JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 4, 111-128 (1962)

Relaxed Variational Problems*

J. WARGA

Avco Research and Development Division, Wilmington, Massachusetts

Submitted by Peter Lax

I. INTRODUCTION

Let R be an arbitrary set, T the closed interval $[t_0, t_1]$ of the real axis, E_n the euclidean n-space, V an open set in E_n, A a closed set in V, and B a closed set in $A \times A$. We assume given the vector function $g(x, t, \rho) = (g^1(x, t, \rho), \cdots, g^n(x, t, \rho))$ in E_n, defined for all $(x, t, \rho) \in V \times T \times R$.

We begin by considering the variational problem of minimizing $x^1(t_1)$ subject to the following conditions[1]:

$$\frac{dx(t)}{dt} = \dot{x}(t) = g(x(t), t, \rho(t)) \qquad \text{a.e. in } T \tag{1.1}$$

$$(x(t_0), \qquad x(t_1)) \in B \tag{1.2}$$

$$x(t) \in A, \qquad t \in T. \tag{1.3}$$

Condition (1.1) may be expressed in a slightly different form. Let $G(x, t) = \{g \in E_n \mid g = g(x, t, \rho) \text{ for some } \rho \in R\}$, $(x, t) \in A \times T$. Then condition (1.1) can be written as

$$\dot{x}(t) \in G(x(t), t) \qquad \text{a.e. in } T. \tag{1.1 Orig.}$$

The set $G(x, t)$ is thus the set of all "permissible" values of $\dot{x}$ while passing through the point x at the time t.

We shall refer to the above variational problem as the "*original problem.*" We shall also introduce the associated "*relaxed problem*" which we define as follows:

Let $F(x, t)$, $(x, t) \in A \times T$, be the convex closure[2] of $G(x, t)$.The relaxed

* Partly supported by Contract AF04(647)-305.

[1] The more general problem, in which t_0 and t_1 are not fixed but only restricted by $(t_0, t_1) \in Z \subset T \times T$ can, by a simple transformation, be reduced to the form here considered.

[2] i.e., the convex hull of the closure.

problem consists in minimizing $x^1(t_1)$ subject to conditions (1.2), (1.3), and

$$\dot{x}(t) \in F(x(t), t), \qquad \text{a.e. in } T. \tag{1.1 Relaxed}$$

The problem is "relaxed" in the sense that the permissible set of choices of $\dot{x}(t)$ is enlarged from $G(x(t), t)$ to $F(x(t), t)$.

Let us designate any absolutely continuous vector function $x(t)$ satisfying conditions (1.1 Orig.) [resp. (1.1 Relaxed)], (1.2), and (1.3) as an "*original (resp. relaxed) admissible curve*"; an original (resp. relaxed) admissible curve which minimizes $x^1(t_1)$ will be referred to as an "*original (resp. relaxed) minimizing curve.*" We prove (Theorem 2.2) that for a large class of vector functions $g(x, t, \rho)$, any relaxed admissible curve can be uniformly approximated by curves satisfying the differental equations (1.1). We then show (Theorem 3.3) that, in essentially "bounded" problems, a relaxed minimizing curve exists.

Section IV deals with a "proper representation" $f(x, t, \sigma)$ of $F(x, t)$. It is a mapping of some set $S = \{\sigma\}$ onto $F(x, t)$ with properties specified in (4.0.1), (4.0.2), and (4.0.3). In a paper to follow we shall use such a proper representation to establish "constructive" necessary conditions for minimum in the relaxed problem in which $A = E_n$. Once a relaxed minimizing curve is determined, the construction of Theorem 2.2 can be carried out to approximate that curve by solutions of the differential equations (1.1).

The case for replacing the original problem with the corresponding relaxed problem is made even stronger, and acquires a practical as well as a theoretical significance, when we consider the almost trivial illustrative example discussed at the end of Section II. The original problem of that example admits an original minimizing curve $y(t)$ and the corresponding relaxed problem admits a relaxed minimizing curve $x(t)$. These curves do not coincide and $x^1(t_1) < y^1(t_1)$. The implications, bearing in mind Theorem 2.2, are clear.

Replacing solutions of (1.1 Orig.) with those of (1.1 Relaxed) may be considered as in the spirit of the "generalized curve" approach of Young [1, 2]. Young was able to establish the existence of and necessary conditions satisfied by a generalized curve minimizing the integral $\int_{t_0}^{t_1} f^1(x, \dot{x}, t)\, dt$ and pointed out various possible generalizations. McShane [3-5] extended Young's results to the Problem of Bolza and also studied conditions insuring the existence of an ordinary minimizing curve. Necessary conditions for minimum in the original problem were studied by Pontryagin, Boltyanskii, and Gamkrelidze [6, 7, 10] in the special case $A = E_n$, $x(t_0)$ fixed, $x(t_1)$ unrestricted, and additional results for the case $A \neq E_n$, $x(t_0)$ and $x(t_1)$ fixed, were announced by Gamkrelidze [9]. The Dynamic Programming approach was introduced and discussed by Bellman [11]. Linear problems (in which R is an r-dimensional euclidean cube or polyhedron and $g(x, t, \rho)$ is linear in both x and ρ, while $A = E_n$ and $x(t_0)$ and $x(t_1)$ are either fixed or unrestricted) were studied by Bushaw [12], LaSalle [13], Bellman *et al.* [14], Krasovskii [15], Gamkrelidze [8], and Pontryagin [6]. In these linear problems the original problem is at the same time relaxed.

The author wishes to express his appreciation to Drs. H. Furstenberg and S. Katz, with whom some of the ideas here contained were discussed and to Dr. F. W. Diederich, whose interest in optimized trajectories provided a stimulus for these investigations.

II. Relaxed Admissible Curves as Uniform Limits of Solutions of Eq. (1.1)

The following lemma is well known:

Lemma 2.1. *Let C in E_n be compact and convex, let $\tilde{T}$ be a measurable subset of T, $|\tilde{T}| \neq 0$,*[1] *and let $a(t) \in C$, $t \in \tilde{T}$. Assume that $(1/|\tilde{T}|) \int_{\tilde{T}} a(t)\, dt$ exists. Then*

$$b = \frac{1}{|\tilde{T}|} \int_{\tilde{T}} a(t)\, dt \in C.$$

Theorem 2.2. *Assume that there exist constants K and M, a finite or denumerable collection of disjoint (Lebesgue) measurable sets T_r, $T_r \subset T$, $r = 1, 2, \cdots$, $|\cup_r T_r| = t_1 - t_0$, and a function $\epsilon(h)$ converging to 0 with h such that*

$$|g(x, t, \rho) - g(x, \tau, \rho)| \leq \epsilon(t - \tau), \qquad (x, t, \rho) \in V \times T_r \times R,$$

$$(x, \tau, \rho) \in V \times T_r \times R, \qquad r = 1, 2, \cdots. \tag{2.2.1}$$

$$|g(x, t, \rho) - g(y, t, \rho)| \leq K |x - y|, \qquad (t, \rho) \in T \times R, \qquad x \in V, \qquad y \in V \tag{2.2.2}$$

$$|g(x, t, \rho)| \leq M, \qquad (x, t, \rho) \in V \times T \times R \tag{2.2.3}$$

Then every absolutely continuous curve $x(t)$, $t \in T$, satisfying relations (1.1 Relaxed) is the uniform limit of curves $y_N(t)$, $N = 1, 2, \cdots$, satisfying differential equations (1.1) and such that $y_N(t_0) = x(t_0)$, $N = 1, 2, \cdots$.

[1] Two vertical bars will be used to represent the Lebesgue measure of a subset of T, the euclidean length of a vector and the absolute value of a scalar.

PROOF. Let N be a fixed positive integer and let

$$h = \frac{t_1 - t_0}{N}$$

$$t_k = t_0 + kh, \qquad k = 0, 1, \cdots, N$$

$$T_{k,r} = T_r \cap (t_k, t_{k+1})$$

[where (t_k, t_{k+1}) is the open interval $t_k < t < t_{k+1}$],

$$k = 0, 1, \cdots, N-1; \qquad r = 1, 2, \cdots$$

$$a_{k,r} = \frac{1}{|T_{k,r}|} \int_{T_{k,r}} \dot{x}(t)\, dt, \qquad k = 0, 1, \cdots, N-1; \qquad r = 1, 2, \cdots,$$

whenever $|T_{k,r}| \neq 0$.

Since $\dot{x}(t)$ exists over a set T^* of measure $t_1 - t_0$ we may assume, without loss of generality, that $\dot{x}(t)$ exists for all $t \in \cup_r T_r$ (otherwise we would replace T_r by $T_r \cap T^*$). Since $\dot{x}(t) \in F(x(t), t)$, $t \in T^*$, and $F(x(t), t)$ is the convex closure of $G(x(t), t)$, it follows by a theorem of Caratheodory [16] that there exist points $\rho_j(t) \in R$ and nonnegative numbers $\alpha_j(t)$, $j = 1, \cdots, n+1$ such that

$$\sum_{j=1}^{n+1} \alpha_j(t) \equiv 1, \qquad t \in T^* \tag{2.2.4}$$

$$\left| \sum_{j=1}^{n+1} \alpha_j(t)\, g(x(t), t, \rho_j(t)) - \dot{x}(t) \right| \leq h, \qquad t \in T^*. \tag{2.2.5}$$

Let now $t_{k,r}$ be arbitrary points in $T_{k,r}$, $k = 0, 1, \cdots, N-1$; $r = 1, 2, \cdots$, if $|T_{k,r}| \neq 0$. We observe that

$$|x(t_{k,r}) - x(t)| \leq M\, |t_{k,r} - t| \leq Mh, \quad t \in T_{k,r}, \quad |T_{k,r}| \neq 0, \tag{2.2.6}$$

since $\dot{x}(t) \in F(x(t), t)$, hence, by (2.2.3), $|\dot{x}(t)| \leq M$ a.e. in T. Thus, by (2.2.1), (2.2.2), (2.2.4), (2.2.5), and (2.2.6), for all $t \in T_{k,r}$, $|T_{k,r}| \neq 0$,

$$\begin{aligned}
&\left| \sum_{j=1}^{n+1} \alpha_j(t)\, g(x(t_{k,r}), t_{k,r}, \rho_j(t)) - \dot{x}(t) \right| \\
&\leq \left| \sum_{j=1}^{n+1} \alpha_j(t)[g(x(t_{k,r}), t_{k,r}, \rho_j(t)) - g(x(t_{k,r}), t, \rho_j(t))] \right| \\
&+ \left| \sum_{j=1}^{n+1} \alpha_j(t)[g(x(t_{k,r}), t, \rho_j(t)) - g(x(t), t, \rho_j(t))] \right| \\
&+ \left| \sum_{j=1}^{n+1} \alpha_j(t)[g(x(t), t, \rho_j(t)) - \dot{x}(t)] \right| \leq \epsilon(h) + (KM+1)h.
\end{aligned}$$

Let $F \subset E_n$. We shall represent by $U(F, \delta)$ the union of all n-dimensional balls in E_n with centers in F and with radius δ. Clearly $U(F, \delta)$ is compact and convex if F has these properties.

It follows from the last inequality that

$$\dot{x}(t) \in U(F(x(t_{k,r}), t_{k,r}), \quad \epsilon(h) + (KM + 1)h), \qquad t \in T_{k,r}, \qquad | T_{k,r} | \neq 0,$$

and, by Lemma 2.1,

$$a_{k,r} \in U(F(x(t_{k,r}), t_{k,r}), \quad \epsilon(h) + (KM + 1)h) \qquad \text{if} \qquad | T_{k,r} | \neq 0. \quad (2.2.7)$$

Let $b_{k,r}$ be the point in $F(x(t_{k,r}), t_{k,r})$ nearest $a_{k,r}$. Since $F(x(t_{k,r}), t_{k,r})$ is nonempty, compact, and convex, $b_{k,r}$ exists and is unique. Since $F(x(t_{k,r}), t_{k,r})$ is the convex closure of $G(x(t_{k,r}), t_{k,r})$, there exist (again referring to Caratheodory's theorem [16]) $\rho_{k,r,j}$, $\alpha_{k,r,j}$, $j = 1, \cdots, n + 1$, such that

$$\alpha_{k,r,j} \geq 0, \qquad j = 1, \cdots, n + 1, \qquad \sum_{j=1}^{n+1} \alpha_{k,r,j} = 1$$

$$\left| b_{k,r} - \sum_{j=1}^{n+1} \alpha_{k,r,j}\, g(x(t_{k,r}), t_{k,r}, \rho_{k,r,j}) \right| \leq h$$

hence, by (2.2.7) and by the definition of $b_{k,r}$,

$$\left| a_{k,r} - \sum_{j=1}^{n+1} \alpha_{k,r,j}\, g(x(t_{k,r}), t_{k,r}, \rho_{k,r,j}) \right| \leq \epsilon(h) + (KM + 2)h,$$
$$| T_{k,r} | \neq 0 \qquad (2.2.8)$$

For all k, r such that $| T_{k,r} | \neq 0$, let points $t_{k,r,j}$, $j = 0, 1, \cdots, n + 1$, be defined by

$$t_{k,r,0} = t_k, \qquad t_{k,r,n+1} = t_{k+1}$$
$$t_k \leq t_{k,r,j} \leq t_{k,r,j+1} \leq t_{k+1}, \qquad j = 0, 1, \cdots, n$$
$$| T_{k,r} \cap [t_{k,r,j}, t_{k,r,j+1}] | = \alpha_{k,r,j} \, | T_{k,r} |, \qquad j = 0, 1, \cdots, n,$$

where $[t_{k,r,j}, t_{k,r,j+1}]$ is the closed interval.

It can be easily seen that such points $t_{k,r,j}$ exist.

Let

$$\rho_N(t) = \rho(t) = \rho_{k,r,j}, \qquad t \in T_{k,r}, \qquad t_{k,r,j} < t < t_{k,r,j+1},$$
$$r = 1, 2, \cdots; \qquad j = 0, 1, \cdots, n + 1. \qquad (2.2.9)$$

This relation defines $\rho(t)$ a.e. in T.

We observe that, for $| T_{k,r} | \neq 0$,

$$| T_{k,r} | \sum_{j=1}^{n+1} \alpha_{k,r,j}\, g(x(t_{k,r}), t_{k,r}, \rho_{k,r,j}) = \int_{T_{k,r}} g(x(t_{k,r}), t_{k,r}, \rho(t))\, dt,$$

hence, by (2.2.8),

$$\left| a_{k,r} - \frac{1}{| T_{k,r} |} \int_{T_{k,r}} g(x(t_{k,r}), t_{k,r}, \rho(t))\, dt \right| \leq \epsilon(h) + (KM + 2)h. \quad (2.2.10)$$

Consider now the system

$$\begin{aligned} \dot{y}_N = \dot{y} &= g(y, t, \rho(t)) \\ y(t_0) &= x(t_0). \end{aligned} \quad (2.2.11)$$

Since $\rho(t)$ is, by definition, a step function over every measurable set T_r, $r = 1, 2, \cdots$, assumptions (2.2.1) and (2.2.3) imply that $g(y, t, \rho(t))$ is, for every fixed $y \in V$, a measurable function bounded by M. By (2.2.2), $| g(x, t, \rho(t)) - g(y, t, \rho(t)) | \leq K | x - y |$ for all $x \in V$, $y \in V$, $t \in T$. It follows, by a well known existence theorem [17], that system (2.2.11) has a unique solution $y(t)$ which can be extended up to the boundary of $V \times T$. Let $\bar{t}$ be the largest value not exceeding t_1 such that $y(t)$ exists for all t, $t_0 \leq t \leq \bar{t}$ and let

$$u_N(t) = u(t) = y(t) - x(t), \qquad t_0 \leq t \leq \bar{t}.$$

Then, for $t_0 \leq t_k \leq t_{k+1} \leq \bar{t}$

$$\begin{aligned} u(t_{k+1}) &= u(t_k) + \int_{t_k}^{t_{k+1}} \{g(y(\tau), \tau, \rho(\tau)) - \dot{x}(\tau)\}\, d\tau \\ &= u(t_k) + \sum_{r=1}^{\infty} \int_{T_{k,r}} \{g(y(\tau), \tau, \rho(\tau)) - \dot{x}(\tau)\}\, d\tau \end{aligned} \quad (2.2.12)$$

Now, for $\tau \in T_{k,r}$, $| T_{k,r} | \neq 0$,

$$\begin{aligned} & g(y(\tau), \tau, \rho(\tau)) - \dot{x}(\tau) \\ &= \{g(x(t_{k,r}), t_{k,r}, \rho(\tau)) - \dot{x}(\tau)\} + \{g(y(t_{k,r}), t_{k,r}, \rho(\tau)) - g(x(t_{k,r}), t_{k,r}, \rho(\tau))\} \\ &\quad + \{g(y(\tau), t_{k,r}, \rho(\tau)) - g(y(t_{k,r}), t_{k,r}, \rho(\tau))\} \\ &\quad + \{g(y(\tau), \tau, \rho(\tau)) - g(y(\tau), t_{k,r}, \rho(\tau))\} \end{aligned} \quad (2.2.13)$$

hence, by (2.2.1), (2.2.2), (2.2.3), (2.2.10), and (2.2.11), for $|T_{k,r}| \neq 0$,

$$\left| \int_{T_{k,r}} \{g(y(\tau), \tau, \rho(\tau)) - \dot{x}(\tau)\} \, d\tau \right|$$

$$\leq 2 \, | T_{k,r} | \, \epsilon(h) + (KM + 2)h \cdot | T_{k,r} | + K \, | T_{k,r} | \, | u(t_{k,r}) |$$
$$+ K \int_{T_{k,r}} | y(\tau) - y(t_{k,r}) | \, d\tau$$
$$\leq 2 \, | T_{k,r} | \, \epsilon(h) + (KM + 2)h \cdot | T_{k,r} | + K \, | T_{k,r} | \, | u(t_{k,r}) |$$
$$+ KM \, | T_{k,r} | \, h. \tag{2.2.14}$$

But

$$| u(t_{k,r}) | = | y(t_{k,r}) - x(t_{k,r}) | \leq | y(t_k) - x(t_k) | + 2Mh,$$

hence, by (2.2.12), (2.2.13) and (2.2.14),

$$| u(t_{k+1}) | \leq | u(t_k) | + 2h\epsilon(h) + (KM + 2)h^2 + Kh \, | u(t_k) | + 3KMh^2$$
$$= (1 + Kh) \, | u(t_k) | + 2h\epsilon(h) + (4KM + 2)h^2, \qquad k = 0, 1, \cdots, N-1. \tag{2.2.15}$$

Since $u(t_0) = x(t_0) - y(t_0) = 0$, it follows easily from (2.2.15) that

$$| u(t_k) | \leq \frac{1}{K} \{(1 + Kh)^k - 1\} \{(4KM + 2)h + 2\epsilon(h)\}, \qquad k = 1, 2, \cdots, N$$

and since

$$| u(t) | \leq | u(t_k) | + 2Mh \qquad \text{if} \qquad | t - t_k | \leq h,$$

it follows easily that

$$| u_N(t) | = | u(t) | \leq c_1 \left[\frac{t_1 - t_0}{N} + \epsilon \left(\frac{t_1 - t_0}{N} \right) \right], \qquad t_0 \leq t \leq \bar{t}, \tag{2.2.16}$$

where

$$c_1 = \frac{1}{K} (4KM + 2M + 2) \{ e^{K(t_1 - t_0)} - 1 \}.$$

If $\bar{t} = t_1$, the theorem follows immediately. Otherwise, since V is an open set containing the continuous curve $x(t)$ defined over the closed interval T, there exists a $\delta > 0$ such that the closed set

$$Y = \bigcup_{t \in T} \{ y \in E_n \mid | y - x(t) | \leq \delta \}$$

belongs to the open set V.

Let now N_1 be such that

$$c_1 \left[\frac{t_1 - t_0}{N} + \epsilon\left(\frac{t_1 - t_0}{N}\right)\right] \leq \delta \qquad \text{for all} \qquad N \geq N_1.$$

Then it follows from (2.2.16) that

$$y_N(t) \in Y, \qquad t_0 \leq t \leq \bar{t} < t_1, \qquad N \geq N_1,$$

hence $y_N(t)$, $t_0 \leq t \leq \bar{t}$ is in the interior of $V \times T$ contradicting the existence theorem which states that $y(t)$ can be extended up to the boundary of $V \times T$. It follows that for all $N \geq N_1$, $y_N(t)$ is defined for $t_0 \leq t \leq t_1$ and the inequality (2.2.16) holds over T. The theorem now follows directly. As a corrollary of Theorem 2.2 we state

THEOREM 2.3. *Let $g(x, t, \rho)$ satisfy the assumptions of Theorem 2.2 and let $A = E_n$, $B = A_0 \times E_n$, where A_0 is a closed set in E_n. Assume that there exists an original minimizing curve $x(t)$. Then $x(t)$ is also a relaxed minimizing curve.*

Theorem 2.3 does not, in general, remain valid when the assumption $B = A_0 \times E_n$ is dropped. This can be demonstrated by the following counter-example:[4] let R be the real interval $-1 \leq \rho \leq 1$, $A = E_3$, $t_0 = 0$, $t_1 > 0$,

$$\begin{aligned}
\dot{x}^1 &= g^1(x, t, \rho) = (x^2)^2 - (\rho)^2 \\
\dot{x}^2 &= g^2(x, t, \rho) = \rho \\
\dot{x}^3 &= g^3(x, t, \rho) = (x^2)^4
\end{aligned}$$

and let the relation $(x(t_0), x(t_1)) \in B$ be defined by

$$x(0) = (0, 0, 0), \qquad x^3(t_1) = 0.$$

Then there exists an original minimizing curve $y(t) \equiv (0, 0, 0)$. In fact, it is the only original admissible curve. Indeed,

$$x^3(0) = x^3(t_1) = 0 \qquad \text{and} \qquad \dot{x}^3 = (x^2)^4 \geq 0$$

imply $x^2(t) \equiv 0$, hence $\rho(t) \equiv 0$.

Consider now a special relaxed admissible curve. The point

$$h(x) = \tfrac{1}{2} g(x, t, 1) + \tfrac{1}{2} g(x, t, -1) = ((x^2)^2 - 1, 0, (x^2)^4)$$

[4] A superscript will represent here a power only if it follows a parenthesis; thus $\rho \cdot \rho = (\rho)^2$, etc.

clearly belongs to $F(x, t)$. Let now $x(t)$ be the solution of

$$\dot{x} = h(x)$$
$$x(0) = (0, 0, 0).$$

We can easily verify that $x(t) = (-t, 0,0)$. Since $x^3(t_1) = 0$, this is a relaxed admissible curve. It is, in fact, a relaxed minimizing curve since $\dot{x}^1 = (x^2)^2 - (\rho)^2 \geq -1$ implies $x^1(t_1) \geq -t_1$.

Thus $-t_1 = x^1(t_1) < y^1(t_1) = 0$ and there exists an original minimizing curve which is not a relaxed minimizing curve.

We can approximate $x(t)$ by solutions of $\dot{x} = g(x, t, \rho)$, setting $x(0) = (0, 0, 0)$ and

$$\rho(t) = (-1)^k, \frac{kt_1}{N} \leq t < \frac{(k+1)t_1}{N}, k = 0, 1, \cdots, N-1,$$

and letting $N \to \infty$.

III. Existence of a Relaxed Minimizing Curve

Let $U(F, \delta)$, $F \subset E_n$, $\delta > 0$, be defined as in Theorem 2.2, that is, as the union of all n-dimensional balls in E_n with centers in F and radius δ. We shall say that $F(x, t)$ is "*quasi-continuous at* (x, t)" if there exists a function $\eta(\delta, x, t) > 0$ defined for positive δ and such that

$$| t - \tau | + | x - y | \leq \eta(\delta, x, t)$$

implies

$$F(y, \tau) \subset U(F(x, t), \delta).$$

Theorem 3.1. *Assume that there exists a measurable subset T' of T, $| T' | = t_1 - t_0$, such that $F(x, t)$ is quasi-continuous at (x, t) if $(x, t) \in A \times T'$ and assume that there exists a constant M such that $|f| \leq M$ if $f \in F(x, t)$, $(x, t) \in A \times T$. Then, for every compact set D, the collection of all relaxed admissible curves $x(t)$ such that $x(t) \in D$, $t \in T$, is sequentially compact in the topology of the uniform norm (that is, given any infinite sequence of curves in the collection, there exists a subsequence which converges uniformly to a curve in the collection).*

Proof. Let $x_j(t)$, $t \in T$, $j = 1, 2, \cdots$, be an infinite sequence of relaxed admissible curves contained in D. Since $\dot{x}_j(t) \in F(x_j(t), t)$ a.e. in T, it follows that $| \dot{x}_j(t) | \leq M$ a.e. in T. Since $x_j(t)$ are, by definition, absolutely continuous and since D is bounded, it follows that $x_j(t)$ are uniformly bounded and equicontinuous over T. Thus, by Arzela's theorem, there exists a subsequence [which we shall continue to designate by $x_j(t)$] which converges uniformly to a curve $x(t)$ in D which is Lipschitz-continuous with constant M, hence $\dot{x}(t)$ exists a.e. in T. Since the sets A and B are closed and $x_j(t)$, $j = 1, 2, \cdots$, satisfy conditions (1.2) and (1.3), so does $x(t)$. It remains to show that $\dot{x}(t) \in F(x(t), t)$ a.e. in T.

Let $\epsilon_j(t) = x(t) - x_j(t)$, $t \in T$, $j = 1, 2, \cdots$. For every $t \in T'$, $t < t_1$, for every $h > 0$, $t + h \in T$, and for every $j \geq 1$ we have

$$\begin{aligned}\frac{1}{h}\{x(t+h) - x(t)\} &= \frac{1}{h}\{x(t+h) - x_j(t+h)\} - \frac{1}{h}\{x(t) - x_j(t)\} \\ &\quad + \frac{1}{h}\{x_j(t+h) - x_j(t)\} \\ &= \frac{1}{h}\{\epsilon_j(t+h) - \epsilon_j(t)\} + \frac{1}{h}\int_t^{t+h} \dot{x}_j(\tau)\, d\tau. \qquad (3.1.1)\end{aligned}$$

By assumption, given any $t \in T'$, $t < t_1$, and any $\delta > 0$, there exists $\eta(\delta, t) = \eta(\delta, x(t), t) > 0$ such that

$$F(y, \tau) \subset U(F(x(t), t), \delta) \qquad (3.1.2)$$

provided

$$|t - \tau| + |x(t) - y| \leq \eta(\delta, t).$$

Since $\epsilon_j(\tau)$ converges uniformly to 0 for all $\tau \in T$, there exists an integer $j_0(h) = j_0(h, t, \delta)$ which is the smallest positive integer such that

$$|\epsilon_j(\tau)| \leq \mathrm{Min}\left[\tfrac{1}{4}\delta h, \tfrac{1}{2}\eta\left(\frac{\delta}{2}, t\right)\right] \text{ for all } \tau \in T \text{ and all } j \geq j_0(h). \qquad (3.1.3)$$

Clearly $j_0(h)$ is nondecreasing as $h \to 0$.

Let now

$$0 \leq h \leq \mathrm{Min}\left(\frac{1}{2(M+1)}\eta(\tfrac{1}{2}\delta, t), t_1 - t\right). \qquad (3.1.4)$$

We have, for all τ, $t \leq \tau \leq t + h$,

$$|x_j(\tau) - x(t)| \leq |x_j(\tau) - x_j(t)| + |\epsilon_j(t)| \qquad (3.1.5)$$

As previously observed, we have

$$|x_j(\tau) - x_j(t)| \leq M\,|\tau - t|,$$

hence, by (3.1.3), (3.1.4), and (3.1.5),

$$\begin{aligned} |\tau - t| + |x_j(\tau) - x(t)| &\leq |\tau - t| + M|\tau - t| + \tfrac{1}{2}\eta\left(\frac{\delta}{2}, t\right) \\ &\leq (M+1)h + \tfrac{1}{2}\eta\left(\frac{\delta}{2}, t\right) \\ &\leq \eta\left(\frac{\delta}{2}, t\right) \qquad \text{for all} \quad j > j_0(h). \end{aligned} \tag{3.1.6}$$

Thus, by (3.1.2) and (3.1.6)

$$F(x_j(\tau), \tau) \subset U\left(F(x(t), t), \frac{\delta}{2}\right), \qquad t \leq \tau \leq t + h.$$

Now $\dot{x}_j(\tau) \in F(x_j(\tau), \tau)$ a.e. in T and we can easily verify that $U(C, \epsilon)$ is compact and convex if C has these properties. It follows thus by Lemma 2.1 that

$$\frac{1}{h}\int_t^{t+h} \dot{x}_j(\tau)\, d\tau \in U\left(F(x(t), t), \frac{\delta}{2}\right), \qquad j \geq j_0(h). \tag{3.1.7}$$

Now, by (3.1.3),

$$\left|\frac{1}{h}\{\epsilon_j(t+h) - \epsilon_j(t)\}\right| \leq \frac{1}{h}|\epsilon_j(t+h)| + \frac{1}{h}|\epsilon_j(t)| \leq \frac{\delta}{2}, \qquad j \geq j_0(h)$$

hence, by (3.1.1) and (3.1.7),

$$\frac{1}{h}\{x(t+h) - x(t)\} \in U(F(x(t), t), \delta) \tag{3.1.8}$$

for any $t \in T'$ and any $\delta > 0$ provided h satisfies relation (3.1.4).

Since $x(t)$ is absolutely continuous, its derivative $\dot{x}(t)$ exists a.e. in T and

$$\dot{x}(t) = \lim_{h \to 0} \frac{1}{h}\{x(t+h) - x(t)\} \in U(F(x(t), t), \delta)$$

for every $\delta > 0$ and for almost all $t \in T'$.

It follows that there exists a set T^* of measure $t_1 - t_0$ and contained in T such that

$$\dot{x}(t) \in F(x(t), t), \qquad t \in T^*.$$

This completes the proof of the theorem.

THEOREM 3.2. *Assume that there exists a constant M and a subset T' of T of (Lebesgue) measure $t_1 - t_0$ such that*

$$| g(x, t, \rho) | \leq M, \qquad (x, t, \rho) \in A \times T \times R,$$

and that $g(x, t, \rho)$ is continuous in (x, t) uniformly in ρ for $(x, t) \in A \times T'$. Then $F(x, t)$ satisfies the assumptions of Theorem 3.1.

PROOF. By assumption, if $g \in G(x, t)$ and $(x, t) \in A \times T$, then $| g | \leq M$. Thus, if $\bar{g}$ is in the closure of $G(x, t)$, then $| \bar{g} | \leq M$. It follows that every $f' \in F(x, t)$ is such that $| f' | \leq M$. Let now δ be any positive number and let $(x, t) \in A \times T'$. Then there exists $\eta = \eta(\delta, x, t) > 0$ such that, for every $\rho \in R$,

$$| g(y, \tau, \rho) - g(x, t, \rho) | \leq \frac{\delta}{2} \tag{3.2.1}$$

provided

$$| t - \tau | + | x - y | \leq \eta. \tag{3.2.2}$$

Let now f be any point of $F(y, \tau)$, where y and τ satisfy (3.2.2). Then, by the theorem of Caratheodory [16], there exist points $\rho_j \in R, j = 1, \cdots, n + 1$, and nonnegative numbers α_j, such that

$$\sum_{j=1}^{n+1} \alpha_j = 1,$$

and

$$\left| \sum_{j=1}^{n+1} \alpha_j g(y, \tau, \rho_j) - f \right| \leq \frac{\delta}{2}. \tag{3.2.3}$$

From (3.2.1) and (3.2.3), it follows that

$$\left| \sum_{j=1}^{n+1} \alpha_j g(x, t, \rho_j) - f \right| \leq \delta.$$

Clearly

$$\sum_{j=1}^{n+1} \alpha_j g(x, t, \rho_j) \in F(x, t).$$

It follows that $f \in U(F(x(t), t), \delta)$. Since f is an arbitrary point of $F(y, \tau)$, we conclude that $F(x, t)$ is quasi-continuous at (x, t) for all $(x, t) \in A \times T'$.

As a corollary of Theorems 3.1 and 3.2 we deduce

THEOREM 3.3. *If there exists a constant M, a compact set $D \subset A$, a subset T' of T of measure $t_1 - t_0$ and a relaxed admissible curve $y(t)$ such that*

$$| g(x, t, \rho) | \leq M, \qquad (x, t, \rho) \in D \times T \times R \tag{3.3.1}$$

$g(x, t, \rho)$ is continuous in (x, t) uniformly in ρ for all $(x, t) \in D \times T'$ (3.3.2)

$$x^1(t_1) < y^1(t_1) \textit{ implies } x(t) \in D,\ t \in T,$$
$$\textit{for every relaxed admissible curve } x(t), \tag{3.3.3}$$

then there exists a relaxed minimizing curve.

In particular, if the set A is bounded, it suffices to verify only conditions (3.3.1) and (3.3.2), setting $D = A$.

IV. PROPER REPRESENTATION OF $F(x, t)$. THE YOUNG REPRESENTATION

The vector function $g(x, t, \rho)$ provides a parametric representation of the set $G(x, t)$ in the sense that if maps R onto $G(x, t)$. An analogous representation of $F(x, t)$, with certain special properties, will later [18] prove useful in establishing "constructive" necessary conditions for minimum in a relaxed problem.

Specifically, we shall say that a vector function $f(x, t, \sigma)$ in E_n, where σ belongs to some abstract set S, is a *proper representation* of $F(x, t)$ if

(4.0.1) $F(x, t) = \{ f \mid f = f(x, t, \sigma) \text{ for some } \sigma \in S\}$, $(x, t) \in V \times T$.

(4.0.2) For every absolutely continuous curve $x(t)$ satisfying relation (1.1 Relaxed) there exists a function $\sigma(t) \in S$, $t \in T$, such that $\dot{x}(t) = f(x(t), t, \sigma(t))$, a.e. in T and $f(x, t, \sigma(\tau))$ is, for all $x \in V$ and almost all $t \in T$, a (Lebesgue) measurable function of τ, $\tau \in T$.

(4.0.3) There exists a set $T', T' \subset T$, $| T' | = t_1 - t_0$, and constants M and K such that

(4.0.3.1) $f^i(x, t, \sigma)$ and $\partial f^i(x, t, \sigma)/\partial x^j$, $i, j = 1, \cdots, n$, exist and are continuous functions of (x, t) for all $(x, t, \sigma) \in V \times T' \times S$,

(4.0.3.2) $\qquad | f(x, t, \sigma) | \leq M, \qquad (x, t, \sigma) \in V \times T \times S$

(4.0.3.3) $\left| \dfrac{\partial f^i(x, t, \sigma)}{\partial x^j} \right| \leq K, \qquad (x, t, \sigma) \in V \times T' \times S, \qquad i, j = 1, \cdots, n$

Since $F(x, t)$ is compact and convex, such a proper representation can usually be constructed as a mapping of a k-dimensional ball $S(k \leq n)$ in E_n onto $F(x, t)$ provided $g(x, t, \rho)$ is sufficiently "nice." In the general case, however, we shall find it more convenient to construct such a representation as a mapping of a class S of probability measures over R onto $F(x, t)$. This second approach is patterned after Young's definition of generalized curves [1].

Until now the set R could be assumed arbitrary. Let us now assume that

(4.0.4) we can define R as a compact Hausdorff space in such a manner that $g(x, t, \rho)$ is a continuous function of ρ for each $(x, t) \in V \times T$.

We may then define measurable sets in R as elements of the smallest Borel field of sets containing open sets in R. We then define S to be the class of probability measures over R; thus $\sigma \in S$ if $\sigma(D)$ is defined for every measurable set D in R, $\sigma(D) \geq 0$ and $\sigma(R) = 1$.

Let $h(x, t, \sigma) = \int g(x, t, \rho)\, \sigma(d\rho)$. We shall refer to $h(x, t, \sigma)$ as the "*Young representation*" of $F(x, t)$.

We prove

THEOREM 4.1. *Let $g(x, t, \rho)$ satisfy the assumptions of Theorem 2.2 and assumption (4.0.4). Assume, furthermore, that $g^i(x, t, \rho)$ and $\partial g^i(x, t, \rho)/\partial x^j$, $i, j = 1, \cdots, n$, exist and are continuous functions of (x, t) uniformly in ρ for $(x, t, \rho) \in V \times T' \times R$, where $T' \subset T$, $| T' | = t_1 - t_0$. Then the Young representation $h(x, t, \sigma)$ is a proper representation of $F(x, t)$.*

PROOF. Let $\mathscr{A}$ be the Banach space with elements $\phi(\rho)$ which are real-valued continuous functions of ρ with the norm $\| \phi \| = \mathrm{Max}_\rho\, | \phi(\rho) |$. Let $\mathscr{B}$ be the smallest Banach subspace of $\mathscr{A}$ (with the same norm) which includes the elements $g^i(x, t, \rho)$, $i = 1, 2, \cdots, n$; $x \in V$, $t \in T'' = \cup_r (T' \cap T_r)$, where T_r are the sets referred to in Theorem 2.2. Assumptions (2.2.1) and (2.2.2) imply that each element of $\mathscr{B}$ can be uniformly approximated by linear combinations with rational coefficients of the functions $g^i(x_j, t_k, \rho)$, $j = 1, 2, \cdots$; $k = 1, 2, \cdots$, where (x_j, t_k) are points of a dense subset of $V \times T''$. Thus the space $\mathscr{B}$ is separable.

We shall now define a linear functional $k_N(\phi, \tau)$, where $\phi \in \mathscr{B}$, $\tau \in T$, as follows: let $x(t)$ be an absolutely continuous curve satisfying relation (1.1 Relaxed) and let $\rho_N(t)$ be defined as in (2.2.9). Then let

$$k_N(\phi; \tau) = \int_{t_0}^{\tau} \phi(\rho_N(\theta))\, d\theta, \qquad N = 1, 2, \cdots; \qquad \phi \in \mathscr{B}, \qquad \tau \in T. \quad (4.1.1)$$

The linear functional $k_N(\phi; \tau)$ exists for every τ since $\phi(\rho_N(\theta))$ is, by definition of $\rho_N(\theta)$, a step function over each measurable T_r. We have

$$| k_N(\phi; \tau) | \leq (t_1 - t_0)\, \| \phi \|,$$

hence $k_N(\phi; \tau)$ has, for each N and each $\tau \in T$, a norm not exceeding $t_1 - t_0$.

Furthermore,

$$| k_N(\phi; \tau_1) - k_N(\phi; \tau_2) | \leq || \phi || \, | \tau_1 - \tau_2 |,$$

hence $k_N(\phi; \tau)$ is, for each $\phi \in \mathscr{B}$, a Lipschitz-continuous function of τ with Lipschitz constant $\| \phi \|$. Thus, remembering that both $\mathscr{B}$ and T are separable, we may, using the conventional "diagonal subsequence" argument as in Arzela's theorem, find a subsequence of $k_N(\phi; \tau)$, $N = 1, 2, \cdots$, [which we shall continue to designate as $k_N(\phi; \tau)$, $N = 1, 2, \cdots$] which converges to some $k(\phi; \tau)$ for each $\phi \in \mathscr{B}$, $\tau \in T$. Since, for every $\phi \in \mathscr{B}$, $k_N(\phi; \tau)$ are Lipschitz continuous with constant $\| \phi \|$, the limit $k(\phi; \tau)$ has the same property. It follows that, for each $\phi \in \mathscr{B}$, $k(\phi; \tau)$ is differentiable a.e. in T, say for $\tau \in T(\phi)$. We may assume $T(\phi) \subset T''$ [otherwise replacing $T(\phi)$ by $T'' \cap T(\phi)$] and we have $| T(\phi) | = t_1 - t_0$. Let

$$l(\phi; \tau) = \frac{d}{d\tau} k(\phi; \tau), \qquad \tau \in T(\phi)$$

and let ϕ_j, $j = 1, 2, \cdots$, be a dense subset of the separable Banach space $\mathscr{B}$. We have

$$l(\phi_j; \tau) = \frac{d}{d\tau} k(\phi_j; \tau), \qquad j = 1, 2, \cdots, \qquad \tau \in T''' = \bigcap_{j=1}^{\infty} T(\phi_j) \tag{4.1.2}$$

and $| T''' | = t_1 - t_0$. Thus $l(\phi; \tau)$ is defined, for each $\tau \in T'''$, over a dense subset of $\mathscr{B}$. Since $k(\phi; \tau)$ is, for each $\phi \in \mathscr{B}$, Lipschitz continuous with constant $\| \phi \|$, it follows that $l(\phi; \tau)$ is, for each $\tau \in T'''$, a bounded linear functional over a dense subset of $\mathscr{B}$ with norm not exceeding 1. It can thus be extended to the entire space $\mathscr{B}$ and, by the Hahn-Banach theorem, can be further extended to the space $\mathscr{A}$.

It follows then, by the Riesz representation theorem, that there exists, for each $\tau \in T'''$, a measure σ_τ defined over measurable sets of R such that

$$l(\phi; \tau) = \int \phi(\rho) \, \sigma_\tau(d\rho), \qquad \tau \in T'''. \tag{4.1.3}$$

We may assume that the dense subset of $\mathscr{B}$ with elements $\phi_j, j = 1, 2, \cdots$, included, in particular, the functions $g^i(x_j, t_k, \rho)$, $i = 1, 2, \cdots, n; j = 1, 2, \cdots$, $k = 1, 2, \cdots$, where (x_j, t_k) are points of a dense subset of $V \times T'''$. It follows then, by (4.1.2) and (4.1.3) that

$$\int g^i(x_j, t_k, \rho) \, \sigma_\tau(d\rho), \qquad i = 1, 2, \cdots, n; \qquad j, k = 1, 2, \cdots \tag{4.1.4}$$

are measurable functions of τ, being defined as derivatives of absolutely continuous functions.

We can, furthermore, easily verify that, as a direct consequence of its definition, $l(\phi; \tau) \geq 0$ if $\phi(\rho) \geq 0$. Thus it follows from (4.1.3) that $\sigma_\tau \geq 0$. Setting $\phi(\rho) \equiv 1$, we find that $\sigma_\tau(R) = 1$, $\tau \in T'''$. Thus σ_τ is a probability measure. By (2.2.1) and (2.2.2), $h^i(x, t, \sigma_\tau) = \int g^i(x, t, \rho)\, \sigma_\tau(d\rho)$, $i = 1, \cdots, n$, is, for every $\tau \in T'''$, $(x, t) \in V \times T'''$, a limit of expressions of the form (4.1.4), hence is a measurable function of τ.

We can now prove that the Young representation $h(x, t, \sigma)$ is a proper representation of $F(x, t)$. Let $y_N(t)$, $t \in T$, be defined as in (2.2.11), N being restricted to the subsequence over which $k_N(\phi; \tau)$ converges to $k(\phi; \tau)$. By Theorem 2.2,

$$\lim_{N\to\infty} y_N(t) = x(t), \qquad t \in T.$$

Let $t \in T'''$. Then, for $\alpha > 0$ and $t + \alpha \leq t_1$,

$$y_N(t + \alpha) - y_N(t) = \int_t^{t+\alpha} g(y_N(\theta), \theta, \rho_N(\theta))\, d\theta,$$

and, passing to the limit as $N \to \infty$, we have, by (2.2.2),

$$x(t + \alpha) - x(t) = \lim_{N\to\infty} \int_t^{t+\alpha} g(x(\theta), \theta, \rho_N(\theta))\, d\theta. \tag{4.1.5}$$

Since $g(x, t, \rho)$ is continuous in (x, t) uniformly in ρ for all $(x, t) \in V \times T'''$ and since $x(t)$ is uniformly continuous, there exists, for every $\epsilon > 0$, a number $\eta(\epsilon) = \eta(\epsilon, t) > 0$ such that

$$| g(x(\theta), \theta, \rho_N(\theta)) - g(x(t), t, \rho_N(\theta)) | \leq \epsilon, \qquad N = 1, 2, \cdots \tag{4.1.6}$$

provided

$$| \theta - t | \leq \eta(\epsilon).$$

Thus, by (4.1.5) and (4.1.6),

$$\left| x(t + \alpha) - x(t) - \lim_{N\to\infty} \int_t^{t+\alpha} g(x(t), t, \rho_N(\theta))\, d\theta \right| \leq \alpha\epsilon \tag{4.1.7}$$

for $\alpha \leq \eta(\epsilon)$. Now

$$\lim_{N\to\infty} \int_t^{t+\alpha} g^i(x(t), t, \rho_N(\theta))\, d\theta = k(g^i(x(t), t, \rho); t + \alpha) - k(g^i(x(t), t, \rho); t) \tag{4.1.8}$$

and

$$\lim_{\alpha\to 0} \frac{1}{\alpha} \{k(g^i(x(t), t, \rho); t + \alpha) - k(g^i(x(t), t, \rho); t)\} = l(g^i(x(t), t, \rho); t)$$

$$= h^i(x(t), t, \sigma_t) \qquad \text{a.e. in } T. \tag{4.1.9}$$

Also,

$$\lim_{\alpha \to 0} \frac{1}{\alpha} \{x(t + \alpha) - x(t)\} = \dot{x}(t) \qquad \text{a.e. in } T.$$

Dividing (4.1.7) by α and passing to the limit as $\alpha \to 0$, it follows, by (4.1.8) and (4.1.9) that, for every $\epsilon > 0$,

$$| \dot{x}(t) - h(x(t), t, \sigma_t) | \leq \epsilon \qquad \text{a.e. in } T'''$$

hence

$$\dot{x}(t) = h(x(t), t, \sigma_t) \qquad \text{a.e. in } T.$$

This shows that the Young representation $h(x, t, \sigma)$ has property (4.0.2), since we have previously shown that $h(x, t, \sigma_\tau)$ is a measurable function of τ for all $x \in V$ and almost all $t \in T$.

Property (4.0.3) for $h(x, t, \sigma)$ easily follows from its definition. It remains to show that $h(x, t, \sigma)$ maps S onto $F(x, t)$.

Obviously, a measure σ concentrated at one point ρ yields $h(x, t, \sigma) = g(x, t, \rho)$, hence $G(x, t)$ is contained in the map $H(x, t)$ of S under $h(x, t, \sigma)$. Since $h(x, t, \sigma)$ is linear in σ and the class S of probability measures is convex, it follows that the convex hull of $G(x, t)$ is contained in $H(x, t)$. Now $G(x, t)$ is closed, since $g(x, t, \rho)$ is continuous over the compact space R, hence $F(x, t) \subset H(x, t)$. It is also easy to show, by an argument analogous to that of Lemma 2.1, that $H(x, t) \subset F(x, t)$, hence $H(x, t) = F(x, t)$.

This completes the proof of the theorem.

References

1. Young, L. C. Generalized curves and the existence of an attained absolute minimum in the calculus of variations. *Compt. rend. soc. sci. et lettres Varsovie, Cl. III.* **30**, 212-234 (1937).
2. Young, L. C. Necessary conditions in the calculus of variations. *Acta Math.* **69**, 239-258 (1938).
3. McShane, E. J. Generalized curves. *Duke Math. J.* **6**, 513-536 (1940).
4. McShane, E. J. Necessary conditions in generalized-curve problems of the calculus of variations. *Duke Math. J.* **7**, 1-27, 1940.
5. McShane, E. J. Existence theorems for Bolza problems in the calculus of variations. *Duke Math.* **7**, 28-61 (1940).
6. Pontryagin, L. S. Optimum control processes. *Uspekhi Mat. Nauk.* **14**, No. 1 (85) (1959). (Associated Technical Services, Inc. Translation RJ-2192.)
7. Boltyanskii, V. G., Gamkrelidze, R. V. and Pontryagin, L. S. The theory of optimum processes. *Doklady Akad. Nauk SSSR.* **110**, No. 1, 7-10 (1956).
8. Gamkrelidze, R. V. The theory of optimum processes in linear systems. *Doklady Akad. Nauk.* **116**, No. 1, 9-11 (1957).
9. Gamkrelidze, R. V. Optimum-rate processes with bounded phase coordinates. *Doklady Akad. Nauk SSSR.* **125**, No. 3, 475-478 (1959).

10. Boltyanskii, V. G. The maximum principle in the theory of optimum processes. *Doklady Akad. Nauk SSSR.* **119**, No. 6 (1958).
11. Bellman, R. "Dynamic Programming," Chap. IX. Princeton Univ. Press. Princeton, New Jersey, 1957.
12. Bushaw, D. W. Optimal discontinuous forcing terms, *in* "Contributions to the Theory of Nonlinear Oscillations, IV." Princeton Univ. Press. Princeton, New Jersey, 1958.
13. LaSalle, J. P. Time optimal control systems. *Proc. Natl. Acad. Sci. U.S.* **45**, 573-577 (1959).
14. Bellman, R., Glicksberg I. and Gross O. On the "bang-bang" control problem. *Quart. Appl. Math.* **14**, 11-18 (1956).
15. Krasovskii, N. Concerning the theory of optimal control. *Avtomat i Telemekh.* **18**, 960-970 (1957).
16. Eggleston, H. G. "Convexity," pp. 34-38. Cambridge Tracts in Mathematics and Mathematical Physics. No. 47, Cambridge Univ. Press, 1958.
17. McShane, E. J. "Integration," pp. 344-345. Princeton Univ. Press. Princeton, New Jersey, 1944.
18. Warga, J. Necessary conditions for minimum in relaxed variational problems. *J. Math. Anal. and Appl.* **4**, 129-145 (1962).