MULTIVALUED MAPS AND MULTIVALUED DIFFERENTIAL EQUATIONS

by

K. Deimling
Univ. of Paderborn
Fed. Rep. Germany

Technical Report #233
April 1985
Introduction.
Everybody meets multivalued maps very early in his mathematical education, as inverses of maps which are not one-to-one, but in elementary lectures the multivalued aspect is usually suppressed by means of elementary tricks or restrictions which make sense for practical purposes; think of \( f(x) = x^2 \) which has the inverse \( f^{-1}(x) = \{-\sqrt{x}, \sqrt{x}\} \), from \( R_+ = \{ r \in R : r \geq 0 \} \) to \( 2^R \) (the subsets of \( R \)); or think of a linear operator \( T : X \to Y \) with kernel \( N(T) \neq \{0\} \); in which case one uses the trick to consider \( \hat{T} : X/N(T) \to R(T), \) defined by \( \hat{T}x = Tx \) for \( \hat{x} = x+N(T), \) which is one-to-one and therefore has a singlevalued inverse \( \hat{T}^{-1} : R(T) \to X/N(T). \)

In many more advanced topics of analysis it is not so easy to overcome the multivaluedness and therefore one has created a theory of such maps, which we simply call multis. So a multi is a map \( F : D \subseteq X \to 2^Y \) which associates to each \( x \in D \) a subset \( F(x) \) of \( Y. \) The graph of \( F \) is \( \text{graph}(F) = \{(x,y) \in X \times Y : x \in D, y \in F(x)\}. \) Every \( A \subseteq X \times Y \) may be considered as the graph of a multi, namely of \( F \) defined by \( F(x) = \{y \in Y : (x,y) \in A\}. \) The assumption "\( F(x) \neq \emptyset \) for all \( x \in D \)" will be expressed by \( F : D \subseteq X \to 2^Y \setminus \emptyset. \) If one allows \( F(x) = \emptyset \) for some \( x \) then one introduces \( \text{Dom}(F) = \{x \in X : F(x) \neq \emptyset\}. \) The range of a multi \( F \) is \( R(F) = \bigcup_{x \in X} F(x). \) The inverse multi of \( F \) is \( F^{-1} : Y \to 2^X, \) defined by \( F^{-1}(y) = \{x \in X : y \in F(x)\}. \) and we let \( F^{-1}(A) = \{x \in X : F(x) \cap A \neq \emptyset\}; \) notice that this is in general larger than \( \{x \in X : F(x) \subseteq A\}; \) in particular we have \( \text{Dom}(F^{-1}) = R(F). \) In most cases \( X \) and \( Y \) will be Banach spaces (think of \( \mathbb{R}^n \) if you are not familiar with this).
Conceptionally, the simplest idea in the study of multis is to look for selections, i.e. single-valued maps \( f : D \to Y \) such that \( f(x) \in F(x) \) (\( \forall x \in D \)), which allow to reduce a multivalued problem to a single-valued one, to some extent. Of course selections are only useful if they have some nice properties like continuity or measurability. The existence of such selections under reasonably general assumptions is an interesting problem, the first topic in this series of lectures. It is intimately connected to continuity/measurability properties of the multis, but much depends also on the properties of the sets \( F(x) \) (like compactness, convexity, etc.).

Another branch of the theory of multis is concerned with \textbf{monotone multis}, this means (for \( X = Y = \mathbb{R}^n \)) \( (x-\bar{x}, y-\bar{y}) \geq 0 \) for all \( x, \bar{x} \in D \) and \( y \in F(x), \bar{y} \in F(\bar{x}) \), where \( (x, y) = \sum_{i=1}^{n} x_i y_i \). There exists a well-cooked theory of \textbf{singlevalued monotone maps}, some more advanced questions of which show the need for monotone multis; this theory has applications to some problems in partial differential equations. I shall not talk about this topic, since it is more elementary than the first one and covered in several books today.

The second main topic of this course is the existence of solutions to multivalued differential equations \( x' \in F(t, x) \), together with some qualitative theory (periodic solutions, solutions in cones, etc.). Here the simplest idea is again to use selection theorems which allow reduction to the singlevalued case \( x' = f(t, x) \in F(t, x) \). But we also present other methods, for example the Euler-Cauchy-Polygon method or fixed point theorems.

Let us start with some motivating examples.
Example 0.1. (ODEs without uniqueness) Consider the IVP \( x' = f(t, x) \),
\( x(0) = x_0 \in \mathbb{R}^n \), where \( f : [0, a] \times \mathbb{R}^n \to \mathbb{R}^n \) is continuous and bounded
(for simplicity). Then you know that there exists at least one solution
on \( J = [0, a] \), but there may also be many of them. The problem is equi-
valent to finding \( x \in C(J) \) (the continuous functions on \( J \) with norm
\( |x|_0 = \max_J |x(t)| \)) which satisfies

\[
(1) \quad x(t) = x_0 + \int_0^t f(s, x(s))ds \text{ on } J.
\]

So we have the "solution-multi" \( S : \mathbb{R}^n \to 2^{C(J)} \cdot \mathbb{B} \), defined by
\( S(x_0) = \{ x \in C(J) : x \text{ is sol. of (1) on } J \} \). Properties of \( S \) are
interesting in connection with the "section-multi" \( S_t(x_0) = \{ x(t) : x \in S(x_0) \} \), (i.e. \( S_t : \mathbb{R}^n \to 2^{\mathbb{R}^n \cdot \mathbb{B}} \)), which is the composition of \( S \)
and the continuous map \( T_t : C(J) \to \mathbb{R}^n \) defined by \( T(x) = x(t) \) (i.e.
\( S_t = T_t \circ S \)); it is known that \( S(x_0) \) is compact and connected and
therefore \( S_t(x_0) \) has the same properties,
which can be exploited in the search for
\( \omega \)-periodic solutions if \( f \) is \( \omega \)-periodic
in \( t \) (i.e. \( f(t+\omega, x) = f(t, x) \) on \( \mathbb{R} \times \mathbb{R}^n \));
indeed it is easy to see that such a
solution exists iff \( x_0 \in S_\omega(x_0) \) for some
\( x_0 \in \mathbb{R}^n \). This example indicates that it may be useful to know fixed
point theorems for multis.

Examples 0.2. (Stochastic ODEs) In the simplest case this is \( x' = f(t, x, \omega) \),
\( x(0, \omega) = x_0(\omega) \), where a probability measure space \( (\Omega, \mathcal{A}, \mathbb{P}) \), a measurable
\( x_0 : \Omega \to \mathbb{R}^n \) and \( f : J \times \mathbb{R}^n \times \Omega \to \mathbb{R}^n \) are given such that \( f \) is measurable
in \((t, \omega)\) and continuous in \( x \in \mathbb{R}^n \). This problem is again equivalent to
(2) \[ x(t,\omega) = x_0(t,\omega) + \int_0^t f(s,x(s,\omega),\omega) \, ds \]

(suppose \( |f(t,x,\omega)| \leq M \) on \( J \times \mathbb{R}^n \times \Omega \) for simplicity), i.e. to find solutions \( x \) which are continuous in \( t \), but also measurable in \( \omega \) (such an \( x \) is called a stochastic process, the only type of solutions one is interested in from the probabilistic standpoint). So for every \( \omega \) we have (many) solutions, i.e. \( S(\omega) = \{ x \in C(J) : x \text{ is solution of (2) for this } \omega \} \), and the question is whether we can pick (for every \( \omega \in \Omega \)) a solution \( x_\omega \in S(\omega) \) such that \( x_\omega(t) \) is measurable in \( \omega \) (for every fixed \( t \)). This example shows the need for theorems about measurable selections.

**Example 0.3. (Approximation theory)** The basic problem is: Given \( D \subseteq X \) and \( x \in X \), find \( y \in D \) such that \( |x-y| = \rho(x,D) = \inf_{z \in D} |x-z| \) (the shortest distance from \( x \) to \( D \)). For example, if \( D \) is compact, such a best approximation \( y \) exists, but for a given \( x \) there may be many of them, which motivates to introduce the multi (metric projection) \( P \) : \( X \rightarrow 2^D \), defined by \( P(x) = \{ y \in D : |x-y| = \rho(x,D) \} \). Here it is more interesting to know whether \( P \) admits continuous selections (or even better ones). For example, one has a continuous curve \( \Gamma \) (represented by \( x(t) \)) and wants to "project" it continuously to a continuous curve in \( D \). A problem studied by many people is: Given \( X \) and a subspace \( D \subseteq X \) does there exist a continuous projection \( P_0 : X \rightarrow D \) (which may be a nonlinear map), contrary to the elementary situations which one considers in elementary Functional Analysis, where \( P_0 \) is linear). In general, for compact \( D \) it is obvious that \( P(x) \) is compact (if \( D \) is also convex then
\( \text{Example 0.4. (Control theory)} \) The dynamics of a particle is described by

\[ x' = f(t,x,u(t)); \]

where the functions \( u(\cdot) \) are the controls (realizable by means of mechanical or electronic device) taking values in a bounded subset \( U \) of \( \mathbb{R}^n \). One wants to choose a \( u(\cdot) \) such that something becomes optimal for the corresponding trajectories \( x(\cdot) \) of (3). For example, control the flight of a rocket such that it reaches its goal with minimal amount of fuel; or: find \( u(\cdot) \) such that a particle \( x_0 \) outside \( B = \mathbb{R}^n \) comes to \( B \) in shortest time (along the corresponding trajectory determined by (3)). In many examples of this type the optimal \( u(\cdot) \) has a lot of jumps, so one has to work theoretically with measurable functions \( u(\cdot) \). One way (not the only one) to study the existence problem is to consider the multivalued differential equation

\[ x' \in F(t,x) \text{ with } F(t,x) = f(t,x,U) = \{ f(t,x,u) : u \in U \}; \]

if we can solve this one then we have \( x'(t) = f(t,x(t),u(t)) \) and the only question is whether \( u(\cdot) \) is measurable.

\( \text{Example 0.5. (ODEs with discontinuous right hand sides)} \) There are problems (e.g. thermostats which cause a discontinuous supply of fuel to a burner, or hysteresis effects in nonlinear vibrations) which are
described by $x' = f(t, x)$, where $f$ has discontinuities in $x$. In such situations the easiest way to say something about solutions is to consider the multivalued differential equation

$$\text{(5)} \quad x' \in F(t, x) \quad \text{with} \quad F(t, x) = \bigcap_{\varepsilon > 0} \text{conv} (f(t, x + \varepsilon e(o))),$$

where $\varepsilon e(o) = \{x \in \mathbb{R}^n : |x| \leq \varepsilon\}$ and $\text{conv}$ denotes the "closed convex hull"; notice that $F(t, x) = \{f(t, x)\}$ if $f(t, \cdot)$ is continuous at $x$.

**Example 0.6. (Implicit differential equations):** $f(t, x(t), x'(t)) = 0$

can be regarded as

$$x' \in F(t, x) \quad \text{with} \quad F(t, x) = \{v : f(t, x, v) = 0\}.$$

**Example 0.7. (Solutions under constraints) **There are problems (e.g. dynamical models for pricesystems in economics) where only those solutions of $x' = f(t, x)$ are acceptable which at every $t$ satisfy the constraint $x(t) \in K(t)$ (the sets $K(t)$ are given). In the simplest case $K(t) = K$ this means that solutions starting in $K$ should never leave $K$. For example, if $K$ is a cone, say $K = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_i \geq 0 \text{ for } i = 1, \ldots, n\}$, this means that the solutions starting at "nonnegative" $x_0$ remain nonnegative for all $t$, a condition which is reasonable in a lot of concrete applications (biology, chemistry), since a negative population size or negative concentration of chemical stuff is nonsense. It is obvious that such a behavior of the trajectories can only be obtained by building up "barriers" at the boundary $\partial K$ of $K$, i.e. $f$ has to satisfy certain conditions there. Such necessary and sufficient conditions are well known, they can be formulated geometrically (by means of "tangent cones") and therefore such problems can be solved elegantly (also in the general case $K(t)$) by means of multivalued differential equations.
These examples give you a first impression of reasonable assumptions on the sets \( F(x) \), of the need for selections and the need for fixed point theorems for multis. You will hopefully not expect to get such things without further assumptions on \( F \) and \( F(x) \), for example of continuity type, since if the sets \( F(x) \) vary wildly with \( x \) nothing can be said. So we need some preparation for such results which we consider next. These notes are based on Chap. B of [2], the lousy book [1] and some recent papers.

§ 1 Multivalued Maps and Selections.

1.1 Continuity concepts.

Definition 1.1: Let \( X,Y \) be Banach spaces and \( F : D \subseteq X \rightarrow 2^Y \setminus \emptyset \). Then \( F \) is said to be upper semicontinuous (usc for short) in \( D \) if 
\( \{ x \in D : F(x) \subseteq V \} \) is open in \( D \) whenever \( V \subseteq Y \) is open. \( F \) is called lower semicontinuous (lsc for short) in \( D \) if \( F^{-1}(V) \) is open in \( D \) whenever \( V \subseteq Y \) is open. \( F \) is said to be continuous if it is continuous w.r. to the "Hausdorff-metric" \( d_H(A,B) = \max \{ \sup_{x \in A} d_Y(x,B), \sup_{y \in B} d_Y(A,y) \} \).

Discussion: 1.) Remember \( A \subseteq D \) is open in \( D \) iff (by definition) \( A = D \cap O \) with \( O \) open in \( X \).

2.) \( d_H \) is really a metric on the closed bounded subsets. Since we shall almost always assume that \( F(x) \) is at least closed and bounded (which means compact if \( Y = \mathbb{R}^n \)) we can use \( d_H \); continuity w.r. to \( d_H \) then means:
\( D \ni x_n \rightarrow x_o \in D \Rightarrow d_H(F(x_n),F(x_o)) \rightarrow 0 \) as \( n \rightarrow \infty \).

3.) By taking complements in the definition of usc you see that \( F \) is usc on \( D \) iff \( F^{-1}(A) \) is closed in \( D \) whenever \( A \subseteq Y \) is closed; notice that \( F^{-1}(A) = \{ x \in D : A \cap F(x) \neq \emptyset \} = D \setminus \{ x \in D : F(x) \subseteq Y \setminus A \} \). For
singlevalued maps $F$ we have $F^{-1}(A) = \{ x \in D : F(x) \in A \} = D \setminus F^{-1}(Y \setminus A)$, hence usc means $F^{-1}(Y)$ open in $D$ if $V \subseteq Y$ is open, i.e. usc = continuous, and of course also lsc = continuous.

4.) Both concepts (usc and lsc) can be localized: Call $F$ **usc at** $x_0 \in D$ iff to every open $V \ni F(x_0)$ there exists $\delta = \delta(x_0,V) > 0$ such that $F(B_\delta(x_0) \cap D) \subseteq V$; then $F$ is usc in $D$ iff $F$ is usc at $x_0$ for every $x_0 \in D$. Call $F$ **lsc at** $x_0 \in D$ iff for every $y \in F(x_0)$ and every open neighborhood $V$ of $y$ there exists $\delta = \delta(x_0,y,V) > 0$ such that $F(x) \cap V \neq \emptyset$ ($\forall x \in B_\delta(x_0) \cap D$); then $F$ is lsc in $D$ iff $F$ is lsc at $x_0$ for every $x_0 \in D$.

5.) Remember that a $\varphi : D \subseteq X \rightarrow \mathbb{R}$ is called usc if $D \ni x_n \rightarrow x_0 \in D$ implies $\liminf_{n \rightarrow \infty} \varphi(x_n) \leq \varphi(x_0)$ (correspondingly lsc means $\varphi(x_0) \leq \limsup_{n \rightarrow \infty} \varphi(x_n)$); such $\varphi$ may be discontinuous and therefore the multivalued concepts have nothing to do with this. There is only a vague analogy as follows:

Suppose $F$ is usc at $x_0$ and $F(x_0)$ is closed; then $\limsup_{n \rightarrow \infty} F(x_n) = F(x_0)$ whenever $x_n \rightarrow x_0$. Proof: Given $\varepsilon > 0$, $F$ usc at $x_0$ and $x_n \rightarrow x_0$ implies $F(x_n) \subseteq F(x_0) + B_\varepsilon(0)$ for all large $n$, hence $\limsup_{n \rightarrow \infty} F(x_n) = \bigcup_{n \geq 1} F(x_n) \subseteq F(x_0) + B_\varepsilon(0)$ for every $\varepsilon > 0$, and therefore $\subseteq F(x_0)$ since $F(x_0)$ is closed.

Some more consequences of Def. 1 are contained in

**Proposition 1.1.** Let $F(x)$ be closed bdd for all $x \in D$. Then we have

(a) $F$ is lsc if $\sup \rho(y,F(x_n)) \rightarrow 0$ whenever $x_n \rightarrow x_0$. If $F$ is usc then $F(x_0)$ is lsc. If $F(x)$ is compact ($\forall x \in D$) then $F$ is continuous iff $F$ is lsc and usc.

(b) $F$ usc and $D$ closed $\Rightarrow$ graph($F$) closed. If $F(D)$ is compact and $D$ is closed then $F$ is usc iff graph($F$) is closed.
(c) $D$ compact, $F$ usc and $F(x)$ compact ($\forall x \in D \Rightarrow F(D)$ is compact.)

Proof. All parts are easy to prove. For example, the first part of (a): suppose $F^{-1}(V)$ is not open in $D$ for some open $V \subseteq Y$. Then there is an $x_0 \in D$ with $F(x_0) \cap V \neq \emptyset$ and $(x_n) \subseteq D$ such that $x_n \to x_0$ and $F(x_n) \cap V = \emptyset$. Choose $y_0 \in F(x_0) \cap V$, then $B_r(y_0) \subseteq V$ for some $r > 0$ and $B_r(y_0) \cap F(x_n) = \emptyset \forall n$, a contradiction to sup $\{y, F(x_n)\} \to \infty$.

Example 1.1. $F : \mathbb{R} \to 2^\mathbb{R} \setminus \emptyset$, defined by $F(\alpha) = [-1,1]$ and $F(x) = \{x\}$ for $x \neq \alpha$, is usc but not lsc. $G : \mathbb{R} \to 2^\mathbb{R} \setminus \emptyset$, defined by $G(\alpha) = \{\alpha\}$ and $G(x) = [-1,1]$ for $x \neq \alpha$, is lsc but not usc.

Example 1.2. Consider $D \subseteq X$, two bdd functions $\varphi$, $\psi : D \to \mathbb{R}$, such that $\varphi \leq \psi$ and define $F : D \to 2^\mathbb{R} \setminus \emptyset$ by $F(x) = [\varphi(x), \psi(x)]$. Then $F$ is lsc iff $\varphi$ is usc and $\psi$ is lsc. [For example, proof of $\Leftarrow$: If $V \subseteq \mathbb{R}$ open, $(x_n) \subseteq D$, $x_n \to x_0 \in D$ and $F(x_0) \cap V \neq \emptyset$, then $\lim_{n \to \infty} \varphi(x_n) \leq \varphi(x_0) \leq \psi(x_0) \leq \lim_{n \to \infty} \psi(x_n)$, hence $F(x_n) \cap V \neq \emptyset$ for all large $n$, therefore $F^{-1}(V)$ open in $D$.]

Example 1.3. Let $D \subseteq X$ be compact. Then the metric projection (Ex. 0.3) is usc. [Proof: $A \subseteq X$ closed and $x_n \to x_0$ with $P_D(x_n) \cap A = \emptyset \forall n$ implies $P_D(x_0) \cap A = \emptyset$; (notice that $y_n \in P_D(x_n) \cap A$ is such that there exists a convergent subsequence $(y_{n_k})$, since $D$ is compact, and $\rho(\cdot, D)$ is continuous). But $P_D$ need not be lsc (consider $D = \{(x, |x|) : |x| \leq 1\} \subseteq \mathbb{R}^2$), even if $D$ is convex.

Example 1.4. In honour of the great mathematician H. Poincaré, the "section-multi" $S_t$ (Ex. 0.1.) is called Poincaré-"map" and denoted by $P_t$; he was apparently the first one who proved existence of periodic solutions this way. Under the conditions given in Ex. 0.1., $S$ and $P_t$ are usc.
[Proof: \( P_t = T_t \circ S \) with \( T_t \) continuous; hence \( P_t \) is usc if \( S \) is. Now \( A \subset X \), closed, \( u_n \in S(x_0) \cap A \) and \( x_n \to x_0 \) imply: \( u_n' = f(t,u_n) \) in \( J \), \( u_n(0) = x_n \to x_0 \), hence \( u_n \to u \) in \( X \) for some subsequence and some \( u \) (Ascoli-Arzelà), hence \( u \in S(x_0) \cap A \), i.e. \( S \) is usc). However, \( S \) and \( P_t \) need not be lsc. For example, \( P_t \) for \( u' = 2t \sqrt{|u|} \) is given by

\[
P_t(x) = \begin{cases} 
(\sqrt{x} + t)^2 & \text{for } x > 0 \\
(0, (t - \sqrt{|x|})^2) & \text{for } x \in [-t^2, 0) \\
(-t - \sqrt{|x|})^2 & \text{for } x < -t^2
\end{cases}
\]

In this case \( P_t \) and \( S \) are not lsc at \( x = 0 \).

Example 1.5. (cp. Ex. 0.4.) Given \( f : X \times U \to Y \), let \( F(x) = f(x,U) \). If \( f(\cdot,u) \) is continuous (\( \forall u \in U \)) then \( F \) is lsc. If \( U \) is compact and \( f \) is continuous on \( X \times U \) then \( F \) is also usc. (proof is an easy exercise).

Practical Conclusions. The Discussion, Proposition 1.1 and the examples show the following: usc is most easily checked by the condition

"\( x_n \to x_0 \) and \( F(x_n) \cap A = \emptyset \) (\( A \) closed) \( \Rightarrow F(x_0) \cap A = \emptyset \); lsc is most easily checked by" \( x_n \approx x_0 \) and \( F(x_0) \cap V = \emptyset \) (\( V \) open) \( \Rightarrow F(x_n) \cap V = \emptyset \)

for all large \( n \). Continuity (a strong condition!) means

\( \lim_{n \to \infty} F(x_n) = F(x_0) \) as \( x_n \approx x_0 \). For the (rough) imagination it also helps to think of usc as \( \lim_{n \to \infty} F(x_n) \subseteq F(x_0) \) and of lsc as

\( \lim_{n \to \infty} F(x_n) \supseteq F(x_0) \) (similar to the case of real-valued functions: "usc can only jump upwards, lsc can only jump downwards"; replace \( \leq \) by \( \subseteq \)). Both, usc and lsc, play an essential role in finding reasonable selections of multis as we are going to show next.
1.2 Continuous selections

The basic result about existence of such selections is the following theorem of E. Michael (1956).

**Theorem 1.1:** Let $X, Y$ be Banach spaces, $F : D \subset X \to 2^Y \setminus \emptyset$ Isc and $F(x)$ closed convex ($\forall x \in D$). Then $F$ admits a continuous selection.

**Proof.** 1. Given $\varepsilon > 0$, we first find a continuous $f : D \to Y$ such that $f(x) \in F(x) + V$ on $D$ (with $V = B_\varepsilon(o)$). (Here we don’t need that $F(x)$ is closed).

Indeed, let $U_y = F^{-1}(y-V) = \{ x \in D : y \in F(x)+V \}$, which is open in $D$ since $y-V$ is open and $F$ is Isc. Since $D = \bigcup_y U_y$ and $D$ is paracompact, there exists a locally finite refinement $(W_{\lambda})_{\lambda \in \Lambda}$ of $(U_y)_y \in Y$ and a partition of unity $(\varphi_{\lambda})_{\lambda \in \Lambda}$ w.r. to $(W_{\lambda})_{\lambda \in \Lambda}$ (i.e. $D = \bigcup_{\lambda \in \Lambda} W_{\lambda}$). To every $x \in D$ there exists a neighborhood which meets only finitely many of the $W_{\lambda}$, to $\lambda \in \Lambda$ there is $y \in Y$ such that $W_{\lambda} \subset U_y$, $\varphi_{\lambda}$ is continuous (even locally Lipschitz), $0 \leq \varphi_{\lambda}(x) \leq 1$, $\Sigma_{\lambda \in \Lambda} \varphi_{\lambda}(x) = 1$ on $D$, $\text{supp} \varphi_{\lambda} = \{ x \in D : \varphi_{\lambda}(x) > 0 \} \subset W_{\lambda}$. For $\lambda \in \Lambda$ choose $y_{\lambda}$ such that $\varphi_{\lambda} = 0$ in $X \setminus U_{y_{\lambda}}$ and let $f(x) = \Sigma_{\lambda \in \Lambda} \varphi_{\lambda}(x)y_{\lambda}$. This $f$ is continuous (even locally Lipschitz); $\varphi_{\lambda}(x) \neq 0$ implies $x \in U_{y_{\lambda}}$, hence $y_{\lambda} \in F(x)+V$ and therefore $f(x) \in F(x)+V$ (since this set is convex).

2. Let $V_i = B_{1/2^i}(o)$. By induction we define continuous $f_i : D \to V_i$ such that

1. $f_i(x) \in f_{i-1}(x) + 2V_{i-1}$ (for $i \geq 2$), $f_i(x) \in F(x)+V_i$ (for $i \geq 1$)

(both relations on $D$).

For $i = 1$ we have only the second condition, and the existence of such an $f_1$ was the first step. Suppose we have already $f_1, \ldots, f_k$ for some $k \geq 1$. 
Consider \( G(x) = F(x) \cap (f_k(x) + V_k) \); \( G(x) \) is convex, \( G(x) \neq \emptyset \) (notice that \( V_k = -V_k \)) and \( G \) is lsc as you will check easily. Hence, by Step 1, there exists a continuous \( f_{k+1} \) such that \( f_{k+1}(x) \in G(x) + V_{k+1} \) on \( D \), hence \( f_{k+1}(x) \in f_k(x) + V_k + V_{k+1} = f_k(x) + 2V_k \) and \( f_{k+1}(x) \in f(x) + V_{k+1} \).

3. The first part of (1) means: \((f_1)\) is Cauchy, hence (since \( Y \) is complete) \( f_1(x) \to f(x) \) and \( f \) is continuous; the second part of (1) gives \( f(x) \in F(x) \) since \( F(x) \) is closed.

**Remarks.**

1. The result is sharp in the sense that a continuous selection may not exist if either \( Y \) is not complete, or \( F \) is not lsc, or \( F(x) \) is not closed, or \( F(x) \) is not convex; see Michael's original or Chap 8 of [2] for counter-examples.

2. The conditions of Theorem 1 are of course not necessary for the existence of a continuous selection; see also subsequent examples.

3. In special situations you can find continuous selections directly.

For example, suppose that \( F \) is continuous, \( F(x) \) is closed bdd convex (\( \forall x \in D \)) and \( Y \) is uniformly convex (e.g. \( Y = \mathbb{R}^n \) or Hilbert space); then choose any \( y_0 \in Y \) (e.g. \( y_0 = 0 \)) and define \( f(x) = P_{F(x)}(y_0) \) (i.e. the metric projection of \( y_0 \) onto \( F(x) \)). Then \( f \) is a continuous selection; for \( y_0 = 0 \) and \( Y \) Hilbert this is called the "minimal selection" in [1].

On pp. 70-80 of this "book" (a big exaggeration) you will also find other "proposals": choose always the so-called "Cebycev-center" [see § 17 of [2] for the meaning of this] of \( F(x) \), or the barycenter of \( F(x) \) if \( Y = \mathbb{R}^n \) (in the latter case the selection is Lipschitz if \( F \) is Lipschitz w.r. to \( d_H \), \( F(x) \) is compact convex and \( F(D) \) is bounded).

4. The authors of [1] want to give the impression that much more flexibility is gained by introducing first "locally continuously selectionable" (lcs) multis as follows: Call \( F \) lcs at \( x_0 \in D \) if to
every \( y_0 \in F(x_0) \) there exists a \( \delta = \delta(x_0,y_0) > 0 \) and a continuous
\( f : B_\delta(x_0) \cap D \to Y \) such that \( f(x_0) = y_0 \) and \( f(x) \in F(x) \) on \( B_\delta(x_0) \cap D \).

Call \( F \) lcs on \( D \) if it has this property at every \( x_0 \in D \). However,
a simple application of the definitions given so far shows the
following: An \( F \) satisfying the assumptions of Theorem 1 is of course lcs
on \( D \). [notice that lsc is not destroyed if you replace \( F(x_0) \) by \{\( y_0 \}
with \( y_0 \in F(x_0) \} \). Furthermore an lcs \( F \) is already lsc! If \( F \) is lcs
and \( F(x) \) is convex (\( \forall x \in D \)) then \( F \) has a continuous selection [see step 1
of the proof to Theorem 1.1: To \( y \in D \) choose \( z \in F(y) \) and a continuous
\( f_y : B_\delta(y,z) \cap D \to Y \); then \( f(x) = \sum_{y \in D} f_y(x) \) is a continuous
selection]. If \( F \) is lcs at \( x_0 \), \( G : X \to 2^Y \) has open graph and \( F(x_0) \cap G(x_0) \neq \emptyset \)
then \( F \circ G \) is lcs at \( x_0 \).

Let us consider some examples for Theorem 1.1

**Example 1.6.** If \( \varphi, \psi : D \to R \) are bdd, \( \varphi \leq \psi \), \( \varphi \) usc and \( \psi \) lsc then there
is a continuous \( f : D \to R \) such that \( \varphi \leq f \leq \psi \) (see Ex. 1.2).

**Example 1.7.** The Poincaré-map for \( u' = 2\sqrt{|u|} \) (see Ex. 1.4.\) has a lot of
continuous selections although it is not lsc at \( x = 0 \). The corresponding
"solution-multi", which is also not lsc at \( x = 0 \), has exactly one
continuous selection, namely the one obtained by choosing for every \( x \) the
maximal solution through \( x \).

**Example 1.8.** Let \( \Omega \subset R \times R^n \) be open, \( F : \Omega \to 2^{R^n} \) lsc and \( F(t,x) \) closed
convex (\( \forall (t,x) \in \Omega \)). Then the initial value problem \( x' \in F(t,x) \), \( x(t_0) = x_0 \)
(with \( (t_0,x_0) \in \Omega \)) has a \( C^1 \)-solution up to \( \partial \Omega \), since \( F \) admits a continuous
selection \( f \) and \( x' = f(t,x) \), \( x(t_0) = x_0 \) has such a solution.
Remark. On p. 82 of [1] you find the following glorious result (Corollary 1) about lcs multis: "Let $X$ be a metric space, $Y$ a topological vector space, $X_0 \subseteq X$ open, $F : X_0 \to 2^Y$ lcs and $F(x)$ a "convex cone" ($\forall x \in X_0$), i.e. $\lambda y + \mu \tilde{y} \in F(x)$ for all $y, \tilde{y} \in F(x)$ and all $\lambda, \mu \geq 0$. Then there exists a continuous $f : X \to Y$ such that $f(x) \in F(x)$ on $X_0$ and $f(x) = 0$ on $X \setminus X_0$". Of course! (choose $f(x) = 0$).

Conclusion. It should have become clear that existence of continuous selections is essentially dependent on lsc, unless one sees them directly in very special situations (another one of these will be mentioned in § 2). The convexity of $F(x)$ is also essential for a general result. Now, for some interesting results one needs only continuous approximate selections, i.e. "$f(x)$ is only close to $F(y)$ with $y$ close to $x$", and for the existence of these upper semicontinuity is the right concept, as we show next.

1.3. **Approximate selections.**

A simple example of an usc multi without continuous selections is

$$F : \mathbb{R} \to 2^{\mathbb{R}}(0), \text{ defined by }$$

$$F(x) = \begin{cases} (-1) & \text{for } x < 0 \\ [-1,1] & \text{for } x = 0 \\ 1 & \text{for } x > 0 \end{cases}$$

However it is easy to get approximate selections.

**Theorem 1.2.** Let $X,Y$ be Banach spaces; $F : D \subseteq X \to 2^Y \setminus \emptyset$ usc and $F(x)$ convex ($\forall x \in D$). Then to $\varepsilon > 0$ there exists a locally Lipschitz $f_\varepsilon : D \to Y$ such that $f_\varepsilon(D) = \text{conv}(F(D))$ and $f_\varepsilon(x) \in F(D \cap B_\varepsilon(x)) \cup B_\varepsilon(0)$ on $D$. 
In particular: If \( D = \mathbb{R}^n \) with \( n \) open and if \( K \subset \mathbb{R}^n \) is compact then there is a neighborhood \( V_\varepsilon \) of \( K \) with \( V_\varepsilon = \{ x \in \mathbb{R}^n : \rho(x,K) < \varepsilon \} \) and a locally Lipschitz \( \lambda \)-finite dimensional \( f_\varepsilon : V_\varepsilon \to \text{conv}(F(K)) \) such that \( f_\varepsilon(x) \in F(K \cap B_\varepsilon(x)) + B_\varepsilon(o) \) on \( V_\varepsilon \).

**Proof.** Since \( F \) is usc, given \( \varepsilon > 0 \) and \( x \in D \) we find \( \delta = \delta(x,\varepsilon) > 0 \) (w.l.o.g. \( \delta \leq \varepsilon \)) such that \( F(D \cap B_\delta(x)) \subset F(x) + B_\varepsilon(o) \). Choose a locally finite refinement \( (W\lambda_{\lambda \in \Lambda}) \) of \( (D \cap B_r(x)) \) with \( r(x) = \delta(x,\varepsilon)/3 \), a partition of unity \( (\varphi_{\lambda \in \Lambda}, y_{\lambda \in \Lambda} \in F(W_{\lambda})) \), and let

\[
f_\varepsilon(x) = \sum_{\lambda \in \Lambda} \varphi_{\lambda}(x) y_{\lambda}
\]

for \( x \in D \). Evidently, \( f_\varepsilon \) is locally Lipschitz and \( f_\varepsilon(D) \subseteq \text{conv}(F(D)) \). To \( x \in D \) there are only finitely many \( \lambda \) (say \( \lambda_1, \ldots, \lambda_m \)) such that \( \varphi_{\lambda_i}(x) > 0 \), which means \( x \in W_{\lambda_i} \subset B_{r(z_i)}(z_i) \cap \text{int} D \) for some \( z_i \in D \). Let \( r_p = r(z_p) = \max_{1 \leq m} r(z_i) \). Then \( |z_i - z_p| \leq |z_i - x| + |x - z_p| \leq 2r_p \), hence \( B_{r(z_i)}(z_i) \subset B_{3r_p}(z_p) \) for \( 1 \leq m \). Consequently, \( y_{\lambda_i} \in F(W_{\lambda_i}) \subset F(B_{3r_p}(z_p)) \cap \text{int} D \) for \( 1 \leq m \), hence \( f_\varepsilon(x) \in F(z_p) + B_\varepsilon(o) \) (since this set is convex) \( \subset F(D \cap B_\varepsilon(x)) + B_\varepsilon(o) \), since \( |x - z_p| < r_p = \delta(z_p, \varepsilon)/3 < \varepsilon \).

For the second part: Notice that the compact \( K \) can be covered by finitely many \( D_r(x) \), say \( x = x_1, \ldots, x_m \); then \( f_\varepsilon(x) = \sum_{1 \leq m} \varphi_{\lambda_i}(x) y_{\lambda_i} \) is finitely dimensional (i.e. \( f_\varepsilon(D) \subset Y_0 \subset Y \) with \( \dim Y_0 < \infty \)); the refinement is not necessary here.

By means of the second part of Theorem 1.2, it is almost trivial to extend the classical topological degree theory to usc \( F \) with compact convex values (and such that e.g. \( F(D) \) is compact); see Chap. 8 of [2].

In particular one gets useful fixed point theorems, for example

**Theorem 1.3.** Let \( X \) be a Banach space, \( D = X \text{ bdd}, F : D \to 2^X \neq \emptyset \text{ usc,} \)

\( F(D) \) compact and \( F(x) \) closed convex (\( \forall x \in D \)). Suppose also that one of the following conditions holds:

\[
\text{(a) } N^{-1}F(x) \text{ is a compact convex set for each } x \in D \text{; (b) } \text{conv}(F(D)) \text{ is compact; (c) } F(D) \text{ is a uniformly equicontinuous family of functions.}
\]

Then \( f_\varepsilon \) has a fixed point for each \( \varepsilon > 0 \).
(a) D open and \( x_0 + \lambda(x-x_0) \notin F(x) \) on \( \partial D \) for all \( \lambda > 1 \) and some \( x_0 \in D \).

(b) D closed convex and \( F(D) \subseteq D \).

Then \( F \) has a fixed point (i.e. an \( x \in D \) such that \( x \in F(x) \)).

Proof. (a) requires degree theory (see Chaps. 1,2 in [2]).

(b) can easily be reduced to (a), but we can also prove it directly to show how Theorem 1.2 can be applied: W.I.O.G. D compact convex (notice that \( D_0 = \text{conv} F(D) \) is compact convex, \( F(D_0) = D_0 \) and \( F|_{D_0} \) is usc since \( F^{-1}(A) = F^{-1}(A \cap D_0) \); we find continuous \( f_n : D \rightarrow \text{conv} F(D) \subseteq D \) such that \( f_n(x) \in F(D \cap B_{1/n}(x)) + B_{1/n}(0) \) on D; by Schauder's fixed point theorem \( f_n \) has a fixed point \( x_n \in D \), and, W.I.O.G., \( x_n \rightarrow x_0 \in D \) since D is compact. Hence

\[
x_n \in F(D \cap B_{1/n}(x_n)) + B_{1/n}(0) = F(D \cap B_{\delta}(x_0)) + B_{\epsilon}(0)
\]

\[
\subseteq F(x_0) + B_{\epsilon}(x_0) + B_{1/n}(0) \quad (\forall \text{ large } n)
\]

(given \( \epsilon > 0 \), choose \( \delta > 0 \) so that (*) holds by the usc of \( F \) at \( x_0 \), hence \( x_0 \in F(x_0) \) since \( F(x_0) \) is closed.

This theorem can be applied, for example, to find solutions of multi-valued differential equations. As a gag we prove (an extension of Schauder's theorem)

**Corollary 1.1.** Let \( X \) be a Banach space, \( K \subseteq X \) compact convex, \( f : K \rightarrow X \) continuous. Then there exists \( x_0 \in K \) such that

\[
|f(x_0) - x_0| = c(f(x_0), K).
\]

Proof. Let \( P \) be the metric projection onto \( K \); remember that \( P \) is usc, \( P(x) \) is compact convex, and \( P(X) \) is compact. Hence \( F : K \rightarrow 2^K \setminus \emptyset \), defined by \( F(x) = P(f(x)) \), satisfies (b) of Theorem 1.3. Hence, \( x_0 \in F(x_0) \) for some
\(x_0 \in K\), and therefore \(|x_0 - f(x_0)| = \rho(f(x_0), K)\).

Remarks. You have seen that uc is good for approximate (continuous) selections; the convexity of \(F(x)\) in Theorem 1.3 can be weakened, as we shall indicate later; for Theorem 1.2 the convexity was essential. In § 24 of [2] you find Theorem 1.3 applied to a thermostat problem which leads to a multivalued "integrodifferential equation of Volterra type"; experiments and intuition suggest existence of periodic solutions to this equation, but nobody can prove it so far.

Theorem 1.3(b) is called "Kakutani's fixed point theorem" in case \(X = \mathbb{R}^n\); "Ky Fan's theorem" for \(X\) locally convex and \(D\) compact, "Bohnenblust-Karlin theorem" if as stated (In German we have the saying: Names are sound and smoke).

Now, in some concrete problems (Stochastic differential equations, Optimal Control, etc.) one cannot expect to find continuous (approximate) selections, but one hopes to get measurable selections. Since measurability is much weaker than continuity one hopes to (and must) find them under much weaker assumptions on \(F\).

1.4. Measurable selections.

I guess it helps to recall some basic things from measure theory which everybody should know anyway. A pair \((\Omega, \mathcal{A})\) is called a measurable space if \(\Omega\) is a set and \(\mathcal{A}\) is a \(\sigma\)-algebra of subsets of \(\Omega\) (i.e. a subset of \(2^\Omega\) such that \(\emptyset \in \mathcal{A}, A \in \mathcal{A} \Rightarrow \Omega - A \in \mathcal{A}, A_1 \in \mathcal{A} (\forall i \in \mathbb{N}) \Rightarrow \bigcup_{i=1}^\infty A_i \in \mathcal{A}\)). For example, \((\mathbb{R}, d)\) a metric spaces and \(\mathcal{A}\) the corresponding Borel-\(\sigma\)-algebra \(\mathcal{B}(\Omega)\), i.e. the smallest \(\sigma\)-algebra \(\subset 2^{\Omega}\) containing all open sets [notice that \(2^\Omega\) is a \(\sigma\)-algebra containing all open sets, and the intersection of \(\sigma\)-algebras is again a \(\sigma\)-algebra].
In the special case \( n = \mathbb{R}^n \) we also have the \( \sigma \)-algebra \( \mathcal{L}(\mathbb{R}^n) \) of all Lebesgue-measurable subsets of \( \mathbb{R}^n \), and for \( n \in \mathcal{L}(\mathbb{R}^n) \) we have \( \mathcal{L}(\mathbb{R}^n) = \{ M \cap n : M \in \mathcal{L}(\mathbb{R}^n) \} \), the \( \sigma \)-algebra of all Lebesgue-measurable subsets of \( n \). Recall also that \( \mathcal{L}(\mathbb{R}^n) \) is the smallest \( \sigma \)-algebra containing \( \mathcal{B}(\mathbb{R}^n) \) and \( \{ A \subset \mathbb{R}^n : A \in \mathcal{B} \text{ for some } B \in \mathcal{L}(\mathbb{R}^n) \text{ with } \nu_n(B) = 0 \} \), \( \nu_n \) the \( n \)-dimensional Lebesgue-measure.

Given two measurable spaces \((D, \mathcal{K})\) and \((n, \mathcal{L})\), a map \( f : D \to n \) is called \((\mathcal{K}, \mathcal{L})\)-measurable if \( f^{-1}(B) \in \mathcal{K} \) (\( \forall B \in \mathcal{L} \)). [Remember: If \( D \) and \( n \) are metric, continuity means \( f^{-1}(O) \) open whenever \( O \) is open; but now we only require e.g. \( f^{-1}(O) \) is Borel, for example, the intersection of countably many open sets; hence measurability is in general much less stringent than continuity].

In the sequel, image spaces \( n \) will be metric and \( \mathcal{L} = \mathcal{L}(n) \).

To check measurability of \( f \), it is then enough to show \( "f^{-1}(V) \in \mathcal{K}\) for all open \( V\)" or equivalently \( f^{-1}(M) \in \mathcal{K} \) for all closed \( M \) [reason: \( \{ W \subset n : f^{-1}(W) \in \mathcal{K} \} \) is a \( \sigma \)-algebra (as you check easily) containing all open sets, hence containing \( \mathcal{L}(n) \), since this is the smallest one containing all open sets].

If \( f_n : D \to n \) is measurable and \( f_n(x) \to f(x) \) on \( D \) then \( f \) is measurable too (notice: \( f^{-1}(M) = \bigcap_{p \geq 1} \bigcap_{m \in M} f_n^{-1}(M') \) if \( M \) is closed and \( M' = \{ x \in \mathbb{R} : \rho(x, m) < 1/p \} \).

In the special case \( D \in \mathcal{L}(\mathbb{R}^n) \) and \( n = \mathbb{R}^n \), \( f : D \to \mathbb{R}^n \) is said to be Lebesgue-measurable if it is \( (\mathcal{L}(D), \mathcal{L}(\mathbb{R}^n))\)-measurable, i.e. if \( f^{-1}(V) \) is Lebesgue-measurable for all open \( V \). If, in particular, \( m = 1 \), then you only have to check that \( \{ x \in D : f(x) < r \} \in \mathcal{L}(D) \) for every \( r \in \mathbb{R} \), since this implies \( \{ x \in D : f(x) \geq r \} \in \mathcal{L}(D) \), \( f^{-1}([a, b)) \in \mathcal{L}(D) \) and therefore \( f^{-1}(V) \in \mathcal{L}(D) \) for every open \( V \) (since such a \( V \) is a countable union of intervals \( [a, b) = \{ x \in \mathbb{R}^n : a_i \leq x_i < b_i \text{ for } i = 1, \ldots, m \} \).
Now, consider multis $F : D \to 2^\Omega \setminus \emptyset$. We could say again that $F$ is $(\mathcal{A}, \mathcal{D})$-measurable if $F^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{D}$; but then there is no simple characterization since $(\omega \in \Omega : F^{-1}(\omega) \in \mathcal{A})$ need not be a $\sigma$-algebra this time; notice that we only have
\[ D \setminus F^{-1}(\omega) \subseteq F^{-1}(\omega \setminus W) = \{ x \in D : F(x) \cap (\omega \setminus W) \neq \emptyset \}. \]
However, it is still useful to consider multis $F$ such that, for example, $F^{-1}(M) \in \mathcal{A}$ for all closed $M$, or for all open $M$, and this is what some people call a measurable multi.

**Proposition 1.2.** Let $(D, \mathcal{A})$ be a measurable space, $(\Omega, d)$ a metric space and $F : D \to 2^\Omega \setminus \emptyset$. Then
\begin{enumerate}[(a)]  
\item If $F^{-1}(M) \in \mathcal{A}$ for all closed $M \subseteq \Omega$ then $F^{-1}(V) \in \mathcal{A}$ for all $V \subseteq \Omega$. \[ \text{open} \]
\item If $\Omega$ is separable then: $F^{-1}(V) \in \mathcal{A}$ for all open $V \ni \rho(x, F(x))$ is $(\Omega, \mathcal{A}(\Omega))$-measurable for every $x \in \Omega$.
\end{enumerate}

**Proof.** (a) $V$ open $\Rightarrow V = \bigcup_{n \geq 1} M_n$ with $M_n = \{ x \in V : \rho(x, \Omega \setminus V) \geq 1/n \}$ closed. Hence $F^{-1}(V) = \bigcup_{n \geq 1} F^{-1}(M_n) \in \mathcal{A}$. (b) Notice that $F^{-1}(B_\rho(x)) = \{ z \in D : \rho(x, F(z)) < r \}$ and in a separable space every open $V$ is an at most countable union of open balls.

Now the basic result on measurable selections is

**Theorem 1.4.** Let $(D, \mathcal{A})$ be a measurable space, $(\Omega, d)$ a separable metric space and $F : D \to 2^\Omega \setminus \emptyset$ such that $F^{-1}(V) \in \mathcal{A}$ for all open $V \subseteq \Omega$ and $F(z)$ is complete for every $z \in D$. Then $F$ admits an $(\mathcal{A}, \mathcal{A}(\Omega))$-measurable selection.

**Proof.** By induction we define a sequence of measurable $f_n : D \to \Omega$ such that $
\rho(f_n(z), F(z)) < 2^{-n}$ and $d(f_{n+1}(z), f_n(z)) \leq 2^{-n+1}$ on $D$ for all $n$. Then we
are done since both properties and the completeness \( F(z) \) imply the existence of an \( f : D \to \mathbb{N} \) such that \( f_n(z) \to f(z) \) on \( D \) and \( f(z) \in F(z) \) on \( D \); in particular \( f \) is measurable.

Let \( \{x_n : n \geq 1\} \) be dense in \( \mathbb{N} \) (i.e. \( \cdots = n \)). Define \( f_0(z) = x_p \) where \( p \) is the smallest integer such that \( F(z) \cap B_1(x_p) \neq \emptyset \). Since \( f_0^{-1}(x_p) = F^{-1}(B_1(x_p)) \cup F^{-1}(B_1(x_m)) \in \mathcal{A} \) and \( f_0^{-1}(V) \) is an at most countable union of such \( f_0^{-1}(x_p) \), it is clear that \( f_0 \) is measurable.

Suppose we have already \( f_k \). Then \( z \in D_1 = f_k^{-1}(x_i) \) implies \( f_k(z) = x_i \) and \( \rho(f_k(z), F(z)) < 2^{-k} \), i.e. \( F(z) \cap B_{2^{-k}}(x_i) \neq \emptyset \). Therefore we define \( f_{k+1}(z) = x_p \) for \( z \in D_1 \) and \( p \) the smallest integer such that \( F(z) \cap B_{2^{-k}}(x_i) \neq \emptyset \). Thus, \( f_{k+1} \) is defined on \( D = \cup_{i \geq 1} D_i \), it is measurable and we have \( \rho(f_{k+1}(z), F(z)) < 2^{-k-1} \) and \( d(f_{k+1}(z), f_k(z)) = d(x_p, x_i) \leq 2^{-k-1} + 2^{-k} < 2^{-k+1} \) (on \( D_i \) for every \( i \)).

**Example 1.9.** Let \( F : \mathbb{R}^n \to 2^{\mathbb{N}} \setminus \emptyset \) be usc, \( F(\mathbb{R}^n) \) bdd and \( F(x) \) closed \((\forall x)\).

Let \( J = [0,a] \subset \mathbb{R} \), \( u \in C_{\text{loc}}(J) \) and \( M = \{ v \in L^\infty(J) : v(t) \in F(u(t)) \text{ a.e. in } J \} \) (almost everywhere in \( J \) means in \( J \setminus J_0 \) for some \( J_0 \in L^\infty(\mathbb{R}) \) with \( u_1(J_0) = 0 \)). For the thermostat problem mentioned earlier one needs \( M \neq \emptyset \). To show this, it is enough to find a measurable selection \( v \) of \( F \circ u \) (since \( F \) is bdd such a \( v \) is then bdd, i.e. in \( L^\infty(J) \)).

Now it is easy to see that \( F \circ u : J \to 2^{\mathbb{N}} \setminus \emptyset \) is usc, and \((F \circ u)(t)\) is closed for all \( t \in J \), hence complete. Since usc means \((F \circ u)^{-1}(A) \) closed for closed \( A \), Proposition 1.2 (a) shows \((F \circ u)^{-1}(V) \in \mathcal{B}(J) = L(J) \) for all open \( V \subset \mathbb{R} \), and therefore a measurable \( v \) exists.
Example 1.10. Let \( U \subseteq \mathbb{R}^m \) be compact, \( f : \mathbb{R}^n \times U \to \mathbb{R}^n \) continuous and \( x : J = [0,a] \to \mathbb{R}^n \) absolutely continuous such that \( x'(t) \in f(x(t),U) \) a.e. in \( J \). Then there exists a Lebesgue-measurable \( u : J \to U \) such that \( x'(t) = f(x(t),u(t)) \) a.e. in \( J \) (remember Example 0.4.).

To see this we have to find a measurable selection of \( F : J \to \mathcal{P}(U) \), defined by \( F(t) = \{ v \in U : x'(t) = f(x(t),v) \} \) (on the set of measure zero where \( x' \) is possibly not defined, define \( x'(t) = f(x(t),v_0) \) for a fixed \( v_0 \in U \)). Then \( F(t) \neq \emptyset \) on \( J \) and \( x' : J \to \mathbb{R}^n \) is measurable. By Lusin's theorem, given \( \varepsilon > 0 \) we find a closed \( J_\varepsilon \subseteq J \) such that \( \mu_1(J \setminus J_\varepsilon) < \varepsilon \) and \( x|_{J_\varepsilon} \) is continuous. Then \( F|_{J_\varepsilon} \) is usc with closed values (easy to check) and therefore, as in the previous example, there exists a measurable selection \( v_\varepsilon : J \to U \) of \( F|_{J_\varepsilon} \); now the same for \( J \setminus J_\varepsilon \) (with \( \varepsilon/2 \)), and so on, which means \( J = N \cup \bigcup_{i=1}^\infty J_i \) with \( J_i \) closed, \( \mu_1(N) = 0 \), and \( x' \) continuous on \( J_i \); this gives measurable \( v_i \) on \( J_i \); define \( v(t) = v_0 \) on \( N \), \( v(t) = v_i(t) \) on \( J_i \); then \( v \) is a measurable selection of \( F \) on \( J \).

Example 1.11. Consider the stochastic JVP

\[
(2) \quad x' = f(t,x,\omega), \quad x(0,\omega) = x_0(\omega) \quad (t \in J = [0,a], \omega \in \Omega),
\]

with \( f : J \times \mathbb{R}^n \times \Omega \to \mathbb{R}^n \) measurable in \( (t,\omega) \), continuous in \( x \) and

\[
(3) \quad |f(t,x,\omega)| \leq M(t,\omega)(1+|x|)
\]

with \( M(\cdot,\omega) \) measurable (\( \forall \omega \)) and \( \int_J M(t,\omega)dt < \infty (\forall \omega) \), \( x_0 : \Omega \to \mathbb{R}^n \) measurable. Then (2) has a solution which is measurable in \( \omega \) (and of course absolutely continuous in \( t \)). To see this, notice that (2) is equivalent to

\[
(4) \quad x(t,\omega) = x_0(\omega) + \int_0^t f(s,x(s,\omega),\omega)ds
\]
(i.e. we have to find a solution of (4) continuous in \( t \), measurable in \( \omega \)).

The approximate solutions

\[
x_m(t, \omega) = \begin{cases} x_0(\omega) & \text{for } t \leq 0 \\ x_0(\omega) + \int_0^t f(s, x_m(s-1/m, \omega), \omega) \, ds & \text{for } t \geq 0 \end{cases}
\]

are measurable in \( \omega \) (since obtained by successive insertion of measurable functions into \( f \)). By estimate (3) it is easy to see that \((x_m(\cdot, \omega))_{m \geq 1}\) is equicontinuous and bdd in \( C(J) \), for every fixed \( \omega \in \Omega \).

Now, consider \( X = C(J) \) with \( \| \cdot \|_0 \) and define \( F : \Omega \to 2^X \sim \emptyset \) by

\[
F(\omega) = \{ y \in X : y = \lim_{m \to \infty} x_k(\cdot, \omega) \text{ for some sequence } k_m \to \infty \}
\]

\( F(\omega) \) is closed (hence complete since \( X \) is complete). To check \( F^{-1}(V) \in \mathcal{A} \) for all open \( V \) we use the equivalent condition Proposition 1.2(b) and observe (choosing \((t_k)_{k \geq 1}\) dense in \( J \)) that

\[
\{ \omega : p(x, F(\omega)) < r \} = \bigcup_{m \geq 1} \lim_{j \to \infty} \bigcap_{k \geq 1} x_j(t_k, \cdot)^{-1}(B_{r-1/m}(x(t_k))) \in \mathcal{A}
\]

Thus \( F \) admits an \((\mathcal{A}, \mathcal{L})\)-measurable selection \( \tilde{x} : \Omega \to C(J) \) (\( \mathcal{L} \) the Borel \( \sigma \)-algebra of \( X \)). Define \( x(t, \omega) = \tilde{x}(\omega)(t) \). Then \( \tilde{x}(\omega) \in F(\omega) \) implies that \( x(\cdot, \omega) \) is a solution of (4) for this \( \omega \) [notice: every \( y \in F(\omega) \) is a solution of (4) for this \( \omega \) (take limit for \( m \to \infty \)) and \( x(t, \cdot) \) is measurable since it is the composition of \( \tilde{x} \) and the continuous map \( y \in C(J) \to y(t) \in \mathbb{R}^n \).

This is Theorem 1 in [3]; it is only motivated by stoch. ODEs but in fact a general "measurable dependence on parameters" theorem; it is in contrast to the continuous dependence theorem" (\( f \) continuous in \( \omega \to x(t, \cdot) \) continuous?) where one needs: for every \( \omega \in \Omega \), (2) has a unique solution \( x(\cdot, \omega) \) (e.g. \( f \) locally Lipschitz in \( x \)); consider the counter-example

\[
x'(t, \omega) = 2 \max \{0, x\}^{1/2} \omega, x_0(\omega) = 0 \text{ which has no continuous solution on } ([0,1] \times [-1,1]).
\]
Example 1.12. Let $J = [0,a] \subseteq \mathbb{R}$, $F : J \to 2^{\mathbb{R}^n} \setminus \emptyset$ continuous, $F(t)$ closed $(\forall t)$, $\nu : J \to \mathbb{R}^n$ Lebesgue-measurable. Then there is a Lebesgue-measurable $f : J \to \mathbb{R}^n$ such that $f(t) \in F(t)$ a.e. and $|\nu(t) - f(t)| = p(\nu(t), F(t))$ a.e. on $J$.

This is also a consequence of Theorem 1.4. and Proposition 1.2(b), since it means to find a measurable selection of $P_{F(t)}(\nu(t))$; notice e.g. that $p(\nu(t), F(t))$ is measurable (approximate $\nu$ by step functions and use the continuity of $F$ w.r. to $d_H$).

In § 24 of [2] you find more examples (also the control problem mentioned in Ex. c.4.), see also [4].

References


§ 2  Existence Theorems for Multivalued Differential Equations.

We consider the IVP

\[(1) \quad x' \in F(t,x) \text{ on } J = [0,a], \quad x(0) = x_0 \in \mathbb{R}^n\]

with \( F : J \times \mathbb{R}^n \to \mathbb{R}^n \setminus \emptyset \). By a solution \( x \) of (1) we understand an absolutely continuous (a.c.) \( x : J_\delta = [0,\delta] \to \mathbb{R}^n \) (for some \( 0 < \delta \leq a \)) with \( x(0) = x_0 \) and \( x'(t) \in F(t,x(t)) \text{ a.e. in } J_\delta \). [Recall: \( x \) is a.c. iff \( x(t) = x(s) + \int_s^t y(\tau) d\tau \) for some \( y \in L^1(J_\delta) \) and all \( t,s \in J_\delta \); and then \( x'(t) = y(t) \text{ a.e.} \).]

This solution concept is reasonable; in general one cannot expect \( C^1 \)-solutions and on the other hand I don’t know enough about weak conditions on \( F \) which guarantee existence of weaker solutions (say of bounded variation). Since the main emphasis in these lectures lies on the methods of proof, we make assumptions on \( F \) w.r. to \( (t,x) \), i.e. we do not discuss to which extent separate assumptions in \( t \) and \( x \) lead to existence [remember e.g.: in the singlevalued case continuity in \( x \), measurability in \( t \) and \( |f(t,x)| \leq M(t) \) with \( M \in L^1(J) \) are enough for a.c. solutions]; this means that we could restrict ourselves to the autonomous case and treat the general one as \( x' \in F(x_{n+1},x) \), \( x_{n+1}' = 1 \), \( x(0) = x_0 \), \( x_{n+1}(0) = 0 \), as we do perhaps later.

Of course we start with the simple case.

2.1. Convex-valued right hand sides.

Theorem 2.1. Let \( J = [0,a] \subset \mathbb{R} \) and \( F : J \times \mathbb{R}^n \to \mathbb{R}^n \setminus \emptyset \) be such that \( F(t,x) \) is compact convex for all \( (t,x) \). Then

(a) If \( F \) is lsc then (1) has a \( C^1 \)-solution

(b) If \( F \) is usc then (1) has an a.c. solution.

While (a) is trivial (see Example 1.8.), all proofs of (b) require some knowledge about the \( w^* \)-topology of Banach spaces. So let us first recall a few facts about this.
Let $X$ be a real Banach space, $X^*$ its dual (i.e. the continuous linear $x^*: X \to \mathbb{R}$, with norm $|x^*| = \sup\{x^*(x) : |x| \leq 1\}$), $(X^*)^* = X^{**}$ the second dual of $X$. Then $x \in X$ can be considered as an element of $X^{**}$, by defining $(x^*, x) := (x, x^*) = x^*(x)$ (I write the functional always in the second place). Besides the norm topology on $X$ one uses successfully the weak topology $\sigma(X, X^*)$ on $X$; it is defined by the basic neighborhood system of $x = 0$, i.e.

$$U_0 = \{x \in X : |x^*(x)| < \varepsilon \text{ for all } x^n \in E, \varepsilon > 0 \text{ and } E \subset X^* \text{ finite}\}.$$  

The neighborhood system of any $x_0 \in X$ is then defined by $U(x_0) = x_0 + U_0$ and $A \subset X$ is weakly open if to every $x_0 \in A$ there exists $U \in U_0$ such that $x_0 + U \subset A$. Convergence of a sequence $(x_n)$ to $x_0$ in this topology means $(x_n, x^*) \to 0$ (\forall x^* \in X^*); it is called weak convergence and denoted by $x_n \rightharpoonup x_0$.

The same can be done for $X^* : \sigma(X^*, X^{**})$ the weak topology on $X^*$; but more important is $\sigma(X^*, X)$, the weak*-topology of $X^*$, where the basic neighborhood system of $x^* = 0$ is given by

$$U^* = \{x^* \in X^* : |x^*(x)| < \varepsilon \text{ for all } x \in E, \varepsilon > 0 \text{ and } E \subset X \text{ finite}\}.$$  

Convergence of $(x^*_n)$ to $x^*_0$ in this topology means $x^*_n(x) \rightharpoonup x^*_0(x)$ (\forall $x \in X$), the $w^*$-convergence, denoted by $x^*_n \rightharpoonup x^*_0$. However, since these basic neighborhood systems are in general not countable, consideration of sequences for the description of topological concepts (closedness, compactness etc.) is not enough (one has to use nets or filters), but in some interesting cases sequences are really enough.

For example, if $X^*$ is separable and $M = X$ is bounded then $(M, \sigma(X, X^*))$ is metrizable, i.e. there is a metric which describes exactly the weak topology induced on $M$, and therefore sequences are enough (in every
metric space). For our purpose more important is

**Proposition 2.1(a)** Let $M \subseteq X^*$ be closed bounded. Then $(M, \sigma(X^*, X))$ is metrizable iff $X$ is separable.

(b) ** Alaoglu's Theorem:** $M \subseteq X^*$ is $w^*$-compact iff $M$ is bounded and $w^*$-closed.

Now, since $\bar{B}_r(x) \subseteq X^*$ is bounded and $w^*$-closed, it is $w^*$-compact; hence if $X$ is separable and $|x| \leq r \ (\forall n)$, there is subsequence $(x_n^*)$ such that $x_n^* \to x^*$ for some $x^*$ with $|x^*| \leq r$. (This is what we use in the sequel). This is classical linear Functional Analysis and can be found in every reasonable introduction to it, e.g. Taylor, Dunford/Schwartz, Hille/Phillips, Yosida.

**Proof of Theorem 2.1(b).** 1. Since $F$ is usc we have $F(t,x) = F(c, o) + \bar{B}_r(o)$ on $D = J_0 = \bar{B}_c(x_0)$, $\delta > 0$ sufficiently small. Since $F(c, o)$ is bounded, this means $F(t,x) = \bar{B}_c(o)$ on $D$ for some $c > 0$. By Theorem 1.2 there exist continuous $f_n : D \to \text{conv } F(D) \subseteq \bar{B}_c(o)$ such that $f_n(t,x) \in F(D \cap B_1/n(t,x) + B_1/n(o)$. Therefore $x = f_n(t,x)$, $x(\cdot) = x_0$ has a $C^1$-solution $x_n$ on $J_\rho$ for $\rho = \min(\delta, \delta/c)$, by the classical Peano Theorem.

Since $f_n : D \to \bar{B}_c(o)$, $(x_n)$ is equi-continuous and bounded in $C(J_\rho)$, hence w.l.o.g. $x_n \to x$ in $C(J_\rho)$ by Ascoli/Arzelà. Furthermore, $|x_n'| \leq c$ for all $n$, and that is all we know, since $(f_n)$ has no good convergence properties. But:

2. Consider $x_n'$ as element of $L^\infty(J_\rho)$ ("the space of bounded measurable functions"), i.e. $(x_n') = \bar{B}_c(o) \subseteq L^\infty(J_\rho) = (L^1(J_\rho))^*$. Since $L^1(J_\rho)$ is separable, we have w.l.o.g. $x_n' \to y \in L^\infty(J_\rho)$ by Proposition 2.1(b), i.e. $\int_J x_n(t)z(t)dt \to \int_J y(t)z(t)dt \ (\forall z \in L^1(J_\rho))$. Hence
\[ x(t) = x_n(t) = x_0 + \int_0^t x_n(s) ds = x_0 + \int_0^t \int_{\mathbb{R}^n} y(s) x_{\lfloor o, t \rfloor}(s) ds \]

and therefore
\[ y(t) = x'(t) \text{ a.e., } |x'(t)| \leq c \text{ a.e.} \]

Now, since \( L^\infty(J) \subseteq L^1(J) \), consider \((x'_n)\) and \(x'\) as elements of \( L^1(J) \); then we have in particular \((x'_n, z) \to (x', z)\) for all \(z \in L^\infty(J)\), i.e. \(x'_n \to x'\) in \(L^1(J)\).

3. We have
\[ x'_n(t) = f_n(t, x_n(t)) \in F(D \cap B_{1/n}(t, x(t))) + B_{1/n}(0) = F(D \cap B_n(t, x(t))) + B_n(0) \]

for all \(n \geq n_0\) uniformly on \(J\), since \(x_n \to x\) uniformly and \(F\) is u.s.c. Now
\[ C_n = \{ v \in L^1(J) : v(t) \in \text{conv} F(D \cap B_n(t, x(t))) + B_n(0) \text{ a.e.} \} \]

is closed convex [remember that a sequence converging in \(L^1\) has an a.e. convergent subsequence]. But for convex subsets closedness is the same as weak closedness; hence \(x'_n \in C_n\) for all \(n \geq n_0\) implies \(x' \in C_n \quad (\forall n > 0)\), and therefore \(x'(t) \in F(t, x(t))\) a.e. since \(F(t, x(t))\) is closed convex and \(F\) is u.s.c. \(\Box\)

**Example 2.1 (ODEs with discontinuous right hand side).** Remember Example 0.5.

Given \(x' = f(x)\) with \(f : \mathbb{R}^n \to \mathbb{R}^n\) locally bounded, consider

\[ (2) \quad x' \in F(x) = n \lim_{\varepsilon \to 0} \text{conv} f(x + B_{\varepsilon}(0)), \quad x(0) = x_0. \]

Evidently, \(F\) is u.s.c and \(F(x)\) compact convex for all \(x \in \mathbb{R}^n\). Hence (2) has a (local) a.c. solution \(x\). At all continuity points of \(f\) we have \(x'(t) = f(x(t))\) since \(F(x(t)) = \{f(x(t))\}\) there and at \(t\) such that \(f\) is continuous in a neighborhood of \(x(t)\) we have \(x \in C^1\).

**Remarks.** 1. From this local result it is easy to get global ones along traditional lines. In the u.s.c case \((\mathbb{R}^n\) instead of \(B_p(x_0))\) a global solution exists if \(F(t, x) \subseteq B_{r(x)}(0)\) with \(r(x) = c(1 + |x|)\) for some \(c > 0\).
If $F$ is even continuous, then $m(F(t,x)) := p_F(t,x)(o)$ is a continuous selection and therefore $|m(F(t,x))| \leq c(1+|x|)$ is sufficient for global existence. In the usc case it is enough to assume that $m(F(t,x))$ is bounded (choose always $y_{\lambda} = m(F(t_{\lambda},x_{\lambda}))$ with $(t_{\lambda},x_{\lambda}) \in W_{\lambda}$ in the proof to Theorem 1.2).

2. **The fixed point method:** If $x$ is a solution of (1) on $J_{p}$ then $|x'(t)| \leq c$, i.e. $x$ is Lipschitz with constant $c$ (from the proof just given). So let

$$K = \{x \in C(J_{p}) : x \text{ is Lipschitz with constant } c, \; x(o) = x_{o}\}$$

which is compact in $C(J)$. Since a solution satisfies $x' \in F(t,x)$ a.e., it is a fixed point of $G : K \rightarrow 2^K$, defined by

$$G(x) = \{y \in K : y'(t) \in F(t,x(t)) \text{ a.e. in } J_{p}\}$$

Since $x(\cdot)$ is continuous we have $H = F(\cdot,x(\cdot))$ usc on $J$ (if $F$ is usc), and $H(t)$ closed if $F(t,x)$ is closed; hence Theorem 1.4 yields a measurable selection $z$ of $H$ (cp. Ex. 1.9); let $y(t) = x_{o} + \int_{0}^{t} z(s)ds$; then $y \in G(x)$, i.e. $G(x) \neq \emptyset \; (\forall x \in K)$. If $F(t,x)$ is also bounded and convex then the weak convergence argument shows that $G$ is usc and $G(x)$ compact convex ($\forall x \in K$), and therefore $G$ has a fixed point, by Theorem 1.3.

3. **The integral method:** Given $H : J \rightarrow 2^{R_{+}} \neq \emptyset$ define $\int_{J} H(t)dt := \int_{J} h(s)ds$; $h$ an integrable selection of $H$ (which may of course be $\emptyset$ without further assumptions on $H$). Now it is easy to see: If $F$ is usc and $F(t,x)$ compact convex ($\forall (t,x)$) then $x \in C(J)$ is an a.c. solution of $x' \in F(t,x)$ a.e. iff $x(t) = x(s) + \int_{s}^{t} F(\tau,x(\tau))d\tau \; (\forall t,s \in J)$ (non-trivial; = for a closed convex $A$ one has $\int_{A} A d\tau = (t-s)A$; use this and the facts that $x$ is Lipschitz (hence a.e. differentiable) and $F$ is usc). Now consider $x_{n}, f_{n}$ as in the proof to Theorem 2.1(b) and show $p(x(t)-x(s)) \int_{s}^{t} F(\tau,x(\tau))d\tau = 0$ (since $x_{n} \rightarrow x$ uniformly and $F(\cdot,x(\cdot))$ is usc; use Ex. 1.12 for a measurable selection $\psi_{n}$.
of $F(\cdot, x(\cdot))$ such that $|f_n(s, x_n(s)) - v_n(s)| = o(f_n(\cdot), F(s, x(s)))$.

4. Euler-Cauchy-Polygons. $v_0 = m(F(0, x_0)), x_1 = x_0 + p/k v_0$; given

$(\frac{j-1}{k} p, x_{j-1})$ for $j < k$, let $v_{j-1} = m(F(\frac{j-1}{k} p, x_{j-1}))$ and $x_j = x_{j-1} + p/k v_{j-1}$

call $x_k(\cdot)$ the polygon through these points and proceed as in the proof to Theorem 1(b).

Conclusion. No matter which proof you prefer, everything depends
heavily on the convexity assumption about $F(t, x)$.

2.2 The nonconvex case.

If $F(t, x)$ is not convex, the methods used so far do not work directly.
We have to look for modifications. For example, one can try to define
the Euler-Cauchy-Polygons more carefully, so that one can replace the
"weak convergence in $L^1(J)$-argument" by a compactness argument in $B(J)$,
the space of bounded functions on $J$ (remember that we had $|x_n(t)|_0 \leq c$
for all $n$). Recall that $\sup\{|x(t) - x(s)| : s, t \in J\} = \text{diam} x(J)$ is
(sometimes) called the oscillation of $x \in B(J)$ on $J$, and $A \subseteq B(J)$ is said
to be equioscillating if to each $\varepsilon > 0$ there exists a partition
$J_1, \ldots, J_k(\varepsilon)$ of $J$ that $\max_{i \in A} \sup_{J_i} \text{diam} x(J_i) \leq \varepsilon$.

Proposition 2.2. Let $X$ be a Banach space, $A \subseteq B(X)$ equioscillating
and $A(t) = \{x(t) : x \in A\}$ relatively compact ($\forall t \in J$). Then $A$ is
relatively compact.

Proof. Given $\varepsilon > 0$, consider a partition $J_1, \ldots, J_k$ of $J$ such that

$\sup_{A} \max_{i} \text{diam} x(J_i) \leq \varepsilon$; choose $t_i \in J_i$ and notice that $x \to (x(t_1), \ldots, x(t_k))$
is a compact map (by the second assumption) from $A$ to $X^k$, hence the
image can be covered by finitely many ε-balls, i.e. there exist
\( x_1, \ldots, x_n \in A \) (for some \( n \)) such that \( \min_{j \leq n} \max_{t_i \leq k} |x(t_i) - x_j(t_i)| \leq \varepsilon \) (\( \forall x \in A \)).
and therefore \( A \subset \bigcup_{j=1}^{n} B_{3\varepsilon}(x_j) \).

Now, if \( F \) is continuous (hence uniformly continuous on \( J = \mathbb{R}_+(x_0) \)) and \( F(t, x) \) is compact then it is possible to define polygons \( x_n(\cdot) \) such that \( A = (x_n') \) satisfies Proposition 2.2, and therefore a convergent subsequence yields a local solution to (1). This is Theorem 1 in Chap. 2.3 of [1]. We shall not present the proof since by a different method it is possible to show that a solution exists if \( F \) is only I-sc, i.e. we have

**Theorem 2.2.** Let \( J = [0, a] \subset \mathbb{R}_+, x_0 \in \mathbb{R}^n \), \( F : J \times \mathbb{R}_+(x_0) \rightarrow 2^{\mathbb{R}^n} \)
I-sc, \( F(t, x) \) compact (\( \forall (t, x) \)) and \( F(J \times \mathbb{R}_+(x_0)) = B_c(0) \). Then (1) has a (local) a.c. solution.

The idea of proof is this: Consider the compact convex set \( K = \mathcal{C}(J_0) \) from (3) and the multi \( \mathcal{G} : K \rightarrow 2^K \) from (4); we want to show that \( \mathcal{G} \) admits a continuous selection \( \mathcal{G} \); then we are done since \( \mathcal{G} \) has a fixed point by Schauder's theorem, i.e. \( x = \mathcal{G}(x) \in \mathcal{G}(x) \). To find this selection we proceed similar to the proof of Theorem 1.1. This time the situation is on the one hand more complicated (no use of simple convex combinations) but on the other hand it is very special since we have now a compact set \( D = K \) and the very special space \( C(J_0) \). Since \( K \) is compact it is always enough to consider finite open coverings, say \( K = \bigcup_{i=1}^{m} V_i \), and therefore finite partitions of unity, say \( \varphi_1, \ldots, \varphi_m \) with \( \text{supp} \varphi_i \subset V_i \). With such a partition we may define (the only new idea compared to the general proof)
what one might call a continuous partition of $J_\rho$ by setting (for $x \in K$)

$$\tau_0(x) = 0, \quad \tau_i(x) = \tau_{i-1}(x) + \varphi_i(x) \rho, \quad J_i(x) = [\tau_{i-1}(x), \tau_i(x)) \quad (\text{for } i = 1, \ldots, m)$$

(the last interval closed at $\tau_m(x) = \rho$); notice that the $J_i(x)$ are disjoint, some of them may be empty, but their union is $J_\rho$ since $\sum_{i=1}^m \varphi_i(x) = 1$ on $K$.

Now the essentials for an induction proof (like to Theorem 1.1) are contained in

**Proposition 2.3(a)** Let $J_1(\cdot), \ldots, J_m(\cdot)$ be a continuous partition of $J$ and $v_i \in B_R(\sigma) \subseteq L^\infty(J)$ for $i = 1, \ldots, m$. Then $v = \sum_{i=1}^m v_i x_{J_i(\cdot)} \in C^1(L^1(J)(K))$. i.e. $v : K \to L^1(J)$ is continuous.

(b) Let $M = \mathbb{R}^n$ compact and $H : M \to \mathbb{R}^n \setminus \emptyset$ lsc. Then to $\varepsilon > 0$ there exists $n(\varepsilon) > 0$: to each continuous $u$ such that $(t, u(t)) \in M$ on $J_u$ there is a measurable $v : J_u \to \mathbb{R}^n$ such that $\rho(v(t), H(t, x)) \leq \varepsilon$ whenever $(t, x) \in M$ and $|x-u(t)| \leq n(\varepsilon)$.

**Proof (a)** is trivial:

$$\int_J |v(x)(t) - v(y)(t)| dt \leq R \sum_{i=1}^m \int_J |x_{J_i(x)}(t) - x_{J_i(y)}(t)| dt =$$

$$= R \sum_{i=1}^m v_i (J_i(x) \Delta J_i(y)) \to 0$$

for $x \to y$ (remember $A \Delta B = A \setminus B \cup B \setminus A$) since the $\tau_i(\cdot)$ are continuous.

(b) is a little bit more complicated and will be proved later.

**Proof of Theorem 2.2. Step 1:** By Proposition 2.3(b) (with $H = F$ and $\varepsilon = 1$) we find $n_0 = n(1) > 0$, $K \subseteq \cup_{i=1}^{m_0} B_{x_i}(x_i)$ and measurable $v_i$ corresponding to $x_i \in K_i$; let $J_1(\cdot), \ldots, J_{m_0}(\cdot)$ be a corresponding continuous partition of $J_\rho$ and $g_0 = \sum_{i=1}^{m_0} v_i x_{J_i}(\cdot)$.

Then $\sup_{K} \sup_{J} \rho(g_0(x)(t), F(t, x(t))) \leq 1$, since $x_{J_i}(x)(t) = 1$ means $g_0(x)(t) = v_i(t)$ and $\varphi_i(x) \to 0$ (notice $\chi_i = 0$), hence $x \in B_{n_0}(x_i)$ and therefore $\rho(v_i(t), F(t, x(t))) \leq 1$. 

Step 2. We prove by induction that for $n = 0,1,2,\ldots$ we find $g_n = \sum_{i=1}^{m_n} v_i x_j^n i.o.$ such that

(1) $\sup_K \sup_{J_P} \rho(g_n(x)(t), F(t,x(t))) \leq 1/2^n$ (for all $n \geq 0$)

(11) $\sup_K |g_n(x) - g_{n-1}(x)|_{L^1(J_P)} \leq \alpha/2^n$ (for some $\alpha > 0$ and all $n \geq 1$).

Suppose we have (1), (11) for $n = k-1$. Then (by Lusin's Theorem) all $v_i^{k-1}$ are continuous on $J \sim S_k$, for some open $S_k$ with $\mu(S_k) < 1/2^{k+1}$ (we write in the sequel: $J$ for $J_P$, $\mu$ for $\mu_1$). Choose $\delta > 0$ such that

$max_{i=1}^{m_{k-1}} |\tau_i^{k-1}(x) - \tau_i^{k-1}(\bar{x})| < (2^{k+2} m_{k-1})^{-1}$ for $|x-\bar{x}| \leq \delta$, choose finitely many $x_j \in K$ such that $K = \cup U_j(x_j)$ and let

$E_j = \cup_{i=1}^{m_k} \{ t \in J : |t - \tau_i^{k-1}(x_j)| < (2^{k+2} m_{k-1})^{-1} \}.$

Then

(5) $x \in B_{\delta}(x_j)$ and $t \in J \sim E_j \Rightarrow g_{k-1}(x)(t) = g_{k-1}(x_j)(t) = v_i^{k-1}(t)$

for some $i$ (hence $g_k(x)(\cdot)$ is continuous on $J \sim (E_j \cup S_k)$).

Notice: $x_j^{k-1}(x)(t) = 1 \Rightarrow g_{k-1}(x)(t) = v_i^{k-1}(t)$ and $\tau_i^{k-1}(x) \leq t < \tau_i^{k-1}(x)$;

If also $x_j^{k-1}(x_j)(t) = 1$ then $g_{k-1}(x_j)(t) = v_i^{k-1}(t)$ and $\tau_i^{k-1}(x_j) < t < \tau_i^{k-1}(x_j)$; hence $t \notin E_j$ and the choice of $\delta$ imply $p = q$; draw a picture of the intervals. Now, let

$p_j^k(t) = \begin{cases} 2c+1 & \text{on } E_j \cup S_k \\ 1/2^{k-1} & \text{on } J \sim (E_j \cup S_k) \end{cases}$ (w.l.o.g. $c > 1$ in the Theorem)

and

$F_j(t,y) = [g_{k-1}(x_j)(t) + p_j^k(t) B_{1+\lambda}(0)] \cap F(t,y)$ with $\lambda > 0$.

We have $F_j(t,y) \neq \emptyset$ for $t \in J$ and $|y - x_j(t)| \leq \delta$ (indeed: $t \in E_j \cup S_k \Rightarrow F(t,x) \subseteq B_{\delta}(0) \subseteq \cdots$ since $p_j^k(t) = 2c$ and $|g_{k-1}(u_j)| \leq c+1$ (notice that the $|v_i^{k-1}(t)| \leq c+1$); $t \in J \sim (E_j \cup S_k)$: Let $x(\cdot) = x_j(\cdot) + y - x_j(t)$, which
satisfies \( x \in B_\delta(x_j) \) and \( x(t) = y \), hence \( g_{k-1}(x)(t) = g_{k-1}(x_j)(t) \) and
by induction hypothesis \( \rho(x_{k-1}(t), F(t, y)) \leq 1/2^{k-1} \). Furthermore, \( F_j \) is lsc on \( \{(t, y) : t \in E_j, |y - x_j(t)| \leq \delta \} \) [notice: \( \delta_j \) is lsc, \( F \) is lsc, \( \delta_k \) is lsc, \( \delta_j \) is continuous on \( J \subset (E_j \cup S_k) \) and \( F_j(t, y) = F(t, y) \) for \( t \in E_j \cup S_k \)].

Therefore we can apply Proposition 2.3(b) to \( F_j, \delta = 1/2^{k-1} \).

\( n_j = n(1/2^{k+1}) \); choose finitely many \( x_j \) such that \( K \cap B_\delta(x_j) = \bigcup \bigcup (x_j) \).

choose corresponding \( v_j \) and a continuous partition \( (J_j) \) of \( J \) corresponding to \( \bigcup \bigcup B_\delta(x_j) \cap B_{\delta_j}(x_j) \), and let
\[
\bar{g}_k(x) = \sum_{i, j} v_j q_j x_j \Delta J_j(x_j) .
\]

If \( x_{j,i} \) then \( g_k(x)(t) = v_{j,i}(t) \) and \( x \in B_{\delta_j}(x_{j,i}) \), hence \( \rho(x_{j,i}, F_j(t, x(t))) \leq 1/2^{k+1} \) and therefore (i) for \( n = k \) (since \( F_j(t, x) \subset F(t, x) \)).

To see (ii) for \( n = k \), notice that \( \mu(E_j \cup S_k) \subset \mu(E_j) + \mu(S_k) \leq 1/2^{k+1} + 1/2^{k+1} = 1/2^k \) and \( |g_k(x)(t) - g_{k-1}(x)(t)| \leq 2(c+1) \) on \( E_j \cup S_k \); for \( t \in E_j \cup S_k \) we have \( (|y - x_j(t)| \leq \delta \)
\[
F_j(t, y) \subset g_{k-1}(x_j)(t) + \delta \frac{1}{2^{k-1}} \Delta S \subset g_k(x)(t) \Delta S, \quad (x \in B_\delta(x_j))
\]
\[
g_k(x)(t) = v_{j,i}(t) \quad \text{for some } i, \quad \text{hence}
\]
\[
\rho(x_{j,i}, g_k(x)(t) + \delta \frac{1}{2^{k-1}} \Delta S) \leq \rho(x_{j,i}, F_j(t, y)) \leq 1/2^{k+1} ,
\]
and therefore \( |g_k(x)(t) - g_{k-1}(x)(t)| \leq 3/2^k \). This means
\[
|g_k(x) - g_{k-1}(x)|_{L^1(J)} \leq \alpha/2^k \quad \text{for } \alpha = 2(c+1) + 3\delta \text{ and all } x \in K.
\]

**Step 3.** For \( x \in K \), (i) says that \( (g_n(x)) \) is Cauchy in \( L^1(J) \), hence convergent to some \( g(x) \in L^1(J) \). Since an \( L^1(J) \)-convergent sequence has an a.e.-convergent subsequence, (i) shows \( g(x)(t) \in E \), a.e. (in particular \( |g(x)(t)| \leq c \)). Let \( \omega(x)(t) = x_0 + \int_0^t g(x)(s)\,ds \). Then \( \omega : K \to K \) is continuous, by the uniformity w.r. to \( K \) in (i) and (ii) and the fact \( g_n \in C_0^1(J) \).
Proof of Proposition 2.3(b). We write \( z, \rho \) for \( (t,x), (t,z) \) ... Let 

\[ \eta_\varepsilon(z) = \sup(\eta > 0 : \eta \in B_{\eta}(z) \cap M) \]

for \( z \in M \).

[notice: to \( y \in H(z) \) and \( \varepsilon > 0 \) there exists \( \eta > 0 : H(p) \cap B_\varepsilon(y) \neq \emptyset \) for \( \rho \in B_\eta(z) \cap M \), hence \( y \in H(p) + B_\varepsilon(0) \) for these \( \rho \). It is easy to see that \( \eta_\varepsilon \) is continuous [Indeed: \( |z-\hat{z}| < \sigma/3 \) implies \( B_1 = B_{\eta_\varepsilon(z)-2\sigma/3}(\hat{z}) \)]

\[ \cap M = B_{\eta_\varepsilon(z)-\sigma/3}(z) \cap M \]

\[ B_2, \text{ hence } \eta \in B_{\eta_\varepsilon(z)-\sigma/3}(z) \cap M \] implies

\[ \eta \in (H(p)+B_\varepsilon(0)) \neq \emptyset \text{ and therefore } \eta_\varepsilon(z) \geq \eta_\varepsilon(\hat{z})-2\sigma/3 \text{ and interchange of } z \text{ and } \hat{z} \text{ yields } \eta_\varepsilon(z) \geq \eta_\varepsilon(z)-2\sigma/3 \]. Hence \( \eta_\varepsilon(z) \geq \eta(z) \) on \( M \) since \( M \) is compact.

Now it is easy to see that \( H_\varepsilon : M \to \mathbb{R}^n \setminus \emptyset \), defined by \( H_\varepsilon(z) = B_{\eta_\varepsilon(z)}(z) \cap M \) is lsc, and therefore \( A_\varepsilon \), defined by \( A_\varepsilon(z) = \overline{A_\varepsilon(z)} \cap M \), is lsc with closed values. Hence, for a continuous \( u \), \( A_\varepsilon(\cdot, u(\cdot)) \) is lsc with closed values and therefore the measurable selection theorem gives a measurable \( v : v(t) \in A_\varepsilon(t, u(t)) \) on \( J \); this means \( \rho(v(t), H(t,x)) \leq \varepsilon \) for \( (t,x) \in M \) with \( |x-u(t)| \leq \eta(z) \).

Remark. This was essentially the proof given to Theorem 1 in Chap. 2.6 of [1]. The proof given there contains 7 errors (not due to the printer) and the Theorem is preceded by 4 completely superfluous pages on which the same thing is proved under the assumption that \( F \) be continuous. This is only one out of many instances showing that the authors and the responsible editor have lost their mind publishing such a junk.

2.3 A relation between the convex and the nonconvex case.

Under the strong condition that \( F \) is Lipschitz w.r. to \( \delta_H \) it can be shown that the solutions of \( x' \in F(t,x) \) are dense in the set of solutions of \( x' \in \overline{\text{conv}} F(t,x) \); control people call this relaxation, since the convex-valued case is usually much easier.
We start with a comparison result which is at the same time an existence theorem.

**Lemma 2.1.** Let $y : J = [0, a] \rightarrow \mathbb{R}^n$ be a.c., $D = \{(t, x) : t \in J\}$, $|x-y(t)| \leq r$, $F : D \rightarrow \mathbb{R}^n$ be continuous, $F(t, x)$ closed: $(\forall (t, x) \in D)$, $d_{\mathbb{H}}(F(t, x), F(t, \tilde{x})) \leq k(t) |x-\tilde{x}| (\forall (t, x), (t, \tilde{x}) \in D)$ with $k \in L^1(J)$, $|x_0-y(0)| = \varepsilon < r$, $\rho(y'(t), F(t, y(t))) \leq p(t)$ a.e. with $p \in L^1(J)$, $K(t) = \int_0^t k(s)ds$, $\varphi(t) = \alpha e^{K(t)} + \int_0^t e^{K(t)-K(s)} p(s)ds$ and $\alpha \in (0, a]$ such that $\varphi(t) \leq r$ on $J = [0, a]$.

Then (1) has an a.c. solution $x$ on $J_\alpha$ such that $|x(t)-y(t)| \leq \varphi(t)$ in $J_\alpha$.

**Proof.** Consider successive approximation in the following way. By Example 1.12 there is a measurable selection $v_0$ of $F(\cdot, y(\cdot))$ such that $|v_0(t)-y'(t)| = \rho(y'(t), F(t, y(t)))$ a.e. Let $x_1(t) = x_0 + \int_0^t v_0(s)ds$.

and in general, given $x_n$, choose a measurable selection $v_n$ of $F(\cdot, x_n(\cdot))$ such that $|v_n(t)-x_n'(t)| = \rho(x_n'(t), F(t, x_n(t)))$ a.e. and let $x_{n+1}(t) = x_n + \int_0^t v_n(s)ds$.

We have $|x_{n+1}'(t)-x_n'(t)| = \rho(x_n'(t), F(t, x_n(t))) \leq k(t) |x_n(t)-x_{n-1}(t)|$ a.e. $(x_0(t) := y(t))$ and, for example, $|x_1(t)-y(t)| \leq \delta + \int_0^t p(s)ds$, 

$|x_{n+1}(t)-x_n(t)| \leq \int_0^t k(s) |x_n(s)-x_{n-1}(s)|ds$ for $n \geq 1$, so that induction shows

$|x_{n+1}'(t)-x_n'(t)| \leq k(t) [\delta K(t)^{n-1} + \int_0^t \frac{K(t)-K(s)}{(n-1)} p(s)ds]^{n-1} p(s)ds \quad n \geq 1$

$|x_{n+1}(t)-x_n(t)| \leq \delta K(t)^n + \int_0^t \frac{K(t)-K(s)}{n} p(s)ds \quad n \geq 0$.

hence $x_n \rightarrow x$ on $J_\alpha$ (uniformly, for some $x \in C(J)$), by the second inequality.
hence \( x(t) = x_0 + \int_0^t v(s)ds \), \( |x(t)-y(t)| \leq \varphi(t) \), \( x' = v \), a.e. and
\( p(x_n'(t), F(t, x_n(t))) = |v_n(t)-x_n'(t)| \to 0 \) a.e., hence \( x'(t) \in F(t, x(t)) \) a.e.

By means of this lemma it is easy to prove the "Relaxation-Theorem" of A.F. Filippov (we consider only the autonomous case).

**Theorem 2.3.** Let \( F : \mathcal{B}_p(x_0) \subset \mathbb{R}^n \to \mathbb{R}^n \) be Lipschitz and \( F(x) \) compact for all \( x \in \mathcal{B}_p(x_0) \). Let \( x : J = [0, a] \to \mathcal{B}_p(x_0) \) be an a.c. solution of \( x \in \overline{\text{conv}} F(x) \), \( x(0) = x_0 \). Then to \( \varepsilon > 0 \) there is a solution \( y \) on \( J \) of \( y \in F(y) \), \( y(0) = x_0 \) such that \( |x-y|_0 \leq \varepsilon \).

**Proof.** Choose \( n > 0 \) such that \( |x(t)-x_0|+n \leq r \) on \( J \), let \( \varepsilon < n \). By Lemma 2.1 we only have to find an a.c. \( z \) such that \( |z-x_0| \leq \varepsilon / 2 \), \( z(0) = x_0 \), \( \rho(z'(t), F(z(t))) \leq c \varepsilon \) with an appropriate \( c \), since then there is an a.c. solution \( v \) of \( v \in \overline{\text{conv}} F(v) \), \( v(0) = x_0 \) such that \( |v(t)-z(t)| \leq \varphi(a) = \int_0^a k(a-s)c \ vds \leq c \varepsilon / 2 \) (for \( c = k/\{2(e^{ak}-1)\} \)),

where \( k \) is the Lipschitz constant of \( F \).

We have \( F(\mathcal{B}_p(x_0)) \subset \mathcal{B}_m(0) \) for some \( M > 0 \). Hence \( F(x(\cdot)) \) and \( \overline{\text{conv}} F(x(\cdot)) \) are Lipschitz of constant \( 1 = k \cdot M \). Now, consider the partition \( J = (t_i, t_{i+1}] \) with \( t_i = ia/n \) for \( 0 \leq i \leq n-1 \), and let \( C_1 = \overline{\text{conv}} F(x(t_i)) \). Then \( \overline{\text{conv}} F(x(t)) \subset C_1 + \mathcal{B}_{1/n}(0) \) for all \( t \in J_1 \). Let \( S_j (j = 1, \ldots, m) \) be a (disjoint) partition of \( \bigcup_{j \in J_1} \overline{\text{conv}} F(x(t)) \) with \( S_j \) Borel and \( \operatorname{diam}(S_j) \leq \rho \) for all \( j \) (\( \rho \) will be chosen later on). Let \( E_j = (x(\cdot))^{-1}(S_j), x_j = x_{E_j}, y_j \in S_j \) and \( y = \sum_{j=1}^m y_j \).

Since \( x'(t) \in \bigcup_{j \in J_1} S_j \) a.e., we have \( |y(t)-x'(t)| \leq \rho \) a.e. on \( J_1 \), and, since \( y_j \in C_1 + \mathcal{B}_{1/n}(0) \), we find \( \alpha_j > 0 \) with \( \sum_{k} \alpha_{jk} = 1 \) and \( z_{jk} \in F(x(t_i)) \) such that \( |y_j - \sum_{k} \alpha_{jk} z_{jk}| \leq 21a/n \). Now, \( E_j \) can be divided into measurable \( E_{jk} \) such that \( u(E_{jk}) = \alpha_{jk} u(E_j) \) (if \( E_j = [a, \omega] \), consider \( \varphi(t) = \int_a^t x_j(s)ds \), which is continuous increasing and ranges from \( a \) to \( u(E_j) \).
let \( \tau_k = \sup \{ t \in [\alpha, \omega] : \psi(t) \leq u(E_j) \Sigma \alpha_{jp} \} \), \( I_1 = [\alpha, \tau_1] \), \( I_k = (\tau_{k-1}, \tau_k] \) and let \( E_{jk} = E_j \cap I_k \); then \( E_j = U \cup E_{jk} \) (disjoint) and
\[
u(E_{jk}) = \int_{\tau_{k-1}}^{\tau_k} \chi_j(s)ds = \psi(\tau_k) - \psi(\tau_{k-1}) = \nu(E_j)\alpha_{jk}.
\]
Let \( x_{jk} = x_{E_{jk}} \), define \( \nu : J \to \mathbb{R}^n \) as the step function such that \( \nu\big|_{J_1} = \Sigma \chi_j x_{jk} \), and let \( z(t) = x_0 + \int_0^t \nu(s)ds \) (which is Lipschitz of constant \( M \)). Then we have
\[
\left| \int_{J_1} (\nu(s)-y(s))ds \right| = \left| \Sigma \nu(E_j) y_{j_1} - \Sigma \nu(E_j) \alpha_{jk} z_{jk} \right| \leq \Sigma \nu(E_j) |y_{j_1} - \Sigma \alpha_{jk} z_{jk}| \leq \frac{21a^2}{n}.
\]

hence
\[
\left| z(t_j) - x_0 \right| - \int_0^{t_j} y(s)ds \leq \frac{21a^2}{n},
\]

\[
\left| x(t_j) - x_0 - \int_0^{t_j} y(s)ds \right| \leq p - a
\]
and therefore \( |z(t_j) - x(t_j)| \leq 21a^2/n + p - a \). Now, for \( t \in J_1 \):
\[
|z(t) - x(t)| \leq |z(t) - z(t_j)| + |z(t_j) - x(t_j)| + |x(t_j) - x(t)| \leq 21a^2/n + p - a + 2aM/n,
\]
and if \( z'(t) \) exists then \( z'(t) = \nu(t) = z_{jk} \in F(x(t_j)) \), hence
\[
p(z'(t), F(z(t))) \leq k |z(t) - x(t_j)| \leq k \frac{Ma}{n} + 2k1a^2/n + kp - a,
\]
hence all conditions are satisfied for \( p = \varepsilon b \) with \( b \) small and \( n \) large.

The following counter-example of V. Pliss shows that without Lipschitz condition the result may be wrong.

**Example 2.2.** Consider \( F : \mathbb{R}^2 \to \mathbb{R}^2 \), defined by
\[
F(x) = \{ y : y_1 \in (-1,1) \, , \, y_2 = \sqrt{x_2} + |x_1| \}.
\]
Let \( x' \in F(x) \), \( x(0) = 0 \). Then \( x_2' = \sqrt{|x_2|} + |x_1|, x_1 \in (-1,1) \); evidently, \( x_1 \to 0 \) in a neighborhood of \( t = 0 \) is impossible, hence \( |x_1(t)| > 0 \) for some sequence \( \tau_n \to 0 \), and therefore \( |x_1(t)| \geq \varepsilon_n > 0 \) in \( \tau_n \omega_n \) for some \( \omega_n \), hence \( x_2' \geq \sqrt{|x_2|} + \varepsilon_n \) and \( x_2(\tau_n) \geq 0 \), i.e. \( x_2(t) \geq \frac{1}{2} (t - \varepsilon_n)^2 \) in \( \tau_n \omega_n \).
for \( t \geq \omega_n \) we have \( x_2^1 \geq \sqrt{|x_2^2|} \), \( x_2(\omega_n) = \frac{1}{4}(\omega_n - \tau_n)^2 \), hence
\[
x_2(t) \geq \frac{1}{4}(t - \tau_n)^2 \rightarrow \frac{1}{4} t^2, \text{ i.e. } x_2(t) \geq t^2/4 \text{ in } [0, \infty).
\]
The problem \( x' \in \text{conv } F(x), x(0) = 0 \) has the solution \( x(o) = 0 \), which therefore cannot be approximated by solutions of \( x' \in F(x), x(0) = 0 \).

2.4. Further existence results.

(1) Existence of \( C^1 \)-solutions.

Theorem 2 of § 2.2 in [1] yields a local \( C^1 \)-solution of \( x' \in F(x), x(0) = x_0 \) if \( F \) is absolutely continuous and \( F(x) \) is compact
\[(\forall x \in B_r(x_0)) \quad \text{given } y \in F(x_0) \text{ one also finds a } C^1 \text{-solution satisfying}
\]
in addition \( x'(0) = y \).

Here \( F : D \rightarrow \mathbb{R}^n \) is said to be absolutely continuous (w.r. to \( d_H \)) if to every \( \varepsilon > 0 \) there exists \( \delta_\varepsilon > 0 \): \( \sum d_H(F(x_i), F(x_{i+1})) \leq \varepsilon \) for every finite set of points \( x_i \in D \) satisfying \( \sum |x_i - x_{i+1}| \leq \delta_\varepsilon \). Such an \( F \) is evidently uniformly continuous on \( D \), and, of course, Lipschitz continuity implies a.c.

The idea of proof is to define approximate solutions with equicontinuous derivatives: Consider \( t_n^1 = \frac{1}{n} a; x_n(t_0^n) = x_0, y_n(t_0^n) = y; \) if \( x_n(t_n^k) \) and \( y_n(t_n^k) \) are defined for \( k \leq i-1 \), let \( x_n(t_n^i) = x_n(t_n^{i-1}) + \frac{a}{n} y_n(t_n^{i-1}) \) and choose \( y_n(t_n^i) \in \text{cl} F(x_n(t_n^i))(y_n(t_n^{i-1}))(\text{metric projection}) \), and define \( x_n : [0, \infty) \rightarrow \mathbb{R}^n, y_n : [0, \infty) \rightarrow \mathbb{R}^n \) as the polygons through these nodal points. Then it is easy to see that the \( y_n(\cdot) \) are equicontinuous, hence w.1.o.g.

convergent to some \( y \). Then \( x(t) = x_0 + \int_0^t y(s)ds \) is a solution (it is easy to see that the \( x_n(\cdot) \) converge to \( x(\cdot) \) since \( y_n = x_n^1 \) at the \( t_n^1 \)).

Theorem 3 in [1, § 2.2] states (without proof) the same result if \( F \) is Lipschitz but \( F(x) \) is only closed so \( F \) is no longer locally bounded.
(which is the case in Theorem 2 and gives a uniform Lipschitz constant for the $x_n(\cdot)$) and a more careful construction of the $x_n$, controlling the slopes, is necessary).

(11) The set of solutions in the usc, convex-valued case.

Let $J = [0,a]$, $F : D = J \times B_\rho(x_0) \to 2^{\mathbb{R}^n} \setminus \emptyset$ usc, $F(t,x)$ compact convex and $F(D) = B_\rho(x_0)$. Let $r \in (0,R)$ and $a = \min(2^{-1}(R-r)/M)$, $J = [0,a]$. Then every (a.c.) solution of

$$x' \in F(t,x), \; x(0) = y \in B_\rho(x_0)$$

exists on $J$, and (5) has at least one solution, by Theorem 2.1. Let

$$S(y) = \{x \in C(J) : x \text{ is solution of } (6), \; P_t(y) = (x(t), x \in S(y))\}.$$

Like for singlevalued $F$ the maps $S : B_\rho(x_0) \to 2^{C(J)} \setminus \emptyset$ and $P_t : B_\rho(x_0) \to 2^{\mathbb{R}^n} \setminus \emptyset$ are usc. (Since $P_t = T_t \circ S$, with $T_t : C(J) \to \mathbb{R}^n$ continuous, it is enough to show $S$ is usc, and this is obvious by the proof to Theorem 2.1). Furthermore $S(y)$ and $P_t(y)$ are compact, but of course not convex in general. Since one wants to use $P_t$ for finding periodic solutions, via fixed points of $P_t$, this is a handicap, which can however be overcome by suitable approximations to $S$ and $P_t$, showing in particular that $S(y)$ and $P_t(y)$ are connected sets. The procedure of approximating $S$ follows the method which is well known to true specialists in singlevalued ODEs. We first have

**Proposition 2.4.** Let $X,Y$ be Banach spaces, $F : D \subseteq X \to 2^Y \setminus \emptyset$ usc and $F(x)$ closed convex ($\forall x \in D$). Then there exist $F_n : D \to 2^Y \setminus \emptyset$ with

$$F_n(D) = \overline{\text{conv}} F(D), \; F(x) \subseteq F_{n+1}(x) \subseteq F_n(x) \subseteq F_0(x) \quad (\forall n \geq 1, \forall x \in D)$$

and

$$F_n(x) = F(x) + B_\rho(0)$$

for $n \geq n_0(x)$. The $F_n$ can be chosen of the form

$$\Sigma \psi_\lambda(\cdot)C_\lambda$$

with closed convex $C_\lambda \subseteq Y$ and a (locally finite) partition of unity $(\psi_\lambda)_{\lambda \in \Lambda}$.
Proof. Let \( r > 0, r_n = (1/3)^n r \), \((W^n_\lambda)_{\lambda \in A_n}\) a locally finite cover of \( D \) and a refinement of \((B^n_r(x))_{x \in D}, (\psi^n_\lambda)_{\lambda \in A_n}\) a partition of unity w.r. to \((W^n_\lambda)_{\lambda \in A_n}\), \( C^n_\lambda = \text{conv} F(B^n_{2r_n}(x^n_\lambda)) \) with \( x^n_\lambda \) such that \( W^n_\lambda = B^n_{r_n}(x^n_\lambda) \), and \( F_n(x) = \sum_{\lambda \in A_n} \psi^n_\lambda(x) C^n_\lambda \).

Evidently, \( F_n(x) \) is convex and \( F_n(D) = \text{conv} F(D) \). To see \( F_{n+1}(x) = F_n(x) \)

1. \( \Lambda^n_\mu = \{ \lambda \in \Lambda_n : x \in B^n_r(x^n_\lambda) \} \) and notice that the choice of \( x^n_\lambda \) implies \( \psi^n_\lambda(x) = 0 \) if \( \lambda \notin \Lambda^n_\mu \) (remember \( \text{supp} \psi^n_\lambda \subset \overline{W^n_\lambda} \)); now \( \lambda \in \Lambda^n_\mu \) and \( \mu \in \Lambda^n_{n+1} \) means \( |x^n_\lambda - x^n_\mu| < r_n + r_{n+1} = 4/3 r_n \), hence \( B^n_{2r_{n+1}}(x^n_{n+1}) \subset B^n_{2r_n}(x^n_\lambda) \), and therefore \( C^n_{n+1} \subset C^n_\lambda \), hence

\[
F_n(x) = \sum_{\lambda \in \Lambda^n_\mu} \psi^n_\lambda(x) C^n_\lambda \subset C^n_{n+1} \quad \text{(since \( \Sigma \psi^n_\lambda(x) = 1 \)).}
\]

and therefore \( F_{n+1}(x) = \sum_{\mu \in \Lambda^n_{n+1}} \psi^{n+1}_\mu(x) C^{n+1}_\mu \subset F_n(x) \) since \( F_n(x) \) is convex.

The inclusion \( F(x) \subset F_n(x) \) follows similarly. Finally, since \( F \) is usc, we have \( F(y) = F(x) + B_\varepsilon(o) \) for all \( y \in D \) such that \( |y - x| \leq \varepsilon = \varepsilon(x,y) \);

for \( n \geq n_0 \) we have \( r_{n+1} < \delta \), hence for \( \lambda \in \Lambda^n_\mu \): \( B^n_{2r_n}(x^n_\lambda) \subset B^n_\delta(x) \), and therefore \( C^n_\lambda \subset \overline{F(x)} + B_\delta(o) \) (since the latter is closed convex); consequently \( F_n(x) = F(x) + B_\varepsilon(o) \) for \( n \geq n_0(x,y) \).

Remark: This is the correct version of Theorem 1 in § 1.13 of [1], where it is claimed that the \( F_n \) are usc, which is as ridiculous as the authors' talk of "\( \sigma \)-selectionable maps" in this situation.

Now we have

Theorem 2.4. Let \( J = [0,a] \), \( F : D_0 = D_0 \rightarrow B_R(x_0) \rightarrow \mathbb{R}^n \) usc, \( F(t,x) \)
compact convex \((V(t,x))\) and \( F(D) = B_\delta(o) \). Let \( r \in (0,R) \),
\( a = \min(a_0/(R-r)/\delta) \) and \( J = [0,a] \).

Then there exist compact convex sets \( \{ \mathcal{K}_m \}_{m=1}^{k_m} \subset \mathbb{R}^n(J) \) and continues
\[
f_m : \overline{B}_r(x_0) \times U_m \to C(J) \text{ such that } S(y) = f_{m+1}(y, U_{m+1}) = f_m(y, U_m) \text{ on } \overline{B}_r(x_0) \text{ and } f_m(y, U_m) = S(y) + \beta_\varepsilon(o) \text{ for } m \geq m_0(y, \varepsilon). \text{ Moreover, } f_m(y, \cdot)^{-1}(x) \text{ is convex for } x \in S(y). \text{ In particular, } S(y) \text{ (and hence } P_t(y)) \text{ is connected.}
\]

Proof. By Proposition 2.4, with \( D = J \times \overline{B}_r(x_0) \) compact, we find \( F_m \) such that \( F_m(D) \subseteq \text{conv} F(D) \) (compact), \( (F_m(t, x)) \) is decreasing to \( F(t, x) \). Since "everything" is compact, the \( F_m \) are usc and it is enough to consider finite partitions of unity, hence \( F_m(t, x) = \sum_{i=1}^{k_m} \varphi_i^m(t, x)C_i^m \). Since \( F_m(t, x) \) is also compact convex, let \( S_m(y) = \{ x \in C(J) : x \text{ is solution of } x' \in F_m(t, x), x(o) = y \} \). Then \( S(y) = S_{m+1}(y) = S_m(y) \) is a trivial consequence of \( F(t, x) = F_{m+1}(t, x) = F_m(t, x) \). Furthermore \( S_m(y) = S(y) + \beta_\varepsilon(o) \) for \( m \geq m_0(\varepsilon, y) \) is also trivial (indirect proof).

Let \( U_m = k \sum_{i=1}^{k_m} C_i^m(J) \). This is a compact convex subset (metrizable) of \( k \sum_{i=1}^{k_m} L^\infty(J) \) with the \( w^* \)-topology. For \( u \in U_m \) the IVP \( x' = \sum_{i=1}^{k_m} \varphi_i^m(t, x)u_i(t), x(o) = y \) has a unique a.c. solution \( f_m(y, u)(\text{since we may choose the } \varphi_i^m \text{ locally Lipschitz}); \text{ it depends continuously on } y \text{ and } u, \text{ i.e. } f_m : \overline{B}_r(x_0) \times U_m \to C(J) \text{ is continuous.} \)

Evidently \( f_m(y, U_m) = S_m(y) \). But if \( x \in S_m(y), \text{ i.e. } x' = \sum_{i=1}^{k_m} \varphi_i^m(t, x)u_i(t), x(o) = y \), then (by Ex. 1.10) which obviously remains true if the compact \( U \subset \mathbb{R}^m \) is replaced by any compact metric space \( \subset \mathbb{U} \) there exists a measurable selection \( u : J \to U_m \) such that \( x' = \sum_{i=1}^{k_m} \varphi_i^m(t, x)u_i(t) \) a.e., i.e. \( x = f_m(y, u) \) and therefore \( S_m(y) = f_m(y, U_m) \). Finally, \( f_m(y, \cdot)^{-1}(x) \) is evidently convex for \( x \in S(y) \).

Since \( U_m \) (as a convex set) is connected and \( f_m(y, \cdot) \) is continuous, we have \( S_m(y) = f_m(y, U_m) \) connected, hence \( S(y) = \bigcup_{m \geq 1} S_m(y) \) connected, and therefore \( P_t(y) = T_t(S(y)) \) connected too (notice: \( S(y) \) is not connected \( \Leftrightarrow \) (since \( S(y) \) is compact) \( S(y) = M_1 \cup M_2 \) with \( M_1 \) compact and \( M_1 \cap M_2 = \emptyset \), hence the distance between \( M_1 \) and \( M_2 \) is positive and therefore \( S_m(y) \) "has also a gap" for \( m \) sufficiently large, which is a contradiction to
the connectedness of $S_m(y)$. //

**Corollary 2.1.** Let the conditions of Theorem 2.4 about $F$ be satisfied and assume $P_\tau : C \rightarrow 2^C \setminus \emptyset$ for some $\tau > 0$ and some compact convex subset of $B_\tau(x_0)$. Then $P_\tau$ has a fixed point $y \in P_\tau(y)$.

**Proof.** Since $C$ is compact we even have $S_m(y) = S(y) + \overline{B}_ \varepsilon (0)$ for $m \geq m_0(\varepsilon)$ (i.e. uniformly on $C$); use again a trivial indirect proof. Let $P_m = T_\varepsilon \circ S_m$, choose $u_m \in U_m$; then $g_m = T_\varepsilon f_m(\cdot, u_m)$ is a continuous selection of $P_m$. By Corollary 1.1 there exists $y_m \in C$ such that $|g_m(y_m) - y_m| = \rho(g_m(y_m), C)$, and since $C$ is compact we have w.l.o.g. $y_m \rightarrow y_0 \in C$. Hence

$S_m(y_m) = S(y_m) + \overline{B}_ \varepsilon (0) = S(y_0) + \overline{B}_ \varepsilon (0)$ for $m \geq m_0(\varepsilon)$ implies $g_m(y_m) \in T_\varepsilon (S(y_0) + \overline{B}_ \varepsilon (0))$ for $m \geq m_0$, and therefore $\rho(g_m(y_m), C) \rightarrow 0$ (since $P_\tau(C) \subset C$), i.e. $g_m(y_m) \rightarrow y_0$, and therefore $y_0 \in T_\tau S(y_0) = P_\tau(y_0)$. //

**Remarks.** Theorem 2.4 is Theorem 3 in § 2.2 of [1]; Corollary 2 is as wrong as the proof to Corollary 3 in § 2.2 of [1]. In Theorem 2 of § 2.2 these authors prove that $P_\tau(y)$ is connected, by a different proof; this proof is a trivial word by word imitation of a proof given in [5] for singlevalued operator equations, including of course the IVP for singlevalued ODEs; the authors are cheating since they do not mention this reference, and of course it is a waste of paper since the result follows as above anyway. Clearly, Corollary 2.1 can be used in the study of periodic solutions to $x' \in F(t,x)$ if $F$ is $\tau$-periodic in $t$, but we shall not go into details here. Instead we extend a well-known fact for singlevalued ODEs: through every point of $\partial P_\tau(y)$ there exists a solution
such that \( x(t) \in \mathcal{A}P_t(x) \forall t \leq \tau \); in honor of M. Hukuhara this is called the "Hukuhara-property" by some authors. This is again interesting for control theory (see Ex. 6.4): move \( x_0 \) into the target \( x_1 \) in minimal time \( \tau \) such that \( x_1 \in \mathcal{A}P_\tau(x_0) \); under weak extra conditions, guaranteeing that \( \tau \) exists, \( x_1 \) belongs to \( \mathcal{A}P_\tau(x_0) \) and the fact that there is an \( x \in S(x_0) \) such that \( x(t) \in \mathcal{A}P_t(x_0) \) \( \forall t \leq \tau \) gives sometimes necessary conditions for the existence of such optimal solutions.

**Theorem 2.5.** Under the conditions of Theorem 2.4 to \( x_1 \in \mathcal{A}P_\tau(x_0) \) there exists a solution \( x \) of (1) such that \( x(t) \in \mathcal{A}P_t(x_0) \) for 

\[ 0 \leq t \leq \tau \in \mathbb{J}. \]

**Proof.** (a) First of all, it is clear that if \( x(t_0) \in \text{int}(P_{t_0}(x_0)) \) for some \( t_0 < \tau \) then \( B_\delta(x(t)) \subseteq P_t(x_0) \) for \( t \in [t_0, t_0+\delta] \) and some \( \delta > 0 \). \( x \in S(x_0) \) [Otherwise there exist \( t_n \rightarrow t_0 \) and \( y_n \rightarrow x(t_0) \) such that \( y_n \notin P_{t_n}(x_0) \); then a solution \( z_n \) through \( (t_n, y_n) \) (to the left and right) cannot satisfy \( z_n(t_0) \in \mathcal{A}P_{t_0}(x_0) \), since piecing \( z_n(t), t \geq t_0 \) together with a solution from \( 0 \) to \( t_0 \) would be a solution in \([0, t_n]\). Hence \[ |x(t_0) - z_n(t_0)| \leq |x(t_0) - z_n(t_n)| + |z_n(t_n) - z_n(t_0)| \leq |x(t_0) - y_n| + M(t_n - t_0) \] and \[ |x(t_0) - z_n(t_0)| \geq \varepsilon \](n such that \( x(t_0) + B_\varepsilon(0) \subset P_{t_0}(x_0) \)) is a contradiction for large \( n \).

(b) Secondly, it is enough to show: Given \( t_0 < t_1 \) with \( t_0 \) and \( t_1 \in [0, \tau] \) and \( t_1 - t_0 \) sufficiently and \( x_1 \in \mathcal{A}P_{t_1}(x_0) \), there is a solution \( x \) such that \( x(t_1) = x_1 \) and \( x(t_0) \in \mathcal{A}P_{t_0}(x_0) \).

Indeed, consider \( t_1^n = \frac{1}{n^2} (1 = \ldots, n) \) and suppose \( x_n \in S(x_0) \) is such that \( x_n(t_1^n) \in \mathcal{A}P_{t_1^n}(x_0) \) for \( i = 1, \ldots, n \) and all sufficiently large \( n \); then, since \( S(x_0) \) is compact, w.l.o.g. \( x_n \rightarrow x \in S(x_0) \) and clearly \( x(t) \in \mathcal{A}P_t(x_0) \) for all \( t \in [0, \tau] \), by step (a).
(c) Now, if $t_1 - t_0$ is sufficiently small, all solutions (to the left) through $(t_1, x_1)$ exist in $[t_0, t_1]$ and belong to $B_R(x_0)$; let $S$ be this subset of $C(J)$ and $A$ its section at $t_0$ (which is compact and connected). We have to show $A \cap \exists P_{t_0}(x_0) = \emptyset$. If $(R^n \sim P_{t_0}(x_0)) \cup \text{int } (P_{t_0}(x_0)) \neq A$ then we are done. If this union covers $A$ and $A \cap \exists P_{t_0}(x_0)$, then $A$ is contained in one of these sets (since $A$ is connected) namely

\[ A = \text{int}(P_{t_0}(x_0)), \ \text{since } x_1 \in \exists P_{t_1}(x_0). \]

Since $A$ is compact, we have

\[ d(A, R^n \sim P_{t_0}(x_0)) = \alpha > 0. \]

Now, the solution (to the left)-multi is used on a neighborhood $U$ of $x_1$; hence there is $\varepsilon > 0$ such that solutions on $[t_0, t_1]$ starting in $B_\varepsilon(x_1)$ remain in the $\alpha$-neighborhood of the set of solutions through $(t_1, x_1)$. Hence to $y \in B_\varepsilon(x_1) \sim P_{t_1}(x_0)$ there exists $x_2 \in P_{t_0}(x_0)$ which is $z(t_0)$ for a solution on $[t_0, t_1]$ of $z \in F(t, z)$, $z(t_1) = y$. Hence, piecing $z$ together with a solution $x \in S(x_0)$ through $(t_0, x_2)$ we have a solution on $[0, t_1] \times (t_1, y)$, i.e. $y \in P_{t_1}(x_0)$, a contradiction.

References


§ 3 Solutions under constraints.

We consider the problem of finding a.c. solutions to $x' \in F(t, x)$ which satisfy constraints $x(t) \in K(t)$ on $J$, where the sets $K(t)$ are given. We again consider only the finitedimensional case.

3.1 Boundary conditions.

In the sequel $K \neq \emptyset$ will always be a closed subset of $R^n$.

If the IVP $x' = f(t, x)$, $x(t_0) = x_0 \in K$ has a solution $x(t)$ with $x(t) \in K$ on $[t_0, t_0 + \delta)$ then

$$\lim_{h \to 0^+} h^{-1}(x_0 + hf(t_0, x_0)) \in K, \quad h = o(h)$$

i.e. $\lim_{h \to 0^+} h^{-1}(x_0 + hf(t_0, x_0), K) = o$ is a necessary condition for this. If $f$ is continuous then it is not very hard to prove (by means of appropriate Euler-Cauchy-polygons) that the condition

$$\lim_{h \to 0^+} h^{-1}(x + hf(t, x), K) = o \quad (\forall t \in J, x \in K)$$

implies the existence of a solution with values in $K$, as was shown by M. Nagumo in 1942. For fixed $(t, x)$ this is a condition on the vector $v = f(t, x)$ which can also be expressed more geometrically, e.g. by introducing

$$T_K(x) = \cap_{\epsilon > 0} \cap_{\alpha > 0} \cup_{\delta > 0} \{ h^{-1}(K - x) + \delta \epsilon (0) \},$$

since obviously

$$\forall v \in T_K(x) \Rightarrow \lim_{h \to 0^+} h^{-1} p_K(x + hv) = o \quad (p_K(y) = p(y, K)).$$

Some people call $T_K(x)$ the "contingent cone" to $K$ at $x$, which is not justified since for general $K$ this is far from what one calls a cone in (Nonlinear) Functional Analysis, namely a closed convex $C$ such that $\lambda C \subseteq C$ for all $\lambda \geq 0$ and $C \cap (-C) = \{0\}$; if the last condition is missing (i.e. if $C$ is allowed to contain nontrivial subspaces) we call $C$ a cone.
So, by (3), condition (1) is equivalent to \( f(t,x) \in T_K(x) \) and in this form it can be extended to multivalued \( F \), by asking for \( F(t,x) \cap T_K(x) \neq \emptyset \) or the stronger version \( F(t,x) \subseteq T_K(x) \). Some simple consequences of the definitions are contained in

**Proposition 3.1.** Let \( K \subseteq \mathbb{R}^n \) be closed. Then we have

(1) \( x \in T_K(x) \) for \( x \in K \), \( T_K(x) = \emptyset \) for \( x \notin K \), \( T_K(x) = \mathbb{R}^n \) for \( x \in \mathbb{R} = \text{int}(K) \).

(11) \( T_K(x) \) is closed and \( \lambda T_K(x) = T_K(x) \) for all \( \lambda \geq 0 \).

(iii) If \( K \) is also convex then \( T_K(x) \) is a kone (for \( x \in K \)) and with \( N_K(x) = \{ \mathbf{x}^* \in \mathbb{R}^n : \mathbf{x}^*(x) = \sup_{y \in K} \mathbf{x}^*(y) \} \) ("the cone of normals to \( K \) at \( x \); \( x = \mathbb{R}^n \)) we have: \( \mathbf{v} \in T_K(x) \) iff \( \mathbf{x}^*(\mathbf{v}) \leq 0 \) (\( \forall \mathbf{x}^* \in N_K(x) \)).

(iv) If \( K \) is a kone then \( \mathbf{v} \in T_K(x) \) iff \( \mathbf{x}^*(\mathbf{v}) \geq 0 \) for all \( \mathbf{x}^* \in K^* = \{ \mathbf{x}^* \in \mathbb{R}^n : \mathbf{x}^*(y) \geq 0 \ \text{on} \ K \} \) satisfying \( \mathbf{x}^*(x) = 0 \).

(v) If \( K \) is also convex then \( \lim_{h \to 0^+} h^{-1} \rho_K(x + h\mathbf{v}) = 0 \) is equivalent to \( \lim_{h \to 0^+} h^{-1} \rho_K(x + h\mathbf{v}) = 0 \), and it is equivalent to \( \mathbf{v} \in -x + \text{Ext}(K(\lambda)) \), where \( \text{Ext}(K) = \{(1-\lambda)x + \lambda y : \lambda \geq 0, y \in K \} \) is the "inward set" of \( x \in K \" (a geometric interpretation which is used in fixed point theory; see e.g. § 18 in [2]).

(vi) If \( K = B_r(x_0) \) and \( \| \cdot \| \) is the Euclidean-norm then \( \mathbf{v} \in T_K(x) \) (for \( \| x - x_0 \| = r \)) iff \( (x-x_0, \mathbf{v}) \leq 0 \).

**Proof.** (1) is obvious; (11) follows from (3) [remember \( \rho_K(y) - \rho_K(\tilde{y}) \) for closedness and consider \( \lambda^{-1} h \) instead of \( h \) to see \( \lambda T_K(x) = T_K(x) \)].

(iv) is a consequence of (iii); notice that \( \mathbf{x}^*(\mathbf{v}) \leq 0 \) for all \( \mathbf{x}^* \in N_K(x) \) is equivalent to "\( \mathbf{x}^*(\mathbf{v}) \geq 0 \) for all \( \mathbf{x}^* \) such that \( \mathbf{x}^*(x) = \inf_{y \in K} \mathbf{x}^*(y) \)" and since \( \lambda K \subseteq K \) the latter means \( \inf_{y \in K} \mathbf{x}^*(y) = 0 \), hence \( \mathbf{x}^* \in K^* \) and \( \mathbf{x}^*(x) = 0 \).
(vi) follows from (iii) too: Since $(\mathbb{R}^n)^* = \mathbb{R}^n$, $x^*(x) = \sum_{i=1}^n x_i^* x_i$; therefore $x^*(x) = \sup_{y \in \mathbb{R}^n} x^*(y) = x^*(x)^* + r \sup_{y \in \mathbb{B}(x_0)} |y| \leq 1$
means\( \left| x^* - x^*_0 \right| \leq |x - x_0| |x^*| \), and therefore $x^* = x - x_0$, hence $(x - x_0, v) = x^*(v)$ ≤ 0.

Now we prove (iii) and (v): By (i), $T_K(x)$ will be a cone if we can show its convexity, but this follows from the second part: $x^*(\lambda v + (1-\lambda)\tilde{v}) = \lambda x^*(v) + (1-\lambda)x^*(\tilde{v}) \leq 0$ if $v, \tilde{v} \in T_K(x), \lambda \in [0,1]$ and $x^* \in N_K(x)$. We have $T_K(x) = -x + I_K(x)$ for $x \in K: v \in T_K(x) \Leftrightarrow v = y_n + z_n$ with $y_n \to 0$, $y_n \in K$, $|z_n| \leq \epsilon_n h_n$, $\epsilon_n \to 0 \Rightarrow v = x + y_n + h_n z_n$ with $h_n \to 0 \Rightarrow v \in -x + I_K(x)$.

We only have to show

\[ \lim_{h \to 0^+} h^{-1} p_K(x + hv) = 0 \iff v \in -x + I_K(x) \iff x^*(v) \leq 0 \quad (\forall x^* \in N_K(x)) \ . \]

This is very easy (and done in § 18 of [2]): We find $h \in (0,1)$ and $y \in K: |x + hv| - p_K(x+uv) + \epsilon$, hence $|v + hv - ((1-h^{-1})x + h^{-1}y)| \leq h^{-1} p_K(x + hv) + \epsilon$, and therefore $v + x \in I_K(x)$ is obvious. Choose $z \in I_K(x)$ such that $(v + x) - z \leq h^{-1} p_K(x + hv) + \epsilon$ since $K$ is convex there exists $\mu_0 > 0: x + u(z - x) \in K$

This is trivial; Suppose $v \in M = -x + I_K(x)$; since $M$ is closed convex,

there exists $x^*: \sup_{z \in M} x^*(z) < x^*(v)$ (as follows from the separation theorem for convex sets), hence $x^*(v) > \lambda x^*(y - x)$ for all $\lambda \geq 0$, and $y \in K$,

therefore $x^*(v) \leq 0$, a contradiction.

Example 3.1. If $K$ is not convex we may have $\lim_{h \to 0^+} h^{-1} p_K(x + hv) = 0$ but $\lim_{h \to 0^+} h^{-1} p_K(x + hv) > 0$ at many points. Consider, for example, the union $K = \mathbb{R}^2$ of $(0,0)$ and the closed discs of radius $\alpha^n$ and center $(o, o^n \sqrt{2})$, }
where $a = (\sqrt{2}-1)/(\sqrt{2}+1)$. It is easy to see (draw picture) that $T_K(o, o) = \{v : |v_1| \leq v_2\}$ and for $v \in T_K(o, o) \setminus \{(0, 0)\}$ \(\lim\) is different from \(\mathrm{Tim}\). Notice also that for the points $x_n = (o, o^n(\sqrt{2} + 1))$ we have $T_K(x_n) = \{v : v_2 = o\}$; hence $T_K(\cdot)$ is neither usc nor lsc at $x = o$.

**Proposition 3.2.** Let $K \subset \mathbb{R}^n$ be closed convex. Then $N_K(\cdot)$ has closed graph and $T_K(\cdot)$ is lsc on $K$.

**Proof.** Closedness of graph $N_K(\cdot)$ is trivial. If $V$ is open and $v \in T_K(x_0) \cap V$ then $B_\varepsilon(v) \subseteq V$ for some $\varepsilon > 0$; hence $x_n \to x_0$ and $T_K(x_n) \cap V = \emptyset$ implies $|\lambda(y-x_n) - v| \geq \varepsilon$ for all $n$, $\lambda \geq 0$, $y \in K$, and therefore $|\lambda(y-x_0) - v| \geq \varepsilon$, i.e. $v \notin T_K(x_0)$, a contradiction. Thus, $T_K(\cdot)$ is lsc.

The situation considered so far can be generalized as follows: Suppose we have a preorder $\leq$ on $K$, i.e. $x \leq x$ for all $x \in K$ (reflexivity) and $x \leq y, y \leq z \Rightarrow x \leq z$ (transitivity). This can be described by the multi

$Q : K \to 2^K \setminus \emptyset$, defined by $Q(x) = \{y \in K : y \leq x\}$. And conversely: If $Q : K \to 2^K \setminus \emptyset$ is given such that $x \in Q(x)$ on $K$ and $Q(y) = Q(x)$ whenever $x \in K$ and $y \in Q(x)$, then $y \leq x \iff y \in Q(x)$ defines a preorder $\leq$ on $K$.

Given such a $Q$ we may then look for monotone solutions, i.e. $x : J \to K$ such that $s \leq t$ implies $x(t) \in Q(x(s))$. Notice that $Q(x) = K$ is the special case of solutions in $K$.

**Example 3.2.** Consider $\psi_i : K \to \mathbb{R}$ for $i = 1, \ldots, m$ and define $Q(x) = \{y \in K : \psi_i(y) \leq \psi_i(x) \text{ for } i = 1, \ldots, m\}$. To have a monotone solution $x$ then means that all $\psi_i(x(\cdot))$ are decreasing. This is interesting for some economical models, where the $\psi_i$ are “utility functions”, and for stability theory.
Since the case \(Q \neq K\) is not more difficult (under appropriate conditions on \(Q\)) than the case \(Q = K\), we formulate existence theorems for monotone solutions.

3.2 Existence of monotone solutions.

**Theorem 3.1.** Let \(K \subset \mathbb{R}^n\) be closed, \(F : K \to 2^{\mathbb{R}^n}\) usc and F(x) compact convex (\(\forall x \in K\)). Let \(Q : K \to 2^K\) be lsc, graph (Q) closed, \(x \in Q(x)\) (\(\forall x \in K\)) and \(Q(y) \subset Q(x)\) for all \(x \in K\) and \(y \in Q(x)\).

Then for every \(x_0 \in K\) the IVP \(x(t) \in F(x)\), \(x(0) = x_0\) has a solution, satisfying \(x(t) \in Q(x(s))\) whenever \(s \leq t\), i.e. \(F(x) \cap T_Q(x)(x) \neq \emptyset\) on \(K\).

**Proof.** 1. The condition is necessary: Let \(x\) be a solution through
\[x_0 \in K; \text{then } x(t) \in Q(x_0), \text{ hence } F(x_0) \cap T_Q(x)(x_0) = \emptyset\.

Since F is usc, we have \(F(x(t)) \subset F(x_0) \cap T_Q(x)(x_0)\) for \(t \leq t_0\), hence \(F(x_0) \cap T_Q(x)(x_0)\) is closed convex, hence \(\int_0^{t_0} x'(s)ds \in F(x_0) \cap T_Q(x)(x_0)\).

2. Now we prove sufficiency, by means of properly chosen polygons.

Since F is usc, we have \(F(D) \subset B_c(\sigma)\) for \(D = K \cap B_r(x_0)\) with some \(r\) and \(c > 0\).

For each \(y \in D\), given \(k \in \mathbb{N}\), we find \(h_y < 1/k\) and \(v_y \in F(y): F(y) = \rho_Q(y)(y + h_y v_y) < h_y/(2k)\). Since Q is lsc, it is easy to see that \(\rho_Q(x)(x + h_y v_y) < h_y/(2k)\) is usc, hence the sets \(N(y) = \{x \in K : F(y) = \rho_Q(x)(x + h_y v_y) < h_y/(2k)\}\) are open in \(K\) and therefore \(K \cap B_{n_y}(y) \subset N(y)\) for some \(n_y < 1/k\). Since \(D\) is compact, we have
let \( h_i = h_{y_1}, u_i = u_{y_1}, h_0(k) = \min(h_j = j = 1, \ldots, n) \).

Thus \( x \in \mathbb{D} \) belongs to some \( h_0(y_1) \mathbb{n} \mathbb{k} = \mathbb{n}(y_1) \), which means
\[
|v_i - (x_i - x)h_i^{-1}| \leq 1/k \text{ for some } x_i \in Q(x), \text{ hence, with } u_i = (x_i - x)h_i^{-1},
\]

omitting the index \( i \) we have: To \( x \in \mathbb{D} \) there exists \( h \in [\mathbb{n}(k)\mathbb{1}/k] \) and \( u \in \mathbb{R}^n \) such that
\[
(a) \quad x + hu \in Q(x), \quad u \in C = F(D) + \mathbb{B}_1(0).
(b) \quad |x - y| \leq 1/k \text{ and } |v - u| \leq 1/k \text{ for some } y \in \mathbb{D} \text{ and some } v \in F(y).
\]

3. Let \( r = r/(c+1) \). To \( x_0 \) we find \( h_0 \) and \( u_0 \) (according to (a), (b)) such that \( x_1 := x_0 + h_0u_0 \in Q(x_0) \) and \( u_0 \in F(D) \cap \mathbb{B}_1/k(x_0) + \mathbb{B}_1/k(0) \); therefore
\[
|x_1 - x_0| \leq h_0(c+1) \leq r, \text{ i.e. } x_1 \in \mathbb{D}, \text{ if } 1/k \leq r. \text{ Now, we do}
\]
the same for \( x_1 \) as what we did for \( x_0 \), and so on; after finitely many steps we have \( x_i \) \( i = 1, \ldots, n \) for \( i \leq 1 \) and define the
polygons \( x_k(t) \) by
\[
x_k(t) = x_{i-1} + (t - \tau_{i-1})u_{i-1} \text{ on } (\tau_{i-1}, \tau_i) \text{ for } 1 \geq i, \quad x_k(0) = x_0.
\]

For \( t \in (\tau_{i-1}, \tau_i) \) we have \( x_k(t) = u_{i-1} \in F(D) \cap \mathbb{B}_1/k(x_{i-1}) + \mathbb{B}_1/k(0) \) and
\[
|x_k(t) - x_{i-1}| \leq 1/k(c+1), \text{ hence}
\]
\[
x_k(t) \in F(D) \cap \mathbb{B}_1/k(x_k(t)) + \mathbb{B}_1/k(0) \text{ a.e. (for } a = c+2). \]

Therefore, as in the proof to Theorem 2.1: w.l.o.g. \( x_k \to x \) in \( C(J) \),
\( x_k \to x' \) in \( L^1(J) \) and consequently \( x' \in F(x) \) a.e. Now let \( s < t \). Then
for \( k \) large enough we have \( \tau_{kp} \) close to \( s \) and \( \tau_{kq} \) (with \( q > p \)) close
to \( t \). Since we have \( x_k(\tau_{kp}) = x_j \in Q(x_{j-1}) = Q(x_k(\tau_{k-1,j-1})) \), the transitivity of \( Q \) yields \( x_k(\tau_{kp}) \in Q(x_k(\tau_{kp})) \); and therefore the closedness of \( \mathrm{graph}(Q) \) implies in the limit \((x(s), x(t)) \in \mathrm{graph}(Q), \text{ i.e.}
\]
\[
x(t) \in Q(x(s)), \quad /
\]

Remarks. 1. Theorem 3.1 contains Proposition 1, Theorem 1, Proposition 2 and Theorem 3 in Chapt. 4.2 of [1]; Theorem 2 of this Chapter 4.2 tells
you that \( x' = f(x) \), \( x(o) = x_0 \) in a Hilbert space has a solution if \( f \)

is continuous; this is ridiculous since it is well known that Peano's theorem is wrong in every infinite dimensional Banach space; see e.g.

p. 27 in [6].

2. Chap. 4 and Chap. 5 of [1] contain a detailed discussion of the case \( K = L \cap A^{-1}(M) \), where \( L \subseteq \mathbb{R}^n \) and \( M \subseteq \mathbb{R}^m \) are closed and \( A \) is a linear map from \( \mathbb{R}^n \) into \( \mathbb{R}^m \). These results are easy exercises.

3. Theorem 1 in Chap. 4.6 of [1] says that \( x' \in F(x) \), \( x(o) = x_0 \) has a solution in a closed \( K \) if the \( F(x) \) are only compact, but \( F \) is continuous and satisfies also the strong condition \( F(x) \subseteq T_K(x) \) on \( K \). The proof is also given by means of appropriate polygons.

3.3 The timedependent case.

We now want to find solutions of

\[
\begin{align*}
(4) & \quad x' \in F(t,x) \text{ a.e., } x(o) = x_0, \, x(t) \in K(t) \text{ on } J = [o,a].
\end{align*}
\]

For this purpose it is convenient to introduce a "derivative" of a multi, as a natural generalization of the singlevalued case. Remember, for example, that \( (1,0,f_x) \) and \( (0,1,f_y) \) span the "tangent plane" at \( (x_0, f(x_0)) \) if \( f : \mathbb{R}^2 \rightarrow \mathbb{R} \) is differentiable at \( x_0 \), i.e. \( (x_0, f(x_0)) + (u, f_x \cdot u_1 + f_y u_2) \) is the general vector of the tangent plane and the tangent space (the translate through zero) is \( \{(u, f'(x_0)u) : u \in \mathbb{R}^2\} \), the graph of \( f'(x_0) \). Now the substitute for the tangent space is the "contingent cone" and therefore, given \( F : K \subseteq \mathbb{R}^n \rightarrow 2^{\mathbb{R}^m} \setminus \emptyset \), we define the ("contingent") derivative \( DF(x_0, y_0) \) of \( F \) at \( x_0 \) and \( y_0 \in F(x_0) \) as the multi \( DF(x_0, y_0) : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^m} \) whose graph is \( T_{\text{graph}(F)}(x_0, y_0) \), i.e.
(5) \[ DF(x_o, y_o)(u) = \{ v \in \mathbb{R}^n : (u, v) \in T_{\text{graph}(F)}(x_o, y_o) \} \] 

If \( DF(x_o, y_o)(u) \neq \emptyset \) then, in particular, \( \lim_{h \to 0^+} h^{-1} p_K(x_o + hu) = 0 \), i.e. \( u \in T_K(x_o) \) in other words: \( \text{Dom } DF(x_o, y_o) \subseteq T_K(x_o) \). Furthermore, we have for \((x_o, y_o) \in \text{graph } F\)

(6) \[ v \in DF(x_o, y_o)(u) \iff \lim_{h \to 0^+} \frac{F(x_o + h u, y_o) - y_o}{h} \to 0 \text{ for some } h \to 0^+, u \to u. \]

Indeed, \((u, v) \in T_{\text{graph}(F)}(x_o, y_o)\) means \( h_n^{-1} |x_o + h_n u - x| \to 0 \) and \( h_n^{-1} |y_o + h_n v - y| \to 0 \) for some \( h_n \to 0^+, x_n \in K \) and \( y_n \in F(x_n) \); this yields the right hand side with \( u_n = h_n^{-1}(x_n - x_o) \), and vice versa. Clearly, in the singlevalued case we get the old concepts; for example, if \( f \) has a derivative at \( x_c \) in direction \( u \), i.e. \( f'(x_o) = \lim_{t \to 0^+} \frac{f(x_o + tu) - f(x_o)}{t} \) exists, then \( DF(x_o, f(x_o))(u) = \{ f'(x_o) \} \).

Now, we have

**Theorem 3.2.** Let \( K : J = [0, a] \to \mathbb{R}^n \) have closed graph, \( F : \text{graph } K \to \mathbb{R}^n \) usc and \( F(t, x) \) compact convex (for all \((t, x) \in \text{graph } K \)). Then (4) has a (local) solution if

(7) \( F(t, x) \cap \text{DK}(t, x)(\{t\}) \neq \emptyset \quad (\forall (t, x) \in \text{graph } K) \).

**Proof.** \( K_0 = \text{graph } K \) is closed in \( \mathbb{R}^{n+1} \); define \( F_0 : K_0 \to \mathbb{R}^{n+1} \) by

\[ F_0(x_{n+1}, x) = (1, F(x_{n+1}, x)) \] for \((x_{n+1}, x) \in K_0 \). Evidently \( F_0 \) is usc and \( F(x_{n+1}, x) \) is compact convex. Hence, by Theorem 3.1, the autonomous problem \((x_{n+1}, x) \in F_0(x_{n+1}, x), (x_{n+1}(0), x(0)) = (0, x_0) \) has a solution in \( K_0 \) if \( F_0(x_{n+1}, x) \cap \text{DK}_0(x_{n+1}, x)(\{1\}) \neq \emptyset \) on \( K_0 \), i.e. \( F(x_{n+1}, x) \cap \text{DK}(x_{n+1}, x)(\{1\}) \neq \emptyset \) on \( K_0 \).

Clearly, condition (7) is also necessary if one wants to have a local solution through every \((t_0, x_0) \in \text{graph } K \) satisfying \( x(t) \in K(t) \) in \([t_0, t_0 + \delta(t_0, x_0)] \) (define \( K(t) = K(a) \) and \( F(t, x) = F(a, x) \) for \( t > a \)).
Example 3.3. The implicit problem

\[ f(t, x(t), x'(t)) \in F(t, x(t)) \text{ a.e., } x(0) = x_0 \in K(0), x(t) \in K(t) \text{ in } t \geq 0 \]

has a solution if \( K : \mathbb{R}^+ \to 2^{\mathbb{R}^n} \setminus \emptyset \) has closed graph; \( F : \text{graph } K \to 2^{\mathbb{R}^m} \) is usc and \( F(t, x) \) is closed convex \( (\forall (t, x) \in \text{graph } K) \); \( f : \text{graph}(K) \times \mathbb{R}^n \to \mathbb{R}^n \) is continuous and the \( f(t, x, \cdot) \) are affine: for some \( r > 0 \) and all \( (t, x) \in \text{graph } K \) there exists a solution \( v \in \text{DK}(t, x)(1) \cap \mathbb{B}_r(o) \) of \( f(t, x, v) \in F(t, x) \). Indeed, define \( G(t, x) = \{ v \in \mathbb{R}^n : f(t, x, v) \in F(t, x), |v| \leq r \} \). Evidently, \( G \) is usc, bounded on graph \( K \), \( G(t, x) \) is compact convex and \( G(t, x) \cap \text{DK}(t, x)(1) \setminus \emptyset \) on graph \( K \). Hence \( x' \in G(t, x) \), \( x(0) = x_0 \) has a global solution in graph \( K \).

Such problems appear in control theory: \( \mathbb{R}^n \) is the state space of the system, \( \mathbb{R}^m \) the space of observations, one has an observation map \( B : \mathbb{R}^n \to \mathbb{R}^m \) and a feedback map \( C : \mathbb{R}^m \to \mathbb{R}^n \) and then the model is e.g. of type

\[ \begin{aligned}
\text{(8)} \quad & x' \in F(x) + C(dB(x(t))/dt), \quad x(0) = x_0 \\
\end{aligned} \]

(i.e. the observations are used as controls). Suppose, for example, that \( K \subset \mathbb{R}^n \) is closed, \( F : K \to 2^{\mathbb{R}^n} \setminus \emptyset \) is usc with closed convex values, \( C \) is an \( n \times m \) matrix, \( B \) is \( C^1 \) in an open set containing \( K \), and there exists some \( r > 0 \) such that for every \( x \in K \) we find a solution \( v \in \text{DK}(x)(1) \cap \mathbb{B}_r(o) \) of \( v \in F(x) + CB'(x)v \). Then (8) has a solution in \( K \) if \( x_0 \in K \); notice that (8) is \( (1 - CB'(x))x' \in F(x), \quad x(0) = x_0 \).

Remarks. 1. It is trivial that Theorem 3.2 can be extended to the case where one has a preorder described by a reflexive and transitive multi set satisfying conditions like in Theorem 3.1.

2. It is also easy to see that for nonconvex-valued \( F \) the condition \( F(x) \cap \text{TK}(x) \neq \emptyset \) is too weak for solutions in \( K \); consider e.g.

\( K = \mathbb{B}_1(o) \subset \mathbb{R}^2 \), \( F(x) = \{ (-1, 0), (1, 0) \} \) which satisfies \( F(x) \cap \text{TK}(x) \neq \emptyset \)
since $T_k(x) = \{ y : (y,x) \leq 0 \}$ for $|x| = 1$, but $x' \in F(x)$, $x(0) = (0,1)$ or $(0,-1)$ has no solution in $K$.

3. In Chap. 4.7 of [1] the results are extended (under strong assumptions) to multivalued functional differential equations:

$$ x'(t) \in F(t,x(t)) \text{ a.e. in } t \in [0,\alpha], \quad T(x) = \varphi, $$
where $(T(x)) = x(t + \tau)$ for $\tau \in \mathbb{R}$ and $x \in C([-\alpha,\alpha])$.


### 3.4 Equilibria, fixed points, periodic solutions.

Like in the singlevalued case, a zero of $F$ is called an equilibrium for $x' \in F(x)$. Clearly, a theorem on existence of zeros of $F$ yields a fixed point for $G(x) = F(x) + x$, and vise versa. Contrary to the classical situation we shall not assume that $G$ maps $K$ into itself, but that it is only "weakly inward", i.e. $G(x) \cap T_k(x) \neq \emptyset$ with $T_k(x) = x + T_k(x)$; remember the proof of Proposition 3.1. To find a zero of $F$ means to find a constant solution of $x' \in F(x)$, and under certain conditions on $F$ this can be done by means of $\omega$-periodic solutions with $\omega \rightarrow 0$. Therefore we start with a result about $\omega$-periodic solutions.

**Theorem 3.3.** Let $K \subset \mathbb{R}^n$ be compact convex, $F : \mathbb{R}_+ \times K \rightarrow \mathbb{R}^n \setminus \emptyset$ be usc, $F(t)$ compact convex and $F(t + \omega, x) = F(t, x)$ on $\mathbb{R}_+ \times K$, for some $\omega > 0$. Suppose that $F(t, x) \cap T_k(x) \neq \emptyset$ on $\mathbb{R}_+ \times K$. Then $x' \in F(t, x)$ has an $\omega$-periodic solu

**Proof.** 1. Let $J = [0,\omega]$. We have to find a solution $x$ satisfying $x(0) = x$.

We have $F(t,x)$ compact and w.l.o.g. $K = \mathbb{R}^n$, since $K$ has nonempty interior $K$ = span $K$ and consider $F(t,x) = F(t,x) \cap \mathbb{R}^n$. The latter is not always true in infinite dimensional spaces. By Proposition 2.4 (see proof to Theorem 2.4), we find continuous multiv $F_n : J \times K \rightarrow \mathbb{R}^n$ by $F_n(t,x) = F(t,x)$.
the form $F_n(t,x) = \sum_{i=1}^{m} \psi_i(t,x)C_i$ such that $F(t,x) = F_{n+1}(t,x) = F_n(t,x)(\forall n)$. Let $G_n(t,x) = F_n(t,x) + \beta_{1/n}(o)$. Then $G_n$ is 1sc and $G_n(t,x) \cap H_K(x) \neq \emptyset$ on $K$. Hence, by Remark 4 in § 1.2, there is a continuous selection $g_n$ of $G_n \cap H_K$. 2. So we have continuous $g_n : J \times K \rightarrow R$ such that $g_n(t,x) \in T_K(x)$ (and $g_n(t,x) \in G_n(t,x)$). Let $P_K$ be the metric projection onto $K$ (this part is now a typical Hilbert space proof, to make it short) and $\bar{g}_n(t,x) = g_n(t,P_K(x))$. By Theorem 3.1 we know that $u' = g_n(t,u)$, $u(o) = x \in K$ has at least one solution on $J$ (in $K$ of course), since $g_n$ is bounded. This is also a solution of $u' = \bar{g}_n(t,u)$, $u(o) = x$, but $\bar{g}_n$ has the advantage that all solutions of the IVP remain in $K$, in other words: "$K$ is invariant for $u' = \bar{g}_n(t,u)$", as we shall show in the next section (see Theorem 3.5 (11)). Hence we can apply Corollary 2.1 to $P_u$ (the Poincaré operator for $\bar{g}_n$) and find $x_n \in P_u(x_n) \subset K$, i.e. a solution $u_n$ of $u' = g_n(t,u)$ such that $u_n(o) = u_n(o) = x_n$. Now w.l.o.g. $x_n \rightarrow x_0 \in K$ and $u_n \rightarrow u$, a solution of $u' \in F(t,u)$ a.e. satisfying $u(o) = u(o) = x_0$.  

As an immediate consequence of this result, we have

**Theorem 3.4.** Let $K \subset R^n$ be compact convex, $F : K \rightarrow R^n \times \emptyset$ usc and $F(x)$ compact convex ($\forall x \in K$). Then

1. $F(x) \cap H_K(x) \neq \emptyset$ on $K \Rightarrow F$ has a zero
2. $F(x) \cap H_K(x) \neq \emptyset$ on $K \Rightarrow F$ has a fixed point.

**Proof.** It is enough to prove (i) since (ii) means $(Fx-x) \cap H_K(x) \neq \emptyset$ on $K$.

By Theorem 3.3, for every $\omega > 0$ we find a $\omega$-periodic solution of $x' \in F(x)$; let $\omega_n = 1/2^n$ and $x_n(\cdot)$ an $\omega_n$-periodic solution on $R_+$. Hence
w.l.o.g. $x_n(0) \to x_0$ and $x_n(\cdot) \to x(\cdot)$ uniformly on $J = [0, 1]$, $x' \in F(x)$ a.e.,

$|x(t) - x(s)| \leq L|t - s|$ for some $L > 0$ and all $t, s \in J$, $|x(k/2^n) - x_0| \leq \alpha_n$,

$max_j |x_n(t) - x(t)| + |x_n(0) - x_0| = \alpha_n$ and therefore $|x(t) - x_0| \leq L/2^n + \alpha_n$ on $J$

for all $n$, i.e. $x(t) \sim x_0$.

**Remark.** Theorem 3.4 is also true if one replaces "$F(x)$ compact convex" by "$F(x)$ closed convex"; see e.g. Theorem 24.5 in [2] for every Banach space.

3.5 **Invariant sets.**

So far we have studied the question under which conditions on $K$ and $F$ there is at least one solution in $K$. If $F$ is defined on a larger set these conditions do not necessarily imply that all solutions starting in $K$ remain in $K$, i.e. $K$ need not be (forward) invariant for $x' \in F(t, x)$. In the single-valued case some conditions guaranteeing invariance are known, but not really satisfactory; they usually include some uniqueness assumptions for the IVP, see e.g. § 5 of [6], and the only idea of proof is: assume a solution $u$ leaves $K$ and consider e.g. $p_K(u(t))$ to get a contradiction. The same remarks apply to what one has done for $x' \in F(t, x)$, as we are going to indicate.

**Proposition 3.3(1)** Let $K \subset \mathbb{R}^n$ be closed. Then for $y, z \in \mathbb{R}^n$

$\lim_{h \to 0^+} h^{-1}[p_K(y+hz) - p_K(y)] \leq (z, T_K(w))$ for all $w \in P_K(y)$.

(11) Let $K \subset \mathbb{R}^n$ be compact convex. Then $[p_K(x)^2]' = 2(x - p_K(x))$.

**Proof.** (1) For $w \in P_K(y)$ and $v \in T_K(w)$ we have

$h^{-1}(p_K(y+hz) - p_K(y)) \leq h^{-1}(p_K(w+hz) + |w-y| - p_K(y)) = h^{-1}p_K(w+hz)$

$\leq |z - v| \cdot h^{-1}p_K(w+hv)$.

hence application of \( \lim \) and then \( \inf \) gives the result.

(11) Let \( \omega(x) = p_K(x) = |x-P_K(x)|^2 \) and \( \psi(t) = t^{-1}(\varphi(x+th)-\varphi(x)) \). Then
\[
\lim_{t \to 0^+} \sup_{t \to 0^+} t^{-1}(|x+th-P_K(x)|^2 - |x-P_K(x)|^2) = 2\langle x-P_K(x), h \rangle
\]

\[
\lim_{t \to 0^+} \psi(t) \geq \lim_{t \to 0^+} t^{-1}(|x+th-P_K(x+th)|^2 - |x-P_K(x+th)|^2) = 2\langle x-P_K(x), h \rangle
\]
hence \( \lim \psi(t) = 2\langle x-P_K(x), h \rangle \); here we used \( |x-y|^2 = |x|^2 + |y|^2 - 2\langle x,y \rangle \), the definition of \( P_K \), the continuity of \( P_K \) (in fact \( P_K \) satisfies \( |P_K(x)-P_K(y)| \leq |x-y| \); see e.g. Proposition 9.2 in [2]). Hence the directional derivative is continuous linear in \( h \) and continuous in \( x \), and therefore also \( \varphi'(x) = 2\langle x-P_K(x) \rangle \), i.e. \( |\varphi(x+h)-\varphi(x)-2\langle x-P_K(x), h \rangle = o(|h|) \) as \( |h| \to 0 \).

By means of these facts it is very easy to prove

"Little" Theorem 3.5 (i) Let \( K \subseteq \mathbb{R}^N \) be closed, \( \Omega \supseteq K \) open,
\( F : \Omega \to 2^{\mathbb{R}^N} \setminus \emptyset \) Lipschitz and \( F(x) \) compact for all \( x \in K \),
\( F(x) \subset T_K(x) \) on \( K \). Then \( K \) is invariant for \( x' \in F(x) \).

(ii) Let \( K \subseteq \mathbb{R}^N \) be closed convex, \( F : K \to 2^{\mathbb{R}^N} \setminus \emptyset \) usc and \( F(x) \) compact convex (\( \forall x \in K \)), \( F(K) \) compact. Suppose also that \( F(x) \subset T_K(x) \) on \( K \).
Then \( x' \in F(x) \), \( x(o) = x_0 \in K \) has a solution on \( \mathbb{R}_+ \) and \( K \) is invariant for \( x' \in F \circ P_K(x) \).

**Proof.** (1) Let \( x \) be a solution on \([0,a]\) with \( x(0) \in K \), and let
\( \omega(t) = p_K(x(t)) \). Since \( x \) and \( \omega \) are Lipschitz, consider \( t \) such that
\( x'(t) \) and \( \omega'(t) \) exist. Then
\[
\omega'(t) = \lim_{h \to 0} h^{-1} [p_K(x(t)+hx'(t)+o(h))-p_K(x(t))]
\]
\[
= \lim_{h \to 0} h^{-1} [p_K(x(t)+hx'(t))-p_K(x(t))] \leq o(x'(t), T_K(w))
\]
with \( w \in P_K(x(t)) \)
\[ \varphi(x'(t), F(w)) \leq d_{H}(F(x(t)), F(w)) \leq \| x(t) - w \| = \| w(t) \].

hence \( \varphi' \leq \| w \| \) a.e., \( \varphi(o) = 0 \), and therefore \( \varphi(t) \leq 0 \), since \( \varphi \) is a.c.

(ii) By Theorem 3.1 and the compactness of \( \mathbb{F}(K) \) it is clear that 
\( x' \in F(x) \), \( x(o) = x_0 \in K \) has a solution on \( \mathbb{R}_+ \); and the same is therefore true for \( x' \in F(P_K(x)) \), \( x(o) = x_0 \in K \). Now, suppose that \( x \) is a solution of the latter such that \( x(t_1) \notin K \) for some \( t_1 > 0 \), and let \( \varphi(t) = \rho_K(x(t)) \).

Then \( \varphi'(t) = 2(x(t) - P_K(x(t)), x'(t)) \) a.e. (= 0 for \( x(t) \notin K \)), and for \( t \) such that \( x(t) \notin K \) we have \( x'(t) \in F(P_K(x(t))) = T_K(P_K(x(t))) \); since \( (R^n)^* = R^n \) and \( (y-P_K(x), x-P_K(x)) \leq 0 \) for all \( y \in K \), we have \( \sup_{K} (y, x-P_K(x)) = (P_K(x), x-P_K(x)) \), i.e. \( x-P_K(x) \in H_K(P_K(x)) \), hence \( \varphi'(t) \leq 0 \) for all \( t \) such that \( x(t) \notin K \) by Proposition 3.1 (iii). Thus \( \varphi(t) \leq 0 \) a.e. for \( x(t) \notin K \), and therefore \( 0 < \varphi(t_1) = \int_0^{t_1} \varphi'(s) ds \leq 0 \), a contradiction.

3.6 Remarks

1. Control problems. Consider

\( x' = f(x, u(t)), x(o) = x_0 \in K \subset R^n \)

with \( f : K \times U \rightarrow R^n \) continuous, \( U \subset R^m \) compact. Is there a measurable \( u : J \rightarrow U \) such that (9) has a solution (in \( K \) of course)?, and does there exist a continuous \( g : K \rightarrow U \) (a "closed loop or feedback control") such that

\( x' = f(x, g(x)), x(o) = x_0 \in K \)

has a solution (in \( K \))? As a simple consequence of Theorem 3.1 we have the following answers.
Corollary 3.1. Let \( K \subseteq \mathbb{R}^n \) and \( U \subseteq \mathbb{R}^m \) be compact, \( f : K \times U \rightarrow \mathbb{R}^n \) continuous. Then

(i) If \( f(x,u) \) is convex (\( \forall x \in K \)) and \( C(x) = f(x,\cdot)^{-1}(T_K(x)) \neq \emptyset \) on \( K \) then (9) has a solution \((x,u)\) such that \( u(t) \in C(x(t)) \) a.e.

(ii) If \( K \) and \( U \) are also convex and \( f(x,\cdot) \) is affine (\( \forall x \in K \)) then the stronger condition \( f(x,\cdot)^{-1}(-y + T_K(x)) \neq \emptyset \) on \( K \times B_r(o) \) (for some \( r > 0 \)) implies the existence of a feedback control \( g \) such that (10) has a solution.

Proof. (i) is obvious: \( F(x) = f(x,u) \) is compact convex, \( F \) is continuous and \( F(x) \cap T_K(x) \neq \emptyset \) on \( K \). Hence \( x' \in F(x), x(0) = x_0 \) has a solution, and (see Example 1.10) there is a measurable \( u \) such that \( x'(t) = f(x(t),u(t)) \) a.e., and \( u(t) \in C(x(t)) \) a.e. since \( x'(t) \in T_K(x(t)) \) a.e.

(ii) The stronger assumption means \( B_r(o) \subseteq T_K(x) - f(x,u) \) for all \( x \in K \). We also know that \( T_K(\cdot) \) is lsc since \( K \) is convex. We show that \( C(\cdot) \) is lsc. Then \( C(\cdot) \) has a continuous selection \( g \) and \( x' = f(x,g(x)) \), \( x(0) = x_0 \in K \) has a solution since \( f(x,g(x)) \in T_K(x) \).

Let \( u \in C(x_0), x_n \to x_0 \), but suppose \( C(x_n) \cap B_r(o) = \emptyset \) for all \( n \). Since \( f \) is continuous and \( T_K(\cdot) \) is lsc we have \( f(x_n,u) \in T_K(x_n) + B_r(o) \) for \( |u-u_0| \leq \epsilon_n \to 0 \) and therefore \( \lambda_n f(x_n,u) \in \lambda_n T_K(x_n) + \lambda_n \rho_n (0) \subseteq \lambda_n T_K(x_n) + (1-\lambda_n)T_K(x_n) - (1-\lambda_n)f(x_n,u) \)

for \( \lambda_n > 0 \) such that \( \lambda_n \rho_n \leq (1-\lambda_n)^r \), i.e. \( \lambda_n \leq r/(r+\rho_n) \). Since \( T_K(x_n) \) is convex and \( f(x_n,\cdot) \) is affine this means

\[ f(x_n,u_n) \in T_K(x_n), \text{ i.e. } u_n \in C(x_n) \text{ for } u_n = \lambda_n u + (1-\lambda_n) v_n \]

for some \( v_n \in U \). Now, \( |u_n - u_0| \leq \lambda_n \epsilon_n + (1-\lambda_n) \text{diam}(U) \leq 0 \) for large \( n \), if we choose \( \lambda_n \leq r/(r+\rho_n) \) such that \( \lambda_n \to 1 \).

3. Variational differential inequalities. Let \( K \subset \mathbb{R}^n \) be closed convex and \( F : K \to 2^{\mathbb{R}^n} \setminus \emptyset \) be given. If \( F \) doesn’t satisfy \( F(x) \cap T_K(x) = \emptyset \) on \( K \), we may consider

\[
(11) \quad x' \in P_{T_K(x)}(F(x)), \quad x(0) = x_0 \in K
\]

instead of \( x' \in F(x), \quad x(0) = x_0 \). This is of course reasonable only if \( \mathbb{R} \neq \emptyset \) since then (11) coincides with the original problem on \( \mathbb{R} \) and we have only changed \( F \) at \( aK \). This kind of reasoning was introduced in [7] for economic planning procedures. Now, since \( (\mathbb{R}^n)^* = \mathbb{R}^n \) and \( T_K(x) \) is closed convex, (11) can also be formulated as follows

**Proposition 3.4.** Let \( K \subset \mathbb{R}^n \) be closed convex, \( F : K \to 2^{\mathbb{R}^n} \setminus \emptyset \) and

\[
N_K(x) = \{ p \in \mathbb{R}^n : (x, p) = \sup_{K} (y, p) \}.
\]

Then we have

\[
(1) \quad x(\cdot) \text{ is a solution of (11) iff } x(\cdot) \text{ is a solution of (12)}
\]

\[
(12) \quad x' \in F(x) - N_K(x), \quad x(0) = x_0.
\]

(11) \( x(\cdot) \) is a solution of (12) on \( J \) iff for almost all \( t \in J \) there exists \( f(x(t)) \in F(x(t)) \) such that

\[
(13) \quad \sup_{K} (x'(t) - f(x(t)), x(t) - y) \leq 0.
\]

**Proof.** (11) is trivial by definition of \( N_K(x) \).

(1) Recall that for a cone \( C \subset \mathbb{R}^n \) (or any Hilbert space) one has

\[
(14) \quad w \in P_C(v) \quad \Longleftrightarrow \quad w \in C, \quad (y, v-w) \leq 0 \quad (\forall y \in C), \quad (w, v-w) = 0.
\]

[Indeed: \( \Rightarrow \): Since \( C \) is convex, we have \( (y-w, v-w) \leq 0 \) on \( C \); since \( C \subset C \) for \( \lambda \geq 0 \), this implies \( (y, v-w) \leq 0 \) on \( C \), \( (w, v-w) \leq 0 \) and also \( (\lambda v, v-w) \leq 0 \).

\( \Leftarrow \): For \( y \in C \) we have

\[
|v-y|^2 = |v-w|^2 + 2(v-w, w-y) + |w-y|^2 \geq |v-w|^2 + |w-y|^2
\]

and therefore \( f(v, C) = |v-w| \).]
If \( x \) solves (11) then \( w = x'(t) = P_{\mathcal{T}_K(z)}(v) \) for \( z = x(t) \) and some \( v \in F(z) \), hence \((y,v-w) \leq 0 \) on \( T_K(z) = \{ \lambda(k-z) : \lambda \geq 0, k \in K \} \), in particular \((k-z,v-w) \leq 0 \) on \( K \), hence \( v-w \in N_K(z) \).

If \( x \) solves (12) then \( w = x'(t) \in v-N_K(z) \) for some \( v \in F(z) \), hence \( v-w \in N_K(z) \), i.e. \((y,v-w) \leq 0 \) on \( T_K(z) \), and therefore \( p(v,T_K(z))^2 \geq |v-w|^2 + 2(w,v-w) \). But \((w,v-w) = 0 \) [notice: \( w \in T_K(z) \) since the solution is in \( K \), and \( z \) follows from \( x(t-h) = z-hw+o(h) \in K \) for \( h > 0 \) small], hence \( w = P_{\mathcal{T}_K(z)}(v) \).

**Corollary 3.2.** If \( K \subset \mathbb{R}^n \) is compact convex, \( F : K \to \mathbb{R}^n \) is usc and \( F(x) \) is compact convex (\( \forall x \in K \)) then (11) has a solution.

**Proof.** \( F(K) \subset B_c(0) \) for some \( c > 0 \). By Proposition 3.2, \( N_K \) has closed graph, hence \( H(x) = F(x)-B_c(0) \cap N_K(x) \) defines an usc multi (with compact convex values) on \( K \); remember Proposition 1.1(b). By (14) we have

\[
(15) \quad P_{\mathcal{T}K(x)}(v) = v-P_{N_K(x)}(v) \quad \text{on} \quad K \times \mathbb{R}^n
\]

[Indeed: \( w = P_{\mathcal{T}_K(x)}(v) = v-w \in N_K(x) \), \((y,v-w) \leq 0 \) on \( N_K(x) \) by Proposition 3.1(11) and \((v-w,w) = 0 \), i.e. \( v-w = P_{N_K(x)}(v) \) by (14)]. Since \(|P_{N_K(x)}(v)| \leq |v| \leq c \) (notice \( 0 \in N_K(x) \)) this means \( w \in H(x) \) and \( w \in T_K(x) \).

Hence \( x' \in H(x) \subset F(x)-N_K(x) \), \( x(o) = x_0 \in K \) has a solution, by Theorem 3.1.

On pp. 271-273 of [1] this is "applied" to games. To speak of variational differential inequalities in connection with (11) is justified by (13); for example, \((x-x_0, \varphi'(x_0)) \geq 0 \) on \( C \) if the differentiable \( \varphi \) attains its minimum over the convex \( C \) at \( x_0 \in C \); see the remarks in § 26 of [2] for more general cases.
References


Appendix

Problems and suggestions for further work in this field.

1. Extend the existence theorems in § 2 to the case where \( \mathbb{R}^n \) is replaced by
an arbitrary Banach space \( X \), as far as possible. Try at the same time to
separate conditions w.r. to \( t \) and w.r. to \( x \) (see the remark in the intro-
duction to § 2). Some results of this type are known; see e.g. the references
in [1] and in the paper

differential equations. Atti Sem. Mat.Fis.Univ.Modena XXX II,

In this paper you will find the following result (Theorem 2):

**Theorem A 1.** Let \( X \) be separable, \( F : J \times X \to 2^X \setminus \emptyset \) such that \( F(t, \cdot) \) is
continuous, \( F(\cdot, x) \) is measurable, \( F(t, x) \) is compact and \( F(t, x) \subset K(t) \subset \emptyset \)
\( g(t) \mathbb{B}_1(0) \) on \( J \times X \) for some measurable \( K : J \to 2^X \setminus \emptyset \) with \( K(t) \) compact
convex and some \( g \in L^1(J) \). Then \( x' \in F(t, x) \), \( x(0) = 0 \) has an a.c. solution
on \( J \).

The proof given there is extremely long and should be simplified and the
result should be improved.

2. Extend the results of § 3 as far as possible to arbitrary Banach spaces.
Consider in particular the difficult problem of existence of periodic
solutions in closed convex subsets of \( X \), even for the singlevalued case
(under reasonably general assumptions about \( f \); see e.g. § 10 of [2]), perhaps
using ideas such as contained in Proposition 2.4, Theorem 2.4 (which date
back to N. Aronszajn et al.) and Theorem 3.3. An open problem in fixed point
theory is a proof of the following

**Theorem A 2.** Let \( X \) be a Banach space, \( K \subset X \) closed bounded convex,
\( F : K \to 2^X \setminus \emptyset \) usc, \( F(x) \) compact convex, \( F \gamma \)-condensing (\( \gamma \) = Kuratowski's
or Hausdorff's measure of noncompactness) and \( F \) weakly inward (i.e.
\( F(x) \cap \overline{I_K(x)} \neq \emptyset \) on \( K \). Then \( F \) has a fixed point?

See \S\ 18 of [2] and the "last remark" for known special cases.

Concerning the problem \((\ast)\) \( x \in F(t,x) \), \( x(0) = 0 \in r(0) \), \( x(t) \in r(t) \) on \( J \),

the paper [8] contains

**Theorem A 3.** Let \( X \) be a Banach space, \( G = \text{graph} \ r \), \( F(t,x) \neq \emptyset \) compact
on \( U \), \( F \) uniformly continuous on \( G \), \( F(t,x) \subset K(t) \) (with \( K \) weakly usc,
left continuous and \( K(t) \) compact convex), \( \forall \) right continuous and \( G \)
closed, and: \( \forall \varepsilon > 0 \), \( t \in J \), \( x \in r(t) \), \( z \in F(t,x) \) and \( \tau \) with \( 0 < \tau - t \leq \varepsilon \)
there exists \( y \in r(\tau) : (y-x)/((\tau-t) \in \overline{B}(z) \). Then \((\ast)\) has an a.c.
solution on \( J \).

Again the proof is extremely long; try to generalize and simplify. Check
also [4] and more recent contributions of C. Castaings"school", most
probably published in "Seminario d'Analyse Convexe Montpellier (France)".

3. Try to prove reasonable comparison and estimation results for multi-
valued differential equations with possible applications in stability
theory. Chap. 6 of [1] contains only unsatisfactory preliminary results
about the latter, using Lyapunov functions. Consult Prof. J. Eisenfeld
(Arlington) for some models in mathematical biology where interesting
equilibria are to be expected just in some points where the singlevalued
right hand side of the equation has discontinuities, and start a deeper
study of asymptotic behavior.

4. Extend the results from [3] and

stochastic boundary value problems (TR \# 223 UTA; to appear in
Stoch. Anal. Appl.)

to multivalued differential equations, depending measurably on a para-
meter, and discuss its relevance for stochastic control theory. See,
for example, the second part of

control. 2nd Print. Springer-Verlag 1982.

for the latter. Study before (carefully) whether the examples (from
control theory) in § 2, § 3 really yield a big simplification (or
are even necessary) in control theory; see, for example, introductions
to this field such as

J. Wiley 1967.

Springer-Verlag 1982;

see also Example 24.10 in [2] in this light.

5. Read the applications to economics on pp. 243-263 in [1] and study
more in this direction if you are interested.

6. Study applications of multivalued differential equations in problems
where one has hysteresis effects. A recent book on this subject is
(unfortunately in Russian only)

Nauka, Moscow 1983.

Check also western, in particular engineering, literature on this.

7. Check whether Example 3.3 (implicit differential equations) can be
improved (e.g. weaker assumptions w.r. to x') and find more situations
where these results apply while classical ones (local implicit function
theorem etc) do not.
8. Try to improve Chap. 4.7 in [1] on multivalued functional differential equations. Most probably such equations also appear in problems with hysteresis.


Conclusion. By means of these remarks it should have become clear that this course was only an introduction to this interesting subject, that many things remain to be clarified before one can think of writing what is worth to be called a book.

"Last remark": For the present state of the fixed point problem see,