(u,v) \tag{2.1}$$

in the notation

$$\begin{aligned} J_t(u_t, v_t) &:= \langle p_t, u_t\rangle + \langle q_t, v_t\rangle - \langle v_t, D_t u_t\rangle + \varphi_t(u_t) - \psi_t^*(v_t), \\ J_e(u_e, v_e) &:= \langle p_e, u_e\rangle + \langle q_e, v_e\rangle - \langle v_e, D_t u_t\rangle + \varphi_e(u_e) - \psi_e^*(v_e), \end{aligned} \tag{2.2}$$

and with j taken to be the bi-affine functional on $\mathcal{U} \times \mathcal{V}$ that corresponds to the dynamics and is expressed in terms of the trajectories x and y associated with u and v by

$$\begin{aligned} j(u,v) :&= \int_{t_0}^{t_1} \langle y_t, B_t u_t + b_t\rangle\, dt + \langle y_{t_0}, B_e u_e + b_e\rangle \\ &= \int_{t_0}^{t_1} \langle x_t, C_t^* v_t + c_t\rangle\, dt + \langle x_{t_1}, C_e^* v_e + c_e\rangle. \end{aligned} \tag{2.3}$$

(The validity of the equation in (2.3) is proved in [**1**, §6].) Because some of the terms in (2.2) can take on the value ∞ while others are $-\infty$, a convention is necessary to ensure that $J(u,v)$ is well defined. The one we follow is standard in convex analysis: $\infty - \infty = \infty$. This clarifies the meaning of $J_t(u_t, v_t)$ and $J_e(u_e, v_e)$ in all cases in (2.2):

$$\begin{aligned} J_t(u_t, v_t) &= \begin{cases} \text{finite value} & \text{when } u_t \in U_t \text{ and } v_t \in V_t, \\ -\infty & \text{when } u_t \in U_t \text{ and } v_t \notin V_t, \\ \infty & \text{when } u_t \notin U_t, \end{cases} \\ J_e(u_e, v_e) &= \begin{cases} \text{finite value} & \text{when } u_e \in U_e \text{ and } v_e \in V_e, \\ -\infty & \text{when } u_e \in U_e \text{ and } v_e \notin V_e, \\ \infty & \text{when } u_e \notin U_e, \end{cases} \end{aligned} \tag{2.4}$$

where the sets U_t, U_e, V_t, V_e, are the effective domains in (1.2) and (1.5). The convention enters into the formula for $J(u,v)$ in resolving the integral as ∞ whenever the positive part of the integrand (which is always measurable by the argument given in the proof of Proposition 1.1) has integral ∞ while the negative part has integral $-\infty$. (This amounts to writing $J(u,v)$ with the terms $\int_{t_0}^{t_1} \varphi_t(u_t)\, dt$ and $-\int_{t_0}^{t_1} \psi_t^*(v_t)\, dt$ separated out and then invoking the convention $\infty - \infty = \infty$ in forming the overall sum. The first of these terms is unambiguously finite or ∞, as seen in Proposition 1.1, while the second is finite or $-\infty$.)

Proposition 2.1. *The functional J is convex-concave on $\mathcal{U}\times\mathcal{V}$ with finite values on $U\times V$ but infinite values everywhere else. Fpr each $v\in V$, $J(u,v)$ is lower semicontinuous in $u\in\mathcal{U}$, while for each $u\in U$, $J(u,v)$ is upper semicontinuous in $v\in\mathcal{V}$. The objective functionals F and G in $(\mathcal{P})$ and $(\mathcal{Q})$ are given by*

$$F(u)=\inf_{u\in\mathcal{U}} J(u,v)=\inf_{u\in U} J(u,v) \quad \textit{and} \quad G(v)=\sup_{v\in\mathcal{V}} J(u,v)=\sup_{v\in V} J(u,v).$$

Proof. In view of the definitions of U and V in (1.1) and (1.4), the convention adopted in the formula for $J(u,v)$ entails having

$$J(u,v)=\begin{cases}\text{finite value} & \text{when } u\in U \text{ and } v\in V,\\ -\infty & \text{when } u\in U \text{ and } v\notin V,\\ \infty & \text{when } u\notin U.\end{cases} \tag{2.5}$$

The fact that $J(u,v)$ is convex in u and concave in v relative to the product set $U\times V$ is obvious from the convexity of the functions $\varphi_t,\varphi_e,\psi_t^*,\psi_e^*$. The semicontinuity follows from (A1) and (A3) by Fatou's lemma, cf. [**7**].

To establish the formula asserted for $G(v)$, it suffices because of the infinities in (2.5) to prove the first equality in the case of $v\in V$. This is done by taking the first of the forms for $j(u,v)$ in (2.3) and calculating

$$\begin{aligned}\inf_{u\in\mathcal{U}} J(u,v)=&\int_{t_0}^{t_1}[\langle q_t,v_t\rangle-\psi_t(v_t)-\langle y_t,v_t\rangle]\,dt+[\langle q_e,v_e\rangle-\psi_e(v_e)-\langle y_{t_0},v_e\rangle]\\ &+\inf_{u\in\mathcal{U}}\{\int_{t_0}^{t_1}[\langle p_t-B_t^*y_t,u_t\rangle+\varphi_t(u_t)]\,dt+[\langle p_e-B_e^*y_{t_0},u_e\rangle+\varphi_e(u_e)]\}.\end{aligned}$$

The infimum on the right equals $-\int_{t_0}^{t_1}\varphi_t^*(B_t^*y_t-p_t)\,dt-\varphi_e^*(B_e^*y_{t_0}-p_e)$ through the conjugacy formulas

$$\varphi_t^*(r)=\sup_{u\in\mathbb{R}^k}\{\langle r,u\rangle-\varphi_t(u)\} \quad\text{and}\quad \varphi_e^*(r)=\sup_{u_e\in\mathbb{R}^{k_e}}\{\langle r,u_e\rangle-\varphi_e(u_e)\}$$

and the fundamental theorem on conjugates of integral functionals, cf. [**7**, Theorem 3C]. The proof of the formula for $F(u)$ follows the same pattern. (The apparent lack of symmetry in (2.5) is restored though the observation, already made, that only the values of J on $U\times V$ really matter.) $\square$

Theorem 2.2. *Under (A1)–(A3), the optimal values in problems $(\mathcal{P})$ and $(\mathcal{Q})$ always satisfy $\inf(\mathcal{P})\geq\sup(\mathcal{Q})$. A pair $(\bar u,\bar v)$ furnishes a saddle point of J on $\mathcal{U}\times\mathcal{V}$ if and only if $\bar u$ is optimal for $(\mathcal{P})$, $\bar v$ is optimal for $(\mathcal{Q})$, and one actually has $\inf(\mathcal{P})=\sup(\mathcal{Q})$. This saddle point condition is equivalent to the following, where $\bar x$ and $\bar y$ denote the primal and dual trajectories generated by $\bar u$ and $\bar v$:*

$$\begin{aligned}&(\bar u_t,\bar v_t) \textit{ is a saddle point of } J_t(u_t,v_t)-\langle B_t^*\bar y_t,u_t\rangle-\langle C_t\bar x_t,v_t\rangle \textit{ on } U_t\times V_t \textit{ for a.e. } t,\\ &(\bar u_e,\bar v_e) \textit{ is a saddle point of } J_e(u_e,v_e)-\langle B_e^*\bar y_{t_0},u_e\rangle-\langle C_e\bar x_e,v_e\rangle \textit{ on } U_e\times V_e.\end{aligned} \tag{2.6}$$

Proof. Up to the equivalence of the saddle point condition with (2.6), the assertions are well known consequences of the relationship displayed in Proposition 2.1, where primal and dual objectives are derived as “halves” of a minimax problem. The saddle point condition has the means by definition that

$$\bar{u} \in \operatorname*{argmin}_{u \in U} J(u, \bar{v}) \text{ and } \bar{v} \in \operatorname*{argmax}_{v \in V} J(\bar{u}, v).$$

Due to (2.5), it requires $\bar{u} \in U$ and $\bar{v} \in V$. Then in terms of the notation

$$\begin{aligned} \bar{J}_t(u_t, v_t) &:= J_t(u_t, v_t) - \langle B_t^* \bar{y}_t, u_t \rangle - \langle C_t \bar{x}_t, v_t \rangle, \\ \bar{J}_e(u_e, v_t) &:= J_e(u_e, v_e) - \langle B_e^* \bar{y}_{t_0}, u_e \rangle - \langle C_e \bar{x}_{t_1}, v_e \rangle, \end{aligned} \tag{2.7}$$

it reduces by the calculation in the proof of Proposition 2.1 to

$$\begin{aligned} &\bar{u}_t \in \operatorname*{argmin}_{u_t \in \mathbb{R}^k} \bar{J}_t(u_t, \bar{v}_t) \text{ and } \bar{v}_t \in \operatorname*{argmax}_{v_t \in \mathbb{R}^l} \bar{J}_t(\bar{u}_t, v_t) \text{ a.e.,} \\ &\bar{u}_e \in \operatorname*{argmin}_{u_e \in \mathbb{R}^{k_e}} \bar{J}_e(u_e, \bar{v}_e) \text{ and } \bar{v}_e \in \operatorname*{argmax}_{v_e \in \mathbb{R}^{l_e}} \bar{J}_e(\bar{u}_e, v_e). \end{aligned}$$

These relations assert that $(\bar{u}_t, \bar{v}_t)$ is a saddle point of $\bar{J}_t$ on $\mathbb{R}^k \times \mathbb{R}^l$ for a.e. t and $(\bar{u}_e, \bar{v}_e)$ is a saddle point of $\bar{J}_e$ on $\mathbb{R}^{k_e} \times \mathbb{R}^{l_e}$. But $\bar{J}_t$ and $\bar{J}_e$ have the structure (2.4) relative to $U_t \times V_t$ and $U_e \times V_e$. The saddle points in question are therefore expressed equivalently with respect to $U_t \times V_t$ and $U_e \times V_e$. This is all that had to be proved. □

The saddle point conditions in (2.6) will be referred to as the *minimaximum principle* for $(\mathcal{P})$ and $(\mathcal{Q})$. This principle is *always sufficient* for optimality according to the Theorem 2.2, and it is necessary for optimality in any circumstances where we happen to know that $\inf(\mathcal{P}) = \sup(\mathcal{Q})$ and that both problems have solutions. We shall prove in due course that assumption (A4) provides such a circumstance. Our method requires us to examine an auxiliary pair of problems in which trajectories are optimized without direct mention of controls. This will be done in the next section.

The minimaximum principle can be stated in terms of a duality between finite-dimensional optimization problems at every instant of time. With x and y as parameter vectors in $\mathbb{R}^n$, consider the problems

$$(\mathcal{P}_t(x,y)) \qquad \min_{u \in U_t} \{ \langle p_t - B_t^* y, u \rangle + \varphi_t(u) + \psi_t(q_t - C_t x - D_t u) \},$$

$$(\mathcal{Q}_t(x,y)) \qquad \max_{v \in V_t} \{ \langle q_t - C_t x, v \rangle - \psi_t^*(u) - \varphi_t(B_t^* y + D_t^* v - p_t) \},$$

for each $t \in [t_0, t_1]$ and also the problems

$$(\mathcal{P}_e(x,y)) \qquad \min_{u_e \in U_e} \{ \langle p_e - B_e^* y, u_e \rangle + \varphi_e(u_e) + \psi_e(q_e - C_e x - D_e u_e) \},$$

$$(\mathcal{Q}_e(x,y)) \qquad \max_{v_e \in V_e} \{\langle q_e - C_e x, v_e\rangle - \psi_e^*(u_e) - \varphi_e(B_e^* y + D_e^* v_e - p_e)\}.$$

Proposition 2.3. *The minimaximum principle (2.6) is equivalent to the following set of conditions on the controls $\bar u$ and $\bar v$, as expressed through the corresponding trajectories $\bar x$ and $\bar y$:*

(a) $\bar u_t$ solves $(\mathcal{P}_t(\bar x_t, \bar y_t))$, $\bar v_t$ solves $(\mathcal{Q}_t(\bar x_t, \bar y_t))$, and $(\mathcal{P}_t(\bar x_t, \bar y_t)) = (\mathcal{Q}_t(\bar x_t, \bar y_t))$.

(b) $\bar u_e$ solves $(\mathcal{P}_e(\bar x_{t_1}, \bar y_{t_0}))$, $\bar v_e$ solves $(\mathcal{Q}_e(\bar x_{t_1}, \bar y_{t_0}))$, and $(\mathcal{P}_e(\bar x_{t_1}, \bar y_{t_0})) = (\mathcal{Q}_e(\bar x_{t_1}, \bar y_{t_0}))$.

Proof. Elementary minimax theory informs us that $(\bar u_t, \bar v_t)$ has the saddle point property in (2.6) for a given t if and only if $\bar u_t$ minimizes over $u \in U_t$ the function

$$f(u)_t := \sup_{v \in V_t} \{J_t(u,v) - \langle B_t^* \bar y_t, u\rangle - \langle C_t \bar x_t, v\rangle\},$$

$\bar v_t$ maximizes over $v \in V_t$ the function

$$g(v)_t := \inf_{u \in U_t} \{J_t(u,v) - \langle B_t^* \bar y_t, u\rangle - \langle C_t \bar x_t, v\rangle\},$$

and $\inf_{U_t} f_t = \sup_{V_t} g_t$. These functions are calculated from the reciprocal conjugacy formulas

$$\psi_t(s) = \sup_{v \in V_t} \{\langle s, v\rangle - \psi_t^*(v)\} \quad \text{and} \quad \psi_e(s_e) = \sup_{v_e \in V_e} \{\langle s_e, v_e\rangle - \psi_e^*(v_e)\}$$

to be the objectives in $(\mathcal{P}_t(\bar x_t, \bar y_t))$ and $(\mathcal{Q}_t(\bar x_t, \bar y_t))$, respectively. The assertion concerning this pair of problems is therefore valid. The one for $(\mathcal{P}_e(\bar x_{t_1}, \bar y_{t_0}))$ and $(\mathcal{Q}_e(\bar x_{t_1}, \bar y_{t_0}))$ is proved similarly. □

3. Bolza formulations.

Generalized problems of Bolza in the calculus of variations concern trajectories as elements of the space $\mathcal{A}^1 = \mathcal{A}_n^1[t_0, t_1]$ consisting of all the absolutely continuous arcs x in $\mathbb{R}^n$ over $[t_0, t_1]$. (The superscript 1 refers to the fact that the function $t \mapsto \dot x_t$ is an element of $\mathcal{L}_n^1[t_0, t_1]$.) Such problems have the form

$$(\mathcal{P}_B) \qquad \text{minimize } \Phi(x) := \int_{t_0}^{t_1} L_t(x_t, \dot x_t)\, dt + L_e(x_{t_0}, x_{t_1}) \text{ over all } x \in \mathcal{A}^1,$$

where the functions L_t and L_e on $\mathbb{R}^n \times \mathbb{R}^n$ may be extended-real-valued. In the *convex* case, where L_t and L_e are convex functions on $\mathbb{R}^n \times \mathbb{R}^n$, there is a dual problem

$$(\mathcal{Q}_B) \qquad \text{maximize } \Psi(y) := -\int_{t_0}^{t_1} M_t(y_t, \dot y_t)\, dt - M_e(y_{t_0}, y_{t_1}) \text{ over all } y \in \mathcal{A}^1,$$

in which M_t and M_e are derived from L_t and L_e by

$$M_t(y_t, \dot y_t) = L_t^*(\dot y_t, y_t) \quad \text{and} \quad M_e(y_{t_0}, y_{t_1}) = L_e^*(y_{t_0}, -y_{t_1}). \tag{3.1}$$

An extensive duality theory for convex problems of Bolza was developed in [**3**], [**4**]. We intend to apply this theory to gain insights into the relationship between the control problems $(\mathcal{P})$ and $(\mathcal{Q})$. For this purpose we choose to define

$$\begin{aligned} L_t(x,w) &= \inf_{\substack{u\in U_t\\ A_t x+B_t u+b_t=w}} \{\langle p_t,u\rangle + \varphi_t(u) + \psi_t(q_t - C_t x - D_t u) - \langle c_t,x\rangle\},\\ L_e(x_0,x_1) &= \inf_{\substack{u_e\in U_e\\ B_e u_e+b_e=x_0}} \{\langle p_e,u_e\rangle + \varphi_e(u_e) + \psi_e(q_e - C_e x_1 - D_e u_e) - \langle c_e,x_1\rangle\}. \end{aligned} \tag{3.2}$$

(Here x and u are temporarily just dummy vectors in $\mathbb{R}^n$ and $\mathbb{R}^k$, and similarly x_0 and x_1 in $\mathbb{R}^n$.) Our work with these expressions will make use of the concept of the *recession function* associated with a lower semicontinuous, proper, convex function f on $\mathbb{R}^n$, denoted by $\operatorname{rc} f$. Many facts about such recession functions are assembled in [**8**, §8 and §13]. We mention in particular that

$$(\operatorname{rc} f)(z) = \lim_{\lambda\to\infty} [f(\bar z + \lambda z) - f(\bar z)]/\lambda \ \text{ for any } \bar z \in \operatorname{dom} f, \tag{3.3}$$

and that coercivity of f is equivalent to $\operatorname{rc} f$ being the indicator function δ_0 of the origin, where

$$\delta_0(z) = \infty \text{ for } z \neq 0, \quad \delta_0(0) = 0.$$

Proposition 3.1. *Under assumptions (A1)–(A4), the Bolza functional Φ in $(\mathcal{P}_B)$ is well defined on $\mathcal{A}^1$ and is convex. The functions L_t and L_e are themselves lower semicontinuous, proper and convex on $\mathbb{R}^n \times \mathbb{R}^n$, and L_t depends epi-continuously on t. The infima defining L_t and L_e are attained whenever finite, i.e. whenever the given constraints in (3.2) can be satisfied. The recession functions are expressed by*

$$\begin{aligned} (\operatorname{rc} L_t)(x,w) &= (\operatorname{rc}\psi_t)(-C_t x) - \langle c_t,x\rangle + \delta_0(w - A_t x),\\ (\operatorname{rc} L_e)(x_0,x_1) &= (\operatorname{rc}\psi_e)(-C_t x_1) - \langle c_e,x_1\rangle + \delta_0(x_0). \end{aligned} \tag{3.4}$$

Proof. Consider the functions

$$\begin{aligned} K_t(x,w,u) &= \langle p_t,u\rangle + \varphi_t(u) + \psi_t(q_t - C_t x - D_t u)\\ &\qquad - \langle c_t,x\rangle + \delta_0(w - A_t x - B_t u - b_t)\\ K_e(x_0,x_1,u_e) &= \langle p_e,u_e\rangle + \varphi_e(u_e) + \psi_e(q_e - C_e x_1 - D_e u_e)\\ &\qquad - \langle c_e,x_1\rangle + \delta_0(x_0 - B_e u_e + b_e). \end{aligned} \tag{3.5}$$

By virtue of (A1) these are lower semicontinuous, proper, convex functions on $\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^k$ and $\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^{k_e}$. The definitions given for L_t and L_e in (3.2) are equivalent to

$$L_t(x,w) = \inf_{u\in\mathbb{R}^k} K_t(x,w,u) \ \text{ and } \ L_e(x_0,x_1) = \inf_{u_e\in\mathbb{R}^{k_e}} K_e(x_0,x_1,u_e). \tag{3.6}$$

In the language of convex analysis, therefore, L_t is the image of K_t under the projection $(x,w,u)\mapsto(x,w)$, while L_e is the image of K_e under $(x_0,x_1,u_e)\mapsto(x_0,x_1)$. We wish to apply a general theorem about such images, namely [**8**, Theorem 9.2]. This involves a condition on the recession functions of K_t and K_e, which are calculated via (3.3) and the coercivity of φ_t and φ_e in (A4) to be

$$\begin{aligned}(\mathrm{rc}\,K_t)(x,w,u) &= \langle p_t,u\rangle + \delta_0(u) + (\mathrm{rc}\,\psi_t)(-C_tx - D_tu) - \langle c_t,x\rangle + \delta_0(w - A_tx - B_tu)\\ &= \delta_0(u) + (\mathrm{rc}\,\psi_t)(-C_tx) - \langle c_t,x\rangle + \delta_0(w - A_tx),\\ (\mathrm{rc}\,K_e)(x_0,x_1,u_e) &= \langle p_e,u_e\rangle + \delta_0(u_e) + (\mathrm{rc}\,\psi_e)(-C_ex_1 - D_eu_e) - \langle c_e,x_1\rangle + \delta_0(x_0 - B_eu_e)\\ &= \delta_0(u_e) + (\mathrm{rc}\,\psi_e)(-C_ex_1) - \langle c_e,x_1\rangle + \delta_0(x_0).\end{aligned} \tag{3.7}$$

(We make use of the coercivity of φ_t and φ_e in replacing $\mathrm{rc}\,\varphi_t$ and $\mathrm{rc}\,\varphi_e$ by δ_0.) The fact that $(\mathrm{rc}\,K_t)(0,0,u) = 0$ only for $u = 0$, and $(\mathrm{rc}\,K_e)(0,0,u_e) = 0$ only for $u_e = 0$ guarantees by the theorem just cited from [**8**] that L_t and L_e are lower semicontinuous, proper, convex functions for which the infima in (3.6) are always attained (i.e. the ones in (3.2) are attained when the constraints can be satisfied), and that

$$(\mathrm{rc}\,L_t)(x,w) = \inf_{u\in\mathbb{R}^k}(\mathrm{rc}\,K_t)(x,w,u) \quad\text{and}\quad (\mathrm{rc}\,L_e)(x_0,x_1) = \inf_{u_e\in\mathbb{R}^{k_e}}(\mathrm{rc}\,K_e)(x_0,x_1,u_e).$$

The latter formulas are the same as the ones claimed in (3.4) because of the special nature of $\mathrm{rc}\,K_t$ and $\mathrm{re}\,K_e$ in (3.7).

We must verify that L_t depends epi-continuously on t. We shall do this by way of theorems of McLinden and Bergstrom [**11**], showing first that K_t depends epi-continuously on t. Let us write $K_t = K_t^1 + K_t^2 + K_t^3$ with

$$\begin{aligned}K_t^1(x,w,u) &= \langle p_t,u\rangle + \psi_t(q_t - C_tx - D_tu) - \langle c_t,x\rangle,\\ K_t^2(x,w,u) &= \varphi_t(u),\\ K_t^3(x,w,u) &= \delta_0(w - A_tx - B_tu).\end{aligned}$$

The functions in this decomposition are lower semicontinuous, proper and convex on $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^k$. We argue first that each depends epi-continuously on t. This is obvious for K_t^2 because of (A2). It holds for K_t^2 because this is a finite convex function by (A4) whose values depend continuously on t; cf. Proposition 1.3. (A finite convex function depends epi-continuously on t if and only if its value at each point depends continuously on t [**5**, Corollaries 4 and 5].) In the case of K_t^3 the epi-continuity follows from [**11**, Theorem 8] because the linear transformation $(x,w,u) \mapsto w - A_tx - B_tu$ depends continuously on t (by (A3)) and has all of $\mathbb{R}^n$ as its range. We deduce next from [**11**, Theorem 5] that $K_t^2 + K_t^3$ depends epi-continuously on t, because the set $\mathrm{dom}\,K_t^2 - \mathrm{dom}\,K_t^3$ is all of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^k$ and therefore certainly contains the origin in its interior. The same theorem of [**11**] applied to $K_t^1 + (K_t^2 + K_t^3)$ then yields the desired epi-continuity of K_t with respect to t, since $\mathrm{dom}\,K_t^1 - \mathrm{dom}(K_t^2 + K_t^3)$ too is all of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^k$. Recalling now that L_t is the image of K_t under the projection $(x,w,u) \mapsto (x,w)$, and $(\mathrm{rc}\,K_t)(0,0,u) = 0$ only for $u = 0$, we obtain from [**11**, Theorem 7] that L_t depends continuously on t. This property of L_t implies in particular that $L_t(x,w)$ is lower semicontinuous with respect to (t,x,w). The integrand in the formula for the Bolza functional Φ is certainly measurable then. The functional is well defined in this case under the $\infty - \infty = \infty$ convention explained earlier. Its convexity follows from that of L_t and L_e. $\square$

Corollary 3.2. *Under assumptions (A1)–(A4), the dual Bolza functional* Ψ *in* $(\mathcal{Q}_B)$ *is well defined on* $\mathcal{A}_n^1[t_0, t_1]$ *and is concave. The functions* M_t *and* M_e *are lower semicontinuous, proper and convex on* $\mathbb{R}^n \times \mathbb{R}^n$*, and* M_t *depends epi-continuously on* t*. These functions satisfy the reciprocal conjugacy formulas*

$$L_t(x_t, \dot{x}_t) = M_t^*(\dot{x}_t, x_t) \ \text{ and } \ L_e(x_{t_0}, x_{t_1}) = M_e^*(x_{t_0}, -x_{t_1}). \tag{3.8}$$

Proof. This merely invokes the basic properties of the Legendre-Fenchel transform [**8**, §12], including the fact that it preserves epi-convergence of convex functions [**9**]. $\square$

It will be demonstrated now that the dual Bolza functional Ψ bears the same relationship to the elements of the dual control problem $(\mathcal{Q})$ as the primal Bolza functional Φ does through (3.2) to the elements of $(\mathcal{P})$.

Proposition 3.3. *Under (A1)–(A4), the dual functions* M_t *and* M_e *are expressed by*

$$\begin{aligned} -M_t(y,z) &= \sup_{\substack{v \in V_t \\ A_t^* y + C_t^* v + c_t = -z}} \{\langle q_t, v\rangle - \psi_t^*(v) - \varphi_t^*(B_t^* y + D_t^* v - p_t) - \langle b_t, y\rangle\}, \\ -M_e(y_0,y_1) &= \sup_{\substack{v_e \in V_e \\ C_e^* v_e + c_e = y_1}} \{\langle q_e, v_e\rangle - \psi_e^*(v_e) - \varphi_e^*(B_e^* y_0 + D_e^* u_e - p_e) - \langle b_e, y_0\rangle\}, \end{aligned} \tag{3.9}$$

where the suprema are attained whenever the indicated constraints can be satisfied. The recession functions of M_t *and* M_e *are given by*

$$\begin{aligned} (\mathrm{rc}\, M_t)(y,z) &= (\mathrm{rc}\, \varphi_t^*)(-B_t^* y) - \langle b_t, y\rangle + \delta_0(z + A_t^* y), \\ (\mathrm{rc}\, M_e)(y_0,y_1) &= (\mathrm{rc}\, \varphi_e^*)(-B_e^* y_0) - \langle b_e, y_0\rangle + \delta_0(y_1). \end{aligned} \tag{3.10}$$

Proof. Starting toward the proof of the formula for M_t in (3.9), we observe that the definition of M_t in (3.1), which means

$$M_t(y,z) = \sup_{x,w}\{\langle z, x\rangle + \langle y, w\rangle = L_t(x,w)\},$$

can be combined with the specification of L_t in (3.2) to yield

$$\begin{aligned} M_t(y,z) &= \sup_{x,u}\{\langle z, x\rangle + \langle y, A_t x + B_t u + b_t\rangle - \langle p_t, u\rangle - \varphi_t(u) - \psi_t(q_t - C_t x - D_t u) + \langle c_t, x\rangle\} \\ &= \langle b_t, y\rangle - \inf_{x,u}\{f(x,u) - g(E(x,u))\}, \end{aligned} \tag{3.11}$$

where E is the linear transformation given by $E(x,u) = C_t x + D_t u$ and f and g are the convex and concave functions given by

$$\begin{aligned} f(x,u) &= \varphi_t(u) - \langle z + A_t^* y + c_t, x\rangle - \langle B_t^* y + p_t, u\rangle, \\ g(s) &= -\psi_t(q_t - s). \end{aligned}$$

Inasmuch as g is finite everywhere by (A4), we can apply Fenchel's duality theorem as stated in [**8**, Corollary 31.2.1] to write

$$\inf_{x,u}\{f(x,u) - g(E(x,u))\} = \max_v\{g^*(v) - f^*(E^*(v))\},$$

where the "max" indicates attainment. The adjoint linear transformation E^* takes v into the pair $(C_t^* v, D_t^* v)$. Direct calculation of the conjugate functions f^* and g^* yields

$$\begin{aligned} f^*(s,r) &= \delta_0(s + z + A_t^* y + c_t) + \varphi_t^*(r + B_t^* y - p_t), \\ g^*(v) &= \langle q_t, v\rangle - \psi_t^*(v). \end{aligned}$$

Therefore

$$-M_t(y,z) = -\langle b_t, y\rangle + \max_v\{\langle q_t, v\rangle - \psi_t^*(v) - \delta_0(C_t^* v + z + A_t^* y + c_t) - \varphi_t^*(D_t^* v + B_t^* y - p_t)\}.$$

This is equivalent to the formula asserted in (3.9). The argument for M_e is runs parallel. We have from (3.1) that

$$M_e(y_0, y_1) - \sup_{x_0,x_1}\{\langle y_0, x_0\rangle + \langle y_1, x_1\rangle - L_e(x_0,x_1)\},$$

and in combining this with (3.2) we get

$$\begin{aligned} M_e(y_0,y_1) = \sup_{x_0,x_1}\{&\langle y_0, B_e u_e + b_e\rangle - \langle y_1, x_1\rangle - \langle p_e, u_e\rangle - \varphi_e(u_e) \\ &- \psi_e(q_e - C_e x_1 - D_e u_e) - \langle c_e, x_1\rangle\} \\ = \langle b_e, y_0\rangle - &\inf_{x_1,u_e}\{f_e(x_1,u_e) - g_e(E_e(x_1,u_e))\}, \end{aligned}$$

where $E_e(x_1,u_e) = C_e x_1 + D_e u_e$ and

$$\begin{aligned} f_e(x_1,u_e) &= \varphi_e(u_e) + \langle y_1, x_1\rangle - \langle B_e^* y_0 - p_e, u_e\rangle \\ g_e(s_e) &= -\psi_e(q_e - s_e). \end{aligned}$$

Fenchel's duality theorem brings us to

$$-M_e(y_0,y_1) = -\langle b_e, y_0\rangle + \max_{v_e}\{g_e^*(v_e) - f_(^* E_e^*(v_e))\},$$

where $E_e^*(v_e) = (C_e^* v_e, D_e^* v_e)$ and

$$f_e^*(s, r_e) = \delta_0(y_1) + \varphi_e^*(r_e + B_e^* y_0 - p_e),$$
$$g_e^*(v_e) = \langle q_e, v_e\rangle - \psi_e^*(v_e),$$

and this representation is equivalent to the one claimed for M_e in (3.9). Because of the symmetry between the formulas in (3.9) and (3.2), we can obtain the recession function expressions in (3.10) by appealing to Proposition 3.1 in dual form. □

These results prepare us for demonstrating that the Bolza problems $(\mathcal{P}_B)$ and $(\mathcal{Q}_B)$ are reduced representations of control problems quite close to, but somewhat broader than, $(\mathcal{P})$ and $(\mathcal{Q})$. The extended control problems, which we denote by $(\mathcal{P}')$ and $(\mathcal{Q}')$, are obtained simply by replacing $\mathcal{U}$ and $\mathcal{V}$ by the slightly larger control spaces

$$\mathcal{U}' := \{u \,|\, u_e \in \mathbb{R}^{k_e},\ u_t \in \mathbb{R}^k \text{ measurable in } t \text{ with } t \mapsto B_t u_t \text{ summable }\},$$
$$\mathcal{V}' := \{v \,|\, v_e \in \mathbb{R}^{l_e},\ v_t \in \mathbb{R}^l \text{ measurable in } t \text{ with } t \mapsto C_t^* v_t \text{ summable }\}.$$

Thus the extended primal problem is

$$(\mathcal{P}') \quad \begin{aligned} &\text{minimize the functional } F(u) = \\ &\int_{t_0}^{t_1} [\langle p_t, u_t\rangle + \varphi_t(u_t) + \psi_t(q_t - C_t x_t - D_t u_t) - \langle c_t, x_t\rangle]\, dt \\ &\quad + [\langle p_e, u_e\rangle + \varphi_e(u_e) + \psi_e(q_e - C_e x_{t_1} - D_e u_e) - \langle c_e, x_{t_1}\rangle] \\ &\text{over } u \in \mathcal{U}',\ \text{where } x \text{ is determined from } u \text{ by} \\ &\quad \dot{x}_t = A_t x_t + B_t u_t + b_t \ \text{ a.e.}, \quad x_{t_0} = B_e u_e + b_e, \end{aligned}$$

while the extended dual problem is

$$(\mathcal{Q}') \quad \begin{aligned} &\text{maximize the functional } G(v) = \\ &\int_{t_0}^{t_1} [\langle q_t, v_t\rangle - \psi_t^*(v_t) - \varphi_t^*(B_t^* y_t + D_t^* v_t - p_t) - \langle b_t, y_t\rangle]\, dt \\ &\quad + [\langle q_e, v_e\rangle - \psi_e^*(v_e) - \varphi_e^*(B_e^* y_{t_0} + D_e^* v_e - p_e) - \langle b_e, y_{t_0}\rangle] \\ &\text{over } v \in \mathcal{V}',\ \text{where } y \text{ is determined from } v \text{ by} \\ &\quad -\dot{y}_t = A_t^* y_t + C_t^* v_t + c_t \ \text{ a.e.}, \quad y_{t_1} = C_e^* v_e + c_e. \end{aligned}$$

Note that each $u \in \mathcal{U}'$ does determine a unique trajectory $x \in \mathcal{A}^1$ in $(\mathcal{P}')$, and similarly each $v \in \mathcal{V}'$ determines a unique $y \in \mathcal{A}^1$ in $(\mathcal{Q}')$. We shall say in this situation that x and y are *realized* by the controls u and v. For the moment we think of the functionals F and G in the extended sense of $(\mathcal{P}')$ and $(\mathcal{Q}')$ as being defined with the appropriate conventions regarding infinite values, but it will emerge from further analysis that actually $F(u) > -\infty$ and $G(v) < \infty$.

Proposition 3.4. *Assume (A1)–(A4). Then the primal problems* $(\mathcal{P}_B)$ *and* $(\mathcal{P}')$ *are equivalent to each other in the sense that*

$$\Phi(x) = \inf\{F(u) \,\big|\, u \in \mathcal{U}', x \text{ realized by } u\}, \quad \text{with attainment when } \Phi(x) < \infty.$$

Likewise, the dual problems $(\mathcal{Q}_B)$ *and* $(\mathcal{Q}')$ *are equivalent to each other in the sense that*

$$\Psi(y) = \sup\{G(v) \,\big|\, v \in \mathcal{V}', y \text{ realized by } v\}, \quad \text{with attainment when } \Psi(y) > -\infty.$$

Proof. In terms of the functions K_t and K_e in (3.5) define

$$\Upsilon(x,u) = \int_{t_0}^{t_1} K_t(x_t, \dot{x}_t, u_t)\, dt + K_e(x_{t_0}, x_{t_1}).$$

The representations (3.6) lead to

$$\Phi(x) = \min\{\Upsilon(x,u) \,\big|\, u_e \in \mathbb{R}^{k_e}, u_t \in \mathbb{R}^k \text{ measurable in } t\}. \tag{3.12}$$

This is justified by the fundamental result in [**12**, p.316] on control formulations versus Bolza formulations. (The inf-boundedness condition in the hypothesis of that result is fulfilled because of the recession function property of K_t established in (3.7).) Formula (3.12) is equivalent to the assertion made in the present theorem about the primal problems. Symmetry yields the corresponding fact about the dual problems. □

4. Hamiltonian functions and duality.

Further progress in applying the theory of Bolza problems to the original control problems $(\mathcal{P})$ and $(\mathcal{Q})$ will depend on study of the *Hamiltonian* function for problems $(\mathcal{P}_B)$ and $(\mathcal{Q}_B)$, which in general is defined on $\mathbb{R}^n \times \mathbb{R}^n$ by

$$H_t(x,y) = \sup_{w \in \mathbb{R}^n} \{\langle y, w\rangle - L_t(x,w)\}. \tag{4.1}$$

Proposition 4.1. *Under assumptions (A1)–(A4), the Hamiltonian* $H_t(x,y)$ *is finite everywhere, concave in* $x \in \mathbb{R}^n$*, convex in* $y \in \mathbb{R}^n$*, and continuous in* (t,x,y)*.*

Proof. The fact that $H_t(x,y)$ is concave in x and convex in y follows simply from the convexity of $L_t(x,w)$ in (x,w), as in the theory of convex problems of Bolza more generally. The defining equation (4.1) says that $H_t(x,\cdot)$ is the function conjugate to $L_t(x,\cdot)$. For each choice of t and x, $L_t(x,\cdot)$ is not only lower semicontinuous and convex but proper on $\mathbb{R}^n$. This is evident from (3.2) and the finiteness of ψ_t assumed in (A4). Moreover the recession function of $L_t(x,\cdot)$ is $(\operatorname{rc} L_t)(0,\cdot)$ on the general basis of (3.3), and the formula in Proposition 3.1 shows $(\operatorname{rc} L_t)(0,\cdot)$ to be δ_0. Thus $L_t(x,\cdot)$ is coercive, so that its conjugate must be finite everywhere. In other words, $H_t(x,y)$ must be finite for all (t,x,y).

We claim next that for fixed x, $L_t(x,\cdot)$ depends epi-continuously on t. This is equivalent to the assertion that the function $(z,w) \mapsto (L_t + f)(z,w)$ depends epi-continuously on t when for fixed x we define $f(z,w)=\delta_0(z-x)$. Such epi-continuity is justified by [**11**, Theorem 5], because $\operatorname{dom} L_t - \operatorname{dom} f$ is all of $\mathbb{R}^n \times \mathbb{R}^n$. The Legendre-Fenchel transform preserves epi-convergence [**9**], so in passing to the conjugate function $H_t(x,\cdot)$ of $L_t(x,\cdot)$ we have $H_t(x,\cdot)$ depending epi-continuously on t. Because $H_t(x,\cdot)$ is finite everywhere, its epi-continuity with respect to t is the same as the continuity of $H_t(x,y)$ in t for fixed (x,y) [**5**, Corollaries 4 and 5]. This implies the continuity of $H_t(x,y)$ in (t,x,y) by [**8**, Theorem 35.4], due to the concavity-convexity. □

Theorem 4.2. *Assumptions (A1)–(A4) guarantee that the Bolza problems $(\mathcal{P}_B)$ and $(\mathcal{Q}_B)$ both have solutions, and the same for the extended control problems $(\mathcal{P}')$ and $(\mathcal{Q}')$. Moreover,*

$$-\infty < \inf(\mathcal{P}') = \inf(\mathcal{P}_B) = \sup(\mathcal{Q}_B) = \sup(\mathcal{Q}') > \infty.$$

Proof. Only the part concerning the Bolza problems needs to be dealt with, because the rest will then follow immediately from Theorem 3.4. We shall apply the main results of the duality theory for Bolza problems in [**4**, Theorem 1, Theorem 3 and its Corollary 1]. The background for this application is the finiteness of the Hamiltonian as proved in Proposition 4.1, which guarantees by [**4**, Corollary on p.17] that certain basic integrability conditions, called (C_0) and (D_0) in that paper, are fulfilled. The duality results say then that we have $\infty > \inf(\mathcal{P}_B) = \sup(\mathcal{Q}_B) > \infty$ with solutions existing for both problems, provided that the following two criteria are met in terms of arcs x and y in $\mathcal{A}^1$ (this is a slightly specialized case of the results in question):

$$\int_{t_0}^{t_1} (\operatorname{rc} L_t)(x_t,\dot{x}_t)\,dt + (\operatorname{rc} L_e)(x_{t_0},x_{t_1}) \le 0 \text{ only for } x=0,$$

$$\int_{t_0}^{t_1} (\operatorname{rc} M_t)(y_t,\dot{y}_t)\,dt + (\operatorname{rc} M_e)(y_{t_0},y_{t_1}) \le 0 \text{ only for } y=0.$$

The recession function formulas provided in Propositions 3.1 and 3.3 indicate that this is indeed true in the present circumstances, because a linear ordinary differential equation has no solution starting from 0 except the 0 solution. □

The conversion of this duality and existence theorem into one for the original control problems $(\mathcal{P})$ and $(\mathcal{Q})$ will rely on the theory of optimality conditions for convex problems of Bolza as developed in [**3**], [**4**], [**13**]. In addition to a generalized Hamiltonian differential equation involving subgradients of H_t, there is a transversality condition on endpoints that usually is expressed through subgradients of L_e or M_e but will now be posed in new form. This form involves subgradients of what we shall call the *endpoint Hamiltonian*:

$$H_e(x_1,y_0) := \sup_{x_0\in\mathbb{R}^n} \{\langle x_0,y_0\rangle - L_e(x_0,x_1)\}. \tag{4.2}$$

Proposition 4.3. *Under (A1)–(A4), the endpoint Hamiltonian H_e is a finite concave-convex function on $\mathbb{R}^n \times \mathbb{R}^n$.*

Proof. Definition (4.2) expresses $H_e(x_1,\cdot)$ as the function conjugate to $L_e(\cdot,x_1)$. The latter function, as seen from its definition in (3.2), is not identically ∞ for any choice of

x_1 and is therefore by Proposition 3.1 a lower semicontinuous, proper, convex function on $\mathbb{R}^n$. Moreover its recession function is $(\mathrm{rc}\, L_e)(\cdot, 0)$, and this is δ_0 by formula (3.4) in Proposition 3.1. Hence $L_e(\cdot, x_1)$ is coercive. Consequently its conjugate $H_e(x_1, \cdot)$ is finite everywhere [**8**, §13]. This means that $H_e(x_1, y_0)$ is finite for every choice of (x_1, y_0) in $\mathbb{R}^n \times \mathbb{R}^n$. The concavity of $H_e(x_1, y_0)$ in y_0 is a general consequence of the joint convexity of $L_e(x_0, x_1)$ in (x_0, x_1) in (4.2), cf. [**8**, Theorem 33.1]. □

Optimality conditions for the Bolza problems will be presented now in terms of subgradients of the concave-convex functions H_t and H_e. The theory of such subgradients may be found in [**8**, §§35-37].

Theorem 4.4. *For arcs $\bar{x}$ and $\bar{y}$ to be optimal for $(\mathcal{P}_B)$ and $(\mathcal{Q}_B)$ under (A1)–(A3), the following pair of conditions is always sufficient, and when (A4) holds also necessary:*

$$
\begin{aligned}
&(-\dot{\bar{y}}_t, \dot{\bar{x}}_t) \in \partial H_t(\bar{x}_t, \bar{y}_t) \quad \textit{for a.e. } t \in [t_0, t_1], \\
&(\bar{y}_{t_1}, \bar{x}_{t_0}) \in \partial H_e(\bar{x}_{t_1}, \bar{y}_{t_0}).
\end{aligned}
\tag{4.3}
$$

Proof. If the second condition in (4.3) were replaced by the usual transversality condition $(\bar{y}_{t_0}, -\bar{y}_{t_1}) \in \partial L(\bar{x}_{t_0}, x_{t_1})$, the general result would become a special case of [**3**, Theorems 5 and 6], because of the equality of optimal values in Theorem 4.2. It remains only to observe that the stated conditions in terms of H_e and L_e are equivalent to each other by a general fact of subgradient theory in the case of the relationship between H_e and L_e in (4.2), namely [**8**, Theorem 37.5]. □

Corollary 4.5. *Suppose (A1)–(A4) hold. Then for an arc $\bar{x}$ to be optimal in the Bolza problem $(\mathcal{P}_B)$ it is necessary and sufficient that there exist an arc bary such that the Hamiltonian conditions in (4.3) are satisfied. Any such arc y then solves the dual problem $(\mathcal{Q}_B)$.*

Proof. This combines Theorem 4.4 with the existence assertions in Theorem 4.2. □

Generalized Hamiltonian differential equations formulated for convex problems of Bolza as in (4.3) have been studied for their own sake in [**13**] and also, incidentally, play a central role for Bolza problems in the nonconvex case, cf. Clarke [**14**]. We need next to determine the specific form they take relative to the given data structure.

Proposition 4.6. *Under (A1)–(A4) one has*

$$
\begin{aligned}
H_t(x, y) &= \langle y, A_t x\rangle + \langle b_t, y\rangle + \langle c_t, x\rangle + J_t^*(B_t^* y, C_t x), \\
H_e(x_1, y_0) &= \langle b_e, y_0\rangle + \langle c_e, x_1\rangle + J_e^*(B_e^* y_0, C_e x_1),
\end{aligned}
\tag{4.4}
$$

where J_t^* and J_e^* are the concave-convex functions on $\mathbb{R}^n \times \mathbb{R}^n$ conjugate to J_t and J_e in

(2.2) and given by

$$
\begin{aligned}
J_t^*(r,s) &= \sup_{u\in U_t}\inf_{v\in V_t}\langle r,u\rangle + \langle s,v\rangle - J_t(u,v)\} \\
&= \inf_{v\in V_t}\sup_{u\in U_t}\langle r,u\rangle + \langle s,v\rangle - J_t(u,v)\}, \\
J_e^*(r_e,s_e) &= \sup_{u_e\in U_e}\inf_{v_e\in V_e}\langle r_e,u_e\rangle + \langle s_e,v_e\rangle - J_e(u_e,v_e)\} \\
&= \inf_{v_e\in V_e}\sup_{u_e\in U_e}\langle r_e,u_e\rangle + \langle s_e,v_e\rangle - J_e(u_e,v_e)\}.
\end{aligned} \tag{4.5}
$$

These functions are finite everywhere, and $J_t^*(r,s)$ depends continuously on (r,s).

Proof. The conjugacy formulas

$$
\begin{aligned}
\psi_t(q_t - C_t x - D_t u) &= \sup_{v\in\mathbb{R}^l}\{\langle q_t - C_t x - D_t u, v\rangle - \psi_t^*(v)\} \\
&= \sup_{v\in V_t}\{\langle q_t - C_t x - D_t u, v\rangle - \psi_t^*(v)\}, \\
\psi_e(q_e - C_e x_1 - D_e u_e) &= \sup_{v_e\in\mathbb{R}^{l_e}}\{\langle q_e - C_e x_1 - D_e u_e, v_e\rangle - \psi_e^*(v_e)\} \\
&= \sup_{v_e\in V_e}\{\langle q_e - C_e x_1 - D_e u_e, v_e\rangle - \psi_e^*(v_e)\},
\end{aligned}
$$

allow us to rewrite the defining formulas (3.2) for L_t and L_e in the notation (2.2) as

$$
\begin{aligned}
L_t(x,w) &= \inf_{\substack{u\in U_t \\ A_t x + B_t u + b_t = w}}\ \sup_{v\in V_t}\{J_t(u,v) - \langle c_t,x\rangle - \langle C_t x, v\rangle\}, \\
L_e(x_0,x_1) &= \inf_{\substack{u\in U_e \\ B_e u + b_e = x_0}}\ \sup_{v_e\in V_e}\{J_e(u_e,v_e) - \langle c_e,x_1\rangle - \langle C_e x_1, v_e\rangle\}.
\end{aligned}
$$

These expressions can be substituted into the definitions (4.1) of H_t and (4.2) of H_e to obtain

$$
\begin{aligned}
H_t(x,y) &= \sup_{u\in U_t}\{\langle y, A_t x + B_t u + b_t\rangle - \sup_{v\in V_t}\{J_t(u,v) - \langle c_t,x\rangle - \langle C_t x, v\rangle\}\} \\
&= \langle y, A_t x\rangle + \langle b_t, y\rangle + \langle c_t, x\rangle + \sup_{u\in U_e}\inf_{v\in V_e}\{\langle C_t x, v\rangle + \langle B_t^* y_t, u\rangle - J_t(u,v)\}, \\
H_e(x_1,y_0) &= \sup_{u_e\in U_e}\{\langle y_0, B_e u_e + b_e\rangle - \sup_{v_e\in V_e}\{J_e(u_e,v_e) - \langle c_e,x_1\rangle - \langle C_e x_1, v_e\rangle\}\} \\
&= \langle b_e, y_0\rangle + \langle c_t, x_1\rangle + \sup_{u\in U_e}\inf_{v\in V_e}\{\langle C_t x_1, v_e\rangle + \langle B_e^* y_0, u_e\rangle - J_e(u_e,v_e)\}.
\end{aligned}
$$

In the final versions of these formulas, inf and sup can be interchanged because of the coercivity of the functions $\varphi_t, \varphi_e, \psi_t^*, \psi_t^*$, in the definitions (2.2) of J_t and J_e and the structure (2.4). This is justified as a minimax theorem by [**8**, Theorem 37.3], a result which establishes at the same time the finiteness of the expressions (4.5).

Our last task in the proof is to demonstrate that $J_t^*(r,s)$ depends continuously on (t,r,s). This could be carried out in detail with arguments like those that established the continuity of $H_t(x,y)$ in (t,x,y) in Proposition 4.1. There is a shortcut, though. The argument for H_t made no use of any particular properties of the vectors and matrices in (A3) other than their continuous dependence on t. The continuity property would therefore be present in particular if B_t and C_t were identity matrices, in which case the continuity property of H_t reduces to that of J_t. Thus $J_t(r,s)$ must be continuous with respect to (t,r,s) as claimed. $\square$

The next theorem establishes the equivalence between the Hamiltonian and minimax approaches to optimality.

Theorem 4.7. *Under assumptions (A1)–(A4), the Hamiltonian optimality conditions in (4.3) are satisfied by a pair of arcs $\bar{x}$ and $\bar{y}$ in $\mathcal{A}^1$ if and only if $\bar{x}$ and $\bar{y}$ are trajectories in $\mathcal{A}^\infty$ realized by controls $\bar{u} \in \mathcal{U}$ and $\bar{v} \in \mathcal{V}$ that satisfy the minimaximum principle (2.5).*

Proof. This result will be developed from the following formulas for the subgradients of H_t and H_e, which are based on the representations of these functions in Proposition 4.6. (Here ∂_1 and ∂_2 designate subgradients with respect to the first and second arguments of a bivariate function):

$$\begin{aligned}
\partial_1 H_t(x,y) &= A_t^* y + c_t + C_t^* \partial_2 J_t^*(B_t^* y, C_t^* x),\\
\partial_2 H_t(x,y) &= A_t x + b_t + B_t \partial_1 J_t^*(B_t^* y, C_t^* x),\\
\partial_1 H_e(x_1,y_0) &= c_e + C_e^* \partial_2 J_e^*(B_e^* y_0, C_e^* x_1),\\
\partial_2 H_e(x_1,y_0) &= b_e + B_e \partial_1 J_e^*(B_e^* y_0, C_e^* x_1).
\end{aligned}$$

These are obtained by the calculus in [**8**, Theorems 23.8 and 23.9] and are justified by the finiteness of the concave-convex functions J_t and J_e, which was proved in Proposition 4.6. We combine these formulas with the fact that

$$\partial H_t(x,y) = \partial_1 H_t(x,y) \times \partial_2 H_t(x,y) \quad \text{and} \quad \partial H_e(x_1,y_0) = \partial_1 H_e(x_1,y_0) \times \partial_2 H_e(x_0,y_0)$$

(cf. [**8**, §35]) and similarly for J_t^* and J_e^* to see that

$$\begin{aligned}
\partial H_t(x,y) &= \{(A_t^* y + C_t^* v + c_t, A_t x + B_t u + b_t) \,|\, (u,v) \in \partial J_t^*(B_t^* y, C_t x)\},\\
\partial H_e(x_1,y_0) &= \{(C_e^* v_e + c_t, B_e u_e + b_e) \,|\, (u_e,v_e) \in \partial J_e^*(B_e^* y_0, C_e x_1)\}.
\end{aligned} \tag{4.6}$$

A further observation is that

$$\begin{aligned}
\partial J_t^*(B_t^* y, C_t x) &= [\text{ set of all saddlepoints of } J_t(u,v) - \langle B_t^* y, u\rangle - \langle C_t x, v\rangle \text{ on } U_t \times V_t],\\
\partial J_e^*(B_e^* y_0, C_e x_1) &= [\text{ set of all saddlepoints of } J_e(u_e,v_e) - \langle B_e^* y_0, u_e\rangle - \langle C_e x_1, v_e\rangle \text{ on } U_e \times V_e],
\end{aligned} \tag{4.7}$$

which is true by conjugacy [**8**, Theorems 37.5 and 36.6]. As far as endpoints are concerned, we have from (4.6) and (4.7) that

$$\begin{aligned}
(\bar{y}_{t_1}, \bar{x}_{t_0}) \in \partial H_e(\bar{x}_{t_1}, \bar{y}_{t_0}) \Leftrightarrow & \exists (\bar{u}_e, \bar{v}_e) \in \partial J_e^*(B_e^* \bar{y}_{t_1}, C_e \bar{x}_{t_1})\\
& \text{with } \bar{x}_{t_0} = B_e \bar{u}_e + b_e, \;\; \bar{y}_{t_1} = C_e^* \bar{v}_e + c_e.
\end{aligned} \tag{4.8}$$

Similarly, for any t it is true that

$$\begin{aligned}(-\dot{\bar{y}}_t, \dot{\bar{x}}_t) \in \partial H_t(\bar{x}_t, \bar{y}_t) \Leftrightarrow & \exists (\bar{u}_t, \bar{v}_t) \in \partial J_t^*(B_t^* \bar{y}_t, C_t \bar{x}_t) \\ & \text{with } \dot{\bar{x}}_t = A_t \bar{x}_t + B_t \bar{u}_t + b_t, \;\; -\dot{\bar{y}}_t = A_t^* \bar{y}_t + C_t^* \bar{v}_t + c_t.\end{aligned} \tag{4.9}$$

If the minimaximum principle (2.6) holds for some choice of $\bar{u} \in \mathcal{U}$ and $\bar{v} \in \mathcal{V}$, and $\bar{x}$ and $\bar{y}$ are the corresponding state trajectories in the control problems $(\mathcal{P})$ and $(\mathcal{Q})$, we do have (4.7) and (4.8), the latter for a.e. t. Then, according to the formulas we have arrived at, $\bar{x}$ and $\bar{y}$ satisfy the Hamiltonian conditions (4.3) and belong to $\mathcal{A}_n^\infty[t_0, t_1]$ instead of just $\mathcal{A}_n^1[t_0, t_1]$.

Conversely, suppose $\bar{x}$ and $\bar{y}$ are elements of $\mathcal{A}_n^1[t_0, t_1]$ for which the Hamiltonian conditions (2.6) hold. Certainly then there is a choice of $\bar{u}_e$ and $\bar{v}_e$ for which (4.8) holds. We know further that for almost every t we can find u_t and v_t satisfying (4.9). It must be shown that these vectors can be chosen in such a way that the functions $t \mapsto u_t$ and $t \mapsto v_t$ are *measurable and essentially bounded.* Inasmuch as J_t^* is finite everywhere, the subgradient set $\partial J_t^*(r, s)$ is always nonempty and compact (by [**8**, Theorem 23.4] as applied to the separate arguments). The continuity of $J_t^*(r, s)$ in t implies further that the multifunction $(t, r, s) \mapsto \partial J_t^*(r, s)$ is locally bounded and of closed graph [**8**, Theorem 24.5]. Therefore the multifunction $t \mapsto \partial J_t^*(\bar{x}_t, \bar{y}_t)$ is likewise locally bounded and of closed graph, as well and nonempty-valued. A measurable selection then exists and must be essentially bounded (cf. [**7**, Cor. 1C]). This selection is in the form of a function $t \mapsto (\bar{u}_t, \bar{v}_t)$ with exactly the properties we need. □

References.

1. R.T. Rockafellar, "Linear-quadratic programming and optimal control", SIAM J. Control Opt. **25**(1987), 781–814.
2. R.T. Rockafellar, "On the essential boundedness of solutions to problems in piecewise linear-quadratic optimal control", in *Analyse Mathématique et Applications*, F. Murat and O. Pironneau (eds.), Gauthier-Villars, Paris, 1988.
3. R.R. Rockafellar, "Conjugate convex functions in optimal control and the calculus of variations", J. Math. Anal. Appl. **32** (1970), 174-222.
4. R.T. Rockafellar, "Existence and duality theorems for convex problems of Bolza", Trans. Amer. Math. Soc. **139** (1971), 1–40.
5. G. Salinetti and R.J-B Wets, "On the convergenct of sequences of convex sets in finite dimensions", SIAM Review **21** (1979), 18–33.
6. R.J-B Wets, "Convergence of convex functions, variational inequalities and convex optimization problems", in *Variational Inequalities and Complementarity Problems*, R.W. Cottle, F. Giannessi, and J.-L. Lions (eds.), Wiley, 1980, 376–403.
7. 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. No. 543, Springer-Verlag, Berlin, New York, 1976, 157-207.
8. R.T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, Princeton N.J., 1970.

9. R.J. Wijsman, “Convergence of sequences of convex sets, cones and functions, II”, Trans. Amer. Math. Soc. **123** (1966), 32–45.
10. E. Michael, “Continuous selections, I”, Ann. of Math. **63** (1956), 361–382.
11. L. McLinden and R.C. Bergstrom, “Preservation of convergence of convex sets and functions”, Trans. Amer. Math. Soc. **268** (1981), 127–142.
12. R.T. Rockafellar, “Existence theorems for general control problems of Bolza and Lagrange”, Advances in Math. **15** (1975), 312–333.
13. R.T. Rockafellar, “Generalized Hamiltonian equations in convex problems of Lagrange”, Pacific J. Math. **32** (1970), 46–63.
14. F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983.
15. R.T. Rockafellar, “State constraints in convex control problems of Bolza”, SIAM J. Control **10** (1972), 691–715.
16. R.T. Rockafellar, “Dual problems of Lagrange for Arcs of bounded variation”, in *Calculus of Variations and Control*, Academic Press, New York, 1976, 155–192.
17. R.T. Rockafellar, ‘Optimality conditions for convex control problems with nonnegative states and the possibility of jumps”, in *Game Theory and Math. Economics*, O. Moeschlin (ed.), North-Holland, 1981, 339-349.