THE EXISTENCE OF OPTIMAL CONTROLS

Emilio Roxin

1. INTRODUCTION

In the theory of optimal control (see references [1], [6]) it is commonly assumed that the solution of the optimization problem exists, and therefore only necessary conditions are obtained. It is our aim to give here sufficient conditions of rather general type which assure the existence of the solution.

We shall study the system

$$\dot{x}^i = f^i(t, x^1, \cdots, x^n, u^1, \cdots, u^m) \qquad (i = 1, 2, \cdots, n), \tag{1.1}$$

where, as usual, $\dot{x} = dx/dt$. We write

$$x^0 = t, \qquad f^0 = 1, \tag{1.2}$$

and we introduce the vectors

$$\begin{aligned} \hat{x} &= (x^1, \cdots, x^n), \\ x &= (x^0, x^1, \cdots, x^n) = (x^0, \hat{x}), \\ \hat{f} &= (f^1, \cdots, f^n), \\ f &= (f^0, f^1, \cdots, f^n) = (f^0, \hat{f}), \\ u &= (u^1, \cdots, u^m) \end{aligned} \tag{1.3}$$

in euclidean vector space, with the usual norm $\|x\|^2 = \Sigma(x^i)^2$. Equation (1.1) then becomes

$$\dot{\hat{x}} = \hat{f}(t, \hat{x}, u) \tag{1.4}$$

or

$$\dot{x} = f(x, u). \tag{1.5}$$

The following assumptions are made:

i) $\hat{f}(t, \hat{x}, u)$ is defined in $I \times \hat{X} \times U$, where I is the real line (the positive half-line could be used, alternatively), $\hat{X}$ is the $\hat{x}$-space, and U is a compact set in u-space.

Received September 2, 1961.

This research was partially supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract Number AF 49(638)-382. Reproduction in whole or in part is permitted for any purpose of the United States Government.

ii) $\hat{f}(t, \hat{x}, u)$ is continuous in $(\hat{x}, u)$ and integrable with respect to t for each $(\hat{x}, u) \in \hat{X} \times U$.

iii) The Lipschitz condition holds: there exists a constant K such that for any $(t, u) \in I \times U$

$$\| f(t, \hat{x}, u) - f(t, \hat{x}_1, u) \| < K \cdot \| \hat{x} - \hat{x}_1 \| .$$

iv) For all $u \in U$ (uniformly)

$$\| \hat{f}(t, \hat{x}_1, u) \| \leq \mu(t) \cdot g(\| \hat{x} \|) ,$$

where the function $\mu(t)$ is integrable in every finite interval, $g(\| \hat{x} \|)$ is bounded in each bounded region of the x-space, and $g(\| \hat{x} \|) = O(\| \hat{x} \|)$ for $\| x \| \to \infty$.

We shall consider the parameter u as a function of t. A function $u(t)$ will be called an *admissible control function* if

v) $u(t)$ is measurable;

vi) for each t, $u(t) \in U$.

On the set U we impose the convexity condition

vii) $\hat{f}(t, \hat{x}, U) = \{ \hat{f}(t, \hat{x}, u) \mid u \in U \}$ is a convex set for each $(t, \hat{x}) \in I \times \hat{X}$.

We note that since U is compact and $\hat{f}(t, \hat{x}, u)$ is continuous in u, the set $\hat{f}(t, \hat{x}, U)$ is also compact. The same is true of the set $f(x, U)$, which differs from the former by the component $f^0 = 1$.

According to well-known theorems (see [2], [8]), these conditions guarantee the existence and uniqueness of the solution of (1.5) for any admissible control function $u(t)$. In particular, a finite escape time is ruled out by the assumptions made in (iv). The importance of the convexity condition (vii) will appear in the proof of the theorem.

2. THE ATTAINABLE SET

Definition. Given the equation (1.5), we shall say that the point x_1 is *attainable* from x_0 if there exists an admissible control function $u(t)$ defined in the time interval $(t_0, t_1) = (x_0^0, x_1^0)$, where $t_0 \geq t_1$, such that equation (1.5), with $u = u(t)$, together with the initial condition

$$x(t_0) = x_0$$

has the solution $x(t)$ satisfying

$$x(t_1) = x_1 .$$

The set of all points x that are attainable from x_0 will be called the set attainable from x_0, and we shall denote it by R_{x_0}. When there is no possibility of misunderstanding, the subscript x_0 will be dropped.

In many simple cases it is easy to see which points are attainable and which are not; the linear case has been treated in detail in [4], [6]. Here we are concerned with rather general properties of the attainable set.

THEOREM. *If the equation* (1.5) *fulfills the conditions* (i) *to* (vii), *then for each initial* x_0, *the attainable set* R_{x_0} *is a closed set.* (A similar theorem under less general assumptions was proved in [7].)

Without loss of generality we may suppose that $x_0 = 0$, and we write simply R for R_{x_0}.

We shall prove that if $\xi_1, \xi_2, \xi_3, \cdots$ is a sequence of points of X such that $\xi_i \in R$ and $\lim_{i\to\infty} \xi_i = \xi_0$, then $\xi_0 \in R$.

Since ξ_i belongs to R for $i = 1, 2, 3, \cdots$, there exist admissible control functions $u_i(t)$ such that the corresponding solutions $x_i(t)$ of (1.5), which start at $x_i(0) = 0$, satisfy

$$x_i(t) = \int_0^t f(x_i(t), u_i(t))dt \qquad (0 \leq t \leq T_i,\ i = 1, 2, 3, \cdots) \tag{2.1}$$

with

$$x_i(T_i) = \xi_i \quad (T_i = \xi_i^0) \qquad (i = 1, 2, 3, \cdots). \tag{2.2}$$

Now we take some interval (0, T) which includes all the intervals $(0, T_i)$ (this is possible, since $\lim_{i\to\infty} T_i = \xi_0^0 = T^0$), and for $i = 1, 2, 3, \cdots$ we define the functions

$$\phi_i(t) = \begin{cases} f(x_i(t), u_i(t)) & \text{for } 0 \leq t \leq T_i, \\ f(x_i(T_i), u_i(T_i)) = \text{const.} & \text{for } T_i \leq t \leq T. \end{cases} \tag{2.3}$$

We also extend the definition of the $x_i(t)$ to the interval (0, T) by

$$x_i(t) = \int_0^t \phi_i(t)\,dt \qquad (i = 1, 2, 3, \cdots), \tag{2.4}$$

which coincides with (2.1) in $(0, T_i)$.

The functions $x_i(t)$ are uniformly bounded. In fact, assumption (iv) assures us that the equation

$$\frac{d\eta(t)}{dt} = \mu(t)\,g(\eta), \qquad \eta(0) = 0$$

has a solution in the whole interval (0, T), where it is therefore bounded. Besides, the inequality

$$\frac{d}{dt}\|x_i\| \leq \left\|\frac{dx_i}{dt}\right\| = \|f(t, x_i, u_i)\| \leq \mu(t)\,g(\|x_i\|)$$

shows in a well-known manner that since $x_i(0) = 0$,

$$\|x_i(t)\| \leq \eta(t),$$

proving that the $x_i(t)$ are uniformly bounded in $(0, T)$.

Since the $x_i(t)$ are uniformly bounded, we can write

$$\|f(t, x_i, u_i)\| \leq \mu^*(t),$$

where $\mu^*(t)$ is integrable. The functions $f(t, x_i, u_i)$ are measurable and therefore integrable, and the same applies to the $\phi_i(t)$. The inequality

$$\int \|\phi_i(t)\| \, dt \leq \int \mu^*(t) \, dt$$

shows that, considered as elements of the L_1-space of the interval $(0, T)$, the functions $\phi_i(t)$ form a bounded sequence. By weak completeness of L_1 ([3]), we can select a subsequence which converges weakly to a certain measurable function $\phi_0(t)$. In the following, we suppose that the indices have been changed in such a way that $i = 1, 2, 3, \cdots$ refers to the weakly convergent subsequence mentioned above. Therefore, for each measurable set E, $E \subset (0, T)$,

$$\lim_{i \to \infty} \int_E \phi_i(t) \, dt = \int_E \phi_0(t) \, dt. \tag{2.5}$$

Let

$$T_0 = \xi_0^0 = \lim_{i \to \infty} T_i. \tag{2.6}$$

If the measurable set E is contained in an interval $(0, \tau)$ with $\tau < T_0$, then $E \subset (0, T_i)$ for almost all i, and according to (2.3)

$$\lim_{i \to \infty} \int_E f(x_i(t), u_i(t)) \, dt = \int_E \phi_0(t) \, dt, \tag{2.7}$$

this equality being valid in any interval $(0, \tau)$ with $\tau < T_0$.

We now define the function

$$x_0(t) = \int_0^t \phi_0(t) \, dt \qquad (0 \leq t \leq T). \tag{2.8}$$

Combining (2.1) and (2.7), we see that for each value of t less than T_0,

$$\lim_{i \to \infty} x_i(t) = x_0(t) \qquad (0 \leq t \leq T_0).$$

The proof that also

$$\lim_{i \to \infty} x_i(T_i) = x_0(T_0)$$

is immediate, since in

$$\|x_i(T_i) - x_0(T_0)\| \leq \|\int_0^{T_0} [\phi_i(t) - \phi_0(t)]\,dt\| + \int_{T_0}^{T_i} \|\phi_i(t)\|\,dt$$

the first integral tends to zero according to the weak convergence, and the second one because $T_i \to T_0$, the functions $\|\phi_i(t)\|$ being majorized by the integrable function $\mu^*(t)$. Hence

$$x_0(T_0) = \xi_0^0. \tag{2.9}$$

It remains to prove that $x_0(t)$ is the trajectory corresponding to a certain admissible control function $u_0(t)$. According to (2.8), this means that

$$\phi_0(t) = f(x_0(t), u_0(t)) \qquad (0 \leq t \leq T_0). \tag{2.10}$$

From the weak convergence we deduce that, for every vector y of the space X,

$$\limsup[y \cdot \phi_i(t)] \geq y \cdot \phi_0(t) \geq \liminf[y \cdot \phi_i(t)] \tag{2.11}$$

almost everywhere in $(0, T_0)$. In fact, if we suppose, for example, that on a set E of positive measure

$$\limsup[y \cdot \phi_i(t)] < y \cdot \phi_0(t),$$

then we obtain immediately

$$\limsup[y \cdot \int_E \phi_i(t)\,dt] < y \cdot \int_E \phi_0(t)\,dt,$$

in contradiction to (2.5).

For any value of t $(0 \leq t < T_0)$ and almost every i,

$$\phi_i(t) \in \{f(x_i(t), U)\};$$

therefore, for almost every value of t of the interval $(0, T_0)$,

$$\lim_{i\to\infty} \sup[\text{l.u.b.}\,(y \cdot f(x_i(t), U))] \geq y \cdot \phi_0(t) \geq \lim_{i\to\infty} \inf[\text{g.l.b.}\,(y \cdot f(x_i(t), U))].$$

By continuity of the function $f(x, u) = f(t, \hat{x}, u)$ in $(\hat{x}, u)$ (the value of t remains fixed), we may write

$$\text{l.u.b.}\,[y \cdot f(x_0(t), U)] \geq y \cdot \phi_0(t) \geq \text{g.l.b.}\,[y \cdot f(x_0(t), U)]. \tag{2.12}$$

This equation, valid for every vector y, implies that $\phi_0(t)$ belongs to the closed convex set $f(x_0(t), U)$; that is, there exists a value $u_0 \in U$ such that

$$\phi_0(t) = f(x_0(t), u_0). \tag{2.13}$$

In this way we have defined a function $u_0(t)$, on almost every point of the interval $(0, T_0)$. Obviously we may extend the definition to the remaining points in any convenient way (for example, putting $u_0 = 0$ at these points), without affecting the integral of $f(x_0(t), u_0(t))$.

Having proved that $u_0(t)$ takes its values in U, we need only show that the values of $u_0(t)$ can be selected so that, as a function of t, $u_0(t)$ is measurable. In this connection, we observe that it may happen that for some values of t, $u_0(t)$ is not uniquely defined by the value of $f(x_0(t), u_0)$.

We proceed to construct the function $u_0(t)$ in the following manner.

1) Take some $\varepsilon > 0$.

2) The function $\phi_0(t)$ is integrable in $(0, T_0)$. Therefore, we can determine a bounded function $\phi_0'(t)$ which coincides with $\phi_0(t)$ in the interval $(0, T_0)$ with the exception of a set of measure less than ε. We shall call A the set of values of t where $\phi_0(t) = \phi_0'(t)$.

3) Divide the set A into measurable subsets A_i $(i = 1, 2, 3, \cdots, p)$ such that the oscillation of $\phi_0(t)$ on each A_i (this is the l. u. b. $\|\phi_0(t_1) - \phi_0(t_2)\|$ for $t_1, t_2 \in A_i$) is less than ε. This can be done, since $\phi_0(t)$ is measurable and bounded on A.

4) Take a sequence u_i $(i = 1, 2, 3, \cdots)$ which is dense in U (U is a compact set in euclidean space).

5) In each set A_i $(i = 1, 2, \cdots, p)$, take a value $t_i \in A_i$. Call $\alpha_i = \phi_0(t_i)$.

6) Denote by A_{ij} the subset of A_i where

$$\|f(x_0(t), u_j) - \alpha_i\| < 2\varepsilon .$$

The sets A_{ij} are measurable, since $f(x_0(t), u) = f(t, \hat{x}_0(t), u)$ is measurable in t and continuous in $\hat{x}_0$, while $\hat{x}_0(t)$ is also continuous.

7) Now we can prove that

$$\bigcup_{j=1}^{\infty} A_{ij} = A; \qquad (i = 1, 2, \cdots) .$$

In fact, if $t \in A_i$ we have

$$\|\phi_0(t) - \phi_0(t_i)\| = \|\phi_0(t) - \alpha_i\| < \varepsilon .$$

Besides, as proved above, there exists a value $v \in U$ such that $f(x_0(t), v) = \phi_0(t)$. Since $f(x, u)$ is continuous in u and the sequence $\{u_i\}$ is dense, there exists some u_j such that

$$\|f(x_0(t), u_j) - f(x_0(t), v)\| = \|f(x_0(t), u_j) - \phi_0(t)\| < \varepsilon .$$

Hence

$$\|f(x_0(t), u_j) - \alpha_i\| < 2\varepsilon ,$$

and $t \in A_{ij}$ for that value of j.

8) A certain value of t may of course belong to several sets A_{ij}. In order to define the function $u_\varepsilon(t)$, we shall therefore use the sets

$$A'_{ij} = A_{ij} \cap C\left\{\bigcup_{k=1}^{j-1} A_{ik}\right\},$$

where the C indicates the complementary set. Note that $t \in A'_{ij}$ if t belongs to A_{ij} but does not belong to some A_{ik} with $k < j$.

9) We are now in a position to define the function

$$u_\varepsilon(t) = u_j \text{ for } t \in A'_{ij}.$$

As shown above, this function is defined on the whole interval $(0, T_0)$ less a set of measure smaller than ε. It has the property that

$$\|\phi_0(t) - f(x_0(t), u_\varepsilon(t)\| \leq \|\phi_0(t) - \alpha_1\| + \|f(x_0(t), u_j) - \alpha_i\| \leq \varepsilon + 2\varepsilon = 3\varepsilon.$$

10) If we make this construction for a nullsequence of values of ε, we obtain a sequence of functions $u_\varepsilon(t)$ such that

$$\lim_{\varepsilon \to 0} f(x_0(t), u_\varepsilon(t)) = \phi_0(t) \tag{2.14}$$

almost everywhere on $(0, T_0)$.

11) For defining the corresponding limit function $u_0(t)$ we may, for example, put for each component of the vector $u_0(t)$

$$u_0^i(t) = \lim_{\varepsilon \to 0} \sup u_\varepsilon^i(t). \tag{2.15}$$

The value $u_\varepsilon(t)$ belongs to U, and since U is compact, $u_0(t) \in U$ for each t. (The formula (2.15) defines $u_0(t)$ almost everywhere, but the definition may obviously be completed in a convenient way.)

As upper limit of a sequence of measurable functions, $u_0(t)$ is measurable [5].

Finally, by continuity of $f(x, u)$ with respect to u, and by the equalities (2.14) and (2.15),

$$f(x_0(t), u_0(t)) = \phi_0(t) \tag{2.16}$$

almost everywhere, and our theorem is proved.

Remark 1. The convexity condition (vii) on $f(x, U)$ is essential in our theorem. The following example shows that without this condition the result is no longer true, even in very simple cases.

Consider the system

$$\dot{x}^0 = 1, \qquad \dot{x}^1 = \frac{|u|}{1 + (x^2)^2}, \qquad \dot{x}^2 = \frac{u}{1 + (x^2)^2},$$

with the admissibility condition $|u(t)| \leq 1$ for the scalar u.

The set $f(x, U)$ consists of two segments (as shown in Figure 1)

$$\dot{x}^1 = \pm\dot{x}^2, \qquad |\dot{x}^1| = |\dot{x}^2| \leq \frac{1}{1 + (x^2)^2}.$$

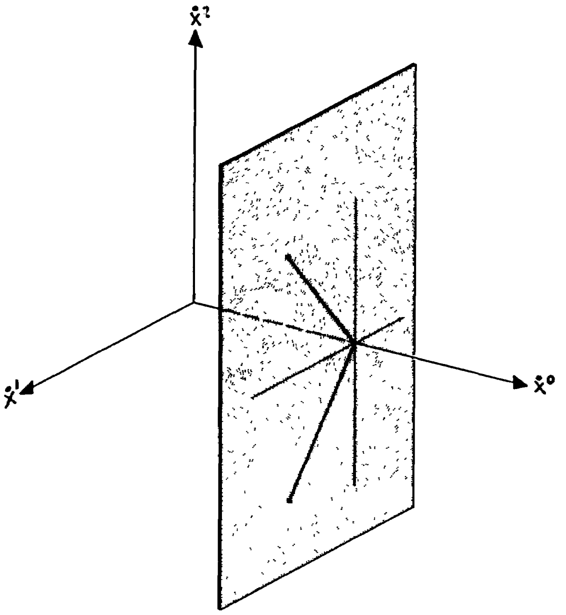

Figure 1

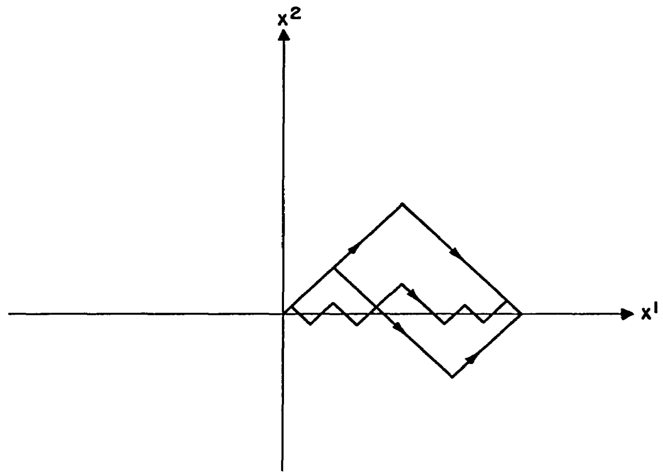

Figure 2

Figure 2 shows several possible trajectories in the (x^1, x^2)-plane, for reaching the point $x^1 = 1$, $x^2 = 0$ from the origin, the path making constantly an angle of $\pm 45^0$ with the x^1-axis.

The corresponding time, that is, the value of x^0 at the end point, will be smaller for trajectories where $(x^2)^2$ is smaller, in other words, for trajectories nearer to the x^1-axis. The lower limit is not attained because the limit trajectory, which is the x^1-axis, is not admissible.

Remark 2. We now give an example of an attainable set which is not closed; here there exist trajectories going to infinity in finite time (condition (iv) is violated).

If in the equation

$$\dot{x} = 2x^2(1 - t) - 1 + u \qquad (|u| \leq 1)$$

we put $u = 1$, we obtain $\dot{x} = 2x^2(1 - t)$. The trajectories of this equation, given by

$$x = \frac{1}{(t - 1)^2 + c},$$

are shown in Figure 3.

If we start at P_0 $(t = 0, x = 1)$, the attainable set is limited on its upper side by the curve for which $c = 0$, because we can always proceed along the curves shown in Figure 3, for $c > 0$. The first branch of the curve $c = 0$, between P_0 and infinity, is attainable, but the branch from infinity to the right is not (without passing through infinity) because once we are on a curve $c = \varepsilon > 0$, we can no longer reach the curve $c = 0$. A point like P is thus not attainable from P_0, in spite of the fact that it is on the boundary of the attainable set.

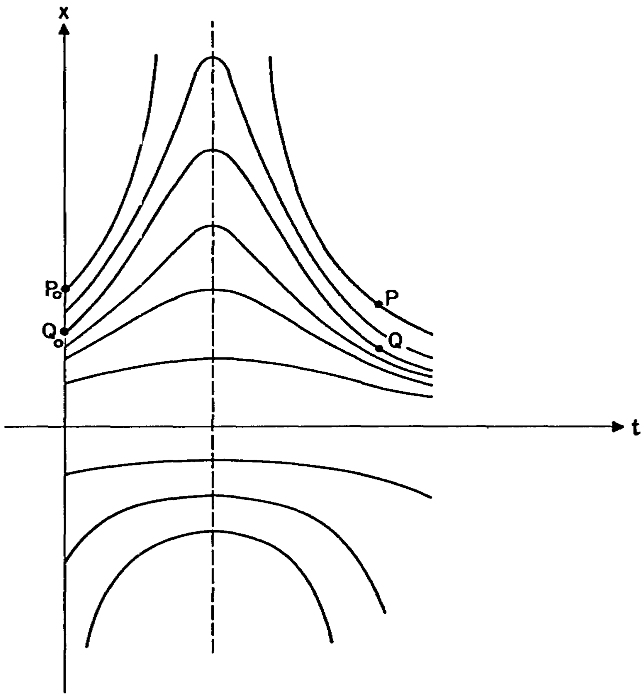

Figure 3

Remark 3. Sometimes equation (1.5) is given in the form

$$\dot{x} = f_1(x) + B(x)v \qquad (v \in V), \tag{2.16}$$

where x and f_1 are n-vectors, v is an m-vector, and B is an $n \times m$ matrix.

In general, equation (1.5) can be written

$$\dot{x} = f(x, u_0) + [f(x, u) - f(x, u_0)] = f_1(x) + v, \tag{2.17}$$

with

$$v \in V(x) = \{f(x, U) - f(x, u_0)\}.$$

The main difference between (2.16) and (2.17) is that in equation (2.17) the set V depends on x. Under the assumption that V(x) is convex for each value of x, it is contained in a linear subspace of some dimension m, so that (2.17) can be written exactly like (2.16). The set V can then be taken independently of x if and only if for each x, the set f(x, U) in (2.17) is the linear image of some fixed set W, more precisely, if in (2.17) we can write

$$v = v(x, w) = B(x)w \qquad (w \in W). \tag{2.18}$$

On the other hand, if we start with the most general statement

$$\dot{x} = f(x, u) \qquad (u(x, t) \in V(x)) \tag{2.19}$$

and suppose that the set V(x) is continuous in x, in the sense that there exists a continuous function $\phi = \phi(x, w)$ such that V(x) is the image of the fixed set W: $V(x) = \phi(x, W)$, then (2.19) can be written also as

$$\dot{x} = f(x, \phi(x, w)) = g(x, w) \qquad (w \in W),$$

which is of the form (1.5).

3. APPLICATION TO PROBLEMS OF OPTIMIZATION

A typical formulation of the optimal control problem is the following.

Given equation (1.5) subject to the conditions stated there, given the initial condition

$$x(0) = x_0 \tag{3.1}$$

and the final condition

$$x(t_1) \in E, \tag{3.2}$$

where E is a given closed set in x-space, we are required to find an admissible control function $u(t)$ and a value $t_1 > 0$ such that (1.5), (3.1) and (3.2) are satisfied and some given function $\psi(x)$, evaluated at the final point $x(t_1)$, has the least possible value.

Sometimes we are interested in minimizing a functional of the form

$$\int_0^{t_1} \phi(x, u)\, dt,$$

but we can easily reduce this case to the preceding one by introducing a new coordinate x^{n+1} related to equation (1.5) by

$$\dot{x}^{n+1} = \phi(x, u),$$

and minimizing $\psi(x) = x^{n+1}$.

If $E \cap R_{x_0}$ is a compact set and $\psi(x)$ is continuous, we can assert that there exists a point $x_1 \in E \cap R_{x_0}$ where $\psi(x)$ attains its minimum. This gives us the solution to our problem. In general, if all the previously mentioned conditions hold, one can only assert that $E \cap R_{x_0}$ is closed, and this is not sufficient for insuring the existence of the minimum of $\psi(x)$. Nevertheless, in many cases one may restrict the attention to some bounded domain, and then the desired result follows.

An important example is the time-optimization problem, where $\psi(x) = x^0 = t$. In this case, if $E \cap R_{x_0}$ is not empty, for example, if $x_2 \in E \cap R_{x_0}$, we may look for the minimum of t in the interval $0 \leq t \leq t_2 = x_2^0$. We call this interval T, and we see easily that $E \cap R_{x_0} \cap T$ is bounded and therefore compact, so that the existence of the minimum follows.

REFERENCES

1. V. G. Boltyanskii, R. V. Gamkrelidze, and L. S. Pontryagin, *The theory of optimal processes. I. The maximum principle,* Izv. Akad. Nauk SSSR. Ser. Mat. 24 (1960), 3-42. (Russian).

2. C. Carathéodory, *Vorlesungen über reelle Funktionen,* Leipzig, 1927; reprinted, Chelsea, New York, 1948.

3. N. Dunford and J. T. Schwartz, *Linear operators, part I,* Interscience Publishers, New York, 1958.

4. R. V. Gamkrelidze, *Theory of time optimal processes for linear systems,* Izv. Akad. Nauk SSSR. Ser. Mat. 22 (1958), 469-476. (Russian).

5. H. Hahn and A. Rosenthal, *Set functions,* Univ. of New Mexico Press, Albuquerque, 1948.

6. J. P. LaSalle, *The time optimal control problem.* Contributions to the theory of nonlinear oscillations, Vol. V, Annals of Mathematics Studies No. 45, Princeton, 1960, pp. 1-24.

7. E. B. Lee and L. Markus, *Optimal control for nonlinear processes,* Arch. Rational Mech. Anal. 8 (1961), 36-58.

8. G. Sansone and R. Conti, *Equazioni differenziali non lineari,* Edizioni Cremonese, Roma, 1956.

RIAS, Baltimore, Maryland