JOURNAL OF DIFFERENTIAL EQUATIONS 32, 149–170 (1979)

Discontinuous Differential Equations, I

OTOMAR HÁJEK*

Department of Mathematics and Statistics, Case Western Reserve University, Cleveland, Ohio 44106

Received January 3, 1978; revised April 3, 1978

The classical notions of solution, to an ordinary differential equation, are sometimes insufficient. Three generalisations have been proposed (Filippov, Krasovskij, and, implicitly, Hermes). The first part of the paper studies these concepts; it is shown that, in very general circumstances, the last two are equivalent (and sometimes one of the alternative descriptions has a decided advantage over the other). Subsequently the results will be applied to optimal control theory and to differential games.

1. INTRODUCTION

There are two problems in 'applied' differential equation theory for which the standard concepts of solution are inadequate.

The first appears in optimal control theory; a typical example concerns a linear equation $\dot{x} = Ax - Bu$ with control values $u(t)$ bounded a priori, and minimisation of time to reach target (or of fuel consumption $\int_0^T |u(s)| \, ds$). One is led to study an appropriate 'feedback equation' $\dot{x} = Ax - BF(x)$ with the term $F(x)$ discontinuous (essentially because the optimal controls $t \mapsto u(t)$ are themselves discontinuous). The crucial question is uniqueness and continuous dependence of solutions to the initial value problem into positive time (existence is ensured otherwise; uniqueness into negative time usually fails, trivially). Obviously none of the usual approaches to uniqueness theorems is even remotely related to this situation. As concerns references, the minimum time problem is classical, [14], [16]; see [10] for L_1-optimisation; the significance of the uniqueness problem is discussed in [11].

The second problem is a basic one in differential games. The general setting includes a differential equation, $\dot{x} = f(x, p, q)$; the two parameters p, q represent the actions of the two players, often with aims that are opposed in some sense and degree. Many examples suggest that a natural mode of play is $p = p(x)$, $q = q(x)$; and the first player's aim is to choose $p(x)$ so that a given goal is

* This paper was prepared while the author held a von Humboldt award at the TH Darmstadt, Fachbereich Mathematik.

0022-0396/79/050149–22$02.00/0

attained, against all 'action' $q(x)$ of his opponent. As before, quite often the 'best' choices are discontinuous; and this makes the model itself questionable, since there is no assurance that any trajectory at all will ensue from a given choice of player controls, $\dot{x} = f(x, p(x), q(x))$. Several ways to by-pass the problem have been suggested. In one, artificial small time-delays τ are introduced, the players take turns, and subsequently one takes $\tau \to 0$ (e.g. [6]); the essence is a partial reduction to difference (rather than differential) equations. Or the information pattern is completely changed: one player is confined to controls depending on time alone, the other to strategies without knowledge of his opponent's future action (the idea is due to Varaiya, and used in [12]). Or, finally, the players are to agree in advance that they will use only those control pairs $p(x)$, $q(x)$ for which there is a solution to $\dot{x} = f(x, p(x), q(x))$ (e.g., [20], [2]); this is obviously open to all manner of objection, but seems closest to the heart of the matter. In this the natural first question is to the existence of solutions. These by-pass techniques are often combined; e.g., Pšeničnyj's ϵ-strategies, or the 'constructive' trajectories of [18]. There ensue complicated and arbitrary seeming definitions; the implied justification is that, it is for these concepts that one can prove the assertions to follow.

Obviously, what is really needed in these problems is a natural and satisfactory theory of discontinuous differential equations. The present paper is a contribution to this; it attempts to bring together the work that has already been done, and develop and apply it.

Some remarks on notation and terminology. For subsets X, Y of Euclidean n-space R^n, $X + Y$ denotes the set $\{x + y : x \in X, y \in Y\}$, B is the open unit ball $\{x : |x| < 1\}$; with some abuse of notation, $x + \epsilon B$ is then the open ϵ-ball centered at x. If one uses this 'operator' notation, another symbol is needed for set-theoretic difference, and we use an inverted slash $X \backslash Y$. For $M \subset R^n$, $\overline{cvx}\, M$ denotes closure of the convex hull of M.

Null sets in R^n are those whose n-dimensional Lebesgue measure is 0. The abbreviations are standard: AC is absolute continuity (and not alternating current), a.e. is "almost everywhere.' In R^1, intervals are to be non-degenerate, i.e., contain more than one point.

For set-valued mappings $t \mapsto S_t$, the Aumann integral $\int_0^1 S_t = \int_0^1 S_t\, dt$ is the set consisting of all $\int_0^1 s(t)\, dt$ for Lebesgue integrable $s(\cdot)$ with every value $s(t) \in S_t$; see [1], [15]. The set-valued mapping $x \mapsto S_x$ is lower semi-continuous iff the set

$$\{x : S_x \cap G \neq \varnothing\}$$

is open for each open G, and the mapping is measurable if this set is measurable, again for each open G. (For more information on these topics see [21].)

A brief description of the contents. Several types of generalised solutions, called Filippov, Krasovskij, and Hermes solutions are described in Section 1; this also contains explanatory material and some examples. In Section 2 we prove a local existence theorem for Hermes solutions (under almost no assumptions at all); the apparatus, Euler–Lebesgue approximants, is then studied further.

The classical closure theorem is carried over to Krasovskij solutions in Section 4; a consequence is that usually each Krasovskij solution is a Hermes solution. The converse problem is solved in Section 5, via a new version of a measurable selection theorem.

The following material is to appear separately: Hermes' concept of stability with respect to measurement, and uniqueness theorems; applications to the two problems with which this Introduction opens: time-optimal control of linear systems, and general strategies in differential games.

Many of the technical results on the 'calculus' of set-valued mappings (in particular, Lemmas 4.2 and 5.4) are mere extensions of known results in case the set mappings involved are measurable [1, 15, 21]. The present generalisation, to possibly non-measurable mappings, is not vacuous, but is actually needed for the main results; and it has rather interesting consequences. Lemma 5.3 is concerned with approximate 'tracking,' over an entire interval, by a 'bang-bang' control. The elementary proof follows that of Hermes ([14], Lemma 2, part (c)). Since far going generalisations of the bang-bang principle are available and known, Hermes' idea should provide much stronger results.

2. Generalised Solutions

First, the classical concepts. We refer either to a single differential equation

$$\dot{x} = f(t, x) \tag{1}$$

or, more generally, to a control system

$$\dot{x} = f(t, x, u), \qquad u \in U \tag{2}$$

(here $f\colon R^1 \times R^n \to R^n$ and $f\colon R^1 \times R^n \times R^m \to R^n$, $U \subset R^m$ respectively). It should be emphasized that there are no assumptions on f; however, we often assume that f is *locally bounded* (i.e., bounded on each bounded subset of its domain), and, occasionally, that f is measurable in t ($t \mapsto f(t, x, u)$ measurable for each fixed x, u).

Definition 2.1. A function $x\colon J \to R^n$ (J an interval in R^1) is a classical or Newton solution ($\mathscr{N}$-solution) to (1) iff

$$\dot{x}(t) = f(t, x(t)) \tag{3}$$

for all $t \in J$. It is called a Carathéodory solution (or $\mathscr{C}$-solution) to (1) iff it is AC on each compact subinterval of J, and (3) holds a.e. in J.

There follow the definitions of generalised solutions. To describe the first two, a little preparation is useful. Assume given a set-valued mapping F, from points $(t, x) \in R^1 \times R^n$ to subsets $F(t, x) \subset R^n$. We define operators **K** and **F**,

$$\begin{aligned} \mathbf{K}F(t, x) &= \bigcap_{\epsilon>0} \overline{cvx}\, F(t, x + \epsilon B), \\ \mathbf{F}F(t, x) &= \bigcap_{\epsilon>0} \bigcap_{\text{null } Z} \overline{cvx}\, F(t, (x + \epsilon B)\backslash Z) \end{aligned} \tag{4}$$

(B is the open unit ball in R^n). This will be applied with $F(t, x) = f(t, x, U)$.

DEFINITION 2.2. Let $x: J \to R^n$ (J an interval in R^1) be AC on each compact subinterval of J. Then x is called a Filippov solution (or $\mathscr{F}$-solution) of (2) iff

$$\dot{x}(t) \in \mathbf{F}f(t, x(t), U) \qquad \text{a.e. in } J,$$

and a Krasovskij solution of (2) iff

$$\dot{x}(t) \in \mathbf{K}f(t, x(t), U) \qquad \text{a.e. in } J.$$

DEFINITION 2.3. Let $x: J \to R^n$ (J an interval in R^1) be AC on each compact subinterval of J. Then x is called a Hermes solution (or $\mathscr{H}$-solution) of (2) iff there exist measurable functions $p_k : J \to R^n$, $u_k : J \to U$ and $\mathscr{C}$-solutions x_k of

$$\dot{y} = f(t, y + p_k(t), u_k(t))$$

such that

$$p_k \to 0, \qquad x_k \to x$$

uniformly on each compact subinterval of J.

DEFINITION 2.4. In addition to (2) consider the corresponding autonomous version:

$$\begin{aligned} \dot{\xi} &= 1, \\ \dot{x} &= f(\xi, x, u), \; u \in U. \end{aligned} \tag{5}$$

Then $x: J \to R^n$ is called a Hermes solution in the extended sense (or $\mathscr{H}^*$-solution) of (2) iff the function $t \mapsto (t, x(t))$, mapping $J \to R^1 \times R^n$, is a Hermes solution of (5) in the sense of 2.3. Analogously for $\mathscr{F}^*$ and $\mathscr{K}^*$ solutions.

These shall be our basic concepts; historical and technical comments, and examples, follow. For simplicity, in these we shall usually refer to (1) rather than (2).

The $\mathscr{C}$-solutions were designed to treat equations (1) where $f(t, x)$ depends continuously on x, but possibly discontinuously on t [3]; e.g., linear equations $\dot{x} - Ax = \varphi(t)$, where the forcing term φ may be discontinuous. In this the concept is spectacularly successful: for instance, in the linear case all is smooth sailing as long as $\varphi(\cdot)$ is Lebesgue integrable, a far weaker assumption than one usually needs.

In Filippov's definition [4], the crucial point is the 'minus a null set Z': the concept is set up to ignore possible misbehaviour of f on sets of small measure (in state space). The ϵB term is introduced to obtain room for this to function, and is ultimately done away with by taking $\epsilon \to 0$ via intersections.

Krasovskij's definition [17], [18, pp. 40–42] uses the ϵ-device above to take into account the behavior of f at nearby points (the fact that the **K** operator is formally simpler than **F** could hardly have been the initial motivation).

For Hermes solutions we refer to [14], but the reader will not find the explicit definition there. Hermes showed that every $\mathscr{F}$-solution (of autonomous (1)) is a uniform limit of $\mathscr{C}$-solutions x_k to 'inner perturbations' p_k, as in Definition 2.3. Subsequently he defined stability with respect to measurement by requiring that every $\mathscr{C}$-solution be uniformly approached by all such $\mathscr{C}$-solutions x_k with the same initial datum. In Definition 2.3 we merely gave the limit a name.

As concerns Definition 2.4, there is obviously no need to introduce the extended version of Newton or Carathéodory solutions. However, starting with Filippov solutions, a distinction appears. This might be ascribed to the undeniable fact that the state and the time variable do play diverse roles in (1), and the effect begins to be felt. (One might even modify the doughnut/coffee-cup description of a topologist to apply to differential equationists.)

The study of differential equations discontinuous in the state variable was begun by Flügge-Lotz ([13]; for an account see [8]).

Next, comments on details in the definitions.

These have been shorn of all assumptions on the right-hand side f. No doubt Leibniz might object to a lack of continuity in f for $\mathscr{N}$-solutions; but the point can be made that, e.g., the requirements usually made *ab initio* in describing $\mathscr{C}$-solutions properly belong in the Carathéodory existence theorem. Similar remarks apply to the other concepts, particularly to that of $\mathscr{F}$-solution.

All the definitions but the first have the prerequisite that the solution be absolutely continuous. The aim is to eliminate the 'singular' functions x having $\dot{x} = 0$ a.e., which thus have no connection with their derivative (within a differential equation, this would be disastrous).

The convex hull device that appears in **K**, **F** and Definition 2.2 may seem artificial. Note, however, that this is all that can be detected by scalar 'observations' $c^T x(\cdot)$ of the solution. E.g., an AC function x has $\dot{x}(t) \in \mathbf{K}f(x(t))$ a.e. iff, for every constant vector c, the derivative $c^T\dot{x}$ lies between lim sup and lim inf of $c^T f(x + \epsilon B)$ as $\epsilon \to 0$. (Similarly for **F**, and essential lim sup, [4].)

The 'inner' perturbations $p_k(\cdot)$ of Definition 2.3, as in $\dot{x} = f(x + p_k(t))$, may be constrasted with 'outer' perturbations $q_k(\cdot)$ in $\dot{x} = f(x) + q_k(t)$. On the one hand, for *continuous* f, every small inner perturbation obviously "is" a small outer one,

$$\dot{x} = f(x + p_k(t)) = f(x) + q_k(t), \qquad q_k = f(x + p_k) - f(x).$$

Even without continuity, the converse holds in the limit: if $\dot{x} = f(x) + q_k(t)$ and we set $y = x - \int^t q_k$, then

$$\dot{y} = f(y + p_k) \quad \text{and} \quad p_k \to 0, \quad y \to x \qquad \text{for} \quad p_k = \int^t q_k .$$

One is tempted to say that inner perturbations are more appropriate for discontinuous right-hand sides; in any case, they are at least as general, in the above sense.

Even if the function f is not measurable, Definition 2.3 requires that the p_k, u_k, and also the $\mathscr{C}$-solutions x_k, be measurable. It turns out (see the next section) that this can always be ensured.

Of the generalised solutions, the most closely related seem to be $\mathscr{K}$ and $\mathscr{H}$-solutions (at least, both have to do with different species of inner perturbation, ϵB and $p_k(t)$); and one is led to conjecture at least one containment relation. Observe, however, that in Definition 2.3 there figure measurable selections $u_k(t) \in U$, and uniformly (rather than pointwise) small state perturbations. All this will be treated in Sections 4 and 5.

Example 2.5. The dynamical system is planar, with

$$\dot{x} = 1, \qquad \dot{y} = \begin{cases} 1 & \text{for } y < 0, \\ -1 & \text{for } y \geqslant 0. \end{cases} \tag{6}$$

Off the x-axis all solutions are $\mathscr{N}$ and $\mathscr{C}$. Next, consider solutions on $[0, +\infty)$ with initial value on the axis, $y(0) = 0$.

There are no $\mathscr{N}$ nor $\mathscr{C}$ solutions (the vector field has 'jammed' the solutions, or the x-axis consists of end-points). Obviously

$$\mathbf{F}f(x_0, 0) = \mathbf{K}f(x_0, 0) = \{(1, y): |y| \leqslant 1\};$$

thus $x(t) = x_0 + t$, $y(t) \equiv 0$ are $\mathscr{F}$ and $\mathscr{K}$ solutions. Since there are no further types, the $\mathscr{K}$ solutions have continuous dependence on initial data (into positive time). Observe that, for the inextensible solutions, we have $\mathscr{C} \cap \mathscr{F} = \varnothing$ (solutions defined over as large a domain as possible).

Example 2.6. We modify (6) on the x-axis only:

$$\dot{x} = \begin{cases} 1 \\ -1, \\ 1 \end{cases} \qquad \dot{y} = \begin{cases} 1 & \text{for } y < 0, \\ 0 & \phantom{\text{for }} y = 0, \\ -1 & \phantom{\text{for }} y > 0, \end{cases} \tag{7}$$

The $\mathscr{F}$-solutions are, of course, as before. Each $\mathbf{K}f(x_0, 0)$ is a triangle with vertices $(1, \pm 1)$, $(-1, 0)$; in particular, any function $x(t)$ with $|\dot{x}(t)| \leqslant 1$ and $y(t) \equiv 0$ defines a $\mathscr{K}$-solution. One of these is an $\mathscr{N}$ and $\mathscr{C}$-solution: $x(t) = x_0 - t$, $y(t) \equiv 0$. Observe that $\mathscr{C}$-solutions and $\mathscr{F}$-solutions have continuous dependence on initial data, $\mathscr{K}$-solutions do not; and that $\mathscr{C}$-solutions need not be $\mathscr{F}$-solutions (thus the Filippov solutions are *not* generalisations of classical solutions).

The example is essentially the same as that of [4], Example p. 104. Granted that uniqueness of $\mathscr{F}$-solutions is an important property, the fact that $\mathscr{C}$-solutions need not be $\mathscr{F}$-solutions can be disastrous. The reader is referred to [14], Example 2, p. 159: even in a linear system with a cost-optimal feedback control for $\mathscr{C}$-solutions, if one allows $\mathscr{F}$-solutions only, there need be no optimal controls (precisely because the optimal solutions are not in $\mathscr{F}$). This is quite surprising since $\mathscr{F}$-solutions do satisfy a rather strong version of a closure theorem [4, p. 109].

EXAMPLES 2.7. Let $N \subset R^1$ be a non-measurable set which is dense and also has dense complement [13, p. 70]; let χ be its characteristic function. For the equation

$$\dot{x} = \chi(x) \tag{8}$$

obviously all $\mathbf{K}\chi = [0, 1]$; thus the $\mathscr{K}$-solutions are precisely the non-decreasing functions with Lipschitz constant 1. Constants $x_0 \notin N$ are $\mathscr{C}$-solutions, and probably there no others.

Later we will show that the $\mathscr{H}$-solutions are precisely these $\mathscr{K}$-solutions; these are AC, therefore measurable, and so are the perturbations $p_k(\cdot)$ in Definition 2.3. One might imagine that they have to thread their way carefully within the non-measurable function $\chi(x)$ of (8).

Consider also the equation

$$\dot{x} = \chi(t). \tag{9}$$

Then there are no $\mathscr{C}$, $\mathscr{F}$, nor $\mathscr{K}$-solutions (the derivative of an AC function is measurable). On the other hand, if (9) were made autonomous as in (4), the situation essentially reverts to (8), and there do exist $\mathscr{F}^*$ and $\mathscr{K}^*$-solutions.

These last examples were rather extreme. The author does not propose to spend his declining years solving non-measurable differential equations. But it is quite useful not to have to check for measurability always; especially since (as subsequent proof will show) nothing is gained by requiring it.

We shall not deal with Newton solutions (note $\mathscr{N} \subset \mathscr{C}$ if Newton solutions are AC locally, e.g. if f in (1) is locally bounded). At this stage we have the following general containments, for solutions of a single equation (1) or system (2).

$$\begin{array}{ccccc} \mathscr{F}^* \subset \mathscr{K}^* & & & & \mathscr{H}^* \\ \cup \quad\quad \cup & & & & \cup \\ \mathscr{F} \subset \mathscr{K} & & & & \mathscr{H} \\ & \supset & & \subset & \\ & & \mathscr{C} & & \end{array} \tag{10}$$

Indeed, obviously $\mathscr{F} \subset \mathscr{F}^*$, etc.; and $\mathscr{F} \subset \mathscr{K}$ follows from $\mathbf{F}f \subset \mathbf{K}f$, which then also yields $\mathscr{F}^* \subset \mathscr{K}^*$. Similarly, $f \in \mathbf{K}f$ provides $\mathscr{C} \subset \mathscr{K}$, and $p_k \equiv 0$ in Definition 2.3 yields $\mathscr{C} \subset \mathscr{H}$.

The following presents a simple condition (a weak form of piecewise continuity) for $\mathscr{F} = \mathscr{K}$. This is useful since a lot is known about $\mathscr{F}$-solutions, thanks to Filippov [4], while the information on $\mathscr{K}$-solutions is rather meager. The extension of the result to systems (2) is immediate.

LEMMA 2.8. *Consider $\dot{x} = f(x)$ under the following assumption on f: there exists a disjoint decomposition*

$$R^n = \bigcup M_i \qquad \textit{with} \quad M_i \subset \overline{\operatorname{Int} M_i}\,, \tag{11}$$

and continuous f_i: $R^n \to R^n$ such that $f = f_i$ on M_i. Then each $\mathscr{K}$-solution is an $\mathscr{F}$-solution (so that $\mathscr{F} = \mathscr{K}$).

Proof. It suffices to show $\mathbf{K}f \subset \mathbf{F}f$, or that $f(x + \epsilon B) \subset \overline{f((x + \epsilon B)\backslash Z)}$ for each $x \in R^n$, $\epsilon > 0$, null set Z. Take any $y \in x + \epsilon B$, find k so that $y \in M_k$ (thus $f(y) = f_k(y)$). From (11), there exist $y_j \to y$ in Int M_k; obviously we may even take $y_j \in \operatorname{Int} M_k \backslash Z$ and, of course, $y_j \in x + \epsilon B$. By continuity, $f_k(y_j) \to f_k(y)$, so that $f(y)$ is in the closure of $f((x + \epsilon B)\backslash Z)$ as asserted.

We note that the condition in (11) is essential; e.g. in (7), the set where $\dot{x} = -1$ has empty interior, (11) is not satisfied; and, in fact, $\mathscr{F} \neq \mathscr{K}$. For an instance, see [7, p. 461]: on the switching surface, the feedback control is arbitrarily defined as 0 (with disastrous consequences, [9]).

3. EXISTENCE AND APPROXIMATION OF HERMES SOLUTIONS

The section is concerned with differential equations

$$\dot{x} = f(t, x) \tag{1}$$

rather than control systems. Here the function f: $R^1 \times R^n \to R^n$ is said to be *measurable in t* iff $t \mapsto f(t, x)$ is measurable for each fixed $x \in R^n$; this is satisfied trivially if, as in Corollary 3.2, (1) is autonomous.

THEOREM 3.1 (Existence for $\mathscr{H}$-solutions). *Let f be locally bounded and measurable in t. Then, for any initial data t_0, x_0, there is a Hermes solution $x(\cdot)$ if (1) with $x(t_0) = x_0$, defined at least on some interval $(t_0 - \epsilon, t_0 + \epsilon)$.*

Proof. In the usual manner we need only consider $t_0 = 0$, $x_0 = 0$; treat only a right neighbourhood $[0, \epsilon)$ of $t_0 = 0$; and assume that f is bounded globally, by some constant φ: otherwise one replaces f by a bounded function which coincides with f inside an appropriate neighbourhood of the nominal position.

For each $\delta = 1, \frac{1}{2}, \frac{1}{3}, \ldots$ construct the analogue of the Euler polygonal arc, a function $y(\cdot)$ as follows. For $j = 0, 1, \ldots$ we set $t_j = j\delta$, $y_0 = 0$, and then

$$y(t) = y_j + \int_{t_j}^{t} f(s, y_j)\, ds \qquad \text{in } [t_j, t_{j+1}],$$

where $y_{j+1} = y(t_{j+1})$. Then $y(\cdot)$ has Lipschitz constant φ, and

$$\dot{y}(t) = f(t, y(t) + p(t)) \quad \text{a.e.},$$

where

$$| p(t)| = | y_j - y(t)| \leqslant \varphi \mid t - t_j \mid \leqslant \varphi\delta.$$

Thus $y(\cdot)$ is a $\mathscr{C}$-solution of the perturbed equation (2), with inner perturbations $p \to 0$ uniformly as $\delta \to 0+$. Since all $y(0) = 0$ and φ is a common Lipschitz constant, the theorem of Arzelà and Ascoli applies. Thus some subsequence of the $y = y_\delta$ converges uniformly; by definition, the limit is a $\mathscr{H}$-solution of (1).

COROLLARY 3.2. *In (1) let f be locally bounded. If (1) is autonomous, then locally there are $\mathscr{H}$-solutions to any initial data; even if (1) is not autonomous, then locally there are at least $\mathscr{H}^*$-solutions to prescribed initial data.*

By inspection of the proof (construction of $y(\cdot)$) it is easily seen that the standard extendability results apply: $x(\cdot)$ is defined globally on R^1 if f is bounded, or has linear growth, etc.

It is now natural to ask whether the $\mathscr{H}$-solutions obtained by the construction above, approximation by "polygonal arcs," are in some way special, or constitute a yet further interesting class of generalised solutions. A negative answer (in essentially the most general case) is provided by Proposition 3.5, with apparatus prepared in

LEMMA 3.3. *Let x: $[0, 1] \to R^n$ satisfy a Lipschitz condition. Then, for every $\epsilon > 0$, there is a closed null-set Z and a Lipschitz function y such that*

$$| x(t) - y(t)| \leqslant \epsilon \qquad on \quad [0, 1],$$

and which is 'piecewise linear' in the sense that

$$\dot{y}(t) \equiv \dot{x}(\alpha)$$

for all t in any component (α, β) *of* $[0, 1] \backslash Z$.

Proof. The Lipschitz function x is AC, so the derivative $\dot{x}(\alpha)$ exists outside a null-set Z_0. Thus, for each $\alpha \notin Z_0$, sufficiently small $\beta - \alpha > 0$, and all $t \in [\alpha, \beta]$, we have

$$| x(t) - x(\alpha) - (t - \alpha)\, \dot{x}(\alpha) | \leqslant \epsilon(t - \alpha).$$

Therefore there is a disjoint (hence, countable) collection of such intervals $[\alpha_k, \beta_k]$, covering $[0, 1] \backslash Z_0$ up to a null-set. In particular,

$$[0, 1] = \bigcup (\alpha_k, \beta_k) \cup Z, \qquad \text{meas}\, Z = 0, \quad Z = \bar{Z}$$

(we have included the end-points of the intervals in the null-set Z).

Now, let $y(\cdot)$ be linear in each (α_k, β_k), or rather

$$y(t) = c_k + x(\alpha_k) + (t - \alpha_k)\, \dot{x}(\alpha_k) \qquad \text{for} \quad t \in [\alpha_k, \beta_k]; \tag{3.1}$$

The constants c_k are as yet undetermined. We wish to make $y(\cdot)$ at least continuous. If $t_i \to t+$ with all t, t_i in the union of our intervals,

$$\alpha_i \leqslant t_i \leqslant \beta_i, \quad \alpha_k \leqslant t \leqslant \beta_k, \qquad \text{but} \quad [\alpha_i, \beta_i] \neq [\alpha_k, \beta_k],$$

then necessarily $t = \beta_k$, and

$$\beta_k < \alpha_i < \beta_i \to \beta_k$$

(since the closed intervals are disjoint, with full measure of union). Analogously if $t_j \to t-$. We obtain two conditions, right and left respectively:

$$\lim c_i = c_k + (x(\alpha_k) - x(\beta_k) + (\beta_k - \alpha_k)\, \dot{x}(\alpha_k)),$$
$$\lim c_j = c_k.$$

To achieve this, set up the saltus-function

$$c(t) = \sum_{\beta_i < t} (x(\alpha_i) - x(\beta_i) + (\beta_i - \alpha_i)\, \dot{x}(\alpha_i)). \tag{4}$$

It is well-defined: absolute convergence is ensured by estimate (3): $\sum | \cdots | \leqslant \epsilon(t - 0) \leqslant \epsilon$. It is left-continuous, and right-discontinuities occur, at most, at the points β_k. Finally, we redefine $y(\cdot)$ throughout $[0, 1]$ by

$$y(t) = c(t) + \begin{cases} x(\alpha_k) + (t - \alpha_k)\, \dot{x}(\alpha_k) & \text{if} \quad t \in [\alpha_k, \beta_k] \\ 0 & \text{if} \quad t \notin \bigcup [\alpha_i, \beta_i]; \end{cases} \tag{5}$$

then $y(\cdot)$ is continuous, and "piece-wise" linear, with (3.1) satisfied for $c_k = c(\alpha_k) = c(t)$.

Let us now estimate $x - y$, first at times t in some $[\alpha_k, \beta_k]$:

$$x(t) - y(t) = (x(t) - x(\alpha_k) - (t - \alpha_k)\,\dot{x}(\alpha_k)) + c(t),$$
$$|x(t) - y(t)| \leqslant \epsilon(t - \alpha_k) + \sum \epsilon(\beta_i - \alpha_i) \leqslant \epsilon(t - 0) \leqslant \epsilon$$

by (3). From continuity of x and y, the same estimate extends over [0, 1] entire.

Next, let us show that y is a Lipschitz function. (Note that $|\dot{y}(t)| = |\dot{x}(\alpha)| \leqslant \lambda$ a.e. is not sufficient: y could still be singular, or have a singular component.) Take $t > s$ in [0, 1], both in some of our intervals:

$$\alpha_j \leqslant s \leqslant \beta_j\,, \qquad \alpha_k \leqslant t \leqslant \beta_k\,.$$

The case $j = k$ is trivial; assume, therefore, $\beta_j < \alpha_k$. Then

$$y(t) - y(s) = c(t) - c(s) + x(\alpha_k) + (t - \alpha_k)\,\dot{x}(\alpha_k) - x(\alpha_j) - (s - \alpha_j)\,\dot{x}(\alpha_j).$$

Here the term (see (4))

$$c(t) - c(s) = \sum_{s \leqslant \beta_i < t} = (j^h\text{th}) + \sum_{s < \alpha_i < \beta_i < t}\,,$$

so that

$$\begin{aligned} y(t) - y(s) = {} & x(\alpha_j) - x(\beta_j) + (\beta_j - \alpha_j)\,\dot{x}(\alpha_j) \\ & + \sum_{s < \alpha_i < \beta_i < t} (\cdot) + x(\alpha_k) + (t - \alpha_k)\,\dot{x}(\alpha_k) \\ & - x(\alpha_j) - (s - \alpha_j)\,\dot{x}(\alpha_j). \end{aligned}$$

Here the terms $x(\alpha_j)$ cancel, as do the $\alpha_j \dot{x}(\alpha_j)$; the sum is split into two, each converging absolutely; and terms $x(t) - x(t)$, $x(s) - x(s)$ are added:

$$\begin{aligned} = {} & x(s) - x(\beta_j) + \sum (x(\alpha_i) - x(\beta_i)) - x(t) + x(\alpha_k) \\ & + x(t) - x(s) \\ & + (\beta_j - s)\,\dot{x}(\alpha_j) + \sum (\beta_i - \alpha_i)\,\dot{x}(\alpha_i) + (t - \alpha_k)\,\dot{x}(\alpha_k). \end{aligned} \tag{6}$$

If λ is a Lipschitz constant for $x(\cdot)$, then it now follows that

$$|y(t) - y(s)| \leqslant (t - s)\lambda + (t - s)\lambda + (t - s)\lambda;$$

again, the estimate is extended over [0, 1] by continuity, $y(\cdot)$ has Lipschitz constant 3λ.

This, together with (4) and (5), establish the assertion.

There is a generalisation, with Lipschitz condition on x, y replaced by absolute continuity only (more work is then done on (6)); we shall not need this. For readers familiar with the concepts of Riemann and Lebesgue sums, there is a curious consequence:

COROLLARY 3.4. *For every integrable* φ: $[0, 1] \to R^n$,

$$\int_0^1 \varphi(t)\, dt = \lim_{\mathscr{D}} \sum_k \varphi(\alpha_k)(\beta_k - \alpha_k),$$

limit taken over suitable disjoint decompositions $\mathscr{D}$:

$$[0, 1] = \bigcup (\alpha_k, \beta_k) \cup Z, \qquad \text{meas } Z = 0. \tag{7}$$

Proof. Set $x = \int_0^t \varphi$ and use the generalisation of Lemma 3.3.

We shall say that y: $J \to R^n$ is an Euler-Lebesgue *approximant*, for a differential equation $\dot{y} = g(t, y)$, iff $y(\cdot)$ is AC, and there exists a disjoint decomposition $\mathscr{D}$: $J = \bigcup (\alpha_k, \beta_k) \cup Z$ such that Z is a null-set, and

$$\dot{y}(t) \equiv g(\alpha_k, y(\alpha_k)) \qquad \text{for} \quad \alpha_k < t < \beta_k . \tag{8}$$

In this connection, the mesh size of $\mathscr{D}$ is $\sup_k(\beta_k - \alpha_k)$.

PROPOSITION 3.5. *In* (1) *let* f *be locally bounded. Then every* $\mathscr{H}$*-solution* $x(\cdot)$ (*on a compact interval*) *is the uniform limit of Euler-Lebesgue approximants* y_k *to perturbed equations*

$$\dot{y} = f(t, y + p_k(t)), \qquad p_k \to 0 \textit{ uniformly}. \tag{9}$$

Every uniform limit of Euler-Lebesgue approximants to (9), *with mesh sizes tending to* 0, *is an* $\mathscr{H}^*$*-solution of* (1).

Proofs. Since $\mathscr{H}$-solutions are themselves uniform limits of $\mathscr{C}$-solutions (of suitably perturbed equations), it suffices to prove the first assertion for $x(\cdot)$ a $\mathscr{C}$-solution, on $[0, 1]$.

Then $x(\cdot)$ is AC: from the assumptions, the points $f(t, x(t))$ admit a common bound, so that x satisfies a Lipschitz condition. Apply the construction from Lemma 3.3. We have (7), $y(\cdot)$ is AC, and on each component (α_k, β_k),

$$\begin{aligned} \dot{y}(t) \equiv \dot{x}(\alpha_k) &= f(\alpha_k, x(\alpha_k)) \\ &= f(\alpha_k, y(\alpha_k) + p(\alpha_k)), \end{aligned}$$

where $p = x - y$ satisfies $|p(t)| = |x(t) - y(t)| \leqslant \epsilon$ by (5).

For the second assertion, note that every $E - L$ approximant for (9) is, according to (8), a $\mathscr{C}$-solution to

$$\dot{y} = f(t + (\alpha_k - t), y + p_k(t)),$$

where $\alpha_k - t \to 0$ uniformly if the mesh sizes tend to 0.

4. Closure Theorem for Krasovskij Solutions

LEMMA 4.1. *Let $t \mapsto C_t$ be a set-valued mapping whose values C_t are compact, convex subset of R^n, all contained in a common ball. Then $\int_0^1 C_t$ is compact and convex.*

Proof. The set is convex by Richter's theorem (e.g., [16], p. 26); obviously it is bounded. To prove it is closed, assume $c_k(t) \in C_t$ a.e. and $\int_0^1 c_k(t)\,dt = x_k \to x$. All $c_k(\cdot)$ belong to a closed ball of finite radius in L_2-space; this ball is weak star compact, and the weak star topology restricted to this ball is metrisable. Thus we have *sequential* weak star compactness, and may assume that the original sequence converges to an L_2 function $c(\cdot)$ (in the weak star topology). It follows that $\int_0^1 c(t)\,dt = x$. Also, there exists a sequence $b_k(\cdot)$ of suitable finite convex combinations of the $c_k(\cdot)$ which converges to $c(\cdot)$ a.e. Convexity of the sets C_t of values yields that each $b_k(t) \in C_t$ a.e., and compactness of C_t yields $c(t) \in C_t$ a.e. Thus indeed $x \in \int_0^1 C_t$.

Remark. In the case that $t \mapsto C_t$ is also measurable, this is a special case of a theorem due to Aumann (e.g., [16], p. 29); and our proof is a minor modification of part of this. Absence of measurability assumptions will be important later.

LEMMA 4.2. *For $t \in [0, 1]$ let $t \mapsto S_t$ be a set-valued mapping, whose values are all contained in a common ball of R^n. If x: $[0, 1] \to R^n$ has*

$$x(t) - x(s) \in \int_s^t S_r \tag{2}$$

(for all $t > s$ in $[0, 1]$), then $x(\cdot)$ is AC and satisfies

$$\dot{x}(t) \in \overline{cvx}\, S_t \qquad \text{a.e.} \tag{3}$$

In particular,

$$x(t) = x(0) + \int_0^t \dot{x}(r)\,dr \in x(0) + \int_0^t \overline{cvx}\, S_r$$

for all $t \in [0, 1]$.

Proof. From uniform boundedness, $x(\cdot)$ satisfies a Lipschitz condition; thus it is *AC*. Take any countable *dense* set $c_1, c_2, \ldots$ in R^n (these will be used as 'test directions'). Then, for any fixed $x \in R^n$ and $S \subset R^n$ it is true that $x \in cvx\, S$ if (from density), and only if,

$$c_k{}^T x \leqslant \sup c^T S \qquad \text{for each} \quad k = 1, 2, \ldots .$$

Now assume that (3) fails, on a set of positive measure. Then, on a possibly smaller set M, still of positive measure, we have more: there exists a single test direction $c = c_k$, and $\epsilon > 0$, such that

$$\sup c^T S_t < c^T \dot{x}(t) - 2\epsilon \qquad (t \in M) \tag{4}$$

(here $t \mapsto \sup c^T S_t$ may not be measurable; and we proceed to bypass this). Set $\varphi(t) = c^T \dot{x}(t) - 3\epsilon$, and take any point t which (i) belongs to M, (ii) is a point of metric density of M, and (iii) is a Lebesgue point of φ (recall $x \in AC$, so $\dot{x}$ and φ are L_1). By definition of derivatives, for arbitrarily small $h > 0$.

$$\varphi(t) + \epsilon = c^T \dot{x}(t) - \epsilon \leqslant \frac{1}{h} c^T(x(t+h) - x(t)). \tag{5}$$

From (2), there exists $t \mapsto s(t) \in S_t$ with

$$x(t+h) - x(t) = \int_t^{t+h} s(r)\, dr$$

($s(\cdot)$ possibly depending on t and h). Then

$$c^T(x(t+h) - x(t)) = \int_t^{t+h} c^T s(r)\, dr = \int_{I\cap M} + \int_{I\setminus M}$$

with $I = [t, t+h]$. On $I \cap M$ we estimate via (4),

$$c^T(x(t+h) - x(t)) \leqslant \int_{I\cap M} \varphi(r)\, dr + \int_{I\setminus M} c^T s(r)\, dr$$

$$= \int_t^{t+h} \varphi(r)\, dr + \int_{I\setminus M} (c^T s(r) - \varphi(r))\, dr.$$

Note that both $c^T s(\cdot)$ and $\varphi(\cdot)$ are uniformly bounded; if α denotes a bound,

$$\leqslant \int_t^{t+h} \varphi(r)\, dr + 2\alpha \operatorname{meas}([t, t+h]\setminus M).$$

We use this estimate in (5),

$$\varphi(t) + \epsilon \leqslant \frac{1}{h}\int_t^{t+h} \varphi(r)\, dr + 2\alpha \frac{\operatorname{meas}([t, t+h]\setminus M)}{\operatorname{meas}[t, t+h]}$$

for arbitrarily small h. As $h \to 0+$, from our choice of t (see above), the last term tends to 0, and the integral term to $\varphi(t)$. The contradiction $\varphi(t) + \epsilon \leqslant \varphi(t)$ concludes the proof.

THEOREM 4.3 (Closure theorem for $\mathscr{K}$-solutions). *Let $x_k(\cdot)$ be Krasovskij solutions of*

$$\dot{y} = f(t, y + p_k(t, y), u_k(t, y)) + q_k(t, y)$$

on [0, 1], *with*

$$u_k(t, y) \in U, \textit{ and } \quad p_k \to 0,\ q_k \to 0 \textit{ uniformly};$$

assume that f *is locally bounded, and* U *bounded.*

4.3.1. *If the* $x_k(\cdot)$ *converge uniformly, then the limit function is a Krasovskij solution of* $\dot{x} = f(t, x, u)$, $u \in U$.

4.3.2. *Unless* $x_k(0) \to \infty$, *some subsequence of the* $x_k(\cdot)$ *does converge uniformly, at least on some* $[0, \epsilon]$ (*with* $\epsilon > 0$ *depending only on* f, U, $\liminf |x_k(0)|$).

Proof. First assume $x_k \to x$ uniformly; choose $\epsilon > 0$, $\delta > 0$ arbitrarily. Then, for large indices,

$$|p_k(\cdot)| < \epsilon, \qquad |q_k(\cdot)| < \delta, \qquad |x_k(\cdot) - x(\cdot)| < \epsilon$$

on [0, 1], so that, a.e.,

$$\begin{aligned} \dot{x}_k(t) &\in \overline{cvx}\,(f(t, x_k(t) + \epsilon B, U) + \delta B) \\ &\subset \overline{cvx}\,(f(t, x(t) + 2\epsilon B, U) + \delta B). \end{aligned}$$

It follows that, for any $t > s$ in [0, 1],

$$x_k(t) - x_k(s) \in \int_s^t \overline{cvx}\,(f(r, x(r) + 2\epsilon B, U) + \delta B)\, dr.$$

The assumptions of Lemma 4.1 are satisfied; hence, on taking limits as $k \to \infty$,

$$x(t) - x(s) \in \int_s^t \overline{cvx}\,(\cdots).$$

By Lemma 4.2, $x(\cdot)$ is AC, and

$$\dot{x}(t) \in \overline{cvx}\,(f(t, x(t) + 2\epsilon B, U) + \delta B) \qquad \text{a.e.}$$

Take limits over a sequence $\delta \to 0$ to obtain

$$\dot{x}(t) \in \overline{cvx}\, f(t, x(t) + 2\epsilon B, U) \qquad \text{a.e.,}$$

and then over a sequence $\epsilon \to 0$ to verify

$$\dot{x}(t) \in \mathbf{K} f(t, x(t), U) \quad \text{a.e.}$$

Thus indeed $x(\cdot)$ is a $\mathscr{K}$-solution.

For the second assertion assume only $x_k(0) \to \infty$. Then $x_{k_i}(0) \to x_0$ for a subsequence $y_i = x_{k_i}$ and point x_0. We may now include $y_i(0) - x_0$ into the inner perturbations $p(\cdot)$, and assume from the outset that all $y_i(0) = 0$. Choose any $\alpha > 0$, set $\beta = \sup |f([0, 1], (\alpha + 1)B, U)| + 1$. Then, for $t \in [0, \beta/\alpha]$ (and large indices i: $|p_i| < 1 > |q_i|$), one has $y_i(t) \in \alpha B$, since $|y(t)| < B$ there. Thus all $y_i(t)$ remain in a bounded portion of R^n, and their derivatives are uniformly bounded. From the theorem of Arzelà and Ascoli, a subsequence of the y_i (a subsequence of the x_k) converges uniformly on $[0, \beta/\alpha]$. This concludes both proofs.

Corollary 4.4. *For $\dot{x} = f(t, x, u)$, $u \in U$, with f locally bounded and U bounded, each Hermes solution is a Krasovskij solution.*

Proof. Take any $\mathscr{H}$-solution $x(\cdot)$, and the appropriate $\mathscr{C}$-solutions $x_k \to x$ of $\dot{y} = f(t, y + p_k(t), u)$ ($p_k \to 0$ uniformly). Trivially, the $x_k(\cdot)$ are $\mathscr{K}$-solutions; thus by Theorem 4.3, $x = \lim x_k$ is a $\mathscr{K}$-solution of our equation.

5. $\mathscr{K}^* \subset \mathscr{H}^*$

The source of some results on non-measurable equations is the following trivial miracle.

Lemma 5.1. *Let $x \mapsto F(x)$ be a set-mapping, from points $x \in R^k$ to subsets of R^n. Then the set-mapping*

$$x \mapsto F(x + G)$$

is lower semi-continuous if $G \subset R^k$ is open.

Proof. For any (open) $H \subset R^n$,

$$\{x: F(x + G) \cap H \neq \varnothing\} = (-G) + \{y: F(y) \cap H \neq \varnothing\}.$$

Since G is open, so is $-G$, and thus also $(-G)+$ (any set).

Lemma 5.2 (Moveable Carathéodory's Theorem). *Assume that all values $f(x) \in \operatorname{cvx} F(x)$, where $f: R^k \to R^n$ is measurable, $x \mapsto F(x)$ is a measurable set-valued mapping, and each value $F(x)$ is closed (in R^n). Then*

$$f(x) = \sum_0^n \alpha_k(x) f_k(x), \qquad f_k(x) \in F(x),$$

$$\alpha_k(x) \geqslant 0, \qquad \sum_0^n \alpha_k(x) = 1$$

for suitable measurable α_k: $R^k \to R_1$, f_k: $R^k \to R_n$.

(The proof appears in [16] as part of Theorem 8.4, pp. 29–30. The version of Filippov's Lemma used here, with closed values instead of compact ones, is in [12], p. 118.)

For the next assertion the notation is as follows: E^m is the standard m-simplex in R^{m+1},

$$E^m = \left\{(\alpha_0, ..., \alpha_m): \sum_0^m \alpha_k = 1, \text{ all } \alpha_k \geqslant 0\right\},$$

and the set of its verices is

$$V^m = \{e_1, ..., e_{m+1}: e_k \text{ basic unit vector in } R^{m+1}\}.$$

LEMMA 5.3 (Bang-bang tracking). *Let $t \mapsto B(t)$ be an integrable $n \times (m+1)$ matrix, and $\epsilon > 0$. For every solution of*

$$\dot{x}(t) = B(t)\, u(t), \quad u(t) \in E^m, \qquad x(t_0) = x_0$$

there is a solution of

$$\dot{y}(t) = B(t)\, v(t), \quad v(t) \in V^m, \qquad y(t_0) = x_0$$

with all $| x(t) - y(t)| < \epsilon$.

Proof. Assume $t_0 = 0$, choose $\delta > 0$ (to be fixed later). Apply the bang-bang principle (Theorem 8.3 in [16]) to each interval $[k\delta, (k+1)\delta)$, $k = 0, \pm 1, ...$. There results a measurable control $v(\cdot)$ with all values in V^m, and a corresponding solution $y(\cdot)$ such that

$$y(k\delta) = x(k\delta) \qquad \text{for} \quad k = 0, \pm 1, ... \tag{1}$$

To estimate $x - y$ at t, find closest $k\delta$, and use (1):

$$\begin{aligned} | x(t) - y(t)| &= |(x(t) - x(k\delta)) - (y(t) - y(k\delta))| \\ &= \left| \int_{k\delta}^t B(x)(u(s) - v(s))\, ds \right| \leqslant \int_{k\delta}^{(k+1)\delta} | B(s)|\, ds (2m)^{1/2} < \epsilon \end{aligned}$$

once we choose δ so that the L_1-function B, or rather its AC integral, satisfies

$$\sup_t \int_t^{t+\delta} | B(s)|\, ds < \epsilon \cdot (2m)^{-1/2}.$$

Remarks. There is no limit version of Lemma 5.2. E.g. for $n = 1$, $B \equiv 1$, $x_0 = 0$ and $[-1, 1]$ in place of E^1, consider $u \equiv 0$. Then controls $v(\cdot)$ with all $| v(t)| = 1$ do approximate $x \equiv 0$, but cannot match exactly. The proof follows that of Hermes ([14], Lemma 2, part (c)). There is an obvious extension to linear control systems $\dot{x} = A(t)x + B(t)u$.

LEMMA 5.4 (Approximate measurable selection). *Assume that*

$$p(t) \in f(q(t) + Q, U) + P \quad a.e. \tag{2}$$

where the functions $p, q\colon R^1 \to R^n$ *are measurable,* $Q \subset R^n$, $U \subset R^m$, $P \subset R^n$, $f\colon R^n \times R^m \to R^n$.

Then, for any $\epsilon > 0$, *there exist measurable functions* a, b, u *such that*

$$\begin{aligned} p(t) &= f(q(t) + a(t), u(t)) + b(t) \qquad \text{a.e.,} \\ a(t) &\in Q + \epsilon B, \qquad b(t) \in P + \epsilon B, \qquad u(t) \in U. \end{aligned} \tag{3}$$

(B is the unit ball in R^n).

Proof. Set $\delta = \epsilon/2$. Since p, q are measurable, they can be uniformly approximated by 'simple' functions:

$$\begin{aligned} |p(t) - p'(t)| &< \delta, \qquad |q(t) - q'(t)| < \delta, \\ p'(t) &= \sum p_k c_k(t), \qquad q(t) = \sum q_k c_k(t), \end{aligned}$$

with constants p_k, q_k in R^n, and the $c_k(\cdot)$ characteristic functions of measurable sets $E_k \subset R^1$, same for both p and q. From (2)

$$p'(t) \in f(q'(t) + Q', U) + P' \quad \text{a.e.}$$

(here $Q' = Q + \delta B$, $P' = \delta B$). At $t \in E_k$

$$p_k \in f(q_k + Q', U) + P'. \tag{4}$$

For each k choose (constant) solutions to (4),

$$p_k = f(q_k + a_k', u_k) + b_k',$$

$$a_k' \in Q', \qquad u_k \in U, \qquad b_k' \in P',$$

and use these to define simple functions

$$a_k'(t) = \sum a_k' c_k(t), \qquad b'(t) = \sum b_k' c_k(t), \qquad u(t) = \sum u_k c_k(t).$$

Obviously these are measurable, and map into Q', U, P' respectively. Thus

$$p'(t) = f(q'(t) + a'(t), u(t)) + b'(t) \quad \text{a.e.,}$$

and (3) holds on absorbing the differences $p' - p$, $q' - q$ into the terms b, a.

Remarks. Note that $t \mapsto f(q(t) + a(t), u(t))$ is measurable even if f is not. The result is an *approximate* implicit function theorem: in the result there appear ϵ-terms even if $Q = P = 0$ initially. There is a significant difference between our result and, e.g., Filippov's Lemma. There the selection is 'pointwise,' or stroboscopic in the terminology of [12]. Here it is not even almost-stroboscopic. Some generalisations are immediate: the domain of t could be R^k, U might well be replaced by measurable $t \mapsto U_t$, and $f(q, u)$ replaced by $f(t, q, u)$ depending measurably on t. Finally, R^n and R^m could be replaced by a separable metric space, and R^1 by a set with a σ-algebra of sets (relative to which one would interpret measurability).

THEOREM 5.5. *In the autonomous control system*

$$\dot{x} = f(x, u), \qquad u \in U \tag{5}$$

let f be locally bounded, and U bounded. Then every Krasovskij solution is a Hermes solution.

Proof 5.5.1. Consider any Krasovskij solution $x(\cdot)$ of (5) on a compact interval J. We wish to find measurable $p_k: J \to R_n$ with $p_k \to 0$ uniformly, and $\mathscr{C}$-solutions x_k of $\dot{y} = f(y + p_k(t)u)$ such that $x_k \to x$ uniformly.

5.5.2. Choose $\epsilon > 0$. Then, a.e.,

$$\dot{x}(t) \in \mathbf{K}f(x(t), U) \subset \overline{cvx} f(x(t) + \epsilon B, U)$$
$$\subset cvx(f(\cdots) + \epsilon B) \subset cvx(\overline{f(\cdots)} + \epsilon \bar{B}).$$

We intend to apply Carathéodory's theorem in the version of Lemma 5.2. For this we need to verify measurability of the set-mapping

$$t \mapsto \overline{f(x(t) + \epsilon B, U)} + \epsilon \bar{B},$$

and this follows from Lemma 5.1, with $F(x + G) = f(x + \epsilon B, U)$, and continuity of $t \mapsto x(t)$.

5.5.3. From Lemma 5.1,

$$\dot{x}(t) = \sum_0^n \alpha_k(t)\, b_k(t) \qquad \text{a.e.}$$

with measurable $\alpha_k \geqslant 0$, $\sum \alpha_k = 1$, and also measurable $b_k: J \to R^n$, whose values satisfy

$$b_k(t) \in \overline{f(x(t) + \epsilon B, U)} + \epsilon \bar{B} \subset f(x(t) + \epsilon B, U) + 2\epsilon B.$$

We propose to apply bang-bang tracking to (6), $u(\cdot)$ being the vector function with coordinates $\alpha_0, \ldots, \alpha_n$. That $t \mapsto B(t)$, with columns $b_k(t)$, is integrable follows from the boundedness assumptions on f, U, J, and continuity of $x(\cdot)$.

5.5.4. From Lemma 5.3 it now follows that there exists an AC function y: $J \to R^n$, and integrable a: $J \to R^n$, such that

$$|x(t) - y(t)| < \epsilon \qquad (\text{all } t \in J),$$
$$\dot{y}(t) = a(t) \in f(x(t) + \epsilon B, U) + 2\epsilon B \quad \text{a.e.}$$

From (7), $x(t) + \epsilon B \subset y(t) + 2\epsilon B$, so that

$$\dot{y}(t) \in f(y(t) + 2\epsilon B, U) + 2\epsilon B \quad \text{a.e.}$$

5.5.5. We may apply Lemma 5.4 directly:

$$\dot{y}(t) = f(y(t) + w_1(t), u(t)) + w_2(t) \quad \text{a.e.} \tag{8}$$

with w_1, w_2: $J \to 3\epsilon B$, u: $J \to U$ all measurable. Next, the term w_2 is absorbed: write $z(t) = y(t) - \int^t w_2(s)\, ds$, $\lambda =$ length J, to obtain

$$\begin{gathered} |y(t) - z(t)| \leqslant 3\epsilon\lambda, \\ y(t) + w_1(t) = z(t) + w(t), \qquad |w(t)| \leqslant 3\epsilon(1 + \lambda), \end{gathered} \tag{9}$$

still with $w(\cdot)$ measurable. Then from (8),

$$\dot{z}(t) = f(z(t) + w(t), u(t)) \quad \text{a.e.}$$

Thus $z(\cdot)$ is a $\mathscr{C}$-solution; finally, from (7) and (9), $|z(t) - x(t)| \leqslant \epsilon(1 + 3\lambda)$. Since $\epsilon > 0$ was arbitrary, the $\mathscr{C}$-solutions $z(\cdot)$ and inner perturbations $w(\cdot)$ do have the required properties.

COROLLARY 5.6. *For autonomous control systems $\dot{x} = f(x, u)$, $u \in U$, with bounded U and locally bounded f, the Hermes and the Krasovskij solutions coincide.*

COROLLARY 5.7. *For control systems $\dot{x} = f(t, x, u)$, $u \in U$, with bounded U and locally bounded f, the Hermes and the Krasovskij solutions, both in the extended sense, coincide.*

Proofs (Corollary 4.4 and Theorem 5.5). The scheme for generalised solutions is now as follows:

$$\begin{array}{ccccc} \mathscr{F}^* & \subset & \mathscr{K}^* & = & \mathscr{H}^* \\ \cup & & \cup & & \cup \\ \mathscr{F} & \subset & \mathscr{K} & \supset & \mathscr{H} \\ & & & & \cup \\ & & & & \mathscr{C} \end{array}$$

none of the inclusions is an equality in general. The last vertical $\mathscr{H} \subset \mathscr{H}^*$ may be completed with a middle term, limits of Euler-Lebesgue approximants.

Some remarks in conclusion.

The difference between $\mathscr{K}$ and $\mathscr{F}$-solutions is that, in the construction of **F** (Section 2, (4)) some small sets have been subtracted. It might be instructive to omit other types of small set, e.g. closed nowhere dense, or first category. The first questions to be asked are whether there is an analogue to $\mathbf{F}f(x) \neq \varnothing$ (under mild conditions, cf. [4], Lemma 1), and also to a closure theorem.

We have shown $\mathscr{H} \subset \mathscr{K}$; the converse containment holds for autonomous systems, or under a measurability condition. Can the gap be closed?

The $\mathscr{F}$-solutions of say $\dot{x} = f(x)$ obviously do what they were designed to—ignore misbehaviour of f on null sets. There is still the question, what else do they do. In detail, suppose a $\mathscr{K}$-solution $x(\cdot)$ remains a $\mathscr{K}$-solution for every equation $\dot{x} = g(x)$ with $\{f \neq g\}$ a null set; is then $x(\cdot)$ necessarily an $\mathscr{F}$-solution?

In the preparation of this paper, the author is indebted to many of his colleagues for their views, comments, and suggestions; in particular, to Professor George Leitmann, and to the referee.

References

1. R. J. Aumann, Integrals of set-valued functions, *J. Math. Anal. Appl.* **12** (1965), 1–12.
2. A. Blaquière, F. Gérard, and G. Leitmann, "Quantitative and Qualitative Games," Academic Press, New York, 1969.
3. C. Carathéodory, "Vorlesungen über reele Funktionen," 1st ed., Teubner, Leipzig 1918.
4. A. F. Filippov, Differential equations with discontinuous right-hand side (in Russian), *Mat. Sb.* **51** (1960).
5. I. Flügge-Lotz, "Discontinuous Automatic Control," Princeton Univ. Press, Princeton, N. J., 1953.
6. A. Friedman, "Differential Games," Wiley–Interscience, New York, 1971.
7. R. V. Gamkrelidze, Theory of time-optimal processes in linear systems (in Russian), *Izv. Akad. Nauk SSSR* **22** (1958), 449–477.
8. O. Hájek, "Dynamical Systems in the Plane," Academic Press, London, 1968.
9. O. Hájek, Terminal manifolds and switching locus, *Math. Systems Theory* **6** (1973), 289–301.
10. O. Hájek, L_1-Optimization in Linear systems with bounded controls, *J. Optimization Theory Appl.*, in press.
11. O. Hájek, Inverted paradoxes in optimal control theory, Preprint no. 275, TH Darmstadt, Fachbereich Mathematik, 1976.
12. O. Hájek, "Pursuit Games," Academic Press, New York, 1975.
13. P. R. Halmos, "Measure Theory," Van Nostrand Reinhold, New York, 1950.
14. H. Hermes, Discontinuous vector fields and feedback control, *in* "Differential Equations and Dynamical Systems" (J. K. Hale and J. P. LaSalle, Eds.), pp. 155–165, Academic Press, New York, 1967.
15. H. Hermes, Calculus of set valued functions and control, *J. Math. Mech.* **8** (1968), 47–60.
16. H. Hermes and J. P. LaSalle, "Functional Analysis and Time Optimal Control," Academic Press, New York, 1969.
17. N. N. Krasovskij, "Game-Theoretic Problems of Capture" (in Russian), Nauka, Moscow, 1970.

18. N. N. Krakovskij and A. I. Subbotin, "Positional Differential Games" (in Russian), Nauka, Moscow, 1974.
19. E. B. Lee and L. Markus, "Foundations of Optimal Control Theory," Wiley, New York, 1967.
20. G. Leitmann, "Cooperative and Non-Cooperative Many-Player Differential Games," Springer–Verlag, Vienna, 1974.
21. D. H. Wagner, Survey of measurable selection theorems, *SIAM J. Control Optimization* 15 (1977), 859–903.