An introduction to Mathematical Theory of Control

Vasile Staicu

University of Aveiro

UNICA, May 2018

Introduction

- Since the beginnings of Calculus, differential equations have provided an effective mathematical model for a wide variety of physical phenomena.
- Consider a system whose state can be described by a finite number of real valued parameters, say $x = (x_1, x_2, ..., x_n)$.
- If the rate of change $\dot{x} = \frac{dx}{dt}$ is entirely determined by the state x itself, then the evolution of the system can be modelled by the ordinary differential equation

$$\dot{x} = g(x). \tag{1}$$

- If the state of the system is known at some initial time t_0, the future behavior for $t > t_0$ can then be determined by solving a problem of Cauchy consisting of (1) together with the initial condition

$$x(t_0) = x_0. \tag{2}$$

Introduction

- We are here taking a spectator's point of view: the mathematical model allows us to understand a portion of the physical word and predict its future evolution, but we have no means of altering its behavior in any way.
- Celestial mechanics provides a typical example of this situation.
- We can accurately calculate the orbits of moons and planets and exactly predict time and location of eclipses, but we cannot change them in the slightest amount.
- This classical point of view assumes that physical systems are isolated and all the information contained in them is completely accessible to the external observer.
- The mathematical model describes the evolution of the system taking into account its internal strengths and observations on the systems made from abroad but there is no way to interfere in this evolution.

Introduction

- The failure of this framework turns out to be clear in front of the development of technology and engineering. The invention of machines and tools and their dissemination put the researcher in front of the reality of systems that can be easily guided from abroad in order to regulate its operation based on a pre-established project and with the specific objective to be realized.
- Mathematical models of these systems should take into account also the forces or impulses coming from abroad.
- Thus, comes the need to accept the *modern point of view*, where any physical system (machine, or chemical, biological or social process) is defined in a certain region of space and has its own dynamics in base which develops, exchanging continuously information with the external environment.

Introduction

- Such exchange of information occurs through certain channels that are named generically *inputs* (if the information coming from the exterior to the system) and *outputs* (the information will be to the outside of the system).
- Graphically we can represent a system like

$$u(t) \longmapsto \blacksquare \longmapsto x(t)$$

- The mathematical control theory is the theoretical basis for the study of physical systems equipped with inputs and outputs and provides a different paradigm.
- We now assume the presence of an external agent, namely a "controller" who can actively influence the evolution of the system.

- This new situation is modelled by a control system,

$$\dot{x} = f(x, u), \; u(.) \in \mathcal{U} \tag{3}$$

where $\mathcal{U}$ is a family of admissible control functions.

- In this case, the rate of change $\dot{x}(t)$ depends not only of the state x itself, but also on some external parameters, say $u = (u_1, u_2, ..., u_m)$, which can also vary in time.
- The control function $u(.)$ subject to some constraints, will be chosen by a controller in order to modify the evolution of the system and achieve certain preassigned goals:
    - steer the system from a state to another,
    - maximize the terminal value of one of the parameters,
    - minimize a certain cost functional, etc.

Introduction

- In a standard setting, we are given a set of control values $U \subset \mathbb{R}^m$.
- The family of *admissible control functions* is defined as

$$\mathcal{U} := \{u : \mathbb{R} \to \mathbb{R}^m : u \text{ measurable, } u(t) \in U \text{ for a. e. } t\}$$

- The control system (3) can be written as a differential inclusion, namely

$$x' \in F(x) \tag{4}$$

where, for each x, the set of possible velocities is given by

$$F(x) := \{y : y = f(x, u) \text{ for some } u \in U\}. \tag{5}$$

Introduction

- An absolutely continuous function $x : [0, T] \to \mathbb{R}^n$ is said to be an admissible solution (or trajectory) of the control system (3) if there exists a measurable control $u : [0, T] \to U$ such that

$$x'(t) = f(x(t), u(t)) \text{ for almost every } t \in [0, T]. \tag{6}$$

- By solution (or trajectory) of the differential inclusion (4) we mean any absolutely continuous function $x : [0, T] \to \mathbb{R}^n$ such that

$$x'(t) \in F(x(t)) \text{ for almost every } t \in [0, T]. \tag{7}$$

- Clearly, every admissible trajectory of the control system (3) is also a solution of the differential inclusion (4).

Introduction

- Under some regularity assumptions on f, given any solution of (4), one can select a measurable control $u : [0, T] \to U$ such that

$$x'(t) = f(x(t), u(t)) \text{ for almost every } t \in [0, T],$$

that is, x is a trajectory of the control system (3).

- Differential inclusions often provide a convenient approach for the analysis of control systems, as we will see in what follows.

Introduction

- The control law can be assigned in two basically different ways:
  - in "open loop" form, as a function of time: $t \longrightarrow u(t)$, and
  - in "closed loop" or "feedback", as a function of the state: $x \longrightarrow u(x)$.
- Implementing an open loop control $u = u(t)$ is in a sense easier, since the only information needed is provided by a clock, measuring time.
- On the other hand, to implement a closed loop control $u = u(x)$ one constantly needs to measure the state x of the system.

Introduction

- Designing a feedback control, however yields to some distinct advantages. The feedback controls are more robust in the presence of random perturbations.
- For example, assume that we seek a control $u(.)$ which steers the system described by (3) from an initial state P to the origin. This can be achieved say by an open loop control $t \longrightarrow u(t)$.
- In many practical situations, however, the evolution is influenced by additional disturbance which cannot be predicted in advance. The actual behavior of the system will be thus be governed by

$$\dot{x} = f(x, u) + \eta(t), \tag{8}$$

where $t \longrightarrow \eta(t)$ is a perturbation term. In this case, if the open loop control $u = u(t)$ steers the system (3) from an initial state P to the origin, this same control function may not accomplish the same task with (8), when a perturbation is present.

Introduction

- Alternatively, one can solve the problem of steering the system described by (3) from an initial state P to the origin by mean of a closed loop control. In this case, we would seek a control function $u = u(x)$ such that all the trajectories of the ordinary differential equation

$$\dot{x} = g(x) := f(x, u(x)) \tag{9}$$

approach the origin as $t \to \infty$.

- This approach is less sensitive to the presence of external disturbances.

- Two examples will be presented below,

Example 1: boat on the river

- Consider a river with a straight course, and using a set of planar coordinates, assume that it occupies the horizontal strip

$$S := \left\{(x_1, x_2) \in \mathbb{R}^2 : -\infty < x_1 < +\infty, \ -1 < x_2 < 1\right\}.$$

Moreover, assume that speed of the water is given by the velocity vector

$$v(x_1, x_2) = \left(1 - x_2^2, 0\right).$$

If a boat on the river is merely dragged by the current, its position will be determined by the differential equation

$$(\dot{x}_1, \dot{x}_2) = \left(1 - x_2^2, 0\right).$$

On the other hand, if the boat is powered by an engine, then its motion can be modelled by the control system

$$(\dot{x}_1, \dot{x}_2) = \left(1 - x_2^2 + u_1, u_2\right), \tag{10}$$

where $u = (u_1, u_2)$ describe the velocity of the boat relative to the water.

Example 1

- The set $\mathcal{U}$ of admissible controls consists of all measurable functions $u : \mathbb{R} \to \mathbb{R}^2$ taking values inside the closed disc

$$U = \left\{ (\omega_1, \omega_2) \in \mathbb{R}^2 : \sqrt{\omega_1^2 + \omega_2^2} \leq M \right\}, \tag{11}$$

where M accounts for the maximum speed (in any direction) that can be produced by the engine.

- Given an initial condition $(x_1, x_2)(0) = (\bar{x}_1, \bar{x}_2)$ solving (10) one finds

$$\begin{aligned} x_1(t) &= \bar{x}_1 + t + \int_0^t u_1(s)\, ds - \int_0^t \left(\bar{x}_2 + \int_0^s u_2(\tau)\, d\tau \right)^2 ds, \\ x_2(t) &= \bar{x}_2 + \int_0^t u_2(s)\, ds \qquad (-1 \leq x_2 \leq 1). \end{aligned}$$

Example 1

- In particular, the constant control $u = (u_1, u_2) = \left(-\frac{2}{3}, 1\right)$ takes the boat from a point $(\bar{x}_1, -1)$ on one side of the river to the point $(\bar{x}_1, 1)$ on the opposite side, in two units of time.
- Observe that if $M > 0$ then the boat can be steered from any point P of the river to any other point of the river (exercise: justify !) and that the admissible trajectories of the control system $(10) - (11)$ coincide with the solutions of the differential inclusion

$$(\dot{x}_1, \dot{x}_2) \in F(x_1, x_2) = \left\{ (y_1, y_2) : \sqrt{\left(y_1 - 1 + x_2^2\right)^2 + y_2^2} \leq M \right\}.$$

Example 2: cart on a rail

- Consider a cart which can move without friction along a straight rail. For simplicity assume that it has unit mass.
- Let $y(0) = \bar{y}$ be its initial position and $v(0) = \bar{v}$ be its initial velocity.
- If no forces are present, its future position is simply given by

$$y(t) = \bar{y} + t\bar{v}.$$

- Next, assume that a controller is able to push the cart, with an external force $u = u(t)$.

Example 2

- The evolution of the system is then determined by the second order equation (determined by the Newton law)

$$\ddot{y} = u(t). \tag{12}$$

- Calling now $x_1(t) = y(t)$ and $x_2(t) = v(t)$, respectively the position and the velocity of the chart at time t, the equation (12) can be written as a first order control system

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u \end{cases} \tag{13}$$

Given the initial condition $x_1(0) = \bar{y}$, $x_2(0) = \bar{v}$, solving (13) one finds

$$x_1(t) = \bar{y} + t\bar{v} + \int_0^t (t-s)\, u(s)\, ds, \;\; x_2(t) = \bar{v} + \int_0^t u(s)\, ds.$$

Example 2

- Assuming that the force satisfies the constraint

$$|u(t)| \leq 1,$$

the control system (13) is equivalent to the differential inclusion

$$(\dot{x}_1, \dot{x}_2) \in F(x_1, x_2) = \{(x_2, \omega) : -1 \leq \omega \leq 1\}.$$

Example 2

- We now consider the problem of steering the system at the origin $(0,0) \in \mathbb{R}^2$.
- In other words, we want the chart to be eventually at the origin with zero speed.
- For example, if the initial condition $(\bar{y}, \bar{v}) = (2,2)$, this goal is achieved by the open-loop control

$$\bar{u}(t) = \begin{cases} -1 & \text{if} \quad 0 \leq t \leq 4 \\ 1 & \text{if} \quad 4 \leq t \leq 6 \\ 0 & \text{if} \quad t \geq 6. \end{cases}$$

  A direct calculation shows that $(x_1(t), x_2(t)) = (0,0)$ for $t \geq 6$.
- However, the above control would not accomplish the same task in connection with $(\bar{y}, \bar{v}) \neq (2,2)$ (exercise: verify!)

Example 2

- A related problem is that of asymptotic stabilization: we seek a feedback control function $u = u(x_1, x_2)$ such that for every initial data $(\bar{y}, \bar{v})$ the corresponding solution of the Cauchy problem

$$(\dot{x}_1, \dot{x}_2) = (x_2, u(x_1, x_2)), \quad (x_1, x_2)(0) = (\bar{y}, \bar{v})$$

approaches the origin as $t \longrightarrow \infty$, that is,

$$\lim_{t \longrightarrow \infty} (x_1, x_2)(t) = (0, 0).$$

- There are several feedback controls which accomplish this task.
- One is

$$u(x_1, x_2) = -x_1 - x_2.$$

Questions of interest

- In connection with a control system of the general form (3) several mathematical questions can be formulated.
- A first set of problems is concerned with the *dynamics of the system*.
- Given an initial state $\bar{x}$, one would like to determine which other states $x \in \mathbb{R}^n$ can be reached using the various admissible controls $u \in \mathcal{U}$.
- More precisely, given a control function $u = u(t)$, call $x(., u)$ the solution of the Cauchy problem

$$\dot{x}(t) = f(x(t), u(t)), \ x(0) = \bar{x},$$

and define the reachable set at time t as

$$R(t) = \{x(t, u) : u \in \mathcal{U}\}.$$

Questions of interest

- The closure, boundedness, convexity and the dimension of the set $R(t)$ provide useful information on the control system.
- In addition, it is interesting to study whether $R(t)$ is a neighborhood of the initial point $\bar{x}$ for all $t > 0$. In the positive case, the system is said to be *small time locally controllable at* $\bar{x}$.
- Another important case is when $\bigcup_{t>0} R(t)$ includes the entire space $\mathbb{R}^n$. We then say that the system is *globally controllable.*
- The dependence of the reachable set $R(t)$ on the time t and on the set of admissible controls $\mathcal{U}$ is also a subject of investigation.
- One may ask whether the same points in $R(t)$ can be reached by using controls which are piecewise constant, or take values within the set of extreme points of U.
- Being able to perform the same tasks by mean of a smaller set of control functions, easier to implement, is quite relevant in practical applications.

Questions of interest

- A second very important area of control theory is concerned with the *optimal control*.
- In many applications, among all strategies which accomplish a certain task, one seeks an optimal one, based on a given performance criterion.
- A performance criterion can be defined by an integral functional of the form
$$J = \int_0^T L(t, x, u)\, dt, \tag{14}$$
and the values of J will have to be optimized among the admissible trajectories of (3), with a number of initial or terminal constraints.

Questions of interest

- For example, among all controls which steer the system from the initial point $\bar{x}$ to some point on the target set $\mathcal{T}$ at time T, we may seek the one that minimize the cost functional (14).
- This problem is formulated as

$$\min_{u\in\mathcal{U}} \int_0^T L(t,x,u)\,dt, \tag{15}$$

subject to

$$\dot{x} = f(x,u), \; x(0) = \bar{x}, \; x(T) \in \mathcal{T}. \tag{16}$$

- Observe that if (3) takes the simplest form

$$\dot{x} = u, \; u(t) \in U = \mathbb{R}^n$$

and if $\mathcal{T} = \{\bar{y}\}$ consists on just one point, then we do not have any constraint on the derivative $\dot{x}$.

Questions of interest

- Our problem of optimal control thus reduces to the standard problem of the Calculus of Variations
$$\min_{x(.)} \int_0^T L(t, x, \dot{x})\, dt, \;\; x(0) = \bar{x}, \;\; x(T) = \bar{y}. \tag{17}$$
- Roughly speaking, the main difference between the problem $(15) - (16)$ and (17) is that in (17) the derivative $\dot{x}$ is not restricted, while in $(15) - (16)$ the derivative is constrained within the set
$$F(x) := \{y : y = f(x, u) \;\text{ for some } u \in U\}.$$
- The basic mathematical theory of optimal control has been concerned with the following main topics: existence of optimal controls; necessary conditions for the optimality of a control; sufficient conditions for optimality.
- The next section will be dedicated to Carathéodory solutions to ordinary differential equations which arise from control systems, and to which the classical notion of solution is not appropriate.

Normed vector spaces

- A *real vector space* is a nonempty set X endowed with two operations: $+ : X \times X \to X$ and $\cdot : \mathbb{R} \times X \to X$ satisfying the following conditions:

1. $(x+y)+z = x+(y+z)$, $\forall x, y, z \in X$;
2. $x+y = y+x$, $\forall x, y \in X$;
3. there exists $0 \in X$ such that $x+0 = x$, $\forall x \in X$;
4. for every $x \in X$ there exists $-x \in X$ such that $x+(-x) = 0$ $\forall x \in X$;
5. $(\alpha+\beta)\cdot x = \alpha \cdot x + \beta \cdot x$, $\forall \alpha, \beta \in \mathbb{R}$, $\forall x \in X$;
6. $\alpha \cdot (x+y) = \alpha \cdot x + \alpha \cdot y$, $\forall \alpha \in \mathbb{R}$, $\forall x, y \in X$;
7. $1 \cdot x = x$, $\forall x \in X$ (where 1 is the unit real number)
8. $\alpha \cdot (\beta \cdot x) = (\alpha\beta)\cdot x$, $\forall \alpha, \beta \in \mathbb{R}$, $\forall x \in X$.

- We denote by $(X, +, \cdot)$ a vector space with the operations $+$ and $\cdot$.

Normed vector spaces

- A subset A of a vector space $(X, +, \cdot)$ is said to be *convex* if

$$\lambda x + (1 - \lambda) y \in A, , \forall x, y \in A \text{ and } \forall \lambda \in [0, 1] .$$

- Let $(X, +, \cdot)$ be a (real) vector space. A norm on X is a mapping $\|.\| : X \longrightarrow \mathbb{R}$ with the following properties:
  (i) $\|x\| = 0$ if and only if $x = 0$;
  (ii) $\|\alpha \cdot x\| = |\alpha| \, \|x\| \; \forall \alpha \in \mathbb{R}, \forall x \in X$;
  (iii) $\|x + y\| \leq \|x\| + \|y\| \; \forall x, y \in X$.
- If $\|.\|$ is a norm on X then we say that $(X, \|.\|)$ is a *vector normed space*, or simply a *normed space*.

Normed vector spaces

- Let $(X, \|.\|)$ be a normed space and for $x_0 \in X$ and $r > 0$ let

$$B_r(x_0) = \{x \in X : \|x - x_0\| < r\}$$

be the *open ball* centered at x_0 with radius r, and

$$\overline{B}_r(x_0) = \{x \in X : \|x - x_0\| \leq r\}$$

be the *closed ball* centered at x_0 with radius r.

- Both the open and the closed balls centered at x_0 with radius r are convex set

Indeed, if $x, y \in B_r(x_0)$ and $0 \leq \lambda \leq 1$ then

$$\begin{aligned} \|\lambda x + (1-\lambda) y - x_0\| &\leq \|\lambda (x - x_0)\| + \|(1-\lambda)(y - x_0)\| \\ &= \lambda \|x - x_0\| + (1-\lambda) \|y - x_0\| \\ &< \lambda r + (1-\lambda) r = r \end{aligned}$$

hence $\lambda x + (1-\lambda) y \in B_r(x_0)$, and $B_r(x_0)$ is a convex set. The proof for $\overline{B}_r(x_0)$ is similar.

Normed vector spaces are metric spaces

- Recall that if $(X, \|.\|)$ is a normed space then $d : X \times X \to \mathbb{R}$ defined by $d(x, y) = \|x - y\|$ is a distance (or metric) on X, that is satisfies the following properties:
  (a) $d(x, y) = 0$ if and only if $x = y$;
  (b) $d(x, y) = d(y, x)$ $\forall x, y \in X$;
  (c) $d(x, y) \leq d(x, z) + d(z, y)$ $\forall x, y, z \in X$.
- If $d : X \times X \to \mathbb{R}$ satisfies (a), (b) and (c) then d is called a metric on X and (X, d) is said to be a *metric space*.

Banach spaces

- Let $(X, \|.\|)$ be a normed space.
- A sequence $(x_n)_{n\geq 1} \subset X$ is said to be convergent if there exists $x \in X$ such that: for every $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ so that $\|x_n - x\| < \varepsilon$ whenever $n \geq n_\varepsilon$.
- A sequence $(x_n)_{n\geq 1} \subset X$ is said to a Cauchy sequence if for every $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ so that $\|x_n - x_m\| < \varepsilon$ whenever $n, m \geq n_\varepsilon$.
- We say that a normed space is complete if every Cauchy sequence is convergent.
- A complete normed vector space is called a Banach space.

Examples of Banach spaces

- The finite dimensional space $\mathbb{R}^n$ with the Euclidian norm

$$\|x\| = \sqrt{\sum_{i=1}^{n} x_i}, \text{ if } x = (x_1, ..., x_n) \in \mathbb{R}^n;$$

- The space $C([a,b], \mathbb{R}^n)$ of all continuous functions $f : [a,b] \to \mathbb{R}^n$, with norm

$$\|f\|_\infty = \sup_{t\in[a,b]} \|f(t)\|;$$

- Given a subset $\Omega \subset \mathbb{R}^m$, we say that a map $f : \Omega \to \mathbb{R}^n$ is *Lipschitz continuous* if there exists a constant L such that

$$\|f(x) - f(y)\| \leq L \|x - y\|, \ \forall x, \ y \in \Omega.$$

- The space $Lip(\Omega, \mathbb{R}^n)$ of all these Lipschitz continuous mappings is a Banach space with norm

$$\|f\|_{Lip} = \sup_{x\in\Omega} \|f(x)\| + \sup_{x\neq y} \frac{\|f(x) - f(y)\|}{\|x - y\|}.$$

Examples of normed space which is not Banach spaces

- The space $X = \mathcal{P}([0,1])$ of polynomial functions $p : [0,1] \rightarrow \mathbb{R}$ with the norm
$$\|p\|_\infty = \sup_{t \in [0,1]} \|p(t)\|$$
is a normed space but not a Banach space
- Indeed, the sequence of polynomials $p_n(t) = \sum_{k=0}^{n} \frac{t^k}{k!}$ is a Cauchy sequence converges uniformly on $[0,1]$ hence also with respect to $\|.\|_\infty$ to the continuous function $f(t) = e^t$.
- Hence it is a Cauchy sequence in $(C([0,1], \mathbb{R}), \|.\|_\infty)$, therefore also in $(\mathcal{P}([0,1]), \|.\|_\infty)$.
- But its limit $f(t) = e^t$, is not polynomial, that is,it has no limit in the space $X = \mathcal{P}([0,1])$.

Banach contraction mapping principle

- The next Banach's theorem show that for an equation, a unique solution exists and depends continuously on the parameters that describe the problem if the equation can be written in the form

$$x = \Phi(\lambda, x) \tag{18}$$

and the map $x \rightarrow \Phi(\lambda, x)$ is a contraction for each given value of the parameter λ, that is, for some $\varkappa < 1$,

$$\|\Phi(\lambda, x) - \Phi(\lambda, y)\| \leq \varkappa \|x - y\|, \ \forall \lambda \in \Lambda, \forall x, \ y \in X.$$

Banach contraction mapping principle

- **Theorem** (contraction mapping principle) Let $(X, \|.\|)$ be a Banach space, Λ be a metric space and let $\Phi : \Lambda \times X \to X$ be a continuous mapping such that, for some $\varkappa < 1$,

$$\|\Phi(\lambda, x) - \Phi(\lambda, y)\| \leq \varkappa \|x - y\|, \ \forall \lambda \in \Lambda, \forall x, \ y \in X. \tag{19}$$

Then:
(i) For each $\lambda \in \Lambda$ there exists a unique $x(\lambda) \in X$ such that

$$x(\lambda) = \Phi(\lambda, x(\lambda)) \tag{20}$$

(ii) The map $\lambda \to x(\lambda)$ is continuous
(iii) For any $\lambda \in \Lambda$, $y \in X$ one has

$$\|y - x(\lambda)\| \leq \frac{1}{1-\varkappa} \|y - \Phi(\lambda, y)\|. \tag{21}$$

Proof of Banach contraction principle

- Fix any point $y \in X$. For each $\lambda \in \Lambda$, consider the sequence $y_k = y_k(\lambda)$ defined by

$$y_0 = y, y_{k+1} = \Phi(\lambda, y_k) \text{ for } k \geq 0.$$

- By induction, for every $k \geq 0$ one checks that

$$\|y_{k+1} - y_k\| \leq \varkappa^k \|y_1 - y_0\| = \varkappa^k \|y - \Phi(\lambda, y)\|. \tag{22}$$

- Since $\varkappa < 1$, the sequence $(y_k)_{k\geq 0}$ is Cauchy in the Banach space X, hence it converges to some limit point, which we call $x(\lambda)$.

- By the continuity of $\Phi(\lambda, .)$ we have

$$x(\lambda) = \lim_{k\to\infty} y_{k+1} = \lim_{k\to\infty} \Phi(\lambda, y_k) = \Phi(\lambda, x(\lambda)),$$

hence (20) holds.

- The uniqueness of $x(\lambda)$ is proved observing that, if

$$x_1 = \Phi(\lambda, x_1), \; x_2 = \Phi(\lambda, x_2),$$

then by (19) it follows that

$$\|x_1 - x_2\| = \|\Phi(\lambda, x_1) - \Phi(\lambda, x_2)\| \leq \varkappa \|x_1 - x_2\|.$$

The assumption $\varkappa < 1$ implies

$$x_1 = x_2.$$

Proof of Banach contraction principle

- Next, observe that (22) yields

$$\begin{aligned}\|y_{k+1}-y\| &\le \sum_{i=0}^{k}\|y_{i+1}-y_i\| \le \|y-\Phi(\lambda,y)\|\sum_{i=0}^{k}\varkappa^i \\ &\le \|y-\Phi(\lambda,y)\|\sum_{i=0}^{\infty}\varkappa^i = \frac{\|y-\Phi(\lambda,y)\|}{1-\varkappa}.\end{aligned}$$

  and letting $k\to\infty$ we obtain (21).
- To show the continuity of $\lambda\to x(\lambda)$ let $(\lambda_n)_{n\ge 1}$ be a sequence of parameters converging to λ^*.

Proof of Banach contraction principle

- Using (21) with $\lambda = \lambda_n$ and $y = x(\lambda^*)$ we obtain

$$\begin{aligned} \|x(\lambda^*) - x(\lambda_n)\| &\leq \frac{1}{1-\varkappa} \|x(\lambda^*) - \Phi(\lambda_n, x(\lambda^*))\| \\ &= \frac{1}{1-\varkappa} \|\Phi(\lambda^*, x(\lambda^*)) - \Phi(\lambda_n, x(\lambda^*))\| \end{aligned} \quad (23)$$

  and since $\lambda \to \Phi(\lambda, x)$ is continuous, the right-hand side of (23) tends to zero as $n \to \infty$.
- Hence $x(\lambda_n) \to x(\lambda^*)$ proving the continuity of $\lambda \to x(\lambda)$.

Elements of Lebesgue measure theory

- If $I = (a, b)$ is an open interval of $\mathbb{R}$ then the Lebesgue measure of I is defined by
$$\mu(I) = b - a.$$
- For any open set $A \subset \mathbb{R}$ the Lebesgue measure is defined by
$$\mu(A) = \sup\left\{\sum_{j\in J}\mu(I_j) : \bigcup_{j\in J} I_j \subset A,\ I_j \text{ open int., } I_j \cap I_k = \varnothing \text{ if } j \neq k\right\}.$$
- If $F \subset \mathbb{R}$ is compact (that is, closed and bounded), then there exists an open and bounded interval $I = (\alpha, \beta)$ such that $F \subset I$. Define then the Lebesgue measure of F by
$$\mu(F) = \mu(I) - \mu(I \backslash F).$$

- If $X \subset \mathbb{R}$ is any subset, then we define

$$\mu_*(X) = \sup\{\mu(F) : F \text{ compact } \subseteq X\}$$

and

$$\mu^*(X) = \inf\{\mu(A) : A \text{ compact open, bounded and with } X \subseteq A\}.$$

We say that X is Lebesgue measurable if $\mu_*(X) = \mu^*(X)$ and the common value is called the Lebesgue measure of X and denoted by $\mu(X)$:

$$\mu(X) := \mu_*(X) = \mu^*(X).$$

- One has the following:
Proposition: A subset $X \subseteq \mathbb{R}$ is Lebesgue measurable if and only if for every $\varepsilon > 0$ there exists a compact $K_\varepsilon \subseteq \mathbb{R}$ and an open set $A_\varepsilon \subseteq \mathbb{R}$ such that

$$K_\varepsilon \subseteq X \subseteq A_\varepsilon \text{ and } \mu(A_\varepsilon \backslash K_\varepsilon) < \varepsilon.$$

Elements of Lebesgue measure theory

- **Example:** A finite or countable set $A \subseteq \mathbb{R}$ is Lebesgue measurable and has a null measure.
- We denote by $\mathcal{L}$ the family of all Lebesgue measurable subsets of $\mathbb{R}$ and if I is an interval, we denote by $\mathcal{L}(I)$ the family of all measurable subsets of I :
$$\mathcal{L}(I) = \{A \in \mathcal{L} : A \subseteq I\}.$$
- Let $I \subseteq \mathbb{R}$ be an interval. A function $f : I \to \mathbb{R}^n$ is said to be Lebesgue measurable if for every open set $V \subseteq \mathbb{R}^n$ the preimage $f^{-1}(V) = \{t \in I : f(t) \in V\}$ is a Lebesgue measurable subset of I.
- Recall that $f : I \to \mathbb{R}^n$ is said to be continuous if at each $t_0 \in I$, one has
$$f(t_0) = \lim_{t \to t_0} f(t).$$
$f : I \to \mathbb{R}^n$ is said to be lower semicontinuous if at each $t_0 \in I$, one has
$$f(t_0) \leq \liminf_{t \to t_0} f(t).$$

Elements of Lebesgue measure theory

- **Remark:** Every continuous or lower semicontinuous function, is measurable.
- We say that a property P holds *almost everywhere* (a.e. for short) on a set $A \subseteq \mathbb{R}$ if there exists a set $N \subseteq \mathbb{R}$ such that $\mu(N) = 0$ and every point of $A \backslash N$ has the property P.
- Let $(f_n)_{n \geq 1}$ be a sequence of measurable functions, $f_n : I \longrightarrow \mathbb{R}^n$ and if

  $$f_n(x) \rightarrow f(x) \text{ for a.e. } x \in I$$

  then the function f is measurable.
- If $f : [a, b] \rightarrow \mathbb{R}^n$ is measurable and if $\|f(t)\| \leq \phi(t)$ for all $t \in [a, b]$ for some integrable scalar function ϕ, then f itself is integrable.

- If f and $\widehat{f}$ differ only on a set of measure zero are identified. With this equivalence relation, the space of integrable functions $f : [a, b] \to \mathbb{R}^n$ is written $L^1([a, b], \mathbb{R}^n)$ and it is a Banach space with norms

$$\|f\|_{L^1} = \int_a^b \|f(t)\| \, dt.$$

- **Lebesgue dominated convergence theorem:** If $(f_n)_{n \geq 1}$ is a sequence of measurable functions which converges a.e. to f and if there exists an integrable function ψ such that $\|f_n(t)\| \leq \psi(t)$ for all $n \geq 1$ and $t \in [a, b]$, then

$$\int_a^b f(t) \, dt = \lim_{n \to \infty} \int_a^b f_n(t) \, dt$$

- **Characterization of measurable functions:** For any function $f\,[a, b] \to \mathbb{R}^n$, the following statements are equivalent:
  (i) f is measurable
  (ii) For every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subseteq [a, b]$ with

$$\mu\left([a, b] \setminus K_\varepsilon\right) < \varepsilon$$

  and such that the restriction

$$f\,\big|_{K_\varepsilon} \text{ is continuous.}$$

- **Absolutely continuous functions:** A function $f : [a, b] \to \mathbb{R}^n$ is said to be absolutely continuous if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for any finite family of disjoint intervals $\{]s_i, t_i] : 1 \leq i \leq N\}$ with total length

$$\sum_{i=1}^{N} |t_i - s_i| < \delta$$

one has

$$\sum_{i=1}^{N} |f(t_i) - f(s_i)| < \varepsilon.$$

- If $f : [a, b] \to \mathbb{R}^n$ is absolutely continuous, then its derivative f' is defined almost everywhere and it satisfies

$$f(t) = f(t_0) + \int_{t_0}^{t} f'(s)\, ds \text{ for all } t,\ t_0 \in [a, b].$$

- Every Lipschitz continuous function f defined on $[a, b]$ is absolutely continuous.
- The set of all absolutely continuous functions $f : [a, b] \to \mathbb{R}^n$ is denoted by

$$AC([a, b], \mathbb{R}^n).$$

Elements of multivalued analysis

- Assume (X, d) is a metric space. For $x \in X$ and $A \subseteq X$ let $d(x, A)$ be the distance from x to A defined by

$$d(x, A) = \inf \{d(x, y) : y \in A\}.$$

- If $A \subseteq X$ and $\varepsilon > 0$ then $B(A, \varepsilon)$ denotes the $\varepsilon-$ neighborhood of A defined by

$$\overline{B}(A, \varepsilon) = \{x \in X : d(x, A) \leq \varepsilon\}.$$

- If $A = \{x_0\}$ then $B(A, \varepsilon)$ is the closed ball centered at x_0 with radius ε :

$$\overline{B}(x_0, \varepsilon) = \{x \in X : d(x, x_0) \leq \varepsilon\}.$$

- If A, B are two closed, bounded and nonempty subsets of X then the Hausdorff-Pompeiu distance between A and B is defined by

$$d_H(A, B) = \max\left\{\sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A)\right\}.$$

- One has that

$$\begin{aligned} d_H(A, B) &= \inf\left\{\varepsilon > 0 : B \subset \overline{B}(A, \varepsilon) \text{ and } A \subset \overline{B}(B, \varepsilon)\right\} \\ &= \max\left\{d(a, B), d(b, A) : a \in A \text{ and } b \in B\right\}. \end{aligned}$$

- The space $C(X)$ of all closed, bounded and nonempty subsets of X endowed with the Hausdorff-Pompeiu distance d_H is a metric space.

- Let X and Y be two sets and let $2^Y = \{A : A \subseteq Y\}$ be the family of all subsets of Y.
- By multifunction or multivalued map or set-valued map from X to Y we mean a map or correspondence which associate to each $x \in X$ a unique subset $F(x)$ of Y.
- The set $F(x)$ is called the value of F at x.
- A multifunction F from X to Y will be denoted by $F : X \to 2^Y$.
- The set
$$Gr(F) = \{(x, y) : x \in X, y \in F(x)\}$$
is called the *graph* of the multifunction $F : X \to 2^Y$.

Elements of multivalued analysis

- Let X, Y be metric spaces. We say that F is closed, bounded, compact valued if its values are all closed, bounded or compact.
- We say that $F : X \to 2^Y$ has closed graph if $Gr(F)$ is a closed subset of the product metric space $X \times Y$, which is equivalent to:

$$\text{if: } x_n \in X,\ x_n \to x,\ y_n \in F(x_n),\ y_n \to y \text{ then } y \in F(x).$$

Elements of multivalued analysis

- Let $F : X \longrightarrow 2^Y$ be a multivalued map with closed and bounded values. We say that F is Hausdorff continuous at $x_0 \in X$ if

$$\lim_{x \longrightarrow x_0} d_H \left(F(x), F(x_0) \right) = 0,$$

  that is, for every $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that $d_H \left(F(x), F(x_0) \right) < \varepsilon$ whenever $x \in B(x_0, \delta_\varepsilon)$.

- Let $I \subseteq \mathbb{R}$ and let $F : I \longrightarrow 2^{\mathbb{R}^n}$ be a multifunction. A function $f : I \longrightarrow \mathbb{R}^n$ is said to be a selection of F if

$$f(t) \in F(t) \text{ for all } t \in I.$$

- If $f : I \longrightarrow \mathbb{R}^n$ is a selection of F and if it is a measurable function then we say that f is a measurable selection of F.

- In order to introduce the *lexicografic measurable selection*, we need to define the *lexicographic order* relation on $\mathbb{R}^n$: if $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$ then we say that $x < y$ if one of the following situations happens:

$$x_1 < y_1 \text{ or}$$
$$x_1 = y_1 \text{ and } x_2 < y_2 \text{ or}$$
$$.....$$
$$x_1 = y_1, ..., x_{n-1} = y_{n-1}, \text{ and } x_n < y_n.$$

- One has the following:
Proposition: If $K \subseteq \mathbb{R}^n$ is compact subset of $\mathbb{R}^n$ then there there exists a first element ξ of K with respect to the lexicografic order, that is, there exists $\xi \in K$ such that $\xi < x$ for all $x \in K$.

Elements of multivalued analysis

- If $F : I \longrightarrow 2^{\mathbb{R}^n}$ is a multifunction with compact values then we can define the lexicographic selection $t \longrightarrow \xi(t)$ where $\xi(t) \in F(t)$ is the first element of the compact set $F(t)$ with respect to the lexicographic order.
- One has the following:
  Proposition If $I = [a, b]$ and if $F : I \longrightarrow 2^{\mathbb{R}^n}$ is a compact valued multifunction, with closed graph and if $t \longrightarrow \xi(t)$ is the lexicografic selection of F then $t \longrightarrow \xi(t)$ is a measurable selection.

Carathéodory solutions

- As we already seen in the basic model of a control system a control system,

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

as soon as a control function $u = u(t)$ from the family $\mathcal{U}$ of admissible controls is assigned, the evolution can be determined by solving the ordinary differential equation (ode, for short)

$$\dot{x} = g(t, x) \tag{24}$$

where

$$g(t, x) := f(x, u(t)). \tag{25}$$

Classical solutions

- In the classical theory of ordinary differential equations like (24) we assume that the function g is continuous with respect to both the variable and the appropriate notion of solution is the one of "classical solution".
- **Definition:** Let $\Omega \subseteq \mathbb{R}\times\mathbb{R}^n$ be an open set and let $g : \Omega \rightarrow \mathbb{R}^n$ be a function which define the equation (24) . A function $\varphi : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$, with I an interval, is called a *classical solution* of (24) if $\varphi \in C^1 (I, \mathbb{R}^n)$, $G(\varphi) := \{(t, \varphi(t)) : t \in I\} \subseteq \Omega$ and

$$\varphi'(t) = g(t, \varphi(t)) \text{ for all } t \in I.$$

Carathéodory solutions

- For the case of control systems, the function g defined by (25) follows to be only measurable with respect to t variable, so the notion of classical solution is not appropriate. We will use instead the notion of Carathéodory solution.
- **Definition:** Let Ω be an open set in $\mathbb{R} \times \mathbb{R}^n$ and let $g : \Omega \to \mathbb{R}^n$ be a function which define an equation like (24) . A function $\varphi : I \subseteq \mathbb{R} \to \mathbb{R}^n$, with I an interval, is called a Carathéodory solution of (24) if $\varphi \in AC\left(I, \mathbb{R}^n\right)$, $G\left(\varphi\right) := \left\{\left(t, \varphi\left(t\right)\right) : t \in I\right\} \subseteq \Omega$ and

$$\varphi'\left(t\right) = g\left(t, \varphi\left(t\right)\right) \text{ a.e } t \in I. \tag{26}$$

Carathéodory solutions: some remarks

- (i) An absolutely continuous function $\varphi : [t_0, t_1] \subseteq \mathbb{R}^n$, is a Carathéodory solution of (24) if $G(\varphi) := \{(t, \varphi(t)) : t \in [t_0, t_1]\} \subseteq \Omega$ and

$$x(t) = x(t_0) + \int_{t_0}^{t} g(s, x(s))\, ds \text{ for every } t \in \ [t_0, t_1].$$

- (ii) Any classical solution is a Carathéodory solution
  (iii) The function $\varphi : [-1, 1] \to \mathbb{R}$ defined by

$$\varphi(t) = \begin{cases} t & \text{if} \quad t \geq 0 \\ -t & \text{if} \quad t < 0 \end{cases}$$

  is a Carathéodory solution but not a classical solutions of

$$|x'| = 1, x(0) = 0.$$

Existence of Carathéodory solutions

- We assume that the function $g : \Omega \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ satisfies the following conditions:

(A) For every x the function $t \to g(t,x)$ defined on the section $\Omega_x = \{t : (t,x) \in \Omega\}$ is measurable.
For every t, the function $x \to g(t,x)$ defined on the section $\Omega_t = \{x : (t,x) \in \Omega\}$ is continuous

(B) For every compact $K \subset \Omega$ there exist constants C_K, L_K such that

$$\begin{array}{c} \|g(t,x)\| \leq C_K, \ \|g(t,x) - g(t,y)\| \leq L_K \|x-y\|, \\ \text{for all } (t,x), (t,y) \in K. \end{array} . \tag{27}$$

- For some $(t_0, x_0) \in \Omega$ we consider the Cauchy problem (g, t_0, x_0) :

$$\dot{x} = g(t,x), \ x(t_0) = x_0. \tag{28}$$

Existence of Carathéodory solutions

- **Theorem** (existence) Given a map $g : \Omega \to \mathbb{R}^n$, consider the Cauchy problem (g, t_0, x_0) in (28) for some $(t_0, x_0) \in \Omega$.
  (i) If g satisfies (A), (B), then there exists $\varepsilon > 0$ such that (28) has a local solution $x(.)$ defined for $t \in [t_0, t_0 + \varepsilon]$
  (ii) Assume, in addition that the function g is defined in the entire space $\mathbb{R} \times \mathbb{R}^n$ and there exist constants C, L such that

$$\|g(t,x)\| \leq C, \ \ \|g(t,x) - g(t,y)\| \leq L\|x-y\| \tag{29}$$

  for all $(t,x), (t,y) \in \mathbb{R} \times \mathbb{R}^n$.
  Then for every $T > t_0$, the Cauchy problem (28) has a global solution $x(.)$ defined for all $t \in [t_0, T]$. Moreover, the solution depends continuously on the initial data x_0.

Proof of the Existence of Carathéodory solutions

- Assume that g satisfies (A) in $\mathbb{R} \times \mathbb{R}^n$ and (29). Let $T > t_0$. We shall construct a forward solution $x = x_{x_0} : [t_0, T] \to \mathbb{R}^n$.
- In view of applying Theorem 1 (the Banach contraction mapping principle) we define $\Lambda = \mathbb{R}^n$ (the initial value $x_0 \in \mathbb{R}^n$ plays here the role of a parameter). Let X be the space $C\left([t_0, T], \mathbb{R}^n\right)$ endowed with the weighted norm

$$\|x\|_+ = \sup_{t \in [t_0, T]} e^{\cdot -2Lt} \|x(t)\|,$$

which is equivalent to the usual $C\left([t_0, T], \mathbb{R}^n\right)$ norm defined by $\|x\|_\infty = \sup_{t \in [t_0, T]} \|x(t)\|$,and define $\Phi : \Lambda \times X \to X$ by setting

$$\Phi(x_0, w)(t) = x_0 + \int_{t_0}^{t} g(s, w(s))\, ds, \;\; t \in [t_0, T]. \qquad (30)$$

Proof of the Existence of Carathéodory solutions

- To prove that Φ is well defined, for each $w \in X$ we need to show that the composite function $s \longrightarrow g(s, w(s))$ is integrable. Let $\delta = T - t_0$.
- Given a function $w \in X$ we consider the sequence of piecewise constant functions $(w_n)_{n\in N}$ with

$$w_n(t) = w\left(t_0 + k\frac{\delta}{n}\right), \ t \in \left[t_0 + k\frac{\delta}{n}, t_0 + (k+1)\frac{\delta}{n}\right], \ 0 \le k \le n-1$$

- By assumption (A), the maps $t \longrightarrow g(t, w_n(t))$ are all measurable.
- Moreover, (29) implies that

$$\lim_{n\to\infty} \|g(t, w(t)) - g(t, w_n(t))\| \le \lim_{n\to\infty} L\|w(t) - w_n(t)\| = 0,$$

hence $t \longrightarrow g(t, w(t))$ is measurable, being a pointwise limit of a sequence of measurable functions.

Proof of the Existence of Carathéodory solutions

- Since

$$\|g(s, w(s))\| \leq C \text{ for all } s,$$

it follows that the integral in (30) is well defined and depends continuously on t.

- Hence Φ is well defined and takes values inside X.
- The continuity of the map $x_0 \longrightarrow \Phi(x_0, w)$ is obvious. To prove that $w \longrightarrow \Phi(x_0, w)$ is a contraction in $\left(X, \|.\|_+\right)$, let w, $w' \in X$ and let

$$\theta = \left\|w - w'\right\|_+ . \tag{31}$$

- By the definition of $\|.\|_+$ and (31), we have

$$\left\|w(s) - w'(s)\right\| \leq \theta e^{.2Ls} \text{ for all } s \in [t_0, T].$$

Proof of the Existence of Carathéodory solutions

- Moreover, by (29), for all $t \in [t_0, T]$,

$$
\begin{aligned}
& e^{\cdot -2Lt} \left\| \Phi\left(x_0, w(t)\right) - \Phi\left(x_0, w'(t)\right) \right\| \\
& = e^{-2Lt} \left\| \int_{t_0}^{t} \left[g\left(s, w(s)\right) - g\left(s, w'(s)\right) \right] ds \right\| \\
& \leq e^{-2Lt} \left\| \int_{t_0}^{t} L \left\| w(s) - w'(s) \right\| ds \right\| \\
& \leq e^{-2Lt} \int_{t_0}^{t} L\theta e^{2Ls} ds < \frac{\delta}{2}.
\end{aligned}
$$

- Therefore,

$$
\left\| \Phi\left(x_0, w\right) - \Phi\left(x_0, w'\right) \right\|_{+} \leq \frac{1}{2} \left\| w - w' \right\|_{+}.
$$

Proof of the Existence of Carathéodory solutions

- We can now apply the Banach contraction mapping principle, obtaining the existence of a unique continuous mapping $x_0 \longrightarrow x_{x_0}$ such that

$$x_{x_0} = \Phi\left(x_0, x_{x_0}\right)$$

that is,

$$x_{x_0}\left(t\right) = x_0 + \int_{t_0}^{t} g\left(s, x_{x_0}\left(s\right)\right) ds, \;\; t \in \left[t_0, T\right].$$

- This means that x_{x_0} is the required solutions, and this achieves the proof of (ii).

Proof of the Existence of Carathéodory solutions

- To prove (i) concerning local existence of solution without the additional assumption (29) , choose $\varepsilon > 0$ small enough so that the set

$$K = \{(t, x) : |t - t_0| \leq \varepsilon, \|x - x_0\| \leq \varepsilon\}$$

is entirely contained in Ω.

- Then consider a smooth cut-off function $\phi : \mathbb{R} \times \mathbb{R}^n \to [0, 1]$ such that $\phi \equiv 1$ on K, while $\phi \equiv 0$ outside some larger compact set K', with $K \subset K' \subset \Omega$. Observe that the function

$$g^+ (t, x) = \begin{cases} \phi (t, x) \, g (t, x) & \text{if} \quad (t, x) \in \Omega \\ 0 & \text{if} \quad (t, x) \notin \Omega \end{cases}$$

satisfies (A) and (B) , together with the extra assumption (29) , because it vanishes outside the compact set K'.

Proof of the Existence of Carathéodory solutions

- By the previous steps, there exists a solution $x(.)$ to the Cauchy problem
$$x' = g^+(t,x), \; x(t_0) = x_0$$
defined on arbitrary large interval $[t_0, T]$.
- We now recall that K is a neighborhood of (t_0, x_0). Therefore for some $\varepsilon > 0$ sufficiently small, the point $(t, x(t))$ remains in K as $t \in [t_0, t_0 + \varepsilon]$.
- Since g and g^+ coincide on K, the function $x(.)$ thus provides a local solution to the original problem (g, t_0, x_0) on the interval $[t_0, t_0 + \varepsilon]$.

Uniqueness of Carathéodory solutions

- The next lemma represents a main ingredient in several uniqueness proofs and it is a simplified version of Gronwall's lemma:
- **Lemma** (Gronwall) Let $z(.)$ be an absolutely continuous nonnegative function such that

$$z(t_0) \leq \gamma, \ \dot{z}(t) \leq \alpha(t) z(t) + \beta(t) \ \text{ for a.e. } \ t \in [0, T], \qquad (32)$$

for some integrable functions α, β and some constant $\gamma \geq 0$. Then z satisfies the inequality

$$z(t) \leq \gamma e^{\int_{t_0}^{t} \alpha(s) ds} + \int_{t_0}^{t} \beta(s) e^{\int_{s}^{t} \alpha(u) du} ds \text{ for all } t \in [0, T]. \qquad (33)$$

- **Theorem** (uniqueness) Let $g : \Omega \to \mathbb{R}^n$ satisfies (A) and (B), as in the existence Theorem. Let $x_1(.)$, $x_2(.)$ be solutions of (28), defined on the intervals $[t_0, t_1]$, $[t_0, t_2]$ respectively. If $T = \min\{t_1, t_2\}$, then

$$x_1(t) = x_2(t) \ \text{ for all } \ t \in [t_0, T].$$

Proof of the uniqueness of Carathéodory solutions

- Since g satisfies assumption (B), for the compact set

$$K = \{(t, x_1(t)) : t \in [t_0, T]\} \cup \{(t, x_2(t)) : t \in [t_0, T]\}$$

  there exists $L_K \geq 0$ such that

$$\|g(t,x) - g(t,y)\| \leq L_K \|x - y\| \text{ for all } (t,x),\ (t,y) \in K.$$

- For all $t \in [t_0, T] = [t_0, t_1] \cap [t_0, t_2]$ let

$$z(t) = \|x_1(t) - x_2(t)\|.$$

- One has that $z(.)$ is absolutely continuous and

$$\dot{z}(t) \leq \|\dot{x}_1(t) - \dot{x}_2(t)\| \leq L_K z(t) \text{ for almost all } t \in [t_0, T].$$

  Applying Gronwall's lemma with $\alpha = L_K$, $\beta = \gamma = 0$ we obtain that $z(t) \leq 0$ for all $t \in [t_0, T]$.
- Hence

$$z(t) = \|x_1(t) - x_2(t)\| = 0 \text{ for all } t \in [t_0, T].$$

Characterization of maximal solutions

- Let $g : \Omega \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$. In general, a solution to the Cauchy problem
$$\dot{x} = g(t, x), \; x(t_0) = x_0 \tag{34}$$
is defined only locally, for t in a neighborhood of the initial time t_0.
- If a solution cannot be extended beyong a certain time T, two cases may arise:
(i) as $t \to T^-$, the point $(t, x(t))$ approaches the boundary $\partial\Omega$ of the domain Ω;
(ii) as $t \to T^-$, the solution blows up, that is $\|x(t)\| \to \infty$.

- **Theorem** Let $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ be an open set, $g : \Omega \to \mathbb{R}^n$ satisfies the basic assumptions (A) and (B), $(t_0, x_0) \in \Omega$ and let $T = \sup\{\tau : (34) \text{ has a solution defined on } [t_0, \tau]\}$. Then either $T = +\infty$ or

$$\lim_{t \to T^-} \left[\|x(t)\| + \frac{1}{d((t, x(t)), \partial\Omega)} \right] = +\infty, \tag{35}$$

where $d((t, x), \partial\Omega) = \inf\{|t - s| + \|x - y\| : (s, y) \in \partial\Omega\}$.

Characterization of maximal solutions

- **Proof:** Assume that $T < \infty$. If (35) does not hold, then there exist $M > 0$, $\varepsilon > 0$ and a sequence $t_n \longrightarrow T^-$ as $n \longrightarrow \infty$, such that

$$\|x(t_n)\| \leq M, \ d((t_n, x(t_n)), \partial\Omega) \geq \varepsilon.$$

By possibly taking a subsequence, we can assume that $x(t_n)$ converges to a point x_∞ as $n \longrightarrow \infty$, with $(T, x_\infty) \in \Omega$.

- Let $\rho > 0$ so small that the cylinder

$$K = \{(t, x) : |t - T| \leq \rho, \|x - x_0\| \leq \rho\}$$

is entirely contained in Ω.

Characterization of maximal solutions

- Consider a smooth cut-off function $\phi : \mathbb{R} \times \mathbb{R}^n \to [0,1]$ such that $\phi \equiv 1$ on K, while $\phi \equiv 0$ outside some larger compact set K', with $K \subset K' \subset \Omega$, and define

$$g^+(t,x) = \begin{cases} \phi(t,x)\, g(t,x) & \text{if} \quad (t,x) \in \Omega \\ 0 & \text{if} \quad (t,x) \notin \Omega \end{cases} \tag{36}$$

- Then $g^+ \equiv g$ on K, and in addition g^+ satisfies the global bounds C, L such that

$$\left\| g^+(t,x) \right\| \le C, \ \left\| g^+(t,x) - g^+(t,y) \right\| \le L \left\| x - y \right\|$$

for all $(t,x), (t,y) \in \mathbb{R} \times \mathbb{R}^n$, for some constants L, $C \ge 1$.

Characterization of maximal solutions

- Fix $\delta > 0$ small enough so that

$$(2C+1)\,\delta \leq \rho,$$

and choose n so large that

$$\|x(t_n) - x_\infty\| \leq \delta, \;\; T - t_n \leq \delta.$$

- By (ii) of the existence Theorem, the Cauchy problem

$$y' = g^+(t, y), \;\; x(t_n) = x(t_n)$$

has a solution $y(.)$ defined on $[t_n, T+\delta]$.

Characterization of maximal solutions

- We can now define an extension $\widehat{x}$ of x by setting

$$\widehat{x}(t) = \begin{cases} x(t) & \text{if} \quad t_0 \leq t \leq t_n \\ y(t) & \text{if} \quad t_n \leq t \leq T + \delta. \end{cases}$$

- Since $\|y'(t)\| \leq C$ for $t \in [t_n, T+\delta]$, we have

$$\begin{aligned} \|y(t) - x_\infty\| &\leq \|y(t) - x(t_n)\| + \|x(t_n) - x_\infty\| \\ &\leq C(t - t_n) + \delta \leq 2C\delta + \delta \leq \rho. \end{aligned}$$

- Therefore, for all $t \in [t_n, T+\delta]$, the point $(t, y(t))$ remains inside the compact K where g and g^+ coincide.
- The function $\widehat{x}(.)$ thus provide a solution of the original problem (34), defined on a strictly larger interval $[t_0, T+\delta]$. This contradicts the maximality of T, thus proving the theorem.

- We consider the control system

$$x' = f(t, x, u), \ \ u(.) \in \mathcal{U} \tag{37}$$

where the set of admissible controls is defined as

$$\mathcal{U} := \{u : \mathbb{R} \to \mathbb{R}^m : u \text{ measurable, } u(t) \in U \text{ for all } t\} \tag{38}$$

- We shall assume the following basic hypothesis:
(H) : The set $U \subset \mathbb{R}^m$ of control values is compact, $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ is an open set, the function $f : \Omega \times U \to \mathbb{R}^n$ is continuous in all variables and continuously differentiable with respect to x.

Trajectory of nonlinear control systems

- **Definition:** An absolutely continuous function $x : [a, b] \longrightarrow \mathbb{R}^n$ is said to be a solution (or trajectory) of (37) if its graph $\{(t, x(t)) : t \in [0, T]\}$ is entirely contained in Ω, and if there exists a measurable control $u : [a, b] \longrightarrow U$ such that

$$x'(t) = f(t, x(t), u(t)) \text{ for almost every } t \in [a, b].$$

An equivalent differential inclusion

- In connection with (37), define the multifunction $F : \Omega \to \mathcal{P}(\mathbb{R}^n)$ by

$$F(t,x) = \{f(t,x,\omega) : \omega \in U\} \tag{39}$$

and consider the differential inclusion

$$x' \in F(t,x). \tag{40}$$

- **Definition:** We call *solution* or *trajectory* of the differential inclusion (40) any absolutely continuous function $x : [a,b] \to \mathbb{R}^n$ whose graph $\{(t, x(t)) : t \in [a,b]\}$ is entirely contained in Ω, and

$$x'(t) \in F(t, x(t)) \text{ for almost every } t \in [a,b].$$

An equivalent differential inclusion

- Observe that, for each (t, x), the set of admissible velocities x' in (37) is given in parametrized form, as the image of a fixed set $U \subseteq \mathbb{R}^n$.
- On the other hand, when we study the differential inclusion (40), we do not assume any parametrization.
- The next result due to Filippov, shows that the two approaches are essentially equivalent.
- **Filippov's lemma** An absolutely continuous function $x : [a, b] \longrightarrow \mathbb{R}^n$ is a trajectory of the control system (37) if and only if it is a trajectory of the differential inclusion (40) with $F(t, x)$ defined by (39).

Proof of Filippov's lemma

- If $x : [a, b] \to \mathbb{R}^n$ is a trajectory of the control system (37) then its graph is entirely contained in Ω, and if there exists a measurable control $u : [a, b] \to U$ such that

$$x'(t) = f(t, x(t), u(t)) \text{ for almost every } t \in [a, b].$$

- Then

$$x'(t) = f(t, x(t), u(t)) \in F(t, x(t)) \text{ for almost every } t \in [a, b],$$

hence it is a trajectory of the differential inclusion (40).

- Assume now that $x : [a, b] \to \mathbb{R}^n$ is a trajectory of the differential inclusion (40). Then

$$x'(t) \in F(t, x(t)) \text{ for almost every } t \in [a, b].$$

Proof of Filippov's lemma

- Fix any $\overline{\omega} \in U$ and define the multifunction $W : [a, b] \to 2^{\mathbb{R}^n}$ by

$$W(t) = \begin{cases} \{\omega \in U : f(t, x(t), \omega) = x'(t)\} & \text{if } x'(t) \in F(t, x(t)) \\ \{\overline{\omega}\} & \text{otherwise.} \end{cases}$$

- Remark that $W(t)$ is a compact subset and $W(t) = \{\overline{\omega}\}$ for t in a set of null Lebesgue measure.
- Let $u(t)$ be the first element of $W(t)$. Then by Proposition 2 of Section 2 one has that $t \to u(t)$ is measurable, $u(t) \in W(t)$ hence

$$x'(t) = f(t, x(t), u(t)),$$

and $x(.)$ is a trajectory of the control system (37).

Fundamental properties of trajectories

- Let f, U satisfy the basic assumption (H) and let $(0, \overline{x}) \in \Omega$.
- For any measurable control $u : [0, T] \to U$, the Cauchy problem

$$x' = f(t, x, u(t)), x(0) = \overline{x} \tag{41}$$

is equivalent to

$$x' = g(t, x), x(0) = \overline{x}. \tag{42}$$

where $g(t, x) = f(t, x, u(t))$ is sufficiently regular to have a local solution $x(., u) : [0, \delta] \to \mathbb{R}^n$ of (42).

Fundamental properties of trajectories

- To study how the solution $x(., u)$ depends on u, we first consider the globally bounded case, assuming
  (H^*) : The set $U \subset \mathbb{R}^m$ of control values is compact, the function $f : \mathbb{R} \times \mathbb{R}^n \times U \to \mathbb{R}^n$ is continuous in all variables and continuously differentiable with respect to x and satisfies

  $$\|f(t, x, \omega)\| \leq C, \|D_x f(t, x, \omega)\| \leq L, \tag{43}$$

  for some constants C, $L \geq 0$ and all $(t, x, \omega) \in \mathbb{R} \times \mathbb{R}^n \times U$.
- Then we have the following global existence and continuous dependence result:

Fundamental properties of trajectories

- **Theorem** (imput-output map)**:** Let the assumption (H^*) be satisfied. Then, for every $T > 0$ and $u \in \mathcal{U}$ the Cauchy problem (41) has a unique solution $x(., u) : [0, T] \to \mathbb{R}^n$.
  The imput-output map $u \to x(., u)$ is continuous from $L^1([0, T], \mathbb{R}^m)$ to $\mathcal{C}([0, T], \mathbb{R}^n)$ (the space of continuous functions from $[0, T]$ into $\mathbb{R}^n$.
- The continuous dependence of trajectories on the control is a useful basic result.
- Consider now a sequence of admissible controls $(u_n)_{n\in\mathbb{N}}$ such that the sequence of the corresponding solutions $(x(., u_n))_{n\in\mathbb{N}}$ converge to $x(.)$ uniformly on $[0, T]$.
- A natural question is whether the function x is a solution of the original problem.

- **Example:** Consider the system

$$x'(t) = u(t), \ x(0) = 0, \ u(t) \in \{-1, 1\} \ a.e. \tag{44}$$

and for $t \in \mathbb{R}$ define

$$u_n(t) = \begin{cases} 1 \text{ if } \sin(nt) \geq 0 \\ -1 \text{ if } \sin(nt) < 0. \end{cases}$$

Then $x(., u_n)$ converges to $x \equiv 0$ uniformly for all $t \in \mathbb{R}$ but $x \equiv 0$ is not a trajectory of (44) .

- We can conclude that the set of trajectories of (44) is not closed.

Fundamental properties of trajectories

- The closure of the set of trajectories of a control system is best studied within the framework of differential inclusions.
- Let $F : \mathbb{R} \times \mathbb{R}^n \to 2^{\mathbb{R}^n}$,

$$F(t, x) = \{f(t, x, \omega) : \omega \in U\} \tag{45}$$

be the multifunction associated to the control system (41) and let

$$x' \in F(t, x) \tag{46}$$

be the differential inclusion defined by F (which is equivalent to to the control system (41), by Filippov's lemma).
- We have the following result
Theorem (*Closure of the set of trajectories*)**:** Assume the multifunction $F : \mathbb{R} \times \mathbb{R}^n \to 2^{\mathbb{R}^n}$ is Hausdorff continuous with compact convex values. Then the set of trajectories of (46) is closed in $C([0, T], \mathbb{R}^n)$.

Fundamental properties of trajectories

- **Corollary**: Let the basic assumption (H) hold. Let $(x_n(.))_{n\in\mathbb{N}}$ be a sequence of solutions to

$$x' = f(t,x,u)\,,\;\; u(t)\in U \tag{47}$$

converging to $x(.)$ uniformly on $[0,T]$. If the graph $\{(t,x(t)) : t\in[0,T]\}$ is entirely contained in Ω and the sets

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

are convex, then $x(.)$ is a trajectory of the control (47)

Fundamental properties of trajectories

- The question here is what happens when the set $\mathcal{U}$ of admissible controls is changed.
- For example, assume that instead of using all measurable controls $u : [0, T] \to [-1, 1]$ we can only use piecewise constant controls, or say, control taking only the two values $+1$, -1.
- Do these limitations substantially reduce our ability to control the system?

  Some results in this direction are presented below,
- **Theorem** (*density of trajectories with piecewise constant controls*) Let f, U satisfy the basic hypothesis (H) . Then the family of trajectories of (47) corresponding to piecewise constant controls is dense in the set of all solutions of (47) (with measurable controls).

- **Proof.** Indeed, given any control $u \in \mathcal{U}$ assume that the corresponding solution $x(., u)$ of

$$x' = f(t, x, u(t)), x(0) = \overline{x}$$

is defined on $[0, T]$. Construct a sequence of piecewise constant controls $u_n \in \mathcal{U}$ converging to u in L^1. If f satisfies the global bounds (H^*), by the continuity of the imput-output map $u \to x(., u)$ the corresponding trajectories $x(., u_n)$ converge to $x(., u)$ uniformly on $[0, T]$.

- To prove that the uniform convergence holds also in the general case, it suffices to consider an auxiliary function f^+ defined by

$$f^+(t, x) = \begin{cases} \phi(t, x) f(t, x, u) & \text{if} \quad (t, x) \in \Omega \\ 0 & \text{if} \quad (t, x) \notin \Omega \end{cases}$$

which satisfies (H^*) and coincide with f for (t, x) in a neighborhood of the graph

$$\{(t, x(t, u)) : t \in [0, T]\}.$$

- The result then holds for the system

$$x' = f^{+}(t, x, u(t)), x(0) = \bar{x}$$

hence for the original system

$$x' = f(t, x, u(t)), x(0) = \bar{x}$$

as well.

- **Theorem (approximation)** Let f, U satisfy the basic hypothesis (H). Let U' be a closed subset of U such that

$$\{f(t, x, \omega) : \omega \in U\} \subset co \left\{f(t, x, \omega) : \omega \in U'\right\} \ \forall (t, x) \in \Omega. \tag{48}$$

Then every trajectory of

$$x' = f(t, x, u(t)), x(0) = \bar{x}, \ u(t) \in U$$

can be uniformly approximated by trajectories of

$$x' = f(t, x, u(t)), x(0) = \bar{x}, \ u(t) \in U'.$$

- **Remark.** An important case where the key assumption (48) holds is the following. Assume that the control system is linear with respect to u :

$$x' = h(t,x) + A(t,x) \cdot u,$$

each $A(t,x)$ being a $n \times m$ matrix.
Let U' be a compact subset of U and call F and F' the corresponding multifunctions:

$$\begin{aligned} F(t,x) &= \{f(t,x,\omega) : \omega \in U\} \text{ and} \\ F'(t,x) &= \{f(t,x,\omega) : \omega \in U'\}. \end{aligned}$$

Then by linearity,

$$U \subset coU' \Longrightarrow F(t,x) \subset coF'(t,x).$$

- This implication, however, is usualy false when f is nonlinear.

- **Example:** Let h, g be smooth vector fields on $\mathbb{R}^n$. Then the set of trajectories of

$$x' = h(x) + u(t)\, g(x), \; x(0) = 0, \; u(t) \in \{-1, 1\} \text{ a.e.}$$

is dense on the set of trajectories of

$$x' = h(x) + u(t)\, g(x), \; x(0) = 0, \; u(t) \in [-1, 1] \text{ a.e.}$$

because

$$co\{-1, 1\} = [-1, 1].$$

- **Example:** On $\mathbb{R}^2$ consider the systems

$$\left(x_1', x_2'\right) = \left(u, 1 - u^2\right), \ u(t) \in [-1, 1], \ (x_1, x_2)(0) = (0, 0) \quad (49)$$

$$\left(x_1', x_2'\right) = \left(u, 1 - u^2\right), \ u(t) \in \{-1, 1\}, \ (x_1, x_2)(0) = (0, 0). \quad (50)$$

Then $coU' = U$, but the set of solutions of (50) is not dense on the set of solutions of (49).
Indeed, taking $u(t) = 0$ one obtain a solution of (49)

$$(x_1, x_2)(t) = (0, t),$$

This solution cannot be approximated by trajectories of (50), because $u(t) \in \{-1, 1\}$ implies $x_2(t) = 0$

Reachable sets, controllability

- For $\Omega \subseteq \mathbb{R}^n$, $U \subseteq \mathbb{R}^m$ and $f : \Omega \times U \to \mathbb{R}^n$ consider the nonlinear control system
$$x' = f(t, x, u), \ \ u(.) \in \mathcal{U} \tag{51}$$
where the set of admissible controls is defined as
$$\mathcal{U} := \{u : [0, t_u] \to U : u \text{ measurable}\}. \tag{52}$$
- For $u \in \mathcal{U}$ and $(0, x_0) \in \Omega$ let $x(., x_0, u)$ be the solution of the Cauchy problem
$$x' = f(t, x, u(t)), \ \ x(0) = x_0. \tag{53}$$
- **Definition.** Let $x_1, x_2 \in \mathbb{R}^n$. We say that x_2 is reachable from x_1 or that x_1 is transferable to x_2 at time t if there exists an admissible control $u : [0, t] \to U$ such that
$$x_2 = x(t, x_1, u).$$

- The set

$$R(t, x_0) = \{x \in \mathbb{R}^n : x = x(t, x_0, u) \text{ for some } u \in \mathcal{U}\}$$

is called the reachable set from x_0 at time t.
- For $K \subseteq \mathbb{R}^n$ we define

$$R(t, K) = \bigcup_{x_0 \in K} R(t, x_0) \tag{54}$$

the set of points reachable at time t from a point of K.

- **Definition** We say that the system (51) is locally controllable at x_0 if x_0 is an interior point of $R(t, x_0)$ for each $t > 0$:

$$x_0 \in int\ R(t, x_0) \text{ for all } t > 0,$$

- We say that the system (51) is globally controllable at x_0 if

$$\bigcup_{t>0} R(t, x_0) = \mathbb{R}^n.$$

Minimum time function

- Let $R(0) = \bigcup\limits_{t>0} R(t,0)$ be the set of all points reachable from the origin $0 \in \mathbb{R}^n$.
- The function $T : \mathbb{R}^n \to [0,\infty]$ defined by

$$T(x) = \begin{cases} \inf \{t \geq 0 : x \in R(t,0)\} & \text{if} \quad x \in R(0) \\ +\infty & \text{if} \quad x \in \mathbb{R}^n \backslash R(0) \end{cases}$$

is said to be the "minimum time function" function.

Minimum time problem

- The **minimum time problem** consists in:
  (i) Determine $T(x)$ for each $x \in R(0)$,
  (ii) Determine $u^* \in \mathcal{U}$ such that

$$x = x(T(x), 0, u^*). \tag{55}$$

- A control $u \in \mathcal{U}$ satisfying (55) is called an optimal control with respect to x.

- **Theorem** (The Bellman principle of optimality) If $u^* \in \mathcal{U}$ is an optimal control with respect to $x_1 \in R(0)$ and if $T_2 < T(x_1)$ and $x_2 = x(T_2, 0, u^*)$ then $u^* \in \mathcal{U}$ is an optimal control also with respect to x_2.
- **Proof:** Assume that is possible to reach x_2 from the origin at time $\widetilde{T} < T_2$, following the trajectory corresponding to a control $\widetilde{u} \in \mathcal{U}$, that is

$$x_2 = x\left(\widetilde{T}, 0, \widetilde{u}\right).$$

The Bellman principle of optimality

- Define the control

$$\widehat{u}(t) = \begin{cases} \widetilde{u}(t) & \text{if} \quad t \in \left[0, \widetilde{T}\right] \\ u^*\left(t + T_2 - \widetilde{T}\right) & \text{if} \quad \widetilde{T} \le t \le T(x_1) - \left(T_2 - \widetilde{T}\right) \end{cases}$$

and observe that the trajectory corresponding to $\widehat{u}$ allows to reach x_1 from the origin 0 at time

$$T(x_1) - \left(T_2 - \widetilde{T}\right) < T(x_1),$$

in contradiction with the definition of $T(x_1)$.

Compactness of the reachable sets

- The following result concerning the compactness of reachable set will be useful to prove existence of optimal controls.
- **Theorem (Compactness).** Let $\Omega \subseteq \mathbb{R} \times \mathbb{R}^n$, $U \subseteq \mathbb{R}^m$ and $f : \Omega \times U \to \mathbb{R}^n$ satisfy the basic hypothesis (H). Assume that there exists a compact set $\widehat{\Omega} \subseteq \Omega$ such that the graphs of all trajectories defined on $[0, T]$ of the control system (51) starting from $x_0 \in \mathbb{R}^n$ are contained in $\widehat{\Omega}$ and that all the sets $F(t, x) = \{f(t, x, \omega) : \omega \in U\}$ are convex, then, for every $\tau \in [0, T]$, the reachable set $R(\tau, x_0)$ is compact.

Mayer optimal control problem

- We consider the control system

$$x' = f(t, x, u), \quad u(.) \in \mathcal{U} \tag{56}$$

where

$$\mathcal{U} := \{u : \mathbb{R} \to \mathbb{R}^m : u \text{ measurable, } u(t) \in U \ \forall t\}. \tag{57}$$

- Given an initial state x_0, a set of admissible terminal conditions $S \subseteq \mathbb{R} \times \mathbb{R}^n$, and a cost function $\phi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ we consider the optimization problem

$$\min \{\phi(T, x(T, u)) : u \in \mathcal{U} \text{ defined on } [0, T], T > 0\} \tag{58}$$

with the initial and terminal constraints

$$x(0) = x_0, \ (T, x(T)) \in S. \tag{59}$$

Mayer optimal control problem

- When as in (58), the performance criterion depends only on terminal time T and on the terminal point $x(T)$ of the trajectory, we say that the problem is in Mayer form.
- Remark that the maximization problem

$$\max \{\psi(T, x(T, u)) : u \in \mathcal{U} \text{ defined on } [0, T], T > 0\}$$

is equivalent to (58) choosing $\phi = -\psi$.

Existence of optimal controls for Mayer problem

- Under suitable hypotheses, we shall prove the existence of an admissible control $u^* \in \mathcal{U}$ whose corresponding trajectory $x^*(.) = x(., u^*)$ satisfies the constraints (58), that is,

$$x^*(0) = x_0, \text{ and } (T, x^*(T)) \in S,$$

and yields the minimum in (58), that is,

$$\phi(T, x^*(T)) = \min\{\phi(T, x(T, u)) : u \in \mathcal{U} \text{ defined on } [0, T], T > 0\}$$

The key assumption will be the convexity of the sets

$$F(t, x) = \{f(t, x, \omega) : \omega \in U\} \tag{60}$$

which guarantees the closure of the set of trajectories of (56).

Existence of optimal controls for Mayer problem

- We start by considering the simplest case, assuming that the terminal time $T > 0$ is fixed, and the problem $(58) - (59)$ takes the form

$$\min \{\phi(x(T,u)) : u \in \mathcal{U}\}, \ x(0) = x_0, \ x(T) \in S,$$

where ϕ is continuous and S is a closed subset of $\mathbb{R}^n$.

- If the assumptions of Theorem giving the compactness of reachable set hold and the set of trajectories reaching S is nonempty, then an optimal control exists.

Existence of optimal controls for Mayer problem

- Indeed, by Theorem on compactness of reachable sets, the set $R(T)$ is compact, hence there exists a point $x_{\min} \in R(T) \cap S$ where ϕ is minimal, that is,

$$\phi(x_{\min}) = \min\{\phi(x(T, u)) : u \in \mathcal{U}\}.$$

- Any control $u^* \in \mathcal{U}$ that steers the system to the point $x_{\min}$, i.e., such that

$$x(T, u^*) = x_{\min}$$

  is clearly optimal.

Existence of optimal controls for Mayer problem

- A similar result holds for the more general problem $(56)-(59)$.
- On the control system (56) we make the following assumptions, somewhat stronger than the assumption (H). We assume: $\left(\widehat{H}\right)$: The set $U \subset \mathbb{R}^m$ of control values is compact, the function $f : \{0, \infty) \times \mathbb{R}^n \times U \to \mathbb{R}^n$ is continuous in all variables, continuously differentiable with respect to to x and satisfies

$$\|f(t,x,u)\| \leq C(1 + \|x\|), \tag{61}$$

for some constants $C \geq 0$ and all $(t,x,u) \in \mathbb{R} \times \mathbb{R}^n \times U$.

Existence of optimal controls for Mayer problem

- Then we have the following existence result:
- **Theorem.** Let the assumption $\left(\widehat{H}\right)$ hold. Assume that the sets of velocities in (60) are convex, the cost function ϕ is continuous, and the target set S is compact and contained in the strip $[0, T] \times \mathbb{R}^n$. If some trajectory $x(.)$ satisfying the constraints (59) exists, then the problem $(56) - (59)$ has an optimal solution.
- **Proof:** By assumption, there exists at least one admissible trajectory reaching the target set S.
  Therefore, we can construct a sequence of controls $u_n : [0, T_n] \to U$ whose corresponding trajectories $x_n(.)$ starting at $\overline{x}$ satisfy

$$(T_n, x_n(T_n)) \in S, \quad \lim_{n\to\infty} \phi(T_n, x_n(T_n)) = \inf_{u,T} \phi(T, x(T, u)). \quad (62)$$

Existence of optimal controls for Mayer problem

- Since $S \subseteq [0, \overline{T}] \times \mathbb{R}^n$, we have $T_n \leq \overline{T}$ for every n. We can now prolong each function x_n to the entire interval $[0, \overline{T}]$ by setting

$$x_n(t) = x_n(T_n) \text{ for } t \in [T_n, \overline{T}].$$

- By the assumption $(\widehat{H})$, all trajectories satisfy the uniform bound

$$|x_n(t, u)| \leq \left(e^{Ct} - 1\right) + e^{Ct} \|x_0\| \text{ for } t \in [0, \overline{T}]. \tag{63}$$

Since f is uniformly bounded on bounded sets, the sequence $x_n(.)$ is uniformly Lipschitz continuous. Using Ascoli's compactness theorem,by possible taking a subsequence, we can assume that

$$T_n \to T^* \text{ for some } T^* \leq \overline{T} \text{ and } x_n(.) \to x^*(.) \text{ uniformly on } [0, T^*].$$

Existence of optimal controls for Mayer problem

- Because if the convexity of the sets $F(t,x)$, we have that $x^*(.)$ is an admissible trajectory of (56), that is, there exists a control $u^* : [0, T^*] \to U$ such that

$$\frac{dx^*(t)}{dt} = f(t, x^*(t), u^*(t)) \text{ for a.e. } t \in [0, T^*].$$

Clearly $x^*(0) = x_0$. Since S is closed, the first relation in (62) implies

$$(T^*, x^*(T^*)) = \lim_{n\to\infty} (T_n, x_n(T_n)) \in S.$$

- Finally, from (62) and the continuity of ϕ it follows

$$\phi(T^*, x^*(T^*)) = \lim_{n\to\infty} \phi(T_n, x_n(T_n)) = \inf_{u,T} \phi(T, x(T,u))$$

therefore, the control u^* is optimal.

- **Remark:** The above proof is a typical example of the Direct Method for proving existence of optimal solutions. The basic steps are:
  1. Construct a minimizing sequence $x_n(.)$;
  2. Show that some subsequence converge to a function $x^*(.)$;
  3. Prove that $x^*(.)$ is an admissible trajectory and satisfies the bondary conditions;
  4. Prove that $x^*(.)$ attainsthe minimum value of the optimization probem.

- **Remarks:**

  1. The assumption that S is compact is used to ensure that the domains $[0, T_n]$ are uniformly bounded. The theorem still holds if S is closed and $\phi(T, x) \to \infty$ as $T \to \infty$.
  2. The continuity of ϕ can be replaced by lower semicontinuity.

Necessary conditions

- The aim of this section is to derive necessary conditions in order that a trajectory $x^* = x(., u^*)$ be optimal for the Mayer problem

$$\max \{\psi(x(T, u)) : u \in \mathcal{U}\} \tag{64}$$

subject to

$$x' = f(t, x(t), u(t)), x(0) = \overline{x} \tag{65}$$

the set of admissible controls being

$$\mathcal{U} := \{u : [0, T] \to U : u \text{ measurable}\}. \tag{66}$$

- As usual $\Omega \subseteq \mathbb{R} \times \mathbb{R}^n$ is an open set, $U \subseteq \mathbb{R}^m$, $f : \Omega \times U \to \mathbb{R}^n$ is continuous on $\Omega \times U$ with continuous Jacobian matrix $D_x f$ of first order derivatives with respect to x. The function $\psi : \mathbb{R}^n \to \mathbb{R}$ is differentiable. No assumption is made on the set $U \subseteq \mathbb{R}^m$. In particular we may well have $U = \mathbb{R}^m$.

Pontryagin maximum principle

- For the above problem, where the terminal time T is fixed and no constraints are imposed on the terminal point the Pontryagin maximum principle (for short, PMP)can be stated in a particularly simple form.
- **Theorem (**PMP, free terminal point**)** In connection with the maximization problem $(64) - (66)$, let u^* be a bounded admissible control whose corresponding trajectory $x^* = x\left(., u^*\right)$ is optimal. Call $p : [0, T] \rightarrow \mathbb{R}^n$ be the solution of the adjoint linear equation

$$p'(t) = -p(t) \cdot D_x f\left(t, x^*(t), u(t)\right), \;\; p(T) = \nabla\psi\left(x^*(T)\right). \tag{67}$$

Then the maximality condition

$$p(t) f\left(t, x^*(t), u^*(t)\right) = \max\left\{p(t) f\left(t, x^*(t), \omega\right) : \omega \in U\right\} \tag{68}$$

holds at almost every time $t \in [0, T]$.

- In the above theorem x, f, v represent column vectors, $D_x f$ is a $n \times n$ matrix, while p is a row vector.

- In coordinates, the equations (67) and (68) can be rewritten as

$$p_i'(t) = -\sum_{j=1}^{n} p_j(t) \frac{df_j(t)}{dx_i}(t, x^*(t), u^*(t)), \;\; p_i(T) = \frac{d\psi}{dx_i}(x^*(T)) \tag{69}$$

$$\begin{aligned} &\sum_{i=1}^{n} p_i(t) f_i(t)(t, x^*(t), u^*(t)) = \\ &= \max\left\{\sum_{i=1}^{n} p_i(t) f_i(t)(t, x^*(t), \omega) : \omega \in U\right\} \end{aligned} \tag{70}$$

- **Remark.** The pevious theorem remain valid if one replace max by min .

The previous result motivates a practical method for finding optimal solutions for the problem $(64)-(66)$:

1. First, define the function $u = u(t,x,p) \in U$ by mean of the equality

$$pf(t,x,u(t,x,p)) = \max\{pf(t,x,\omega) : \omega \in U\} \tag{71}$$

2. Then solve the system of $2n$ equations

$$\begin{cases} x' = f(t,x,u(t,x,p)) \\ p' = -pD_x f(t,x,u(t,x,p)) \end{cases} \tag{72}$$

   with boundary conditions

$$x(0) = \overline{x}, \;\; p(T) = \nabla\psi(x(T)). \tag{73}$$

3. By the previous theorem, if a bounded optimal control exists, it must be found among the solutions of $(72)-(73)$.

Remarks

In general this method encounters two main dificulties:

1. The map $(t, x, p) \to u(t, x, p)$, implicitely defined by the maximality condition (71), may be multivalued or discontinuous.
2. The problem $(72) - (73)$ is not a Cauchy problem but a (usually harder) two points boundary value problem. Indeed, the initial value $p(0)$ is not explicitely known. Instead, an equation is given involving the terminal values $p(T)$ and $x(T)$.
3. In particular cases, the equations for p and x can be uncoupled, and a solution is easily found

Example: linear pendulum with external force

- Let q be the position of a linearized pendulum, controlled by an external force u, with magnitude constraint
$$|u(t)| \leq 1, \forall t.$$
- Assuming that the initial position and velocities are both zero, the motion is determined by the quations
$$q''(t) + q(t) = u(t), \ \ q(0) = q'(0) = 0.$$
- We wish to maximize the displacement $q(T)$ at a fixed terminal time T.
- Introducing the variables $x_1 = q$, $x_2 = q'$, the opimization problem can be formulated as
$$\max \{x_1(T, u) : u \in \mathcal{U}\}$$

where the dynamics is described by
$$\begin{cases} x_1' = x_2 & x_1(0) = 0 \\ x_2' = -x_1 + u & x_2(0) = 0 \end{cases}$$

- The set of admissible controls is

$$\mathcal{U} = \{u : [0, T] \rightarrow [-1, 1] : u \text{ measurable}\}$$

- In this case, the adjoint equations (70) take the form

$$\begin{cases} p_1' = p_2 & p_1(T) = 1 \\ p_2' = -p_1 + u & p_2(T) = 0. \end{cases}$$

- These equations can be solved for p independently of x, yielding

$$p_1(t) = \cos(T - t), \; p_2(t) = \sin(T - t).$$

- By (71), the optimal control u^*satisfies

$$p_1 x_2 + p_2(-x_1 + u^*) = \max\{p_1 x_2 + p_2(-x_1 + \omega) : |\omega| \leq 1\}.$$

- Therefore, the optimal control is given by

$$u^*(t) = sign(p_2(t)) = sign(\sin(T - t)).$$

References

- J. P. Aubin and A. Cellina, Differential inclusions, Springer-Verlag, 1984;
- A. Bressan and B. Picoli, Introduction to the Mathemtical Theory of Control, AIMS Series in Applied Mathematics, 2007;
- F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, P. R. Wolenski, Nonsmooth analysis and Control Theory, Springer-Verlag, New York, 1998.
- A. F. Filippov, Differential equations with discontinuous right-hand side, Kluwer, Doldrecht, 1988.
- L. Cesari, Optimisation - Theory and Applications, Springer, 1983;

References

- E. B. Lee and L. Marcus, Foundations of Optimal control theory, Willey, 1967;
- P. D. Loewen, Optimal control via nonsmooth analysis, C. R. M. Proceedings & Lecture Notes, Amer. Math. Soc., Providence, 1993.
- J. Macki, A. Strauss, Introduction to optimal control theory, SpringerVerlag, New York, 1982.
- L. S. Pontryagin, V. Boltyanskii, R. V. Gamkrelidze, E. F. Mishenko, The Mathematical Teory of Optimal processes, Wiley, 1962