Mini-course

Elements of Multivalued and Nonsmooth Analysis

Vladimir V. Goncharov

Departamento de Matemática, Universidade de Évora
PORTUGAL

Università di Verona
Facoltà di Scienze Matematiche, Fisiche e Naturali
November 2-14, 2015

Program of the mini-course

Lection I	Introduction
Lections II	Elements of Convex Analysis
Lection III	Introduction to Multivalued Analysis
Lection V	Some advanced properties of multifunctions
Lections V	Differential Inclusions
Lecture VI	Viability Theory

Outline

There are two ways to characterize a function $f : X \to Y$ (two kinds of its properties):

- differential or locally
- integral or globally

Depending on X and Y (usually topological vector spaces) the differential characteristics admit various sense but they always mean a rate of changement of one variable (function $f(x)$) with respect to other (argument x)

For example,

- usual derivative, if $X = Y = \mathbb{R}$
- partial derivatives, gradient, if $X = \mathbb{R}^n$, $Y = \mathbb{R}$
- divergence, curl, jacobean etc., if $X = Y = \mathbb{R}^n$

Each differential characteristics can be interpreted in three ways:

(a) physically as the rate of changement (see above) that in different applications has a proper more concrete sense (e.g., velocity if x is time or something like that)

(b) geometrically as slope of the tangent line, position of the tangent plane, degree of extension (contraction) of a solid under some forces etc.

(c) analitically as the possibility to approximate the function f at a neighbourhood of a given point by some simpler function (affine one)

Due to these interpretations the smoothness means

(**a**) existence of an instantaneous velocity (or, in general, rate of changement of some variable) at a given point

(**b**) possibility to pass a tangent line (plane) to the graph at a given point

(**c**) possibility to approximate the function by a linear one near a given point; existence of a certain (continuous) limit etc.

Respectively, nonsmoothness means the lack of the above properties

In other words,

(a) at some time moment a velocity fails to exist: the material point suddenly stops, accelerates or changes direction of the movement; in other (physical) interpretation: the lack of elasticity in a certain material that results appearance of some cracks, splits and so on

(b) there are some "acute" points of the graph (some peaks, edges etc.)

(c) it is impossible to approximate by an affine function due to non existence of limit at a given point

Let $f : \mathbb{R} \to \mathbb{R}$. There are possible various situations

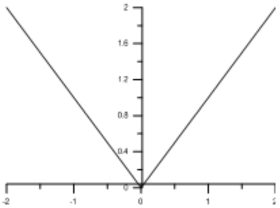
1). Left- and right-sided derivatives $f'_{\pm}(x)$ exist, are finite and different

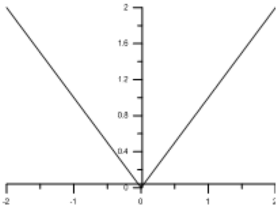
$$f'_{\pm}(x) = \lim_{h\to 0\pm} \frac{f(x+h) - f(x)}{h}$$

Example

$$f(x) = |x| = \begin{cases} x & \text{if} \quad x \geq 0 \\ -x & \text{if} \quad x < 0 \end{cases}$$

$$f'_{+}(0) = 1; \qquad f'_{-}(0) = -1$$


2
1.6
1.2
0.8
0.4
0
-2
-1
0
1



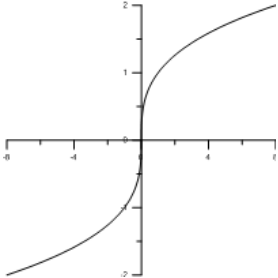
2). Left- and right-sided derivatives $f'_{\pm}(x)$ exist but can be infinite

Examples

a).

$$f(x) = \sqrt[3]{x}$$

$$f'_{\pm}(0) = \lim_{h \to 0\pm} \frac{\sqrt[3]{h}}{h} = \lim_{h \to 0\pm} \frac{1}{h^{2/3}} = +\infty$$

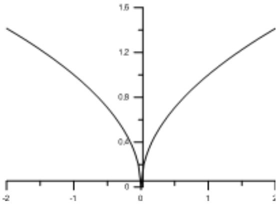


b).

$$f(x) = \sqrt{|x|}$$

$$f'_+(0) = \lim_{h\to 0+} \frac{\sqrt{|h|}}{h} = \lim_{h\to 0+} \frac{1}{\sqrt{|h|}} = +\infty;$$

$$f'_-(0) = \lim_{h\to 0-} \frac{\sqrt{|h|}}{h} = \lim_{h\to 0-} \frac{\sqrt{|h|}}{-|h|} = \lim_{h\to 0+} \left(-\frac{1}{\sqrt{|h|}}\right) = -\infty,$$



3). One of the one-sided derivatives (or both) does not exist (neither finite nor infinite)

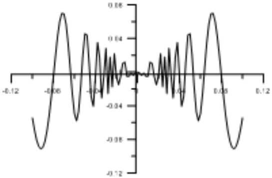
This means that the limit (called derivative number)

$$\lim_{n\to\infty} \frac{f(x+h_n) - f(x)}{h_n}$$

depends on choice of a sequence $h_n \to 0$

Example

$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{if} \quad x \neq 0; \\ 0 & \text{if} \quad x = 0, \end{cases}$$



Given $a \in [-1, 1]$ let us define

$$h_n = \frac{1}{\arcsin a + 2\pi n} \to 0, \quad n \to \infty$$

Then

$$\lim_{n\to\infty} \frac{f(h_n) - f(0)}{h_n} = \lim_{n\to\infty} \frac{h_n \sin(\arcsin\ a + 2\pi n)}{h_n} = a$$

Thus the set of all derivative numbers (so called **contingent derivative**) of the function $f(\cdot)$ is $[-1, 1]$

One can write

$$\text{Cont}\, f(x) = \begin{cases} \left\{\sin\frac{1}{x} - \frac{1}{x}\cos\frac{1}{x}\right\} & \text{se} \quad x \neq 0; \\ [-1, 1] & \text{se} \quad x = 0. \end{cases}$$

Sometimes another so named **paratingent derivative** can be useful. It is defined as the set of all limits

$$\lim_{n\to\infty}\frac{f(x_n)-f(y_n)}{x_n-y_n}$$

where $\{x_n\}$ and $\{y_n\}$ are arbitrary sequences tending to x $(x_n\neq y_n)$

Always $\text{Cont}\, f(x)\subset \text{Parat}\, f(x)$ but the reverse inclusion can fail

In the latter example setting

$$x_n=\frac{1}{\pi/2+2\pi n}\quad \text{e}\quad y_n=\frac{1}{2}x_n$$

we see that

$$\lim_{n\to\infty}\frac{f(x_n)-f(y_n)}{x_n-y_n}=\lim_{n\to\infty}\frac{x_n\sin(\pi/2+2\pi n)-1/2x_n\sin(\pi+4\pi n)}{1/2x_n}$$

$$=2\notin \text{Cont}\, f(0)$$

In this case indeed

$$\text{Parat}\, f\,(0) = \left]-\infty, +\infty\right[$$

To see this it is enough for each $l \in \mathbb{R}$ choose a and b such that $l = 2a - b$ and define

$$x_n := \frac{1}{\arcsin\ a + 2\pi n} \quad \text{e} \quad y_n := \frac{1}{\arcsin\ b + 4\pi n}$$

Exercise 1.1

Finish the proof

If $f : \mathbb{R} \to \mathbb{R}$ is lipschitzean then both $\mathrm{Cont}\, f\,(x)$ and $\mathrm{Parat}\, f\,(x)$ are bounded but nevertheless $\mathrm{Parat}\, f\,(x)$ can be larger as well

Example

$$f(x) = |x|$$

In fact,

$$\mathrm{Cont}\, f\,(0) = \{-1, 1\} \text{ while } \mathrm{Parat}\, f\,(0) = [-1, 1]$$

To see this take $l \in]-1, 1]$, an arbitrary sequence $x_n \to 0+$ and set $y_n = -ax_n$ where

$$a := \frac{1-l}{1+l}$$

Excercise 1.2

Show that

$$\lim_{n\to\infty} \frac{f(x_n) - f(y_n)}{x_n - y_n} = l$$

Observe more that $f(\cdot)$ can be even differentiable at x with $\mathrm{Parat}\, f\,(x)$ not singleton

Example

$$f\,(x) = \begin{cases} x^{3/2} \sin\left(\frac{1}{x}\right) & \text{if} \quad x \neq 0; \\ 0 & \text{if} \quad x = 0, \end{cases}$$

Here the derivative $f'(0)$ exists but $\mathrm{Parat}\, f\,(0) =]-\infty, +\infty[$

Exercise 1.3

Prove this

Observe that this function is not lipschitzean, and its derivative is not continuous at 0,

$$f'(x) = \begin{cases} \frac{3}{2}x^{1/2}\sin\left(\frac{1}{x}\right) + \frac{1}{\sqrt{x}}cos\left(\frac{1}{x}\right) & \text{if} \quad x \neq 0; \\ 0 & \text{if} \quad x = 0, \end{cases}$$

Whenever the function $f(\cdot)$ is **continuously differentiable (smooth)** at x one has

$$\operatorname{Parat} f\,(x) = \operatorname{Cont} f\,(x) = \left\{f'\,(x)\right\}$$

Indeed, given $\{x_n\}$ and $\{y_n\}$ tending to x with $x_n \neq y_n$ we find z_n between x_n and y_n such that

$$f(x_n) - f(y_n) = f'(z_n)(x_n - y_n)$$

(Langrange Theorem), and then by the continuity:

$$\lim_{n\to\infty} \frac{f\,(x_n) - f\,(y_n)}{x_n - y_n} = \lim_{n\to\infty} f'\,(z_n) = f'\,(x)$$

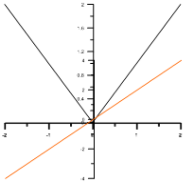
2 Examples. Contingent and paratingent derivatives

There is another approach to generalized differentiation of nonsmooth functions

For the function, e.g., $f(x) = |x|$ let us consider (geometrically) the set of all lines passing below the graph of f and "touching" it at 0

Analitically, the set of slopes of those lines (called subdifferential $\partial f(x)$) will be introduced in sequel for the class of convex functions. In our example

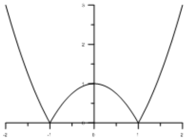
$$\partial f(0) = \mathrm{Parat}\, f(0) = [-1, 1]$$



However, this definition has essential defect: there is a lot of situations when $\partial f(x)$ does not describe well the local structure of the function

Example

$$f(x) = |x^2 - 1|$$



We have $\partial f(-1) = [-1, 0]$ and $\partial f(1) = [0, 1]$ while

$$\partial f(x) = \emptyset$$

whenever $x \in]-1, 1[$ (in spite of the continuous differentiability)

In fact, the subdifferential $\partial f(x)$ characterizes well only so named convex functions (see Lection II)

We say that the lines with slopes from $\partial f(x)$ support the function $f(\cdot)$ at x

In general, a definition of the (generalized) derivative should combine two approaches:

- "supporting" the graph of $f(\cdot)$ (from below or from above) at least locally at some given point
- approximating $f(\cdot)$ near a given point by a simpler function

Mapping studied in the Nonsmooth Analysis can be defined not necessarily in $\mathbb{R}$ but in $\mathbb{R}^n$ or in infinite dimensional Hilbert or Banach spaces

So, we consider $f : X \to \mathbb{R}$ (the case of operators $f : X \to Y$ is out of our objectives)

The motivation comes from, e.g., Calculus of Variations

The basic problem of Calculus of Variations is minimizing the functional

$$f : x(\cdot) \mapsto \int_{t_0}^{t_1} \varphi(t, x(t), \dot{x}(t))\, dt$$

on a set of functions $x : [t_0, t_1] \to \mathbb{R}^n$ satisfying some suplementary conditions (end-point, isoperimetric, holonomic, nonholonomic etc.)

In classic theory (it goes back to J. Bernoulli, L. Euler etc.) the functional f is supposed to be defined on the space $\mathbf{C}^2([t_0, t_1], \mathbb{R}^n)$ of functions twice continuously differentiable on $[t_0, t_1]$ that excludes from consideration a lot of real applications

Necessary optimality condition (famous Euler-Lagrange equation) in classic form requires differentiability of the integrand $\varphi(\cdot, \cdot, \cdot)$ up to the second order (under this assumption the functional f is differentiable in the sense of Fréchet)

All of this is very restrictive and needs to be extended to nonsmooth case

Nowadays, the functional f usually is considered to be defined in a more general space $\mathbf{AC}([t_0, t_1], \mathbb{R}^n)$ of all absolutely continuous functions $x : [t_0, t_1] \to \mathbb{R}^n$

The functional f can be defined also in a space of functions depending on various variables, say

$$f : u(\cdot) \mapsto \int_{\Omega} \Phi\left(x, u(x), \nabla u(x)\right)\, dx$$

Here $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is an admissible function satisfying some boundary (or other) conditions

For instance, the famous Newton's problem on minimum resistence leads to minimization of such kind functional with the integrand

$$\Phi(x, u, \xi) = \frac{1}{1 + |\xi|^2}$$

under some appropriate physically reasonable constraints

Due to necessity to consider resistence of various (not only smooth) solids, nowadays one supposes the functional f to be defined on the (maximally general) space $X = \mathbf{W}^{1,p}(\Omega, \mathbb{R})$ of Sobolev functions $u(\cdot)$, which are integrable (of the order $p \geq 1$) together with their gradients ∇u (in the sense of distributions)

Extending more one can consider the functional f above defined on the space $X = \mathbf{W}^{1,p}(\Omega, \mathbb{R}^m)$ with $m > 1$. Here $\nabla u(x)$ is the (generalized) Jacobi matrix of the function $u(\cdot)$, which can be treated as the deformation of a solid

Variational problems with such functionals have a lot of applications in Elasticity, Plasticity, Theory of Phase Tranfers etc. Certainly, the integrand Φ as well as the functional f may be nonsmooth

So, Nonsmooth Analysis is one of the sourses and motivations of **Multivalued Analysis** since each generalization of the usual derivative (gradient and so on) is a set (multivalued object)

Another sourse is Game Theory, which nowadays has a lot of applications in various fields of Engineering and Economics

We have two players (may be more) A and B (for instance, 2 factories, which have economic relations with each other; two competitive species of animals etc.)

Assume that the player A can choose its strategy x from some set S_A while B chooses a strategy $y \in S_B$

Furthermore, let us given two utility functions $f_A(x, y)$ e $f_B(x, y)$ that mean the profit obtained by the player A or B, respectively, after realization of the strategies x and y

Assuming that the players have no information about behaviour of other, the main problem is **how to choose strategies x and y in order to guarantee a maximal possible profit?**

First of all each player should minimize the risks coming from behaviour (unknown) of the other player. Namely, they define so called marginal functions

$$\bar{f}_A(x) := \inf\{f_A(x,y) : y \in S_B\}$$

$$\bar{f}_B(y) := \inf\{f_B(x,y) : x \in S_A\}$$

and then the respective guaranteed profit

$$f_A^* = \sup\left\{\bar{f}_A(x) : x \in S_A\right\}$$

$$f_B^* = \sup\left\{\bar{f}_B(y) : y \in S_B\right\}$$

Observe that analysing the set of strategies of the "adversary", which tends to minimize the profit

$$\mathfrak{M}_A(x) := \left\{y \in S_B : f_A(x,y) = \bar{f}_A(x)\right\}$$

the player A can diminish his riscs (so, augment the profit), e.g., excluding strategies hardly realizable

Similarly, the player B does considering the set of strategies

$$\mathfrak{M}_B(y) := \left\{x \in S_A : f_B(x,y) = \bar{f}_B(y)\right\}$$

So, the (multivalued) mappings $\mathfrak{M}_A(\cdot)$ and $\mathfrak{M}_B(\cdot)$ called marginal mappings are very useful as well

There is a special class of games (antagonistic games or games with zero sum) where the gain of one of players equals the loss of other, i.e.,

$$f_B(x, y) = -f_A(x, y)$$

In such a case the positive function, say $f(x, y) = f_A(x, y) = -f_B(x, y)$, is called cost of the game, and

$$f_A^* = \sup_{x \in S_A} \inf_{y \in S_B} f(x, y);$$
$$-f_B^* = \inf_{y \in S_B} \sup_{x \in S_A} f(x, y)$$

mean the profit of the player A and the loss of B, respectively

Observe that always

$$f_A^* \leq -f_B^*$$

If, instead, the equality holds

$$f_A^* + f_B^* = 0$$

then the game has an equilibrium

If, moreover, there exists a point $(x^*, y^*) \in S_A \times S_B$ such that

$$f(x^*, y^*) = f_A^*$$

then (x^*, y^*) is said to be a saddle point of the game $(S_A, S_B, f(x, y))$

Multivalued Analysis includes Differential and Integral Calculus as extension of the respective classic Calculus as well as some specific problems (e.g., continuous selections or parametrization)

The counterpart of the Differential Equations Theory on multivalued level is Theory of Differential Inclusions, which studies such objects:

$$\dot{x}(t) \in F(t, x(t))$$

where $F(t, x)$ is a multivalued mapping

Among numerous fields leading to differential inclusions we touch Differential Equations with discontinuous right-hand side and Optimal Control

Let us consider the Cauchy problem

$$\dot{x} = f(t, x), \quad x(t_0) = x_0$$

where the function $f : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}^n$ can be discontinuous w.r.t. x at various points including x_0. Such equations often appear in problems of Mechanics

In the classic sense the problem above may have no solutions. Let, for instance, $t_0 = 0,\ x_0 = 0$ and

$$f(t, x) = \begin{cases} 1 & \text{se} \quad x \leq 0; \\ -1 & \text{se} \quad x > 0. \end{cases}$$

If a solution $x(\cdot)$ is such that $x(t) > 0$ in a neighbourhood of some t^*, say for all $t \in [t^* - \delta, t^* + \delta]$, then $\dot{x}(t) = -1$ and

$$x(t) = -\int_{t^*-\delta}^{t} ds = t^* - \delta - t \leq 0$$

Similarly, we have contradiction assuming that $x(t^*) < 0$

Lection I Introduction
5 Differential equations with discontinuous right-hand side

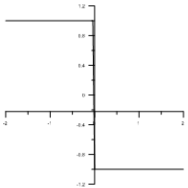
In order to overcome this inconvenience A.F.Filippov in 1960^{th} proposed to relax somehow the problem by considering the Differential Inclusion

$$\dot{x} \in F(t,x), \quad x(t_0) = x_0$$

where

$$F(t,x) := \text{co}\left\{ \lim_{n\to\infty} f(t,x_n) : x_n \to x \right\}$$

Such problem for DI with convex-valued upper semicontinuous right-hand side always admits a solution



Another sourse of Differential Inclusions is Optimal Control Theory that gets a lot of applicatios in numerous fields of technology, natural sciences, economics, even medicine etc.

Suppose that some (physical, biological, economic etc.) process is governed by the differential system

$$\dot{x} = f\left(t, x, u\right), \qquad x(t_0) = x_0,$$

containing a parameter u in the right-hand side. This means that the process can be controlled by substituting in the place of u some (measurable) function $u : [t_0, t_1] \to \mathbb{R}^r$

Assume that the control function $u(\cdot)$ admits its values in some set $U(t, x)$ (possibly depending on the system state x as well)

The problem is **to find an admissible control function $u^*(\cdot)$ and the respective trajectory $x^*(\cdot),\ x^*(t_0) = x_0$, of the equation**

$$\dot{x} = f\left(t, x, u^*(t)\right).$$

which gives minimum to some functional

$$\mathcal{I}\left(x, u\right) = \Phi\left(x\left(t_1\right)\right)$$

It is so called Optimal Control Problem in the **Maier form** (or with **terminal functional**)

The case of more general **Bolza functional**

$$\mathcal{I}(x,u) = \Phi(x(t_1)) + \int_{t_0}^{t_1} \varphi(t, x(t), u(t))\, dt$$

can be easily reduced to a terminal one

Exercise 1.4

Make this reduction

Denoting by

$$F(t,x) := \{f(t,x,u) : u \in U(t,x)\}$$

the set of velocities we naturally associate to our control system the Differential Inclusion

$$\dot{x}(t) \in F(t,x(t))$$

So, we should only minimize the terminal functional

$$\mathcal{I}(x,u) = \Phi(x(t_1))$$

among all the solutions of DI and find a trajectory $x^*(\cdot)$

Applying then Filippov's Lemma (which is a consequence of the measurable selection Theorem, see Lection III) we can construct a measurable control function $u^*(\cdot)$ such that

$$\dot{x}^*(t) = f(t, x^*(t), u^*(t))$$

for almost all $t \in [t_0, t_1]$

In order to minimize a functional on the solution set of DI the notion of attainable set is relevant

Namely, denoting by $\mathcal{H}_F(t_0, x_0)$ the family of all solutions $x(\cdot),\ x(t_0) = x_0$, of DI, the set

$$\mathcal{H}_F(t_0, x_0)(\tau) := \{x(\tau) : x(\cdot) \in \mathcal{H}_F(t_0, x_0)\}$$

is said to be attainable set of DI at the time moment τ

So, the Optimal Control problem is reduced in some sense to the (finite dimensional) minimization problem:

$$\text{Minimize } \{\Phi(x) : x \in \mathcal{H}_F(x_0)(t_1)\}$$

Consequently, we should

- study properties of the attainable sets
- reconstruct a trajectory $x^*(\cdot)$ of DI by its initial and terminal positions

Further development of Optimal Control Theory leads to Differential Games where there are two (or more) control functions corresponding to each of the players (say A and B)

Namely, let us assume that two players (two factories in economic relation; two adversaries in a military conflict etc.) at each time moment $t \in [t_0, t_1]$ have resources $x^1(t)$ and $x^2(t)$, respectively, that satisfy the differential equation

$$\dot{x}(t) = f(t, x(t), u, v),$$
$$x(t_0) = \left(x_0^1, x_0^2\right)$$

where $x(t) = \left(x^1(t), x^2(t)\right)$, and the control parameters u and v admit values in some sets $U(t,x) \subset \mathbb{R}^{r_1}$ and $V(t,x) \subset \mathbb{R}^{r_2}$, respectively

Assuming the antagonistic character of the game, define, furthermore, the cost functional

$$\mathcal{J}\left(x,u,v\right)=\Psi\left(x\left(t_{1}\right)\right)+\int_{t_{0}}^{t_{1}}\psi\left(t,x\left(t\right),u\left(t\right),v\left(t\right)\right)\,dt$$

Thus, the problem is

- for the player A to minimize $\mathcal{J}\left(x,u,v\right)$ among all the strategies $v(\cdot)$ of the player B and then maximize a profit
- for the player B to maximize $\mathcal{J}\left(x,u,v\right)$ among all the strategies $u(\cdot)$ of the player A and then minimize a loss

If in an Optimal Control Problem or in a Differential Game above a phase constraint

$$x(t) \in K$$

appears then this problem can be reduced to so called **viability problem** for Differential Inclusion:

$$\begin{array}{rcl} \dot{x}(t) & \in & F(t, x(t)); \\ x(t) & \in & K; \\ x(t_0) & = & x_0 \in K \end{array}$$

Here K is a (locally) closed set, which can be given, e.g., by means of finite number of algebric equalities and inequalities (K can depend also on t)

For existence of a (viable) solution in the problem above one needs to impose some suplementary tangential hypothesis

In the **next Lection** we consider the simplest class of nonsmooth objects: convex functions and sets

Convex Analysis is the basis of all modern Analysis, combines methods of Abstract Functional Analysis and Geometry

The principal feature of Convex Analysis is **duality**. So, our goal is to explain the relations between various dual objects (conjugate functions, polar sets and so on)

One of the properties (Krein-Milman theorem) will be proved

Then we introduce so important concepts of Convex Analysis as subdifferential of a convex function and normal and tangent cones to a convex set. Some possible generalizations to nonconvex sets will be given

Lection III is devoted to very brief survey of the Multivalued Analysis. We introduce the most current continuity concepts for multivalued mappings concentrating our efforts on the **continuous selection problem**

Then we pay attention to measurability properties of multifunctions and to the concept of the multivalued (Aumann) integral

We will prove two very important theorems on multivalued mappings (Michael Theorem on continuous selections and Kuratowski and Ryll-Nardžewski theorem on measurable choice)

As a consequence of the latter result we formulate Filippov's Lemma on implicit functions we talked already about

In the first part of Lecture IV we prove the fundamental theorem of Multivalued Analysis, so called A.A.Lyapunov's Theorem on the range of vector measure

Then we pass to multivalued mappings, which admit values in functional spaces (of integrable functions), in particular, to mappings with so called decomposability property

We are interested in continuous selections of such mappings (another version of Michael's Theorem)

Here we give a sketch of the nice and suggestive proof of so called Fryzskowski's selections Theorem

In the Lection V we introduce the notion of Differential Inclusion and of its Carathéodory type solution

Further, we give survey of the most significative methods for resolving of the inclusions and sketch of proofs of some important existence theorems

Finally, the last Lection VI will be devoted to Viability Theory or to Differential Inclusions with phase constraints

We conclude, applying one of the methods presented in the previous lecture (namely, method of continuous selections and fixed points) to a viability problem

For thorough studying of the subject I would recommend the following books and some papers:

- Rockafellar R.T. *Convex Analysis*, Princeton University Press, Princeton, New Jersey (1972)
- Ekeland I. & Temam R. *Convex Analysis and Variational Problems*, North-Holland, Amsterdam (1976)
- Kuratowski K. *Topology*, Vol. I, PWN -Polish Scientific Publishers & Academic Press, London (1958)
- Aubin J.-P. & Frankowska H. *Set-Valued Analysis*, Birkhauser, Boston (1990)
- Aubin J.-P. & Cellina A. *Differential Inclusions*, Springer, Berlin (1984)

- Michael E. *Continuous selections I*, Ann. Math. **63** (1956), 361-381
- Himmelberg C.J. *Measurable relations*, Fund. Math. **87** (1975), 53-72
- Fryszkowski A. *Continuous selections for a class of non-convex multivalued maps*, Studia Math. **76** (1983), 163-174
- Goncharov V.V. *Existence of solutions of a class of differential inclusions on a compact set*, Sib. Math. J. **31** (1990), 727-732

Outline

We start with **convex functions** $f : X \to \mathbb{R} \cup \{+\infty\}$ (for convenience it is allowed to admit infinite values as well)

Here X can be any Hilbert, Banach or a Topological Vector Space

Definition

A function $f(\cdot)$ is said to be convex if for each $x, y \in X$ and each $0 \leq \lambda \leq 1$ the inequality

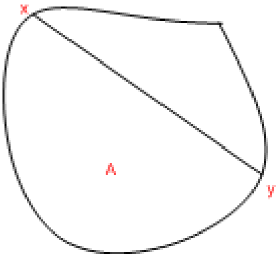
$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y). \qquad (1)$$

holds. If in (1) the strict inequality takes place whenever $x \neq y$ and $0 < \lambda < 1$ then we say that $f(\cdot)$ is strictly convex

We consider also **convex sets** as a counterpart to convex functions

Definition

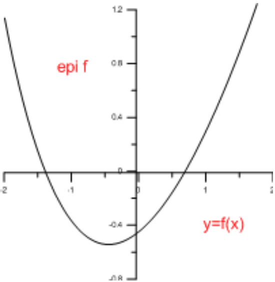
A set A is said to be convex if $\lambda x + (1 - \lambda) y \in A$ for each $x, y \in A$ and $0 \leq \lambda \leq 1$



The convex functions and sets are usually studied together because to each (convex) function $f : X \to \mathbb{R} \cup \{+\infty\}$ one can associate the (convex) set

$$\operatorname{epi} f := \{(x, a) : a \geq f(x)\}$$

(epigraph of $f(\cdot)$)



On the other hand, to each (convex) set $A \subset X$ one can associate the (convex) indicator function

$$\mathbf{I}_A(x) = \begin{cases} 0 & \text{if} \quad x \in A \\ +\infty & \text{if} \quad x \notin A \end{cases}$$

Observe, moreover, the following simple fact

Proposition

The function $f : X \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous if and only if its epigraph $\mathrm{epi}\, f$ is closed

Consider also the (effective) domain

$$\mathrm{dom}\, f := \{x \in X : f(x) < +\infty\}$$

and say that $f(\cdot)$ is proper if $\mathrm{dom}\, f \neq \emptyset$

For topology of the convex sets and functions we refer to the books given in the bibliography. Here instead we only emphasize two remarkable properties:

- if a convex function $f : X \to \mathbb{R} \cup \{+\infty\}$ is upper bounded in a neighbourhood of some point $x_0 \in \operatorname{dom} f$ then it is continuous at x_0
- for very large class of spaces X (including all Banach spaces) a convex lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ is continuous (and even locally lipschitzean) on the interior of the effective domain $\operatorname{dom} f$

The main geometric property of convex sets is the linear **separation**, i.e., the possibility to separate disjoint sets by affine manifolds (lines, planes etc.)

In a finite dimensional space this is almost obvious (geometric, algebraic) fact, but, in general, it is derived from **Hahn-Banach Theorem**, which is the basic principle of Functional Analysis

Its geometric formulation can be given as follows

Mazur's Theorem

Let X be a Locally Convex Space (LCS). If $C \subset X$ is convex, open and such that $C \cap L = \emptyset$ for some affine manifold $L \subset X$ then there exists a closed (affine) hyperplane $H \supset L$ with $C \cap H = \emptyset$

Each affine hyperplane can be written analytically as

$$H = \{x \in X : \varphi(x) = \alpha\} \qquad (*)$$

for some linear continuous functional φ ($\varphi \in X'$ where X' is the dual LCS)

For the sake of symmetry we denote such functional φ by x' and in the place of $\varphi(x)$ write $\langle x, x' \rangle$

To (*) we associate naturally two (closed) half-spaces

$$H_\alpha^+ = \left\{x \in X : \left\langle x, x' \right\rangle \le \alpha\right\} \quad \text{and} \quad H_\alpha^- := \left\{x \in X : \left\langle x, x' \right\rangle \ge \alpha\right\}$$

The respective open half-spaces will be denoted by $\mathring{H}_\alpha^+$ and $\mathring{H}_\alpha^-$, resp.

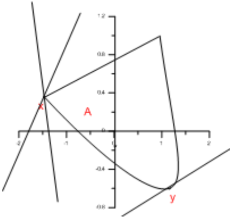
So, let us remind two main **Separation Theorems**, which give basis of the whole Convex Analysis

Separation Theorems

I Let $A, B \subset X$ be convex nonempty sets such that A (or B) is open and $A \cap B = \emptyset$. Then there exist $x' \in X'$, $x' \neq 0$, and $\alpha \in \mathbb{R}$ (a hyperplane $H \subset X$ associated to x' and α) such that $A \subset H_\alpha^+(x')$ and $B \subset H_\alpha^-(x')$ (we say that H separates the sets A and B)

II Let $A, B \subset X$ be convex nonempty sets such that A is compact, B is closed and $A \cap B = \emptyset$. Then there exist $x' \in X'$, $x' \neq 0$, and $\alpha \in \mathbb{R}$ (a hyperplane $H \subset X$ associated to them) such that $A \subset \mathring{H}_\alpha^+(x')$ and $B \subset \mathring{H}_\alpha^-(x')$ (in this case H separates the sets A and B strictly)

If $A \subset X$ is convex, closed and $\operatorname{int} A \neq \emptyset$ then each point $x \in \partial A$ (boundary of A) can be (nonstrictly) separated from intA by some closed hyperplane H called **supporting hyperplane** (see Separation Theorem I)



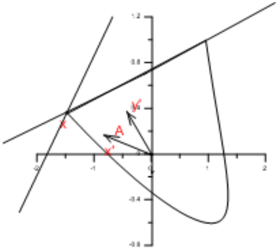
We see that at some points supporting hyperplane is unique (at the point y in pic.), at others no (at the point x)

In the first case we say that the set A (or its boundary) is smooth at y

Lection II Elements of Convex Analysis
2 Separation of convex sets. Support function

Otherwise, if we fix a hyperplane H (an "orthogonal" vector x', which defines H) then it can "touch" (be supporting) the convex set A at unique point x or no (see pic.)

In the first case we say that A is strictly convex (or rotund) at x

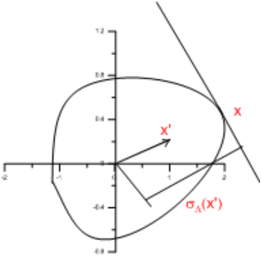


Here we have the first duality of Convex Analysis between rotundity and smoothness

For quantitative description of convex sets we need to introduce the support function $\sigma_A : X' \to \mathbb{R} \cup \{+\infty\}$ associated to A:

$$\sigma_A\left(x'\right) := \sup\left\{\left\langle x, x'\right\rangle : x \in A\right\}, \quad x' \in X'$$

The following picture illustrates the geometric sense of the support function (the maximal distance, for which one should move the plane in the direction x' in order to touch the boundary of A)



We have naturally

$$A \subset \{x \in X : \langle x, x' \rangle \leq \sigma_A(x')\}$$

However, if we take all of vectors $x' \in X'$ (in a normed space it is enough to choose those with $\|x'\| = 1$) we obtain complete representation of A

Representation of a convex closed set

If X is a normed space with the norm $\|\cdot\|$ then the equality

$$A = \bigcap_{\|x'\|=1} \{x \in X : \langle x, x' \rangle \leq \sigma_A(x')\}$$

holds

This formula gives the "external" representation of a convex closed set as the intersection of (supporting) half-spaces

Another "internal" representation is given by famous **Krein-Milman Theorem** proved in finite dimensions by H. Minkowski in the initial of XX century

To formulate this Theorem let us return to the duality between rotundity and smoothness considered above and make it more precise

There are two dual approaches to study a convex (closed) set

- To fix $x \in \partial A$ and consider the set

$$F^x := \left\{x' \in \partial \mathbf{B} : \langle x, x' \rangle = \sigma_A(x')\right\}$$

  If F^x is a singleton then we have smoothness at x. If it is not then we come to the notion of the normal cone (considered below)

- To fix $x' \in X'$ with $\|x'\| = 1$ and consider the set

$$F_{x'} := \left\{x \in A : \langle x, x' \rangle = \sigma_A(x')\right\}$$

  called exposed face of A. If $F_{x'}$ is a singleton (called exposed point) then we have rotundity w.r.t. x'

There is another type of faces besides the exposed ones, which are used in Krein-Milman theorem

Definition

A convex subset $F \subset A$ is said to be extremal face of A if for each $x, y \in A$ such that $]x, y[\cap F \neq \emptyset$ we have $x, y \in F$

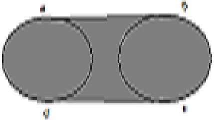
We say that a point $x \in \partial A$ is extremal point of A if $F = \{x\}$ is its (0-dimensional) extremal face

Exercise 2.1

Prove that each exposed face (point) is also an extremal one

The opposite implication is, in general, false already in $\mathbb{R}^2$ as the following example shows

Example



Another important property of extremal faces (unlike exposed ones) is the **transitivity**

- F is an extremal face of A & G is an extremal face of $F \Longrightarrow G$ is an extremal face of A

Now we are able to formulate the basic Theorem

Krein-Milman Theorem

Let $A \subset X$ be a nonempty convex compact set (X is a Locally Convex Space). Then we have

- $\operatorname{ext} A \neq \emptyset$
- $A = \overline{\operatorname{co}} \operatorname{ext} A$

Here it is important that X is an arbitrary LCS because afterwards this theorem will be applied to Banach spaces with the weak topology, which is not normable

Observe that the set of extreme points of a compact set may be not closed already in the space $\mathbb{R}^3$ as the following example shows

Example $A = \operatorname{co}(B \cup C)$ where

$$
\begin{aligned}
B &= \{(x_1, x_2, x_3) : x_1^2 + x_2^2 \leq 1,\ x_3 = 0\} \\
C &= \{(x_1, x_2, x_3) : \max(|x_1|, |x_3|) \leq 1,\ x_2 = 0\}
\end{aligned}
$$

In finite dimensions, nevertheless, the convex hull $\operatorname{co} \operatorname{ext} A$ is always closed, and the closure in the Krein-Milman Theorem can be omitted (this is so named Minkowski Theorem proved at the beginning of XX century)

Exercise 2.2

Prove that for each convex compact $A \subset \mathbb{R}^2$ the set $\operatorname{ext} A$ is closed

Hypothesis of the compactness of the set A in the Krein-Milman Theorem is essential

Exercise 2.3

Prove that the unit closed ball in the space of all summable functions $L^1(T, \mathbb{R})$ has no extreme points. Here $T = [a, b]$ is a segment of the number line

By the way, it follows from the assertion above that the space $L^1(T, \mathbb{R})$ can not be conjugate for some Banach space (in particular, it is not reflexive)

Let us define now a construction for functions similar to the support function for sets

Namely, fix $(x', a) \in (X \times \mathbb{R})' = X' \times \mathbb{R}$ and consider $\sigma_{\mathrm{epi}\, f}(x', a)$. Since $\mathrm{epi}\, f$ is upper unbounded, we obviously have

$$\sigma_{\mathrm{epi}\, f}(x', a) = +\infty \ \text{ whenever } \ a > 0$$

The case $a = 0$ characterizes $\mathrm{dom}\, f$ but not properly $f(\cdot)$

So, after normalizing (dividing by $|a|$) we get $f^* : X' \to \mathbb{R} \cup \{+\infty\}$,

$$f^*(x') := \sigma_{\mathrm{epi}\, f}(x', -1) = \sup_{x \in X} (\langle x, x' \rangle - f(x))$$

called **conjugate function** (or **Legendre-Fenchel transform**) of $f(\cdot)$

The conjugate function and specially the **double conjugation** (called also **Γ-regularization** of the function $f : X \to \mathbb{R} \cup \{+\infty\}$) are very important for various fields of Analysis and for Applications

If the space X is reflexive then the second conjugate function $f^{**} = (f^*)^*$ is defined on the same space X and admits the following (equivalent) characterizations

- $f^{**}(x)$ is the pointwise supremum of all the affine functions below $f(\cdot)$
- $f^{**}(x)$ is the greatest convex lower semicontinuous function among those below $f(\cdot)$
- $\operatorname{epi} f^{**} = \overline{\operatorname{co}} \operatorname{epi} f$ for all $x \in \operatorname{int} \operatorname{dom} f^{**}$

Exercise 2.4

Calculate the conjugation of the following functions

(a) $f:\mathbb{R}\to\mathbb{R},\ f(x)=\frac{|x|^p}{p}$, with $p>1$

(b) $f:\mathbb{R}\to\mathbb{R},\ f(x)=e^{|x|}$

(c) $f:\mathbb{R}^2\to\mathbb{R},\ f(x)=e^{x_1+2x_2}$

Furthermore, the following formula holds:

$$f^{**}(x) = \inf \left\{ \sum_{i=1}^{k} \lambda_i f(x_i) : \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1,\ x_i \in X,\ \sum_{i=1}^{k} \lambda_i x_i = x \right\}$$

If $X = \mathbb{R}^n$ then in the above formula one can set $k = n + 1$

In particular, from these characterizations one deduces

Theorem on double conjugation

For each proper convex lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ (and only for that) we have

$$f^{**}(x) = f(x) \text{ for all } x \in X$$

The second conjugate function has a lot of applications in Calculus of Variations. For instance, in order to resolve the minimization problem

$$\text{minimize}\left\{\int_{\Omega} f(\nabla u(x))\,dx : u(\cdot)\in u_0(\cdot)+\mathbf{W}_0^{1,1}(\Omega)\right\},$$

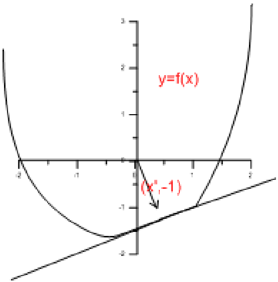
where $f(\cdot)$ is, in general, nonconvex integrand, one usually minimizes first the relaxed functional

$$\int_{\Omega} f^{**}(\nabla u(x))\,dx$$

and then by using obtained minimizer $\hat{u}(\cdot)$ constructs a minimizer of the original one

Similarly, as in the case of sets we have alternative

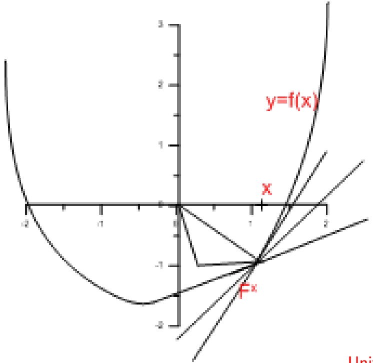
- Given $x' \in X'$ (with $\|x'\| = 1$ in the case of a normed space) consider the set
$$F_{x'} := \{(x, f(x)) : \langle x, x' \rangle - f(x) = f^*(x')\},$$
which is nothing else than an exposed face of $\operatorname{epi} f$. If $F_{x'}$ is a singleton then the function $f(\cdot)$ is strictly convex w.r.t. x'



- Otherwise, given $x \in \operatorname{dom} f$ consider the set

$$F^x := \{x' \in X' : \langle x, x' \rangle - f(x) = f^*(x')\}$$

(the set of all "directions", in which the respective ("orthogonal") hyperplane touches $\operatorname{epi} f$ at x)



Recalling the definition of the conjugate function we have

Definition

The set

$$\partial f(x) := \left\{x' \in X' : f^*(x') = \langle x, x' \rangle - f(x)\right\}$$
$$= \left\{x' \in X' : f(y) \geq f(x) + \langle x - y, x' \rangle \ \ \forall y \in X\right\}$$

is said to be subdifferential of $f(\cdot)$ at x

Definition

In general, a function $f : X \to \mathbb{R} \cup \{+\infty\}$ (not necessarily convex nor lsc) is said to be subdifferentiable at $x \in X$ if $\partial f(x) \neq \emptyset$

In fact, subdifferentiability is equivalent to (local) convexity

Proposition

$f(\cdot)$ is subdifferentiable at $x \in \operatorname{int}\operatorname{dom} f$ iff

$$f(x) = f^{**}(x)$$

In this case $\partial f(x) = \partial f^{**}(x)$

In other words, subdifferential does not distinguish a function $f(\cdot)$ near x from its "convex envelope"

Let us emphasize now some important properties of ∂f for a convex lsc function $f : X \to \mathbb{R} \cup \{+\infty\}$

- $\partial f(x)$ is always convex and closed subset of X'
- if $x \in \operatorname{int} \operatorname{dom} f$ then $\partial f(x)$ is nonempty bounded, consequently, weakly compact in X'
- for any sequence $\{(x_n, x_n')\} \subset X \times X'$ such that $x_n \to x$, $\{x_n'\}$ converges weakly to $x' \in X'$ and $x_n' \in \partial f(x_n)$ we always have $x' \in \partial f(x)$ (the graph of ∂f is strongly$\times$weakly sequentially closed)
- for all $x, y \in \operatorname{dom} f$ and all $x' \in \partial f(x)$, $y' \in \partial f(y)$ the inequaity

$$\langle x - y, x' - y' \rangle \geq 0$$

holds. We say that the mapping $x \mapsto \partial f(x)$ is monotone (or accretive)

The condition of minimum of a convex function can be easily expressed in terms of the subdifferential

$x \in X$ is a minimum point of a convex lsc function $f(\cdot)$ iff one of the conditions below holds

- $0 \in \partial f(x)$
- $x \in \partial f^*(0)$

The last condition is equivalent to the first one due to the following assertion

$x' \in \partial f(x)$ iff $x \in \partial f^*(x')$ iff

$$f(x) + f^*(x') = \langle x, x' \rangle$$

Subdifferential calculus for convex functions includes the rules

- if $\lambda > 0$ then $\partial(\lambda f)(x) = \lambda\partial f(x)$
- always $\partial f(x) + \partial g(x) \subset \partial(f+g)(x)$
- **(the sum rule)** the equality

$$\partial f(x) + \partial g(x) = \partial(f+g)(x)$$

holds true whenever

there exists a point $\bar{x} \in \mathrm{dom}\, f \cap \mathrm{dom}\, g$ **at which either the function** f **or** g **is continuous**

Similarly to convex lsc functions there is duality between a convex set and its polar

Definition

The set

$$A^0 := \left\{x' \in X' : \langle x, x' \rangle \leq 1 \ \ \forall x \in A\right\} \\ = \left\{x' \in X' : \sigma_A(x') \leq 1\right\}$$

is said to be polar to the set $A \subset X$ (A is not necessarily convex nor closed)

Let us list the main properties of the polar sets

1. if $A, B \subset X$ are such that $A \subset B$ then $B^0 \subset A^0$
2. if $A \subset X$ and $\lambda \neq 0$ then $(\lambda A)^0 = \lambda^{-1} A^0$
3. for each family $\{A_\alpha\}_{\alpha \in I}$ of subsets of X one has

$$\left(\bigcup_{\alpha \in I} A_\alpha \right)^0 = \bigcap_{\alpha \in I} A_\alpha^0$$

4. if each $A_\alpha \subset X$ is closed convex and contains the origin then

$$\left(\bigcap_{\alpha \in I} A_\alpha \right)^0 = \overline{\mathrm{co}} \left(\bigcup_{\alpha \in I} A_\alpha^0 \right)$$

The main property of polar sets is the following

Bipolar Theorem

For each $A \subset X$ (X is a reflexive Banach space) we have

$$A^{00} = \overline{\text{co}}\,(A \cup \{0\})$$

Thus $A = A^{00}$ iff A is closed, convex and $0 \in A$

Exercise 2.6

Justify the properties 1.-4. of the polar sets. Proving the equality 4. use the Bipolar Theorem

Exercise 2.7

Construct the polars to the following sets:

(a) $A = \{x = (x_1, x_2) \in \mathbb{R}^2 : (x_1 - 1)^2 + (x_2 - 1)^2 \leq 4\}$

(b) $A = \mathrm{co}\,\{(0, 1), (2, 0), (0, -1), (-2, 0)\}$

(c) $A = \{x = (x_1, x_2) \in \mathbb{R}^2 : |x_2| \leq 1 - x_1^2,\ |x_1| \leq 1\}$

Returning now to the first approach to convex sets introduce the notion of normal cone at a point $x \in A$ as

$$\begin{aligned} \mathbf{N}_A(x) &:= \{x' \in X' : \langle x, x' \rangle = \sigma_A(x')\} \\ &= \{x' \in X' : \langle y - x, x' \rangle \leq 0 \ \ \forall y \in X\} \end{aligned}$$

The normal cone can be also interpreted as $\partial \mathbf{I}_A(x)$

$\mathbf{N}_A(x)$ is a convex closed cone, which equals $\{0\}$ if $x \in \operatorname{int} A$

The normal cone is always non trivial ($\neq \{0\}$) whenever $x \in \partial A$ and $\operatorname{int} A \neq \emptyset$ (it follows from the first separation theorem)

Another important object associated to a convex closed set, which has a lot of applications (in particular, in Viability Theory), is tangent cone

It is defined (unlike the normal cone) in the same space X as A (not in dual):

$$\mathbf{T}_A(x) := (\mathbf{N}_A(x))^0 = \{v \in X : \langle v, x' \rangle \leq 0 \;\; \forall x' \in \mathbf{N}_A(x)\}$$

It is nontrivial closed convex cone, and $\mathbf{T}_A(x) = X$ whenever $x \in \operatorname{int} A$

The normal and tangent cones can be nicely characterized by means of the distance function to the set (in a normed space)

Denoting by $\mathrm{d}_A(\cdot)$ the distance from a point to the set A in X we have

I $\mathbf{N}_A(x) \cap \overline{\mathbf{B}} = \partial \mathrm{d}_A(x)$

II $\mathbf{N}_A(x) = \bigcup_{n=1}^{\infty} n\, \partial \mathrm{d}_A(x) = \bigcup_{\lambda>0} \lambda\, \partial \mathrm{d}_A(x)$

III $\mathbf{T}_A(x) = \left\{ v \in X : \lim_{\lambda \to 0+} \frac{1}{\lambda} \mathrm{d}_A(x + \lambda v) = 0 \right\}$

IV $\mathbf{T}_A(x) = \overline{\bigcup_{\lambda>0} \frac{A-x}{\lambda}}$

If $A \subset X$ is a closed nonconvex set then there is a multiplicity of notions of normal as well as tangent cones

Let us mention some of the tangent cones, which will be used mostly in viability theorems. We give only some definitions without comments, interpretation and properties

- $\mathbf{T}_A^b(x) = \left\{ v \in X : \liminf_{\lambda \to 0+} \frac{1}{\lambda} \mathrm{d}_A(x + \lambda v) = 0 \right\}$
  Bouligand's tangent (ou contingent) cone
- $\mathbf{T}_A^a(x) = \left\{ v \in X : \lim_{\lambda \to 0+} \frac{1}{\lambda} \mathrm{d}_A(x + \lambda v) = 0 \right\}$ Adjacent cone
- $\mathbf{T}_A^c(x) = \left\{ v \in X : \lim_{\lambda \to 0+, y \to x, y \in A} \frac{1}{\lambda} \mathrm{d}_A(y + \lambda v) = 0 \right\}$
  Clarke's tangent cone

Outline

We will use the following notations and basic definitions:

- 2^X is the family of all nonempty closed subsets $A \subset X$
- $\mathrm{comp}\, X$ $(\mathrm{conv}\, X)$ is the family of all compact (respectively, compact and convex) sets $A \subset X$
- $F : X \to 2^Y$ (or $F : X \rightrightarrows Y$) is a multivalued mapping (or multifunction)
- $\mathrm{dom}\, F := \{x \in X : F(x) \neq \emptyset\ \}$ is said to be the domain of F
- $\mathrm{graph}\, F := \{(x, y) \in X \times Y : y \in F(x)\}$ is the graph of F

- $F^{-1}(C) := \{x \in X : F(x) \subset C\}$ is said to be the small preimage of $C \subset Y$ (under the mapping F)
- $F_{-1}(C) := \{x \in X : F(x) \cap C \neq \emptyset\}$ is the total preimage of C
- $F(A) := \bigcup_{x \in A} F(x)$ is the total image of $A \subset X$
- $F^{-1} : Y \rightrightarrows X, \ F^{-1}(y) := \{x \in X : y \in F(x)\}$ is said to be the inverse mapping of F

The subdifferential $\partial f(x)$, the normal and tangent cones $N_A(x)$ and $T_A(x)$ are examples of multivalued mappings

Another example is the mapping $\xi \mapsto \mathcal{H}_f(\xi)$ associating to each $\xi \in X$ the **set of solutions to the Cauchy problem**

$$\dot{x}(t) = f(t, x(t)), \qquad x(t_0) = \xi$$

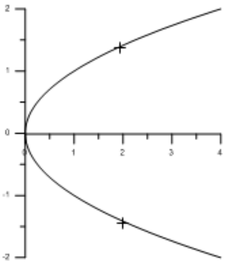
Similarly, to the control system

$$\begin{aligned} \dot{x} &= f(t, x, u) \\ x(t_0) &= \xi \\ u &\in U(t, x) \end{aligned}$$

one can also associate the mulivalued mapping $\xi \mapsto \mathcal{S}(\xi)$ where $S(\xi)$ is the set of all the trajectories of this system

Furthermore, the **inverse mapping** for an arbitrary (single-valued) mapping is, in general, multivalued (if there is no injectivity)

For instance, the function $y = x^2$ is "invertible" just for $x \geq 0$, although x can admit negative values as well. In general, (multivalued) inverse mapping here has the form $x = F(y) = \{\sqrt{y}, -\sqrt{y}\},\ y \geq 0$



2 Vietoris semicontinuity

The first definition of continuity for multivalued mappings is due to L. Vietoris. It is associated with some topology in the space 2^X (so called exponential or Vietoris topology)

Definition

$F : X \rightrightarrows Y$ is said to be **upper semicontinuous at a point $x_0 \in X$ (by Vietoris)** if for each open set $V \supset F(x_0)$ there exists a neighbourhood U of x_0 such that $V \supset F(x)$ for all $x \in U$

$F : X \rightrightarrows Y$ is said to be **lower semicontinuous at a point $x_0 \in X$ (by Vietoris)** if for each open set $V \subset Y$ with $F(x_0) \cap V \neq \emptyset$ there exists a neighbourhood U of x_0 such that $F(x) \cap V \neq \emptyset$ for all $x \in U$

$F : X \rightrightarrows Y$ is said to be **continuous at $x_0 \in X$ (by Vietoris)** if it is both lower and upper semicontinuous by Vietoris at this point

2 Vietoris semicontinuity

Let's give the global version of semicontinuity

Definition

Naturally, $F : X \rightrightarrows Y$ is upper (lower) semicontinuous by Vietoris if it is upper (lower) semicontinuous at each point $x_0 \in X$

Otherwise, $F : X \rightrightarrows Y$ is upper (lower) semicontinuous by Vietoris if for each open $V \subset Y$ the small preimage $F^{-1}(V)$ (respectively, the total preimage $F_{-1}(V)$) is open in X

Let's give the global version of semicontinuity

Definition

Naturally, $F : X \rightrightarrows Y$ is upper (lower) semicontinuous by Vietoris if it is upper (lower) semicontinuous at each point $x_0 \in X$

Otherwise, $F : X \rightrightarrows Y$ is upper (lower) semicontinuous by Vietoris if for each open $V \subset Y$ the small preimage $F^{-1}(V)$ (respectively, the total preimage $F_{-1}(V)$) is open in X

Upper semicontinuity of multifunctions is connected with the other important property: closedness of the graph. Namely,

- each upper semicontinuous multifunction $F : X \to 2^Y$ has the closed graph (at least, if Y is a metric space)
- converse is true whenever F admits values in a common compact set

Furthermore, assuming Y to be a metric space with the distance $\mathrm{d}(\cdot,\cdot)$, in the family of bounded sets from 2^Y one can define the metrics (so called **Pompeiu-Hausdorff metrics**) by the formula

$$\mathfrak{D}(A,B) := \max\left\{\sup_{x\in A}\inf_{y\in B}\mathrm{d}(x,y), \sup_{y\in B}\inf_{x\in A}\mathrm{d}(x,y)\right\}$$

If Y is a normed space with the closed unit ball $\overline{\mathbf{B}}$ then we have another representation

$$\mathfrak{D}(A,B) = \inf\left\{\varepsilon>0 : A\subset B+\varepsilon\overline{\mathbf{B}} \text{ and } B\subset A+\varepsilon\overline{\mathbf{B}}\right\}$$

If X is a Banach space then

there exists a homeomorphism (even an isometry) between the space $\operatorname{conv} X$ endowed with the Pompeiu-Hausdorff metrics $\mathfrak{D}(\cdot,\cdot)$ and the subspace (indeed, a closed cone) in $\mathcal{C}(X')$ of all continuous functions $X' \to \mathbb{R}$ (with the usual $\sup$-norm): $A \mapsto \sigma_A(\cdot)$. Namely,

$$\mathfrak{D}(A,B) = \sup_{\|x'\|=1} \left|\sigma_A(x') - \sigma_B(x')\right|$$

This metrics (consisting, in fact, of two semimetrics) suggests another definition of the semicontinuity

Definition

$F : X \rightrightarrows Y$ is said to be **upper semicontinuous (by Hausdorff)** at a point $x_0 \in X$ if for any $\varepsilon > 0$ there exists a neighbourhood U of x_0 such that

$$F(x) \subset F(x_0) + \varepsilon\overline{\mathbf{B}} \qquad \forall x \in U$$

$F : X \rightrightarrows Y$ is said to be **lower semicontinuous (by Hausdorff)** at a point $x_0 \in X$ if for any $\varepsilon > 0$ there exists a neighbourhood U of x_0 such that

$$F(x_0) \subset F(x) + \varepsilon\overline{\mathbf{B}} \qquad \forall x \in U$$

In general, two (upper, lower) semicontinuity concepts (by Vietoris or by Hausdorff) are different but they coincide if Y is a Banach space and $F : X \to \mathrm{comp}\, Y$

Observe that speaking about multivalued mappings, the new operations appear, namely, one can consider the **union or intersection** of given multifunctions. So, the natural question arises: if some mappings $F_1, F_2 : X \rightrightarrows Y$ are upper (lower) semicontinuous then what can we say about the mappings $x \mapsto F_1(x) \cup F_2(x)$ and $x \mapsto F_1(x) \cap F_2(x)$?

It turns out that the union $x \mapsto F_1(x) \cup F_2(x)$ is **upper** (**lower**) semicontinuous whenever both F_1 and $F_2 : X \to 2^Y$ are **upper** (respectively, **lower**) semicontinuous

As about the intersection the situation is much complicated. On one hand, we have

always the intersection $x \mapsto F_1(x) \cap F_2(x)$ is **upper** semicontinuous whenever both F_1 and $F_2 : X \to 2^Y$ are **upper** semicontinuous (provided just that Y is normal, thus for normed spaces it is OK)

The latter implication is not true for intersections as one can see from the following simple example

Example

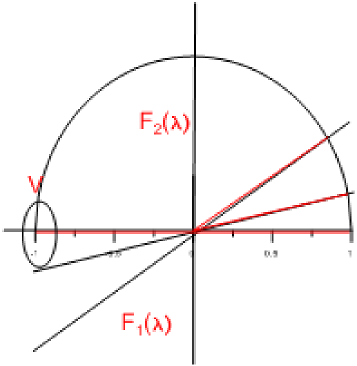
Let $F_1 : \mathbb{R} \to \mathbb{R}^2$ be the multifunction, which associates to each $\lambda \in [0, 1]$ the segment of the line in $\mathbb{R}^2$:

$$F_1(\lambda) := \{(x_1, x_2) : x_2 = \lambda x_1, \ \ -1 \le x_1 \le 1\}$$

and

$$F_2(\lambda) := \{(x_1, x_2) : x_1^2 + x_2^2 \le 1, \, x_2 \ge 0\}, \qquad \lambda \in \mathbb{R}$$

Then F_1 and F_2 are even **continuous** (F_2 is constant) but the intersection is not lower semicontinuous at $\lambda = 0$ (see pic.)


$F_2(\lambda)$
V
$F_1(\lambda)$

However, let us mention the case very important for applications when intersection inherits the lower semicontinuity

Theorem

Let Y be a metric space with the distance $\mathrm{d}(\cdot,\cdot)$ and $F: X \to 2^Y$ be a **lower semicontinuous** (by Vietoris) multifunction. Assume that a continuous (single-valued) function $f: X \to Y$ and a lower semicontinuous real valued function $\varphi: X \to]0,+\infty[$ are such that

$$\Phi(x) := F(x) \cap \mathbf{B}(f(x), \varphi(x)) = \{y \in F(x) : \mathrm{d}(y, f(x)) < \varphi(x)\}$$

is not empty for all x. Then the set-valued mapping $\Phi(\cdot)$ as well as $x \mapsto \overline{\Phi(x)}$ are **lower semicontinuous** (by Vietoris)

Notice that the **strict** inequality here (the ball is **open**) is extremely important

5 Continuous selections. Michael's theorem

The **continuous selections problem** is very important for Multivalued Analysis. It is new one, i.e., has no counterpart in the Classic Analysis

A single-valued mapping $f : X \to Y$ is said to be **selection** of the multifunction $F : X \rightrightarrows Y$ if $f(x) \in F(x)$ for each $x \in X$

5 Continuous selections. Michael's theorem

The **continuous selections problem** is very important for Multivalued Analysis. It is new one, i.e., has no counterpart in the Classic Analysis

A single-valued mapping $f : X \to Y$ is said to be **selection** of the multifunction $F : X \rightrightarrows Y$ if $f(x) \in F(x)$ for each $x \in X$

A fundamental result was obtained by E. Michael in 1950^{th}

Michael's Theorem

Let X be an arbitrary **paracompact** topological space (in particular, metric space) and Y be a **Banach space**. Assume that $F : X \rightrightarrows Y$ is a **lower semicontinuous** (by Vietoris) multifunction admiting **nonempty closed and convex** values

Then a **continuous selection** $f : X \to Y$ of F exists and can be chosen such that $f(x_0) = y_0$ where $(x_0, y_0) \in \operatorname{graph} F$ is a given point

Let us pay attention to all of the hypotheses of Michael's Theorem. Indeed, all of them are essential and can not be dropped

The first example of a **continuous** multifunction $F : \mathbb{R} \rightrightarrows \mathbb{R}^2$ with closed but nonconvex values, which has no any continuous selection, was constructed by A.F. Filippov at the beginning of 1960^{th}

Furthermore, the **existence of a continuous selection** passing through an *a priori* given point of the graph is a **sufficient condition** for the lower semicontinuity

So, continuous selections give a characterization of **lower** semicontinuity similarly as closedness of the graph characterizes **upper** semicontinuity

Let us pass to the Integral Calculus for multivalued mappings and, first of all, define measurability concepts for multifunctions

Definition

Let T be a measurable space with a σ-algebra $\mathfrak{M}$ of measurable sets (e.g., we may consider $T = [0, 1]$ with the σ-algebra $\mathfrak{L}$ of Lebesgue measurable sets or the σ-algebra of Borel sets). Let also X be any topological (or metric) space

We say that the mapping $F : T \rightrightarrows X$ is measurable (weakly measurable) if for any open $U \subset X$ the preimage

$$F^{-1}(U) = \{t \in T : F(t) \subset U\}$$

(resp., $F_{-1}(U) = \{t \in T : F(t) \cap U \neq \emptyset\}$) is measurable

It is easy to show that

- measurability of F implies weak measurability if X is any metric space
- measurability and weak measurability are equivalent if X is metric and F admits compact values

Furthermore, in the Definition we may take a closed set $C \subset X$ in the place of open U, just changing the preimages ($F_{-1}(C)$ in the case of measurability and $F^{-1}(C)$ in the case of weak measurability)

Usually, one considers a metric separable space X where some nice properties of measurable functions hold (such as Lusin's Theorem)

Notice that if X is metric separable then the weak measurability of $F : T \rightrightarrows X$ implies that the real function $t \mapsto \mathrm{d}_{F(t)}(x)$ is measurable for each $x \in X$

Let us give some example of a measurable multivalued mapping, which is useful for proving, e.g., Filippov's Lemma

Example

Let X and Y be metric spaces (X is separable), $U \subset Y$ be an open set and $f : T \times X \to Y$ be a single-valued function such that

- $t \mapsto f(t,x)$ is measurable for each $x \in X$
- $x \mapsto f(t,x)$ is continuous for each $t \in T$
  (f T is the space with a measure then in the latter assumption we may require the continuity for a.e. $t \in T$)

Then the mapping $F : T \rightrightarrows X$,

$$F(t) := \{x \in X : f(t,x) \in U\}$$

is **measurable**

The main theorem on measurable selections of a multivalued mapping is the following

Kuratowski and Ryll-Nardžewski Measurable Selection Theorem

Let X be a separable complete metric space. Let also $F : T \rightrightarrows X$ be a weakly measurable multifunction whose values are nonempty and closed. Then F admits a measurable selection $f(t) \in F(t)$.

In fact, there exists a countable family of measurable selections $\{f_n\}$ such that

$$F(t) = \overline{\{f_n(t) : n \geq 1\}}$$

(this is so called Castaing representation)

Now as a consequence of the above theorem we obtain

Filippov's Lemma on implicit functions

Let X and Y be some metric spaces (X is separable and complete). Let also a (single-valued) function $f : T \times X \to Y$ be such that

- $t \mapsto f(t, x)$ is measurable for each $x \in X$
- $x \mapsto f(t, x)$ is continuous for each $t \in T$

and $U : T \rightrightarrows X$ be a measurable multifunction with compact values. If, furthermore, $v : T \to X$ is some measurable function satisfying

$$v(t) \in f(t, U(t)) := \{f(t, u) : u \in U(t)\}, \; t \in T,$$

then one can choose a measurable selection $u(t) \in U(t)$ such that

$$v(t) = f(t, u(t)), \; t \in T$$

Let now T be a space with some nonatomic complete measure μ (we may consider the segment $T = [0, 1]$ with Lebesgue measure denoted by μ_0 or dt). Then we can define the integral of a multivalued mapping

Definition

Assume that the multifunction $F : T \to 2^X$ is weakly measurable and integrably bounded, i.e., there is a nonnegative summable function $l(\cdot) \in L^1(T, X)$ such that $\sup\limits_{v \in F(t)} \|v\| \leq l(t)$ for a.e. $t \in T$. Then the Aumann integral of the mapping F on T is defined as

$$\left\{ \int_T f(t)\, dt : f(\cdot) \ \text{ is a measurable selection of } \ F(\cdot) \ \text{ on } \ T \right\}$$

By the measurable selections theorem the Aumann integral denoted further by

$$\int_T F(t)\,dt$$

is always a nonempty bounded set

It turns out that in the case $X = \mathbb{R}^n$ the Aumann integral is closed, therefore compact subset of $\mathbb{R}^n$

This follows from Dunford-Pettis Theorem if F admits convex values

Otherwise, it is a consequence of the famous A.A.Lyapunov's Theorem

Outline

A.A.Lyapunov's Theorem on the range of vector measure

Let $f : T \to \mathbb{R}^n$ be an integrable function ($f(\cdot) \in L^1(T, \mathbb{R}^n)$)

Then the set

$$\Sigma := \left\{ \int_E f(t)\, dt : E \in \mathfrak{M} \right\}$$

is **convex and compact** in $\mathbb{R}^n$. More precisely, $\Sigma = \int_T F(t)\, dt$ where $F : T \rightrightarrows \mathbb{R}^n$ is the measurable multifunction,

$$F(t) := \{\lambda f(t) : 0 \le \lambda \le 1\}, \qquad t \in T$$

Here $\mathfrak{M}$ is the σ-algebra of Lebesgue measurable subsets of $T = [0, 1]$

The integral $\int\limits_T F(t)\,dt$ is convex set because the multifunction $F : T \rightrightarrows \mathbb{R}^n$ admits convex values. It is also compact by Dunford-Pettis Theorem as was said above. So, we should prove just the equality

$$\Sigma = \int\limits_T F(t)\,dt$$

We represent

$$\int\limits_T F(t)\,dt = \left\{ \int\limits_T \alpha(t) f(t)\,dt : \alpha(\cdot) \in W \right\} = Q(W)$$

Here

$$W := \{\alpha(\cdot) \in L^\infty(T, \mathbb{R}) : 0 \leq \alpha(t) \leq 1 \text{ for a.e. } t \in T\}$$

and $Q : L^\infty(T, \mathbb{R}) \to \mathbb{R}$ is the linear continuous functional,

$$Q : \alpha(\cdot) \mapsto \int_T \alpha(t) f(t)\, dt$$

It is obvious that $\Sigma \subset Q(W)$

In order to prove the opposite inclusion let us fix $x \in Q(W)$ and consider the preimage

$$W_x := W \cap Q^{-1}(x)$$

The set W_x is nonempty and **weakly compact** in the space $L^\infty(T,\mathbb{R}) = (L^1(T,\mathbb{R}))'$ (by Banach-Alaoglu Theorem)

Applying Krein-Milman Theorem we find $\alpha_0(\cdot) \in \operatorname{ext} W_x$ and prove that $\alpha_0(\cdot)$ can admit only values 0 and 1 a.e.

Assuming the contrary, we find $\varepsilon > 0$ such that $\mu(\Delta_\varepsilon) > 0$ where

$$\Delta_\varepsilon := \{t \in T : \varepsilon \le \alpha_0(t) \le 1 - \varepsilon\}$$

Represent the function $f(\cdot)$ as

$$f(t) = (f_1(t), f_2(t), ..., f_n(t))$$

where all $f_i(\cdot)$ are summable on T and, consequently, on Δ_ε

Denote by Λ the linear (finite dimensional) subspace of $L^1(\Delta_\varepsilon, \mathbb{R})$ generated by the restrictions of $f_i(\cdot),\ i = 1, 2, ..., n$, onto Δ_ε

Since Λ is closed and different from the whole space $L^1(\Delta_\varepsilon, \mathbb{R})$, applying Hahn-Banach Theorem we find a linear continuous nontrivial functional on $L^1(\Delta_\varepsilon, \mathbb{R})$, which is equal to zero on Λ

This functional can be identified with some function $\beta(\cdot) \in L^\infty(\Delta_\varepsilon, \mathbb{R})$ such that

$$\|\beta(\cdot)\| = 1 \ \text{ and } \ \int_{\Delta_\varepsilon} \beta(t) f(t)\, dt = 0$$

Extend the function $\beta(\cdot)$ out of Δ_ε by setting $\beta(t) = 0,\ t \in T \setminus \Delta_\varepsilon$, and observe that

- $0 \leq \alpha_0(t) \pm \varepsilon\beta(t) \leq 1$ for all $t \in T$
- $Q(\alpha_0 \pm \varepsilon\beta) = x$

Thus, $\alpha_0 \pm \varepsilon\beta \in W_x$ and $\alpha_0 = 1/2(\alpha_0 + \varepsilon\beta) + 1/2(\alpha_0 - \varepsilon\beta)$ contradicting the choice of α_0

By using A.A.Lyapunov's Theorem we can easily prove the **convexity of the integral**

$$\int_T F(t)\, dt$$

If $f(\cdot)$ and $g(\cdot)$ are measurable selections of $F(\cdot)$ on T and $0 \leq \lambda \leq 1$ then we should find another measurable selection $\varphi(\cdot)$ such that

$$\lambda \int_T f(t)\, dt + (1-\lambda) \int_T g(t)\, dt = \int_T \varphi(t)\, dt$$

By A.A.Lyapunov's Theorem there exists a measurable set $E \in \mathfrak{M}$ such that

$$\lambda \int_T (f(t) - g(t))\, dt = \int_E (f(t) - g(t))\, dt$$

Finally, we set

$$\varphi(t) = f(t)\chi_E(t) + g(t)\chi_{T\setminus E}(t) \in F(t)$$

where $\chi_E(\cdot)$ is the characteristic function of the set E

In fact, in a finite dimensional space a more precise result than the convexity can be proved

Aumann Theorem

If $F : T \to \operatorname{comp} \mathbb{R}^n$ is measurable and integrably bounded then $t \mapsto \operatorname{co} F(t)$ is also measurable and integrably bounded (with compact values) and

$$\int_T \operatorname{co} F(t)\, dt = \int_T F(t)\, dt$$

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a lower semicontinuous function satisfying the superlinear growth assumption:

$$f(\xi) \geq \Phi(\|\xi\|)$$

where $\Phi : \mathbb{R}^+ \to \mathbb{R}^+$ is a real function such that

$$\lim_{r\to+\infty} \frac{\Phi(r)}{r} = +\infty$$

Consider the variational problem

$$\text{Minimize}\left\{\int_T f(x'(t))\,dt : x(\cdot)\in\mathcal{W}\right\}$$

where

$$\mathcal{W} := \{x(\cdot)\in \mathbf{AC}(T,\mathbb{R}^n) : x(0)=x_0,\ x(1)=x_1\}$$

One of the approaches for resolving this problem is to **relax** it, namely, to reduce it to the convexified one

$$\text{Minimize}\left\{\int_T f^{**}(x'(t))\,dt : x(\cdot)\in\mathcal{W}\right\}$$

Let $\bar{x}(\cdot)$ **be a minimizer in this relaxed problem** existing by the famous Tonelli's Theorem. Then we represent

$$f^{**}(\bar{x}'(t))$$

through the **infimum** (see the properties of double conjugations above)

More exactly, we represent the point $g(t) := (\bar{x}'(t), f^{**}(\bar{x}'(t))$ through the extremal points of the (bounded) n-dimensional face of $\operatorname{epi} f^{**}$, which contains $g(t)$

Boundedness follows from the growth assumption

Applying then the Kuratowski and Ryll-Nardzevski Theorem we find measurable functions $v_k : T \to \mathbb{R}^n$ and $\lambda_k : T \to [0,1]$, $k = 1, 2, ..., n+1$, such that

$$\sum_{k=1}^{n+1} \lambda_k(t) = 1$$

$$\begin{aligned} \bar{x}'(t) &= \sum_{k=1}^{n+1} \lambda_k(t) v_k(t) \\ f^{**}(\bar{x}'(t)) &= \sum_{k=1}^{n+1} \lambda_k(t) f(v_k(t)) \end{aligned}$$

In particular, if $n = 1$ then we find two measurable functions $v_1, v_2 : T \to \mathbb{R}$ and measurable $\lambda : T \to [0, 1]$ such that

$$\begin{aligned} \bar{x}'(t) &= \lambda(t)v_1(t) + (1 - \lambda(t))v_2(t) \\ f^{**}(\bar{x}'(t)) &= \lambda(t)f(v_1(t)) + (1 - \lambda(t))f(v_2(t)) \end{aligned}$$

By **A.A.Lyapunov's theorem**, considering the measurable function

$$t \mapsto g(t) := \begin{pmatrix} v_1(t) - v_2(t) \\ f(v_1(t)) - f(v_2(t)) \end{pmatrix}$$

we find a measurable set $E \in \mathfrak{M}$ such that

$$\int_T \lambda(t) g(t)\, dt = \int_E g(t)\, dt$$

After setting

$$v(t) := v_1(t)\chi_E(t) + v_2(t)\chi_{T\setminus E}(t)$$

it is easy to see that the function $x(\cdot)$,

$$x(t) := x_0 + \int_{t_0}^{t} v(s)\, ds, \qquad t \in T,$$

belongs to $\mathcal{W}$ and minimizes the original functional

Now we deduce from A.A.Lyapunov's Theorem an important property of a vector measure, which will be useful in sequel

Let us consider a vector measure $\mu : \mathfrak{M} \to \mathbb{R}^n$ represented by

$$\mu(E) = \int_E f(t)\, dt, \quad E \in \mathfrak{M}$$

with some $f(\cdot) \in L^1(T, \mathbb{R}^n)$

Its norm (in the space of measures $\mathcal{M}(\mathfrak{M}, \mathbb{R}^n)$) or so called total variation

$$\|\mu\|_M := \int_T |f(t)|\, dt$$

Darboux property

There exist measurable sets $A_\alpha \in \mathfrak{M}$, $\alpha \in [0,1]$, such that

- $A_\alpha \subset A_\beta$ if $\alpha \leq \beta$
- $\mu(A_\alpha) = \alpha\mu(T)$

First, given $A \in \mathfrak{M}$ and applying A.A.Lyapunov's Theorem to the σ-algebra $\mathfrak{M}_A := \{A \cap E : E \in \mathfrak{M}\}$ (taking into account that $\mu(\emptyset) = 0$) we find $B \in \mathfrak{M}$ such that

$$\mu(B) = \frac{1}{2}\mu(A)$$

Let us prove the statement for the binary fractions $\frac{k}{2^n}$, $k = 1, 2, ..., 2^n$; $n = 1, 2, ...$ (denote by D the set of these fractions)

For $n = 1$ and $k = 0, 1, 2$ such sets are already constructed above. By induction let us assume that $A_{\frac{k}{2^n}}$ with the above properties are found for all $1 \le k \le 2^n$

So, we should construct sets $A_{\frac{k}{2^{n+1}}}$ just for odd numbers $1 \le k \le 2^{n+1}$

Define $A := A_{\frac{k+1}{2^{n+1}}} \setminus A_{\frac{k-1}{2^{n+1}}}$ and find $B \in \mathfrak{M}$ with

$$\mu(B) = \frac{1}{2}\mu(A) = \frac{1}{2}\left(\mu\left(A_{\frac{k+1}{2^{n+1}}}\right) - \mu\left(A_{\frac{k-1}{2^{n+1}}}\right)\right)$$
$$= \frac{1}{2}\left(\frac{k+1}{2^{n+1}} - \frac{k-1}{2^{n+1}}\right)\mu(T) = \frac{1}{2^{n+1}}\mu(T)$$

We set

$$A_{\frac{k}{2^{n+1}}} := B \cup A_{\frac{k-1}{2^{n+1}}}$$

and see that

- monotonicity property continues to hold
- $\mu\left(A_{\frac{k}{2^{n+1}}}\right) = \frac{k}{2^{n+1}}\mu(T)$

Now given $\alpha \in [0,1]$ let us define

$$A_\alpha = \bigcup_{r\in D,\, r\leq\alpha} A_r$$

Obviously $A_\alpha \in \mathfrak{M}$ and the family $\{A_\alpha\}$ is increasing

Taking an arbitrary increasing sequence $\{r_n\} \subset D$ with $r_n \to \alpha-$ we have

$$\mu(A_\alpha) = \lim_{n\to\infty} \mu\left(A_{r_n}\right) = \lim_{n\to\infty} r_n \mu(T) = \alpha\mu(T)$$

We'll use the following nice consequence of the Darboux property of the **Lebesgue measure** μ_0:

If the functions $f : K \to L_1(T, X)$ and $z : K \to [0, 1]$ are continuous at a point $x_0 \in K$ (here K is an arbitary metric or even topological space, X is a Banach space), then the function $g : x \mapsto f(x)\chi_{A_{z(x)}}$ is continuous at x_0 as well

We deduce continuity of $g(\cdot)$ at x_0 from the following estimates:

$$\begin{aligned} \|g(x) - g(x_0)\| &\leq \|f(x)\chi_{A_{z(x)}} - f(x_0)\chi_{A_{z(x)}}\| \\ &+ \|f(x_0)\chi_{A_{z(x)}} - f(x_0)\chi_{A_{z(x_0)}}\| \leq \|f(x) - f(x_0)\| \\ &+ \int_{A_{z(x)}\Delta A_{z(x_0)}} |f(x_0)(t)|\,dt \end{aligned}$$

Then we use continuity of $f(\cdot)$, the Lebesgue integrability of the function $t \mapsto |f(x_0)(t)|$ and the obvious equality

$$\mu_0 \left(A_{z(x)} \Delta A_{z(x_0)}\right) = |z(x) - z(x_0)|$$

Finally, we need the following

Parametrized Darboux property

Let $x \mapsto \mu_x$ be a continuous mapping from a compact metric space K to the space of measures $\mathcal{M}(\mathfrak{M}, \mathbb{R}^n)$
Then for any $\varepsilon > 0$ there exists a family of measurable sets $\{A_\alpha\} \subset \mathfrak{M}$ satisfying the Darboux property above for the Lebesgue measure μ_0 (or for an a priori given finite dimensional measure) such that

- $|\mu_x(A_\alpha) - \alpha\mu_x(T)| \leq \varepsilon$

for all $\alpha \in [0, 1]$ and all $x \in K$

It is obtained by using the compactness argument (recall that the space K is compact)

Speaking about A.A.Lyapunov's and Aumann's Theorems we already considered the measurable selections, which are constructed starting from given ones by using the **"concatenation" procedure**, i.e., a new function $v(\cdot)$ is obtained as equal to $v_1(\cdot)$ on a measurable subset, and to $v_2(\cdot)$ on its complement

Definition

We say that a set $\mathcal{U} \subset L^1(T, X)$ is **decomposable** if for each functions $u(\cdot)$ and $v(\cdot)$ from $\mathcal{U}$ we have also

$$u\chi_E + v\chi_{T\setminus E} \in \mathcal{U}$$

whenever $E \in \mathfrak{M}$

It is known that a set $\mathcal{U} \subset L^1(T,X)$ is decomposable iff it is the set of measurable selections of some measurable integrably bounded multivalued mapping

It turns out that the decomposability property can substitute the convexity in Michael's selections Theorem

Fryszkowski's Theorem

Let K be a **compact metric space** and X be a separable Banach one. Then each lower semicontinuous multifunction $F : K \rightrightarrows L^1(T,X)$ with nonempty closed decomposable values (we say that the mapping $F(\cdot)$ is decomposable) admits a continuous selection (passing through an arbitrary point of the graph)

This Theorem was proved by A.Fryszkowski in 1983 and afterwards was extended first by A.Bressan and G.Colombo to the case of an arbitrary metric space K (in 1986) and then by S.Ageev and D.Repovs to the paracompact case (2000)

Sketch of the proof

The first step is construction of an approximate multivalued mapping. Namely, given $\varepsilon > 0$ we want to find two continuous (single-valued) mappings $f : K \to L^1(T, X)$ and $r : K \to L^1(T, \mathbb{R}^+)$ such that for all $x \in K$

- $\int_T r(x)(t)\, dt \le \varepsilon$
- $F(x) \cap \left\{ u(\cdot) \in L^1(T, X) : \|u(t) - f(x)(t)\| < r(x)(t) \text{ a.e. on } T \right\} \neq \emptyset$

Compare with the respective construction in Michael's Theorem

By using lower semicontinuity of $F(\cdot)$ (by Vietoris) and Egorov's and Lusin's Theorems on measurable functions we prove that this intersection is lower semicontinuous as well. Consequently, its closure (in $L^1(T, X)$) is lower semicontinuous and admits closed decomposable values

Thus, we construct a sequence of lower semicontinuous multifunctions $F_n(\cdot)$ with nonempty closed decomposable values such that for all $x \in K$

- $F_1(x) = F(x)$
- $F_1(x) \supset F_2(x) \supset ... \supset F_n(x) \supset ...$
- $F_{n+1}(x) = \overline{\{u(\cdot) \in F_n(x) : \|u(t) - f_n(x)(t)\| < r_n(x)(t) \text{ a.e. on} T\}}$ where $\{f_n\}$ is a sequence of continuous functions $K \to L^1(T, X)$ and $\{r_n\}$ is a sequence of continuous functions $K \to L^1(T, \mathbb{R}^+)$ with

$$\int_T r_n(x)(t)\, dt \le \frac{1}{2^n}$$

Here the overbar means the closure in $L^1(T, X)$

Fix now $x \in K$ and take an arbitrary measurable function $u_n^x(\cdot) \in F_n(x)$

Then for any $m > n$ due to motonicity we have $u_m^x(\cdot) \in F_{n+1}(x)$ and

$$\|u_m^x(t) - f_n(x)(t)\| \leq r_n(x)(t) \text{ for a.e. } t \in T$$

On the other hand.

$$\|u_m^x(t) - f_{m-1}(x)(t)\| \leq r_{m-1}(x)(t) \text{ for a.e. } t \in T$$

Adding these two inequalities and integrating on T we arrive at

$$\int_T \|f_n(x)(t) - f_{m-1}(x)(t)\| \leq \frac{1}{2^n} + \frac{1}{2^{m-1}} \qquad (*)$$

Since the space $L^1(T,X)$ is complete, the sequence $\{f_n(x)\}$ converges (for a fixed $x \in K$) to some (integrable) function $f(x) : T \to X$

Passing to the limit as $m \to \infty$ in (*), we have

$$\int_T \|f_n(x)(t) - f(x)(t)\| \leq \frac{1}{2^n}, \quad x \in K,$$

and, hence, the convergence is uniform, implying that $f : K \to L^1(T,X)$ is continuous

Finally, $f(x) = \lim_{n\to\infty} u_x^n(\cdot) \in F_0(x) = F(x), \ x \in K$, due to the closedness of $F(x)$

Thus, in order to conclude the proof we should find continuous functions $f : K \to L^1(T,X)$ and $r : K \to L^1(T,\mathbb{R}^+)$ with the approximate property above

Let us recall that given a closed set $\mathcal{A} \subset L^1(T,\mathbb{R})$ its essential infimum is defined as the infimum of $\mathcal{A}$ in the **lattice** of the real measurable functions (defined on T) with the partial order

$$a(\cdot) \le b(\cdot) \text{ iff } a(t) \le b(t) \text{ for a.e. } t \in T$$

i.e., $\operatorname{ess\,inf} \mathcal{A}$ is a measurable function $a_0(\cdot)$ such that

- $a_0(t) \le a(t)$ for a.e. $t \in T$ whenever $a(\cdot) \in \mathcal{A}$
- if for some measurable function $b(\cdot) \in \mathcal{A}$ we have $b(t) \le a(t)$ for a.e. $t \in T$ whenever $a(\cdot) \in \mathcal{A}$ then $b(t) \le a_0(t)$ a.e. on T

It is well known that $\operatorname{ess\,inf}\mathcal{A}$ always exists, unique (up to changes on sets of null measure) and

$$\operatorname{ess\,inf}\mathcal{A} = \inf_n a_n(t), \quad t \in T$$

for some sequence $\{a_n(\cdot)\} \subset \mathcal{A}$

If all functions from $\mathcal{A}$ are nonnegative then due to the closedness of $\mathcal{A}$ by Lebesgue Dominated Convergence Theorem we deduce that $\operatorname{ess\,inf}\mathcal{A} \in \mathcal{A}$ whereas some other properties follow from the decomposability of $\mathcal{A}$

Given a decomposable set $\mathcal{U} \subset L^1(T, X)$ (X is a separable Banach space) let us denote by

$$\psi(t) := \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}} \|u(t)\|$$

i.e., $\psi(\cdot)$ is the essential infimum of the set $\mathcal{A} := \{\|u(\cdot)\| : u(\cdot) \in \mathcal{U}\}$ as defined above

Then there exists a sequence $\{u_n(\cdot)\} \subset \mathcal{U}$ decreasing in norm, i.e., such that

$$\|u_1(t)\| \geq \|u_2(t)\| \geq ... \geq \|u_n(t)\| \geq ... \text{ for a.e. } t \in T$$

and

$$\psi(t) := \lim_{n\to\infty} \|u_n(t)\|, \quad t \in T$$

In construction of $\{u_n(\cdot)\}$ we strongly use decomposability of the set $\mathcal{U}$

8 Properties of the essential infimum

Another useful property of a decomposable set $\mathcal{U}$ is the following

$$\psi(t) = \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}} \|u(t)\| = \|u^*(t)\|$$

for some function $u^*(\cdot) \in \mathcal{U}$

Let us consider now a lower semicontinuous mapping $F : K \to L^1(T, X)$ with closed decomposable values and set $\psi_x(t) := \operatorname*{ess\,inf}_{u(\cdot)\in F(x)} |u(t)|$

Theorem

The mapping $G : K \to L^1(T, \mathbb{R})$,

$$G(x) := \{v(\cdot) \in L^1(T, \mathbb{R}) : v(t) \geq \psi_x(t) \text{ for a.e. } t \in T\}$$

admits closed and **convex** values and is (Vietoris) lower semicontinuous

Proof

To see this let us take $x_0 \in K$, a sequence $\{x_n\} \subset K$ converging to x and $v_0 \in G(x_0)$

Then there exists $u_0(\cdot) \in F(x_0)$ with

$$v_0(t) \geq \|u_0(t)\| = \psi_{x_0}(t) \text{ for a.e. } t \in T$$

Due to l.s.c. of $F(x)$ there exists a sequence $u_n(\cdot) \in F(x_n)$ converging to $u_0(\cdot)$ in $L^1(T, X)$

Therefore the sequence $v_n := \|u_n\| - \|u_0\| + v_0$ converges to v_0 and

$$v_n(t) \geq \|u_n(t)\| \geq \psi_{x_n}(t) \text{ for a.e. } t \in T$$

So, now we are ready to construct an approximate mapping for $F : K \to L^1(T, X)$

Given $x_0 \in K$ and $u_0 \in F(x_0)$ let us denote by G_{u_0} the multivalued mapping from Theorem above where we substitute $F - u_0$ in the place of F

Then by l.s.c. of G_{u_0} applying Michael's selections Theorem we find a continuous real function $\varphi_{(x_0,u_0)}$ defined on K such that

- $\varphi_{(x_0,u_0)}(x)(t) \geq \operatorname*{ess\,inf}_{u \in F(x)} \|u(t) - u_0(t)\|$ a.e. on T for all $x \in K$
- $\varphi_{(x_0,u_0)}(x_0) = 0$

Fixed $\varepsilon > 0$ consider the open neighbourhood of x_0:

$$V_{(x_0,u_0)} := \left\{ x : \int_T \varphi_{(x_0,u_0)}(x)(t)\, dt < \varepsilon/4 \right\}$$

and choose by the compactness of K a finite number of the sets $V_i := V_{(x_i,u_i)}$ for some $x_i \in K,\ u_i \in F(x_i)$, $i = 1, ..., p$, such that

$$K = \bigcup_{i=1}^{p} V_i$$

Let $\{e_i(\cdot)\}$ be a continuous partition of unity corresponding to $\{V_i\}$

Then we consider the parametrized vector measure with the density (depending on $x \in K$) given by

$$(\varphi_1(x)(\cdot), ..., \varphi_p(x)(\cdot))$$

where $\varphi_i(x) := \varphi_{(x_i,u_i)}(x)$

For this measure we construct a family of measurable sets $\{A_\alpha\}_{\alpha\in[0,1]}$ with the properties

- $A_\alpha \subset A_\beta$ whenever $\alpha \leq \beta$
- $\mu_0(A_\alpha) = \alpha\mu_0(T)$ (μ_0 is the Lebesgue measure on T)
- $\left|\int\limits_{A_\alpha} \varphi_i(x)(t)\,dt - \alpha\int\limits_T \varphi_i(x)(t)\,dt\right| \leq \frac{\varepsilon}{4p},\ \ x \in K,\ i = 1, ..., p$

Introduce now the continuous functions $z_i : K \to [0,1]$,

$$z_i(x) = e_1(x) + e_2(x) + ... + e_i(x), \quad i = 1, ..., p,$$

and define

$$f(x)(t) := \sum_{i=1}^{p} u_i \chi_{A_{z_i(x)} \setminus A_{z_{i-1}(x)}}(t)$$

$$r(x)(t) := \sum_{i=1}^{p} \left(\varphi_i(x) + \frac{\varepsilon}{4} \right) \chi_{A_{z_i(x)} \setminus A_{z_{i-1}(x)}}(t)$$

The functions $f : K \to L^1(T, X)$ and $r : K \to L^1(T, \mathbb{R})$ are continuous as shown above

The inequality

$$\int_T r(x)(t)\,dt \le \varepsilon$$

follows easily from the last property of the family $\{A_\alpha\}$

Finally, by the properties of $\operatorname{ess\,inf}$ (see above) for each $u_i(\cdot)$, $i = 1, 2, .., p$ (recall that $u_i \in F(x_i)$) and each $x \in K$ there exists a function $u_x^i(\cdot) \in F(x)$ with

$$\|u_x^i(t) - u_i(t)\| = \operatorname*{ess\,inf}_{u(\cdot)\in F(x)} \|u(t) - u_i(t)\| \text{ for a.e. } t \in T$$

By decomposability the function

$$u_x(t) := \sum_{i=1}^{p} u_x^i \chi_{A_{z_i(x)} \setminus A_{z_{i-1}(x)}}(t)$$

belongs to $F(x),\ x \in K$

From the properties of the family $\{A_\alpha\}$ we easily deduce that

$$\|u_x(t) - f(x)(t)\| < r(x)(t) \text{ for a.e. } t \in T$$

and, hence,

$$\{u(\cdot) \in F(x) : \|u(t) - f(x)(t)\| < r(x)(t) \text{ a.e. on } T\} \neq \emptyset$$

for all $x \in K$, and everything is proved

Outline

Let us consider the segment $T = [0, 1]$ endowed with the Lebesgue measure dt and the σ-algebra $\mathfrak{M}$ of Lebesgue measurable sets. Assume that $F : T \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a multifunction with nonempty closed values

Definition

The expression

$$\dot{x} \in F(t, x) \qquad \text{(DI)}$$

is said to be Differential Inclusion

An absolutely continuous function $x : T \to \mathbb{R}^n$, which satisfies the relation $\dot{x}(t) \in F(t, x(t))$ for a.e. $t \in T$ is said to be **Carathéodory type solution** of the inclusion (DI)

1 Basic definitions. Cauchy problem

In what follows we will consider also DI with the right-hand side defined on an infitite dimensional Banach space X in the place of $\mathbb{R}^n$

Notice that in this case the definition of (Carathéodory type) solution should be changed

Indeed, we should require that to be solution an absolutely continuous function $x : T \to X$ must have derivative at almost each point $t \in T$, which must be Lebesgue integrable on T (then from absolute continuity one conclude also that the function $x(\cdot)$ can be recovered from its derivative $\dot{x}(\cdot)$ by using the Newton-Leibnitz formula)

However, the above property (existence a.e. of the integrable derivative) holds for an arbitrary absolutely continuous function not only in $\mathbb{R}^n$ but in very large class of Banach spaces (so called spaces with Radon-Nikodym property), which includes all reflexive spaces (in particular, all Hilbert ones)

2 Existence theorems. Convex l.s.c. case

If F admits **convex** values and is **lower semicontinuous** w.r.t. both variables (t, x) then given point (x_0, v_0) with $v_0 \in F(0, x_0)$ by Michael's Theorem we can find a continuous selection $f(t, x) \in F(t, x)$ such that $f(0, x_0) = v_0$

Applying now Peano's Theorem we find a classic (i.e., continuously differentiable) solution of the Cauchy problem

$$\dot{x} = f(t, x), \qquad x(0) = x_0$$

Then $x(\cdot)$ is a **classic solution of the Differential Inclusion**, satisfying the conditions $x(0) = x_0$ and $\dot{x}(0) = v_0$ (with fixed initial state and initial velocity)

All methods for proving existence of a solution can be divided in two kinds:

- direct methods
- indirect methods

Direct methods are those where a solution is obtained as limit of a sequence of some good functions, which can be treated as approximative solutions with some *a priori* given exactness

Indirect methods instead refer to some either analitical or topological results in order to establish existence of a solution such as Schauder fixed point Theorem or Contractions Principle

To illustrate this let us recall the basic existence theorems for ODE (Peano and Picard-Lindelöff Theorems), which can be proved by using both direct and indirect methods

In turn the direct methods can be distinguished according with the fundamental principles, which are on the base of these methods

Indeed, there are two different fundamental priciples in Analysis, and any eistence theorem is somehow based on one of them:

- Compactness Principle when a solution is an eventual limit (accumulation) point of some approximate sequence, which may be not unique
- Completeness Principle when a solution is the (unique) limit (!) of the convergent Cauchy sequence

For instance, considering the Cauchy problem for ODE

$$\dot{x} \in f(t,x), \;\; x(0) = x_0$$

we have two possibilities:

(i) assuming continuity of f in x (and measurability w.r.t. t + integrable boundedness), we prove existence of a solution (not unique), constructing so called Euler polygons (or other approximations) and applying to them compactness argument due to Arzelà-Askoli Theorem

(ii) assuming lipschitzeanity of f in x (and measurability w.r.t. t as well), we constract a sequence of successive approimations, which converges in a complete space (of continuous functions) being a Cauchy sequence there

All of these methods are applicable in the case of Differential Inclusions as well with some modifications due to multivaluedness of the right-hand side

The question, which method can be used, depends on the hypotheses imposed to the right-hand side

As about indirect methods, besides the Fixed point Theorems, which are relevant also here, we will use the **new tool** concerned with continuous selections and also a topological tool based on famous Baire Category Theorem

4 Euler polygons. Compactness argument

This method can be applied whenever the right-hand side F is upper semicontinuous in x and admits convex compact values

For the sake of simplicity we consider only the autonomous case (when F depends only on x and is upper semicontinuous). In fact, a solution exists also in general case when F depends on t in a measurable way (and it is integrably bounded) but in this case the algorithm of proving should be a bit more complicate

Theorem

Let us assume that $F : \mathbb{R}^n \to \mathrm{conv}(\mathbb{R}^n)$ is upper semicontinuous and bounded, i.e.,

$$\sup_{v \in F(x)} |v| \leq R \tag{B}$$

for all $x \in \mathbb{R}^n$ with some constant $M > 0$. Then given $x_0 \in \mathbb{R}^n$ there exists a Carathéodory solution $x(\cdot)$ of DI such that $x(0) = x_0$

Observe that without the boundedness hypothesis (B) we can also prove existence but just of a local solution (defined on some interval $[0,\delta],\ 0<\delta<1$)

Sketch of the proof

Fix $n=1,2,\ldots$ and divide the segment $T=[0,1]$ into n parts by points

$$t_i^n := i/n,\ i=1,2,\ldots,n$$

First, we set $x_0^n = x_0$, choose $v_0^n \in F(0,x_0^n)$ and define $x_1^n := x_0^n + \frac{1}{n}v_0^n$

Then, by induction in i construct two finite sequences $\{x_i^n\}$ and $\{v_i^n\}$ such that

$$x_{i+1}^n = x_i^n + \frac{1}{n}v_i^n \quad \text{and} \quad v_i^n \in F(t_i^n, x_i^n),$$

$i=0,1,\ldots,n-1$

Define continuous (polygonal) function

$$x_n(t) := x_i^n + (t - t_i^n)\, v_i^n, \quad t \in [t_i^n, t_{i+1}^n]$$

We have

$$\dot{x}_n(t) \in F\left(t_i^n, x_n(t_i^n)\right), \quad t \in]t_i^n, t_{i+1}^n[\qquad (*)$$

4 Euler polygons. Compactness argument

The sequence $\{x_n(\cdot)\}$ is uniformly bounded and equicontinuous on T $(|\dot{x}_n(t)| \leq R, \ \ t \in T)$

By Arzela-Askoli Theorem $\{x_n(\cdot)\}$ is relatively compact in $C(T, \mathbb{R}^n)$, so without of generality assume that it converges (uniformly) to some continuous function $x(\cdot)$

On the other hand, by Danford-Pettis Theorem the sequence $\{\dot{x}_n(\cdot)\}$ is weakly relatively compact in $L^1(T, \mathbb{R}^n)$. So, without loss of generality assume that $\dot{x}_n(\cdot) \rightharpoonup v(\cdot), \ \ n \to \infty$, with some $v(\cdot) \in L^1(T, \mathbb{R}^n)$

Hence,

$$x_n(t) = x_0 + \int_0^t \dot{x}(\tau)\, d\tau \to x_0 + \int_0^t v(\tau), \ \ n \to \infty, \ \ t \in T$$

Consequently, the function $x(\cdot)$ is absolutely continuous on T and $\dot{x}(t) = v(t)$ for a.e. $t \in T$

Now, let us apply Mazur's Lemma and find a sequence of convex combinations of $\{\dot{x}(\cdot)\}$, converging to $v(\cdot)$ strongly in $L^1(T, \mathbb{R}^n)$

Namely, let

$$v_n(\cdot) := \sum_{k=k_n}^{\infty} \lambda_k^n \dot{x}_k(\cdot) \to \dot{x}(\cdot), \quad n \to \infty$$

Here $0 \leq \lambda_k^n \leq 1$ with $\sum_{k=k_n}^{\infty} \lambda_k^n = 1$ and only finite number of $\lambda_k^n, \; k \geq k_n$, are positive

Without loss of generality we assume that $v_n(t) \to \dot{x}(t), \; t \to \infty$, a.e.

4 Euler polygons. Compactness argument

Denote now by $\mathcal{T}$ the set (of full measure) of all $t \in T$, where the convergence above takes place and where all the inclusions (*) hold

Fix $t \in \mathcal{T}$. Then given $\varepsilon > 0$ by (Hausdorff) upper semicontinuity of F we find $\delta > 0$ such that for each $x \in \mathbb{R}^n$ with $|x - x(t)| \leq \delta$ we have

$$F(x) \subset F(x(t)) + \varepsilon \overline{B}$$

Let us choose $n \geq 1$ so large that

$$|x_n(t) - x_n(t_i^n)| \leq \frac{1}{n} R \leq \delta, \; i = 0, 1, ..., n$$

Hence, taking into account (*) we find

$$\dot{x}(t) \in F(t, x(t)) + \varepsilon \overline{B}$$

Since $\varepsilon > 0$ is arbitrary, we have

$$\dot{x}(t) \in F(t, x(t)),$$

and $x(\cdot)$, $x(0) = x_0$, is a requied solution

5 Variational problem. Tonelli's Theorem

The similar reasoning can be applied, for instance, proving existence of a minimizer in a variational problem

$$\text{Minimize} \quad \left\{ \int_T f(t, \dot{x}(t))\, dt : x(\cdot) \in \mathcal{W} \right\}$$

where $\mathcal{W}$ is the set of all absolutely continuous functions $x : T \to \mathbb{R}^n$ with $x(0) = x_0$ and $x(1) = x_1$

We can consider a minimization problem where some restriction on derivative (in the form of differential inclusion $\dot{x} \in F(t, x)$ assuming that F is u.s.c. with convex compact values) as well as some phase constraints are involved

We assume the following hypotheses:

(i) the mapping $(t, v) \mapsto f(t, v)$ is continuous

(ii) for each $t \in T$ the mapping $v \mapsto f(t, v)$ is convex

(iii) for some $p > 1$ there exist constants $\alpha_1 > 0$, $\alpha_2 \geq 0$, $\beta \in \mathbb{R}$ and an integrable function $\gamma : T \to \mathbb{R}^+$ such that

$$\alpha_1 |v|^p + \beta \leq f(t, v) \leq \alpha_2 |v|^p + \gamma(t)$$

for a.e. $t \in T$ and all v (the first inequality is so called coercivity condition of the rank p)

Sketch of the proof

Let us denote by

$$I := \inf \left\{ \int_T f(t, \dot{x}(t))\, dt : x(\cdot) \in \mathcal{W} \right\} < +\infty$$

Find a minimizing sequence $\{x_n(\cdot)\} \subset \mathcal{W}$ satisfying

$$\int_T f(t, \dot{x}_n(t))\, dt \leq I + \frac{1}{2^n}$$

From the coercivity assumption we have

$$\int_T |\dot{x}_n(t)|^p \, dt \leq \frac{I + 1 - \beta}{\alpha_1}, \quad n \geq 1$$

So, the sequence $\{\dot{x}_n(\cdot)\}$ is bounded in the space $L^p(T, \mathbb{R}^n)$ and, by Alaoglu-Banach Theorem, it is relatively weakly compact

Assume without loss of generality that $\{\dot{x}_n(\cdot)\}$ converges weakly to $v(\cdot) \in L^p(T, \mathbb{R}^n)$

Then for each given $t \in T$ the sequence $\{x_n(t)\}$ converges to

$$x(t) = x_0 + \int_0^t v(\tau) \, d\tau$$

In particular, $x(1) = x_1$ and so $x(\cdot) \in \mathcal{W}$

It remains thus to prove that $x(\cdot)$ is a minimizer

To this end by using again Mazur's Lemma (as in the proof of the existence theorem for DI) we choose a subsequence of convex combinations converging to $v(\cdot)$ strongly (in the space $L^p(T, \mathbb{R}^n)$) and (without loss of generality) almost everywhere on T:

$$v_n(t) := \sum_{k=k_n}^{\infty} \lambda_k^n \dot{x}_k(t) \to v(t) \text{ for all } t \in \mathcal{T}$$

where $\mathcal{T} \subset T$ with $\mu_0(\mathcal{T}) = \mu_0(T)$

Here for each n just finite family of numbers λ_k^n $(k \geq k_n)$ are positive

By using convexity of f in the second variable we have

$$f(t,v_n(t)) \le \sum_{k=k_n}^{\infty} \lambda_k^n f(t,\dot{x}_k(t))$$

Integrating on T and continuing the latter inequality we arrive at:

$$\int_T f(t,v_n(t))\,dt \le \sum_{k=k_n}^{\infty} \lambda_k^n \int_T f(t,\dot{x}_k(t))\,dt \le I + \sum_{k=k_n}^{\infty} \frac{1}{2^k} \le I + \frac{1}{2^{n-1}}$$

On the other hand, due to continuity of f

$$f(t,v_n(t)) \to f(t,\dot{x}(t)),\ n \to \infty$$

for a.e. $t \in T$

Finally, by Lebesgue's Dominated convergence Theorem we arrive at

$$I \leq \int_T f(t, \dot{x}(t))\, dt = \lim_{n\to\infty} \int_T f(t, \dot{x}_n(t))\, dt \leq I$$

We use here the second inequality from the hypothesis (iii)

So, infimum is attended on the function $x(\cdot)$

The method of successive approximations based on the completeness argument is used for Differential Inclusions with (not necessarily convex-valued) lipschitzean right-hand side (compare with Picard-Lindelöff Theorem for ODE) even in an arbitrary Banach space

Theorem

Let us assume that $F : T \times X \to \operatorname{comp} X$ is such that

- $t \mapsto F(t,x)$ is measurable for each $x \in X$
- there exists an integrable function $k(\cdot) \in L^1(T, \mathbb{R}^+)$ such that

$$\mathfrak{D}(F(t,x), F(t,y)) \leq k(t)\|x - y\|$$

for all $x, y \in X$ and a.e. $t \in T$

Let us take more an arbitrary absolutely continuous function $y : T \to X$, $y(0) = y_0$

Then there exists a solution $x(\cdot)$, $x(0) = x_0$, of DI such that

(a) $\|x(t) - y(t)\| \le \xi(t)$ for all $t \in T$

(b) $\|\dot{x}(t) - \dot{y}(t)\| \le k(t)\xi(t) + \rho(t)$ for a.e. $t \in T$

where

$$\xi(t) := \|x_0 - y_0\| \exp m(t) + \int_0^t \rho(s) \exp(m(t) - m(s))\, ds$$

$$m(t) := \int_0^t k(s)\, ds$$

$$\rho(t) := \mathrm{d}_{F(t,y(t))}(\dot{y}(t))$$

Traditionally the estimates (a)-(b) are called Filippov-Gronwall inequalities

Sketch of the proof

We construct a sequence of approximate solutions $x_n(\cdot)$ by some recursive procedure

Let us choose a measurable selection $v_0(t)$ of $t \mapsto F(t, y(t))$ with

$$\|v_0(t) - \dot{y}(t)\| = \rho(t)$$

(projection of the drivative $\dot{y}(\cdot)$ on the set $F(t, y(t))$)

Then define

$$x_1(t) = x_0 + \int_0^t v_0(t)\, dt$$

and choose a measurable selection $v_1(t)$ of the mapping $t \mapsto F(t, x_1(t))$ such that

$$\|v_1(t) - \dot{x}_1(t)\| = \mathrm{d}_{F(t,x_1(t))}(\dot{x}_1(t))$$

Denoting by

$$x_2(t) = x_0 + \int_0^t v_1(t)\,dt$$

we continue this process and find a sequence of measurable functions $\{v_n(\cdot)\}$ and a sequence of absolutely continuous functions $\{x_n(\cdot)\}$ s. t.

(i) $v_n(t) \in F(t, x_n(t))$ for a.e. $t \in T$

(ii) $\|v_n(t) - \dot{x}_n(t)\| = \mathrm{d}_{F(t,x_n(t))}(\dot{x}_n(t))$

(iii) $x_n(t) = x_0 + \int_0^t v_{n-1}(t)\,dt, \quad n = 1, 2, \ldots$

From (ii) by using lipschitzeanity assumption we obtain

$$\begin{aligned}\|\dot{x}_{n+1}(t) - \dot{x}_n(t)\| &\leq \mathfrak{D}(F(t, x_{n-1}(t)), F(t, x_n(t))) \\ &\leq k(t)\|x_n(t) - x_{n-1}(t))\|\end{aligned}$$

By successive integrating of these inequalities (for $n = 1, 2, ...$) and applying the recursive procedure we obtain the estimates for $\|x_{n+1}(t) - x_n(t)\|$, which will imply that $\{x_n(\cdot)\}$ is a Cauchy sequence in the complete space $\mathbf{C}(T, X)$ and that the sequence of derivatives is a Cauchy sequence in the space $L^1(T, X)$

So they converge to a solution $x(\cdot)$ and to its derivative, respectively

Then, passing to limits in respective inequalities, we easily obtain the estimates (a) and (b)

This is an indirect approach, which allows by using Fryzkowski selections Theorem to avoid the convexity assumption in the case when the right-hand side is just l.s.c. in x

Theorem

Let us assume the mapping $F : T \times \mathbb{R}^n \to \mathrm{comp}(\mathbb{R}^n)$ to be such that

- $x \mapsto F(t, x)$ is lower semicontinuous for a.e. $t \in T$
- F is superpositionally measurable
- $\|v\| \leq l(t)$ for all $v \in F(t, x)$, $x \in X$ and a.e. $t \in T$

Then for each $x_0 \in X$ there exists a solution $x(\cdot)$, $x(0) = x_0$, of DI

In order to obtain the superposional measurability one usually assumes F to be jointly measurable in both variables t and x w.r.t. the σ-algebra $\mathfrak{M} \otimes \mathfrak{B}$ (here $\mathfrak{B}$ is the σ-algebra of Borel subsets of X)

Let us denote by

$$\mathcal{K} := \{x(\cdot) \in \mathbf{AC}(T, \mathbb{R}^n) : \|\dot{x}(t)\| \le l(t) \;\text{ for all }\; t \in T\},$$

which is relatively compact in $\mathbf{C}(T, \mathbb{R}^n)$ by Arzela-Askoli Theorem and also closed, consequently, compact

Let us consider the so called (multivalued) **Nemytsky operator** $\mathfrak{F} : \mathcal{K} \to L^1(T, \mathbb{R}^n)$,

$$\mathfrak{F}(x) := \{v(\cdot) \in L^1(T, \mathbb{R}^n) : v(t) \in F(t, x(t)) \;\text{ for a.e. }\; t \in T\}$$

The values $\mathfrak{F}(x)$ are obviously **decomposable and closed**

It is easy to show that the multivalued mapping $\mathfrak{F}$ is (Vietoris) lower semicontinuous

Applying now Fryzkowski Theorem (on the compact space $\mathcal{K}$) we find a **continuous selection**

$$g(x) \in \mathfrak{F}(x), \qquad x(\cdot) \in \mathcal{K}$$

Let us define the continuous mapping $f : \mathcal{K} \to \mathcal{K}$,

$$f(x)(t) := x_0 + \int_0^t g(x)(s)\, ds, \qquad t \in T$$

Since $\mathcal{K}$ is compact, convex and $f(\cdot)$ is continuous, by Schauder's fixed point Theorem we find a function $x^*(\cdot) \in \mathcal{K}$ such that $x^* = f(x^*)$

Then, obviously, $x^*(0) = f(x^*)(0) = x_0$, and

$$\dot{x}^*(t) = g(x^*)(t) \in F(t, x^*(t)) \qquad \text{for a.e. } \ t \in T$$

So, $x^*(\cdot)$ is a required solution of the Cauchy problem

An interesting approach to search solutions of DI was proposed first by A.Cellina in 1980 and then was developed in works by F.De Blasi, G.Pianigiani, A.Bressan and others

It is based on a topological argument following from the so called Baire category Theorem and allows to find a solution, whose derivatives not just belong to the right-hand side but pass through its extremal points

So, being the sets $F(t,x)$ convex and compact (in $\mathbb{R}^n$), we search a solution $x(\cdot)$, $x(0)=x_0$, of the Differential Inclusion

$$\dot{x}(t) \in \mathrm{ext}\, F(t,x(t))$$

For the sake of simplicity in what follows we will consider only autonomous case (when F does not depend on t)

8 Extremal solutions. Baire category approach

Baire category Theorem

If X is a complete metric space then there is no a nonempty open set $G \subset X$, which can be represented as a countable union of rare (or nowhere dense) sets (usualy one says that each open set is of the second category)

The dual formulation

If X is a complete metric space and $\{U_n\}$ is a sequence of open dense sets in X then the intersection

$$\bigcap_{n=1}^{\infty} U_n$$

is also dense in X (it is a dense G_δ-set)

8 Extremal solutions. Baire category approach

Denote by $\mathcal{H}_F(x_0)$ the set of all solutions to the convex problem, which is nonempty (we already know this) and closed in $\mathbf{C}(T, \mathbb{R}^n)$, so it is a complete metric space

We are going to construct a sequence $\{\mathcal{H}_k\}$ of open dense subsets of $\mathcal{H}_F(x_0)$ such that

$$\bigcap_{k=1}^{\infty} \mathcal{H}_k \subset \mathcal{H}_{\mathrm{ext}\,F}(x_0)$$

Then Baire Category Theorem formulated above gives that $\bigcap_{k=1}^{\infty} \mathcal{H}_k$ is dense in $\mathcal{H}_F(x_0)$ as well, so it is **nonempty**

Open and dense sets $\mathcal{H}_k$ can be constructed by using the so named Choquet functions

To each nonempty compact convex $K \subset \mathbb{R}^n$ one can associate a Choquet function $l(\cdot, K) : \mathbb{R}^n \to \overline{\mathbb{R}}$, which satisfies the following properties:

(i) $l(x, K) = -\infty$ for $x \notin K$, and $0 \le l(x, K) \le \operatorname{diam} K$ for all $x \in K$

(ii) $l(\cdot, K)$ is concave

(iii) $l(\cdot, K)$ is upper semicontinuous. Moreover, it is upper semicontinuous in both variables x and K. In other words, for each sequence $\{K_m\}$ converging to K w.r.t. Hausdorff distance, and for each $\{x_m\}$ converging to x the inequality

$$l(x, K) \ge \limsup_{m \to \infty} l(x_m, K_m)$$

holds

(iv) $l(x, K) = 0$ if and only if $x \in \operatorname{ext} K$

Examples of Choquet function were given by A.Bressan, F.De Blasi and G.Pianigiani also in infinite-dimensional spaces

S.Suslov in 1991 proposed to characterize extreme points via a finite sequence of functions $\{l_i(\cdot, K)\}$ but not via single one (however, his functions have very simple form in difference of usually used Choquet ones)

$$l_i(x, K) := \max \left\{ \langle e_i, y - z \rangle : y, z \in K \text{ and } \frac{y+z}{2} = x \right\}$$

where $\{e_i\}_{i=1}^n$ is orthonormal basis in $\mathbb{R}^n$

Then, in the place of the property (iv) above one should have

- $x \in \operatorname{ext} K$ if and only if $l_i(x, K) = 0$ for all $i = 1, ..., n$

To the Choquet function $l(x,K)$ (or to each of $l_i(x,K)$, $i=1,2,...,n$) one associates first a functional defined on the space of absolutely continuous functions

$$\mathcal{L}(x(\cdot)) := \int_T l(\dot{x}(t), F(x(t)))\, dt$$

which is equal to $-\infty$ out of the solution set $\mathcal{H}_F(x_0)$, and then the family of subsets

$$\mathcal{H}_\eta := \{x(\cdot) \in \mathcal{H}_F(x_0) : \mathcal{L}(x(\cdot)) < \eta\}, \qquad \eta > 0$$

To prove the openess of each $\mathcal{H}_\eta$ it is enough to show upper semicontinuity of $\mathcal{L}(x(\cdot))$ on $\mathbf{C}(T, \mathbb{R}^n)$

But $-\mathcal{L}(x(\cdot))$ is nothing else than the integral functional with the lagrangean, which is convex and lower semicontinuous w.r.t. the derivative, and is lower semicontinuous w.r.t. the state variable

The lower semicontinuity of such functional is one of the elements of proving existence of minimizers in Tonelli Method

In order to prove density we should already use the lipschitzeanity of the right-hand side and resolve, in fact, the following problem:

taking $x(\cdot) \in \mathcal{H}_F(x_0)$, we should find a sequence $\{x_m(\cdot)\} \subset \mathcal{H}_F(x_0)$ such that $x_m(\cdot) \to x(\cdot)$ uniformly in T and $\mathcal{L}(x_m(\cdot)) \to 0$ as $m \to \infty$

By Carathéodory Theorem for a.e. $t \in T$ the derivative $\dot{x}(t)$ can be represented as a convex combination of not more than $n+1$ extreme points; and this representation can be chosen measurable w.r.t. t

In other words, $\dot{x}(t) \in \mathrm{co}\, C(t)$ for some measurable compact-valued mapping $C(t) \subset \mathrm{ext}\, F(x(t))$ $(\mathrm{card}\, C(t) \leq n+1)$

Now by using Aumann Theorem we find a sequence of measurable selections $\{v_m(\cdot)\}$ of the mapping $C(\cdot)$ itself such that $\|x(t) - y_m(t)\| \to 0$ uniformly in T where $y_m(t) := x_0 + \int_0^t v_m(s)\, ds$

The functions $y_m(\cdot)$ could be candidats for approximation of $x(\cdot)$ because their derivatives are extreme points, **but** of the set $F(x(t))$ (not of $F(y_m(t))$), so $y_m(\cdot) \notin \mathcal{H}_F(x_0)$

However, we have the estimate

$$\rho(\dot{y}_m(t), F(y_m(t))) \leq \mathcal{D}(F(x(t)), F(y_m(t))) \leq L\|x(t) - y_m(t)\| \to 0$$

uniformly in $t \in T$. Here we need already lipschitzeanity of F (with a constant $L > 0$)

Hence, for each m there exists a solution $x_m(\cdot)$ of the convexified problem, satisfying Filippov-Gronwall Inequality

$$\|x_m(t) - y_m(t)\| \leq \int_0^t \rho_m(s)\exp(L(t-s))\, ds \to 0$$

uniformly on T where

$$\rho_m(t) := \mathrm{d}_{F(y_m(t))}(\dot{y}_m(t))$$

Thus, on one hand, $x_m(t) \to x(t)$ uniformly in T

On the other hand, fix $t \in T$ such that $\dot{x}_m(t) \in F(x_m(t))$ and $\dot{y}_m(t) \in C(t)$

Since $C(t)$ is compact, without loss of generality assume that $\dot{y}_m(t) \to w \in \operatorname{ext} F(x(t))$

Since also $F(x_m(t)) \to F(x(t))$, due to the upper semicontinuity of the Choquet function we have

$$\limsup_{m\to\infty} l(\dot{x}_m(t), F(x_m(t))) \le l(w, F(x(t))) = 0$$

for a.e. $t \in T$, and the density is proved

Outline

1 Setting of the problem

In general, we consider a multifunction $F : T \times K \rightrightarrows X$ where $X = \mathbb{R}^n$ (or even an arbitrary Banach space) and K is its nonempty closed (or locally closed) subset

Sometimes K is considered depending on t, and so F is defined on the graph of the mapping $t \mapsto K(t)$

Given $x_0 \in K$ we are interested to prove existence of an absolutely continuous function $x(\cdot)$, $x(0) = x_0$, such that

$$x(t) \in K \;\; \forall t \in T$$

and

$$\dot{x}(t) \in F(t, x(t)) \text{ for a.e. } t \in T$$

2 Nagumo's Theorem. Tangential condition

The first result on viability was obtained for ODE by M.Nagumo in 1942

He involved the tangential condition assuming that the right hand side belongs to the contingent cone to a set K (not necessarily convex)

Nagumo's Theorem

Let $K \subset \mathbb{R}^n$ be a nonempty closed set and $f : K \to \mathbb{R}^n$ be continuous satisfying the following condition

$$f(x) \in T_K^b(x) \quad \forall x \in K$$

Then for each $x_0 \in K$ there exists a solution $x(\cdot)$, $x(0) = x_0$, of the differential equation

$$\dot{x} = f(x),$$

certainly, with $x(t) \in K, \; t \in T$

As about Differential Inclusions, the first result on viability was obtained by G.Haddad in 1981 and was concerned with DI with upper semicontinuous convex-valued right-hand side

Haddad's Theorem

Assume that $K \subset \mathbb{R}^n$ is a closed set and $F : T \to K \to \mathrm{conv}(\mathbb{R}^n)$ is measurable in t for each $x \in K$ and is upper semicontinuous in x for a.e. $t \in T$. Let us assume also the tangential condition to be valid in the form:

$$F(t,x) \cap T_K^b(x) \neq \emptyset \quad \forall x \in K$$

Then for each $x_0 \in K$ there exists a solution $x(\cdot)$, $x(0) = x_0$, of the DI

A.Bressan proved in 1983 the viability result for an **autonomous** differential inclusion with lower semicontinuous right-hand side having not necessarily convex values but with a more strong tangential assumption

$$F(x) \subset T_K^b(x)$$

He showed also that under the weak Haddad's condition there is no viability in a very simple case (F is constant but not convex)

The Bressan's result was extended to nonautonomous case by G.Colombo in 1990

The prrofs by A.Bressan and G.Colombo were very technical involving some special approximations similar to Euler polygons

The question is: how can the indirect methods (such as continuous selections method) be applied to the viability problems?

There are some difficulties because this (continuous selections) method not always works

However, it can be applied to the case when the set K is **convex**

Although existence of a solution to the Cauchy problem was already proved by another method (Colombo's result), by using the continuous selections technique we can get solutions with other (so named nonlocal) initial data (for instance, with the periodical condition $x(0) = x(1)$) where the convexity of the set K is very essential

Theorem

Let X be a Banach space and $K \subset X$ be a compact convex set. Let $F : T \times K \to 2^X$ be such that

(i) F is $\mathfrak{M} \otimes \mathfrak{B}$-measurable mapping that implies easily the superpositional meaasurability

(ii) the mapping $x \mapsto F(t, x)$ is lower semicontinuous (by Vietoris) for a.e. $t \in T$

(iii) there exists a summable function $l : T \to \mathbb{R}^+$ such that $\|v\| \leq l(t)$ for all $v \in F(t, x)$, $x \in K$ a.e. on T

(iv) $F(t, x) \subset T_K(x)$ for all $x \in K$

Then, given $x_0 \in K$ there exists a solution $x(\cdot)$, $x(0) = x_0$, of DI such that $x(t) \in K$ for all $t \in T$

Sketch of the proof

First, we consider the set

$$\mathcal{K} := \{x(\cdot) \in L^1(T, X) : x(t) \in K \text{ for a.e. } t \in T\}$$

and its compact convex subset

$$\mathcal{K}^* := \{x(\cdot) \in \mathcal{K} \cap \mathbf{AC}(T, X) : \|\dot{x}(t)\| \leq 2(l(t) + 1) \text{ for a.e. } t \in T\}$$

As usual we prove that the **Nemytski operator** $\mathfrak{F} : \mathcal{K} \rightrightarrows L^1(T, X)$,

$$\mathfrak{F}(x) := \{u(\cdot) \in L^1(T, X) : u(t) \in F(t, x(t)) \text{ for a.e. } t \in T\},$$

is lower semicontinuous (by Vietoris) and admits nonempty closed decomposable values

So, choosing a continuous selection $g(x) \in \mathfrak{F}(x)$, $x \in \mathcal{K}$, we prove then that

$$g(x) \in \mathfrak{T}_{\mathcal{K}}(x) \qquad \forall x \in \mathcal{K}$$

or, in other form,

$$\lim_{\lambda \to 0+} \frac{1}{\lambda} \mathfrak{D}_{\mathcal{K}}(x + \lambda g(x)) = 0$$

(Here $\mathfrak{T}_{\mathcal{K}}(\cdot)$ is a tangent cone to the convex set $\mathcal{K}$ and $\mathfrak{D}_{\mathcal{K}}(\cdot)$ is the distance from the set $\mathcal{K}$ in the space $L^1(T, X)$)

It is easy to show that the convergence above is uniform on compact subsets of $\mathcal{K}$, in particular, on $\mathcal{K}^*$

Hence, we choose a sequence $\lambda_n \to 0+$ such that

$$\mathfrak{D}_{\mathcal{K}}(x + \lambda_n g(x)) \leq \lambda_n / n \qquad \forall x(\cdot) \in \mathcal{K}^*$$

Let us consider now the multifunction $\mathcal{P}_n : \mathcal{K}^* \rightrightarrows L^1(T, X)$,

$$\begin{aligned} \mathcal{P}_n(x) &:= \{v(\cdot) \in \mathcal{K} : \|x(t) + \lambda_n g(x)(t) - v(t)\| \\ &< d_K(x(t) + \lambda_n g(x)(t)) + \lambda_n/n \text{ for a.e. } t \in T\} \end{aligned}$$

It is lower semicontinuous (see above), so by Fryszkowski Theorem there exists a continuous selection $v_n(x) \in \overline{\mathcal{P}(x)}$

Thus, we have the inequality

$$\|x(t) + \lambda_n g(x)(t) - v_n(x)(t)\| \leq d_K(x(t) + \lambda_n g(x)(t)) + \lambda_n/n$$

Integrating on T, dividing by λ_n and changing the integral and the distance, we arrive at

$$\|f(x) - f_n(x)\|_C \leq \frac{1}{\lambda_n}\mathfrak{D}_{\mathcal{K}}(x + \lambda_n g(x)) + \lambda_n/n \leq 2\lambda_n/n$$

where

$$f(x)(t) := x_0 + \int_0^t g(x)(s)\, ds$$

and

$$f_n(x)(t) := x_0 + \int_0^t \frac{v_n(x)(s) - x(s)}{\lambda_n}\, ds, \qquad x(\cdot) \in \mathcal{K}^*$$

Extend the functions $v_n(\cdot)$ (and, respectively, $f_n(\cdot)$) to a little bit larger also compact convex set $\mathcal{K}_n \supset \mathcal{K}^*$,

$$\mathcal{K}_n := \left\{ x(\cdot) \in \mathcal{K} \cap \mathbf{AC}(T, X) : \|\dot{x}(t)\| \leq \frac{\operatorname{diam} K}{\lambda_n} \right\}$$

and define the following exponential operator on $\mathcal{K}_n$:

$$\sigma_n(x)(t) := x_0 \exp(-\frac{1}{\lambda_n} t) + \lambda_n^{-1} \int_0^t v_n(x)(s) \exp\left(\frac{t-s}{\lambda_n}\right) ds$$

It turns out that σ_n maps $\mathcal{K}_n$ to $\mathcal{K}_n$, it is continuous and satisfies the equality

$$\frac{d}{dt}\sigma_n(x)(t) = \frac{v_n(x)(t) - \sigma_n(x)(t)}{\lambda_n}$$

Choosing by Schauder's Theorem a fixed point $x_n(\cdot) \in \mathcal{K}_n$ of σ_n, we have

1. $x_n(t) = \sigma_n(x_n)(t) = f_n(x_n)(t)$
2. $x_n(\cdot) \in \mathcal{K}^*$ for all $n = 1, 2, ...$

It follows from 2. and from the compactness of $\mathcal{K}^*$ that the sequence $\{x_n(\cdot)\}$ has a convergent subsequence

Assuming without loss of generality that it converges to some function $x(\cdot)$, by 1. and by the previous estimates it is easy to show that $x(\cdot)$, $x(0) = x_0$, is the required solution

THE END

THANK YOU FOR ATTENTION

GRAZIE PER L'ATENZIONE