# Conjugate Convex Functions, Duality, and Optimal Control Problems I: Systems Governed by Ordinary Differential Equations†

W. HEINS AND S. K. MITTER
*Systems Research Center, Case Western Reserve University*
*Cleveland, Ohio*

Communicated by Mihajlo D. Mesarovic

ABSTRACT

This paper presents a general and complete duality theory for optimal control of systems governed by linear ordinary differential equations with a convex cost criterion. Existence theorems for optimal control problems are obtained using the duality theory and a direct relationship between duality and Pontryagin's Maximum Principle is exhibited. Finally, applications to decomposition of optimal control problems are presented.

## 1. INTRODUCTION

An important part of mathematical programming, from both the theoretical and the computational points of view, is duality theory. For a class of variational problems, Friedrichs [1] outlined a duality theory utilizing the Legendre Transform. In recent years several papers have been written on a duality theory for optimal control problems [2–4]. However it is the authors' belief that no satisfactory and rigorous duality theory for optimal control problems is available, an exception being the very recent work of Van Slyke and Wets [5].

Fenchel [6, 7] initiated a duality theory for finite-dimensional mathematical programs using the important idea of conjugate convex functions. Over the past few years due to the work of Bröndsted [8], Moreau [9, 10], and especially Rockafellar [11–15], various aspects of the theory of convex functions and conjugate convex functions have reached a satisfying stage of completeness. Rockafellar [16] and Dieter [17] have also given a duality theory for infinite-dimensional convex programs.

† This research was partially supported by NSF Grant GK-3714.

*Information Sciences* **2** (1970), 211– 43

An important part of the theory of convex functions is a theory of differentiability for such functions. The notion of subdifferentiability was first introduced by Moreau [18] and later developed by Rockafellar and Bröndsted [19, 20]. The most recent contribution on the differentiability of convex functions is due to Asplund [21].

Recent work of Gamkrelidze [22, 23], Halkin [24, 25] and Neustadt [26–28] have shown that optimal control problems can be regarded as mathematical programs in infinite-dimensional spaces and very general necessary conditions of optimality for infinite-dimensional mathematical programs have been given. These conditions include all known first order necessary conditions of optimality for control problems and in particular the maximal principle.

Motivated by computational consideration (in particular for state-constrained problems), necessary conditions for optimality have been derived for convex programming problems by Pshenichnyi [29], Demyanov and Rubinov [30] and Kantorovich and colleagues [31]. An integral part of the work of Pshenichnyi and Demyanov has been a differentiability theory for convex functions. Work related to this had been done by Danskin earlier [32]. These workers were apparently unaware of the work of Moreau, Rockafellar, and others on subdifferentiability.

Pontryagin's Maximal Principle may be thought of as a duality result for non-linear non-convex optimal control problems. For linear optimal control problems with a cost function being a linear form, this was demonstrated by Van Slyke and Wets [33]. This has also been done by Rockafellar for some special linear optimal control problems. However the explicit relationship between the Pontryagin Maximal Principle and a mathematical programming type of duality theory has not been shown for a general class of linear optimal control processes.

The main objectives of this paper are:

(i) To present a duality type theory for a general class of finite-dimensional linear optimal control problems with convex cost criteria and convex control and state variable constraints.

(ii) To show the relationship of the duality theory to existence theorems of optimal control and to necessary and sufficient conditions of optimality.

(iii) To show how certain decomposition results may be obtained using this theory.

(iv) To show how the duality theory may be applied to certain control processes described by linear Functional Differential Equations and linear Partial Differential Equations. This will be done in Part II of the paper.

The main tools in this development are the recent theory of convex functions and a duality theory of mathematical programming due to Rockafellar [16]. These tools can also be effectively used to widen the application of the

Carathéodory-Hamilton-Jacobi-Bellman theory for optimal control problems. This will be done in Part III of the paper.

## 2. CONVEX FUNCTIONS ON TOPOLOGICAL VECTOR SPACES

In the sequel we shall use various properties and results related to convex functions. These are summarized in this section. A more complete treatment may be found in [8, 10, 16] and in the references cited in these papers.

Let $E$ and $E^*$ be real vector spaces in duality with respect to a certain real bilinear function $\langle \cdot, \cdot \rangle$. We shall assume that $E$ and $E^*$ have been assigned locally convex Hausdorff topologies compatible with this duality, so that elements of each space can be identified with continuous linear functionals on the other. $E$ and $E^*$ will then be referred to as topologically paired spaces.

### 2.1. *Properties of Convex Functions*

*Definition* 2.1. An infinite valued convex function $f$ on $E$ is an everywhere defined function with range in $[-\infty, +\infty]$ whose (upper) epigraph

$$\operatorname{epi}(f) = \{(x, \mu) | x \in E, \quad \mu \in R, \quad \mu \geqslant f(x)\}$$

is a convex set in $E \oplus R$.

If $f$ does not assume both $-\infty$ and $+\infty$ as values this definition of a convex function is equivalent to

$$f(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f(x_1) + (1-\lambda) f(x_2), \quad \forall x_1 \in E, \forall x_2 \in E, \quad 0 \leqslant \lambda \leqslant 1.$$

*Definition* 2.2. The set,

$$\operatorname{dom}(f) = \{x \in E | f(x) < +\infty\}$$

is the effective domain of $f$.

Note that $\operatorname{dom}(f)$ is the projection of the epigraph $\operatorname{epi}(f)$ of $f$ on $E$.

*Definition* 2.3. A convex function $f$ on $E$ is said to be proper if $f(x) > -\infty$ for all $x \in E$ and $f(x) < +\infty$ for at least one $x \in E$.

If $f$ is a proper convex function, then $\operatorname{dom}(f)$ is a non-empty convex set and $f$ is finite there. On the other hand given a finite valued convex function $f$ on a non-empty convex set $C$ in $E$, we obtain a proper convex function $f_0$ with effective domain $C$ by

$$f_0(x) = \begin{cases} f(x), & x \in C, \\ +\infty, & x \notin C. \end{cases}$$

*Definition* 2.4. The indicator function $\psi_C$ of a non-empty convex set $C$ in $E$ is defined as

$$\psi_C(x) = \begin{cases} 0, & x \in C, \\ +\infty, & x \notin C. \end{cases}$$

Thus, the indicator function of a non-empty convex set is a proper convex function.

*Definition* 2.5. A convex function $f$ on $E$ is lower semi-continuous (l.s.c.) if, for each $\mu \in R$, the convex level set

$$\{x \in E \mid f(x) \leqslant \mu\}$$

is a closed set in $E$.

Lower semi-continuity of convex functions is a constructive property. Given any convex function $f$ on $E$, we may construct a l.s.c. convex function $\bar{f}$ on $E$ by taking

$$\bar{f}(x) = \liminf_{z \to x} f(z), \quad \forall x \in E.$$

### 2.2. *Conjugate Convex Functions*

*Definition* 2.6. Let $f$ be a proper convex function on $E$. Its conjugate function $f^*$ on $E^*$ (with respect to the given bilinear function $\langle \cdot, \cdot \rangle$) is defined by

$$f^*(x^*) = \sup_{x \in E} \{\langle x, x^* \rangle - f(x)\}, \quad \forall x^* \in E^*.$$

The function $f^*$ is a l.s.c. convex function but not necessarily proper. However, if $f$ is a l.s.c. proper convex function, then $f^*$ is also l.s.c. proper convex and

$$(f^*)^* = f.$$

Thus, a one-to-one correspondence between the l.s.c. proper convex functions on $E$ and those on $E^*$ is defined by the formulas

$$\begin{cases} f^*(x^*) = \sup_{x \in E} \{\langle x, x^* \rangle - f(x)\}, \\ f(x) = \sup_{x^* \in E^*} \{\langle x, x^* \rangle - f^*(x^*)\}. \end{cases} \tag{2.1}$$

Functions $f$ and $f^*$ satisfying (2.1) are said to be conjugate to each other.

*Definition* 2.7. An element $x^* \in E^*$ is said to be a subgradient of the convex function $f$ at the point $x$ if

$$f(y) \geqslant f(x) + \langle y - x, x^* \rangle, \quad \forall y \in E.$$

The set of all subgradients at $x$, denoted by $\partial f(x)$ is a weak* closed convex set in $E^*$ which might be empty. If $\partial f(x)$ is non-empty, the convex function $f$ is said to be subdifferentiable at $x$. If $f$ is differentiable in the sense of Fréchet, $\partial f(x)$ consists of a single point, namely, the gradient $\nabla f(x)$ of $f$ at $x$. Further, if $f(x)$ is finite the one-sided directional derivative

$$f'(x; z) = \lim_{\lambda \downarrow 0} \frac{f(x + \lambda z) - f(x)}{\lambda}$$

exists, although it might be infinite and it is a positively homogeneous convex function of $z$. Then,

$$x^* \in \partial f(x) \Leftrightarrow f'(x;z) \geqslant \langle z, x^* \rangle, \quad \forall z \in E,$$

and such an $x^*$ exists if and only if $f'(x;z)$ is bounded below in $z$ in some neighborhood of 0. A useful fact regarding l.s.c. proper convex functions on a Banach space is the following.

THEOREM 2.8. *Let $E$ be a Banach space and $f$ a l.s.c. proper convex function on $E$ with effective domain $C$. Assume $C$ has non-empty relative interior* $\mathrm{ri}(C)$. *Then $f$ is continuous on* $\mathrm{ri}(C)$. ||

A function $g$ on $E$ is said to be concave if $-g$ is convex. The theory of concave functions therefore parallels that of convex functions with certain natural changes. In particular,

$$g^*(y^*) = \inf_{y \in E} \{\langle y, y^* \rangle - g(y)\},$$
$$g(y) = \inf_{y^* \in E^*} \{\langle y, y^* \rangle - g^*(y^*)\}$$

define a one-to-one correspondence between the upper semi-continuous (u.s.c.) proper concave functions on $E$ and those on $E^*$.

The following property of conjugate functions which follows easily from the definitions will be frequently used and is stated as a theorem.

THEOREM 2.9. *If $f$ is a l.s.c. proper convex function on $E$ and $g$ is an u.s.c. proper concave function on $E$ then,*

$$x^* \in \partial f(x) \leftrightarrow x \in \partial f^*(x^*) \Leftrightarrow f(x) + f^*(x^*) = \langle x, x^* \rangle$$
$$x^* \in \partial g(x) \leftrightarrow x \in \partial g^*(x^*) \Leftrightarrow g(x) + g^*(x^*) = \langle x, x^* \rangle. \ ||$$

2.3. *Integrals with Convex Integrands*

We give some results on integrals of convex functions of the type

$$\int_T f(t, x(t))\, dt, \ \int_T f^*(t, x^*(t))\, dt$$

where $f(t,x)$ is a l.s.c. proper convex function of $x$ for each $t$ and thus in general not continuous in $x$ and $f^*(t,x^*)$ is the conjugate of $f(t,x)$ for each $t$. In particular we would like to know the relationship between these integrals (if defined in an appropriate manner) regarded as functionals on the spaces to which the curves $x(t)$ and $x^*(t)$ belong. The following facts were proved in [34] in case $f$ is defined on $R^{1+n}$ but the results can be readily extended to separable Hilbert spaces as encountered in optimal control theory. In fact, the extension will follow if one formulates Lemmas 1 and 2 of [34] for a separable Hilbert space and furthermore recalls that the results of [9], used in [34], were proved for a Hilbert space rather than for $R^n$.

The following two lemmas correspond to Lemmas 1 and 2 of reference [34].

*Definition* 2.10. Let $T$ be a measure space with $\sigma$-finite measure $dt$ and $L$ be a real vector space of measurable functions $u$ from $T$ to a separable Hilbert space $H$ and consider,

$$I_f(u) = \int_T f(t, u(t))\, dt$$

where $f$ is a function from $T \times H$ to $[-\infty, +\infty]$. Then $f$ is called a normal convex integrand if it satisfies the conditions:

(i) $f(t,x)$ is a l.s.c. proper convex function on $H$ for each fixed $t$.

(ii) There is a countable collection $\mathscr{U}$ of measurable functions $u$ from $T$ to $H$ such that

(a) for each $u \in \mathscr{U}, f(t,u(t))$ is measurable in $t$,

(b) for each $t$, $\mathscr{U}_t \cap \operatorname{dom}(f(t,x))$ is dense in $\operatorname{dom}(f(t,x))$, where

$$\mathscr{U}_t = \{u(t) | u \in \mathscr{U}\}.$$

LEMMA 2.11. *Suppose $f(t,x) = F(x)$ for all $t$, where $F$ is a l.s.c. proper convex function on $H$. Then $F$ is a normal convex integrand.*

*Proof.* By the separability of $H$, there exists a countable dense subset $D$ of the non-empty convex effective domain of $F$.

Let $\mathscr{U}$ consist of the constant functions on $T$ with values in $D$. Then $\mathscr{U}$ satisfies conditions (a), (b) in Definition 2.10 and since $F$ is l.s.c. and proper, $F$ is a normal integrand. ||

LEMMA 2.12. *Let the function $f(t,x)$ on $T \times H$ have values in $[-\infty, +\infty]$ such that $f(t,x)$ is measurable in $t$ for each fixed $x$ and for each $t$, $f(t,x)$ is a l.s.c. proper convex function in $x$ with interior points in its effective domain. Then $f$ is a normal convex integrand.*

*Proof.* Let $D$ be a countable dense subset of $H$ and let $\mathscr{U}$ be the set of constant functions on $T$ with values in $D$.

Then $\mathscr{U}$ satisfies condition (a) of Definition 2.10. Further $D$ has a dense intersection with the interior of $\operatorname{dom}(f(t,x))$ and therefore with $\operatorname{dom} f(t,x)$ because $\operatorname{dom}(f(t,x))$ as a convex set with non-empty interior has no isolated points. ||

*Definition* 2.13. Let $T$ be a measure space with a $\sigma$-finite measure $dt$, $H$ be a separable Hilbert space, and $\mathscr{L}$ a real vector space of measurable functions from $T$ to $H$. Then, $\mathscr{L}$ is said to be decomposable if it satisfies the following conditions:

(i) $\mathscr{L}$ contains every bounded measurable function from $T$ to $H$ which vanishes outside a set of finite measure.

(ii) If $u \in \mathscr{L}$ and $C$ is a set of finite measure in $T$, then $\mathscr{L}$ contains $\chi_C u$ where $\chi_C$ is the characteristic function of $C$.

In other words, if $\mathscr{L}$ is decomposable one can alter functions in $\mathscr{L}$ arbitrarily in a bounded fashion on every set of finite measure. Namely, subtract $\chi_C u$ from $u$ and add any bounded measurable function vanishing outside $C$. If $\mathscr{L}^*$ is topologically paired to $\mathscr{L}$ with respect to $\langle\cdot,\cdot\rangle$ such that $\langle u(t), u^*(t)\rangle_H$ is summable in $t$ for every $u \in \mathscr{L}$, $u^* \in \mathscr{L}^*$, then condition (i) of Definition 2.13 also implies that the functions in $\mathscr{L}^*$ are summable on sets of finite measure. An important class of function spaces which are decomposable in this sense are the $L^p(0,T;R^m)$ spaces.

Finally we give the following theorem which relates the integrals of conjugate normal convex integrands as conjugate functionals.

THEOREM 2.14. *Let $\mathscr{L}$ and $\mathscr{L}^*$ be topologically paired by means of the summable inner product on $H$, that is*

$$\langle u, u^*\rangle = \int_T \langle u(t), u^*(t)\rangle_H \, dt \quad \forall u \in \mathscr{L}, \forall u^* \in \mathscr{L}^*$$

*and suppose $\mathscr{L}$, $\mathscr{L}^*$ are decomposable. Let $f$ be a normal convex integrand such that $f(t,u(t))$ is summable in $t$ for at least one $u \in \mathscr{L}$ and $f^*(t,u^*(t))$ is summable in $t$ for at least one $u^* \in \mathscr{L}^*$. Then the functionals $I_f$ on $\mathscr{L}$ and $I_{f^*}$ on $\mathscr{L}^*$ where*

$$I_f(u) = \int_T f(t,u(t))\, dt \quad I_{f^*}(u^*) = \int_T f^*(t,u^*(t))\, dt$$

*are proper convex functions conjugate to each other.*

*Proof.* See [34]. ||

## 3. DUALITY THEORY FOR ABSTRACT LINEAR OPTIMAL PROCESSES (ROCKAFELLAR)

Let $E$ and $F$ be Banach spaces and let $u \mapsto \Lambda u$ mapping $E$ to $F$ be a continuous linear transformation. $E$ is to be thought of as the *control space* and $F$ the *response space* of an abstract control process. Let $C \subseteq E$ and $D \subseteq F$ be convex sets. $C$ will be referred to as the *control restraint set* and $D$ the *response restraint set*. A control $u \in C$ for which the corresponding response $\Lambda u \in D$ is termed an *admissible controller*.

(3.1) Let $H(u) = f(u) - g(\Lambda u)$ be a function with values in $[-\infty, +\infty]$ and consider the *primal problem*

$$(P) \quad \begin{cases} \text{minimize } H(u) = f(u) - g(\Lambda u) \\ \text{subject to } u \in C \text{ and } \Lambda u \in D. \end{cases}$$

A controller $\bar{u}$ is a *solution* to $(P)$ if and only if $\bar{u} \in C$, $\Lambda \bar{u} \in D$ and the infimum of $H(u)$ is finite and attained at $\bar{u}$. Such a $\bar{u}$ is called an *optimal control.* For the problem $(P)$ we make the assumption that $f$ is a l.s.c. proper convex function with effective domain $C$ and $g$ is an u.s.c. proper concave function with effective domain $D$. Then the minimand in $(P)$ is a proper convex function or identically $+\infty$.

*Remark.* If $f$ is a finite convex function which does not already have $C$ for its effective domain, we may define the new function $f_0$ such that

$$f_0(u) = \begin{cases} f(u), & u \in C, \\ +\infty, & u \notin C, \end{cases}$$

and if necessary lower the values of $f_0$ on the boundary of $C$ so that it becomes l.s.c. Similarly a finite concave function $g$ can be constructively modified to obtain an u.s.c. proper concave function with $D$ as its effective domain. ||

To define the dual problem $(P^*)$ of $(P)$ let $E^*$ be a real Banach space, topologically paired with $E$ with respect to a bilinear real valued function $\langle \cdot, \cdot \rangle$ on $E \times E^*$. That is, the elements of each space can be identified with continuous linear functionals on the other by means of $\langle \cdot, \cdot \rangle$. Further let $F^*$ be a real Banach space topologically paired to $F$ with respect to $\langle \cdot, \cdot \rangle$. In most of the cases $E^*$, $F^*$ are the dual spaces of $E$, $F$ but there are interesting exceptions.

The dual $(P^*)$ of $(P)$ is defined as

$$(P^*) \quad \begin{cases} \text{maximize } H^*(x^*) = g^*(x^*) - f^*(\Lambda^* x^*) \\ \text{subject to } x^* \in D^*, \Lambda^* x^* \in C^*, \end{cases}$$

where $f^*$ is the conjugate of $f$ with effective domain $C^*$, $g^*$ is the conjugate of $g$ with effective $D^*$, and $\Lambda^*$ is the adjoint transformation of $\Lambda$. Of course, $f^*$, $g^*$, $\Lambda^*$ are defined with respect to $\langle \cdot, \cdot \rangle$. Thus the maximand in $(P^*)$ is a proper concave function or identically $-\infty$. Note that $H^*$ is not the conjugate function of $H$.

$\bar{x}^*$ is a solution to $(P^*)$ if and only if $\bar{x}^* \in D^*$, $\Lambda^* \bar{x}^* \in C^*$ and the supremum of $H^*(x^*)$ is attained at $\bar{x}^*$, in which case the supremum is finite.

We note that the minimization in $(P)$ can be carried out over all of $E$ (unconstrained problem) because of the fact that $C$ and $D$ are the effective domains of $f$ and $g$, respectively. For the same reason the maximization in $(P^*)$ can be taken over all of $F^*$. ||

The following results in this section are all due to Rockafellar [16]. We state the theorems that we use in later sections.

LEMMA 3.1. $\inf H(u) \geqslant \sup H^*(x^*)$. ||

In duality theory the concept of stability is very important, where a stably set process $(P)$ is defined as follows.

Consider the perturbed primal problem $(P(z))$ for some $z \in F$ where

$$(P(z)) \quad \begin{cases} \text{minimize } H(u, z) = f(u) - g(\Lambda u - z) \\ u \in C,\ \Lambda u \in D. \end{cases}$$

*Definition* 3.2. If $\inf H(u,0) = \inf H(u)$ is finite then the process $(P)$ is said to be stably set if, in some neighborhood $N$ of the origin in $F$,

$$\lim_{\epsilon \downarrow 0} \frac{\inf H(u, \epsilon z) - \inf H(u)}{\epsilon} > \infty, \quad \forall z \in N.$$

We shall adopt the convention that in case $\inf H(u)$ is $-\infty(+\infty)$ we say that $(P)$ is stably set (unstably set).

It can be proved that the function $p$ defined as

$$p(z) = \inf_u H(u, z)$$

is a convex function on $F$, so that its one-sided directional derivative at 0,

$$p'(0; z) = \lim_{\epsilon \downarrow 0} \frac{\inf H(u, \epsilon z) - \inf H(u)}{\epsilon}$$

exists for all $z$ although it may be infinite. Therefore the limit in the definition above is well defined for every $z$.

A sufficient condition for the stability of problem $(P)$ in control-theoretic terms may be given. For this purpose we introduce

*Definition* 3.3. The linear system defined by the equation $x = \Lambda u$ is said to be reachable to the response restraint set $D$ if there exists a $u \in C$ such that $\Lambda u \in D$ where $C$ is the control restraint set.

THEOREM 3.4. *If the linear system $x = \Lambda u$ is reachable to* $\operatorname{int}(D)$ *(the interior of $D$) that is there exists a $u$ at which $f$ is finite and $\Lambda u$ is in* $\operatorname{int}(D)$ *then $(P)$ is stably set and* $\inf_u H(u, z)$ *is a continuous function in some neighborhood of the origin.* ||

There are problems of interest in control theory which satisfy this sufficient condition and hence are automatically stable set.

*Example* 1. For the problem $(P)$ if $0 \in C$ and $0 \in \operatorname{int}(D)$ then $P$ is stably set.

*Example* 2. Minimize $f(u) - g(\Lambda u)$
subject to $u \in C$.

Here $g$ is finite everywhere in $F$ (no constraints on the response). If $C$ is non-empty then this problem is stably set. ||

The following theorems give the relationship between ($P$) and its dual ($P^*$). By $\min H(u)$ (resp. $\max H^*(x^*)$) we mean that the infimum of $H(u)$ (resp. supremum of $H^*(x^*)$) are attained at some $u$ (resp. $x^*$).

THEOREM 3.5. *The process* ($P$) *is stably set if and only if*

$$\inf H(u) = \max H^*(x^*).$$

*Dually* ($P^*$) *is stably set if and only if*

$$\min H(u) = \sup H^*(x^*).\ \|$$

THEOREM 3.6. ($P$) *is stably set and has a solution if and only if* ($P^*$) *is stably set and has a solution.* ||

Optimal controllers, that is solutions to ($P$), can be characterized by certain subdifferentiability conditions on $f$ and $g$ in analogy with the case in ordinary calculus when $f$ and $g$ are differentiable. To get a better insight into the nature of the next theorem note that the convex function $f - g \circ \Lambda$ attains a finite minimum in precisely those points $\bar{u}$ where 0 is a subgradient. Thus optimal controllers $\bar{u}$ satisfy

$$0 \in \partial(f - g \circ \Lambda)(\bar{u})$$

We will prove in a moment that

$$0 \in \partial f(u) - \Lambda^*(\partial g(\Lambda u)) \Leftrightarrow \Lambda u \in \partial g^*(x^*), \quad \Lambda^* x^* \in \partial f(u) \tag{3.2}$$

for some $x^* \in F^*$.

Provided

$$0 \in \partial(f - g \circ \Lambda)(u) \Leftrightarrow 0 \in \partial f(u) - \Lambda^*(\partial g(\Lambda u)) \tag{3.3}$$

holds, the right-hand side of (3.2) constitutes a convenient form for characterizing optimal controllers. Indeed, (3.3) is true under the stability condition; see the next theorem.

To verify (3.2) assume that $0 \in \partial f(u) - \Lambda^*(\partial g(\Lambda u))$ for some $u \in E$. Then there exists a $u^* \in \partial f(u)$ and $x^* \in \partial g(\Lambda u)$ such that $u^* = \Lambda^* x^*$, and $x^* \in \partial g(\Lambda u) \to \Lambda u \in \partial g^*(x^*)$. On the other hand if $u$, $x^*$ satisfy the right-hand side of (3.2) then $x^* \in \partial g(\Lambda u) \to \Lambda^* x^* \in \Lambda^*(\partial g(\Lambda u))$ so that $0 \in \partial f(u) - \Lambda^*(\partial g(\Lambda u))$. This proves (3.2). ||

THEOREM 3.7. *The process* ($P$) *and its dual* ($P^*$) *are stably set, with solutions* $\bar{u}$, $\bar{x}^*$, *respectively, if and only if* $\bar{u}$, $\bar{x}^*$ *satisfy* $\Lambda \bar{u} \in \partial g^*(\bar{x}^*)$ *and* $\Lambda^* \bar{x}^* \in \partial f(\bar{u})$. ||

## 4. OPTIMAL CONTROL FOR FINITE-DIMENSIONAL LINEAR PROCESSES WITH CONVEX COST CRITERIA

The duality theory developed in the previous section will now be applied to a class of finite-dimensional linear differential processes with convex cost criteria. In Section 4.1 the data and hypotheses of the problem are given and the functions $f$, $g$, $f*$, $g^*$, $\Lambda$, $\Lambda^*$ are explicitly determined. An existence and uniqueness theorem which makes use of the duality theory is presented. The hypotheses on the existence theorem appear to be weaker than those appearing in the literature. Introducing the adjoint differential equation we arrive at Pontryagin's Maximal Principle in generalized form and in so doing reveal the existence of a close relationship between duality and the maximal principle.

### 4.1. *Formulation of the Problem*

Consider the linear control process

$$(\mathscr{L}) \qquad dx/dt = A(t)\,x(t) + B(t)\,u(t),$$

where $\forall t \in [0,T]$, $x(t) \in R^n$, $u(t) \in R^m$ and $A(t)$, $B(t)$ are $n \times n$ and $n \times m$ matrices which are continuous for $\forall t \in [0,T]$.

Let $L^p[0,T;R^n]$ denote the space (equivalence classes) of Lebesgue measurable functions $v(t)$ with values in $R^n$ such that

$$\|v\|_p = \left\{\int_0^T \|v(t)\|_{R^n}^p\,dt\right\}^{1/p} < \infty, \quad 1 \leqslant p < \infty$$

and

$$\|v\|_p = \text{ess. sup.}\,[\|v(t)\|_{R^n} | t \in [0,T]] < \infty, \quad p = \infty$$

where $\|v(t)\|_{R^n}$ is an appropriate norm on $R^n$.

Whenever there is no possibility of confusion, we shall drop the subscripts on the various norms.

The data of the problem are as follows: Let $\phi(u)$ denote the solution of $(\mathscr{L})$ on $[0,T]$ corresponding to some control function $u$.

(i) $\phi(u)(0) = x \in G_0$, where $G_0$ is a convex subset of $R^n$.

(ii) The solution $\phi(u)$ of $(\mathscr{L})$ is in $L^p[0,T;R^n]$, $1 \leqslant p \leqslant \infty$ and are required to lie in the convex subset of $L^p[0,T;R^n]$, where

(4.1) $X =$
$$\{\phi \in L^p[0,T;R^n] | \phi(0) \in G_0,\ \phi(t) \in G_t \subseteq R^n \text{ a.e. on } ]0,T],\ G_t \text{ is convex}\},$$

(iii) The controllers $u$ are in $L^r[0,T;R^m]$, $1 \leqslant r \leqslant \infty$, and the class of admissible controllers $\mathscr{U}$ is a convex subset of $L^r[0,T;R^m]$, where

(4.2) $\mathscr{U} = \{u \in L^r[0,T;R^m] | u(t) \in \Omega \subseteq R^m$ a.e. on $[0,T]$, $\Omega$ is convex, and the response $\phi(u) \in X\}$.

It is required to choose a controller $\bar{u} \in \mathscr{U}$ such that

(4.3) $$H\{(u,x)\} = l_0(x) + l_1(\phi(u)(T)) + \int_0^T [h(t,u(t)) + k(t,\phi(u)(t))]\,dt$$

is minimized.

The hypotheses are:

(i) $l_0$ and $l_1$ are l.s.c. proper convex functions on $R^n$ with effective domain $G_0$ and $G_T$, respectively.

(ii) $h(t,z)$ is measurable in $t$ for each fixed $z \in R^m$ and for each $t$, $h(t,z)$ is a l.s.c. convex function in $z$ with effective domain $\Omega$.

(iii) $h(t,u(t))$ is summable in $t$ for all $u \in \mathscr{U}$ and $h^*(t,u^*(t))$ is summable in $t$ for at least one

$$u^* \in L^s[0,T;(R^m)^*],\dagger \quad \frac{1}{r} + \frac{1}{s} = 1.$$

(iv) Assumptions analogous to (ii) and (iii) hold for the function $k(t,\phi(t))$ and its conjugate

$$k^*(t,\phi^*(t)), \quad \phi \in L^p[0,T;R^n], \quad \phi^* \in L^q[0,T;(R^n)^*], \quad \frac{1}{p} + \frac{1}{q} = 1.$$

(v) $\Omega$ has a non-empty interior and $G_t$ has a non-empty interior $\forall t \in ]0,T]$.

The problem which we have defined will be referred to as (OC).

The problem defined above may be reformulated to correspond to problem $(P)$ of Section 3 by considering a linear transformation $\Lambda$ on $L^r[0,T;R^m] \oplus R^n$ rather than on $L^r[0,T;R^m]$. For this purpose define the functions

(4.4) $$f\colon L^r[0,T;R^m] \oplus R^n \to R$$
$$:\ (u,x) \mapsto l_0(x) + \int_0^T h(t,u(t))\,dt,$$

(4.5) $$g\colon L^p[0,T;R^n] \oplus R^n \to R$$
$$:\ (\phi,y) \mapsto -l_1(y) - \int_0^T k(t,\phi(t))\,dt,$$

† Superscript * denotes the dual space.

and the linear, bounded transformation

$$(4.6) \qquad \Lambda: L^r[0,T;R^m] \oplus R^n \to L^p[0,T;R^n] \oplus R^n$$

$$: (u,x) \mapsto (\phi(u), \phi(u)(T)), \text{ where } \phi(u) \text{ is defined by}$$

$$(4.7) \quad \phi(u)(t) = \Phi(t,0)\,x + \Phi(t,0) \int_0^t \Phi^{-1}(\tau,0)\,B(\tau)\,u(\tau)\,d\tau, \quad t \in [0,T],$$

and $\Phi$ satisfies the matrix equation

$$(4.8) \qquad d\Phi/dt = A(t)\,\Phi(t,0); \quad \Phi(0,0) = I.$$

Assuming that $\mathscr{U}$ is non-empty, then in view of hypotheses (ii), (v), and (iii), Lemma 2.12, and Theorem 2.14, it follows that $h$ is a normal convex integrand and $f$ is a proper convex function with effective domain $\mathscr{U} \oplus G_0$. Furthermore $f$ and $f^*$ are conjugate convex functions, where

$$(4.9) \quad f^*((u^*, x^*)) = l_0{}^*(x^*) + \int_0^T h^*(t, u^*(t))\,dt,$$

$$\forall (u^*, x^*) \in L^s[0, T; (R^m)^*] \oplus (R^n)^*.$$

The function $f^*$ is proper by hypothesis (iii) and it is automatically l.s.c. as the conjugate of the proper convex function $f$, which in turn implies that $f$ is also l.s.c.. $f^*$ has a non-empty effective domain in $L^s[0,T;(R^m)^*] \oplus (R^n)^*$ which we shall denote by $\mathscr{U}^* \oplus G_0{}^*$.

In a similar manner it may be shown that $g$ is an u.s.c. proper concave function with effective domain $X \oplus G_T$ with an u.s.c. proper conjugate function $g^*$, where

$$(4.10) \quad g^*((\phi^*, y^*)) = -l_1{}^*(-y^*) - \int_0^T k^*(t, -\phi^*(t))\,dt,$$

$$\forall (\phi^*, y^*) \in L^q[0, T; (R^n)^*] \oplus (R^n)^*.$$

The non-empty effective domain of $g^*$ in $L^q[0,T;(R^n)^*] \oplus (R^n)^*$ will be denoted by $X^* \oplus G_T{}^*$.

With the definitions of $f$, $g$, and $\Lambda$ as in (4.4), (4.5), and (4.6) the optimal control problem defined in this section may be reformulated as:

$$(P) \quad \text{minimize } H((u,x)) = f((u,x)) - g(\Lambda(u,x))$$

subject to

$$(u,x) \in \mathscr{U} \oplus G_0 \quad \text{and} \quad \Lambda(u,x) \in X \oplus G_T.$$

The dual problem $(P^*)$ is accordingly given by,

$$(P^*) \quad \text{maximize } H^*((\phi^*, y^*)) = g^*((\phi^*, y^*)) - f^*(\Lambda^*(\phi^*, y^*))$$

subject to

$$(\phi^*, y^*) \in X^* \oplus G_T^* \quad \text{and} \quad \Lambda^*(\phi^*, y^*) \in \mathscr{U}^* \oplus G_0^*.$$

A straightforward computation shows

$$(4.11) \quad \Lambda^*: \; L^q[0,T;(R^n)^*] \oplus (R^n)^* \to L^s[0,T;(R^m)^*] \oplus (R^n)^*: \quad (\phi^*, y^*) \mapsto (u^*, x^*),$$

where

$$(4.12) \quad \begin{cases} u^*(t) = B^*(t)\left[\int_t^T \Phi^*(\tau,t)\phi^*(\tau)\,d\tau + \Phi^*(T,t)y^*\right], \\ x^* = \int_0^T \Phi^*(t,0)\phi^*(t)\,dt + \Phi^*(T,0)y^*, \end{cases}$$

in which $B^*$, $\Phi^*$ are the transposed matrices of $B$ and $\Phi$, respectively. Thus the Rockafellar Duality Theory of Section 3 is applicable to this problem with obvious identifications. This theory is now used to obtain existence results for optimal control and the maximal principle. ||

### 4.2. *Existence Theorem for Optimal Control*

THEOREM 4.1. *Given the optimal control problem (OC), assume in addition:*

(i) *$\Omega$ is bounded and $G_0$ is compact,*
(ii) *for each $t \in [0,T]$,*

$$h(t,z) \geqslant c(t), \quad \forall z \in R^m,$$

*where $c(t)$ is summable on $[0,T]$,*
(iii) *the set $\mathscr{U}$ is non-empty.*

*Then there exists a pair $(\bar{u},\bar{x}) \in \mathscr{U} \oplus G_0$ such that*

$$\min\,[H(u,x)\,|\,u \in \mathscr{U}, x \in G_0] = H((\bar{u},\bar{x})).$$

*Moreover, if $f$ and $l_0$ are strictly convex on their effective domains, then $(\bar{u},\bar{x})$ is a unique optimal pair.*

*Proof.* For arbitrary $(u^*,y^*) \in L^s[0,T;(R^m)^*] \oplus (R^n)^*$ from (4.9)

$$\begin{aligned} f^*((u^*,y^*)) &= l_0^*(y^*) + \int_0^T h^*(t,u^*(t))\,dt \\ &= \sup\,[\langle y,y^*\rangle - l_0(y)\,|\,y \in R^n] \\ &\quad + \int_0^T \sup\,[\langle z,u^*(t)\rangle - h(t,z)\,|\,z \in R^m]\,dt \end{aligned}$$

$$f^*((u^*, y^*)) = \sup [\langle y, y^* \rangle - l_0(y) | y \in G_0] + \int_0^T \sup [\langle z, u^*(t) \rangle - h(t, z) | z \in \Omega] \, dt$$

$$\leqq m + M \int_0^T \|u^*(t)\| \, dt - \int_0^T c(t) \, dt,$$

where $m$ and $M$ are finite constants $< +\infty$.

Therefore the effective domain of $f^*$ is all of $L^s[0,T;R^m] \oplus R^n$. The function $g^*$ defined by (4.10) is proper and hence there exists a $(\bar{\phi}^*, \bar{y}^*)$ at which $g^*$ is finite and $\Lambda^*(\bar{\phi}^*, \bar{y}^*)$ is evidently in the interior of dom$(f^*)$. From the dual version of Theorem 3.4 it follows that $(P^*)$ is stably set.

From Theorem 3.5 it follows that there exists a pair

$$(\bar{u}, \bar{x}) \in L^r[0, T; R^m] \oplus R^n$$

such that $\inf H(u,x)$ is attained at $(\bar{u}, \bar{x})$. It remains to show that $(\bar{u}, \bar{x}) \in \mathscr{U} \oplus G_0$.

$X \oplus G_T$ is reachable by hypothesis, which implies that there is an admissible controller $u \in \mathscr{U}$ and an $x \in G_0$ with corresponding response $\phi(u)$ such that

$$(u, x) \in \mathrm{dom}\,(f)$$

and

$$\Lambda(u, x) = (\phi(u), \phi(u)(T)) \in X \oplus G_T = \mathrm{dom}\,(g).$$

Thus for this controller $u$ and initial state $x, f((u,x)) - g(\Lambda(u,x))$ is finite and therefore

(4.13) $$\min_{(u,\, x)} H(u, x) < +\infty.$$

Furthermore

$$H^*((\bar{\phi}^*, \bar{y}^*)) = g^*((\bar{\phi}^*, \bar{y}^*)) - f^*(\Lambda^*(\bar{\phi}^*, \bar{y}^*))$$

is finite so that

(4.14) $$\sup_{(\phi^*,\, y^*)} H^*((\phi^*, y^*)) > -\infty.$$

From Lemma (3.1), (4.13), and (4.14) it follows that

(4.15) $$+\infty > \min_{(u,\, x)} H((u, x)) \geqslant \sup_{(\phi^*,\, y^*)} H^*((\phi^*, y^*)) > -\infty$$

and hence

$$\min_{(u,\, x)} H((u, x)) = H((\bar{u}, \bar{x}))$$

is finite and thus necessarily $(\bar{u}, \bar{x}) \in \mathscr{U} \oplus G_0$. The uniqueness of $(\bar{u}, \bar{x})$ when $l_0$ and $h$ are strictly convex follows easily from a contradiction argument. ||

COROLLARY 4.2. *Consider the optimal control problem* $(OC)$ *with no constraints on the responses, that is,* $X \oplus G_T = L^p[0,T;R^n] \oplus R^n$. *Assume in addition that conditions* (i) *and* (ii) *of Theorem* 4.1 *are satisfied. Then there exists a pair* $(\bar{u},\bar{x}) \in \mathscr{U} \oplus G_0$ *such that* $\min[H((u,x))|u \in \mathscr{U}, x \in G_0] = H((\bar{u},\bar{x}))$. *Moreover, if* $f$ *and* $l_0$ *are strictly convex on their effective domains, then* $(\bar{u},\bar{x})$ *is a unique optimal pair.*

*Proof.* Condition (iii) of Theorem 4.1 is automatically satisfied. ||

*Remark.* It is clear that this theorem includes other existence theorems for linear optimal control problems with convex cost criteria which require $\Omega$ to be compact [37, Theorem 12, page 232]. Moreover, the case in which $\Omega$ is not bounded but closed and convex could be handled by putting hypotheses on the functions $h$ which makes the effective domain of $f^*$ all of $L^s[0,T;(R^m)^*] \oplus (R^n)^*$. ||

4.3. *Duality Theory and Pontryagin's Maximal Principle*

We would now like to demonstrate the relationship between the duality theory and Pontryagin's Maximal Principle for the class of problems considered in this section. For a somewhat different approach to this, see also Rockafellar [15]. Our development is done under weaker differentiability assumptions than that reported in the literature [37, Theorem 14, page 235]. These differentiability requirements have also been relaxed in recent work of Neustadt [28]. We now make the assumption that the functions $l_0$, $l_1$, and $k$ are subdifferentiable in $x$. We note that this is a weak assumption since $l_0$, $l_1$, and $k$ being convex functions on $R^n$ are subdifferentiable throughout the relative interior of their effective domains [11, page 137, Theorem 13.6].

THEOREM 4.3 (*Integral Maximal Principle*). *For the optimal control process* $(OC)$ *assume that there exists a controller* $u \in \mathscr{U}$ *and an* $x \in G_0$ *with corresponding response* $\phi(u)$, $\phi(u)(0) = x$, *such that* $(\phi(u), \phi(u)(T))$ *lies in the interior of* $X \oplus G_T$. *Then a pair* $(\bar{u},\bar{x})$ *is optimal with respect to the set of admissible controllers* $\mathscr{U}$ *and the set of allowable initial states* $G_0$ *if and only if there exists a*

$$\bar{y}^* \in (R^n)^*, \quad \bar{\phi}^* \in L^q[0,T;(R^n)^*]$$

*such that*

$$l_1(y) \geqslant l_1(\phi(\bar{u})(T)) + \langle y - \phi(\bar{u})(T), -\bar{y}^*\rangle, \quad \forall y \in R^n, \tag{4.16}$$

$$\int_0^T k(t,\phi(t))\,dt \geqslant \int_0^T k(t,\phi(\bar{u})(t))\,dt + \int_0^T [\langle \phi(t) - \phi(\bar{u})(t), -\bar{\phi}^*(t)\rangle]\,dt, \quad \forall \phi \in L^p[0,T;R^n], \tag{4.17}$$

$$l_0(y) \geqslant l_0(\bar{x}) + \langle y - \bar{x}, \bar{x}^*\rangle, \quad \forall y \in R^n, \tag{4.18}$$

*and*

$$(4.19)\quad \int_0^T h(t,u(t))\,dt \geqslant \int_0^T h(t,\bar{u}(t))\,dt + \int_0^T [\langle u(t)-\bar{u}(t),\bar{u}^*(t)\rangle]\,dt,$$

$$\forall u \in L^r[0,T;R^m],$$

*where* $\Lambda^*(\bar{\phi}^*,\bar{y}^*) = (\bar{u}^*,\bar{x}^*)$ *according to* (4.11), (4.12).

*Proof. Necessity.* By hypothesis, there exists a $u \in \mathscr{U}$ and $x \in G_0$ with response $\phi(u)$, where $(\phi(u),\phi(u)(T)) \in \operatorname{int}(X+G_T)$. It follows from Theorem 3.4 that problem $(P)$ corresponding to (OC) is stably set.

Assume $(\bar{u},\bar{x})$ is a solution to $(P)$. Hence from Theorem 3.6 $(P^*)$ is stably set and has a solution $(\bar{\phi}^*,\bar{y}^*)$. Moreover $(\bar{u},\bar{x})$ and $(\bar{\phi}^*,\bar{y}^*)$ satisfy

$$(4.20)\qquad \Lambda((\bar{u},\bar{x})) \in \partial g^*((\bar{\phi}^*,\bar{y}^*)),\quad \Lambda^*((\bar{\phi}^*,\bar{y}^*)) \in \partial f((\bar{u},\bar{x})).$$

From Theorem 3.7,

$$\Lambda(\bar{u},\bar{x}) = (\phi(\bar{u}),\phi(\bar{u})(T))\,(\text{see } 4.6,\, 4.7) \in \partial g^*((\bar{\phi}^*,\bar{y}^*)),$$

which implies

$$(\bar{\phi}^*,\bar{y}^*) \in \partial g((\phi(\bar{u}),\phi(\bar{u})(T))).$$

Hence, by Theorem 2.9,

$$g((\phi,y)) \leqslant g((\phi(\bar{u}),\phi(\bar{u})(T))) + \langle(\phi,y)-(\phi(\bar{u}),\phi(\bar{u})(T)),(\bar{\phi}^*,\bar{y}^*)\rangle$$

or

$$(4.21)\quad -l_1(y) - \int_0^T k(t,\phi(t))\,dt \leqslant -l_1(\phi(\bar{u})(T)) - \int_0^T k(t,\phi(\bar{u})(t))\,dt$$
$$+ \int_0^T [\langle \phi(t)-\phi(\bar{u})(t),\bar{\phi}^*(t)\rangle]\,dt$$
$$+ \langle y-\phi(\bar{u})(T),\bar{y}^*\rangle,$$
$$\forall(\phi,y) \in L^p[0,T;R^n] \oplus R^n.$$

Since $\phi$ is independent of $y$, it follows that

$$(4.22)\qquad l_1(y) \geqslant l_1(\phi(\bar{u})(T)) + \langle y-\phi(\bar{u})(T),-y^*\rangle,\quad \forall y \in R^n$$

and

$$(4.23)\qquad \int_0^T k(t,\phi(t))\,dt \geqslant \int_0^T k(t,\phi(\bar{u})(t))\,dt$$
$$+ \int_0^T [\langle \phi(t)-\phi(\bar{u})(t),-\bar{\phi}^*(t)\rangle]\,dt,$$
$$\forall \phi \in L^p[0,T;R^n].$$

Further,

$$\Lambda^*(\bar{\phi}^*, \bar{y}^*) = (\bar{u}^*, \bar{x}^*) \text{ (see 4.11, 4.12)} \in \partial f((\bar{u}, \bar{x}))$$

holds if and only if

$$l_0(y) + \int_0^T h(t, u(t))\,dt \geqslant l_0(\bar{x}) + \int_0^T h(t, \bar{u}(t))\,dt$$

$$+ \int_0^T [\langle u(t) - \bar{u}(t), \bar{u}^*(t)\rangle]\,dt$$

$$+ \langle y - \bar{x}, \bar{x}^*\rangle, \quad \forall (u, y) \in L^r[0, T; R^m] \oplus R^n.$$

Since $y$ is independent of $u$, it follows that

$$(4.24) \qquad l_0(y) \geqslant l_0(\bar{x}) + \langle y - \bar{x}, \bar{x}^*\rangle, \quad \forall y \in R^n$$

and

$$(4.25) \quad \int_0^T h(t, u(t))\,dt \geqslant \int_0^T h(t, \bar{u}(t))\,dt + \int_0^T [\langle u(t) - \bar{u}(t), \bar{u}^*(t)\rangle]\,dt, \quad \forall u \in L^r[0, T; R^m]$$

This proves the necessity part.

*Sufficiency.* The sufficiency part follows at once from Theorem 3.7. ||

*Remark.* In optimal control problem (OC) the set of admissible controllers and state constraint set are given locally (pointwise in time) by (4.2) and (4.1), respectively. Theorem 4.3 is true even if the constraints were specified in a global manner, that is, as convex sets in $L^r[0,T;R^m]$ and $L^p[0,T;R^n]$. For admissible controllers and state constraints given by (4.2) and (4.1), we can pass to a local form of the maximal principle. This is done in the subsequent theorem.

For the proof of the subsequent theorem we need the concept of a Lebesgue point.

Let $g$: $[0,T] \to R^n$ be such that $g \in L^1[0,T;R^n]$. Let $t \in [0,T]$ be fixed and let $0_j$ be a neighborhood of $t$. Let $|0_j|$ denote the measure of $0_j$. Then all points $t \in [0,T]$ for which

$$\frac{1}{|0_j|} \int_{0_j} g(\sigma)\,d\sigma \to g(t) \quad \text{as } |0_j| \to 0$$

are termed the Lebesgue points of the function $g$. It is known that the complement of the set of Lebesgue points of $g$ has measure zero. ||

THEOREM 4.4 (*Maximal Principle*). *Consider the control process* $(OC)$ *and assume that the hypothesis of Theorem* 4.3 *hold. Then a pair* $(\bar{u},\bar{x})$ *is optimal with respect to the set of admissible controllers* $\mathcal{U}$ *and the set of admissible initial states* $G_0$ *if and only if there exists an* $\bar{\eta}(t) \in (R^n)^*$ *satisfying*

$$d\eta/dt = -A^*(t)\,\eta(t) + x^*(t,\phi(\bar{u})(t)), \tag{4.26}$$

$$-\bar{\eta}(T) \in \partial l_1(\phi(\bar{u})(T)), \tag{4.27}$$

$$\bar{\eta}(0) \in \partial l_0(\bar{x}), \tag{4.28}$$

*where* $A^*(t)$ *is the transposed matrix of* $A(t)$, $x^*(t,\phi(\bar{u})(t)) \in \partial k(t,\phi(\bar{u})(t))$ *almost everywhere in* $[0,T]$ *and such that the maximal principle*

$$\langle B(t)\,\bar{u}(t),\eta(t)\rangle - h(t,\bar{u}(t)) = \max\,[\langle B(t)\,z,\bar{\eta}(t)\rangle - h(t,z) | z \in \Omega] \tag{4.29}$$

*holds almost everywhere in* $[0,T]$.

*Proof*. We shall prove that the necessary and sufficient conditions (4.16)–(4.19) are equivalent to the conditions (4.26)–(4.29).

Let $s \in ]0,T[$ which will be chosen appropriately later, and let $0_j$ be a neighborhood of $s$. Define

$$\phi_j(t) = \begin{cases} \phi(t) & \text{if } t \in 0_j,\ \phi \text{ arbitrary in } L^p[0,T;R^n], \\ \phi(\bar{u})(t) & \text{if } t \in ]0,T[ \sim 0_j. \end{cases} \tag{4.30}$$

Hence $\phi_j \in L^p[0,T;R^n]$. Since (4.17) is true for all $\phi \in L^p[0,T;R^n]$, we have in particular

$$\int_0^T [k(t,\phi_j(t)) - k(t,\phi(\bar{u})(t)) + \langle \phi_j(t) - \phi(\bar{u})(t), \bar{\phi}^*(t)\rangle]\,dt \geqslant 0, \tag{4.31}$$

and using (4.30), (4.31) reduces to

$$\frac{1}{|0_j|}\int_{0_j} [k(t,\phi(t)) - k(t,\phi(\bar{u})(t)) + \langle \phi(t) - \phi(\bar{u})(t), \bar{\phi}^*(t)\rangle]\,dt \geqslant 0. \tag{4.32}$$

Now choose $s$ to be a Lebesgue point of

$$k(t,\phi(t)) - k(t,\phi(\bar{u})(t)) + \langle \phi(t) - \phi(\bar{u})(t), \bar{\phi}^*(t)\rangle.$$

Then passing to the limit and utilizing the property of Lebesgue points, we conclude from (4.32)

$$k(s,\phi(s)) - k(s,\phi(\bar{u})(s)) + \langle \phi(s) - \phi(\bar{u})(s), \bar{\phi}^*(s)\rangle \geqslant 0, \tag{4.33}$$

and this is true for all Lebesgue points $s$.

Now $\phi \in L^p[0,T;R^n]$ was arbitrary. Hence (4.33) may be written as

$$k(t,y) \geqslant k(t,\phi(\bar{u})(t)) + \langle y - \phi(\bar{u})(t), -\bar{\phi}^*(t)\rangle$$

almost everywhere on $[0,T]$, for each $y \in R^n$.

Hence

(4.34) $$-\bar{\phi}^*(t) \in \partial k(t, \phi(\bar{u})(t)) \text{ almost everywhere on } [0, T].$$

A similar argument proves

(4.35) $h(t,z) \geqslant h(t,\bar{u}(t)) + \langle z - \bar{u}(t)\ \bar{u}^*(t)\rangle$ almost everywhere on $[0,T]$ for each $z \in R^m$.

That is,

(4.36) $$\bar{u}^*(t) \in \partial h(t, \bar{u}(t)) \text{ almost everywhere on } [0, T].$$

Let

(4.37) $$\bar{\eta}(t) = \int_t^T \Phi^*(\tau, t)\,\bar{\phi}^*(\tau)\,d\tau + \Phi^*(T, t)\,\bar{y}^*,$$

where $\Phi^*$ is the transposed matrix of $\Phi$.

Therefore

(4.38) $$\bar{\eta}(T) = \bar{y}^*,$$

and from (4.12)

(4.39) $$\begin{cases} \bar{\eta}(0) = \bar{x}^*, \\ \bar{u}^*(t) = B^*(t)\,\bar{\eta}(t). \end{cases}$$

It is easy to see that $\bar{\eta}(t)$ satisfies

(4.40) $$d\eta/dt = -A^*(t)\,\eta(t) + x^*(t, \phi(\bar{u})(t)), \quad -\bar{\eta}(T) \in \partial l_1(\phi(\bar{u})(T)), \quad \bar{\eta}(0) \in \partial l_0(\bar{x}),$$

where $x^*(t,\phi(\bar{u})(t)) = \bar{\phi}^*(t) \in \partial k(t, \phi(\bar{u})(t))$ almost everywhere in $[0,T]$. From (4.36), with $\bar{u}^*(t) = B^*(t)\,\bar{\eta}(t)$, we obtain

$$\langle B(t)\,\bar{u}(t), \bar{\eta}(t)\rangle - h(t, \bar{u}(t)) = \max\,[\langle B(t)\,z, \bar{\eta}(t)\rangle - h(t,z) | z \in \Omega],$$

which holds almost everywhere in $[0,T]$.

This proves the necessity part of the theorem.

*Sufficiency*. Suppose there is a vector $\bar{\eta}(t)$ satisfying

$$d\eta/dt = -A^*(t)\,\eta(t) + \bar{\phi}^*(t), \quad -\bar{\eta}(T) \in \partial l_1(\phi(\bar{u})(T)), \quad \bar{\eta}(0) \in \partial l_0(\bar{x}),$$

and

$$\bar{\phi}^*(t) \in \partial k(t, \phi(\bar{u})(t))$$

almost everywhere.

In the above $\phi$ $(\bar{u})$ is the response with $\phi(\bar{u})(0) = \bar{x}$, of the controller $\bar{u}$ determined by the maximal principle

(4.41) $$\langle B(t)\,\bar{u}(t), \bar{\eta}(t)\rangle - h(t, \bar{u}(t)) = \max\,[\langle B(t)\,z, \bar{\eta}(t)\rangle - h(t,z) | z \in \Omega]$$

almost everywhere.

Since $\Omega$ is the effective domain of $h(t,z)$, (4.41) may be written as

$$\langle B(t)\bar{u}(t), \bar{\eta}(t)\rangle - h(t,\bar{u}(t)) = \sup\,[\langle B(t)z, \bar{\eta}(t)\rangle - h(t,z) \mid z \in R^m]$$
$$= h^*(t, B^*(t)\bar{\eta}(t)) \text{ almost everywhere.}$$

From this and from Theorem 2.9 it readily follows that

$$B^*(t)\bar{\eta}(t) \in \partial h(t,\bar{u}(t)) \text{ almost everywhere.}$$

As before we write $\Lambda^*(-\bar{\phi}^*, \bar{\eta}(T)) = (\bar{u}^*, \bar{x}^*)$, where $\bar{u}^*(t) = B^*(t)\bar{\eta}(t)$ and $\bar{x}^* = \bar{\eta}(0)$ according to (4.11), (4.12).

Hence it follows that

$$\Lambda^*(-\bar{\phi}^*, \bar{\eta}(T)) \in \partial f((\bar{u},\bar{x})). \tag{4.42}$$

Furthermore $\bar{\phi}^*(t) \in \partial k(t, \phi(\bar{u})(t))$ and $-\bar{\eta}(T) \in \partial l_1(\phi(\bar{u})(T))$ imply $\phi(\bar{u})(t) \in \partial k^*(t, \bar{\phi}^*(t))$ and $\phi(\bar{u})(T) \in \partial l_1{}^*(-\bar{\eta}(T))$, respectively, by Theorem 2.9. Therefore we must have, since $\Lambda(\bar{u},\bar{x}) = (\phi(\bar{u}), \phi(\bar{u})(T))$, that

$$\Lambda(\bar{u},\bar{x}) \in \partial g^*(-\bar{\phi}^*, \bar{\eta}(T)). \tag{4.43}$$

Hence from (4.42), (4.43) and Theorem 3.7, it follows that $(\bar{u},\bar{x})$ is a solution to $(P)$, that is, $(\bar{u},\bar{x})$ is an optimal pair. ||

*Remark.* If the initial state $x = x_0$ is fixed, we show below that the condition $\bar{\eta}(0) \in \partial l_0(x_0)$ is automatically satisfied.

For fixed $x = x_0, l_0(x_0)$ is a constant and could be left out of the cost function. But to correspond to the formulation of the problem (OC), we define $l_0$ as

$$l_0(x) = \begin{cases} 0, & x = x_0, \\ +\infty, & x \neq x_0, \end{cases}$$

We thus have the same problem formulation as in Theorem 4.2 with $G_0 = \{x_0\}$.

But

$$\bar{\eta}(0) \in \partial l_0(x_0) \text{ if and only if } l_0(x) \geqslant l_0(x_0) + \langle x - x_0, \bar{\eta}(0)\rangle, \quad \forall x \in R^n,$$

and it follows from the definition of $l_0$ that $\partial l(x_0)$ is all of $(R^n)^*$ so that indeed $\bar{\eta}(0) \in \partial l_0(x_0)$ is always satisfied. Similar considerations are useful in the state-constrained problem presented in Section 4.5.

### 4.4. *An Application to a Reachability Problem*

In this section we show that certain problems of reachability can be transformed into convex optimization problems and solved using methods developed in this paper.

For the linear system $(\mathscr{L})$ defined in Section 4.1 the set of admissible controllers will be defined as

$$\mathscr{U} = \{u \in L^r[0,T;R^m] \mid u(t) \in \Omega \subseteq R^m \text{ a.e. on } [0,T], \Omega \text{ is convex and closed}\}.$$

Assume that $G_0 = \{0\}$ and the set $X$ defined by (4.2) is a closed convex set in $L_p[0,T;R^n]$, $1 < p < \infty$.

It is required to find necessary and sufficient conditions for the existence of a $u \in \mathcal{U}$ such that the corresponding response $\phi(u) \in X$, and $\phi(u)(0) = 0$. We shall first transform this problem into a convex optimization problem.

Let $f$ be the proper convex indicator function of the closed convex set $\mathcal{U}$, that is,

$$f(u) = \begin{cases} 0, & \text{if } u \in \mathcal{U}, \\ +\infty, & \text{if } u \neq \mathcal{U}, \end{cases}$$

Then

$$f^*(u^*) = \sup\,[\langle u, u^*\rangle : u \in \mathcal{U}].$$

Let $g$ be the negative indicator function of the non-empty convex set $X$,

$$g(\phi) = \begin{cases} 0, & \text{if } \phi \in X, \\ -\infty, & \text{if } \phi \notin X. \end{cases}$$

$$g^*(\phi^*) = \inf\,[\langle \phi, \phi^*\rangle : \phi \in X].$$

Define the linear continuous map

$$\Lambda\colon\; L^r[0,T;R^m] \to L^p[0,T;R^n]$$

such that

$$(\Lambda u)(t) = \int_0^t \Phi(t,\tau)\,B(\tau)\,u(\tau)\,d\tau,$$

$\forall t \in [0,T]$, where $\Phi$ is the fundamental matrix.

Then the adjoint $\Lambda^*$ of $\Lambda$ is

$$\Lambda^*\colon\; L^q[0,T;(R^n)^*] \to L^s[0,T;(R^m)^*], \quad \frac{1}{p}+\frac{1}{q} = 1, \frac{1}{r}+\frac{1}{s} = 1$$

such that

$$(\Lambda^*\phi^*)(t) = B^*(t)\int_t^T \Phi^*(\tau,t)\,\phi^*(\tau)\,d\tau, \quad \forall t \in [0,T],$$

$$= B^*(t)\,\eta(t), \quad \text{where } \eta(t) = \int_t^T \Phi^*(\tau,t)\,\phi^*(\tau)\,d\tau.$$

The reachability problem can now be written as

$$(P)\ \text{minimize}\, f(u) - g(\Lambda u), \quad u \in L^r[0,T;R^m].$$

and its dual is

$$(P^*)\ \text{maximize}\, g^*(\phi^*) - f^*(\Lambda^*\phi^*), \quad \phi^* \in L^q[0,T;(R^n)^*].$$

Since $\mathscr{U}$ and $X$ are closed, $f$ is l.s.c. and $g$ is u.s.c. and therefore we may apply the general results, as far as they do not rely on the property that the problem is stably set.

From the sufficiency part of Theorem 4.4 we can see that $\bar{u}$ is a solution to $(P)$ and $\bar{\phi}^*$ a solution to $(P^*)$ if there exist a vector $\bar{\eta}(t)$ such that

$$\langle B(t)\,\bar{u}(t), \bar{\eta}(t)\rangle \geqslant \langle B(t)\,z, \bar{\eta}(t)\rangle, \quad \forall z \in \Omega, \text{ almost everywhere on } [0,T]$$

where $\bar{\eta}(t)$ satisfies

$$d\eta/dt = -A^*(t)\,\eta(t) + \bar{\phi}^*(t), \quad \eta(T) = 0$$

and $\bar{\phi}^*(t)$ is a solution of

$$\langle \phi(\bar{u})(t), \bar{\phi}^*(t)\rangle = \min\,[\langle x, \phi^*(t)\rangle | x \in G_t] \text{ almost everywhere on } [0,T]$$

If $0 \in \mathscr{U}$ and $0 \in \operatorname{int}(X)$, then $(P)$ is stably set from Theorem 3.4 and the above conditions are also necessary from Theorem 4.4.

### 4.5. *Example of a State-Constrained Problem*

Consider the following state-constrained optimal control problem.

$$(\mathscr{L}) \quad dx/dt = A(t)\,x(t) + B(t)\,u(t), \quad \phi(u)(0) = c,$$
$$C(u) = \tfrac{1}{2}\|\phi(u)(T)\|^2,$$
$$\mathscr{U} = \{u \in L^2[0,T;R^m] \,|\, \|u(t)\| \leqslant \rho \text{ a.e. on } [0,T], \rho > 0\},$$
$$\mathscr{X} = \{\phi \in L^2[0,T;R^n] \,|\, \phi(0) = c, \|c\| < \beta, \|\phi(t)\| \leqslant \beta \text{ a.e. on } ]0,T], \beta > 0\}.$$

Let $\Omega = \{z \in R^m \,|\, \|z\| \leqslant \rho\}$,

$\mathscr{G} = \{x \in R^n \,|\, \|x\| \leqslant \beta\}$.

The problem is equivalent to minimizing

$$\tfrac{1}{2}\|\phi(u)(T)\|^2 + \psi_c(x) + \int_0^T \{\psi_\Omega(u(t)) + \psi_g(\phi(u)(t))\}\,dt.$$

subject to

$$dx/dt = A(t)\,x(t) + B(t)\,u(t), \quad \phi(u)(0) = c,$$

where $\psi_c$, $\psi_\Omega$, and $\psi_{\mathscr{G}}$ are the indicator functions of the sets $\{c\}$, $\Omega$, and $\mathscr{G}$, respectively.

Since $(0,c) \in \mathscr{U} \oplus \{c\}$, and $(c,c) \in \operatorname{int}(X \oplus R^n)$ the problem is stably set. Therefore we may apply Theorem 4.4 to obtain the necessary and sufficient conditions of optimality. The main computation involves the subgradients of indicator functions of closed balls in $R^n$.

It is known that $\partial\psi_\Omega(z)$ is the normal cone to $\Omega$ at $z$,

$$\partial\psi_\Omega(z) = \{z^* \in (R^m)^*\,\rho\|z^*\| \leqslant \langle z, z^*\rangle\}.$$

and similarly

$$\partial\psi_{\mathscr{G}}(x) = \{x^* \in (R^n)^* | \beta\|x^*\| \leqslant \langle x, x^*\rangle\}.$$

Hence conditions (4.26)–(4.29) read that $\bar{u}$ is an optimal controller with response $\phi(\bar{u})$ if and only if there is a vector $\bar{\eta}(t)$ satisfying

$$d\eta/dt = -A^*(t)\,\eta(t) + x^*(t), -\bar{\eta}(T) = \phi(\bar{u})(T),$$

where $x^*(t)$ is any element of the set $\{x^* \in (R^n)^* | \beta\|x^*\| \leqslant \langle\phi(\bar{u})(t), x^*\rangle\}$, and the maximal principle holds

$$\langle B(t)\,\bar{u}(t), \bar{\eta}(t)\rangle = \max\,[\langle B(t)\,z, \bar{\eta}(t)\rangle : \|z\| \leqslant \rho]$$

$$\langle \bar{u}(t), B^*(t)\,\bar{\eta}(t)\rangle = \rho\|B^*(t)\,\bar{\eta}(t)\|, \text{ a.e.},$$

from which

$$\bar{u}(t) = \rho\frac{B^*(t)\,\bar{\eta}(t)}{\|B^*(t)\,\bar{\eta}(t)\|}, \quad \text{provided } \|B^*(t)\,\bar{\eta}(t)\| \neq 0.$$

## 5. INTERCONNECTED CONTROL PROCESSES

In practice one frequently encounters control processes of large dimension which, however, have the structure of mutually interconnected subprocesses of smaller dimension. In this section we consider a linear interconnected optimal control process with a cost function of separable type and show how the duality theory developed in this paper is of help in finding a symmetric subprocess structure of the interconnected process and its dual. In particular we obtain a decomposition theorem which relates optimal controllers to solutions of the individual dual subprocesses. For related work see [35, 3, 36] and the bibliography of these papers. ||

### 5.1. *Formulation of the Problem*

Consider the linear interconnected control process

$$(\mathscr{L})\ \begin{cases} dx_i/dt = A_i(t)\,x_i(t) + B_i(t)\,u_i(t) + v_i(t), \\ x_i(0) = 0, \end{cases} \qquad i = 1, 2, \ldots, N,$$

$$(\mathscr{I})\ v_i(t) = \sum_{\substack{j=1\\ j\neq i}}^{N} F_{ij}(t)\,x_j(t), \qquad i = 1, 2, \ldots, N,$$

where $v_i, x_i$ are elements of $L^p[0, T; R^n]$, $1 < p \leqslant \infty$, $u_i \in L^r[0, T; R^{m_i}]$, $1 < r \leqslant \infty$ and $A_i(t)$, $B_i(t)$, $F_{ij}(t)$ are continuous matrices of appropriate order $i, j = 1, 2, \ldots, N$.

Let $u = (u_1, \ldots, u_N)$.

The responses $\phi_i(u)$ of $(\mathscr{L})$ are constrained to lie in the convex subsets $X_i$ of $L^p[0, T; R^{n_i}]$ where

(5.1) $X_i = \{\phi_i \in L^p[0, T; R^{n_i}] | \phi_i(0) = 0, \phi_i(t) \in G_i(t) \subseteq R^{n_i}$ a.e. in $]0, T]$, $G_i(t)$ is convex$\}$, $\quad i = 1, 2, \ldots, N$.

Let $\mathscr{U}_i$ be the convex subset in $L^r[0,T;R^{m_i}]$ defined by

(5.2) $\mathscr{U}_i = \{u_i \in L^r[0,T;R^{m_i}] | u_i(t) \in \Omega_i \subseteq R^{m_i} \text{ a.e. in } [0,T], \Omega_i \text{ is convex}\},$
$i = 1, 2, \ldots, N.$

The class $\mathscr{U}$ of admissible controllers $u$ is defined as

(5.3) $\mathscr{U} = \{u \in L^r[0,T;R^m] | u_i \in \mathscr{U}_i, \text{ responses } \phi_i(u) \in X_i, \quad i = 1, 2, \ldots, N\},$

where $m = \sum_{i=1}^{N} m_i$.

It is assumed that $\Omega_i$ and $G_i(t)$, $t \in ]0,T]$ have non-empty interiors, $i = 1, 2, \ldots, N$.

The cost function is

$$C(u) = \sum_{i=1}^{N} \int_0^T [h_i(t, u_i(t)) + k_i(t, \phi_i(u)(t))]\, dt. \tag{5.4}$$

The functions $h_i$ and $k_i$ satisfy hypothesis (ii), (iii), and (iv) of Section 4.1. By defining

$$x(t) = (x_1(t) \cdots x_N(t))^T, \qquad u(t) = (u_1(t) \cdots u_N(t))^T,$$

$$A(t) = \begin{pmatrix} A_1(t)\, F_{12}(t) \cdots F_{1N}(t) \\ F_{21}(t)\, A_2(t) \cdots F_{2N}(t) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ F_{N1}(t)\, F_{N2}(t) \cdots A_N(t) \end{pmatrix}, \qquad B(t) = \begin{pmatrix} B_1(t) & & 0 \\ & \ddots & \\ 0 & & B_N(t) \end{pmatrix},$$

$(\mathscr{L})$ may be written as

$$\begin{cases} dx/dt = A(t)\, x(t) + B(t)\, u(t), \\ x(0) = 0. \end{cases} \tag{5.5}$$

It follows from the hypotheses that the functions $f_i$, $i = 1, 2, \ldots, N$

$$f_i(u_i) = \int_0^T h_i(t, u_i(t))\, dt \tag{5.6}$$

are l.s.c. proper convex functions on $L^r[0,T;R^{m_i}]$ with effective domains $\mathscr{U}_i$ and conjugate functions

$$f_i^*(u_i^*) = \int_0^T h_i^*(t, u_i^*(t))\, dt \tag{5.7}$$

on

$$L^s[0,T;(R^{m_i})^*], \; \frac{1}{r} + \frac{1}{s} = 1.$$

In particular $f_i^*$ has a non-empty effective domain which we denote by $\mathscr{U}_i^*$.

Let $f$ be defined by

$$f\colon L^r[0,T;R^{m_1}] \oplus \cdots \oplus L_r[0,T;R^{m_N}] \to \mathscr{R}$$

such that

$$(5.8)\quad f(u) = (f_1 \oplus \cdots \oplus f_N)(u) = f_1(u_1) + \cdots + f_N(u_N), \quad \forall u \in L^r[0,T;R^m].$$

It is clear that $f$ is a l.s.c. proper convex function with effective domain $\mathscr{U} = \mathscr{U}_1 \oplus \cdots \oplus \mathscr{U}_N$ and conjugate $f^*$ where

$$(5.9)\qquad f^*(u^*) = (f_1^* \oplus \cdots \oplus f_N^*)(u^*), \quad \forall u^* \in L^s[0,T;(R^m)^*].$$

$f^*$ has effective domain $\mathscr{U}^* = \mathscr{U}_1^* \oplus \cdots \oplus \mathscr{U}_N^*$.

Also from the hypothesis it follows that the functions $g_i$, $i = 1, \ldots, N$,

$$(5.10)\qquad g_i(\phi_i) = -\int_0^T k_i(t, \phi_i(t))\, dt$$

are u.s.c. proper concave functions on $L^p[0,T;R^{n_i}]$ with effective domains $X_i$ and conjugate functions

$$g_i^*(\phi_i^*) = -\int_0^T k^*(t, -\phi_i^*(t))\, dt$$

on

$$L^q[0,T;(R^{n_i})^*], \quad \frac{1}{p} + \frac{1}{q} = 1.$$

In particular $g_i^*$ has non-empty effective domain which we denote by $X_i^*$.

Let $g$ be the function defined by

$$g\colon L^p[0,T;R^{n_1}] \oplus \cdots \oplus L^p[0,T;R^{n_N}] \to \mathscr{R}$$

such that

$$(5.11)\qquad g(\phi) = (g_1 \oplus \cdots \oplus g_N)(\phi) = g_1(\phi_1) + \cdots + g_N(\phi_N),$$

$$\forall \phi \in L^p[0,T;R^n], n = \sum_{i=1}^N n_i.$$

Then $g$ is an u.s.c. proper concave function with effective domain $X = X_1 \oplus \cdots \oplus X_N$ and its conjugate $g^*$ given by

$$(5.12)\quad g^*(\phi^*) = (g_1^* \oplus \cdots \oplus g_N^*)(\phi^*), \quad \forall \phi^* \in L^q[0,T;(R^n)^*],$$

with effective domain $X^* = X_1^* \oplus \cdots \oplus X_N^*$. Finally define the map $\Lambda$ by,

$$\Lambda\colon L^r[0,T;R^{m_1}] \oplus \cdots \oplus L^r[0,T;R^{m_N}] \to L^p[0,T;R^{n_1}] \oplus \cdots \oplus L^p[0,T;R^{n_N}]$$

such that

$$(5.13)\qquad \Lambda u = \phi(y), \quad \forall u,$$

and $\phi(u)$ is defined by

(5.14) $$\phi(u)(t) = \int_0^t \Phi(t,\tau)\, B(\tau)\, u(\tau)\, dt, \quad t \in [0,T],$$

where $\Phi$ is the fundamental matrix of (5.5). Thus $\Phi$ satisfies

$$d\Phi/dt = A(t)\,\Phi, \quad \Phi(0) = I,$$

which is a system of $N \times N$ matrix differential equations

(5.15) $$\begin{cases} d\Phi_{11}/dt = A_1(t)\,\Phi_{11} + F_{12}(t)\,\Phi_{21} + \cdots + F_{1N}(t)\,\Phi_{N1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ d\Phi_{NN}/dt = F_{N1}(t)\,\Phi_{1N} + F_{N2}(t)\,\Phi_{2N} + \cdots + A_N(t)\,\Phi_{NN}, \end{cases}$$

$$\Phi_{ij}(0) = \begin{cases} I, & i = j, \\ 0, & i \neq j. \end{cases}$$

The interconnected control process $(\mathscr{L})$ and $(\mathscr{I})$ in terms of $f$, $g$, $\Lambda$ defined by (5.8), (5.11) and (5.13) (5.14) becomes

$$(P) \quad \text{minimize } f(u) - g(\Lambda u), \text{ subject to } u \in \mathscr{U}, \Lambda u \in X,$$

and consequently the dual $(P^*)$ is

$$(P^*) \quad \text{maximize } g^*(\phi^*) - f^*(\Lambda^* \phi^*), \text{ subject to } \phi^* \in X^*, \Lambda^* \phi^* \in \mathscr{U}^*$$

where $f^*$, $g^*$ are defined by (5.9), (5.12) and $\Lambda^*$ is the adjoint of $\Lambda$.

Thus the general duality theory of Section 3 can be applied to the interconnected control problem. ||

We shall write $(P)$ and $(P^*)$ in a form which emphasizes the subproblem structure of both $(P)$ and $(P^*)$.

For $u \in L^r[0,T;R^m]$, $\Lambda u = (\phi_1, \ldots, \phi_N)$, define the maps $\Lambda_i$,

$$\Lambda_i\colon\ L^r[0,T;R^m] \to L^p[0,T;R^{n_i}]$$

such that

(5.16) $$\Lambda_i u = \phi_i, \quad i = 1,2,\ldots, N;$$

similarly for

$$\Lambda^* \phi^* = (u_1^*, \ldots, u_N^*), \quad \phi^* \in L^q[0,T;(R^n)^*],$$

define

$$\Lambda_i^*\colon\ L^q[0,T;(R^n)^*] \to L^s[0,T;(R^{m_i})^*]$$

such that

(5.17) $$\Lambda_i^* \phi^* = u_i^*, \quad i = 1,2,\ldots, N.$$

Note that $\Lambda_i^*$ is not the adjoint of $\Lambda_i$. From the above definitions, it follows that

(5.18) $$\langle \Lambda u, \phi^* \rangle = \sum_{i=1}^{N} [\langle \Lambda_i u, \phi_i^* \rangle]; \quad \langle u, \Lambda^* \phi^* \rangle = \sum_{i=1}^{N} [\langle u_i, \Lambda_i^* \phi^* \rangle].$$

We can write $(P)$ as

$$(P)\quad \inf_{u\in\mathscr{U}}\{f(u)-g(\Lambda u)\} = \inf_{u\in\mathscr{U}}[f(u) - \inf_{\phi^*\in X^*}\{\langle \Lambda u,\phi^*\rangle - g^*(\phi^*)\}]$$
$$= \inf_{u\in\mathscr{U}}[f(u) + \sup_{\phi^*\in X^*}\{g^*(\phi^*) - \langle \Lambda u,\phi^*\rangle\}].$$

Or, in view of (5.12) and (5.18),

$$(5.19)\qquad (P)\quad \inf_{u\in\mathscr{U}}[f(u) + \sum_{i=1}^{N}\sup_{\phi_i^*\in X_i^*}\{g_i^*(\phi_i^*) - \langle \Lambda_i u,\phi_i^*\rangle\}].$$

Similarly we can write $(P^*)$ in the form

$$(5.20)\qquad (P^*)\quad \sup_{\phi^*\in X^*}[g^*(\phi^*) + \sum_{i=1}^{N}\inf_{u_i\in\mathscr{U}_i}\{f_i(u_i) - \langle u_i, \Lambda_i^*\phi^*\rangle\}].$$

For fixed $u$, $(P)$ has $N$ separated subproblems $(P_i)$ of the form

$$(5.21)\qquad (P_i)\quad \sup_{\phi_i^*\in X_i^*}\{g_i^*(\phi_i^*) - \langle \Lambda_i u,\phi_i^*\rangle\},\quad i=1,2,\ldots,N,$$

where, for fixed $\phi^*$, the dual $(P^*)$ has $N$ separated subproblems $(P_i^*)$ of the form

$$(5.22)\qquad (P_i^*)\quad \inf_{u_i\in\mathscr{U}_i}\{f_i(u_i) - \langle u_i, \Lambda_i^*\phi^*\rangle\},\quad i=1,2,\ldots,N.$$

Note that $(P)$ as well as $(P^*)$ has the same subproblem structure. ‖

Before we can relate the solutions of the subproblems $(P_i)$, $(P_i^*)$ to those of the overall problems $(P)$ and $(P^*)$ we need some results.

Introduce the function $K(u,\phi^*)$ where

$$(5.23)\qquad K(u,\phi^*) = f(u) + g^*(\phi^*) - \langle \Lambda u,\phi^*\rangle.$$

Then, $K$ is a l.s.c. proper convex function of $u$, with effective domain $\mathscr{U}$ and an u.s.c. proper concave of $\phi^*$ with effective domain $X^*$.

A point $(\bar{u},\bar{\phi}^*)$, $\bar{u}\in\mathscr{U}$, $\bar{\phi}^*\in X^*$ is a saddle-point of $K(u,\phi^*)$ if

$$K(\bar{u},\bar{\phi}^*) = \min_{u\in\mathscr{U}} K(u,\bar{\phi}^*) = \max_{\phi^*\in X^*} K(\bar{u},\phi^*).$$

THEOREM 5.1. *The function $K(u,\phi^*)$ defined in* (5.23) *has a saddle-point $(\bar{u},\bar{\phi}^*)$ if and only if $(P)$ and $(P^*)$ are stably set in which case $\bar{u}$ is a solution (optimal controller) of $(P)$ and $\bar{\phi}^*$ is a solution of $(P^*)$.*

*Proof.* See [16]. ‖

We can now prove the decomposition theorem for the interconnected control problem $P$ and its dual $P^*$.

THEOREM 5.2. *Assume $(P)$ and $(P^*)$ are both stably set. If for fixed $\bar{\phi}^*$, $\bar{u}_i$ solves the subproblem* (5.22), *$i=1,\ldots,N$ of $(P^*)$ then $\bar{u}=(\bar{u}_1,\ldots,\bar{u}_N)$ is a solution* *to (P) if and only if $\bar{\phi}^*$ is a solution to $(P^*)$. Dually, if for fixed $\bar{u}$, $\bar{\phi}_i^*$ solves the subproblem* (5.21), $i = 1, \ldots, N$, *of (P) then $\bar{\phi}^* = (\bar{\phi}_1^*, \ldots, \bar{\phi}_N^*)$ is a solution of $(P^*)$ if and only if $\bar{u}$ is a solution of (P).*

*Proof.* Let for some $\bar{\phi}^* \in X^*$, $\bar{u}_i$ be a solution to $(P_i^*)$, $i = 1, \ldots, N$.

Because $\bar{u}_i$ solves $(P_i^*)$ it readily follows that

$$-f_i^*(\Lambda_i^* \bar{\phi}^*) = f_i(\bar{u}_i) - \langle \bar{u}_i, \Lambda_i^* \bar{\phi}^* \rangle, \quad i = 1, \ldots, N, \tag{5.24}$$

and thus

$$-f^*(\Lambda^* \bar{\phi}^*) = f(\bar{u}) - \langle \bar{u}, \Lambda^* \bar{\phi}^* \rangle. \tag{5.25}$$

Therefore,

$$\begin{aligned} K(\bar{u}, \bar{\phi}^*) &= g^*(\bar{\phi}^*) + f(\bar{u}) - \langle \Lambda \bar{u}, \bar{\phi}^* \rangle \\ &= g^*(\bar{\phi}^*) - f^*(\Lambda^* \bar{\phi}^*) \leqslant \sup_{\phi^* \in X^*} K(\bar{u}, \phi^*). \end{aligned} \tag{5.26}$$

Assume $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_N)$ is a solution to $(P)$. Since $(P)$ is stably set and has a solution $\bar{u}$, it follows from Theorem 3.6 that $(P^*)$ has a solution and thus, by Theorem 5.1, we must have

$$\min(P) = \max_{\phi^* \in X^*} K(\bar{u}, \phi^*). \tag{5.27}$$

Hence, from (5.26) and (5.27),

$$g^*(\bar{\phi}^*) - f^*(\Lambda^* \bar{\phi}^*) \leqslant \min(P).$$

But, by Lemma 3.1,

$$\min(P) \geqslant g^*(\bar{\phi}^*) - f^*(\Lambda^* \bar{\phi}^*),$$

so that

$$\min(P) = g^*(\bar{\phi}^*) - f^*(\Lambda^* \bar{\phi}^*),$$

which implies that $\bar{\phi}^*$ is a solution of $(P^*)$.

On the other hand, assume $\bar{\phi}^*$ is a solution of $(P^*)$. Then,

$$\max(P^*) = g^*(\bar{\phi}^*) - f^*(\Lambda^* \bar{\phi}^*).$$

Or in view of (5.25), (5.26)

$$\max(P^*) = g^*(\bar{\phi}^*) + f(\bar{u}) - \langle \Lambda \bar{u}, \bar{\phi}^* \rangle = K(\bar{u}, \bar{\phi}^*).$$

But $(P^*)$ is stably set and has a solution implies $(P)$ has a solution, by Theorem 3.5, such that

$$\min(P) = \max(P^*).$$

That is (see (5.19) and (5.20)),

$$\min_{u \in \mathcal{U}} \max_{\phi^* \in X^*} K(u, \phi^*) = \max_{\phi^* \in X^*} \min_{u \in \mathcal{U}} K(u, \phi^*)$$

and in view of (5.26)

$$\min_{u \in \mathcal{U}} \max_{\phi^* \in X^*} K(u, \phi^*) = \max_{\phi^* \in X^*} \min_{u \in \mathcal{U}} K(u, \phi^*) = K(\bar{u}, \bar{\phi}^*).$$

Hence, $(\bar{u},\bar{\phi}^*)$ is a saddle-point of $K(u,\phi^*)$, implying that $\bar{u}$ is a solution of $(P)$ by Theorem 5.1. This proves the first part of the theorem and the second part follows dually. ‖

The following theorem guarantees that the subproblems $(P_i)$ and $(P_i^*)$ in (5.21) and (5.22) are stably set and do have solutions.

THEOREM 5.3. *Given the interconnected control Problem* $(P)$ *and its dual* $(P^*)$. *Assume that, for* $i = 1, \ldots, N$,

(i) *the effective domain* $\Omega_i$ *of* $h_i(t,z_i)$ *is bounded,*
(ii) $h_i(t,z_i) \geqslant H_i(t)$, $\forall z_i \in \Omega_i$, *where* $H_i$ *is a summable function on* $[0,T]$,
(iii) *the effective domain* $G_i^*(t)$ *of* $k_i^*(t,y_i^*)$ *is bounded,*
(iv) $k_i^*(t,y_i^*) \geqslant K_i(t)$, $\forall y_i^* \in G_i^*(t)$, *where* $K_i$ *is a summable function on* $[0,T]$.

*Then, the subproblems* $(P_i)$ *in* (5.21) *and* $(P_i^*)$ *in* (5.22) *are stably set and have solutions.*

*Proof.*

$$(P_i) \qquad \sup_{\phi_i^* \in X_i^*} \{g_i^*(\phi_i^*) - \langle \Lambda_i u, \phi_i^* \rangle\}, \text{ for some } u,$$

has a dual process

$$\inf_{\phi_i} \{l_i(\phi_i) - g_i(\phi_i)\},$$

where

$$l_i(\phi_i) = \sup_{\phi_i^*} \{\langle \phi_i, \phi_i^* \rangle - \langle \Lambda_i u, \phi_i^* \rangle\}$$
$$= \sup_{\phi_i^*} \langle \phi_i - \Lambda_i u, \phi_i^* \rangle = \begin{cases} 0, & \phi_i = \Lambda_i u, \\ +\infty, & \phi_i \neq \Lambda_i u. \end{cases}$$

Further, according to (5.10)

$$g_i(\phi_i) = -\int_0^T k(t, \phi_i(t))\, dt$$
$$= -\int_0^T \left\{ \sup_{y_i^* \in G_i^*} \langle \phi_i(t), y_i^* \rangle - k^*(t, y_i^*) \right\} dt$$
$$= \int_0^T \left\{ \inf_{y_i^* \in G_i^*} \langle -\phi_i(t), y_i^* \rangle + k^*(t, y_i^*(t) \right\} dt$$
$$\geqslant \int_0^T \left\{ \inf_{y_i^* \in G_i^*} \langle -\phi_i(t), y_i \rangle + K_i(t) \right\} dt > -\infty.$$

We also know that $g_i$ is a proper concave function so that $g_i$ is finite on all of $L^p[0,T;R^{n_i}]$. Thus the dual problem of $(P_i)$ is stably set by Theorem 3.4. Clearly $\Lambda_i u$ is also the solution to this dual problem from which it follows by Theorem 3.6 that $(P_i)$ is stably set and has a solution.

$$(P_i^*) \quad \inf_{u_i \in \mathscr{U}_i} \{f_i(u_i) - \langle u_i, \Lambda_i^* \phi^* \rangle\}, \quad \text{for some } \phi^*,$$

has a dual,

$$\sup_{u_i^*} \{p_i^*(u_i^*) - f_i^*(u_i^*)\},$$

where

$$\begin{aligned} p_i^*(u_i^*) &= \inf_{u_i} \{\langle u_i, u_i^* \rangle - \langle u_i, \Lambda_i^* \phi^* \rangle\} \\ &= \inf_{u_i} \langle u_i, u_i^* - \Lambda_i^* \phi^* \rangle = \begin{cases} 0, & u_i = \Lambda_i^* \phi^*, \\ -\infty, & u_i \neq \Lambda_i^* \phi^*. \end{cases} \end{aligned}$$

From (5.6) we have

$$\begin{aligned} f_i^*(u_i^*) &= \int_0^T h_i^*(t, u_i^*(t))\, dt \\ &= \int_0^T \left\{ \sup_{z_i \in \Omega_i} \langle z_i, u_i^*(t) \rangle - h_i(t, z_i) \right\} dt \\ &\leqslant \int_0^T \left\{ \sup_{z_i \in \Omega_i} \langle z_i, u_i^*(t) \rangle - H_i(t) \right\} dt < +\infty. \end{aligned}$$

Because $f_i^*$ is also proper it follows that $f_i^*$ is finite on all of $L^s[0,T;(R^{m_i})^*]$. It follows by Theorem 3.4 that the dual of $(P_i^*)$ is stably set. Furthermore it has $\Lambda_i^* \phi^*$ as a solution so that $(P_i^*)$ is stably set and has a solution by Theorem 3.6. ‖

## ACKNOWLEDGMENT

The authors are grateful to R. T. Rockafellar for making available preprints of some of his papers.

## REFERENCES

1 R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. I, Wiley (Interscience), New York, 1953.

2 J. D. Pearson, Reciprocity and duality in control programming problems, *J. Math. Anal. Appl.*, **10** (1965), 383–408.

3 J. D. Pearson, Duality and a decomposition technique, *SIAM J. Control*, **4** (1966), 164–172.
4 B. Mond and M. Hanson, Duality for control problems, *SIAM J. Control*, **6** (1968), 114–120.
5 R. M. Van Slyke and R. Wets, A duality theory for abstract mathematical programs with applications to optimal control theory, *J. Math. Anal. Appl.*, **22** (1968), 679–706.
6 W. Fenchel, On conjugate convex functions, *Canad. J. Math.*, **1** (1949), 73–77.
7 W. Fenchel, *Convex Cones, Sets and Functions*, Lecture Notes, Princeton University, Spring 1951.
8 A. Bröndsted, Conjugate convex functions in topological vector spaces, *Mat.-Fys. Medd. Danske Vid. Selsk.*, **34** (1964), 1–27.
9 J. J. Moreau, Proximité et dualité dans un éspace Hilbertien, *Bull. Soc. Math. France*, **93** (1965), 273–299.
10 J. J. Moreau, *Fonctionelles Convexes*, Lecture Notes, Séminaire College de France, 1966.
11 R. T. Rockafellar, *Convex Analysis*, Lecture Notes, Princeton University, Spring 1966.
12 R. T. Rockafellar, Extension of Fenchel's duality theory for convex functions, *Duke Math. J.*, **33** (1966), 81–90.
13 R. T. Rockafellar, Conjugate and Legendre transforms of convex functions, *Canad. J. Math.*, **19** (1967), 200–205.
14 R. T. Rockafellar, Level sets and continuity of conjugate convex functions, *Trans. Amer. Math. Soc.*, **123** (1966), 46–63.
15 R. T. Rockafellar, *Lectures on Duality and Convex Functions*, International Summer School on Mathematical Systems Theory and Mathematical Economics, Varenna, Italy, 1967.
16 R. T. Rockafellar, Duality and stability in extremum problems involving convex functions, *Pacific J. Math.*, **21** (1967), 167–187.
17 U. Dieter, Optimierungsaufgaben in topologischen Vektorraumen, I: Dualitätstheorie, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **5** (1966), 89–117.
18 J. J. Moreau, Fonctionelles sous-differentiables, *C. R. Acad. Sci. Paris*, **257** (1963), 4117–4119.
19 R. T. Rockafellar, Characterization of the subdifferentials of convex functions, *Pacific J. Math.*, **17** (1966), 497–510.
20 A. Bröndsted and R. T. Rockafellar, On the subdifferentiability of convex functions, *Proc. Amer. Math. Soc.*, **16** (1965), 605–611.
21 E. Asplund, Frechet differentiability of convex functions, *Acta Math.*, **121** (1968), 31–47.
22 R. V. Gamkrelidze, On some extremal problems in the theory of differential equations with applications to the theory of optimal control, *SIAM J. Control*, **3** (1965), 106–128.
23 R. V. Gamkrelidze and G. L. Kharatishvilli, Extremal problems in linear topological spaces, *Math. Systems Theory*, **1** (1967), 229–256.
24 H. Halkin, *Non-Linear Non-Convex Programming In Finite Dimensional Spaces, in Mathematical Theory of Control* (A. V. Balakrishnan and L. W. Neustadt, eds.), Academic Press, New York, 1967.
25 H. Halkin and L. W. Neustadt, General necessary conditions for optimization problems, *Proc. Natl. Acad. Sci. U.S.A.*, **56** (1966), 1066–1071.
26 L. W. Neustadt, An abstract variational theory with applications to a broad class of optimization problems, I. General theory, *SIAM J. Control*, **4** (1966), 533–535.
27 L. W. Neustadt, An abstract variational theory with applications to a broad class of optimization problems, II. Applications, *SIAM J. Control*, **5** (1967), 90–137.
28 L. W. Neustadt, A general theory of extremals, *J. Computer Systems Sci.*, **3** (1969), 57–92.

29 B. N. Pshenichnyi, Linear optimal control problems, *SIAM J. Control*, **4** (1966), 577–593.
30 V. F. Demyanov and A. M. Rubinov, Minimization of functionals in normed spaces, *SIAM J. Control*, **6** (1968), 73–88.
31 G. P. Akhilov, L. V. Kantorovich, and G. SH-Rubinshtein, Extremal states and extremal controls, *SIAM J. Control*, **5** (1967), 600–608.
32 J. M. Danskin, *The Theory of Max-Min and Its Application to Weapons Allocation Problems*, Springer-Verlag, Berlin, New York, 1967.
33 R. M. Van Slyke and R. Wets, Programming under uncertainty and stochastic optimal control, *SIAM J. Control*, **4** (1966), 179–193.
34 R. T. Rockafellar, Integrals which are convex functionals, *Pacific J. Math.*, **24** (1968), 525–539.
35 L. S. Lasdon, *Duality and Decomposition in Mathematical Programming*, Case Western Reserve University, Systems Research Center Report No. 119-C-67-52, 1967.
36 M. D. Mesarovic, D. Macko, and Y. Takahara, *A Theory of Multi-Level Systems*, book to be published by Academic Press.
37 E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*, Wiley, New York, 1967.

*Received September 30, 1969*