JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 13, 251-264 (1966)

A Note on the Minimum Effort Control Problem*

W. A. PORTER AND J. P. WILLIAMS

Department of Electrical Engineering and Institute of Science and Technology, The University of Michigan, Ann Arbor, Michigan

Submitted by Lotfi Zadeh

1. INTRODUCTION

By a *continuous linear system* we shall mean a system with input u and output x, governed for $t \geqslant t_0$ by the system of integral equations

$$x(t) = \Phi(t, t_0)\, x^0 + \Phi(t, t_0) \int_{t_0}^{t} \Phi(t_0, s)\, B(s)\, u(s)\, ds. \tag{1}$$

Here $u(t)$ and $x(t)$ are (real or complex) vector functions with m and n components respectively and $\Phi(t, s)$ denotes the system transition matrix. In [1] Neustadt studied various "effort" functions $\epsilon(u)$ associated with such a system. In particular he showed that if the time T is fixed and effort is defined by

$$\epsilon(u) = \left(\int_0^T \sum_{j=1}^{m} | u_j(t) |^p \, dt\right)^{1/p} \qquad 1 < p < \infty$$

then to each target state $x(T)$ there corresponds a unique minimum effort control $u^*(t)$ which transfers x from x^0 to $x(T)$ in time T. The precise value $\epsilon(u^*)$ of the minimum effort was computed as well as the explicit form of the control vector $u^*(t)$.

In this note we will formulate and solve a generalization of Neustadt's problem. The result yields the existence and uniqueness of a minimum effort control and its precise form for a wide class of effort functions and includes the cases of discrete and composite (discrete-continuous) linear systems.

* The sponsorship of this research was provided by the National Science Foundation under Contract Number GP-624 and U.S. Air Force Contract AF-33(657)-11501.

2. The Minimum Effort Problem

To motivate what follows, let us consider the system of Eq. (1). For convenience we suppose the system is initially at rest so that $x^0 = 0$. Let B denote the cartesian product

$$L_p(\tau) \times L_p(\tau) \times \cdots \times L_p(\tau), \qquad \tau = [t_0, T], \qquad 1 < p < \infty$$

where $L_p(\tau)$ consists, as usual, of those complex valued Lebesgue measurable functions on τ whose pth power is integrable. Then to each $u \in B$ there corresponds a unique x satisfying Equation (1). In particular, at time T we have

$$x(T) = \Phi(T, t_0) \int_{t_0}^{T} \Phi(t_0, s)\, B(s)\, u(s)\, ds.$$

With K denoting either the real or complex numbers, this leads us to define a transformation S from B to K^n by writing $Su = x(T)$. It is easy to verify that S is linear. Moreover, with any choice of product norms on B and K^n, S is bounded. Since it is clear that

$$\| u \| = \left(\int_{t_0}^{T} \sum_{i=1}^{m} | u_i(t) |^p \, dt \right)^{1/p} \qquad 1 < p < \infty$$

defines a norm on B, we see that a natural generalization of the control problem of Neustadt is the following.

Problem. Let B and R be Banach spaces and T a bounded linear transformation from B into R. For each ξ in the range of T find an element $u \in B$ satisfying $Tu = \xi$ which minimizes $\| u \|$.

Consider the set $T^{-1}(\xi)$ of all pre-images of ξ under T. The solution to the general minimum effort problem must then answer the following questions: Does $T^{-1}(\xi)$ contain an element of minimum norm? If so, is this element unique? Finally, if both these answers are yes, and if we write $T^{\dagger}\xi$ for the unique minimum pre-image of ξ under T, what is the nature of the function $T^{\dagger}$ so defined, and more specifically, how can one compute its values?

Initially, we allow B to be an arbitrary (real or complex) Banach space. After having answered the first two questions we will see the need of requiring two additional properties of B (namely reflexivity and rotundity) to insure the existence of the minimum energy function $T^{\dagger}$ associated with T. For convenience in studying $T^{\dagger}$ we will then impose a third restriction on B (smoothness). As regards T, we require that it be *onto* R. This amounts to assuming that T has a closed range and hence in particular, if T has a finite dimensional range, results in no loss of generality.

We begin with two examples which show that some additional restriction on B is needed.

EXAMPLE 1. Let C denote the set of all real (or complex) valued continuous functions or the interval $0 \leqslant t \leqslant 1$ which vanish at $t = 0$. Then C is a closed subspace of the usual Banach space of continuous functions on [0, 1], and hence is a Banach space. Let T be the bounded linear transformation from C to K defined by

$$Tu = \int_0^1 u(t)\,dt.$$

Then it is easy to see that

(1) $\inf\{\| u \| : Tu = 1\} = 1$.

(2) $| Tu | < 1$ if $u \in C$ has norm 1.

It follows that the vector (number) 1 does not have a minimum pre-image under T.

EXAMPLE 2. Let D denote the plane equipped with the norm

$$\| x \| = | x_1 | + | x_2 | \qquad \text{if} \qquad x = (x_1, x_2).$$

On D we define the linear transformation T by

$$Tx = x_1 + x_2 .$$

It is obvious that $\| T \| = 1$ and hence that any $x \in D$ satisfying $Tx = 1$ has norm $\geqslant 1$. It follows that *both* of the vectors (0, 1) and (1, 0) are minimum pre-images of 1 under T.

In short, the minimum effort function $T^{\dagger}$ associated with T can fail to exist by virtue of either a lack of or an overabundance of minimum pre-images. It is worth observing that the space C above is not reflexive and the space D has a "flat" unit ball (connect the points (0, 1), (1, 0), (− 1, 0), (0, − 1)). We now proceed to remedy both these defects in B.

DEFINITION. Let $U = \{x : \| x \| \leqslant 1]$ be the unit ball in B and ∂U the boundary of U. B is called *rotund* [2] or *strictly convex* [3] if one of the following equivalent conditions satisfied:

(1) ∂U contains no line segments.

(2) $\| x_1 + x_2 \| = \| x_1 \| + \| x_2 \|$ implies $x_2 = \lambda x_1$ or $x_1 = \lambda x_2$ for some $\lambda \geqslant 0$.[1]

[1] Observe that it follows from (2) that rotundity is preserved by any linear isometry.

(3) For each bounded linear functional φ on B there is at most one $x \in U$ with $\langle x, \varphi \rangle = \varphi(x) = \| \varphi \|$.

(4) Each convex subset C of B has at most one minimum element (i.e., there is at most one vector $x \in C$ satisfying $\| x \| \leqslant \| z \|$ for all $z \in C$.

The following lemma lists some examples of rotund Banach spaces.

LEMMA 1. (1) *Any Hilbert space is rotund.*

(2) *The spaces l_p, L_p are rotund for $1 < p < \infty$.*

(3) *If $B_1, \cdots, B_n$ are rotund Banach spaces, then so is*

$$B = B_1 \times B_2 \times \cdots \times B_n$$

when the norm of $x = (x_1, x_2, \cdots, x_n)$ in B is defined by either of

$$\| x \| = \Big(\sum_i \| x_i \|^p \Big)^{1/p} \qquad 1 < p < \infty$$

$$\| x \| = \Big(\sum_{i,j} a_{ij} \| x_i \| \, \| x_j \| \Big)^{1/2}$$

where $[a_{ij}]$ is a strictly positive $n \times n$ matrix each of whose entries is nonnegative.

PROOF. The first assertion follows immediately from (2) of the above definition and the parallelogram law. The second is well-known and may be found in [4; p. 211] for example. The proof of (3) is straightforward but somewhat detailed and hence will be omitted.

Observe now that because T is linear and continuous the set $T^{-1}(\xi)$ is convex and closed for each $\xi \in R$. The following theorem therefore gives necessary and sufficient conditions on B for our first two questions to be answered affirmatively for *every* T on B.

THEOREM 1. *Let B be a Banach space. Then each closed convex set C in B has at least one (at most one) minimum element if and only if B is reflexive (rotund).*

PROOF. Property (4) above establishes half of the theorem. Suppose then that B is reflexive. Then U is weakly compact and consequently, if

$$\alpha = \inf \{ \| z \| : z \in C \},$$

the sets

$$C_n = \{ z \in C : \| z \| \leqslant \alpha + 1/n \} \qquad (n = 1, 2, \cdots)$$

form a decreasing sequence of non-empty, weakly compact subsets of B and therefore have nonempty intersection. The fact that reflexivity of B is also necessary was recently shown by Phelps [5].

Henceforth we assume that B is reflexive and rotund and focus attention on the function $T^{\dagger}$.

3. The Minimum Effort Function

We begin by examining a special case.

THEOREM 2. *If $B = H$ is a Hilbert space, N is the null space of T and $M = N^{\perp}$, then $T^{\dagger} = T_M^{-1}$ is the inverse of the restriction of T to M.*

PROOF. The transformation T_M is $1 - 1$, continuous, and onto the Banach space R and hence, by the Closed Graph Theorem, is invertible. Let ξ be a fixed vector in R and write $u_\xi = T_M^{-1}(\xi)$. If $u \in H$ is any pre-image of ξ, then

$$u = (u - u_\xi) + u_\xi$$

is the unique decomposition of u in $N \oplus M$ and hence

$$\| u \|^2 = \| u - u_\xi \|^2 + \| u_\xi \|^2 \geqslant \| u_\xi \|^2$$

The result follows from the definition of $T^{\dagger}$.

It is clear that the proof and even the statement of Theorem 2 makes no sense in B. As a matter of fact, it turns out that the function $T^{\dagger}$ will not in general be linear, and different techniques are necessary.

If E is a Banach space then the Hahn-Banach theorem shows that to each non-zero x in E there corresponds at least one $\varphi \in E^*$ such that

$$\| \varphi \| = 1, \qquad \langle x, \varphi \rangle = \| x \|$$

If E is reflexive this result applied to E^* shows that to each $\varphi \neq 0$ in E^* there corresponds at least one $x \in E$ such that

$$\| x \| = 1, \qquad \langle x, \varphi \rangle = \| \varphi \|$$

To insure that for each $\varphi \neq 0$ in E^* the corresponding element x in E is unique it is sufficient (and in fact, necessary) that E be rotund. Thus if E is a rotund reflexive Banach space and φ is a continuous linear functional on E, then φ is not only bounded on the unit ball of E, but in fact attains its supremum, and does so uniquely.

The preceeding remarks show that with a rotund reflexive Banach space B

we are justified in writing $\bar{\varphi}$ for the unique vector in B of norm 1 satisfying $\langle \bar{\varphi}, \varphi \rangle = \| \varphi \|$ and in referring to $\bar{\varphi}$ as the *extremal* of φ. We adopt the convention that the extremal 0 of the 0 functional is the 0 vector in B.

Now let $x \neq 0$ be a vector in B. Regarding x as a linear functional on B the Hahn-Banach produced φ shows that x attains its supremum on the unit ball of B, and that rotundity of B^* is necessary and sufficient for x to attain its supremum uniquely. Thus, requiring that both B and B^* be rotund (and reflexive) we can denote this unique φ by $\bar{x}$ and speak of the *extremal* of x. Since the conjugate of any of the spaces of Lemma 1 is another of the same type it is clear that each of these is still a possible candidate for B.

A Banach space E is called *smooth* if at each point of ∂U there is exactly one supporting hyperplane of U. Day [2; p. 112] notes that the following properties are equivalent:

(1) E is smooth.

(2) For each $x \in \partial U$ there is at most one $\varphi \in E$ such that $\| \varphi \| = 1$ and $\varphi(x) = 1$.

(3) The functional $x \to \| x \|$ has a Gateaux differential at each point of ∂U; that is,

$$\lim_{\epsilon \to 0} \frac{\| x + \epsilon h \| - \| x \|}{\epsilon}$$

exists for each $x \in \partial U$ and $h \in E$.[2]

In addition, it is not difficult to see that for any Banach space E, E is smooth (rotund) if E^* is rotund (smooth). It follows from this that if E is reflexive, E^* is rotund if and only if E is smooth. Accordingly, to enable the dual use of the term extremal in B, we henceforth require that B be rotund, reflexive, and smooth. (This latter hypothesis will be seen to be dispensible.) We note the following properties of the extremal operation:

(i) $\bar{\bar{x}} = x/\| x \|$ for $x \neq 0$ in B,

(ii) $\bar{\bar{\varphi}} = \varphi/\| \varphi \|$ for $\varphi \neq 0$ in B,

(iii) $\overline{\lambda x} = (| \lambda |/\lambda)\, \bar{x}$ any complex scalar λ.

The proof of the following theorem is straightforward.

THEOREM 2. *Let x be given in B. The Gateaux derivative of the norm at x*

$$G(x, h) = \lim_{\epsilon \to 0} \frac{\| x + \epsilon h \| - \| x \|}{\epsilon}$$

[2] This shows that any isometric copy of a smooth space is smooth.

exists for each $h \in B$ and the mapping $h \to G(x, h)$ defines a real linear functional on B of norm 1 which assumes the value $\| x \|$ at x. Consequently, if B is a real *linear space this is the extremal $\bar{x}$ of x. In general this is the real part of the extremal of x:*

$$G(x, h) = \text{Re}\, \langle h, \bar{x} \rangle \qquad (\text{all } h \in B).$$

Recall that the conjugate T^* of T is the bounded linear transformation from R^* to B^* defined for $\varphi \in R^*$ by

$$\langle u, T^*\varphi \rangle = \langle Tu, \varphi \rangle \qquad u \in B.$$

That is, $T^*\varphi$ is the linear functional on B whose value at u is the number $\langle Tu, \varphi \rangle$. The Hahn-Banach theorem shows that $\| T^* \| = \| T \|$. The fact that T is onto R shows that T^* is one-to-one.

The next result deals with another special case.

LEMMA 2. *Suppose that for some $\xi \in R$ we have $\| T^\dagger \xi \| = \| \xi \|$. Then $T^\dagger \xi$ is given by the formula.*

$$T^\dagger \xi = \| \xi \| \, \overline{T^* \bar{\xi}}.$$

Here, if the norm on R is not smooth, $\bar{\xi}$ is understood to be any extremal of ξ.

PROOF. Without loss of generality we may assume that $\| T \| = 1$. Then $\| T^* \| = 1$ hence $\| T^*(\bar{\xi}) \| \leqslant 1$. This, together with

$$\langle T^\dagger(\xi), T^*(\bar{\xi}) \rangle = \langle \xi, \bar{\xi} \rangle = \| \xi \| = \| T^\dagger(\xi) \|$$

shows that

$$T^*(\bar{\xi}) = \overline{T^\dagger(\xi)}.$$

Taking extremals we obtain the desired formula.

REMARK. The formula in the preceeding lemma yields $T^\dagger(\xi)$ to within a positive constant in terms of the extremal operations on R and B. It is generally an easy task to write an explicit formula for construction of extremals. For example, consider the product $B = B_1 \times B_2 \times \cdots B_n$ where the B_i are rotund and B is normed as in Lemma 1. Each bounded linear functional φ on B may be identified with an n-tuple $(\varphi_1, \varphi_2, \cdots, \varphi_n)$ where $\varphi_i \in B_i^*$. Let $\bar{\varphi}_i$ be the extremal of φ_i in B_i. Then it is easy to verify that with the p-norm on B the extremal $\bar{\varphi}$ of φ is given by

$$\bar{\varphi} = (\alpha_1 \bar{\varphi}_1, \alpha_2 \bar{\varphi}_2, \cdots, \alpha_n \bar{\varphi}_n)$$

$$\alpha_i = \alpha^{-1} \| \varphi_i \|^{q-1}, \qquad \alpha = \left(\sum \| \varphi_i \|^q \right)^{1/q}, \qquad q = \frac{p}{p-1}.$$

Similarly, with the matrix norm on B, φ has the form

$$\bar{\varphi} = (\beta_1\bar{\varphi}_1 , \beta_2\bar{\varphi}_2 , \cdots, \beta_n\bar{\varphi}_n)$$

$$\beta_i = \beta^{-1} \sum_j b_{ij} \| \varphi_j \| , \qquad \beta = \sum_{i,j} b_{ij} \| \varphi_i \| \| \varphi_j \|$$

where $[b_{ij}]$ is the inverse of the matrix $[a_{ij}]$.

These formulas imply that the conjugate space B^* is (isometrically isomorphic to) the product $B_1^* \times B_2^* \times \cdots \times B_n^*$ with the respective norms of $\varphi = (\varphi_1 , \varphi_2 , \cdots, \varphi_n)$ given by

$$\| \varphi \| = \Big(\sum_i \| \varphi_i \|^q\Big)^{1/q}, \qquad \| \varphi \| = \sum_{i,j} b_{ij} \| \varphi_i \| \| \varphi_j \| .$$

It follows that if each B_i is also reflexive and smooth, so that each $x = (x_1 , x_2 , \cdots x_n)$ in B has an extremal $\bar{x}$ in B^*, then

$$\bar{x} = (\beta_1\bar{x}_1 , \beta_2\bar{x}_2 , \cdots, \beta_n\bar{x}_n)$$

$$\beta_i = (\beta^{-1} \| x_i \|)^{p-1} \qquad \beta = \Big(\sum_i \| x_i \|^p\Big)^{1/p}$$

and

$$\bar{x} = (\delta_1\bar{x}_1 , \delta_2\bar{x}_2 , \cdots, \delta_n\bar{x}_n)$$

$$\delta_i = \delta^{-1} \sum_j a_{ij} \| x_j \| , \qquad \delta = \sum_{i,j} a_{ij} \| x_i \| \| x_j \| .$$

In particular, if $B = l_{p,n}$ is the space of complex n-tuple $\xi = (\xi_1 , \xi_2 , \cdots, \xi_n)$ with the norm

$$\| \xi \| = \Big(\sum_i \| \xi_i \|^p\Big)^{1/p}$$

then

$$\bar{\xi} = (\eta_1 , \eta_2 , \cdots, \eta_n)$$

where

$$\eta_i = \begin{cases} \dfrac{\bar{\xi}_i}{| \xi_i |}\left(\dfrac{| \xi_i |}{\| \xi \|}\right)^{p-1} & \text{if} \quad \xi_i \neq 0 \\ 0 & \text{if} \quad \xi_i = 0. \end{cases}$$

A precisely analogous formula holds in L_p $(1 < p < \infty)$.

Observe also that if T arises from a linear system in the sense that for a system input u, Tu is the value of the output state vector at some fixed instant, then its range is finite dimensional so that T^*, being a linear transformation on a finite dimensional space, will be given by a matrix. Thus, evaluation of $T^{\dagger}(\xi)$ is reduces to familiar computations. Finally, note that the preceeding remarks in particular determine the extremal operations in (suitably normed) input spaces of the form

$$B = l_{p_1,n_1} \times l_{p_2,n_2} \times \cdots \times l_{p_k,n_k} \times L_{q_1} \times L_{q_2} \times \cdots \times L_{q_l}$$

where $1 \leqslant n_i \leqslant \infty$ and $1 < p_i , q_i < \infty$. In other words, T may represent systems with digital and/or functional inputs.

LEMMA 3. *Let $C = T(U)$ be the image of the unit ball in B. Then C is a convex, circled,*[3] *weakly compact, neighborhood of* 0 *in R.*

PROOF. Since T is linear, C is convex and circled. The Opening Mapping Theorem shows that $T(U)$ contains a multiple of the unit ball in R, and hence is a neighborhood of 0. Finaly, it is known [6; p. 115] that a continuous linear mapping from one Banach space into another remains continuous when both spaces are equipped with their weak topologies. Since U is weakly compact in B it follows that $T(U)$ is weakly compact in R.

It follows from Lemma 3 that C is *radial* at 0. That is, for each $\xi \in R$ there is a scalar $\lambda > 0$ such that $\xi \in \lambda C$. Hence [7] the Minkowski functional p given by

$$p(\xi) = \inf \{\lambda > 0 : \xi \in \lambda C\}$$

is defined and finite on all of R. Since C is convex and circled the functional p is subadditive and absolutely homogeneous.

$$p(\xi + \zeta) \leqslant p(\xi) + p(\zeta) \qquad \xi, \zeta \in R$$
$$p(\lambda\xi) = | \lambda | p(\zeta).$$

The next lemma lists a few facts we will need.

LEMMA 4. (i) *The interior of C consists of those $\xi \in R$ for which $p(\xi) < 1$.*

(ii) *$\partial C = \{\xi \in R : p(\xi) = 1\}$ is the boundary of C.*

(iii) *$\partial C \subset T(\partial U)$.*

PROOF. The assertions (i) and (ii) are well known and follow directly from the definition of p. As for (iii), if $\xi \in \partial C$, then $\xi \in C$ and hence $\xi = Tu$ for some $u \in U$. Since by (ii), $p(\xi) = 1$, we have $\lambda^{-1}\xi \notin C$ for all $\lambda < 1$. But then $\lambda^{-1}u \notin U$ for all $\lambda < 1$. This means that $\| u \| \geqslant 1$, and since $u \in U$, that $\| u \| = 1$.

[3] A set C in a vector space E is *circled* if $\lambda C \subset C$ for all $| \lambda | \leqslant 1$.

REMARK. It is easy to construct examples to show that the reverse inclusion (iii) is not valid in general.

COROLLARY. *The functional p is a norm on R equivalent to the given norm. In fact for some constant $k > 0$ we have*

$$\frac{1}{\| T \|} \| \xi \| \leqslant p(\xi) \leqslant k \| \xi \| \qquad (\xi \in R).$$

PROOF. Suppose $\| \xi \| > 0$ and let λ be any positive scalar with $\xi \in \lambda C$. Then $1/\lambda \, \xi \in T(U)$ and hence

$$\left\| \frac{1}{\lambda} \xi \right\| \leqslant \| T \| .$$

This implies that

$$p(\xi) \geqslant \frac{\| \xi \|}{\| T \|}$$

and hence that p is a norm on R.

By Lemma 3, $C = \{\xi : p(\xi) \leqslant 1\}$ is a neighborhood of 0 in R and hence there is an $\epsilon > 0$ such that $p(\zeta) \leqslant 1$ if $\| \zeta \| \leqslant \epsilon$. Hence $p(\xi) \leqslant (1/\epsilon) \| \xi \|$ for all $\xi \in R$.

We are now able to obtain the promised characterization of $T^{\dagger}(\xi)$. If N is a *real* linear functional on a real vector space E we will say that a subset C of E *lies to the left* of the hyperplane $H = \{\xi \in E : \langle \xi, N \rangle = \alpha_0\}$ provided that $\langle \xi, N \rangle \leqslant \alpha_0$ for all $\xi \in C$. H *supports* C if it meets C and if C lies entirely on one side of H. A geometric form of the Hahn-Banach Theorem, valid in any topological vector space, asserts that a closed convex set with nonempty interior has a supporting hyperplane through each of its boundary points [8; p. 72].

THEOREM 4. *Let $\xi_0 \neq 0$ be a given vector in R and let $\alpha = p(\xi_0)^{-1}$. Then there exists a unique vector N in the unit sphere of R^* such that*

$$T^{\dagger}(\xi_0) = p(\xi_0) \, \overline{T^* N}.$$

The functional N is uniquely determined by the conditions

(i) $\| N \| = 1$.

(ii) *C lies to the left of the hyperplane $H = \{\xi \in R : \langle \xi, N \rangle = \alpha\langle \xi_0 , N\rangle\}$. If B is a complex space this last requirement is to be interpreted as saying that*

$$\operatorname{Re} \langle \xi, N \rangle \leqslant \operatorname{Re} \langle \alpha \xi_0 , N \rangle \qquad \text{all} \qquad \xi \in C.$$

PROOF. Suppose first that B is real. Since C is closed, convex, and has nonempty interior it follows that there is a supporting hyperplane of C at $\alpha\xi_0$ and hence a functional N satisfying (i) and (ii). Note that since $0 \in C$, N is nonnegative at $\alpha\xi_0$.

To prove the theorem it evidently suffices to prove:

(a) $\varphi \in R^*$ satisfies $T^\dagger(\xi_0) = p(\xi_0)\,\overline{T^*\varphi}$ if and only if (ii) holds for φ.

(b) There is at most one φ of norm 1 satisfying $T^\dagger(\xi_0) = p(\xi_0)\,\overline{T^*\varphi}$.

The proof of (b) follows from the fact that the mapping $\varphi \to \overline{T^*\varphi}$ is one-to-one from the unit sphere of R^* into the unit sphere of B.

Suppose next that $T^\dagger(\xi_0) = p(\xi_0)\,\overline{T^*\varphi}$ for some $\varphi \in R^*$. Then

$$\xi_0 = T^\dagger(\xi_0) = \alpha^{-1}T(\overline{T^*\varphi})$$

and hence

$$\begin{aligned}\langle \xi_0, \varphi\rangle &= \langle T(\overline{T^*\varphi}), \varphi\rangle = \langle \overline{T^*\varphi}, T^*\varphi\rangle \\ &= \| T^*\varphi \| \geqslant \langle u, T^*\varphi\rangle = \langle Tu, \varphi\rangle\end{aligned}$$

for all $u \in U$ and since $C = T(U)$ this shows that φ satisfies (ii). (Note that since φ is a real functional, the number $\langle u, T^*\varphi\rangle$ is real for any $u \in U$.)

Finally, suppose $\varphi \in R^*$ satisfies (ii). Since $\alpha\xi_0 \in \partial C$ there is a $u_0 \in \partial U$ with $Tu_0 = \alpha\xi_0$. Then

$$\langle u_0, T^*\varphi\rangle = \langle \alpha\xi_0, \varphi\rangle = |\langle \alpha\xi_0, \varphi\rangle| = |\langle u_0 T^*\varphi\rangle|.$$

Hence by definition of the norm of the functional $T^*\varphi$ on B we have

$$\| T^*\varphi \| = \sup_{u\in U} |\langle u, T^*\varphi\rangle| \geqslant \langle u_0, T^*\varphi\rangle$$

and since $\overline{T^*\varphi} \in U$,

$$\langle \xi_0, \varphi\rangle \geqslant \langle T(\overline{T^*\varphi}), \varphi\rangle = \| T^*\varphi \|.$$

We conclude that $\langle u_0, T^*\varphi\rangle = \| T^*\varphi \|$ and hence that $u_0 = \overline{T^*\varphi}$. Thus the vector $\alpha^{-1}\overline{T^*\varphi}$ is a pre-image (under T) of ξ_0 and to prove that this is $T^\dagger(\xi_0)$ it remains only to show that any $u \in B$ satisfying $Tu = \xi_0$ has a norm of at least α^{-1}. This however follows from

$$\langle u, T^*\varphi\rangle = \langle \xi_0, \varphi\rangle = \alpha^{-1}\langle \alpha\xi_0, \varphi\rangle = \alpha^{-1} \| T^*\varphi \|$$

and the fact that

$$\| u \| = \sup_{f\in B^*} \frac{|\langle u, f\rangle|}{\| f \|}.$$

Suppose now that B is a complex space. Then [7; p. 118] the boundary point $\alpha\xi_0$ of C can be separated from C by a complex linear functional N in the sense that

$$\text{Re}\,\langle \xi, N\rangle \leqslant \text{Re}\,\alpha\langle \xi_0, N\rangle \qquad \text{all} \qquad \xi \in C.$$

The remainder of the argument now proceeds as before.

REMARK. The unique vector N in R^* satisfying (i) and (ii) deserves, in a natural way, to be called the *outward normal* to C at $\alpha\xi_0$. We have shown that there is an outward normal to C at each of its boundary points.

Observe also that it follows from the theorem that $\| T^\dagger(\xi) \| = p(\xi)$. Since the latter function is a (uniformly) continuous function, we see that the minimum effort associated with each state $\xi \in R$ is a continuous function of ξ: if two vectors ξ_1, ξ_2 in R are close, and if u_1 and u_2 are their minimum pre-images under T, then the norms of u_1 and u_2 are correspondingly close.

It is easy to show that in case $B = H$ is a Hilbert space, the formulas $T^\dagger(\xi) = T_M^{-1}\,\xi$ and $T^\dagger(\xi) = p(\xi)\,\overline{T^*N}$ are consistent.

LEMMA 5. *For each $\xi \in R$, set $|\xi| = p(\xi)$. Then $|\;|$ is a norm on R, equivalent to the given norm. Let R_1 denote the space R equipped with the norm $|\;|$. Then R_1 is rotund and smooth.*

The proof of Lemma 5 is left to the reader. Now it follows from Lemma 5 and Theorem 4 that the definition $|\xi| = p(\xi)$ yields a norm on R for which

$$|\xi| = \| T^\dagger(\xi) \|$$

holds identically in ξ. It therefore follows from Lemma 2 that

$$T^\dagger(\xi) = p(\xi)\,\overline{T^*(\xi')} \qquad \xi \in R$$

where ξ' denotes the extremal of ξ relative to the norm $|\;|$. That is, ξ' is characterized by the equations

$$\sup_{p(\zeta)=1} |\langle \zeta, \xi'\rangle| = 1, \qquad \langle \xi, \xi'\rangle = p(\xi).$$

Since by Lemma 1(c) applied to R^*,

$$\overline{T^*(\xi'/\|\xi'\|)} = \overline{T^*(\xi')}$$

we have proven part of the following:

THEOREM 5. *Let ξ be a fixed boundary point of C and let N be the outward normal to C at ξ. Then*

(i) $N = \xi'/\|\xi'\|$ *where ξ' is the extremal of ξ relative to the norm $p(\xi)$ on R.*

(ii) *N is the unique vector φ in R^* of norm* 1 *satisfying* $\| T^*\varphi \| = \langle \xi, \varphi \rangle$.

(iii) $N = \xi' / \| \xi' \|$ *where ξ' is the bounded linear functional on R whose real part is defined for $\zeta \in R$ by*

$$\operatorname{Re} \langle \zeta, \xi' \rangle = \lim_{\epsilon \to 0} \frac{| \xi + \epsilon\zeta | - | \xi |}{\epsilon} = \lim_{\epsilon \to 0} \frac{p(\xi + \epsilon\zeta) - p(\xi)}{\epsilon}.$$

Proof. If $\varphi \in R^*$ satisfies $\| T^*\varphi \| = \langle \xi, \phi \rangle$, then for any $\zeta \in C$, we may *choose $u \in U$* so that $Tu = \zeta$ to obtain

$$\langle \zeta, \varphi \rangle = \langle Tu, \varphi \rangle = \langle u, T^*\varphi \rangle \leqslant \| T^*\varphi \| = \langle \xi, \varphi \rangle$$

and hence, by Theorem 4, φ is a positive multiple of N. This proves (ii).

Now consider (iii). We observe that since R is smooth its norm has a Gateaux derivative at each point on the boundary of its unit ball. That is,

$$G(\xi, \zeta) = \lim_{\epsilon \to 0} \frac{| \xi + \epsilon\zeta | - | \xi |}{\epsilon}$$

exists for each $\xi \in \partial C$ and ζ in R. Assertion (iii) now follows from Theorem 2.

4. Discussion

It is clear from the preceding results that once one knows the set C relatively simple computations furnish (a) the minimum effort T needs to reach any given state ξ in R and (b) the precise pre-image $T^\dagger(\xi)$ of ξ whose effort is this minimum value. Indeed the boundary of the set αC is a "level surface" consisting of those states $\xi \in R$ which T can obtain with a minimum energy of precisely α, and the outward normals to C determine (to within a positive multiple) the class of minimum energy inputs. However, even in the relatively simple case in which B is finite dimensional, the equation $C = T(U)$ is unsuitable for specifying C. It is therefore, natural to seek a simpler way to determine C. For example, if C is a multiple of the unit ball in R we need only one parameter to specify C completely; if C is an ellipsoid we need only to determine the size of its semiaxes, and so on. In any event the conditions of Theorem 5 are sufficient to compute N by iterative procedures if necessary.

References

1. L. W. Neustadt. Minimum effort control systems. *J. Soc. Indust. Appl. Math. Control, Ser. A.* **1** (1962).
2. M. M. Day. "Normed Linear Spaces." Springer, Berlin, 1962.
3. J. A. Clarkson. Uniformly convex spaces. *Trans. Am. Math. Soc.* **40** (1936).

4. L. LIUSTERNIK AND V. SOBOLEV. "Elements of Functional Analysis." Ungar, New York, 1961.
5. R. R. PHELPS. Uniqueness of Hahn-Banach extensions and unique best approximation. *Trans. Am. Math. Soc.* **95** (1960), 238-255.
6. N. BOURBAKI. "Espaces Vectorials Topologiques," Chaps. III, V. Hermann, Paris, 1955.
7. J. L. KELLY AND I. NAMIOKA. "Linear Topological Spaces." Van Nostrand, Princeton, New Jersey, 1963.
8. N. BOURBAKI. "Espaces Vectoriels Topologiques," Chaps. I, II. Hermann, Paris, 1953.