AN IMPLICIT FUNCTION THEOREM WITH AN APPLICATION TO CONTROL THEORY

Richard Datko

Filippov [1] and Roxin [2] have proved the existence of time optimal controls for systems of the form $\dot{x} = f(x, t, u)$, where f is assumed to satisfy a convexity condition with respect to u for fixed (x, t) and u is constrained to lie in some compact set in R^m. Both results depend on the establishment of the following type of theorem: Suppose $f: R^1 \times R^m \to R^n$ is a continuous mapping and $\{\phi_n(t)\} = \{f(t, u_n(t))\}$ is an infinite sequence of measurable mappings from $R^1 \to R^n$, where for each n, u_n is a measurable mapping from R^1 into some compact set A in R^m, and suppose there exists a measurable map $\phi: R^1 \to R^n$ such that on some bounded interval $\lim_{n\to\infty} \phi_n(t) = \phi(t)$ (a.e.). Then there exists a measurable mapping $u: R^1 \to A$ such that $\phi(t) = f(t, u(t))$ (a.e.).

In this paper we prove an analogous result in which we dispense to some extent with the demand that $u(t)$ lie in a compact set in R^m. The method of proof can be used to obtain both Filippov's and Roxin's results. As an application we prove a general existence theorem for nonlinear time optimal controls.

THEOREM 1. *Let* $f: I \times R^m \to R^n$ *be a continuous mapping, where* I *is the interval* $[0, 1)$ *in* R^+, *let* $\Phi: I \to R^n$ *be a measurable mapping, and suppase* $(K_t)_{t \in I}$ *is an expanding family of compact sets in* R^m *such that* $\Phi(t)$ *is in the set* $f(\{t\} \times K_t)$ *for each* $t \in I$. *Then there exists a measurable mapping* $u: I \to R^m$ *such that* $u(t)$ *is in* K_t *for each* t *in* I *and* $f(t, u(t)) = \Phi(t)$ a.e. *in* I.

Remark. Filippov, in his lemma in [1], assumes I to be compact in R^+ and K_t to be compact in R^m and upper-semicontinuous with respect to t in I. By this he means that to each $\varepsilon > 0$ there corresponds a $\delta(t_0, \varepsilon) > 0$ such that K_t lies in $S(K_{t_0}, \varepsilon)$ (that is, K_t lies in an ε-neighborhood of K_{t_0}) for every t in I with $|t - t_0| < \delta$.

Proof of the theorem. We can find a mapping $c: I \to R^m$, which is not necessarily measurable, such that $c(t)$ is in K_t and $f(t, c(t)) = \Phi(t)$ a.e. in I. To see this we define, for each t in I,

$$L_t = \{c \in K_t \mid f(t, c) = \Phi(t)\}.$$

Since f is continuous in $I \times R^m$, L_t is closed and hence compact for t in I. Let

$$c_1(t) = \inf\{c_1 \mid (c_1, \cdots, c_m) \in L_t\},$$

and define successively, for $1 \leq i \leq m - 1$,

$$c_{i+1}(t) = \inf\{c_{i+1} \mid (c_1(t), \cdots, c_i(t), c_{i+1}, \cdots, c_m) \in L_t\}.$$

Received February 24, 1964.
This research was supported by the U. S. Army Research Office (Durham) under Contract DA-36-034-AMC-0221X.

Then $(c_1(t), \cdots, c_m(t))$ is a unique point in L_t, and hence $t \to c(t)$ is a mapping of I into R^m for which $c(t)$ is in K_t and $f(t, c(t)) = \Phi(t)$ in I.

Let $0 < \varepsilon < 1$, put $T = 1 - \varepsilon$, and let $J = [0, T]$. Note that, by assumption, $\{t\} \times K_t$ is in $J \times K_T$ for every t in J; hence

$$\Phi(t) \in f(\{t\} \times K_t) \in f(J \times K_T)$$

for each t in J. Therefore $t \to \|\Phi(t)\|$ (where $\|\ \|$ is any convenient norm) is bounded in J. This implies, by Lusin's theorem (see [3] for example), that there exists a closed set A in J such that the measure of J - A is less than ε and that the restriction of Φ to A is continuous. Clearly the measure of I - A is less than 2ε. Since A and K_T are compact sets in R^1 and R^m, respectively, it follows that Φ and f are uniformly continuous in A and $A \times K_T$, respectively. Thus given any $h > 0$, there exists a $\delta(h) > 0$ such that, for each t and t^1 in A with $|t - t^1| < \delta$, and for each c in K_T,

$$\|\Phi(t) - \Phi(t^1)\| < \frac{h}{2}, \qquad \|f(t, c) - f(t^1, c)\| < \frac{h}{2}. \tag{1.1}$$

For each $n = 1, 2, \cdots$, define

$$J_i^n = [a_{i-1}^n, a_i^n),$$

where

$$a_0^n = 0, \qquad a_i^n = i2^{-n} \quad (i = 1, 2, \cdots, 2^n),$$

so that

$$\bigcup_{i=1}^{2^n} J_i^n = I \qquad \text{and} \qquad J_i^n \cap J_k^n = \emptyset \quad (i \neq k).$$

Let $t_i^n = \inf J_i^n \cap A$, and define $u^n(t) = c(t_i^n)$ for every t in $J_i^n \cap A$. Then each u^n is a measurable mapping of $A \to R^n$. Let $f^n(t) = f(t, u^n(t))$ for each t in A. We claim that

$$\lim_{n \to \infty} f^n(t) = \Phi(t)$$

uniformly in A. Indeed if t is in A, then t is in $J_{i(n)}^n \cap A$ for $n = 1, 2, \cdots$, and

$$\|f^n(t) - \Phi(t)\| \leq \|f(t, c(t_{i(n)}^n)) - f(t_{i(n)}^n, c(t_{i(n)}^n))\| + \|\Phi(t_{i(n)}^n) - \Phi(t)\|, \tag{1.2}$$

by the way we have constructed the values of u^n. Given any $h > 0$, we may therefore choose an $n_0(h) > 0$ such that $2^{-n} < \delta(h)$ for every $n \geq n_0$. From (1.1) and (1.2) we deduce that

$$\|f^n(t) - \phi(t)\| < \delta \tag{1.3}$$

for $n \geq n_0$ and each t in A.

Define, for t in A,

$$u_1(t) = \limsup_{n \to \infty} u_1^n(t), \tag{1.4}$$

and set $u_1(t) = 0$ for $t \notin I - A$. By construction, u_1 is a real-valued measurable function in I. We now construct a mapping $\bar{u}$ of $A \to R^m$ such that, for each t in A, $\bar{u}_1(t) = u_1(t)$, $\bar{u}(t)$ is in K_t, and $f(t, \bar{u}(t)) = \Phi(t)$. For each $t \in A$, there is a subsequence of $\{u_1^n(t)\}$, again denoted by $\{u^n(t)\}$, with the properties that

$$\lim_{n \to \infty} u_1^n(t) = u_1(t), \qquad \lim_{n \to \infty} u_k^n(t) = \bar{u}_k(t)$$

for $k = 2, \cdots, m$, and

$$(u_1(t), \bar{u}_2(t), \cdots, \bar{u}_m(t)) \in K_t.$$

The continuity of f in $I \times R^m$ and (1.3) imply that $f(t, \bar{u}(t)) = \Phi(t)$ for each t in A.

Since u_1 is a real-valued measurable function on I that is bounded in A, we can again apply Lusin's theorem and find a closed subset A_1 of A such that the measure of $A - A_1$ is less than ε and the restriction of u_1 to A_1 is continuous and hence, by compactness of A_1, uniformly continuous. Note that the measure of $I - A_1$ is less than 3ε. Therefore, we can repeat the above construction and obtain a new mapping $\bar{u}$ of A_1 into R^m for which $\bar{u}_1$ coincides with the restriction of u_1 to A_1, where u_1 is defined by (1.4), $\bar{u}_2$ is a real-valued measurable function in I, $\bar{u}(t)$ is in K_t, and $f(t, \bar{u}(t)) = \Phi(t)$ for each t in A_1.

It follows that after m such steps we shall have constructed a measurable mapping u of a closed set A_{m-1} in I into R^m such that the measure of $I - A_{m-1}$ is less than $(m + 2)\varepsilon$, u(t) is in K_t, and $f(t_1, u(t)) = \Phi(t)$ for every t in A_{m-1}.

In the above construction ε is arbitrary; therefore, given a null sequence $\{\varepsilon_n\}$ of positive constants, we can find a sequence $\{F_n\}$ of closed subsets in I and a sequence $\{u_n\}$ of measurable mappings of $\{F_n\}$ in R^n with the following properties:

(i) $\mu(I - F_n) < \varepsilon_n$ (μ is Lebesgue measure),

(ii) $u_n(t) \in K_t$ for each t in F_n,

(iii) $f(t, u_n(t)) = \Phi(t)$ for each t in F_n.

Let $E = \bigcup_{n=1}^{\infty} F_n$. We assert that the measure of $I - E$ is equal to zero. Suppose that, to the contrary,

$$\mu(I - E) = \alpha > 0.$$

Since F_n is in E and $F_n \cap (I - E) = \emptyset$, we see that

$$\begin{aligned}\mu(I) &\geq \mu(F_n) + \mu(I - E) = \mu(I) - \mu(I - F_n) + \mu(I - E) \\ &> \mu(I) - \varepsilon_n + \alpha \\ &> \mu(I)\end{aligned}$$

as soon as $\varepsilon_n < \alpha$, which is a contradiction.

Define

$$\overline{F}_1 = F_1, \quad \overline{F}_n = F_n - \bigcup_{j=1}^{n-1} F_j \quad (n = 2, 3, \cdots),$$

so that $E = \bigcup_{n=1}^{\infty} \overline{F}_n$. Let U be a mapping of E into R^n, whose restriction to $\overline{F}_n$ coincides with u_n and for which $U(t) \in K_t$ for each $t \in I - E$. Then U is a measurable mapping of I into R^n with all the properties claimed in the theorem.

COROLLARY 1.1. *Let* $I = [0, T)$ *be an arbitrary interval in* R^1, *let*

$$f: I \times R^m \to R^n$$

be a continuous mapping, and let $\Phi: I \to R^n$ *be a measurable mapping. Suppose there exists a partition* (I_k) $(1 \leq k \leq p)$ *of* I *into intervals* $I_k = [T_{k-1}, T_k)$ *and a family* $\{K_t\}$ $(t \in I)$ *of sets in* R^m *such that, in each open interval* (T_{k-1}, T_k), $\{K_t\}$ *is an expanding or contracting family of compact sets, and* $\Phi(t)$ *is in* $f(\{t\} \times K_t)$ *for every* t *in* $\bigcup_{k=1}^{p} (T_{k-1}, T_k)$. *Then there exists a measurable mapping* $u: I \to R^m$ *such that* u(t) *is in* K_t *for each* t *in* I *and* $f(t, u(t)) = \Phi(t)$ a.e. *on* I.

Proof. It is clear that the construction of Theorem 1.1 carries over verbatim to the case where the family $\{K_t\}$ is expanding in some interval I_k. If the family is contracting, we may introduce the intervals

$$J_i^n = (a_{i-1}^n, a_i^n],$$

where $a_0^n = T_{k-1}$, $a_i^n = T_{k-1} + (T_k - T_{k-1}) i 2^{-n}$ for $i = 1, 2, \cdots, 2^n$, and define the constants t_i^n by a supremum rather than an infimum.

Remark. The proof of Theorem 1.1 can readily be simplified to yield the original Filippov lemma, if the family $\{K_t\}$ $(t \in I)$ is upper-semicontinuous with respect to t in I. Observe that here I is not required to be compact, as in Filippov's lemma.

We shall apply our result to a problem in control theory. First we state three assumptions.

1. There exists a family of sets $\{K_t\}$ $(t \in R^+)$ and a partition $\{J_K\}$ of R^+, with $J_K = [t_{k-1}, t_k)$, such that $\{K_t\}$ is either an expanding or contracting family of compact sets in each open interval (t_{k-1}, t_k).

2. U is the family of all measurable mappings of R^+ into R^m with the property that u(t) is in K_t for each t in R^+.

3. $f: R^n \times R^+ \times R^m \to R^n$ is continuous, and for each $B > 0$ there exist positive constants $K_1(B)$ and $K_2(B)$, and $L^1(R)$-integrable real-valued functions μ_B and h_B such that

$$\|f(x_1, t, u(t)) - f(x_2, t, u(t))\| \leq K_1(B)\mu_B(t)\|x_1 - x_2\|, \quad (2.1)$$
$$\|f(x_1, t, u(t))\| \leq K_2(B) h_B(t),$$

(where $\| \ \|$ is any norm equivalent to the usual euclidean norm) for any x_1 and x_2 with $\|x_1\| \leq B$, $\|x_2\| \leq B$, and any u in U and any t in R^+.

Observe that for each u in U Condition 3 guarantees the existence of a unique solution x_u of the differential equation

$$\dot{x} = f(x_1, t, u(t)) \tag{2.2}$$

that is defined on some interval $[0, T)$, satisfies the initial condition $x_u(0) = 0$, and has the integral representation

$$x_u(t) = \int_0^t f(x_u(s), s, u(s))ds \qquad (0 \leq t \leq T).$$

Definition. For each nonnegative real number B, $R_0(B) = \{(x_u(t), t)\}$ is the set of points in $R^n \times R^+$ such that x_u is a solution of (2.2) with $x_u(0) = 0$ and $\|x_u(\tau)\| \leq B$ for each τ in $[0, t]$.

THEOREM 2. *If conditions* (1), (2) *and* (3) *are fulfilled and* $f(\{x\} \times \{t\} \times K_t)$ *is convex for each* (x, t) *in* $R^n \times R^+$, *then the set* $R_0(B)$ *is closed in* $R^n \times R^+$.

Proof. Assume that $\{(x_{u_n}(t_n), t_n)\}$ converges to a point (x_0, t_0) in $R^n \times R^+$, where $(x_{u_n}(\tau), \tau)$ is in $R_0(B)$ for $n = 1, 2, \cdots$ and $0 \leq \tau \leq t_n$; for convenience denote x_{u_n} by x_n. In order to prove that (x_0, t_0) is in $R_0(B)$, we must show that there exists a u in U and a solution x_u of (2.2), defined in $[0, t_0] = I$, such that $x_u(0) = 0$, $\|x_u(t)\| \leq B$ on I, and $x_u(t_0) = x_0$.

Let

$$(2.3) \quad \begin{aligned} &\text{(a)} \quad \Phi_n(t) = f(x_n(t), t, u_n(t)) \qquad &&\text{if } 0 \leq t \leq t_n, \\ &\text{(b)} \quad \Phi_n(t) = f(x_n(t_n), t_n, u_n(t_n)) \qquad &&\text{if } t_n \leq t \leq t_0, \end{aligned}$$

where the sets $\{K_{t_n}\}$ may be assumed to be compact if infinitely many t_n are not equal to t_0. For if $t_n \neq t_0$ for infinitely many n, we can always find a subsequence $\{t_{n_i}\}$ such that $K_{t_{n_i}}$ is compact for each t_{n_i}. On the other hand, if $t_n = t_0$ for infinitely many n, we may without loss of generality consider only (2.3a).

By assumption (3),

$$\|\Phi_n(t)\| \leq K_2(B)\, h(t) \tag{2.4}$$

for all t in I.

Hence there exists a subsequence in $\{\Phi_n\}$ that converges weakly in $L^1[0, t_0]$ to an integrable mapping Φ. For convenience, let this be the original sequence. In particular, if we define the vector functions x and $\{\bar{x}_n\}$ on I by the relations

$$(2.5) \quad \begin{cases} x(t) = \int_0^t \Phi(s)\, ds, \\ \bar{x}_n(t) = x_n(t) & \text{if } 0 \leq t \leq t_n, \\ \bar{x}_n(t) = x_n(t_n) + \int_{t_n}^t \Phi_n(s)\, ds & \text{if } t_n \leq t \leq t_0, \end{cases}$$

we see that x is an absolutely continuous vector-valued function on I and $\{\bar{x}_n\}$ converges pointwise to x for every t in I.

Since

$$(2.6)\qquad x(t) = \lim_{n\to\infty} \int_0^t \Phi_n(s)\,ds = \lim_{n\to\infty} \bar{x}_n(t)$$

for every t in I and since, by (2.4) and (2.6), the $\{\bar{x}_n\}$ form an equi-continuous family on I, we deduce that the convergence is uniform on I.

On I, $\|x(t)\| \le B$. Otherwise it would be true that $\|x(t)\| \le B$ on some subinterval $[0, t_1]$ in I, and $\|x(t)\| > B$ for some t in every interval $(t_1, T]$ $(t_1 < T < t_0)$. In this case choose t_2 so that $t_1 < t_2 < t_0$. Then, on $[0, t_2]$, $\bar{x}_n = x_n$ for infinitely many n, by (2.5) and the convergence of $\{t_n\}$ to t_0. Hence $\{x_n\}$ converges to x uniformly on $[0, t_2]$, and since $t_2 < t_1$ for infinitely many n, it follows that $\|x_n(t)\| \le B$ on $[0, t_2]$ for infinitely many n. Thus $\|x(t)\| \le B$ for t in $[0, t_2]$, which contradicts the assumption that $[0, t_1]$ is the maximal subinterval of I with this property.

We claim that $x(t_0) = x_0$. Suppose the contrary; then

$$(2.7)\qquad \|x_0 - x(t_0)\| = \varepsilon_0 > 0,$$

and also

$$(2.8)\qquad \|x_0 - x(t_0)\| \le \|x_0 - \bar{x}_n(t_n)\| + \|\bar{x}_n(t_n) - x(t_n)\| + \|x(t_n) - x(t_0)\|.$$

Since $\{x_n(t_n)\} = \{\bar{x}_n(t_n)\} \to x_0$, $\{\bar{x}_n(t)\} \to x(t)$ uniformly on I, and $\{t_n\} \to t_0$, the right-hand side of (4.8) can be made less than ε_0 for n sufficiently large. This contradicts (2.7).

Since $\{J_k\}$ is a partition of R^+, there exists an integer $N > 0$ such that

$$I \subset \bigcup_{k=1}^{N} J_k = J.$$

Thus K_t is a compact set in R^m for each t in $I \cap (J - \{t_k\})$ $(1 \le k \le N)$, and hence the set $f(\{x\} \times \{t\} \times K_t)$ is a compact convex set for each such t.

We claim that $\Phi(t)$ is in $f(\{x(t)\} \times \{t\} \times K_t)$ a.e. on I. The weak convergence of $\{\Phi_n\}$ to Φ on I implies that for every y in R^n

$$\limsup\,(y \cdot \Phi_n(t)) \ge (y \cdot \Phi(t)) \ge \liminf\,(y \cdot \Phi_n(t)) \quad \text{a.e. on } I.$$

This is an immediate consequence of the following inequality [3, p. 114]. Over any measurable set E in I,

$$\int_E \limsup\,[y \cdot \Phi_n(s)] \ge \limsup \left[\int_E y \cdot \Phi_n(s)\,ds \right] = \int_E y \cdot \Phi(s)\,ds$$

$$= \liminf \left[\int_E y \cdot \Phi_n(s)\,ds \right] \ge \int_E \liminf\,[y \cdot \Phi_n(s)]\,ds.$$

Hence if it should happen that $y \cdot \Phi(t) > \lim\inf[y \cdot \Phi_n(t)]$ on a measurable set E with positive measure, the above inequality would be contradicted. The same argument holds if $\lim\inf[y \cdot \Phi_n(t)] > y \cdot \Phi(t)$ on a set of positive measure.

By (2.3), $\Phi_n(t)$ is in $f(\{x_n(t)\} \times \{t\} \times K_t)$ for $n = 1, 2, \cdots$. Since f is continuous in (x, u) and $\{\bar{x}_n(t)\}$ converges to x(t) uniformly on I, it follows that

$$\text{(2.10)} \qquad \begin{aligned} \sup[y \cdot f(\{x(t)\} \times \{t\} \times K_t)] &\geq y \cdot \Phi(t) \\ &\geq \inf[y \cdot f(\{x(t)\} \times \{t\} \times K_t)] \end{aligned}$$

for every y in R^n and almost all t on I.

Let E be the set of t for which (2.10) is true, and define

$$\overline{\Phi}(t) = \Phi(t) \text{ on } E, \qquad \overline{\Phi}(t) = f(x(t), t, u_t) \text{ on } I - E,$$

where, for t in I - E, u_t is any point in K_t. Note that the Lebesgue measure of I - E is zero. Then

$$x(t) = \int_0^t \Phi(s)\,ds = \int_0^t \overline{\Phi}(s)\,ds$$

for $t \in I$.

By the corollary to our theorem, there exists a u in U such that

$$\Phi(t) = f(x(t), t, u(t)) \quad \text{a.e. on } I.$$

Hence (x_0, t_0) is in $R_0(B)$, which shows that $R_0(B)$ is closed.

REFERENCES

1. A. F. Filippov, *On certain questions in the theory of optimal control,* Vestnik Moskov. Univ. Ser. Mat., Mekh., Astr., Fiz., Khim. 2 (1959), 25-32.

2. E. Roxin, *The existence of optimal controls*, Michigan Math. J. 9 (1962), 109-119.

3. P. R. Halmos, *Measure theory,* Van Nostrand and Co., New York, 1950.

RIAS, Baltimore, Maryland