Viscosity solutions and optimal control problems

Tien Khai Nguyen, Department of Mathematics, NCSU.

1 Optimal control

1.1 Ordinary differential equations and control dynamics

1. A brief review on ordinary differential equations. The time evolution of a system, whose state is described by a finite number of parameters, can be usually modeled by an O.D.E.

$$\begin{cases} \dot{x}(t) \;=\; f(x(t)), \qquad a.e.\ t \in [0, +\infty[, \\ \\ x(0) \;=\; x_0, \end{cases} \tag{1.1}$$

where

- $x : [0, +\infty) \to \mathbb{R}^n$ is the state variable depended on time t;
- $f : \mathbb{R}^n \to \mathbb{R}^n$ is the dynamics;
- x_0 is the initial state.

Definition 1.1 *(Absolutely continuous) A map $x : [a, b] \to \mathbb{R}^n$ is absolutely continuous if for every $\varepsilon, \delta > 0$ such that whenever a finite sequence of pairwise disjoint sub-intervals $(s_k, t_k) \subset [a, b]$ for $k = 1, 2, ...n$ satisfies*

$$\sum_{k=1}^{n} |t_k - s_k| \;\leq\; \delta$$

then it holds

$$\sum_{k=1}^{n} |x(t_k) - x(s_k)| \;\leq\; \varepsilon\,.$$

Denote by

$$AC([a, b], R^n) \;\doteq\; \{x : [a, b] \to \mathbb{R}^n \mid x \text{ is absolutely continuous}\}\,.$$

Notice that every Lipschitz function $x : [a, b] \to \mathbb{R}^n$ is in $AC([a, b], R^n)$. However, the converse of this statement is false in general. Indeed, the followings hold:

Lemma 1.2 *For any $x(\cdot) \in AC([a, b], R^n)$, it holds that it derivative $\dot{x}$ is almost everywhere defined on $[a, b]$ and*

$$x(t) \;=\; x(t_0) + \int_{t_0}^{t} \dot{x}(s)\; ds \qquad \textit{for all } t_0, t \in [a, b]\,.$$

Conversely, given a function $g \in \mathbf{L}^1([a,b],\mathbb{R}^n)$*, the function* $y : [a,b] \to \mathbb{R}^n$ *which is defined by*

$$y(t) \;=\; y(a) + \int_a^t g(s)\; ds \qquad \textit{for all } t \in [a,b]$$

belongs to $AC([a,b],R^n)$ *and*

$$\dot{y}(t) \;=\; g(t) \qquad \text{for } a.e.\ t \in [a,b]\,.$$

Roughly speaking the lemma establishes that a map is absolutely continuous if and only if it coincides with the integral of its derivative. Hence, one could provide an alternative definition for absolutely continuous functions.

Definition 1.3 *(Absolutely continuous) A map* $x : [a,b] \to \mathbb{R}^n$ *is absolutely continuous if and only if* x *is differential almost everywhere on* $[a,b]$ *and*

$$\dot{x}(t) \;=\; x(a) + \int_a^t \dot{x}(s)\; ds\,.$$

In general, the continuity and the almost everywhere differentiability are not sufficient to guarantee the absolute continuity. Indeed, it is well-known that one can

Problem 1. *Construct a (uniformly) continuous and strictly increasing function* $z : [a,b] \to \mathbb{R}$ *such that* z *is differentiable and equal to zero almost everywhere.*

Thus, $z(b) - z(a) > 0 = \displaystyle\int_a^b \dot{z}(s)\; ds$ and it yields that z is not absolutely continuous.

Definition 1.4 (Carathéodory solution) *Given a vector field* $f : \mathbb{R}^n \to \mathbb{R}^n$*, a map* $x : [a,b] \to \mathbb{R}^n$ *is a Carathéodory solution to the ordinary differential equation*

$$\dot{x}(t) \;=\; f(x(t))$$

on $[a,b]$ *if* $x(\cdot)$ *is absolutely continuous and*

$$x(t) \;=\; x(a) + \int_a^t f(x(s))\; ds \qquad \textit{for all } t \in [a,b]\,.$$

It is clear that if $x(\cdot)$ is a Carathéodory solution of the ODE (1.1) then all $a \le t_1 < t_2 \le b$, it holds

$$x(t_2) \;=\; x(t_1) + \int_{t_1}^{t_2} f(x(s))\; ds\,.$$

In particular, it yields the semigroup property

$$x(t+s) \;=\; x(t) \circ x(s) \qquad \text{for all } a \le s, t \le s+t \le b$$

where $x(t) \circ x(s)$ is the value of the solution of the ODE (1.1) with initial data $x(s)$ at time t.

Theorem 1.5 *(Existence result) Assume that the dynamics f is uniformly Lipschitz, i.e.,*

$$\|f(y) - f(x)\| \;\leq\; L_f \cdot \|y - x\| \qquad \textit{for all } x, y \in \mathbb{R}^n$$

for some constant $L_f > 0$. For any initial data $x_0 \in \mathbb{R}^n$, the ODE (1.1) admits a Carathéodory solution.

Let us now introduce an useful lemma which allows to derive the stability result for the ODE (1.1).

Lemma 1.6 *(Gronwall's inequality) Let $z : [0, T] \to [0, +\infty)$ be an absolutely continuously function such that*

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

and $z(0) = z_0$. Then it holds

$$z(t) \;=\; z_0 \cdot e^{\int_0^t \alpha(s)\ ds} + \int_0^t \beta(s) \cdot e^{\int_s^t a(\tau)\ d\tau}\ ds\,.$$

As a consequence of the above lemma, we have the following stability results of the ODE (1.1).

Proposition 1.6.1 *Under the Lipschitz continuity assumption on f in theorem 1.27, the followings hold:*

(i). (Boundedness) For any given initial data x_0, let $y^{x_0}(t)$ be the solution of (1.1). Then

$$\|y^{x_0}(t)\| \;\leq\; \|x_0\| + \|f(x_0)\| \cdot e^{L_f t}\,. \tag{1.2}$$

(ii) (Stability) Given any $x_1, x_2 \in \mathbb{R}$, it holds

$$\|y^{x_2}(t) - y^{x_1}(t)\| \;\leq\; e^{L_f t} \cdot \|x_2 - x_1\| \tag{1.3}$$

for all $t > 0$. In particular, the ODE (1.1) admits a unique solution.

Proof. (i). For any $t > 0$, it holds

$$\frac{d}{dt}\|y^{x_0}(t) - x_0\| \;\leq\; \|f(y^{x_0}(t))\| \;\leq\; \|f(x_0)\| + L_f \cdot \|y^{x_0}(t) - x_0\|.$$

Thus, the Gronwall's inequality implies that

$$\|y^{x_0}(t) - x_0\| \;\leq\; \|f(x_0)\| \cdot \frac{e^{L_f \cdot t}}{L_f}$$

and this yields (1.2)

$$\|y^{x_0}(t)\| \;\leq\; \|x_0\| + \|f(x_0)\| \cdot \frac{e^{L_f \cdot t}}{L_f}.$$

(ii). Similarly, one estimates

$$\frac{d}{dt}\|y^{x_2}(t) - y^{x_1}(t)\| \;\leq\; \|f(y^{x_2}(t)) - f(y^{x_1}(t))\| \;\leq\; L_f \cdot \|y^{x_2}(t) - y^{x_1}(t)\|$$

and the Gronwall's inequality yields (1.3). □

The first order tangent vector. Let $x(t)$ be the solution of the ODE

$$\dot{x}(t) \;=\; f(x(t))$$

Consider a family of nearby solutions, says $t \to x_\varepsilon(t)$. Assume that a given time $t = 0$, one has

$$v_0 \;=\; \lim_{\varepsilon \to 0} \frac{x_\varepsilon(0) - x(0)}{\varepsilon}\,.$$

Proposition 1.6.2 *In addition to the uniformly Lipschitz continuous on f, we assume that f is also continuously differentiable. Then the first order tangent vector*

$$v(t) \;\doteq\; \lim_{\varepsilon \to 0} \frac{x_\varepsilon(t) - x(t)}{\varepsilon}$$

is well defined for all $t \in [0, T]$. Morever, $v(t)$ is the solution of the affine ODE

$$\dot{v}(t) \;=\; Df(x(t)) \cdot v(t), \qquad\qquad v(0) \;=\; v_0\,. \tag{1.4}$$

Using the Laudau notation, we can write

$$x_\varepsilon(t) \;=\; x(t) + \varepsilon \cdot v(t) + o(\varepsilon)$$

where $o(\varepsilon)$ denotes an infinitesimal of higher order with respect to ε. Therefore, one can formally write

$$\dot{x}_\varepsilon(t) \;=\; \dot{x}(t) + \varepsilon \cdot \dot{v}(t) + o(\varepsilon) \;=\; f(x_\varepsilon(t)) \;=\; \dot{x}(t) + Df(x(t)) \cdot \varepsilon v(t) + o(\varepsilon)\,.$$

Adjoint system. It is useful to consider the adjoint system of (1.4)

$$\dot{p}(t) \;=\; -\,p(t) Df(x(t)) \tag{1.5}$$

where $p(t)$ is a row vector. A direct computation yields

$$\begin{aligned}\frac{d}{dt}\,[p(t) \cdot v(t)] \;&=\; \dot{p}(t) \cdot v(t) + p(t) \cdot \dot{v}(t)\\ &=\; -p(t) \cdot Df(x(t))v(t) + p(t)Df(x(t))v(t) \;=\; 0\,.\end{aligned}$$

This implies that then the product $p(t) \cdot v(t)$ is constant in time.

2. Control systems. In some cases, the system can be in influenced also by the external input of a controller. An appropriate model is then provided by a control system, having the form

$$\begin{cases} \dot{x}(t) &= f(x,u), \\ \\ x(0) &= x_0 \,. \end{cases} \tag{1.6}$$

Here $x_0 \in \mathbb{R}^n$ is the initial state and

- $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ is the dynamics of the control system
- $U \subset \mathbb{R}^m$ is the control set
- $u : [0, +\infty[\to U$ is a control function.

Remark. *If we set*

$$F(x) \doteq f(x,U) = \{f(x,u) \mid u \in U\} \,,$$

then the control system (1.6) can be rewritten as an differential inclusion

$$\dot{x} \in F(x), \qquad\qquad x(0) = x_0 \,.$$

There are two types of control:

- If $u = u(t)$ is assigned as a function of time, we say that u is an open-loop control.
- If $u = u(x)$ is assigned as a functions of state variable u, we say that u is a closed-loop (feedback) control,

Let us first consider the open-loop controls. We will write

$$f(x,u) = \begin{pmatrix} f_1(x,u) \\ \vdots \\ f_n(x,u) \end{pmatrix} \qquad \text{and} \qquad x(t) = \begin{pmatrix} x^1(t) \\ \vdots \\ x^n(t) \end{pmatrix}.$$

The set of admissible controls is denoted by

$$\mathcal{U}_{ad} := \big\{ u : [0, +\infty[\to U \mid u \text{ is measurable } \big\}, \tag{1.7}$$

we will also write that

$$u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{pmatrix}.$$

Differently from ODE (1.1), a solution of the control system (1.6) depends on initial state x_0 and the choice of admissible control u.

Definition 1.1 *Given any $x_0 \in \mathbb{R}^n$ and $u \in \mathcal{U}_{ad}$, a solution of (1.6) denoted by $y^{x_0,u}(\cdot)$ is called a trajectory of* (1.6) *starting from x_0 associated with the control u.*

We shall assume that our control system satisfies the following standard hypotheses:

STANDARD HYPOTHESES (F)

(F1). The control set U is compact.

(F2). The function f is continuous. Moreover, there exists a constant $K_1 > 0$ such that

$$|f(y,u) - f(x,u)| \leq L_f \cdot |y - x|, \quad \text{for all } x, y \in \mathbb{R}^n, u \in U.$$

The following holds:

Theorem 1.7 *Under assumption* **(F)**, *given any initial data x_0 and admissible control $u \in \mathcal{U}_{ad}$, the ODE (1.6) admits a unique absolute continuous solution denote by $y^{x_0,u}$ such that*

$$y^{x_0,u}(t) \;=\; x_0 + \int_0^t f(y^{x_0,u}(s), u(s))\, ds \qquad \textit{for all } t \in [0, +\infty)\,.$$

Moreover, the followings hold:

(i) (Boundedness) For any $x_0 \in \mathbb{R}^n$ and $t > 0$,

$$\|y^{x_0,u}(t)\| \;\leq\; e^{L_f \cdot t} \cdot \|x_0\| + \frac{e^{L_f \cdot t} - 1}{L_f} \cdot M$$

with $M = \max_{u \in U} |f(0,u)|$.

(ii) (Stability) For any $x_1, x_2 \in \mathbb{R}^n$ and $t > 0$, the distance between $y^{x_1,u}(t)$ and $y^{x_2,u}(t)$

$$\|y^{x_1,u}(t) - y^{x_2,u}(t)\| \;\leq\; e^{L_f \cdot t} \cdot \|x_1 - x_2\|.$$

Let's introduce the cost function $P : \mathbb{R}^n \times \mathcal{U}_{ad} \times AC([0,\infty), \mathbb{R}^n) \to \mathbb{R}$ which depends on an initial data $x_0 \in R^n$ and an admissible control $u \in \mathcal{U}_{ad}$.

Optimization problem: *Our goal is to seek for an optimal control $u^* \in \mathcal{U}_{ad}$ which minimizes the cost function among all admissible controls, i.e.,*

$$P\left[x_0, u^*, y^{x_0,u^*}\right] \;\leq\; P\left[x_0, u, y^{x_0,u}\right] \qquad\qquad u \in \mathcal{U}_{ad}\,.$$

The problem of

$$Minimizing_{u \in \mathcal{U}_{ad}}\; P\left[x_0, u\right] \quad \text{subject to the control system (1.6)}\,.$$

is called an optimal control problem.

1.2 Optimal control problems.

1.2.1 Standard problems

1. The minimum time problem. The aim of this problem is to minimize the amount of time for the system to reach a given target set $\mathcal{T}$ which is closed subset of $\mathbb{R}^n$. More precisely, for a fixed initial data $x_0 \in \mathbb{R}^n \backslash \mathcal{T}$, denote by

$$\theta(x_0, u) \doteq \min\{t \geq 0 \mid y^{x_0,u} \in \mathcal{T}\}$$

Of course, $\theta(x_0, u)$ is in $[0, +\infty]$, and $\theta(x_0, u)$ is the time taken for the trajectory $y^{x_0,u}$ to reach the target $\mathcal{T}$, provided $\theta(x_0, u) < +\infty$. *The minimum time* $T(x_0)$ to reach the target $\mathcal{T}$ for x_0 is defined by

$$T(x_0) = \inf_{u \in \mathcal{U}_{ad}} \theta(x_0, u) . \tag{1.8}$$

In general, $T(x_0)$ can be $+\infty$, i.e., the point x_0 can not reach to the target $\mathcal{T}$ from the dynamics (1.6). For a fixed time $t > 0$, denote by

$$\mathcal{R}(t) = \{x \in \mathbb{R} \mid T(x) \leq t\}$$

the set of point can reach the target before time t. It is important to consider the *reachable set*

$$\mathcal{R} = \bigcup_{t>0} \mathcal{R}(t)$$

the set of point which can reach to the target in finite time.

Some basic questions:

- **(Controllability)** Given a point $x_0 \in \mathbb{R}^n \backslash \mathcal{T}$, does x_0 belong to $\mathcal{R}$?
- **(Existence and uniquiness)** Given $x_0 \in \mathcal{R}$, is there an admissible control u^* such that
$$\theta(x_0, u) = T(x_0).$$
If the above equality hold, u^* is called an optimal control steering x_0 to the target $\mathcal{T}$ in a minimum amount of time. Is u^* unique?
- (**Necessary conditions)** Can we construct optimal controls by deriving a set of necessary conditions and compute T?
- **(Regularity theory)** Study the regularity properties of the minimum time function T.

Let us consider a simple example.

Example 1: (Rocket railroad car) Imagine a railroad car powered by rocket engines on each side. We introduce the variables

- $x(t)$ is the position of the rocket railroad car on the train track at time t
- $v(t)$ is the velocity of the rocket rail road car at time t
- $F(t)$ is the force from the rocket engines at time t

where $-1 \leq F(t) \leq 1$ and the sign of $F(t)$ depends on which engine is firing.

Our goal: *is to construct $F(\cdot)$ in order to drive the rocket railroad car to the origin 0 with zero velocity in a minimum amount of time.*

Mathematical model: Assuming that the rocket railroad car has mass m, the motion of law is

$$\ddot{x}(t) = \frac{F(t)}{m} := u(t) \tag{1.9}$$

where $u(\cdot)$ is understood as a control function. For simplicity, we will also assume that $m = 1$. The motion equation of the rocket car is

$$\begin{cases} \ddot{x}(t) &= u(t), \\ x(0) &= x_0 \quad \text{and} \quad v(0) = v_0 \end{cases} \tag{1.10}$$

where $u(\cdot) \in \mathcal{U} = [-1, 1]$, x_0 is the position of the rocket railroad car at time 0 and v_0 is the velocity of the rocket railroad car at x_0. By setting

$$z(t) = \begin{pmatrix} x(t) \\ v(t) \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.11}$$

we can rewrite (1.10) as the first order control system:

$$\begin{cases} \dot{z}(t) &= A \cdot z(t) + u(t) \cdot b \\ z(0) &= z_0 \doteq (x_0, v_0)^T. \end{cases} \tag{1.12}$$

The cost function is

$$P[z_0, u(\cdot)] = \int_0^{\theta(z_0,u)} 1 \, ds = \theta$$

where $\theta(z, u)$ is the first time such that $z(\theta) = (0, 0)^T$.
The goal is to find $u^* \in \mathcal{U}_{ad}$ such that

$$P[z_0, u^*(\cdot)] \ \leq \ P[z_0, u(\cdot)], \quad \text{for all } u \in \mathcal{U}_{ad}.$$

In this case, $P[z, u^*(\cdot)] = T(z_0)$ is the minimum time needed to stear z_0 to $(0, 0)^T$.

Problem 2. Prove that the set $\mathcal{R}(t)$ is convex and compact.

Problem 3. *Identify the reachable set $\mathcal{R}$.*

Problem 4. *Given a point $x \in \mathbb{R}^2 \backslash \{0\}$, is there a unique optimal control?*

Problem 5. *Compute the minimum time function T.*

2. The finite time horizon problem: Bolza and Mayer problems. Given $x_0 \in \mathbb{R}^n$ and control $u \in \mathcal{U}_{ad}$, consider the cost functional

$$P[x_0, u, T] \doteq \int_0^T r(u(t), y^{x_0,u}(t))\, dt + g(y^{x_0,u}(T)) \tag{1.13}$$

where

- $r : U \times \mathbb{R}^n \to \mathbb{R}$ is the running cost;
- $g : \mathbb{R}^n \to \mathbb{R}$ is the terminal cost;
- T is the terminal time.

The problem of

$$\text{minimizing}_{u\in\mathcal{U}_{ad}} \; P[x_0, u, T] \quad \text{subjects to the system (1.6)} \tag{\textbf{BP}}$$

is called a *Bolza problem.*

In particular, if the running cost $r \equiv 0$ then **(BP)** becomes

$$\text{minimizing}_{u\in\mathcal{U}_{ad}} \; g(y^{x_0,u}(T)) \quad \text{subjects to the system (1.6)} \tag{\textbf{MP}}$$

and is called a *Mayer problem.*

Problem 6. *Can one rewrite a Bolza problem as a Mayer problem?*

Goal: *For an given initial data x_0 and a terminal time $T > 0$, a natural question is to seek for an optimal control u^* which minimizes the cost function $P[x_0, u, T]$.*

If an optimal control u^* does exist, the value function is denoted by

$$V(T, x_0) \doteq \inf_{u\in\mathcal{U}_{ad}} P[x_0, u, T] = P[x_0, u^*, T]\,.$$

Example 2: *(A classical problem in calculus of variations)* Consider a linear control system

$$\begin{cases} \dot{x}(t) = u(t) \quad a.e.\ t \in [0, T] \\ \\ x(0) = \bar{x}, \end{cases} \tag{1.14}$$

where $x : [0, +\infty[\to \mathbb{R}^n$ and $u(\cdot) \in \mathcal{U}_{ad}$. Here, the admissible control $\mathcal{U}_{ad}^T$ is denoted by

$$\mathcal{U}_{ad}^T = \left\{ u : [0,T] \to \mathbb{R}^n \mid u \in \mathbf{L}^1_{loc}([0,T], \mathbb{R}^n) \right\}. \tag{1.15}$$

In this case, the set of admissible trajectories is

$$\mathcal{A}_{T,\bar{x}} = \left\{ y(\cdot) \in AC([0,T], \mathbb{R}^n) \mid y(0) = \bar{x} \right\}$$

which is the set of all absolutely continuous functions defined on the interval $[0,T]$ with initial state $\bar{x}$.

Let us now introduce

$$L : \mathbb{R}^n \to \mathbb{R} \times \mathbb{R}^n \quad \text{and} \quad g : \mathbb{R}^n \to \mathbb{R} \tag{1.16}$$

are respectively the continuous *running cost (Lagrangian)* and the continuous *terminal cost*. A classical problem in calculus of variations

$$Minimize_{u \in \mathcal{U}_{ad}^T} \int_0^T L(y^{u,\bar{x}}(t), u(t))\, dt + g(y^{u,\bar{x}}(T))\,. \tag{1.17}$$

Example 3: (Minimal surfaces of revolution) Consider in the space $\mathbb{R}^3$ the two circles

$$\begin{cases} z^2 + y^2 = R_1 \\ \\ x = a_1 \end{cases} \quad \text{and} \quad \begin{cases} z^2 + y^2 = R_2 \\ \\ x = a_2 \end{cases} \tag{1.18}$$

where $a_1 < a_2$. Given Let $\mathcal{A}_{ad}$ be the set of functions $\xi : [a_1, a_2] \to \mathbb{R}^3$ such that $\xi(x) = \begin{pmatrix} x \\ 0 \\ r(x) \end{pmatrix}$ where $r(\cdot) : [a_1, a_2] \to \mathbb{R}^+$ is smooth and satisfies that $r(a_1) = R_1$ and $r(a_2) = R_2$. For each $\xi \in \mathcal{A}_{ad}$, we denote by

$$S_\xi = \left\{ (x,y,z)^T \;\middle|\; a_1 \le x \le a_2, z^2 + y^2 = r(x) \right\} \tag{1.19}$$

the surface of revolution generated by ξ. The area of S_ξ is

$$Area(S_\xi) = 2\pi \int_{a_1}^{a_2} r(x)\sqrt{r'(x)^2 + 1}\, dx. \tag{1.20}$$

Our goal: Finding $\xi^*(\cdot) \in \mathcal{A}_{ad}$ such that

$$Area(S_{\xi^*}) \le Area(S_\xi) \quad \text{for all } \xi(\cdot) \in \mathcal{A}_{ad}. \tag{1.21}$$

We can reformulate the problem into a control problem. Indeed, we consider the constant control system

$$\begin{cases} \dot{r}(t) \; = \; u(t), \\ \\ r(a_1) \; = \; R_1, \end{cases} \tag{1.22}$$

where $u(\cdot) \in \mathcal{U}_{ad}$ which is denoted by

$$\mathcal{U}_{ad} = \left\{ u \in C^1([a_1, a_2], \mathbb{R}^+) \;\Big|\; \int_{a_1}^{a_2} u(s)ds = R_2 - R_1 \right\}. \tag{1.23}$$

The payoff functional is

$$P[u(\cdot)] = 2\pi \int_{R_1}^{R_2} r(s)\sqrt{1 + u^2(s)}\; ds. \tag{1.24}$$

The goal is to find $u^*(\cdot) \in \mathcal{U}_{ad}$ such that

$$P[u^*(\cdot)] \leq P[u(\cdot)] \quad \text{for all } u \in \mathcal{U}_{ad}. \tag{1.25}$$

3. The infinite horizon problem with discount. Given $x_0 \in \mathbb{R}^n$ and control $u \in \mathcal{U}_{ad}$, consider the infinite horizon cost functional with discount

$$J[x_0, u] \;=\; \int_0^{+\infty} e^{-\lambda \cdot t} \cdot L(y^{x_0,u}(t), u(t))\; dt \tag{1.26}$$

where $\lambda > 0$ is a given *discount rate* and L is the *running cost* fulfills the following assumption:

(L1) The function $L : \mathbb{R}^n \times U \to \mathbb{R}$ continuous bounded and continuous, more precisely there exist a modulus $\omega_L(\cdot)$ and a constant M_L such that

$$|L(x,u) - L(y,u)| \;\leq\; \omega_L(|x-y|) \qquad \text{and} \qquad |L(x,u)| \;\leq\; M_L$$

for all $x, y \in \mathbb{R}^n$ and $u \in U$.

Our goal *is seek for an optimal control u^* which minimizes the cost functional. If u^* does exists, one needs to calculate the value function*

$$V(x_0) \;\doteq\; minimize_{u\in\mathcal{U}_{ad}}\; J[x,u]\,.$$

Example 4. *(Optimal harvesting of renewable natural resources)* Denote by $x(t)$ the size of fish population at time t, subject to harvesting. This evolves according to the ODE

$$\dot{x} \;=\; \alpha x(M-x) - bxu, \qquad u(t) \in [0, u_{\max}]\,.$$

Here $\dot{x}(t) = \frac{d}{dt}x(t)$ is the derivative with respect to time t and

- M describes the maximum population sustained by the habitat;
- α is is a reproduction rate;
- b measures the efficiency the harvesting effort;

- The control $u(t)$ accounts for the fishing effort, while the product $bx(t)u(t)$ is the actual catch at time t.

We consider the optimal harvesting problem in infinite time horizon, exponentially discounted:

$$maximize: \quad \int_0^\infty e^{-\gamma \cdot t} \cdot [px(t) - cu(t)]\, dt$$

where p is the market price of fish and c is the unitary cost of the harvesting effort.

Main questions:

- *What is the best harvesting strategy? More precisely, how should the fishing effort $u = u(x)$ depend on the current population size x, in order to achieve the maximum profit, over time?*

- *Study what happens to the population size as $t \to \infty$, when this optimal harvesting policy is implemented, How does this limit depend on the coefficients α, γ and c?*

Problem 8. *Can one remove some of the constants by rescaling variables. Namely, assume*

$$y \;=\; c_1 x, \qquad \tau \;=\; c_2 t \qquad \text{and} \qquad v \;=\; c_3 u\,.$$

Rewrite the ODE in terms of the new variables. Choose the constants c_1, c_2, c_3 so that the new equations become

$$\frac{d}{d\tau} y \;=\; y(1-y) - yv, \qquad v \in [0, v_{\max}] = [0, c_3 u_{\max}]\,.$$

Basing on the reformulated problem, one needs to

Problem 9. *Write an ODE satisfied by the value function*

$$V(y) \;=\; [\textit{maximum payoff that can be achieved when the initial population is } x(0) = y]$$

and solves it.

1.2.2 Non-standard problems

1. A producer vs. consumer games. Consider a game between a producer and a consumer. In this model, the following variables are all functions of time $t > 0$

- R = amount of product still unsold, held in reserve,

- p = unit price,

- u_1 = production rate,
- u_2 = consumption rate.

The evolution equations are

$$\begin{cases} \dot{R} & = & u_1 - u_2 \,, \\ \dot{p} & = & -\ln\left(\frac{R}{R_0}\right) p \,. \end{cases} \tag{1.27}$$

The first equation simply says that the amount of product in stock changes depending of the difference between production and consumption. By the second equation, there is a level R_0 of reserves which is considered "appropriate". If the reserves fall below R_0, shortages are predicted and the price increases. If the reserves increase above R_0, a glut is expected and the price decreases The payoff functionals, in infinite time horizon, are

$$\begin{aligned} J_1 &= \int_0^\infty e^{-\gamma t}[pu_2 - c(u_1)]\,dt\,, \\ J_2 &= \int_0^\infty e^{-\gamma t}[\varphi(u_2) - pu_2]\,dt\,. \end{aligned} \tag{1.28}$$

What is the optimal solution to the problem? This question leads to the concept of Nash equilibrium and a study of system of PDEs.

2. Optimal debt management. An accurate description of the debt structure would require a knowledge of the size of the various loans, the interest rate charged on each loan, and the expiration date of each loan. To simply the models, let us denote by

- $x(t)$ = debt size
- $u(t) \in [0,1]$ = payment rate, as a fraction of the income
- $L(u)$ = = cost to the borrower for implementing the control
- B = bankruptcy cost to the borrower.

Given initial debt $\bar{x}$, the borrower seeks to minimize

$$\begin{aligned} J[u,\bar{x}] &\doteq E\left[\int_0^{T_B} e^{-rt}L(u(t))\,dt + B\,e^{-rT_B}\right] \\ &= [\text{cost of servicing the debt}] + [\text{bankruptcy cost}]. \end{aligned}$$

When an investor buy a bond of unit nominal value at time $t = 0$, he receives a stream of payments for all future times. The repayment rate is

$$\psi(t) = (\lambda + r)\cdot e^{-\lambda t}$$

- λ = the rate at which the borrower pays back the principal

- r = the interest payed on bonds.

If bankruptcy does not occur, the total payoff the lender, exponentially discounted in time is

$$\Psi = \int_0^\infty e^{-rt}(r+\lambda)e^{-\lambda t}\,dt = 1.$$

However, if the borrower goes bankrupt at time $T_B < +\infty$, a lender recover

$$\theta(x(T_B)) \in [0,1) = \text{fraction of his outstanding capital}.$$

The payoff to an investor will be

$$\Psi = \int_0^{T_B} (r+\lambda)e^{-(r+\lambda)t}dt + e^{-(r+\lambda)T_B}\theta(x(T_B))).$$

If $\theta < 1$ then $\Psi < 1$. To offset this possible loss, the investor buys a bond at the discounted price $p \in [0,1]$. Assuming that lenders are risk-neutral, we have

$$p = E\Big[\int_0^{T_B} (r+\lambda)e^{-(r+\lambda)t}dt + e^{-(r+\lambda)T_B}\theta(x(T_B)))\Big].$$

Thus, the optimization problem for the borrower is

$$\text{Minimize } J[u,\bar{x}] \doteq \left[\int_0^{T_B} e^{-rt}L(u(t))\,dt + B\,e^{-rT_B}\right]$$

subject to

$$\dot{x}(t) = -\lambda\cdot x(t) + \frac{(\lambda+r)\cdot x - u(t)}{p(t)}$$

where the discounted bond price is

$$p(t) = E\left[\int_t^{T_B} (\lambda+r)e^{-(\lambda+r)(\tau-t)}d\tau + e^{-(r+\lambda)(T_B-t)}\theta(x(T_B)))\right]$$

and T_B is random bankruptcy time.

Model 1: Given $\rho(x)$ the instantaneous bankruptcy risk, T_B is defined as

$$\text{Prob}\,\{T_B \in [t, t+\varepsilon] \mid T_B > t\} = \rho(x(t))\cdot\varepsilon + o(\varepsilon).$$

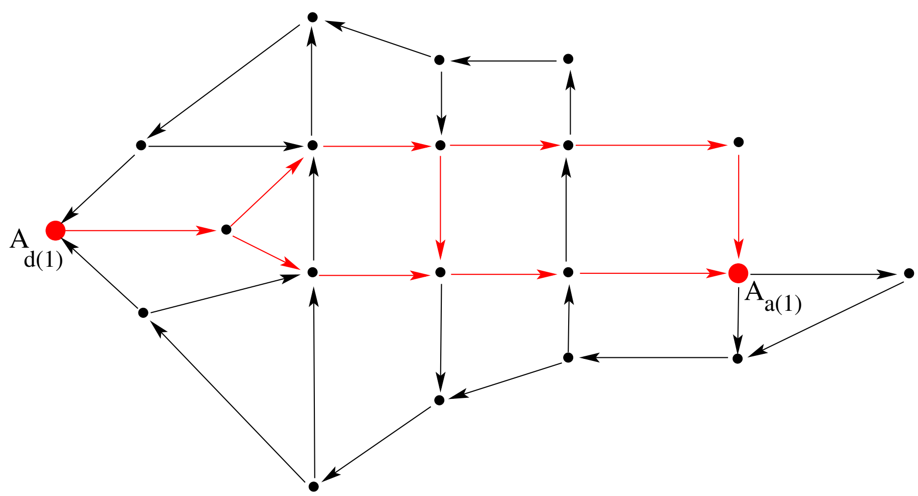
and the yearly income is fixed.

Model 2: The yearly income $Y(t)$ of the borrower is governed by a stochastic evolution equation

$$dY(t) = \mu Y(t)\,dt + \sigma Y(t)\,dW.$$

There is not bankruptcy risk $\rho(x)$ but the borrower can declare bankruptcy at any time he likes.

3. Optimal decision problem on traffic flows. Consider n groups of drivers with different origins and destinations, and different costs. Drivers in the k-th group depart from $A_{d(k)}$ and arrive to $A_{a(k)}$ can use different path $\Gamma_1, \Gamma_2, ...$ to reach to destination. More precisely, let us denote by



- G_k = total number of drivers in the group k for $k = 1, 2, ..., n$;
- Γ_p = viable path to reach destination, for $p = 1, 2, ...N$;
- $t \to u_{k,p}(t)$ = departure rate of k-drivers traveling along the path Γ_p;
- The set of departure rates $\{u_{k,p}\}$ is is **admissible** if

$$u_{k,p}(t) \ \geq \ 0, \qquad \sum_p \int_{-\infty}^{\infty} u_{k,p}(t) \ dt \ = \ G_k \qquad k = 1, 2, ..., n\,.$$

For any $k \in \{1, 2, \ldots, n\}$, a driver β in the k-th group is

$$\varphi_k(\tau^d(\beta)) + \psi(\tau^a(\beta))$$

where $\tau^d(\beta)$ and $\tau^a(\beta)$ are departure and arrival time of driver β, respectively.

Goal: Seeks for a globally optimal admissible family $\{\bar{u}_{k,p}\}$ of departure rates which minimizes the sum of the total costs of all drivers

$$J(\bar{u}) \ = \ \sum_{k,p} \int \left(\varphi_k(t) + \psi_k(\tau_p(t))\right) \ dt.$$

1.3 Existence of optimal open-loop control

This subsection aims to establish an existence result of optimal open-loop control a classical problem in calculus of variation. Given $(T, \bar{x}) \in [0, +\infty[\times\mathbb{R}^n$, we wants to

$$Minimize_{u\in\mathcal{U}_{ad}^T} P[u] \; = \; \int_0^T L(x(t), u(t))\; dt + g(x(T)) \qquad (\mathbf{P})$$

subject to

$$\begin{cases} \dot{x}(t) \; = \; u(t) \quad a.e.\; t \in [0, T] \\ \\ x(0) \; = \; \bar{x}. \end{cases}$$

Here, we shall assume that the following standard assumptions

(L1) For any $R > 0$, there exists $L_R > 0$ such that

$$|L(y, u) - L(x, u)| \leq L_R \cdot |y - x|, \quad \text{for all } u \in \mathbb{R}^n, x, y \in B(0, R).$$

(L2) There exists $l_0 > 0$ and a function $\ell : [0, \infty[\to [0, \infty[$ with $\lim_{r\to\infty} \dfrac{\ell(r)}{r} = +\infty$ and such that

$$L(x, u) \;\geq\; \ell(|u|) - \ell_0, \quad \text{for all } x \in \mathbb{R}^n, u \in \mathbb{R}^n.$$

(L3) For every x, the function $L(x, \cdot)$ is convex.

Theorem 1.1 *Under the standard assumptions* **(L1)**-**(L3)***, assume that g is locally Lipschitz and bounded below the for every $\bar{x} \in \mathbb{R}^n$. Then* **(P)** *admits an bounded optimal control $u^*(\cdot)$, i.e.,*

$$\min_{u\in\mathcal{U}_{ad}} P[u(\cdot)] \; = \; P[u^*(\cdot)].$$

Proof. It is divided into several steps:

1. Let us set

$$-\infty \; < \; \lambda \; := \; \inf_{u\in\in\mathcal{U}_{ad}^T} P[u] \; < \; +\infty$$

and consider a sequence of control functions $u_k(\cdot) \in \mathcal{U}_{ad}^T$ such that $\lim_{k\to\infty} P[u_k] = \lambda$. Thus, there exists $k_0 > 0$ such that

$$P[u_k] \;\leq\; \lambda + 1 \qquad \text{for all } k \geq k_0$$

and (L2) yields

$$\lambda + 1 \;\geq\; \int_0^T \ell(u_k(t))dt - \ell_0 T + \inf_{x\in\mathbb{R}^n} g(x).$$

Since g is bounded below, one has

$$\int_0^T \ell(u_k(t))dt \;\le\; \lambda + 1 + \ell_0 T - \inf_{x\in\mathbb{R}^n} g(x) \;:=\; C_0 \;<\; +\infty.$$

for all $k \ge k_0$. On the other hand recalling that $\lim_{r\to\infty} \dfrac{\ell(r)}{r} = +\infty$, we have

$$\ell(r) \;\ge\; r \qquad \text{for all } r \ge M_0.$$

for some $M_0 > 0$. Thus, it holds

$$\|u_k\|_{\mathbf{L}^1([0,T])} \;\le\; M_0\cdot T + \int_0^T \ell(u_k(t))dt \;\le\; M_0\cdot T + C_0 \;:=\; C_1.$$

for any $k \ge k_0$.

2. Now, let $u \in \mathbf{L}^1([0,T],\mathbb{R}^n)$ be such that $\|u\|_{L^1} \le C_1$. For any $\alpha > 0$, we define

$$u^\alpha(s) \;=\; \begin{cases} u(s) & \text{if } \; |u(s)| \le \alpha, \\ 0 & \text{if } \; |u(s)| > \alpha. \end{cases}$$

Observe that

$$|y^{\bar{x},u^\alpha}(s)|, |y^{\bar{x},u}(s)| \;\le\; R_1 \;:=\; |\bar{x}| + C_1 \qquad \text{for all } s \in [0,T],$$

by setting $I_\alpha = \{s \in [0,T] \mid |u(s)| > \alpha\}$, we estimate

$$\begin{aligned}
&P[u_\alpha(\cdot)] - P[u(\cdot)] \\
&\quad= \int_0^T L(y^{\bar{x},u^\alpha}(s), u^\alpha(s)) - L(y^{\bar{x},u}(s), u(s))ds + g(y^{\bar{x},u^\alpha}(T)) - g(y^{\bar{x},u}(T)) \\
&\quad\le \big(L_{R_1}T + g_{R_1}\big)\cdot \int_0^T |u^\alpha(s) - u(s)|ds + \int_{I_\alpha} L(y^{\bar{x},u}(s),0) - L(y^{\bar{x},u}(s),u(s))ds \\
&\qquad\qquad \le \big(L_{R_1}T + g_{R_1}\big)\cdot \int_{I_\alpha} |u(s)|ds + (K_{R_1} + \ell_0)\cdot |I_\alpha| - \int_{I_\alpha} \ell(|u(s)|)ds
\end{aligned}$$

where g_{R_1} is a Lipschitz constant of g in $B(0,R_1)$ and $K_{R_1} = \sup_{|y|\le R_1} L(y,0)$. If $\alpha > 1$ then

$$P[u_\alpha(\cdot)] - P[u(\cdot)] \;\le\; \Gamma_T \cdot \int_{I_\alpha} |u(s)|ds - \int_{I_\alpha} \ell(|u(s)|)ds$$

with $\Gamma_T := L_{R_1}T + g_{R_1} + K_{R_1} + \ell_0$. Again, from **(L2)**, there exists $\alpha_T > 1$ such that $\ell(r) \ge \Gamma_T \cdot r$ for all $r \ge \alpha_T$. Thus,

$$P[u^{\alpha_T}(\cdot)] - P[u(\cdot)] \;\le\; 0.$$

3. For every $k \geq k_0$, we define

$$v_k(s) \;=\; u_k^{\alpha_T}(s) \;=\; \begin{cases} u_k(s) & \text{if} \quad |u_k(s)| \leq \alpha_T \\ & \\ 0 & \text{if} \quad |u(s)| > \alpha \end{cases} \qquad \text{for all } s \in [0,T].$$

We then have

$$\sup_{k \geq k_0} \|v_k\|_{\mathbf{L}^\infty([0,T])} \;\leq\; \alpha_T \qquad \text{and} \qquad \lim_{k\to\infty} P[v_k] \;=\; \lambda.$$

Since the sequence $\{v_k\}_{k\geq k_0}$ is bounded in $\mathbf{L}^\infty([0,T],\mathbb{R}^n)$, one can construct a subsequence $\{w_k\}_{k\geq 1} \subseteq \{v_k\}_{k\geq k_0}$ such that $\{w_k\}_{k\geq 1}$ converges weakly to $\overline{w}$ in $\mathbf{L}^\infty([0,T],\mathbb{R}^n)$, i.e.,

$$\lim_{k\to\infty} \int_0^T w_k(s)\cdot\varphi(s)ds \;=\; \int_0^T \overline{w}(s)\cdot\varphi(s)ds$$

for all $\varphi \in \mathbf{L}^1([0,T],\mathbb{R}^n)$. In particular, by choosing $\varphi \equiv (1,\dots,1)$, we obtain that $y^{\bar{x},w_k}$ converges uniformly to $y^{\bar{x},\overline{w}}$ and this also implies

$$\lim_{k\to\infty} g(y^{\bar{x},w_k}(T)) \;=\; g(y^{\bar{x},\overline{w}}(T)).$$

Recalling **(L1)**, we have

$$\begin{aligned}
\liminf_{k\to\infty}\left(P[w_k(\cdot)] - P[\overline{w}(\cdot)]\right) \;&=\; \liminf_{k\to\infty} \int_0^T L(y^{\bar{x},w_k}(s), w_k(s)) - L(y^{\bar{x},\overline{w}}(s),\overline{w}(s))ds \\
\geq\; \liminf_{k\to\infty} \int_0^T -L_{R_1}\cdot\left|y^{\bar{x},w_k}(s) - y^{\bar{x},\overline{w}}(s)\right| &+ \left(L(y^{\bar{x},\overline{w}}(s), w_k(s)) - L(y^{\bar{x},\overline{w}}(s),\overline{w}(s))\right)ds \\
&=\; \liminf_{k\to\infty} \int_0^T L(y^{\bar{x},\overline{w}}(s), w_k(s)) - L(y^{\bar{x},\overline{w}}(s),\overline{w}(s))ds.
\end{aligned}$$

Since $L(x,\cdot)$ is convex and $w_k \rightharpoonup \overline{w}$, it holds

$$\liminf_{k\to\infty} \int_0^{\bar{t}} L(y^{\bar{x},\overline{w}}(s), w_k(s)) - L(y^{\bar{x},\overline{w}}(s),\overline{w}(s))ds \;\leq\; 0.$$

and this yields

$$P[\overline{w}(\cdot)] \;\leq\; \liminf_{k\to\infty} P[w_k(\cdot)] \;=\; \lambda.$$

In particular, $u^* = \overline{w}$ is an bounded optimal control of **(P)** and the proof is complete. □

1.4 Pontryagin maximum principle

In this subsection, we wants to derive some necessary conditions for optimality for the following Mayer problem with free terminal point

$$\max_{u\in\mathcal{U}_{ad}} \psi(x(T)) \tag{$\mathbf{P}$}$$

subjects to

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

For the simplicity, we will assume that ψ and f are smooth.

Theorem 1.8 *Let $t\mapsto u^*(t)$ be an optimal control and $x^*(t)$ be the corresponding optimal trajectory for* **(P)***. Denote by $p^*(t)$ the solution of the adjoint equation*

$$\dot{p}^*(t) \ = \ -\,p^*(t)\cdot D_x f(x^*(t),u^*(t)) \qquad \text{with} \qquad p^*(T) \ = \ D\psi(x^*(T)).$$

Then the following holds

$$p^*(t)\cdot f(x^*(t),u^*(t)) \ = \ \max_{w\in U} \ \{p^*(t)\cdot f(x^*(t),w)\} \qquad a.e. \ \ t\in[0,T]$$

Proof. The proof is divided into several steps:

1. (Needle variation) For a fixed $\tau>0$ and $w\in U$, we consider the needle variation $u_\varepsilon\in\mathcal{U}_{ad}$ such that

$$u_\varepsilon(\tau) \ = \ \begin{cases} u^*(\tau) & \text{if} \qquad \tau\notin[\tau-\varepsilon,\tau] \\ \\ w & \text{if} \qquad \tau\in[\tau-\varepsilon,\tau] \end{cases}$$

for ε sufficiently small. The perturbed strategy is denote by

$$x_\varepsilon(t) \ = \ y^{x_0,u_\varepsilon}(t) \qquad \text{for all } t\in[0,T].$$

By the optimality condition, one has that

$$\psi(x^*(T)) \ \geq \ \psi(x_\varepsilon(T)) \qquad \text{for all } \varepsilon>0.$$

Thus, if $\lim_{\varepsilon\to\infty}\frac{x_\varepsilon(T)-x^*(T)}{\varepsilon}=v(T)$ then

$$\lim_{\varepsilon\to\infty}\frac{\psi(x_\varepsilon(T))-\psi(x^*(T))}{\varepsilon} \ = \ D\psi(x^*(T))\cdot v(T) \ \leq \ 0.$$

2. Assume that u^* is continuous at time $t=\tau$. We claim that

$$\lim_{\varepsilon\to0+}\frac{x_\varepsilon(\tau)-x^*(\tau)}{\varepsilon} \ = \ f(x^*(\tau),w))-f(x^*(\tau),u^*(\tau))) \ := \ v(\tau). \tag{1.29}$$

Since $u^*(t) = u_\varepsilon(t)$ for all $t \in [0, \tau - \varepsilon]$, it holds that $x^*(\tau - \varepsilon) = x_\varepsilon(\tau - \varepsilon)$. Using the smoothness of f and continuity of u^* at τ, one can write

$$\begin{aligned} x_\varepsilon(\tau) - x^*(\tau) &= \int_{\tau-\varepsilon}^{\tau} f(x_\varepsilon(s), w) - f(x^*(s), u^*(s))ds \\ &= \int_{\tau-\varepsilon}^{\tau} f(x_\varepsilon(s), w) - f(x^*(s), u^*(\tau))ds + o(\varepsilon) \\ &= \varepsilon \cdot [f(x_\varepsilon(\tau), w) - f(x^*(\tau), u^*(\tau))] + o(\varepsilon) \end{aligned}$$

and this yields (1.29). Since $u_\varepsilon = u^*$ on the remaining interval $[\tau, T]$, the evolution of the tangent vector

$$v(t) := \lim_{\varepsilon \to 0+} \frac{x_\varepsilon(t) - x^*(t)}{\varepsilon} \qquad \text{for all } t \in [\tau, T]$$

governed by the linear equation

$$\dot{v}(t) = D_x f(x^*(t), u^*(t)) \cdot v(t).$$

Recalling that $p^*(t)$ is the solution of the adjoint equation

$$\dot{p}^*(t) = -\, p^*(t) \cdot D_x f(x^*(t), u^*(t)) \qquad \text{with} \qquad p^*(T) = D\psi(x^*(T)),$$

we then have

$$\frac{d}{dt}[p^*(t) \cdot v(t)] = 0 \qquad \text{for all } t \in [\tau, T].$$

In particular,

$$p^*(\tau) \cdot v(\tau) = p^*(T) \cdot v(T) = D\psi(x^*(T)) \cdot v(T) \leq 0$$

and this implies that

$$p^*(\tau) \cdot f(x^*(\tau), u^*(\tau)) = \max_{w \in U} \{p^*(t) \cdot f(x^*(t), w)\}. \tag{1.30}$$

3. One observes that

$$x_\varepsilon(\tau) - x^*(\tau) = \varepsilon \cdot [f(x_\varepsilon(\tau), w) - f(x^*(\tau), u^*(\tau))] + o(\varepsilon)$$

holds if τ is the Lebesgue point of u^*, i.e.,

$$\lim_{\delta \to 0+} \frac{1}{2\delta} \cdot \int_{\tau-\delta}^{\tau+\delta} |u^*(s) - u^*(\tau)|ds = 0.$$

Thus, (1.30) holds for every Lebesgue points of u^*. Thus, the proof is complete by the Lebesgue differentiation theorem. □

Relying on the Maximum Principle, the computation of the optimal control requires two steps:

(i) Given x, p, find $u^*(x,p)$ such that

$$u^*(x,p) \;=\; \operatorname*{argmax}_{w\in U} p\cdot f(x,w);$$

(ii). Solve the two-point boundary value problem

$$\begin{cases} \dot{x} \;=\; f(x, u^*(x,p)) \\ \\ \dot{p} \;=\; -\,p\cdot D_x f(x, u^*(x,p)) \end{cases} \qquad \text{with} \qquad \begin{cases} x(0) \;=\; x_0 \\ \\ p(T) \;=\; \nabla\psi(x(T)). \end{cases}$$

In general, it is not so easy to solve (i)-(ii) since u^* is nonlinear and may be discontinuous of multivalued.

Example 1.(linear pendulum) Consider a linearized pendulum with unit mass. For every $t > 0$, let us denote by

- $q(t) \;=\;$ the position of the pendulum at time t;
- $u(t) \in [-1, 1] =$ an external force at time t.

The equation of motion is

$$\ddot{q}(t) + q(t) \;=\; u(t), \qquad\qquad q(0) \;=\; \dot{q}(0) \;=\; 0. \tag{1.31}$$

Goak: Maximize $q(2)$, the terminal displacement at time $T = 2$.

Let's rewrite (1.31) by setting

$$x(t) \;=\; \begin{pmatrix} q(t) \\ \\ \dot{q}(t) \end{pmatrix}, \qquad\qquad f(x,u) \;=\; \begin{pmatrix} x_2 \\ \\ u - x_1. \end{pmatrix} \tag{1.32}$$

We thus seek for u^* which

$$\max_{u\in\mathcal{U}} \; x_1(2) \qquad \text{subject to} \qquad \dot{x}(t) \;=\; f(x,u).$$

One computes that $D_x f(x,u) \;=\; \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. The linearized equation for a tangent vector $v = (v_1, v_2)^T$ is

$$\dot{v}(s) \;=\; \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} v(s)$$

and the corresponding adjoint vector $p = (p_1, p_2)$ satisfies

$$\dot{p} \;=\; -\,p\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \text{with} \qquad \nabla\psi(x^*(2)) \;=\; (1,0).$$

Solving backward the above ODE, we get

$$p(t) \;=\; (\cos(2-t), \sin(2-t)) \qquad \text{for all } t \in [0,2].$$

For every $t \in [0,T]$, we choose $u^*(t)$ such that

$$u^*(t) \;\in\; \arg\max_{w\in[-1,1]} \left\{\cos(2-t)x_2(t) + \sin(2-t)(-x_1(t)+w)\right\}.$$

It is clear that

$$u^*(t) \;=\; \text{sign}(\sin(T-t)).$$

□

Problem 11(Linear-quadratic optimal control). *Consider the linear control problem*

$$\dot{x} \;=\; Ax + Bu, \qquad\qquad x(0) \;=\; \bar{x}.$$

where $x \in \mathbb{R}^n, u \in \mathbb{R}^m$, $A \in \mathbb{M}^{n\times n}$ and $B \in \mathbb{M}^{n\times m}$. Given two symmetric matrice $Q \in \mathbb{M}^{n\times n}$ and $R \in \mathbb{M}^{m\times m}$, can one write a necessary condition for the optimal control problem

$$\min_{u(\cdot)\in\mathcal{A}_{ad}} \int_0^T \left[x^T Q x + u^T R u\right] dt.$$

The above theorem can be extended to the more optimization problem

$$\max_{u\in\mathcal{U}_{ad}} \left\{ L(t,x(t),u(t))dt + \int_0^T \psi(x(T)) \right\} \qquad (\mathbf{P1})$$

subjects to

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

Theorem 1.9 *Let $t \mapsto u^*(t)$ be an optimal control and $x^*(t)$ be the corresponding optimal trajectory for* **(P1)**. *Denote by $p^*(t)$ the solution of the adjoint equation*

$$\dot{p}^*(t) \;=\; -\,p^*(t)\cdot D_x f(x^*(t),u^*(t)) - DL_x(t,x^*(t),u^*(t)), \quad p^*(T) \;=\; D\psi(x^*(T)).$$

Then the following holds for all most every $t \in [0,T]$

$$p^*(t)\cdot f(x^*(t),u^*(t)) + L(t,x^*(t),u^*(t)) \;=\; \max_{w\in U} \left\{p^*(t)\cdot f(x^*(t),w) + L(t,x^*(t),w)\right\}.$$

Proof. The problem **(P1)** can be written as the folliwng Mayer problem

$$\max_{u\in\mathcal{U}_{ad}} \int_0^T \Psi(x(T)) \qquad \textbf{(P1)}$$

subjects to

$$\dot{y}(t) \;=\; F(t,x(t),u(t)), \qquad y(0) \;=\; (x_0,0).$$

with $y(t)\in\mathbb{R}^{n+1}$ and

$$F(t,x,u) \;=\; (f(t,x,u),L(t,x,u)), \qquad \Psi(y_1,\dots,y_{n+1}) \;=\; \psi(y_1,\dots,y_n)+y_{n+1}.$$

Then one can apply the previous Theorem. □

1.5 Dynamic programming principle

In this section, we will introduce an approach to seek for an optimal feedback control. This will lead to the first order nonlinear partial differential equations of the corresponding value function. The basic tool in this approach is the Dynamic Programming Principle. This principle express the intuitive idea that the minimum cost is achieved if one behaves as follows:

- Let the control system evolve for a small amount of time ε choosing an arbitrary control u and pay the corresponding cost $J[x_0,u]$.

- Denote a new control $u^\sharp$ by

  $$u^\sharp(s) \;=\; u(s) \qquad \text{for all } s\in[0,\varepsilon]\,,$$

  and $u^\sharp(s) = u^*(s)$ for all $s<\varepsilon$ where $u^*(s)$ is the best possible control to minimize the cost function after time $\varepsilon>0$.

- One has that

  $$J[x_0,u] \;\geq\; J(x_0,u^\sharp)\,.$$

Let's recall our control system

$$\begin{cases} \dot{x}(t) \;=\; f(x(t),u(t)), & a.e.\ t\in[0,+\infty[,\\[2ex] x(0) \;=\; x_0\,. \end{cases} \qquad (CS)$$

The basic tool to prove this principle is the following semigroup property for the solutions of system (CS).

Lemma 1.10 *Under standard assumptions* **(F1)-(F2)***, for a given initial data x_0 and control $u\in\mathcal{U}_{ad}$, it holds*

$$y^{x_0,u(\cdot)}(s+t) \;=\; y^{x_s,u(\cdot+s)} \quad \text{with} \quad x_s \;=\; y^{x_0,u}(s)$$

for all $s,t\geq 0$.

Proof. Recalling that f is uniformly Lipschitz, Theorem 1.27 implies that the system (CS) admits a unique solution $y^{x_0,u}(\cdot)$ and has an integral formulation

$$y^{x_0,u}(t) \;=\; x_0 + \int_0^t f(y^{x_0,u}(s), u(s))\; ds \qquad \text{for all } t \geq 0\,.$$

One can write

$$\begin{aligned} y^{x_0,u}(t+s) \;&=\; x_0 + \int_0^{t+s} f(y^{x_0,u}(\tau), u(\tau))\; d\tau \\ &=\; x_0 + \int_0^{s} f(y^{x_0,u}(\tau), u(\tau))\; d\tau + \int_s^{t+s} f(y^{x_0,u}(s+\tau), u(s+\tau))\; d\tau \\ &\qquad\qquad\qquad\qquad\qquad\qquad = \; x_2 + \int_0^t f(y^{x_s,u_s}(\tau), u_s(\tau))\; d\tau \end{aligned}$$

where we set

$$y^{x_s,u_s}(\tau) \;\doteq\; y^{x_0,u}(s+\tau) \qquad \text{and} \qquad u_s(\tau) \;\doteq\; u(s+\tau)$$

and can use this by the uniqueness. □

Remark 1.11 *The following properties of the admissible controls hold*

(i) If $u(\cdot) \in \mathcal{U}_{ad}$ then $u(t+\cdot) \in \mathcal{U}_{ad}$ for all $t \geq 0$;

(ii) For any $u_1 \in \mathcal{U}_{ad}$, $u_2(\cdot) \in \mathcal{U}_{ad}$ and time $t > 0$, the concatenated control

$$u(s) \;\doteq\; \begin{cases} u_1(s) & \text{for } s \in [0,t)\,, \\ u_2(s) & \text{for } s \in [t,+\infty) \end{cases}$$

belongs to $\mathcal{U}_{ad}$.

1. DPP for the minimum time function. Given a closed target set $\mathcal{T} \subset \mathbb{R}^d$ and initial data x_0, the minimum time to reach $\mathcal{T}$ from x_0 is denoted by

$$T(x_0) \;=\; \inf_{u\in\mathcal{U}_{ad}} \{t \geq 0 \mid y^{x_0,u} \in \mathcal{T}\}\,.$$

Proposition 1.11.1 *Under standard assumptions* **(F1)-(F2)***, for a given initial data x_0, the following holds*

$$T(x_0) \;=\; \inf_{u\in\mathcal{U}_{ad}} \;\{s + T(y^{x_0,u}(s))\} \tag{1.33}$$

for all $s \in [0, T_0]$.

Proof. **1.** Given $s > 0$ and control $u \in \mathcal{U}_{ad}$, we first show that

$$T(x_0) \ \leq \ s + T(y^{x_0,u}(s))\,. \tag{1.34}$$

Denote by $x_s \doteq y^{x_0,u}(s)$. One has

$$T(x_s) \ = \ \inf_{v\in\mathcal{U}_{ad}}\{t \geq 0 \mid y^{x_s,v} \in \mathcal{T}\}\,.$$

For any $v \in \mathcal{U}_{ad}$, consider the concatenated control

$$u_v(\tau) \ \doteq \ \begin{cases} u(\tau) & \text{for } \tau \in [0,s)\,, \\ v(s+\tau) & \text{for } \tau \in [s,+\infty)\,. \end{cases}$$

It is clear that $\tilde{u}$ belongs to $\mathcal{U}_{ad}$. From lemma 1.10, it holds

$$y^{x_0,u_v}(s+\tau) \ = \ y^{x_s,v}(\tau) \qquad \text{for all } \tau > 0\,.$$

This implies that

$$s + T(x_s) \ = \ s + \inf_{v\in\mathcal{U}_{ad}}\{t \geq 0 \mid y^{x_s,v} \in \mathcal{T}\} \ = \ \inf_{v\in\mathcal{U}_{ad}}\{t \geq 0 \mid y^{x_0,u_v} \in \mathcal{T}\}$$

and it yields (1.34).

2. To conclude the proof, we show that

$$T(x_0) \ \geq \ s + T(y^{x_0,u}(s))\,. \tag{1.35}$$

Assume that $T(x_0) < +\infty$. By the definition, for any $\varepsilon > 0$, the exists a control u_ε such that the trajectory y^{x,u_ε} reach the target $\mathcal{T}$ before time $T(x_0)+\varepsilon$. This implies that

$$T(x_0)+\varepsilon \ \geq \ s + T(y^{x_0,u_\varepsilon}(s))$$

and it yields

$$T(x_0)+\varepsilon \ \geq \ s + \inf_{u\in\mathcal{U}_{ad}} T(y^{x_0,u_\varepsilon}(s))\,.$$

By letting $\varepsilon \to 0^+$, we obtain (1.35) $\square$

Corollary 1.12 *Assuming that T is a smooth function. We show that T is a solution to the different*

$$H(x,\nabla u(x)) \ = \ 1 \qquad \textit{for all } x \in \mathbb{R}^n\backslash\mathcal{T} \tag{1.36}$$

with

$$H(x,p) \ = \ \sup_{w\in U} \ \langle -p, f(x,w)\rangle.$$

Proof. Indeed, fixed $x_0 \in \mathbb{R}^n \backslash \mathcal{T}$, for any $w \in U$ we consider the constant control

$$u_\varepsilon(s) \;=\; w \qquad \text{for all } s \in [0, \varepsilon[.$$

Let $x_\varepsilon = y^{x_0, u_\varepsilon}(\varepsilon)$, we have

$$x_\varepsilon \;\approx\; x_0 + \varepsilon \cdot f(x_0, w) \qquad \text{and} \qquad T(x_0) \;\leq\; \varepsilon + T(x_\varepsilon).$$

Thus,

$$-1 \;\leq\; \lim_{\varepsilon \to 0+} \frac{T(x_\varepsilon) - T(x)}{\varepsilon} \;=\; \nabla T(x_0) \cdot f(x_0, w)$$

and it yields

$$H(x_0, \nabla T(x_0)) \;\leq\; 1.$$

On the other hand, assume that x^* is an optimal trajectory then

$$T(x_0) \;=\; s + T(x^*(s)) \qquad \text{for all } s > 0$$

and this implies that

$$\nabla T(x_0) \cdot f(x_0, u^*(0)) \;=\; -1.$$

The proof is complete. □

2. DPP for the Bolza problem. Given a running cost $r : U \times \mathbb{R}^n \to \mathbb{R}$ and terminal cost $g : \mathbb{R}^n \to \mathbb{R}$. The value function of the Bolza problem with pay-off P in (1.13) and control system *(CS)* is

$$V(t, x_0) \;=\; \inf_{u \in \mathcal{U}_{ad}} \int_0^t r(u(t), y^{x_0,u}(\tau))\; d\tau + g(y^{x_0,u}(t))$$

for all given initial data $x_0 \in \mathbb{R}^n$.

Proposition 1.12.1 *Under standard assumptions* **(F1)-(F2)**, *for a given initial data $x_0 \in \mathbb{R}^n$ and time $t > 0$, the following holds*

$$V(t, x_0) \;=\; \inf_{u \in \mathcal{U}_{ad}} \left\{ \int_0^s r(u(\tau), y^{x_0,u}(\tau))\; d\tau + V(t - s, y^{x_0,u}(s)) \right\} \tag{1.37}$$

for all $s \in (0, t)$.

Proof. **1.** For any given $u \in \mathcal{U}_{ad}$, we show that

$$V(t, x_0) \;\leq\; \int_0^s r(u(\tau), y^{x_0,u}(\tau))\; d\tau + V(t - s, y^{x_0,u}(s))\,. \tag{1.38}$$

As in the previous proof, denote by $x_s \doteq y^{x_0,u}(s)$. One has

$$V(t - s, x_s) \;=\; \inf_{v \in \mathcal{U}_{ad}} \left\{ \int_0^{t-s} r(v(\tau), y^{x_0,v}(\tau))\; d\tau + g(y^{x_0,v}(t - s)) \right\}.$$

For any $v \in \mathcal{U}_{ad}$, consider the concatenated control

$$u_v(\tau) \doteq \begin{cases} u(\tau) & \text{for } \tau \in [0,s)\,, \\ \\ v(s+\tau) & \text{for } \tau \in [s,+\infty)\,. \end{cases}$$

It is clear that $\tilde{u}$ belongs to $\mathcal{U}_{ad}$. From lemma 1.10, it holds

$$y^{x_0,u_v}(s+\tau) = y^{x_s,v}(\tau) \qquad \text{for all } \tau > 0$$

and

$$y^{x_0,u_v}(\tau) = y^{x_0,u}(\tau) \qquad \text{for all } \tau \in (0,t)\,.$$

Thus,

$$\int_0^t r(u_v(\tau), y^{x_0,u_v}(\tau))\, d\tau + g(y^{x_0,u_v}(t))$$
$$= \int_0^s r(u(\tau), y^{x_0,u}(\tau))\, d\tau + \int_0^{t-s} r(v(\tau), y^{x_0,v}(\tau))\, d\tau + g(y^{x_0,v}(t-s))$$

for all $v \in \mathcal{U}_{ad}$. Therefore,

$$\inf_{v\in\mathcal{U}_{ad}} \int_0^t r(u_v(\tau), y^{x_0,u_v}(\tau))\, d\tau + g(y^{x_0,u_v}(t))$$
$$= \int_0^s r(u(\tau), y^{x_0,u}(\tau))\, d\tau + V(t-s, y^{x_0,u}(s))$$

and it yields (1.38).

2. To complete the proof, we show that

$$V(t,x_0) \geq \int_0^s r(u(\tau), y^{x_0,u}(\tau))\, d\tau + V(t-s, y^{x_0,u}(s))\,. \tag{1.39}$$

For any given $\varepsilon > 0$, there exists a control $u_\varepsilon \in \mathcal{U}_{ad}$ such that

$$V(t,x_0) + \varepsilon \geq \int_0^t r(u_\varepsilon(\tau), y^{x_0,u_\varepsilon}(\tau))\, d\tau + g(y^{x_0,u_\varepsilon}(t))$$
$$\geq \int_0^s r(u_\varepsilon(\tau), y^{x_0,u_\varepsilon}(\tau))\, d\tau + \int_0^{t-s} r(u_\varepsilon(s+\tau), y^{x_s,u_\varepsilon}(s+\tau))\, d\tau + g(y^{x_s,u_\varepsilon}(t-s))$$

where $x_s = y^{x_0,u}(s)$. This implies that

$$V(t,x_0) + \varepsilon \geq \int_0^s r(u_\varepsilon(\tau), y^{x_0,u_\varepsilon}(\tau))\, d\tau + V(t-s, y^{x_0,u_\varepsilon}(s))$$
$$\geq \inf_{v\in\mathcal{U}_{ad}} \left\{ \int_0^{t-s} r(v(\tau), y^{x_0,v}(\tau))\, d\tau + g(y^{x_0,v}(t-s)) \right\}$$

and it yields 1.39. $\square$

Corollary 1.13 *Assume that V is smooth. Then V is the solution to*

$$V_t + H(x, \nabla_x V) \;=\; 0, \qquad V(0, x_0) \;=\; g(x_0)$$

with

$$H(x, p) \;=\; \min_{w \in U} \{r(x, w) + p \cdot f(x, w)\}$$

Proof. I leave it for students. □

3. DPP for the infinite horizon problem. Given a running cost $L : U \times \mathbb{R}^n \to \mathbb{R}$ and a discount rate $\lambda > 0$, the value function of of the infinite time horizon problem with a discount rate is

$$V(x) \;=\; \inf_{u \in \mathcal{U}_{ad}} \left\{ \int_0^{+\infty} e^{-\lambda \cdot t} \cdot L(y^{x_0,u}(t), u(t))\; dt \right\}.$$

The following holds

Proposition 1.13.1 *Under standard assumptions* **(F1)-(F2)***, for a given initial data $x_0 \in \mathbb{R}^n$ and any time $t > 0$. The value function V satisfies*

$$V(x_0) \;=\; \inf_{u \in \mathcal{U}_{ad}} \left\{ \int_0^t e^{-\lambda \cdot s} L(y^{x_0,u}(s), u(s)) ds + e^{-\lambda \cdot t} \cdot V(y^{x_0,u}(t)) \right\}. \tag{1.40}$$

Proof. Fix $x_0 \in \mathbb{R}^n$ and time $t > 0$, for each admissible control $u \in \mathcal{U}_{ad}$, we have

$$\begin{aligned}
P[x_0, u] \;&=\; \int_0^{\infty} e^{-\lambda \cdot t} \cdot L(y^{x_0,u}(s), u(s))\; ds \\
&=\; \int_0^t e^{-\lambda \cdot s} \cdot L(y^{x_0,u}(s), u(s))\; ds + \int_t^{+\infty} e^{-\lambda \cdot s} \cdot L(y^{x_0,u}(s), u(s))\; ds \\
&=\; \int_0^t e^{-\lambda \cdot s} \cdot L(y^{x_0,u}(s), u(s))\; ds + e^{-\lambda \cdot t} \cdot \int_t^{+\infty} e^{-\lambda \cdot (s-t)} \cdot L(y^{x_0,u}(s), u(s))\; ds \\
&=\; \int_0^t e^{-\lambda \cdot s} \cdot L(y^{x_0,u}(s), u(s))\; ds + e^{-\lambda \cdot t} \cdot \int_0^{+\infty} e^{-\lambda \cdot s} \cdot L(y^{x_t,u(t+\cdot)}(s), u(t+s))\; ds
\end{aligned}$$

where $x_t \doteq y^{x_0,u}(t)$. This implies that

$$P[x_0, u] \;\geq\; \int_0^t e^{-\lambda \cdot s} \cdot L(y^{x_0,u}(s), u(s))\; ds + e^{-\lambda \cdot t} \cdot V(y^{x_0,u}(t))$$

and it yields

$$V(x_0) \;\geq\; \inf_{u \in \mathcal{U}_{ad}} \left\{ \int_0^t e^{-\lambda \cdot s} L(y^{x_0,u}(s), u(s)) ds + e^{-\lambda \cdot t} \cdot V(y^{x_0,u}(t)) \right\}.$$

To complete the proof, we need to show that

$$V(x_0) \;\leq\; \inf_{u \in \mathcal{U}_{ad}} \left\{ \int_0^t e^{-\lambda \cdot s} L(y^{x_0,u}(s), u(s)) ds + e^{-\lambda \cdot t} \cdot V(y^{x_0,u}(t)) \right\}. \tag{1.41}$$

For any $u \in \mathcal{U}_{ad}$, we set $x_t = y^{x_0,u}(t)$. For every $\varepsilon > 0$, there exists an control u_ε such that

$$V(x_t) + \varepsilon \ \geq \ \int_0^{+\infty} e^{-\lambda \cdot s} \cdot L(y^{x_t,u_\varepsilon}(s), u_\varepsilon(s))\ ds\,.$$

Denote by

$$\bar{u}_\varepsilon(s) \ \doteq \ \begin{cases} u(s) & \text{for } \tau \in [0,t)\,, \\ \\ u_\varepsilon(s-t) & \text{for } s \in [t,+\infty)\,. \end{cases}$$

By the definition, we have

$$\begin{aligned} V(x_0) \ &\leq \ \int_0^{+\infty} e^{-\lambda \cdot s} \cdot L(y^{x_0,u_\varepsilon}(s), u_\varepsilon(s))ds \\ &= \ \int_0^{t} e^{-\lambda \cdot s} \cdot L(y^{x_0,\bar{u}_\varepsilon}(s), \bar{u}_\varepsilon(s))ds + \int_t^{+\infty} e^{-\lambda \cdot s} \cdot L(y^{x_0,u_\varepsilon}(s), \bar{u}_\varepsilon(s))ds \\ &= \ \int_0^{t} e^{-\lambda \cdot s} \cdot L(y^{x_0,u}(s), u(s))ds + e^{-\lambda \cdot t} \cdot \int_t^{+\infty} e^{-\lambda \cdot (s-t)} L(y^{x_0,\bar{u}_\varepsilon}(s), \bar{u}_\varepsilon(s))ds \\ &= \ \int_0^{t} e^{-\lambda \cdot s} \cdot L(y^{x_0,u}(s), u(s))ds + e^{-\lambda \cdot t} \cdot \int_0^{+\infty} e^{-\lambda \cdot s} L(y^{x_t,u_\varepsilon(s)}, u_\varepsilon(s))\ ds \\ &\leq \ \int_0^{t} e^{-\lambda \cdot s} \cdot L(y^{x_0,u}(s), u(s))ds + e^{-\lambda \cdot t} \cdot [V(x_t) + \varepsilon]\,. \end{aligned}$$

Taking $\varepsilon \to 0$, we obtain that

$$V(x_0) \ \leq \ \int_0^{t} e^{-\lambda \cdot s} \cdot L(y^{x_0,u}(s), u(s))ds + e^{-\lambda \cdot t} \cdot V(y^{x_0,u}(t))$$

for all $u \in \mathcal{U}_{ad}$ and it yields (1.41) $\square$

Corollary 1.14 *Assume that V is smooth. Then V is the solution to*

$$\lambda V \ = \ H(x, \nabla_x V), \qquad V(0,x_0) \ = \ g(x_0)$$

with

$$H(x,p) \ = \ \min_{w \in U} \{r(x,w) + p \cdot f(x,w)\}$$

Proof. For any constant admissible control $\alpha(\cdot) = w \in U$, we have

$$V(x) \ \leq \ \int_0^{t} e^{-\lambda \cdot s} L(y^{x_0,\alpha}(s), w)ds + e^{-\lambda \cdot t} \cdot V(y^{x_0,\alpha}(t)).$$

This implies that

$$-\lambda V(x) + \nabla V(x) \cdot f(x,w) \ = \ \frac{d}{dt}\left[e^{-\lambda \cdot t} \cdot V(y^{x_0,\alpha}(t))\right]_{|t=0} \ \geq \ -\,L(x,w).$$

Thus,

$$\lambda V(x) \ \leq \ \min_{w \in U} \{V(x) \cdot f(x,w) + L(x,w)\} \ = \ H(x, \nabla V(x)).$$

The opposite site is trivial. $\square$

1.6 Recovering the optimal control from the value function

1. Infinite horizon problem. Let V be the value function of the optimization problem

$$V(x) \;=\; \inf_{u\in\mathcal{U}_{ad}} \left\{ \int_0^{+\infty} e^{-\lambda\cdot t}\cdot L(y^{x_0,u}(t),u(t))\;dt \right\}$$

subject to the control system (CS). Assume that V is in C^1, we show how to recover the optimal control.

Given an initial data x_0, the optimal control can be determined as follows:

1. For every control u, we introduce the function

$$\Phi^u(t) \;=\; \int_0^t e^{-\lambda\cdot s}\cdot L(y^{x_0,u}(s),u(s))\;ds + e^{-\lambda t}\cdot V(y^{x_0,u}(t)) \qquad \text{for all } t\geq 0\,. \tag{1.42}$$

It is clear that $\Phi^u(\cdot)$ is a non-decreasing function, i.e.,

$$\Phi^u(t_1)\;\leq\;\Phi^u(t_2) \qquad \text{for all } 0\leq t_1\leq t_2\,.$$

Indeed, we have

$$\begin{aligned} e^{\lambda\cdot t_1}\cdot[\Phi^u(t_1)-\Phi^u(t_2)] \;&=\; V(y^{x_0,u}(t_1)) \\ &- \left(\int_0^{t_2-t_1} e^{-\lambda\cdot s}\cdot L(y^{x_0,u}(t_1+s),u(t_1+s))ds + e^{\lambda(t_2-t_1)V(y^{x_0,u}(t_2))}\right) \;\leq\; 0. \end{aligned}$$

Moreover, u is an optimal control if and only if the function $\Phi^u(\cdot)$ is constant. In this case, we compute

$$\begin{aligned} 0 \;=\; \frac{d}{dt}\,\Phi^u(t) \;&=\; e^{-\lambda\cdot t}\cdot L(y^{x_0,u}(t),u(t)) \\ &\quad -\lambda e^{-\lambda t}\cdot V(y^{x_0,u}(t)) + e^{-\lambda t}\cdot\nabla V(y^{x_0,u}(t))\cdot f(y^{x_0,u}(t)), \end{aligned}$$

and this implies that

$$\lambda\cdot V(y^{x_0,u}(t)) \;=\; L(y^{x_0,u}(t),u(t)) + \nabla V(y^{x_0,u}(t))\cdot f(y^{x_0,u}(t))\,. \tag{1.43}$$

for a.e. $t\geq 0$.

2. Given any $\bar{t}>0$ and control $w\in U$, let us consider

$$v_w(\tau) \;\doteq\; \begin{cases} u(s) & \text{for } s\in[0,\bar{t})\,, \\ \\ w & \text{for } \tau\in[\bar{t},+\infty)\,. \end{cases}$$

Recalling that the function $\Phi^{v_w}(t)$ is monotone non-decreasing. This implies that

$$\begin{aligned} 0 \;\leq\; \frac{d}{dt}\,\Phi^{v_w}(t) \;&=\; e^{-\lambda\cdot t}\cdot\big[\nabla V(y^{x_0,v_w}(t))\cdot f(y^{x_0,v_w}(t),v_w(t)) \\ &\quad + L(y^{x_0,v_w}(t),w(t)) - \lambda\cdot V(y^{x_0,v_w}(t))\big] \qquad \text{for all } t\geq 0. \end{aligned}$$

In particular,

$$\nabla V(y^{x_0,v_w}(\bar{t}))\cdot f(y^{x_0,v_w}(\bar{t})+L(y^{x_0,v_w}(t),w(t))-\lambda\cdot V(y^{x_0,v_w}(\bar{t}))\ \geq\ 0\,.$$

Since

$$y^{x_0,v_w}(\bar{t})\ =\ y^{x_0,u}(\bar{t})\ \doteq\ x_{\bar{t}}\qquad\text{and}\qquad v_w(\bar{t})\ =\ w,$$

we have

$$L(x_{\bar{t}},w)+\nabla V(x_{\bar{t}})\cdot f(x_{\bar{t}},w)-\lambda\cdot V(x_{\bar{t}})\ \geq\ 0\,. \tag{1.44}$$

3. From (1.43) and (1.44), one obtain that

$$L(x_{\bar{t}},w)+\nabla V(x_{\bar{t}})\cdot f(x_{\bar{t}},w)\ \geq\ L(x_{\bar{t}},u(\bar{t}))+\nabla V(x_{\bar{t}})\cdot f(x_{\bar{t}},u(\bar{t}))\,.$$

Therefore, if $(u^*(t),x^*(t))$ be an optimal control and corresponding optimal trajectory pair then the followings hold

$$u^*(t)\ =\ \underset{w\in U}{\operatorname{argmin}}\ \{L(x^*(t),w)+\nabla V(x^*(t),w)\cdot f(x^*(t),w)\}$$

and in particular

$$u^*(0)\ =\ \underset{w\in U}{\operatorname{argmin}}\ \{L(x_0,w)+\nabla V(x_0,w)\cdot f(x_0,w)\}\ .$$

Therefore, if the minimum is attained at a unique point, this uniquely determines the optimal control, in feedback form the function

$$u^*(x)\ =\ \underset{w\in U}{\operatorname{argmin}}\ \{L(x,w)+\nabla V(x,w)\cdot f(x,w)\}$$

4. Let's introduction the Hamilton function

$$H(x,p)\ =\ \min_{w\in U}\ \{L(x,w)+\nabla p\cdot f(x,w)\}\,.$$

If the value function V is differentiable at $\bar{x}$ then V solves the Hamilton-Jacobi equation

$$\lambda\cdot V(x)\ =\ H(x,\nabla V(x))$$

at the point $\bar{x}$. □

2 Viscosity solutions

2.1 The method of characteristics

Given an open set $\Omega\subset\mathbb{R}^n$, consider the first order PDEs

$$\begin{cases} H(x,u,\nabla u)\ =\ 0 & \text{for all } x\in\mathbb{R}^n\\ \\ u(x)\ =\ g(x) & x\in\partial\Omega.\end{cases} \tag{2.1}$$

Assume that H and g is $\mathcal{C}^1$ function, we want to construct a solution in a neighborhood of $\partial\Omega$ by the classical method of characteristics, i.e., determining the value of u along a suitable curve $s \mapsto x(s)$ starting from $\partial\Omega$ by solving a suitable system of ODEs. Let's introduce the variable

$$p(x) \ = \ \nabla u(x) \ = \ (u_{x_1}, u_{x_2}, \ldots, u_{n_n}),$$

we seek a system of ODEs describing how u and p change along $x(\cdot)$. We compute

$$\frac{d}{dt}u(x(t)) \ = \ \sum_{i=1}^{n} u_{x_i}(x(t)) \cdot \dot{x}_i(t)$$

and

$$\frac{d}{dt}p_j(x(t)) \ = \ \sum_{i=1}^{n} p_{x_i x_j}(x(t))\dot{x}_i(t) \ = \ \sum_{i=1}^{n} u_{x_i x_j}(x(t))\dot{x}_i(t).$$

for all $j \in \{1, 2, \ldots, n\}$. On the other hand, differentiating (4.1) w.r.t x_j, one gets

$$\frac{\partial H}{\partial x_j} + \frac{\partial H}{\partial u}u_{x_j} + \sum_{i=1}^{n} \frac{\partial H}{\partial p_i} u_{x_i x_j} \ = \ 0$$

ant this implies that

$$\sum_{i=1}^{n} \frac{\partial H}{\partial p_i} \cdot u_{x_i x_j} \ = \ -\frac{\partial H}{\partial x_j} - \frac{\partial H}{\partial u} \cdot u_{x_j}.$$

The idea's of the method of characteristics is to make the terms involving second derivatives disappear by a good choice of x. In this case, we will choose $\dot{x}_i = \frac{\partial H}{\partial p_i}$ and obtain the following Cauchy problem

$$\begin{cases} \dot{x} &= \dfrac{\partial H}{\partial p} \\ \\ \dot{u} &= p \cdot \dfrac{\partial H}{\partial p}, \\ \\ \dot{p} &= -\dfrac{\partial H}{\partial x} - \dfrac{\partial H}{\partial u} \cdot p \end{cases} \qquad \text{with} \qquad \begin{cases} x(0) &= y \\ u(0) &= u(y) \\ p(0) &= \nabla u(y) \end{cases} \qquad \text{for all } y \in \partial\Omega. \quad (2.2)$$

Here, we wrote

$$u(s) \ = \ u(x(s)) \qquad \text{and} \qquad p(s) \ = \ p(x(s)).$$

Solving the above Cauchy problem for every $y \in \partial\Omega$ could provide a solution to (4.1) in a neighborhood of the boundary of Ω. □

Example 1. Given $\Omega \subset \mathbb{R}^2$ be a bounded open set with smooth boundary, consider a Eikonal equation

$$\begin{cases} |\nabla u|^2 & = & 0 & x \in \Omega \\ u & = & 0 & \text{on } \partial\Omega. \end{cases} \tag{2.3}$$

In this case, the function

$$H(x, u, p) \;=\; p_1^2 + p_2^2 - 1.$$

The associated characteristic system of ODEs is

$$\begin{cases} \dot{x} & = & 2p \\ \dot{u} & = & 2|p|^2 \;=\; 2 \\ \dot{p} & = & 0 \end{cases} \qquad \text{with} \qquad \begin{cases} x(0) & = & y \in \partial\Omega \\ u(0) & = & 0 \\ p(0) & = & \mathbf{n}(y) \end{cases} \tag{2.4}$$

where $\mathbf{n}(y)$ is the internal unit normal to the set Ω at point y. Solving this system of ODEs, we get

$$x(s) \;=\; y + 2\mathbf{n}(y)s \qquad \text{and} \qquad u(x(s)) \;=\; 2s.$$

Assume that Ω is smooth. The solution u is a neighborhood of $\partial\Omega$ is

$$u(x) \;=\; |x - y| \;=\; d_{\partial\Omega}(x).$$

However, if Ω is bounded then there will be a set $\Sigma \subset \Omega$ where for every $\bar{x} \in \Sigma$,

$$d_{\partial\Omega}(\bar{x}) \;=\; |\bar{x} - y_1| \;=\; |\bar{x} - y_2|$$

for some $y_1 \neq y_2 \in \partial\Omega$. This shows that (2.3) does not admits a global C^1 solution in general. One should consider solutions in a generalized sense. By Rademacher's theorem, every Lipschitz real valued function on $\mathbb{R}^n$ is differentiable almost everywhere. This leads to a natural definition

Definition 2.1 *The function $u : \overline{\Omega} \to \mathbb{R}$ is a generalized solution to the Cauchy problem (4.1) if it is Lipschitz, satisfies the boundary conditions, and solves the PDE almost everywhere in Ω.*

This concept is fine for the existence but it does not leads to a useful uniqueness result. Indeed, consider the case of Eikonal equation (2.3) with $\Omega = (-1, 1)$. One can easily check that both $u_1(x) = 1 - |x|$ and

$$u_2(x) \;=\; \frac{1}{2} \cdot \left[\chi_{(-1,0]} \times \left(1 - \left| x + \frac{1}{2} \right|^2 \right) + \chi_{[0,1)} \times \left(1 - \left| x - \frac{1}{2} \right|^2 \right) \right]$$

are generalized solution to (4.1). More general, all piecewise affine functions with slopes in $\{-1, 1\}$ are generalized solutions. Notice that these solutions except for the distance function u_1 has a local minimum in the interior of $]-1, 1[$. Thus, u_1

is the only one that can be obtained as a vanishing viscosity limit. Indeed, assume that there exists a family of C^2 solutions to the viscous equations

$$|u_\varepsilon'|^2 - 1 \;=\; \varepsilon \cdot u_\varepsilon''$$

such that $\lim_{\varepsilon \to 0^+} \|u_\varepsilon - \bar{u}\|_\infty = 0$ for some continuous function $\bar{u}$. Assume that $\bar{u}$ have a local minimum $x_0 \in]-1,1[$. Then

$$\bar{u}(x_0 \pm \delta) \;>\; \bar{u}(x_0)$$

for some $\delta > 0$. Thus, for $\varepsilon > 0$ sufficiently small, the function u_ε has a local minimum $x_\varepsilon \in [x_0 - \delta, x_0 + \delta]$ and

$$u_\varepsilon'(x_\varepsilon) \;=\; 0 \qquad \text{and} \qquad u_\varepsilon''(x_\varepsilon) \;\geq\; 0.$$

This yields a contradiction that

$$-1 \;=\; |u_\varepsilon'(x_\varepsilon)|^2 - 1 \;=\; \varepsilon \cdot u_\varepsilon''(x_\varepsilon) \;\geq\; 0.$$

In general, one could looks for a solution to (4.1) by vanishing viscosity method, i.e., let u_ε be a C^2 solution to the viscous equation

$$H(x, u_\varepsilon, \nabla u_\varepsilon) \;=\; \varepsilon \cdot \Delta u_\varepsilon.$$

Show that

- u_ε is locally bounded in Ω, uniformly w.r.t ε;
- the sequence $\{u_\varepsilon\}_{\varepsilon \geq 0}$ is locally equicontinuous.

The Ascoli's Theorem implies that there exists a subsequence sequence of $\{u_\varepsilon\}$ converges locally uniformly to u which could be the unique solution to (4.1).

2.2 Viscosity solutions via touching functions

Given $\Omega \subseteq \mathbb{R}^n$ open, consider the first order PDE

$$H(x, u, \nabla u) \;=\; 0 \qquad\qquad x \in \Omega. \tag{2.5}$$

Here the function $H : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is a continuous (nonlinear) function.

Definition 2.2 *A function $u \in \mathcal{C}(\Omega)$ is called:*

- *a viscosity subsolution of (2.5) if for every $\varphi \in \mathcal{C}^1(\Omega)$ such that $u - \varphi$ has a local maximum at x_0, it holds*

$$H(x_0, u(x_0), \nabla\varphi(x_0)) \;\leq\; 0;$$

- *a viscosity supersolution of (2.5) if for every $\varphi \in \mathcal{C}^1(\Omega)$ such that $u - \varphi$ has a local minimum at x_0, it holds*

$$H(x_0, u(x_0), \nabla\varphi(x_0)) \;\geq\; 0;$$

- *a viscosity solution of (2.5) if it is both a sub-solution and super-solution.*

The above definition of viscosity solution is naturally motivated by the properties of vanishing viscosity limits.

Theorem 2.3 *Let u_ε be a C^2 solution to the viscous equation*

$$H(x, u_\varepsilon, \nabla u_\varepsilon) \;=\; \varepsilon \cdot \Delta u_\varepsilon. \tag{2.6}$$

Assume that the sequence $\{u_\varepsilon\}_{\varepsilon \geq 0}$ converges uniformly to u on Ω. Then u is a viscosity solution of (2.5).

Proof. We only need to show that u is a sub-viscosity solution of (2.5). The fact that u is a supersolution is proved in an entirely similar way. Given $\varphi \in \mathcal{C}^1(\Omega)$ such that $u - \varphi$ has a local maximum at x_0, we show that

$$H(x_0, u(x_0), \nabla\varphi(x_0)) \;\leq\; 0 \tag{2.7}$$

Consider an alternative function

$$\tilde{\varphi}(y) \;=\; \varphi(y) + |y - x_0|^2.$$

such that $u - \tilde{\varphi}$ has a **strictly** local maximum at x_0, and $\nabla\tilde{\varphi}(x_0) = \nabla\varphi(x_0)$. Thus, since u_ε converges uniformly to u in Ω, one can show that for any given $\delta > 0$, there exists $0 < \rho < \delta$ and a $\mathcal{C}^2$ function ψ such that

(i) For every $y \in B(x_0, \rho_\delta)$, it holds

$$|\nabla\varphi(y) - \nabla\varphi(x_0)| \;\leq\; \delta, \qquad\qquad |\nabla\varphi(y) - \nabla\psi(y)| \;\leq\; \delta;$$

(ii) For every $\varepsilon > 0$ sufficiently small, $u_\varepsilon - \psi$ has a local maximum at a point $x_\varepsilon \in B(x_0, \rho_\delta)$.

By the continuity of u and H, it holds

$$\sup_{y \in B(x_0, \rho_\delta)} |H(y, u(y), \psi(y)) - H(x_0, u(x_0), \nabla u(x_0))| \;=\; O(\delta). \tag{2.8}$$

From (ii), one gets

$$\nabla u_\varepsilon(x_\varepsilon) \;=\; \nabla\psi(x_\varepsilon) \qquad \text{and} \qquad \Delta u_\varepsilon(x_\varepsilon) \;\leq\; \Delta\psi(x_\varepsilon).$$

Thus, (2.6) implies that

$$H(x_\varepsilon, u_\varepsilon(x_\varepsilon), \nabla\psi(x_\varepsilon)) \;=\; H(x_\varepsilon, u_\varepsilon(x_\varepsilon), \nabla u_\varepsilon(x_\varepsilon)) \;=\; \varepsilon\Delta u_\varepsilon(x_\varepsilon) \;\leq\; \varepsilon\Delta\psi_\varepsilon(x_\varepsilon).$$

There exists a sequence $(\varepsilon_m)_{m\geq 1}$ converge to $0+$ such that $\lim_{m\to+\infty} x_{\varepsilon_m} = \bar{x} \in B(x_0,\rho)$, we have

$$H(\bar{x}, u(\bar{x}), \nabla\psi(\bar{x})) \ \leq \ \delta.$$

Thus, (2.8) implies that

$$H(x_0, u(x_0), \nabla u(x_0)) \ \leq \ O(\delta).$$

Taking δ to $0+$, we then obtain (2.9). □

2.3 Generalized differentials

We are now introducing a basic concept of generalized differentials in nonsmooth analysis which can be used to define a viscosity solution and plays important role in regularity theory.

Definition 2.1 *Let f be a real valued function defined on the open set $\Omega \subset \mathbb{R}^n$. For any $x \in \Omega$, the sets*

$$D^- f(x) = \left\{ p \in \mathbb{R}^n \;\middle|\; \liminf_{y\to x} \frac{f(y) - f(x) - \langle p, y-x\rangle}{|y-x|} \geq 0 \right\}$$

$$D^+ f(x) = \left\{ p \in \mathbb{R}^n \;\middle|\; \limsup_{y\to x} \frac{f(y) - f(x) - \langle p, y-x\rangle}{|y-x|} \leq 0 \right\}$$

are called, respectively, the (Fréchet) sub-differential and super-differential of f at x.

In order to get a better felling on the above concepts, Let's denote by

$$\mathrm{Epi}(f) \ = \ \left\{ (y,\beta) \in \mathbb{R}^n \times \mathbb{R} \;\middle|\; \beta \geq f(y) \right\}$$

the epigraph of f, and

$$\mathrm{Hyp}(f) \ = \ \left\{ (x,\beta) \in \mathbb{R}^n \times \mathbb{R} \;\middle|\; \beta \leq f(x) \right\}$$

the hyograph of f. One can show that

- p is a sub-differential of u at x iff the vector $(p,-1)$ is a Fréchet normal vector to $\mathrm{Epi}(f)$ at a point $(x, f(x))$, denote by $(p,-1) \in N^F_{\mathrm{Epi}(f)}(x, f(x))$, i.e.,

$$\limsup_{\mathrm{Epi}(f)\ni(y,\beta)\to(x,f(x))} \left\langle (p,-1), \frac{(y,\beta) - (x,f(x))}{\|(y,\beta)-(x,f(x))\|} \right\rangle \ \leq \ 0$$

- p is a super-differential of u at x the vector $(p,-1)$ is a Fréchet normal vector to $\mathrm{Epi}(f)$ at a point $(x, f(x))$, denote by $(-p,1) \in N^F_{\mathrm{Hyp}(f)}(x, f(x))$, i.e.,

$$\limsup_{\mathrm{Hyp}(f)\ni(y,\beta)\to(x,f(x))} \left\langle (-p,1), \frac{(y,\beta) - (x,f(x))}{\|(y,\beta)-(x,f(x))\|} \right\rangle \ \leq \ 0$$

In other words, one says that

- p is a sub-differential of u at x iff the hyperplane $y \mapsto f(x) + p \cdot (y - x)$ is tangent from below the graph of u at point x.
- p is a super-differential of u at x iff the hyperplane $y \mapsto f(x) + p \cdot (y - x)$ is tangent from above the graph of u at point x.

Example: Consider the distance function to the set $]-\infty, -1] \cup [1, +\infty[$

$$d(x) \ = \ \begin{cases} 0 & \text{if} \quad |x| \geq 1 \\ \\ 1 - |x| & \text{if} \quad |x| \leq 1. \end{cases}$$

In this case, one computes that

$$Df^+(0) \ - \ [-1,1], \quad Df^-(0) \ = \ \emptyset, \quad Df^+(-1) \ = \ \emptyset \quad \text{and} \quad Df^+(-1) \ = \ [0,1].$$

The following characterization of sub- and super-differential is very useful:

Lemma 2.4 *Let u be continuous in Ω. Then*

(i) p is a super-differential of f at x ($p \in D^+f(x)$) iff there exists $\varphi \in \mathcal{C}^1(\Omega)$ such that $f - \varphi$ has a strict local maximum at x and $\nabla\varphi(x) = p$;

(ii) p is a sub-differential of f at x ($p \in D^-f(x)$) iff there exists $\varphi \in \mathcal{C}^1(\Omega)$ such that $f - \varphi$ has a strict local minimum at x and $\nabla\varphi(x) = p$.

Proof. Assume that there exists $\varphi \in \mathcal{C}^1(\Omega)$ such that $f - \varphi$ has a local maximum at x and $\nabla\varphi(x) = p$. In this case, we have

$$f(y) - f(x) \ \leq \ \varphi(y) - \varphi(x) \ = \ p \cdot (y - x) + O(|y - x|)$$

and this implies that

$$\limsup_{y \to x} \frac{f(y) - f(x) - \langle p, y - x \rangle}{|y - x|} \ \leq \ 0.$$

Thus, p is in $Df^+(x)$.

Conversely, assume that $p \in Df^+(x)$. Consider a non-decreasing function $\rho : [0, +\infty[\to [0, +\infty]$ such that

$$\rho(r) \ = \ \max\left\{0, \sup_{0 < |x-y| < r} \left| \frac{f(y) - f(x) - \langle p, y - x \rangle}{|y - x|} \right| \right\}.$$

Since $p \in Df^+(x)$, it holds

$$\rho(0) \;=\; \lim_{r\to 0+} \sup_{0<|x-y|<r} \left| \frac{f(y)-f(x)-\langle p, y-x\rangle}{|y-x|} \right| \;=\; 0.$$

The function $\varphi : \Omega \to \mathbb{R}$ is defined by

$$\varphi(y) \;=\; f(x) + p\cdot(y-x) + \int_0^{2|y-x|} \rho(r)dr + |y-x|^2.$$

Since $\displaystyle\int_0^{2|y-x|} \rho(r)dr = |y-x|\cdot O(|y-x|)$, it holds that $\varphi \in \mathcal{C}^1(\Omega)$ and $\nabla\varphi(x) = p$. On the other hand, we estimate

$$\begin{aligned} f(y)-\varphi(y) \;&=\; [f(y)-f(x)-p\cdot(y-x)] - \int_0^{2|y-x|}\rho(r)dr - |y-x|^2 \\ &\leq\; \rho(|y-x|)\cdot|y-x| - \int_{|y-x|}^{2|y-x|}\rho(r)dr - |y-x|^2 \;\leq\; 0 \end{aligned}$$

and thus $f-\varphi$ has a local strictly maximum at x. □

Corollary 2.5 *A function $u \in \mathcal{C}(\Omega)$ is*

- *a viscosity subsolution of (2.5) if*

$$H(x_0, u(x_0), p) \;\leq\; 0 \qquad \text{for all } x_0 \in \Omega, p \in D^+u(x_0) \tag{2.9}$$

- *a viscosity supersolution of (2.5) if f*

$$H(x_0, u(x_0), p) \;\geq\; 0 \qquad \text{for all } x_0 \in \Omega, p \in D^-u(x_0) \tag{2.10}$$

- *a viscosity solution of (2.5) if (2.9) and (2.10) hold.*

Lemma 2.6 *Let $f \in \mathcal{C}(\Omega)$ and $\varphi \in \mathcal{C}^1(\Omega)$ be such that $f-\varphi$ has a strict local maximum at x. If $(f_n)_{n\geq 1}$ converges to f uniformly, then $f_n-\varphi$ has a local maximum x_n for every $n \geq 1$ such that*

$$\lim_{n\to\infty} x_n \;=\; x \qquad \text{and} \qquad u_m(x_n) \;=\; u(x).$$

Proof. For every $\delta > 0$ sufficiently small, there exists $\varepsilon_\delta > 0$ such that

$$\sup_{|y-x|=\delta} [f(y)-\varphi(y)] \;\leq\; f(x)-\varphi(x)-\varepsilon_\delta.$$

Since $(f_n)_{n\geq 1}$ converges to f uniformly, one has

$$\sup_{|y-x|\leq\delta} |f_m(y)-f(y)| \;\leq\; \frac{\varepsilon_\delta}{2} \qquad \text{for all } n \geq N_\delta$$

for some $N_\delta > 0$. Hence, for all $n \geq N_\delta$, it holds

$$\begin{aligned}\sup_{|y-x|=\delta}[f_m(y)-\varphi(y)] &\leq \sup_{|y-x|=\delta} f(y)-\varphi(y)+\sup_{|y-x|\leq\delta}|f(y)-f_m(y)| \\ &\leq f(x)-\varphi(x)-\frac{\varepsilon_\delta}{2}\end{aligned}$$

and this implies that the function $f_m - \varphi$ has a local maximum y_m in $B(x,\delta)$ for all $m \geq N_\delta$. In particular, for every $\delta = \frac{1}{m}$, let $N_m \in \mathbb{N}$ be a smallest natural number such that $f_k - \varphi$ has a local maximum y_k^m in $B(x, 1/m)$ for every $k \geq N_m$. The sequence $(x_m)_{n\geq 1}$ is constructed by

$$x_k = y_m^k \qquad \text{for all } N_m \leq k < N_{m+1}.$$

□

To complete this subsection, let us recall the basic properties of superdifferential and subdifferential of f.

Proposition 2.1 *Let $f : \Omega \to \mathbb{R}^n$ and $x \in \Omega$. Then, the following properties hold:*

(i) $D^+f(x) = -D^-(-f)(x)$.

(ii) $D^+f(x)$ *and* $D^-f(x)$ *are convex (possibly empty).*

(iii) $D^+f(x)$ *and* $D^-f(x)$ *are both nonempty if and only if* f *is differentiable at* x. *In this case, we have that*

$$D^+f(x) = D^-f(x) = Df(x).$$

(iv) The sets of points where sub-differential or super-differential exists

$$\Omega^+ = \{x \in \Omega \mid D^+f(x) \neq \emptyset\}, \qquad \Omega^+ = \{x \in \Omega \mid D^-f(x) \neq \emptyset\}$$

are dense in Ω.

(v) Both the following sets

$$S^+ = \{x \in \Omega \mid D^+f(x) \text{ has more than two elements}\}$$

and

$$S^- = \{x \in \Omega \mid D^+f(x) \text{ has more than two elements}\}$$

are Hausdorff $(n-1)$*-rectifiable.*

Proof. (i) and (ii) are trivial. Let now us prove (iii).

1. Assume that both f is differentiable at x then it is clear that

$$\nabla f(x) \in D^+f(x) \bigcap D^-f(x).$$

For any $p \in D^+f(x)$, there exists $\varphi \in \mathcal{C}^1(\Omega)$ such that $f - \varphi$ has a local maximum at x and $\nabla\varphi(x) = p$. In particular, one has that $0 = \nabla(f - \varphi)(x)$. Hence,

$$\nabla f(x) \ = \ \nabla\varphi(x) \ = \ p.$$

and this yields $D^+f(x) = \nabla f(x)$. Similarly, one can also have that $D^-f(x) = \nabla f(x)$.

Assume that both $D^+f(x)$ and $D^-f(x)$ are non-empty. For $p^\pm \in D^\pm f(x)$, there exist $\varphi, \psi \in \mathcal{C}^1(\Omega)$ such that

$$\nabla\varphi(x) \ = \ p^+, \qquad \nabla\psi(x) \ = \ p^-, \qquad (f-\varphi)(x) \ = \ (f-\psi)(x) \ = \ 0$$

$f - \varphi$ has a local maximum at x, and $f - \psi$ has a local minimum at x. This implies that

$$\varphi(y) \ \leq \ f(y) \ \leq \ \psi(y) \qquad \text{for all } y \in B(x, \delta)$$

for some $\delta > 0$. In particular,

$$\psi(y) - \varphi(y) \ \geq \ 0 \ = \ \psi(x) - \varphi(x) \qquad \text{for all } y \in B(x, \delta)$$

and it yields $\nabla\psi(x) = \nabla\varphi(x)$. Thus,

$$p^+ \ = \ \nabla\varphi(x) \ = \ \nabla\psi(x) \ = \ p^-$$

and f is differentiable at x and $Df(x) = p^+ = p^-$.

2. Let's show that Ω^+ is dense in Ω. The case Ω^- is entirely similar. For every $\bar{x} \in \Omega$ and $\varepsilon > 0$, we need to show that there exists $y \in B(\bar{x}, \varepsilon)$ such that $D^+f(y)$ is non-empty. Let' introduce a smooth function $\varphi : B(x_0, \varepsilon) \to]0, +\infty[$ defined by

$$\varphi(x) \ = \ \frac{1}{\varepsilon^2 - \|x - x_0\|^2} \qquad \text{for all } x \in B(x_0, \varepsilon).$$

Since

$$\lim_{|x-x_0| \to 0+} \varphi(x) \ = \ +\infty,$$

the function $f - \varphi$ has a local maximum at y in $B(x_0, \varepsilon)$. This implies that

$$\nabla\varphi(y) \ \in \ Df(y)$$

and the proof is complete.

3. We will prove (iii) later. $\square$

2.4 Comparison principle

2.4.1 Static problems

Consider the Hamilton-Jacobi equation in a open bounded domain $\Omega \subset \mathbb{R}^n$

$$u(x) + H(x, \nabla u) \;=\; 0 \qquad x \in \Omega \tag{2.11}$$

where $u : \Omega \to \mathbb{R}$ and the Hamiltonian $H : \Omega \times \mathbb{R}^n \to \mathbb{R}$ is uniformly continuous in x variable and satisfies the equicontinuity assumption

$$|H(x,p) - H(y,p)| \;\leq\; \omega(|x-y|(1+|p|)) \tag{2.12}$$

for some continuous function $\omega : [0, +\infty[\to [0 + \infty[$ with $\omega(0) = 0$.

Theorem 2.7 *Let $\underline{u}, \bar{u} \in \mathcal{C}(\overline{\Omega})$ be viscosity sub- and super solutions of (2.28). If*

$$\underline{u}(x) \;\leq\; \bar{u}(x) \qquad \textit{for all } x \in \partial\Omega, \tag{2.13}$$

then $\underline{u}(x) \leq \bar{u}(x)$ for all $x \in \Omega$.

Proof. 1. We need to show that

$$\max_{x \in \overline{\Omega}} [\underline{u}(x) - \bar{u}(x)] \;\leq\; 0.$$

Assume by a contradiction that there exists $x_0 \in \Omega$ such that

$$\underline{u}(x_0) - \bar{u}(x_0) \;=\; \max_{x \in \overline{\Omega}} [\underline{u}(x) - \bar{u}(x)] \;>\; 0. \tag{2.14}$$

If both $\underline{u}$ and $\bar{u}$ are differentiable at x_0 then it holds

$$\nabla \underline{u}(x_0) \;=\; D^+ \underline{u}(x_0) \;=\; D^- \bar{u}(x_0) \;=\; \nabla \bar{u}(x_0).$$

Thus,

$$\underline{u}(x_0) + H(x_0, \nabla \underline{u}(x_0)) \;\leq\; 0 \;\leq\; \bar{u}(x_0) + H(x_0, \nabla \bar{u}(x_0))$$

and this yields a contradiction.

However, the main difficulty is when both $\underline{u}$ and $\bar{u}$ are not differentiable at x_0. To handle this case, the idea is to seek for nearby points x_ε such that

$$\underline{u}(x_\varepsilon) - \bar{u}(x_\varepsilon) \;>\; 0 \qquad \text{and} \qquad p_\varepsilon \;\in\; D^+ \underline{u}(x_\varepsilon) \bigcap D^- \bar{u}(x_\varepsilon). \tag{2.15}$$

Hence,

$$\underline{u}(x_\varepsilon) + H(x_\varepsilon, p_\varepsilon) \;\leq\; 0 \;\leq\; \bar{u}(x_\varepsilon) + H(x_\varepsilon, p_\varepsilon)$$

and this yields a contradiction again.

<u>Main question:</u> How to find x_ε such that (2.15) holds?

2. To find x_ε, one use a classical technique of doubling of variables. The key idea is to look at the continuous function of two variables $\Phi_\varepsilon : \overline{\Omega}\times\overline{\Omega} \to \mathbb{R}$ defined as

$$\Phi_\varepsilon(x,y) \;=\; \underline{u}(x) - \bar{u}(y) - \frac{1}{2\varepsilon}\cdot|x-y|^2 \qquad \text{for all } (x,y)\in\overline{\Omega}\times\overline{\Omega}.$$

We claim that Φ_ε attains a maximum in $\Omega\times\Omega$ for ε sufficiently small. Indeed, let $(x_\varepsilon, y_\varepsilon)$ be a maximum of Φ_ε, i.e.,

$$\Phi_\varepsilon(x_\varepsilon,y_\varepsilon) \;\geq\; \Phi_\varepsilon(x,y) \qquad \text{for all } (x,y)\in\overline{\Omega}\times\overline{\Omega}.$$

Since $\underline{u}(x_0) - \bar{u}(x_0) > 0$, it holds

$$\Phi_\varepsilon(x_\varepsilon,y_\varepsilon) \;\geq\; \Phi_\varepsilon(x_0,x_0) \;=\; \underline{u}(x_0) - \bar{u}(x_0) \;:=\; \delta \;>\; 0.$$

In particular, set $M := \max\{\|\underline{u}\|_\infty, \|\bar{u}\|_\infty\}$, we have

$$|x_\varepsilon - y_\varepsilon| \;\leq\; 2\sqrt{M\varepsilon}. \tag{2.16}$$

On the other hand, by the uniform continuity of $\bar{u}$ and $\underline{u}$, for $\varepsilon>0$ sufficiently small, it holds

$$\max\{|\underline{u}(x)-\underline{u}(y)|, |\overline{u}(x)-\overline{u}(y)|\} \;\leq\; \frac{\delta}{2}$$

for every $|x-y|\leq\sqrt{2M\varepsilon}$. Thus,

$$\begin{aligned}\delta \;\leq\; \Phi_\varepsilon(x_\varepsilon,y_\varepsilon) \;&\leq\; \underline{u}(x_\varepsilon)-\overline{u}(y_\varepsilon) \;\leq\; \min\begin{cases}\underline{u}(x_\varepsilon)-\bar{u}(x_\varepsilon)+|\bar{u}(x_\varepsilon)-\bar{u}(y_\varepsilon)|\\[2ex] \underline{u}(y_\varepsilon)-\bar{u}(y_\varepsilon)+|\underline{u}(x_\varepsilon)-\underline{u}(y_\varepsilon)|\end{cases}\\ &\leq\; \min\{\underline{u}(x_\varepsilon)-\bar{u}(x_\varepsilon), \underline{u}(y_\varepsilon)-\bar{u}(y_\varepsilon)\}+\frac{\delta}{2}.\end{aligned}$$

and this implies that

$$\min\{\underline{u}(x_\varepsilon)-\bar{u}(x_\varepsilon), \underline{u}(y_\varepsilon)-\bar{u}(y_\varepsilon)\} \;\geq\; \frac{\delta}{2} \;>\; 0.$$

From (2.22), one gets that both x_ε and y_ε are not in the boundary of Ω.

3. Let $\varphi_\varepsilon : \overline{\Omega}\to\mathbb{R}$ and $\psi_\varepsilon : \overline{\Omega}\to\mathbb{R}$ be such that

$$\varphi_\varepsilon(x) \;=\; \underline{u}(x) - \Phi_\varepsilon(x,y_\varepsilon) \;=\; \bar{u}(y_\varepsilon) + \frac{1}{2\varepsilon}\cdot|x-y_\varepsilon|^2$$

and

$$\psi_\varepsilon(y) \;=\; \bar{u}(y) + \Phi(x_\varepsilon,y) \;=\; \underline{u}(x_\varepsilon) - \frac{1}{2\varepsilon}|x-y_\varepsilon|^2$$

It is clear that $\underline{u}-\varphi_\varepsilon = \Phi_\varepsilon(\cdot,y_\varepsilon)$ has a maximum at x_ε and and $\bar{u}-\psi_\varepsilon = -\Phi_\varepsilon(x_\varepsilon,\cdot)$ has a minimum at y_ε. Thus,

$$p_\varepsilon \;:=\; \frac{x_\varepsilon - y_\varepsilon}{\varepsilon} = \nabla\varphi(x_\varepsilon) \;=\; \nabla\psi(y_\varepsilon) \;\in\; D^+\underline{u}(x_\varepsilon)\bigcap D^-\bar{u}(y_\varepsilon).$$

Recalling that $\underline{u}$ and $\bar{u}$ are viscosity sub- and super solutions of (2.28), we have

$$\underline{u}(x_\varepsilon) + H(x_\varepsilon, p_\varepsilon) \;\le\; 0 \;\le\; \bar{u}(y_\varepsilon) + H(y_\varepsilon, p_\varepsilon).$$

In particular, the condition (2.12) yields

$$\begin{aligned} \delta \;&\le\; \Phi_\varepsilon(x_\varepsilon, y_\varepsilon) \;\le\; \underline{u}(x_\varepsilon) - \bar{u}(y_\varepsilon) \;\le\; H(y_\varepsilon, p_\varepsilon) - H(x_\varepsilon, p_\varepsilon) \\ &\le\; \omega\left(|x_\varepsilon - y_\varepsilon| \cdot \left(1 + \frac{|x_\varepsilon - y_\varepsilon|}{\varepsilon}\right)\right). \end{aligned}$$

4. To obtain a contradiction, we will show that

$$\lim_{\varepsilon \to 0+} \frac{|x_\varepsilon - y_\varepsilon|}{\varepsilon} \;=\; 0. \tag{2.17}$$

Indeed, since $\Phi_\varepsilon(x_\varepsilon, x_\varepsilon) \le \Phi_\varepsilon(x_\varepsilon, y_\varepsilon)$, one has

$$\underline{u}(x_\varepsilon) - \bar{u}(x_\varepsilon) \;\le\; \underline{u}(x_\varepsilon) - \bar{u}(y_\varepsilon) - \frac{1}{2\varepsilon} \cdot |x_\varepsilon - y_\varepsilon|^2$$

and this implies that

$$\frac{1}{2\varepsilon} \cdot |x_\varepsilon - y_\varepsilon|^2 \;\le\; |\bar{u}(x_\varepsilon) - \bar{u}(y_\varepsilon)|.$$

Thus, (2.16) and the uniform continuity of $\bar{u}$ yields (2.17). □

As a consequence, one obtains a uniqueness result for the boundary problem

$$\begin{cases} u + H(x, \nabla u) = \; 0 & x \in \Omega \\ \\ \qquad\qquad u \;\equiv\; g & x \in \partial\Omega. \end{cases} \tag{2.18}$$

Corollary 2.8 *Under the same assumptions in Theorem 2.7, the boundary problem (2.18) has at most one viscosity solution.*

2.4.2 Time dependent problems

Consider the Cauchy problem

$$\begin{cases} u_t + H(t, x, \nabla u) = \; 0 & (t, x) \in]0, T[\times \mathbb{R}^n \\ \\ \qquad u(0, x) \;=\; g(x) & x \in \mathbb{R}^n. \end{cases} \tag{2.19}$$

where $u : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ and the Hamiltonian $H : [0, +\infty[\times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfies the Lipschitz continuity assumptions, i.e., there exists a constant $C > 0$ such that

$$|H(t, x, p) - H(s, y, p)| \;\le\; C \cdot (|t - s| + |x - y|) \cdot (1 + |p|) \tag{2.20}$$

and

$$|H(t, x, p) - H(t, x, q)| \;\le\; C \cdot |p - q| \tag{2.21}$$

for all $t \in [0, T]$, $x, y, p, q \in \mathbb{R}^n$.

Theorem 2.9 *Let $\underline{u}, \bar{u} \in \mathcal{C}([0,T] \times \mathbb{R}^n)$ be bounded and uniformly continuous viscosity sub- and super solutions of (2.28). If*

$$\underline{u}(0,x) \ \leq \ \bar{u}(0,x) \qquad \text{for all } x \in \mathbb{R}^n, \tag{2.22}$$

then

$$\underline{u}(t,x) \ \leq \ \bar{u}(t,x) \qquad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^n. \tag{2.23}$$

Proof. We will use the same techniques in the proof of Theorem 2.7. Assume by a contradiction that (2.23) fails, i.e.,

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^n} [\underline{u}(t,x) - \bar{u}(t,x)] \ > \ 0.$$

In particular, there is $\lambda > 0$ such that

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^n} [\underline{u}(t,x) - \bar{u}(t,x) - 2\lambda t] \ := \ \delta > 0 \tag{2.24}$$

Smooth case: If there exists $(t_0, x_0) \in [0,T] \times \mathbb{R}^n$ such that

$$\underline{u}(t_0,x_0) - \bar{u}(t_0,x_0) - 2\lambda t \ = \ \sup_{(t,x)\in[0,T]\times\mathbb{R}^n} [\underline{u}(t,x) - \bar{u}(t,x) - 2\lambda t] \ := \ \delta$$

and both $\underline{u}$ and $\bar{u}$ are differentiable at (t_0, x_0) then

$$\nabla\underline{u}(t_0,x_0) \ = \ \nabla\bar{u}(t_0,x_0), \qquad \underline{u}_t(t_0,x_0) - \bar{u}_t(t_0,x_0) - 2\lambda \ \geq \ 0$$

and

$$\underline{u}_t(t_0,x_0) + H(t_0,x_0,\nabla\underline{u}(t_0,x_0)) \ \leq \ 0 \ \leq \ \bar{u}_t(t_0,x_0) + H(t_0,x_0,\nabla\bar{u}(t_0,x_0)).$$

Thus,

$$2\lambda \ \leq \ \underline{u}_t(t_0,x_0) - \bar{u}_t(t_0,x_0) \ \leq \ 0$$

and this yields a contradiction.

Nonsmooth case: 1. Introduce the function $\Phi_\varepsilon : ([0,T] \times \mathbb{R}^n)^2 \to [0,+\infty[$ such that for all $(t,s,x,y) \in \mathbb{R}^2 \times \mathbb{R}^{2n}$

$$\Phi_\varepsilon(t,x,s,y) \ = \ \underline{u}(t,x) - \bar{u}(s,y) - \lambda(s+t) - \frac{1}{\varepsilon^2} \cdot \left(|t-s|^2 + |x-y|^2\right) - \varepsilon \cdot \left(|x|^2 + |y|^2\right).$$

Since $\bar{u}$ and $\underline{u}$ are bounded, set

$$M \ := \ \max\{\|\bar{u}\|_\infty, \|\underline{u}\|_\infty\},$$

we have

$$\Phi_\varepsilon(t,x,s,y) \ \leq \ 2M - \varepsilon \cdot (|x|^2 + |y|^2) - \frac{1}{\varepsilon^2} \cdot \left(|t-s|^2 + |x-y|^2\right). \tag{2.25}$$

This implies that Φ_ε admits a global maximum at a point $(t_\varepsilon, x_\varepsilon, s_\varepsilon, y_\varepsilon) \in ([0,T]\times\mathbb{R}^n)^2$, and

$$\begin{aligned}
\Phi_\varepsilon(t_\varepsilon, x_\varepsilon, s_\varepsilon, y_\varepsilon) &\geq \sup_{(t,x)\in[0,T]\times\mathbb{R}^n} \Phi_\varepsilon(t,x,t,x) \\
&= \sup_{(t,x)\in[0,T]\times\mathbb{R}^n} \left[\underline{u}(t,x) - \bar{u}(t,x) - 2\lambda t - 2\varepsilon|x|^2\right].
\end{aligned}$$

From (2.24), there exists $(t_1, x_1) \in [0,T]\times\mathbb{R}^n$ such that

$$\underline{u}(t_1,x_1) - \bar{u}(t_1,x_1) - 2\lambda t_1 \ \geq\ \frac{3\delta}{4}.$$

Thus, for every $0 < \varepsilon < \dfrac{\delta}{8|x_1|^2}$, one has

$$\begin{aligned}
&\sup_{(t,x)\in[0,T]\times\mathbb{R}^n} \left[\underline{u}(t,x) - \bar{u}(t,x) - 2\lambda t - 2\varepsilon|x|^2\right] \\
&\qquad\geq\ \underline{u}(t_1,x_1) - \bar{u}(t_1,x_1) - 2\lambda t_1 - 2\cdot\frac{\delta}{8|x_1|^2}|x_1|^2 \ =\ \frac{\delta}{2},
\end{aligned}$$

and this yields

$$\Phi_\varepsilon(t_\varepsilon, x_\varepsilon, s_\varepsilon, y_\varepsilon) \ \geq\ \frac{\delta}{2} \ >\ 0.$$

From (2.25), we get

$$\frac{1}{\varepsilon^2}\cdot\left(|t_\varepsilon - s_\varepsilon|^2 + |x_\varepsilon - y_\varepsilon|^2\right) + \varepsilon\cdot\left(|x_\varepsilon|^2 + |y_\varepsilon|^2\right) \ \leq\ 2M,$$

and

$$\max\{|x_\varepsilon|, |y_\varepsilon|\} \ \leq\ \frac{2M}{\sqrt{\varepsilon}} \qquad \text{and} \qquad \max\{|t_\varepsilon - s_\varepsilon|, |x_\varepsilon - y_\varepsilon|\} \ \leq\ 2M\varepsilon. \tag{2.26}$$

Thus,

$$\begin{aligned}
0 \ \leq\ \Phi_\varepsilon(t_\varepsilon, x_\varepsilon, s_\varepsilon, y_\varepsilon) - \Phi_\varepsilon(t_\varepsilon, x_\varepsilon, t_\varepsilon, x_\varepsilon) \ &=\ \bar{u}(t_\varepsilon, x_\varepsilon) - \bar{u}(s_\varepsilon, y_\varepsilon) \\
&- \lambda\cdot(s_\varepsilon - t_\varepsilon) - \frac{1}{\varepsilon^2}\cdot\left(|t_\varepsilon - s_\varepsilon|^2 + |x_\varepsilon - y_\varepsilon|^2\right) - \varepsilon\cdot\left(|y_\varepsilon|^2 - |x_\varepsilon|^2\right),
\end{aligned}$$

and this implies that

$$\frac{1}{\varepsilon^2}\cdot\left(|t_\varepsilon - s_\varepsilon|^2 + |x_\varepsilon - y_\varepsilon|^2\right) \ \leq\ \bar{u}(t_\varepsilon, x_\varepsilon) - \bar{u}(s_\varepsilon, y_\varepsilon) + \lambda\cdot|t_\varepsilon - s_\varepsilon| + \varepsilon\cdot\left||y_\varepsilon|^2 - |x_\varepsilon|^2\right|.$$

By the uniform continuity of $\bar{u}$ and (2.26), one obtains

$$\lim_{\varepsilon\to 0+} \frac{1}{\varepsilon^2}\cdot\left(|t_\varepsilon - s_\varepsilon|^2 + |x_\varepsilon - y_\varepsilon|^2\right) \ =\ 0. \tag{2.27}$$

On the other hand, from the boundary condition, it holds

$$\frac{\delta}{2} \leq \Phi_\varepsilon(t_\varepsilon, x_\varepsilon, s_\varepsilon, y_\varepsilon) \leq \underline{u}(t_\varepsilon, x_\varepsilon) - \bar{u}(s_\varepsilon, y_\varepsilon) \leq \underline{u}(t_\varepsilon, x_\varepsilon) - \bar{u}(s_\varepsilon, y_\varepsilon) - \underline{u}(0, x_\varepsilon) + \bar{u}(0, x_\varepsilon)$$
$$\leq |\underline{u}(t_\varepsilon, x_\varepsilon) - \underline{u}(0, x_\varepsilon)| + |\bar{u}(s_\varepsilon, y_\varepsilon) - \bar{u}(0, y_\varepsilon)| + |\bar{u}(0, x_\varepsilon) - \bar{u}(0, y_\varepsilon)|.$$

Thus, by the uniform continuity of $\bar{u}, \underline{u}$ and (2.27), we conclude that the maximum of Φ_ε can attain only if both t_ε and s_ε are strictly positive for ε sufficiently small.

2. Let $\varphi_\varepsilon : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ and $\psi_\varepsilon : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ be such that

$$\varphi_\varepsilon(t, x) := \underline{u}(t, x) - \Phi_\varepsilon(t, x, s_\varepsilon, y_\varepsilon) = \bar{u}(s_\varepsilon, y_\varepsilon) + \lambda \cdot (t + s_\varepsilon)$$
$$+ \varepsilon \left(|x|^2 + |y_\varepsilon|^2\right) + \frac{1}{\varepsilon^2} \cdot \left(|t - s_\varepsilon|^2 + |x - y_\varepsilon|^2\right),$$

and

$$\psi_\varepsilon(s, y) = \bar{u}(s, y) + \Phi(t_\varepsilon, x_\varepsilon, s, y) = \underline{u}(t_\varepsilon, x_\varepsilon) - \lambda \cdot (t_\varepsilon + s)$$
$$- \varepsilon \left(|x_\varepsilon|^2 + |y|^2\right) - \frac{1}{\varepsilon^2} \cdot \left(|t - s_\varepsilon|^2 + |x - y_\varepsilon|^2\right).$$

It is clear that $\underline{u} - \varphi_\varepsilon = \Phi_\varepsilon(\cdot, s_\varepsilon, y_\varepsilon)$ has a maximum at $(t_\varepsilon, x_\varepsilon)$ and and $\bar{u} - \psi_\varepsilon = \Phi_\varepsilon(t_\varepsilon, x_\varepsilon, \cdot)$ has a minimum at $(s_\varepsilon, y_\varepsilon)$. Since $\underline{u}$ and $\bar{u}$ are viscosity sub- and super solutions of (2.28)

$$\frac{\partial \varphi_\varepsilon}{\partial t}(t_\varepsilon, x_\varepsilon) + H(t_\varepsilon, x_\varepsilon, \nabla \varphi_\varepsilon(t_\varepsilon, x_\varepsilon)) \leq 0 \leq \frac{\partial \psi_\varepsilon}{\partial t}(s_\varepsilon, y_\varepsilon) + H(s_\varepsilon, y_\varepsilon, \nabla \psi_\varepsilon(s_\varepsilon, y_\varepsilon)).$$

A direct computation yields

$$\lambda + \frac{2(t_\varepsilon - s_\varepsilon)}{\varepsilon^2} + H\left(t_\varepsilon, x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) \leq 0$$
$$\leq -\lambda + \frac{2(t_\varepsilon - s_\varepsilon)}{\varepsilon^2} + H\left(s_\varepsilon, y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right).$$

Using the assumptions (2.20)-(2.21), we estimate

$$2\lambda \leq H\left(s_\varepsilon, y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right) - H\left(t_\varepsilon, x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right)$$
$$\leq C \cdot \left[2\varepsilon |x_\varepsilon - y_\varepsilon| + (|t_\varepsilon - s_\varepsilon| + |y_\varepsilon - x_\varepsilon|) \cdot \left(1 + \frac{2|x_\varepsilon - y_\varepsilon|}{\varepsilon^2} + 2\varepsilon |y_\varepsilon|\right)\right].$$

Taking ε to $0+$, we get that $\lambda \leq 0$ and this implies a contradiction. □

Corollary 2.10 *Under the same assumptions in Theorem 2.9, the Cauchy problem (2.19) has at most one bounded and uniformly continuous viscosity solution*

2.5 Perron's method

2.5.1 Static problems

Hamilton-Jacobi equation in a open bounded domain $\Omega \subset \mathbb{R}^n$

$$u(x) + H(x, \nabla u) = 0 \qquad x \in \Omega \tag{2.28}$$

where $u : \Omega \to \mathbb{R}$ and the Hamiltonian $H : \overline{\Omega} \times \mathbb{R}^n \to \mathbb{R}$ is uniformly bounded and continuous, and satisfied a coercive property

$$\lim_{|p|\to+\infty} \min_{x\in\overline{\Omega}} H(x, p) = +\infty. \tag{2.29}$$

Consider the following constant

$$\gamma_0 := \max_{x\in\overline{\Omega}} |H(x, 0)|.$$

It is clear that $\underline{u}_0 \equiv -\gamma_0$ and $\bar{u}_0 \equiv \gamma_0$ and are viscosity sub- and super-solutions of (2.28). Indeed,

$$\begin{aligned} \underline{u}_0(x) + H(x, \nabla \underline{u}_0(x)) &= -\gamma_0 + H(x, 0) \le 0 \\ &\le \gamma_0 + H(x, 0) = \bar{u}_0(x) + H(x, \nabla \bar{u}_0(x)). \end{aligned}$$

On the other hand, assume that u_1 and u_2 are continuous sub-viscosity solutions of (2.28). Then the function $u = \max\{u_1, u_2\}$ is also continuous sub-viscosity solutions of (2.28). Let's consider the closed set

$$S_0 = \{x \in \Omega \mid u_1(x) = u_2(x)\}.$$

Two case are considered:

- For any given $x \in S_0$, if $p \in D^+u(x)$ then

  $$p \in D^+u_1(x) \bigcap D^+u_2(x).$$

  Since u_1 and u_2 are sub-viscosity solutions of (2.28), it holds

  $$u_i(x) + |H(x, p) \le 0 \qquad i = 1, 2.$$

  Thus,

  $$u(x) + |H(x, p) = \max_{i=1,2} \{u_i(x) + H(x, p)\} \le 0.$$

- For any given $x \in \Omega \backslash S_0$, the continuity of u_1 and u_2 imply that there exists $\delta > 0$ such that

  $$u \equiv u_1 \qquad \text{or} \qquad u \equiv u_1 \quad \text{in } B(x, \delta).$$

  and this implies that

  $$u(x) + |H(x, p)| \le 0.$$

  for all $p \in D^+u(x)$.

The proof is complete. □

With the same argument, the followings holds:

Lemma 2.11 *Let $(u_i)_{i\in I}$ and $(v_i)_{i\in I}$ be families of continuous viscosity sub-solutions and super-solutions of (2.28. Assume that*

$$u \;=\; \sup_{i\in I} u_i \qquad \text{and} \qquad v \;=\; \sup_{i\in I} v_i$$

are continuous. Then, u and v are viscosity sub-solution and super-solution, respectively.

The above observations and the comparison principle lead to a natural question:

Question: Is there one construct a viscosity solution of (2.28) with values in $[-\gamma_0, \gamma_0]$.

Theorem 2.12 *Let $u : \overline{\Omega} \to \mathbb{R}$ be defined by*

$$u(x) \;=\; \sup\left\{\varphi \in \mathcal{C}(\overline{\Omega}, [-\gamma_0, \gamma_0]) \;\middle|\; \varphi \text{ is a subsolution to } (2.28)\right\} \tag{2.30}$$

If u is continuous then u is a viscosity solution to (2.28).

Proof. From Lemma 2.11, we need to show that u is a viscosity super-solution of (2.28). Given any $x_0 \in \Omega$, take $\varphi \in \mathcal{C}^1(\Omega)$ such that $u - \varphi$ has a strict minimum zero at x_0, we need to show that

$$u(x_0) + H(x_0, D\varphi(x_0)) \;\geq\; 0. \tag{2.31}$$

Two cases are considered

- If $\varphi(x_0) = u(x_0) = \gamma_0$ then

  $$\varphi(x) - \varphi(x_0) \;\leq\; u(x) - u(x_0) \;\leq\; 0$$

  for x in a neighborhood of x_0 and this implies (2.31), i.e.,

  $$u(x_0) + H(x_0, D\varphi(x_0)) \;=\; \gamma_0 + H(x_0, 0) \;\geq\; 0.$$

- Otherwise, $u(x_0) < \gamma_0 - \varepsilon$ for some small $\varepsilon_1 > 0$. Assume that (2.31), there exists $0 < \varepsilon < \varepsilon_1$ such that

  $$u(x_0) + H(x_0, D\varphi(x_0)) \;<\; -\varepsilon.$$

  By the continuity of u and φ, there exists $\delta > 0$ such that

  $$u(x) + H(x, D\varphi(x)) \;<\; -\frac{\varepsilon}{2} \qquad \text{and} \qquad u(x) \;<\; \gamma_0 - \frac{\varepsilon}{2} \qquad \text{for all } x \in B(x_0, \delta).$$

  Since $u - \varphi$ has a strict minimum zero at x_0, it holds

  $$\eta \;:=\; \min_{\partial B(x_0,\delta)} [u(x) - \varphi(x)] \;>\; 0.$$

Set $\gamma := \min\{\eta, \varepsilon\}$, we denote by

$$\psi(x) = \begin{cases} u(x) & x \in \Omega \backslash B(x_0, \delta) \\ \\ \max\{u(x), \varphi(x) + \eta/2\} & x \in B(x_0, \delta). \end{cases} \tag{2.32}$$

It is clear that

$$-\gamma_0 \le \psi(x) \le \gamma_0 \qquad \text{for all } x \in \Omega.$$

We now claim that there exists $0 < \delta_1 < \delta$ such that

$$u(x) > \varphi(x) + \frac{\eta}{2} \qquad \text{for all } x \in B(x_0, \delta) \backslash B(x_0, \delta_1).$$

Indeed, for any $x \in \partial B(x_0, \delta)$, there exist $\delta_x > 0$ such that

$$u(x) - \varphi(x) > \frac{\eta_1}{2} \qquad \text{for all } x \in B(x, \delta_x).$$

The compactness property of $\partial B(x_0, \delta)$ implies that it can be covered by a finite number of open ball $B(x, \delta_x/2)$, i.e., there exists $x_1, x_2, \ldots, x_N \in \partial B(x_0, \delta)$ such that

$$\partial B(x_0, \delta) \subset \bigcup_{i=1}^{N} B(x_i, \delta_{x_i}/2).$$

Thus, the constant $\delta_1 > 0$ does exist. In particular, ψ is continuous in Ω and it is easy to see that ψ is also a viscosity sub-solution of (2.28). However, at x_0 we have that

$$u(x_0) \le \varphi(x_0) < \psi(x_0)$$

and this yields a contradiction. $\square$

Corollary 2.13 *The viscosity solution u constructed in 2.36 is Lipschitz.*

Proof. Recalling that u in bounded by γ_0, one has

$$H(x, p) \le \gamma_0 \qquad \text{for all } p \in D^+u(x), x \in \Omega. \tag{2.33}$$

Since H is coercive, it holds that

$$\sup_{x \in \Omega, p \in D^+u(x)} |p| \le C$$

for some constant $C > 0$. We show that

$$u(y) - u(x) \le C \cdot |y - x| \qquad \text{for all } x, y \in \overline{\Omega}$$

and this implies that u is uniformly Lipschitz with a Lipschitz constant C. Given any $\varepsilon > 0$ and $x \in \Omega$, consider the function

$$\varphi(y) = (C + \varepsilon) \cdot |y - x| \qquad \text{for all } y \in \Omega.$$

Assume that $u - \varphi$ attains a max at some x_ε. Two cases are considered:

- If $x_\varepsilon \neq x$ then

$$\nabla\varphi(x_\varepsilon) \;=\; (C+\varepsilon)\cdot\left(\frac{x_\varepsilon - x}{|x_\varepsilon - x|}\right) \;\in\; Du^+(x)$$

  and this yields a contradiction.

- If $x_\varepsilon = x$ then

$$u(y)-\varphi(y) \;\leq\; u(x)-\varphi(x)$$

  and it yields

$$u(y)-u(x) \;\leq\; \varphi(y) \;=\; (C+\varepsilon)\cdot|y-x|$$

The proof is complete. □

2.5.2 Time dependent problems

Consider the Cauchy problem

$$\begin{cases} u_t + H(x,\nabla u) = \; 0 & (t,x)\in]0,\infty[\times\mathbb{R}^n \\ \\ u(0,x) \;=\; g(x) & x\in\mathbb{R}^n. \end{cases} \tag{2.34}$$

where $u : [0,T]\times\mathbb{R}^n \to \mathbb{R}$ and the Hamiltonian H is uniformly bounded and continuous in $\mathbb{R}^n\times B(0,R)$ for any $R>0$ and satisfied a coercive property

$$\lim_{|p|\to+\infty}\inf_{x\in\mathbb{R}^n} H(x,p) \;=\; +\infty. \tag{2.35}$$

Under the above assumption, one has that $\underline{u}_0(t,x) = g(x)-\gamma_0\cdot t$ and $\bar{u}_0(t,x) = g(x)-\gamma_0\cdot t$ and are viscosity sub- and super-solutions of (2.34) where the constant

$$\gamma_0 \;:=\; \max_{x\in\overline{\Omega}}|H(x,0)| \;<\; +\infty$$

Theorem 2.14 *Assume that $g\in C^1(\mathbb{R}^n)$ is uniformly Lipschitz. Let $u : [0,\infty[\times\mathbb{R}^n \to \mathbb{R}$ be defined by*

$$u(x) \;=\; \sup\left\{\varphi\in\mathcal{C}\left([0,\infty]\times\mathbb{R}^n,\mathbb{R}\right)\;\middle|\;\underline{u}_0\leq\varphi\leq\bar{u}_0 \text{ is a subsolution to (2.34)}\right\} \tag{2.36}$$

If u is continuous then u is a viscosity solution to (2.34).

Sketch of proof. We need to show that u is a super viscosity solution of (2.34). Given $\varphi\in\mathcal{C}^1([0,\infty[\times\mathbb{R}^n,\mathbb{R})$ such that $u-\varphi$ has a strict minimum zero at at (t_0,x_0), we show that

$$\varphi_t(t_0,x_0)+H(x_0,\nabla\varphi(t_0,x_0)) \;\geq\; 0. \tag{2.37}$$

Two cases are considered:

- If $\varphi(t_0, x_0) = u(t_0, x_0) = \bar{u}_0(t_0, x_0)$ then (2.34) is trivial.
- Otherwise, assume by a contradiction that $\varphi(t_0, x_0) < \bar{u}_0(t_0, x_0)$ and

$$\varphi_t(t_0, x_0) + H(x_0, \nabla\varphi(t_0, x_0)) \ < \ -2\eta.$$

In this case, there exists $\varepsilon > 0$ and $r > 0$ such that

$$\begin{cases} u(t,x) \ < \ \bar{u}_0(t,x) - \varepsilon & \text{for all } (t,x) \in [t_0 - r, t_0 + r] \times \bar{B}(x_0, r) \\ \\ \varphi(t,x) \ < \ u(t,x) - \varepsilon & \text{for all } (t,x) \in \partial[t_0 - r, t_0 + r] \times \partial\bar{B}(x_0, r) \end{cases}$$

Thus, the function $\psi : [0, +\infty[\times\mathbb{R}^n \to \mathbb{R}$ defined by

$$\psi(x) \ = \ \begin{cases} u(t,x) & (t,x) \in [0, +\infty[\times\mathbb{R}^n \backslash [t_0 - r, t_0 + r] \times \bar{B}(x_0, r) \\ \\ \max\{u, \varphi + \varepsilon/2\} & (t,x) \in [t_0 - r, t_0 + r] \times \bar{B}(x_0, r), \end{cases}$$

is continuous and a subsolution of (2.34) but

$$\psi(t_0, x_0) \ = \ \varphi(t_0, x_0) + \frac{\varepsilon}{2} \ > \ u(t_0, x_0).$$

This implies a contradiction.

$\square$

3 Regularity theory

3.1 Semiconcave functions

In this subsection, we collect some main properties of a semiconcave function with linear modulus.

Definition 3.1 *Let $\Omega \subset \mathbb{R}^n$ be an open set. We say that a function $f : \Omega \to \mathbb{R}$ is semiconcave with linear modulus if there exists $C \geq 0$ such that*

$$\lambda f(x) + (1-\lambda)f(y) - f(\lambda x + (1-\lambda)y) \leq \lambda(1-\lambda) \cdot C \cdot |y - x|^2, \tag{3.1}$$

for all $\lambda \in [0,1]$ such that $[x, y] \subset \Omega$. The constant C is called a semiconcavity constant for f in Ω.

Remark 3.1 *The function $f : \Omega \to \mathbb{R}$ is semiconcave with the semiconcavity constant C in Ω if and only if $f(\cdot) - C \cdot \frac{|\cdot|^2}{2}$ is concave in Ω or $D^2 f \leq C$ in the sense of distributions.*

We now introduce a standard criterion of semiconcave functions.

Proposition 3.1 *Let $f : \Omega \to \mathbb{R}$. Assume that f continuous and*

$$f(x+h) + f(x-h) - 2f(x) \le C \cdot |h|^2 \tag{3.2}$$

for all $[x-h, x+h] \subset \Omega$. Then, f is semiconcave with a semiconcavity constant C in Ω.

Proof. We set

$$g(x) = f(x) - C \cdot |x|^2, \quad \text{for all } x \in \Omega.$$

From (3.8), we have that

$$g(x+h) + g(x-h) - 2g(x) \le 0, \tag{3.3}$$

for all $[x-h, x+h] \subset \Omega$. Moreover, (3.1) follows that

$$\lambda g(y) + (1-\lambda)g(x) - g(\lambda x + (1-\lambda)y) \le 0, \tag{3.4}$$

for all $\lambda \in [0,1]$ such that $[x, y] \subset \Omega$. Now, one can show that (3.3) implies (3.4) for all $\lambda \in \mathbb{Q} \cap [0,1]$. Then, by using the continuity, we obtain (3.4) for all $\lambda \in [0,1]$.
□

Proposition 3.2 *A semiconcave function (with constant C) $f : \Omega \to \mathbb{R}$ is locally Lipschitz continuous in Ω.*

Proof. **1.** As in the previous Proposition, let $g : \Omega \to \mathbb{R}$ be such that

$$g(x) = f(x) - C \cdot |x|^2 \quad \text{for all } x \in \Omega.$$

The function g is concave in Ω, i.e.,

$$\lambda g(y) + (1-\lambda)g(x) \le g(\lambda x + (1-\lambda)y) \qquad \text{for all } \lambda \in [0,1], [x,y] \subset \Omega. \tag{3.5}$$

Given any $x_0 \in \Omega$, we consider a closed cube Q with center x_0 such that $Q \subset \Omega$. Let $x_1, ..., x_{2^n}$ be the vertices of Q and

$$m = \min\{f(x_i) \mid i = 1, ..., 2^n\}.$$

For every $y \in Q$, there exists $0 \le \lambda_1, ..., \lambda_{2^n} \le 1$ such that $\sum_{i=1}^{2^n} \lambda_i = 1$

$$\sum_{i=1}^{2^n} \lambda_i \; = \; 1 \quad \text{and} \quad y \; = \; \sum_{i=1}^{2^n} \lambda_i \cdot x_i.$$

From (3.5), it holds

$$m \;\le\; \sum_{i=1}^{2^n} \lambda_i g(x_i) \;\le\; g(y) \tag{3.6}$$

and this yields

$$f(y) \geq m - C \cdot \|y\|^2 \geq m_0 := m - C \cdot \max_{z \in Q} \|z\|^2 \qquad \text{for all } y \in Q. \tag{3.7}$$

On the other hand, one has

$$g(x) \leq 2g(x_0) - g(2x_0 - x) \leq 2g(x_0) - m \qquad \text{for all } x \in Q$$

and

$$f(x) \ \leq \ 2f(x_0) - C\|x_0\|^2 - m + C \cdot \max_{z \in Q} \|z\|^2.$$

Thus, (3.7) implies that

$$\sup_{x \in Q} |f(y)| \ \leq \ 2f(x_0) + C\|x_0\|^2 + |m| + C \cdot \max_{z \in Q} \|z\|^2$$

2. We claim that f is Lipschitz in $Q_1 = x_0 + \frac{1}{2}(Q - x_0)$. Indeed, given any $x, y \in Q_1$, there exists $x_1 \in \partial Q$ such that $x \in [y, x_1]$ and thus

$$x \ = \ \frac{|y - x|}{|x_1 - y|} \cdot x_1 + \frac{|x - x_1|}{|x_1 - y|} \cdot y.$$

From (3.5), one has

$$\frac{g(x) - g(y)}{|x - y|} \ \leq \ \frac{g(y) - g(x_1)}{|x_1 - y|}.$$

and this implies that

$$\frac{f(x) - f(y)}{|y - x|} \leq \frac{f(y) - f(x_1)}{|x_1 - y|} - C \cdot |x_1 + y| + C \cdot |x + y|$$

Since $f(\cdot)$ is bounded in Q and $|x_1 - y| \leq \frac{\mathrm{diam}(Q)}{4}$, we have

$$\frac{f(x) - f(y)}{|y - x|} \leq L_Q$$

for a suitable constant $L_Q > 0$. Similarly, one gets that

$$\frac{f(y) - f(x)}{|y - x|} \leq L_Q.$$

and this yields

$$|f(y) - f(x)| \leq L_Q \cdot |y - x|, \quad \text{for all } x, y \in Q.$$

The proof is complete. □

Corollary 3.1 *Let $f : \Omega \to \mathbb{R}$. f is semiconcave with a semiconcavity constant C in Ω if and only if f continuous and*

$$f(x+h) + f(x-h) - 2f(x) \leq C \cdot |h|^2 \tag{3.8}$$

for all $[x-h, x+h] \subset \Omega$.

Let us now recall the result of H. Rademacher.

Theorem 3.1 (H. Rademacher) *A locally Lipschitz function $f : \Omega \to \mathbb{R}$ is a.e. differentiable in Ω.*

Hence, we obtains the first result on the differentiability of semiconcave function.

Corollary 3.2 *A semiconcave function $f : \Omega \to \mathbb{R}$ is a.e. differentiable in Ω.*

Moreover, a semiconcave function with linear modulus is a concave function up to a quadratic term. This allows to extend immediately some well-know properties of concave functions.

Theorem 3.2 *Let $f : \Omega \to \mathbb{R}$ be semiconcave. Then the following holds:*

*(i) (**Alexandroff's Theorem**) f is a.e. twice differentiable in Ω, i.e., for a.e. $x \in \Omega$, there exists a vector $p \in \mathbb{R}^n$ and a symmetric matrix B such that*

$$\lim_{y \to x} \frac{f(y) - f(x) - \langle p, y-x \rangle + \langle B(y-x), y-x \rangle}{|y-x|^2} = 0.$$

(ii) The gradient of f, defined almost everywhere in Ω, belongs to the class $BV_{loc}(\Omega, \mathbb{R}^n)$.

Example. (Distance function) Let $S \subset \mathbb{R}^n$ be closed. The distance function from a point to S is defined by

$$d_S(x) = \min_{y \in S} |y - x|, \quad (x \in \mathbb{R}^n)$$

is locally semiconcave in $\mathbb{R}^n \backslash S$.

Exercise. Proving that

(1) $d_S(\cdot)$ is locally semiconcave in $\mathbb{R}^n \backslash S$.

(2) $d_S(\cdot)$ is not locally semiconcave in $\mathbb{R}^n$.

(3) $d_S^2(\cdot)$ is seminconcave with semiconcavity constant 2.

Proposition 3.3 *Let $f : \Omega \to \mathbb{R}$ be semiconcave with semiconcavity constant C. Then, a vector $p \in \mathbb{R}^n$ belongs to $D^+f(x)$ if and only if*

$$f(y) - f(x) - \langle p, y - x \rangle \leq \frac{C}{2} \cdot |y - x|^2$$

for every $y \in \Omega$ such that $[x, y] \subset \Omega$.

Exercise. Proving the above proposition.

Corollary 3.3 *Let $f : \Omega \to \mathbb{R}$ be semiconcave with semiconcavity constant C and let $[x, y] \subset \Omega$. Then, for every $p \in D^+f(x)$, $q \in D^+f(y)$, it holds*

$$\langle q - p, y - x \rangle \leq 2C \cdot |y - x|^2.$$

Before going to give a presentation of superdifferential of a semiconcave function. We introduce the concept of reachable gradient.

Definition 3.2 *Let $f : \Omega \to \mathbb{R}^n$ be locally Lipschitz. For every $x \in \Omega$, we denote by*

$$D^*f(x) = \Big\{ p = \lim_{k\to\infty} Df(x_k) \mid f \text{ is differentiable at } x_k \text{ and } x_k \to x \Big\}.$$

From Rademacher's Thereom, one can see that $D^*f(x)$ is nonempty. In the case of seminconcave function, we also have that

Proposition 3.4 *Let $f : \Omega \to \mathbb{R}$ be semiconcave with semiconcavity constant C and let $x \in \Omega$. Then,*

(i) $D^+f(x) = \mathrm{co}(D^*f(x))$ where $\mathrm{co}(D^*f(x))$ is the convex hull of $D^*f(x)$.

(ii) $D^+f(x)$ is singleton if and only if f is differentiable at x.

(iii) $D^+f(\cdot)$ is upper semicontinuous.

(iv) if $D^+f(y)$ is singleton in the neighborhood $\mathcal{O}_x$ of x then $f(\cdot)$ is C^1 in $\mathcal{O}_x$.

To conclude this subsection, we are now discussing on the singular set of f. We denote by

$$\Sigma_f = \{x \in \Omega \mid f \text{ is not differentiable at } x\}.$$

From proposition 3.4, if f be semiconcave then

$$\Sigma_f = \{x \in \Omega \mid \dim_{\mathcal{H}} D^+f(x) \geq 1\}. \tag{3.9}$$

The followings hold:

Theorem 3.3 *Let $f : \Omega \to \mathbb{R}$ be semiconcave. Then, Σ_f is countable $\mathcal{H}^{(n-1)}$-rectifiable. More generally, if we denote by*

$$\Sigma_f^k = \{x \in \Omega \mid \dim_{\mathcal{H}} D^+f(x) \geq k\}$$

then Σ_f^k is countable $\mathcal{H}^{(n-k)}$-rectifiable.

3.1.1 Semiconcavity and time optimal control

Consider the control systems

$$\begin{cases} \dot{x}(t) \;=\; f(x(t), u(t)), \quad t \in [0, +\infty[\; a.e., \\ \\ x(0) \;=\; x_0, \end{cases} \tag{3.10}$$

where $x_0 \in \mathbb{R}^n$ and

+ $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ is the dynamics of the control system

+ $U \subset \mathbb{R}^m$ is the control set

+ $u : [0, +\infty[\to U$ is a control function.

Standard hypotheses

(H1) $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ is Lipschitz

$$|f(y, u) - f(x, u)| \leq L_1 \cdot |y - x|, \quad \text{for all } x, y \in \mathbb{R}^n, u \in U. \tag{3.11}$$

Moreover, the gradient of f with respect to x exists everywhere and is locally Lipschitz in x, uniformly in u.

(H2) U is compact.

The set of admissible control is

$$\mathcal{U}_{ad} = \big\{u : [0, \infty) \to U \mid u \text{ is measurable}\big\}.$$

For every $u \in \mathcal{U}_{ad}$, we recall that $y^{x_0,u}(\cdot)$ is the trajectory staring from x with control u which is the unique solution of (3.10). The minimum time needed to steer x to the closed target $\mathcal{S}$, regarded as a function of x, is called the minimum time function and is denoted by

$$T_{\mathcal{S}}(x) := \inf \,\{t \geq 0 \mid y^{x,u}(t) \in \mathcal{S}, u \in \mathcal{U}_{ad}\}. \tag{3.12}$$

Now, we define

$$H(x, p) = \sup_{u \in U} \langle p, f(x, u) \rangle. \tag{3.13}$$

By the dynamic programming principle, one can show that $T_S(\cdot)$ is a vicosity solution of *Hamilton-Jacobi-Bellman* equation

$$H(x, \nabla T_S(x)) - 1 = 0, \quad \text{for all } x \in \mathcal{R} \backslash S, \tag{3.14}$$

i.e., for all $x \in \mathcal{R}\backslash$,

$$H(x, p) - 1 \geq 0, \quad \text{for all } p \in D^- T_S(x),$$

$$H(x, p) - 1 \leq 0, \quad \text{for all } p \in D^+ T_S(x),$$

where $\mathcal{R}$ is the reachable set denoted by

$$\mathcal{R} = \{x \in \mathbb{R}^n \mid T_S(x) < \infty\}.$$

In particular, the equation (3.14) hold at all differentiability points of $T_S(x)$. Thus,

$$H(x,p) - 1 = 0, \quad \text{for all } x \in \mathcal{R}\backslash S, p \in D^*T(x).$$

It is well-known that $T_{\mathcal{S}}$ is the unique viscosity solution of (3.14) in $\mathcal{R}\backslash S$ satisfying suitable boundary condition.

We want to study the properties of $T_{\mathcal{S}}$ under the following controllability assumption:

(H3) For very $R > 0$, there exist $\mu_R > 0$ such that for all $x \in (B(0,R) \cap \mathcal{R})\backslash S$, there is $u_x \in U$:

$$f(x,u_x) \cdot \frac{x - \pi_S(x)}{|x - \pi_S(x)|} \le -\mu_R. \tag{3.15}$$

Proposition 3.5 *Assume that system (3.10) satisfies (H1)-(H3). Then, T_S is locally Lipschitz in $\mathbb{R}^d$. Moreover, for every $R > 0$, it holds*

$$T_S(x) \le C_R \cdot d_S(x), \quad \textit{for all } x \in (B(0,R) \cap \mathcal{R})\backslash S$$

for some constant C_R.

Therefore, $T_S(x)$ is differentiable almost everywhere in $\mathcal{R}\backslash S$ and

$$H(x, \nabla T_S(x)) - 1 = 0, \qquad a.e.\ x \in \mathcal{R}\backslash S.$$

We now state the main result of this subsection (see in [**?**]).

Theorem 3.4 *Assume that system (3.10) satisfies (H1)-(H3) and the target S satisfies a ρ_0-internal sphere condition, i.e., for every $x \in \partial S$, there exists x_0 such that $x \in B'(x_0, \rho_0) \subset S$. Then, $T_S(\cdot)$ is locally semiconcave in $\mathcal{R}\backslash S$.*

Sketch of proof. (**The method of middle point**)
Fixing any $x \in \mathcal{R}\backslash S$, let $h \in \mathbb{R}^n$ be such that $[x-h, x+h] \subset \mathcal{R}\backslash S$, one needs to show that

$$T_S(x+h) + T_S(x-h) - 2T_S(x) \le C_x \cdot |h|^2. \tag{3.16}$$

Let $u^*(\cdot)$ be an optimal control steering x to S in time $T_S(x)$. We define

$$y_h^+(t) = y^{x+h,u^*}(t), \quad y(t) = y^{x,u^*}(t) \quad \text{and} \quad y_h^-(t) = y^{x-h,u^*}(t).$$

By the dynamics programming principle, we have that

$$T_S(x+h) + T_S(x-h) - 2T_S(x) \le T(y_h^+(t)) + T(y_h^-(t)) - 2T(y(t)). \tag{3.17}$$

Moreover, observing that

$$|y_h^+(t) + y_h^-(t) - 2y(t)| \leq C \cdot |h|^2. \tag{3.18}$$

From (3.17), (3.18) and the locally Lipschitz continuity of $T_{\mathcal{S}}$, we finally obtain that

$$T_S(x+h)+T_S(x-h)-2T_S(x) \leq T(y_h^+(t))+T(y_h^-(t))-2T\Big(\frac{y_h^+(t)) + T(y_h^-(t)}{2}\Big)+C_1|h|^2.$$

Therefore, we only need to study the semiconcavity property of $T_S(\cdot)$ near to the target S. We leave the rest part for the reader. □

3.2 External sphere condition

3.2.1 Sets with finite perimeter

Let us first recall some basic concepts from geometric measure theory.

Definition 3.3 *Let $A \subseteq \mathbb{R}^d$ and $0 \leq p \leq d$. The p-dimensional Hausdorff measure $\mathcal{H}^p(A)$ is defined by $\mathcal{H}^p(A) = \lim_{\delta \to 0^+} \mathcal{H}^p_\delta(A)$, where*

$$\mathcal{H}^p_\delta(A) = \omega_p \cdot \inf\left\{\sum_{i=1}^{\infty} (\mathrm{diam}(U_i))^p : A \subseteq \bigcup_i U_i,\ \mathrm{diam}(U_i) < \delta\right\},$$

and

$$\omega_p := \frac{2^p\Gamma(\frac{p}{2}+1)}{\pi^{p/2}}, \qquad\qquad \Gamma(p) := \int_0^\infty t^{p-1}\, e^{-t}\, dt.$$

The constant ω_p is chosen so that $\mathcal{H}^p(A)$ equals the Lebesgue measure $\mathcal{L}^p(A)$ if $p \in \mathbb{N}$ and A is a subset of a p-dimensional subspace of $\mathbb{R}^d$.

Moreover,

- The *Hausdorff dimension* $\dim_{\mathcal{H}}(A)$ of A by setting:

  $$\dim_{\mathcal{H}}(A) := \inf\{d \geq 0 :\ \mathcal{H}^d(A) = 0\}.$$

- Let $k \in \mathbb{N}$, we say that $A \subset \mathbb{R}^d$ is *countably k-rectifiable* if

  $$A \subset \mathcal{N} \cup \bigcup_{i=1}^{\infty} S_i$$

  where S_i are suitable Lipschitz k-dimensional surfaces and $\mathcal{N}$ is a $\mathcal{H}^k$-negligible set.

- We say A is k-rectifiable if it is countably k-rectifiable and $\mathcal{H}^k(A) < \infty$, while A is locally k-rectifiable if $A \cap K$ is k-rectifiable for any compact set $K \subset \mathbb{R}^d$.

Lemma 3.1 *Given an open subset Ω of $\mathbb{R}^d$ and a Lipschitz continuous function $f : \Omega \to \mathbb{R}^m$, with Lipschitz rank $L \geq 0$, for every $0 \leq k \leq d$, the estimate $\mathcal{H}^k(f(S)) \leq L^k \mathcal{H}^k(S)$ holds for all $S \subseteq \Omega$.*

The concepts of functions of bounded variation and of sets with finite perimeter will also be used:

Definition 3.4 *Let $\Omega \subset \mathbb{R}^d$ be open, and $u \in L^1(\Omega)$. We say that u is a* function of bounded variation in Ω *(denoted by $u \in BV(\Omega)$) if the distributional derivative of u is representable by a finite Radon measure in Ω, i.e., if*

$$\int_\Omega u \frac{\partial \varphi}{\partial x_i}\, dx = - \int_\Omega \varphi \, dD_i u \text{ for all } \varphi \in C_c^\infty(\Omega),\ i = 1, \ldots, d$$

for some Radon measure $Du = (D_1 u, \ldots, D_d u)$. We denote by $\|Du\|$ the total variation of the vector measure Du, i.e.

$$\|Du\|(\Omega) := \sup \left\{ \int_\Omega u(x) \mathrm{div} \phi(x)\, dx : \ \phi \in C_c^1(\Omega, \mathbb{R}^d),\ \|\phi\|_{L^\infty(\Omega)} \leq 1 \right\}.$$

Accordingly, $u \in L^1_{\mathrm{loc}}(\Omega)$ is a function of locally bounded variation in Ω *(denoted by $u \in BV_{\mathrm{loc}}(\Omega)$) if $u \in BV(U)$ for every $U \subseteq \Omega$ open and bounded with $\overline{U} \subset \Omega$.*

Lemma 3.1 *Let $f \in BV(a,b)$; then there exists a measurable set $I \subseteq (a,b)$ such that $\mathcal{L}^1(I) = b - a$ and*

$$\|Df\|(a,b) \geq |f(t) - f(s)| \qquad \text{for any } t, s \in I\,.$$

Definition 3.5 *Let $E \subset \mathbb{R}^d$ be $\mathcal{L}^d$-measurable, and let $\Omega \subseteq \mathbb{R}^d$ be open. E has* finite perimeter in Ω *if its characteristic function*

$$\chi_E(x) := \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{otherwise}, \end{cases}$$

has bounded variation in Ω, and we say that the perimeter *of E in Ω is $P(E,\Omega) = \|D\chi_E\|(\Omega)$. We say that E has perimeter locally finite in Ω if $P(E,U) < +\infty$ for every open bounded subset U of Ω with $\overline{U} \subset \Omega$.*

Definition 3.6 *Let μ be a Radon measure on $\mathbb{R}^d$, and let M be the union of all open sets $U \subset \mathbb{R}^d$ such that $\mu(U) = 0$; the complement of M is called the support of μ and it is denoted by* $\mathrm{supp}(\mu)$.

The following concept of normal vector was introduced by De Giorgi.

Definition 3.7 *Let Ω be a nonempty open subset of $\mathbb{R}^d$ and $E \subset \mathbb{R}^d$ be a set of finite perimeter in Ω; we call* reduced boundary of E in Ω *the set $\partial^* E$ of all points $x \in \mathrm{supp}(\|D\chi_E\|) \cap \Omega$ such that*

$$\nu_E(x) := \lim_{\rho \to 0^+} \frac{D\chi_E(B(x,\rho))}{\|D\chi_E(B(x,\rho))\|} = \frac{dD\chi_E}{d\|D\chi_E\|}(x)$$

exists in $\mathbb{R}^d$ and satisfies $\|\nu_E(x)\| = 1$. The function $-\nu_E : \partial^* E \to \mathbb{R}^d$ is called the De Giorgi outer normal *to E in x.*

Finally, the following measure-theoretic concepts will be used in our analysis.

Definition 3.8 *Let $E \subset \mathbb{R}^d$ be a Borel set. We set, for $x \in \mathbb{R}^d$ and $0 \le k \le d$,*

$$\delta_E^k(x) = \liminf_{\rho \to 0^+} \frac{\mathcal{H}^k(E \cap B(x,\rho))}{\omega_k \rho^k},$$

where ω_k is the k-dimensional Lebesgue measure of the unit ball in $\mathbb{R}^k$. It is well known that for $k = d$ the limit actually exists and is equal to 1 for $\mathcal{L}^d$-a.e. $x \in E$.

Definition 3.9 *Let $E \subseteq \mathbb{R}^d$ be $\mathcal{L}^d$-measurable. We define:*

$$\begin{aligned} E^0 &:= \{x \in \mathbb{R}^d : \delta_E^d(x) = 0\}, && \textit{the } \text{measure theoretic exterior } \textit{of } E; \\ E^1 &:= \{x \in \mathbb{R}^d : \delta_E^d(x) = 1\}, && \textit{the } \text{measure theoretic interior } \textit{of } E; \\ \partial_M E &:= \mathbb{R}^d \setminus (E^0 \cup E^1), && \textit{the } \text{measure theoretic boundary } \textit{of } E. \end{aligned}$$

Concerning the relations among the above introduced concepts of boundary, we recall the following (see Theorem 3.61, p. 158, in [?]).

Theorem 3.5 (De Giorgi-Federer) *Let Ω be a nonempty open subset of $\mathbb{R}^d$ and $E \subseteq \mathbb{R}^d$ be a set of finite perimeter in Ω. Then*

$$\partial^* E \cap \Omega \subseteq \left\{x \in \mathbb{R}^d : \delta_E^d(x) = 1/2\right\} \subseteq \partial_M E \subseteq \partial E,$$

and

$$\mathcal{H}^{d-1}\Big(\Omega \setminus (E^0 \cup \partial^* E \cup E^1)\Big) = 0.$$

In particular, E has density either 0, or $\frac{1}{2}$, or 1 at $\mathcal{H}^{d-1}-$a.e. $x \in \Omega$, and $\mathcal{H}^{d-1}(\partial_M E \setminus \partial^* E) = 0$.

We conclude this subsection with the following criterion for sets with finite perimeter.

Theorem 3.6 (Federer) *Let Ω be a nonempty open subset of $\mathbb{R}^d$ and $E \subseteq \mathbb{R}^d$ be measurable. If $\mathcal{H}^{d-1}(\partial(\Omega \cap E)) < +\infty$ then $P(E, \Omega) < +\infty$.*

3.2.2 External sphere condition

We now introduce new concepts for sets which is associated with semiconcavity concepts. Basing on these ones, we can extend to study the regularity of a class of continuous functions which is applied to time optimal control.

Definition 3.10 *Let $Q \subset \mathbb{R}^d$ be closed and $v \in \mathbb{R}^d$. We say that v is a proximal normal vector to Q at $x \in \partial Q$, denoted by $v \in N_Q^P(x)$, if there exists a constant $\sigma > 0$ such that*

$$\langle v \ , \ y - x \rangle \leq \sigma \cdot |y - x|^2, \quad \text{for all } y \in Q. \tag{3.19}$$

Equivalently $v \in N_Q^P(x)$ if and only if there exists $\lambda > 0$ such that $\pi_Q(x + \lambda v) = \{x\}$.

Definition 3.11 *Let $Q \subset \mathbb{R}^d$ be closed and $x \in \partial Q$. The vector $v \in N_Q^P(x)$ is realized by a ball of radius ρ if and only if (3.19) satisfies for $\sigma = \frac{|v|}{2\rho}$.*

We are ready to give the main concept for this subsection.

Definition 3.12 *Let $Q \subset \mathbb{R}^d$ be closed and let $\theta(\cdot) : \partial Q \to (0, \infty)$ be continuous. We say that Q satisfies the $\theta(\cdot)$-external sphere condition if and only if for every $x \in \partial Q$, there exists a vector $v_x \neq 0$ such that $v_x \in N_Q^P(x)$ is realized by a ball of radius $\theta(x)$, i.e.,*

$$\left\langle \frac{v_x}{|v_x|}, y - x \right\rangle \leq \frac{1}{2\theta(x)} |y - x|^2.$$

for all $y \in Q$.

We will say that Q satisfies the ρ_0-external sphere condition for a constant $\rho_0 > 0$ if $\rho(\cdot) = \rho_0$. We are now going to study the main properties of sets which satisfies an external sphere condition.

Theorem 3.7 (Locally finite perimeter) *Let $Q \subset \mathbb{R}^d$ be closed. Assuming that Q satisfies the $\theta(\cdot)$-external sphere condition. Then, $\partial Q \cap \overline{\mathcal{O}}$ is finitely $\mathcal{H}^{d-1}$-rectifiable for any bounded, open set $\mathcal{O}$. In particular, Q has locally finite perimeter.*

Proof. Since $\mathcal{O}$ is bounded, we have that $\overline{\mathcal{O}}$ is compact. Therefore, there is a constant $\rho_0 > 0$ such that for every $x \in \partial Q \cap \overline{\mathcal{O}}$, there exists a unit vector $v_x \in N_Q^P(x)$ is realized by a ball of radius ρ_0, i.e.,

$$\langle v_x, y - x \rangle \leq \frac{1}{2\rho_0} |y - x|^2.$$

for all $y \in Q$.

1. By the compactness of $\mathbb{S}^{d-1}$, we can find $M_1 \in \mathbb{N}$ and a finite set $\{v_1, ..., v_{M_1}\} \subset \mathbb{R}^{d-1}$ such that

$$\mathbb{S}^{d-1} \subset \bigcup_{i=1}^{M_1} v_i + \frac{1}{3} B'(0, 1)$$

where $\mathbb{S}^{d-1} = \{v \in \mathbb{R}^d \mid |v| = 1\}$ is the unit sphere with center 0. We partition ∂Q as

$$\partial Q = \bigcup_{i=1}^{M_1} \partial Q_i \tag{3.20}$$

where

$$\partial Q_i \ := \ \Big\{x \in \partial Q \ \Big| \ |v_x - v_i| \leq \frac{1}{3}\Big\}.$$

One has first that

$$\partial Q \cap \overline{\mathcal{O}} = \bigcup_{i=1}^{M} \partial Q_i \cap \overline{\mathcal{O}}. \tag{3.21}$$

Moreover, for every $x \in \partial Q_i \cap \overline{\mathcal{O}}$, it holds

$$\begin{aligned} \langle v_i, y - x\rangle \ &\leq \ \langle v_x, y - x\rangle + |v_i - v_x| \cdot |y - x|, \quad \text{for all } y \in Q \\ &\leq \ \Big(\frac{1}{2\rho_0}|y - x| + \frac{1}{3}\Big) \cdot |y - x| \end{aligned}$$

for all $y \in Q$. Therefore, for every $x, y \in \partial Q_i \cap \overline{\mathcal{O}}$, we have that

$$|\langle v_i, y - x\rangle| \leq \Big(\frac{1}{2\rho_0}|y - x| + \frac{1}{3}\Big) \cdot |y - x|. \tag{3.22}$$

2. We are now going to show that $\partial Q_i \cap \overline{\mathcal{O}}$ is finitely $\mathcal{H}^{d-1}$-rectifiable for all $i \in \{1, .., M\}$. Fixing any $i \in \{1, .., M_1\}$, since $\partial Q_i \cap \overline{\mathcal{O}}$ is compact, there exists $M_2 \in \mathbb{N}$ and $x_1, ..., x_{M_2}$ such that

$$\partial Q_i \cap \overline{\mathcal{O}} \subset \bigcup_{k=1}^{M_2} B'(x_k, \delta),$$

where $\delta = \frac{\rho_0}{6}$. Setting $\partial Q_i^k = \partial Q_i \cap \overline{\mathcal{O}} \cap B'(x_k, \delta)$, we have that

$$\partial Q_i \cap \overline{\mathcal{O}} = \bigcup_{k=1}^{M_2} \partial Q_i^k. \tag{3.23}$$

Moreover, by (3.22) and the choice of δ, we have that for every $x, y \in \partial Q_i^k$, it holds

$$|\langle v_i, y - x\rangle| \leq \frac{1}{2} \cdot |y - x|.$$

Now, let $v_i^\perp$ be the subspace of $\mathbb{R}^d$ which is orthogonal to v_i. Let $\pi_i(\cdot)$ be the projection on $v_i^\perp$. From (3.23), one shows that

$$|\pi_i(y) - \pi_i(x)| \geq \frac{1}{\sqrt{2}} \cdot |y - x|, \quad \text{for all } x, y \in \partial Q_i^k.$$

Thus, $\pi_i : \partial Q_i^k \to v_i^\perp$ is injective. Hence, if we set $A_i^k = \pi_i(\partial Q_i^k) \subset v_i^\perp$, the map $\pi_i^{-1} : A_i^k \to Q_i^k$ is Lipschitz with constant $\sqrt{2}$. Therefore, $\partial Q_i \cap \overline{\mathcal{O}}$ is finitely $\mathcal{H}^{d-1}$-rectifiable. By recalling (3.21), $\partial Q \cap \overline{\mathcal{O}}$ is finitely $\mathcal{H}^{d-1}$-rectifiable.

3. Finally, noting that $\mathcal{H}^{d-1}(A_i^k) < +\infty$, it implies that $\mathcal{H}^{d-1}(\partial Q_i^k) \leq 2^{\frac{d-1}{2}}\mathcal{H}^{d-1}(A_i^k) \leq +\infty$. Recalling (3.21), we obtain that $Q \cap \mathcal{O}$ has finite perimeter. The proof is complete. $\square$

Theorem 3.8 *Let $Q \subset \mathbb{R}^d$ be closed. Assuming that Q satisfies the $\theta(\cdot)$-external sphere condition. For every $k \in \{1, ..., d-1\}$, we denote by*

$$\partial Q^k = \big\{x \in \partial Q \mid \dim_{\mathcal{H}} N_Q^P(x) \geq k\big\}.$$

Then, ∂Q^k is countably $\mathcal{H}^{d-k}$-rectifiable.

Proof. The proof is based of the same technique of the previous theorem. $\square$

Exercise 15. Proving the about theorem for $k = 2$.

We now recall the definition of Fréchet normal vector of a set.

Definition 3.13 *Let $Q \subset \mathbb{R}^d$ be closed and $v \in \mathbb{R}^d$. We say that v is a Fréchet normal vector to Q at x, denoted by $v \in N_Q^F(x)$, if*

$$\limsup_{y \in Q \to x} \left\langle v \ , \ \frac{y-x}{|y-x|} \right\rangle \leq 0. \tag{3.24}$$

Lemma 3.2 *Let $Q \subset \mathbb{R}^d$ be closed. Assuming that Q satisfies a $\theta(\cdot)$-external sphere condition. Then, the map $N_Q^F(\cdot) : \partial Q \rightrightarrows \mathbb{R}^d$ is upper-semicontinuous, i.e.,*

$$\lim_{y \to x} N_Q^F(y) \subseteq N_Q^F(x).$$

Proposition 3.6 *Let $Q \subset \mathbb{R}^d$ be closed. Assuming that Q satisfies a $\theta(\cdot)$-external sphere condition. Then, the set Q is smooth in ∂Q^1, i.e., for every $x \in \partial Q^1$, it holds*

$$\lim_{y \in \partial Q \to x} \left\langle v_x \ , \ \frac{y-x}{|y-x|} \right\rangle = 0,$$

where v_x is the unique unit proximal normal vector to Q at x.

Proof. Assume by a contradiction, there exists a sequence $\{y_n\} \subset \partial Q$ converging to x such that

$$\left\langle -v_x \ , \ \frac{y_n - x}{|y_n - x|} \right\rangle \geq \delta \tag{3.25}$$

for a constant $\delta > 0$ and for all $n \in \mathbb{N}$. Let v_n be the unit proximal normal vector to Q at y_n realized by a ball of radius $\theta(y_n)$. Since y_n converges to x, there exists a constant ρ_0 such that for every n, it holds

$$\langle v_n, z - y_n \rangle \leq \rho_0 \cdot |z - y_n|^2, \quad \text{for all } z \in Q. \tag{3.26}$$

Therefore, v_n must converge to v_x. On the other hand, from the above inequality, we get in particularly that

$$\langle v_n, x - y_n\rangle \leq \rho_0 \cdot |x - y_n|^2.$$

It implies that

$$\left\langle v_n \ , \ \frac{x - y_n}{|x - y_n|}\right\rangle \leq \rho_0 \cdot |x - y_n|.$$

Since $\lim_{n\to\infty} v_n = v_x$ and $\lim_{n\to\infty} y_n = x$, we obtain that

$$\left\langle v_x \ , \ \frac{x - y_n}{|x - y_n|}\right\rangle \leq 0.$$

This contradicts to (3.25). The proof is complete. □

3.2.3 External sphere condition and semiconcavity

In this subsection, we will make a the connection between external sphere condition and semiconcavity. Given any open set Ω in $\mathbb{R}^n$, let $f : \Omega \to \mathbb{R}$ be upper semicontinuous function. Denote by

$$\mathrm{hypo}(f) := \{(x, \beta) \mid x \in \Omega, \beta \leq f(x)\}$$

the hypograph of f. The following holds:

Theorem 3.9 *The function f is locally semiconcave in Ω if and only if f is locally Lipschitz and* hypo(f) *satisfies a $\theta(\cdot)$ external sphere condition.*

Proof. 1. Assume that f is locally semiconcave. From proposition (3.2), we have that f is locally Lipschitz in Ω. We now prove that hypo(f) satisfies a $\theta(\cdot)$ external sphere condition. For every $x \in \Omega$, there exists $v_x \in Df^-(x)$ such that

$$f(y) - f(x) - \langle v_x, y - x\rangle \ \leq \ \frac{C_x}{2} \cdot |y - x|^2, \quad \text{for all } y \in B(x, \delta_x) \tag{3.27}$$

where C_x is a suitable constant and δ_x is a suitable constant such that $B(x, \delta_x) \subset \Omega$. It implies that

$$\big\langle (-v_x, 1), (y - x, f(y) - f(x))\big\rangle \ \leq \ \frac{C_x}{2} \cdot |y - x|^2, \quad \text{for all } y \in B(x, \delta_x).$$

Therefore, there exists $\rho_x > 0$ be such that

$$\left\langle \frac{(-v_x, 1)}{|(-v_x, 1)|}, (y - x, \beta - f(x))\right\rangle \ \leq \ \rho_x \cdot \Big(|y - x|^2 + |\beta - f(x)|^2\Big)$$

Thus, $(-v_x, 1) \in N^P_{\mathrm{hypo}(f)}(x, f(x))$ is realized by a ball of radius $\frac{1}{2\rho_x}$. From here, one can show that hypo(f) satisfies a $\theta(\cdot)$ external sphere condition.

2. For the reversed side, we prefer to leave as an exercise. □

Remark 3.2 *Thereom 3.3 is a particular case of theorem 3.8. Moreover, (ii)-(iv) of proposition 3.4 are consequences of proposition 3.6.*

Let's now study the regularity properties a class of continuous functions whose hypograph satisfies an external sphere condition. From theorem 3.9, such class is a generalization of the class of semiconcave functions and is applied to study the regularity of the minimum time function under a weak controllability condition. We denote by

$$\mathcal{F}_\rho(\Omega, \mathbb{R}) = \{f \in \mathcal{C}(\Omega, \mathbb{R}) \mid \text{hypo}(f) \text{ satifies the } \rho - \text{extermal sphere condition}\}$$

where $\mathcal{C}(\Omega, \mathbb{R})$ is the class of continuous function from Ω to $\mathbb{R}$.

Definition 3.14 *For every $x \in \Omega$, the unit vector $v \in \mathbb{S}^{n-1}$ is a horizontal superdifferential of $f \in \mathcal{F}_\rho(\Omega, \mathbb{R})$ at x, denoted by $v \in \partial^\infty f(x)$, if*

$$(-v, 0) \in N^P_{\text{hypo}(f)}(x, f(x)).$$

Remark 3.2 *Let $f : \Omega \to \mathbb{R}$ be continuous. If f is Lipschitz in a neighborhood of $x \in \Omega$ then the set $\partial^\infty f(x)$ is empty.*

Exercise. *Prove the above Remark.*

From the general function $f \in \mathcal{F}_\rho(\Omega, \mathbb{R})$, one may have that the set $\partial^\infty f(x)$ is non-empty at many points $x \in \Omega$. We set

$$\mathcal{S}_f = \{x \in \Omega \mid \partial^\infty f(x) \neq \varnothing\}.$$

Thanks to the ρ-external sphere condition, the following holds:

Proposition 3.7 *Assuming that $f \in \mathcal{F}_\rho(\Omega, \mathbb{R})$. Then, the set $\mathcal{S}_f$ is closed in Ω.*

Sketch of the proof. 1. For every $x \in \mathcal{S}_f$, there exists $v \in \mathbb{S}^{n-1}$ such that $(-v, 0) \in N^P_{\text{hypo}(f)}(x, f(x))$ is realized by a ball of radius ρ, i.e.,

$$\langle -v, y - x\rangle \leq \rho \cdot (|y - x|^2 + |\beta - f(x)|^2), \quad \text{for all } y \in \Omega, \beta \leq f(y).$$

Indeed, let $(-w, 0) \in N^P_{\text{hypo}(f)}(x, f(x))$, along the ray $x(t) = x - t \cdot w$ $(t > 0)$, by using Clarke's density theorem, one can find a sequence x_n converge to x such that f is differentiable at x_n and $\lim_{n\to\infty} |Df(x_n)| = +\infty$. Moreover, since f is differentiable at x, we have that $(-Df(x_n), 1) \in N^P_{\text{hypo}(f)}(x_n, f(x_n))$ realized by a ball of radius ρ. Therefore, there exists a subsequence $\{x_{n_k}\}$ converge to x such that

$$\lim_{n_k\to\infty} \frac{(-Df(x_{n_k}, 1)}{|(-Df(x_{n_k}, 1)|} = (-v, 0).$$

Thus, $(-v, 0) \in N^P_{\text{hypo}(f)}(x, f(x))$ is realized by a ball of radius ρ.

2. Taking any $x_n \in \mathcal{S}_f$ converging to $x \in \Omega$. From the first step, there exists $v_n \in \mathbb{S}^{n-1}$ such that $(-v_n, 0) \in N^P_{\text{hypo}(f)}(x_n, f(x_n))$ is realized by a ball of radius ρ. Hence, there exists $(-v_x, 0) \in N^P_{\text{hypo}(f)}(x, f(x))$ is realized by a ball of radius ρ. $\square$

In the followings, we would like to estimate the size of $\mathcal{S}_f$ by using the Hausdroff measure.

Lemma 3.3 *For any $x \in \mathcal{S}_f$ such that $N^P_{\text{hypo}(f)}(x, f(x)) = \mathbb{R}^+ \cdot (v, 0)$ for some $v \in \mathbb{S}^{n-1}$, there is $\delta_0 = \delta_0(x) > 0$ such that*

$$\|Df\|_{\text{Sq}(x,\delta)} \geq 2^{n-2} \cdot \delta^{n-\frac{1}{2}}, \quad \textit{for all } 0 < \delta < \delta_0 \tag{3.28}$$

where $\text{Sq}(x, \delta) := \big\{(y_1, \dots, y_n) \in \mathbb{R}^n : \max\limits_{i=1,\dots,n} |y_i - x_i| < \delta\big\}$. *In particular, this would imply that*

$$\|Df\|_{B(x,\delta)} \geq 2^{-\frac{3}{2}} \cdot \delta^{n-\frac{1}{2}}, \quad \textit{for all } 0 < \delta < \delta_0. \tag{3.29}$$

Proof. Without loss of generality we will assume that

$$x = 0 \in \Omega, \quad f(x) = 0 \quad \text{and} \quad N^P_{\text{hypo}(f)}(0, 0) = \mathbb{R}^+(e_1, 0).$$

For any $\delta > 0$, we define

$$\begin{aligned} R_\delta &:= \left\{y = (y_1, \dots, y_n) \in \text{Sq}(0, \delta) : \frac{3}{4}\delta < y_1 < \delta\right\}, \\ S_\delta &:= \{y = (y_1, \dots, y_d) \in \text{Sq}(0, \delta) : -\delta < y_1 < -\delta/2\}. \end{aligned}$$

1. We first claim that there exist $\delta_1 > 0$ such that for every $\delta \in (0, \delta_1)$ it holds

$$\begin{cases} f(y) \ \leq \ -\dfrac{1}{2} \cdot \delta & \text{for all } y \in R_\delta, \\ \\ f(y) \ > \ 0 & \text{for all } y \in S_\delta. \end{cases} \tag{3.30}$$

Indeed, for any $y \in R_\delta$, we have

$$\frac{3}{4} \cdot \delta \ < \ \langle (e_1, 0), (y, \beta) \rangle \ \leq \ \rho \cdot \left(\|y\|^2 + |\beta|^2\right)$$

whenever

$$\frac{3}{4} \cdot \delta \ \leq \ \rho \cdot \left(n\delta^2 + |\beta|^2\right) \quad \text{for all } \beta \leq f(y). \tag{3.31}$$

In particular, this implies that

$$f(y) \ < \ 0 \qquad \text{for all } y \in R_\delta$$

and thus the first inequality of (3.30) holds for $\delta > 0$ sufficiently small.

Let us now prove the second inequality in (3.30). Assume by contradiction that there exist sequences $\{\delta_k\}_{k\geq 1}$ and $\{y_k\}_{k\geq 1}$ such that

$$\delta_k \longrightarrow 0^+, \qquad y_k \in S_{\delta_k} \qquad \text{and} \qquad f(y_k) \leq 0.$$

By the continuity of f, we have that

$$\lim_{k\to\infty} f(y_k) = f(0) = 0.$$

On the other hand, let $(v_k, \alpha_k) \in N^P_{\text{hypo}(f)}(y_k, f(y_k))$ be a normal vector realized by a ball of radius ρ and (v_k, α_k) converges to $(e_1, 0)$. For every $\beta \leq 0 = f(0)$, it holds

$$\left\langle \frac{(v_k, \alpha_k)}{|(v_k, \alpha_k)|}, (0, \beta) - (y_k, f(y_k)) \right\rangle \leq \rho \cdot \left(\|y_k\|^2 + |\beta - f(y_k)|^2 \right).$$

By choosing $\beta = f(y_k)$, we obtain that

$$\langle v_k, -y_k \rangle \leq C \cdot \|y_k\|^2.$$

and a direct computation yields

$$\frac{\delta_k}{2} - \|v_k - e_1\| \sqrt{n}\, \delta_k \leq \langle e_1, -y_k \rangle + \langle v_k - e_1, -y_k \rangle = \langle v_k, -y_k \rangle \leq Cn\, \delta_k^2.$$

Dividing both sides by δ_k and passing to the limit as $k \to \infty$ we obtain a contradiction.

2. The Claim allows us to conclude: indeed, for any $\delta < \delta_0 := \min\{\delta_1, \delta_2\}$ and any $z \in (-\delta, \delta)^{n-1}$ we get

$$|f(y_a, z) - f(y_b, z)| \geq \frac{1}{2}\delta^{\frac{1}{2}} \quad \text{for all } y_a \in]\tfrac{3}{4}\delta, \delta[, y_b \in]-\delta, -\delta/2[.$$

By virtue of Lemma 3.1, for any $z \in (-\delta, \delta)^{d-1}$ there exist $y_a(z) \in]\frac{3}{4}\delta, \delta[$ and $y_b(z) \in]-\delta, -\delta/2[$ such that

$$\|Df_z\|(-\delta, \delta) \geq |f(y_a(z), z) - f(y_b(z), z)| \geq \frac{1}{2}\delta^{\frac{1}{2}}$$

where $f_z := f(\cdot, z)$. We obtain

$$\begin{aligned} \|Df\|(\text{Sq}(0, \delta)) &\geq \int_{]-\delta,\delta[^{n-1}} \|D_{e_1} f\|(z +]-\delta, \delta[e_1) dz \\ &= \int_{]-\delta,\delta[^{n-1}} \|Df_z\|(-\delta, \delta) dz \\ &\geq (2\delta)^{n-1} \cdot \tfrac{1}{2}\delta^{\frac{1}{2}} = 2^{n-2}\delta^{n-\frac{1}{2}}, \end{aligned}$$

where we have denoted by $D_{e_1} f$ the distributional derivative of f along e_1 and by $z +]-\delta, \delta[\cdot e_1$ the line segment joining $(-\delta, z)$ and (δ, z). □

We are ready to prove the main result.

Theorem 3.10 *Let f be in $\mathcal{F}_\rho(\Omega,\mathbb{R})$. Then, $\mathcal{H}^{n-\frac{1}{2}}(\mathcal{S}_f)\cap U$ is finite for every $U\subset\Omega$ open and bounded.*

Proof. We divide the set $\mathcal{S}_f$ into two sets:

$$\mathcal{S}_f=\mathcal{S}_f^1+\mathcal{S}_f^2$$

where

$$\mathcal{S}_f^1:=\Big\{x\in\mathcal{S}_f\mid N^P_{\text{hypo}(f)}(x,f(x))=\mathbb{R}^+(v,0)\Big\},$$
$$S_f^2:=\Big\{x\in\mathcal{S}_f\mid \dim_{\mathcal{H}}N^P_{\text{hypo}(f)}(x,f(x))\geq 2\Big\}.$$

Recalling 3.8, one can show that S_f^2 is $\mathcal{H}^{n-1}$-countably rectifiable. In particular, $\mathcal{H}^{n-\frac{1}{2}}(\mathcal{S}_f^2)=0$. Hence, we only need to prove that $\mathcal{H}^{n-\frac{1}{2}}(\mathcal{S}_f^1)\cap U$ is finite.

We can construct a covering of $S_f^1\cap U$ by setting:

$$\mathcal{B}:=\left\{(x+r\overline{\mathbb{B}}^n):\ x\in S_f^1\cap U,\ r<\frac{\min\{\delta_0(x),\text{dist}(U,\partial\Omega)/2\}}{10}\right\}.$$

Since $\mathcal{B}$ is a fine covering of $S_f^1\cap U$, by using Vitali's covering Theorem, there exists a countable subset of pairwise disjoint balls $\mathcal{B}':=\{x_i+r_i\overline{\mathbb{B}^n}:\ i\in\mathbb{N}\}\subset\mathcal{B}$ such that

$$\bigcup_{B\in\mathcal{B}}B\subseteq\bigcup_{i=1}^{\infty}(x_i+5r_i\mathbb{B}^n),$$

which implies that $\{x_i+5r_i\mathbb{B}^n:\ i\in\mathbb{N}\}$ is a covering of $S_f^1\cap U$ and

$$\bigcup_{i\in\mathbb{N}}(x_i+5r_i\mathbb{B}^n)\subseteq(U+c\mathbb{B}^n)=:W$$

for a suitable constant $c>0$ and thus W is an open bounded subset of Ω.

$$\begin{aligned}|Df|(W)&\geq|Df|\left(\bigcup_{i=1}^{\infty}(x_i+5r_i\mathbb{B}^n)\right)\\&\geq\sum_{i=1}^{\infty}|Df|(x_i+r_i\mathbb{B}^n)\geq\sum_{i=1}^{\infty}2^{-\frac{3}{2}}\cdot r_i^{n-\frac{1}{2}}\geq C\cdot\mathcal{H}^{n-\frac{1}{2}}(\mathcal{S}_f^1).\end{aligned}$$

and this implies that $|Df|(W)<+\infty$. Therefore, $\mathcal{H}^{n-\frac{1}{2}}(\mathcal{S}_f^1)<+\infty$. The proof is complete. □

Corollary 3.4 *Let f be in $\mathcal{F}_\rho(\Omega,\mathbb{R})$. Then, $\mathcal{L}^n(\mathcal{S}_f)=0$.*

We conclude this subsection with the following theorems.

Theorem 3.11 *Let f be in $\mathcal{F}_\rho(\Omega,\mathbb{R})$. Then, f is locally semiconcave in the open set $\Omega\backslash\mathcal{S}_f$. In particular, f is a.e. twice diffrentiable in Ω.*

Excercise 18. Prove the above theorem.

Theorem 3.12 *Let $f:\Omega\to\mathbb{R}$ be continuous. Assuming that* hypo(f) *satisfies a $\theta(\cdot)$-external sphere condition. Then, f is locally semiconcave in the open set $\Omega\backslash\mathcal{S}_f$.*

4 Introduction to scalar conservation laws

Let's consider 1D Hamilton-Jacobi equation

$$V_t + H(V_x) \; = \; 0 \qquad \text{for all } (t,x) \in [0,+\infty[\times\mathbb{R} \tag{4.1}$$

where the Hamiltonian $H : \mathbb{R} \to \mathbb{R}$ is smooth, convex and coercive, i.e.,

$$\lim_{p\to+\infty} \frac{H(p)}{|p|} \; = \; +\infty.$$

It is known that the Cauchy problem (4.1) with a Lipschitz initial data admits a unique Lipschitz viscosity solution V. In particular, V is differentiable almost everywhere and thus we can define

$$u(t,\cdot) \; = \; V_x(t,x) \quad \text{for a.e. } x \in R.$$

for every $t > 0$. Assume that u is smooth then it is a classical solution of the following equation

$$u_t(t,x) + H'(u) \cdot u_x(t,x) \; = \; 0.$$

The above equation can be rewritten in the conservative form

$$u_t + [H(u)]_x \; = \; 0. \tag{4.2}$$

In one dimensional case, (4.1) and (4.15) have a strong connection. Indeed, one can show that if V is a viscosity solution of (4.1) then $u = V_x$ is an entropy admissible solution of (4.15). The concept of entropy admissible solution will be introduced later.

4.1 The method of characteristic and non-smooth solution

In this section, we would like to study the scalar conservation laws in one space variable

$$u_t + [f(u)]_x \; = \; 0 \qquad (t,x) \in [0. + \infty[\times\mathbb{R} \tag{4.3}$$

where

- $f : \mathbb{R} \to \mathbb{R}$ is a given flux;
- $u : [0,+\infty[\times\mathbb{R} \to \mathbb{R}$ is the conserved quantity.

To feel the above equation better, let us give a typical example on traffic flow for (4.3).

Example 1. (Traffic flow) On a single road, let's denote by

- $\rho(t,x)$ is the traffic density at the location x at time t.

- v is the velocity of cars which depends on the traffic density such that

$$v = v(\rho) \qquad \text{with} \qquad \frac{dv}{d\rho} < 0.$$

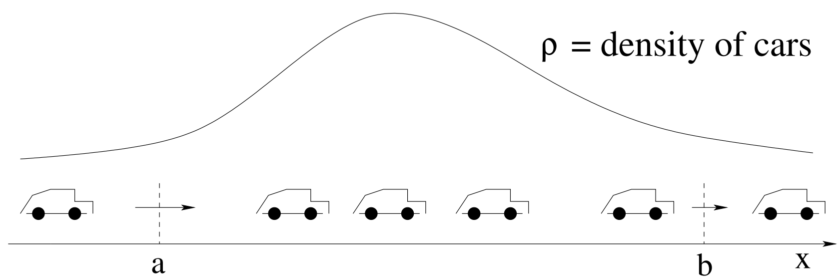
- The flux

$$f(\rho) \doteq \rho \cdot v(\rho)$$

describes the total number of cars crossing the location x at time t.

Giving two locations a and b on the road, the integral

$$\int_a^b \rho(t,x)\, dx = \text{total number of cars in } [a,b] \text{ at time } t\,.$$



We compute

$$\begin{aligned}\frac{d}{dt}\int_a^b \rho(t,x)\, dx &= f(\rho(t,a)) - f(\rho(t,b))\\ &= -\int_a^b \frac{d}{dx} f(\rho(t,x))\, dx\,.\end{aligned}$$

This implies that

$$\int_a^b \rho_t(t,x) + f(\rho(t,x))_x\, dx = 0 \qquad\qquad \text{for all } a < b\,.$$

A PDE for traffic flow

$$\rho_t(t,x) + f(\rho(t,x))_x = 0\,. \tag{4.4}$$

GOAL: *Describe the traffic density at time t. In other words, one would like to fine a solution to (4.4) for a give initial desity* ρ_0.

Assume that $f \in C^1(\mathbb{R})$ and u is a smooth solution of the Cauchy problem

$$\begin{cases} u_t + f(u)_x = 0\,, \\ u(x,0) = \Phi(x) \end{cases} \tag{4.5}$$

In this case, one can write the above equation as a quasilinear equation

$$u_t(t,x) + f'(u(t,x)) \cdot u_x(t,x) \;=\; 0.$$

<u>**The method of characteristic**.</u> Let $x(t)$ be the solution of

$$\dot{x}(t) \;=\; f'(u(x(t),t)), \qquad\qquad x(0) \;=\; \beta\,.$$

The curve $(x(t),t)$ is called a characteristic curve.

Observe that

$$\begin{aligned}\frac{d}{dt}\,u(x(t),t) &= u_t(x(t),t) + \dot{x}(t)\cdot u_x(x(t),t)\\ &= u_t(x(t),t) + f'(u(x(t),t))\cdot u_x(x(t),t) \;=\; 0\,.\end{aligned}$$

This implies that the function u is constant along the *characteristic curve* $(x(t),t)$. In particular, we have

$$u(x(t),t) \;=\; u(x(0),0) \;=\; \Phi(\beta)\,. \tag{4.6}$$

Hence,

$$f'(u(x(t),t)) \;=\; f'(\Phi(\beta))\,,$$

and it yields

$$x(t) \;=\; f'(\Phi(\beta))\cdot t + \beta\,.$$

Recalling (4.6), we obtain the general formula for the solution

$$u(\xi + f'(\Phi(\beta))t, t) \;=\; \Phi(\beta)\,.$$

Remark. *The method can be applied as long as the solution is smooth.*

Example 2. Solve the Burger's equation with initial condition

$$\begin{cases} u_t + \left(\dfrac{u^2}{2}\right)_x &= 0\,,\\ u(x,0) &= x\end{cases}$$

Answer. Since $f'(u) = u$ and $\Phi(x) = x$, one has

$$f'(\Phi(\beta)) \;=\; \beta\,.$$

Thus,

$$u(\beta + \beta\cdot t, t) \;=\; \Phi(\beta) \;=\; \beta\,.$$

Set $x = \beta + \beta\cdot t$, we have

$$\beta \;=\; \frac{x}{1+t}\,.$$

and this implies

$$u(x,t) \;=\; \frac{x}{t+1}\,.$$

for all $(t,x) \in [0,+\infty[\times\mathbb{R}$ $\square$

Example 3 (shock formation in Burgers' equation) Consider the scalar conservation law (inviscid Burgers' equation)

$$\begin{cases} u_t + \left(\dfrac{u^2}{2}\right)_x \;=\; 0\,, \\ u(x,0) \;=\; \bar{u}(x) \;=\; \dfrac{1}{1+x^2}. \end{cases} \tag{4.7}$$

Assume that u is smooth up to time $T>0$. In this case, u must be constant along the characteristic lines in the t-x plane:

$$t \mapsto (t, x + t\bar{u}(x)) \;=\; \left(t, x + \frac{t}{1+x^2}\right).$$

Moreover, these characteristic lines do not intersect before time T. This implies that the continuous map

$$x \mapsto x + \frac{t}{1+x^2}$$

is one-to-one for every $t \le T$. Thus, $x \mapsto x + \dfrac{t}{1+x^2}$ is monotone increasing and

$$\frac{d}{dx}\left(x + \frac{t}{1+x^2}\right) \;=\; 1 - \frac{2tx}{(1+x^2)^2} \;\geq\; 0 \qquad \text{for all } t \in [0,T], x \in \mathbb{R}. \tag{4.8}$$

A direct computation yields

$$\min_{x\in\mathbb{R}}\left(1 - \frac{2tx}{(1+x^2)^2}\right) \;=\; 1 - \frac{t\sqrt{27}}{8} \qquad \text{for all } [0,T].$$

Thus, (4.7) admits a smooth solution up to time $t < \dfrac{8}{\sqrt{27}}$ and then generates a discontinuity at time $t = \dfrac{8}{\sqrt{27}}$. $\square$

4.2 Entropy admissible weak solutions

The above example showed that a basic feature of nonlinear systems of the form (4.3) is that, even for smooth initial data, the solution of the Cauchy problem may develop discontinuities in finite time. In order to prolong solution to (4.3) after the formation of discontinuity we must adopt a weak concept of solution in distributional senses which allow the presence of discontinuities in the solution or in its space derivatives.

4.2.1 Weak solutions

Definition 4.1 *A function* $u \in \mathbf{L}^\infty(]0,T[\times\mathbb{R},\mathbb{R})$ *is a weak solution of the scalar conservation laws*

$$u_t + [f(u)]_x \ = \ 0,$$

if for every $\varphi \in \mathcal{C}^1(]0,T[\times\mathbb{R},\mathbb{R})$ *with compact support, it holds*

$$\int\int_{]0,T\times\mathbb{R}} [u(t,x)\varphi_t(t,x) + f(u(t,x))\varphi_x(t,x)]\, dtdx \ = \ 0. \tag{4.9}$$

Remark 4.2 (Classical solution) *A function* $u \in \mathcal{C}^1(]0,T[\times\mathbb{R},\mathbb{R})$ *is a classical solution of (4.3) if and only if* u *is a weak solution of (4.3).*

Proof. 1. Assume that u is a classical solution of (4.3). For a given $\varphi \in \mathcal{C}^1(]0,T[\times\mathbb{R},\mathbb{R})$ with compact support, we consider the vector field

$$\mathbf{v}(t,x) \ = \ (u(t,x)\cdot\varphi(t,x), f(u(t,x))\cdot\varphi(t,x)).$$

Let $\Omega \subset]0,T[\times\mathbb{R}$ be an open set such that $\mathrm{supp}(v) \subset \Omega$. By the divergence theorem, we have

$$\begin{aligned} 0 \ &= \ \int_{\partial\Omega} \mathbf{v}\cdot\mathbf{n}ds \ = \ \int\int_{]0,T[\times\mathbb{R}} \mathrm{div}\,\mathbf{v}dtdx & (4.10)\\ &= \ \int\int_{]0,T[\times\mathbb{R}} [u_t + f(u)_x]\cdot\varphi dtdx + \int\int_{]0,T[\times\mathbb{R}} u\varphi_t + f(u)\varphi_x dtdx & (4.11)\\ &= \ \int\int_{]0,T[\times\mathbb{R}} u\varphi_t + f(u)\varphi_x dtdx & (4.12)\end{aligned}$$

and this implies that u is a weak solution of (4.3).

2. Assume that u is a weak solution but not a classical solution of (4.3). Since u is smooth, there exists a point (t_0,x_0) such that

$$u_t(t_0,x_0) + [f(u(t_0,x_0))]_x \ \neq \ 0.$$

Without loss of generality, we will assume that the left hand side of the above equation is positive. By the smoothness of u and f, it holds

$$u_t(t,x) + [f(u(t,x))]_x \ > \ 0$$

for every (t,x) in $B_\delta(t_0,x_0) \subset]0,T[\times\mathbb{R}$ for some $\delta > 0$ small. Consider a nonnegative and nonzero function $\varphi \in \mathcal{C}^1(]0,T[\times\mathbb{R},\mathbb{R})$ with a compact support such that

$$\varphi(t,x) \ = \ \begin{cases} 0 & (t,x) \in]0,T[\times\mathbb{R}\backslash B_\delta(t_0,x_0) \\ \\ d_{\partial B_\delta}(t_0,x_0) & (t,x) \in B_\delta(t_0,x_0). \end{cases}$$

Apply this test function to (4.9), we obtain that

$$\begin{aligned} 0 &= \int\int_{]0,T[\times\mathbb{R}} u(t,x)\varphi_t(t,x) + f(u(t,x))\varphi_x(t,x)dtdx \\ &= -\int\int_{]0,T[\times\mathbb{R}} [u_t(t,x) + f(u)_x(t,x)]\cdot\varphi(t,x)dtdx \ < \ 0, \end{aligned}$$

and this yields a contradiction. □

Lemma 4.3 (Closure of set of weak solutions in $\mathbf{L}^1_{\text{loc}}$) *Let $(u_n)_{n\geq 1}$ be a sequence of weak solutions of (4.3) such that*

$$u_n \longrightarrow u \qquad \text{and} \qquad f(u_n) \longrightarrow f(u) \qquad \text{in } \mathbf{L}^1_{\text{loc}}. \tag{4.13}$$

Then the limit function u is also a weak solution of (4.3). Moreover, the same conclusion holds if $u_n \longrightarrow u$ in $\mathbf{L}^1_{\text{loc}}$ and

$$u_n(\mathbb{R}) \ \subseteq \ K \qquad \textit{for all } n\geq 1$$

for some compact set K.

Proof. Assume that (4.13) holds. For every $\varphi\in\mathcal{C}^1(]0,T[\times\mathbb{R},\mathbb{R})$ with compact support, we has

$$\int\int_{]0,T[\times\mathbb{R}} u_n\varphi_t + f(u_n)\varphi_x dtdx \ = \ \int\int_{\Omega} u_n\varphi_t + f(u_n)\varphi_x dtdx$$

and

$$\int\int_{]0,T[\times\mathbb{R}} u\varphi_t + f(u)\varphi_x dtdx \ = \ \int\int_{\Omega} u\varphi_t + f(u)\varphi_x dtdx$$

for some open bounded set. Thus,

$$\begin{aligned} &\limsup_{n\to\infty}\left|\int\int_{]0,T[\times\mathbb{R}} [u_n\varphi_t + f(u_n)\varphi_x] - [u\varphi_t + f(u)\varphi_x]dtdx\right| \\ &\qquad\leq \ \left(\limsup_{n\to\infty}\int\int_{\Omega} |u_n - u| + |f(u_n) - f(u)|dtdx\right)\cdot\|\nabla\varphi\|_\infty \ = \ 0. \end{aligned}$$

and this implies that

$$\int\int_{]0,T[\times\mathbb{R}} u\varphi_t + f(u)\varphi_x dtdx \ = \ \lim_{n\to\infty}\int\int_{]0,T[\times\mathbb{R}} u_n\varphi_t + f(u_n)\varphi_x dtdx \ = \ 0.$$

Therefore, u is a weak solution of (4.3). □

Let us now derive a consequence of conservation form of (4.3). Given any time $t_1, t_2\in]0,T[$, consider a domain

$$\Omega \ = \ \{(t,x) \mid t\in[t_1,t_2], \gamma_1(t) < x < \gamma_2(t)\}$$

for some $\gamma_i : [t_1,t_2]\to\mathbb{R}$ Lipschitz curves.

Remark 4.4 *If u is a classical solution of (4.3) then*

$$\begin{aligned}
&\int_{\gamma_1(t_2)}^{\gamma_2(t_2)} u(t_2,x)dx - \int_{\gamma_1(t_1)}^{\gamma_2(t_1)} u(t_1,x)dx \\
&= \int_{t_1}^{t_2} \dot{\gamma}_2(t)u(t,\gamma_2(t)) - f(u(t,\gamma_2(t)))dt - \int_{t_1}^{t_2} \dot{\gamma}_1(t)u(t,\gamma_1(t)) - f(u(t,\gamma_2(t)))dt.
\end{aligned} \tag{4.14}$$

Proof. Applying the divergence theorem to the vector field

$$\mathbf{v}(t,x) \;=\; (u(t,x), f(u(t,x))),$$

we find that

$$\begin{aligned}
0 \;&=\; \int\int_{\Omega} u_t + [f(u)]_x dtdx \;=\; \int\int_{\Omega} \operatorname{div}\mathbf{v} dtdx \;=\; \int_{\partial\Omega} \mathbf{v}\cdot\mathbf{n}ds \\
&= \int_{t_1}^{t_2} \dot{\gamma}_1(t)u(t,\gamma_1(t)) - f(u(t,\gamma_1(t)))dt + \int_{\gamma_1(t_2)}^{\gamma_2(t_2)} u(t_2,x)dx \\
&\quad - \int_{t_1}^{t_2} \dot{\gamma}_2(t)u(t,\gamma_2(t)) - f(u(t,\gamma_2(t)))dt - \int_{\gamma_1(t_1)}^{\gamma_2(t_1)} u(t_1,x)dx.
\end{aligned}$$

This implies (4.14). □

The formula (4.14) tell us that the variation of the quantity of u contained between γ_1 and γ_2 at different times $t_1 < t_2$ is given by the flow of the vector field $\mathbf{v}$ through the two curves γ_1, γ_2. It also holds for a week solution u provided that

(i) the map $t \mapsto u(t,\cdot)$ is continuous with values in $\mathbf{L}^1_{\text{loc}}$;

(ii) the map $x \mapsto u(t,x)$ is right continuous for all $(t,x) \in]0,T[\times\mathbb{R}$, i.e.,

$$u(t,x) \;=\; \lim_{y\to x^+} u(t,y).$$

Let's now define a weak solution of a Cauchy problem

$$\begin{cases} u_t + [f(u)]_x &= \; 0 \\ u(0,\cdot) &= \; u_0(\cdot) \end{cases} \tag{4.15}$$

for a given $u_0 \in \mathbf{L}^1_{\text{loc}}(\mathbb{R})$.

Definition 4.5 *A function $u : [0,T]\times\mathbb{R} \to \mathbb{R}$ is a weak solution of (4.15) if u is a weak solution of (4.3) in the strip $]0,T[\times\mathbb{R}$ and the map $t \mapsto u(t,\cdot)$ is continuous with values in $\mathbf{L}^1_{\text{loc}}$ for $t \in [0,T]$ with $u(0,\cdot) = u_0(\cdot)$.*

The following holds:

Lemma 4.6 *If $u : [0,T] \times \mathbb{R} \to \mathbb{R}$ is a weak solution of (4.15) then u is also a solution of (4.15) in the distributional sense, i.e., for every $\varphi \in \mathcal{C}^1(]-1,T[\times\mathbb{R},\mathbb{R})$ with compact support, it holds*

$$\int\int_{]0,T\times\mathbb{R}} [u\varphi_t + f(u)\varphi_x]\, dtdx + \int_{-\infty}^{\infty} u_0(x)\varphi(0,x)dx \;=\; 0.$$

Proof. Let $\rho : \mathbb{R} \to [0,1]$ be a smooth function with $\mathrm{supp}(\rho) \subset]0,1[$ and

$$\int_0^1 \rho(s)ds \;=\; 1.$$

For every $\varepsilon > 0$, denote by

$$\rho_\varepsilon \;=\; \frac{1}{\varepsilon}\cdot\rho\left(\frac{t}{\varepsilon}\right) \qquad \text{and} \qquad \beta_\varepsilon(t) \;=\; \int_0^t \rho_\varepsilon(s)ds,$$

we set

$$\varphi^\varepsilon(t,x) \;=\; \beta_\varepsilon(t)\cdot\varphi(t,x) \qquad \text{for all } (t,x) \in [0,T]\times\mathbb{R}.$$

It is clear that $\mathrm{supp}(\varphi^\varepsilon) \subset]0,T[\times\mathbb{R}$ and

$$\lim_{\varepsilon\to 0+} \varphi^\varepsilon(t,x) \;=\; \varphi(t,x) \qquad \text{for all } (t,x) \in [0,T]\times\mathbb{R}.$$

Using the continuity of the map $t \mapsto u(t,\cdot)$, we compute that

$$\begin{aligned} 0 \;&=\; \int\int_{]0,T\times\mathbb{R}} [u\varphi^\varepsilon_t + f(u)\varphi^\varepsilon_x]\, dtdx \\ &=\; \int_0^T \int_{-\infty}^{-\infty} [u\varphi_t + f(u)\varphi_x]\cdot\beta^\varepsilon(t)dxdt + \int_0^T\int_{-\infty}^{\infty} u\varphi\rho_\varepsilon(t)dxdt. \end{aligned}$$

Taking ε to $0+$, we then obtain that

$$\int\int_{]0,T\times\mathbb{R}} [u\varphi_t + f(u)\varphi_x]\, dtdx + \int_{-\infty}^{\infty} u_0(x)\varphi(0,x)dx \;=\; 0.$$

Therefore, u is a solution of (4.15) in the distributional sense. □

4.2.2 Rankine-Hugoniot conditions

Let us first derive conditions which must be satisfied piecewise constant function

$$U(t,x) \;=\; \begin{cases} u^+ & \text{if } x > \lambda\cdot t \\ u^- & \text{if } x < \lambda\cdot t. \end{cases} \tag{4.16}$$

for some $u^\pm, \lambda \in R$, to be a weak solution of (4.3).

Lemma 4.7 *The function U in (4.16) is a weak solution of (4.3) if any only if*

$$\lambda \cdot (u^+ - u^-) \;=\; f(u^+) - f(u^-). \tag{4.17}$$

Proof. For every $\varphi \in \mathcal{C}^1(]0,-\infty[\times\mathbb{R})$ with $\mathrm{supp}(\varphi) \subset \Omega$, we denote by

$$\Omega^+ \;=\; \{(t,x) \in \Omega \mid x > \lambda t\}, \quad \Omega^+ \;=\; \{(t,x) \in \Omega \mid x < \lambda t\}.$$

Normal vectors to the line $x = \lambda t$ are

$$\mathbf{n}^+ \;=\; (\lambda, -1), \qquad \mathbf{n}^-(-\lambda, 1).$$

Consider the vector field

$$\mathbf{v}(t,x) \;=\; (u(t,x)\cdot\varphi(t,x), f(u(t,x))\cdot\varphi(t,x)).$$

By the divergence theorem, we obtain that

$$\begin{aligned}
\int\int_\Omega U\varphi_t + f(U)\varphi_x dxdt \;&=\; \int\int_{\Omega^+} \mathrm{div}\,\mathbf{v} dxdt + \int\int_{\Omega^-} \mathrm{div}\,\mathbf{v} dxdt \\
&=\; \int_{\partial\Omega^+} \mathbf{v}\cdot\mathbf{n}^+ ds + \int_{\partial\Omega^-} \mathbf{v}\cdot\mathbf{n}^- ds \\
&=\; \int [\lambda u^+ - f(u^+)]\cdot\varphi(t,\lambda t)dt + \int [\lambda - u^- + f(u^-)]\cdot\varphi(t,\lambda t)dt \\
&=\; \int \big[\lambda\cdot(u^+ - u^-) - \big(f(u^+) - f(u^-)\big)\big]\cdot\varphi(t,\lambda t)dt.
\end{aligned}$$

Therefore, U is a weak solution of (4.3) if and only if

$$\int \big[\lambda\cdot(u^+ - u^-) - \big(f(u^+) - f(u^-)\big)\big]\cdot\varphi(t,\lambda t)dt \;=\; 0$$

for all $\varphi \in \mathcal{C}^1(]0,-\infty[\times\mathbb{R})$ with compact support. It is equivalent to (4.17). □

Remark 4.8 *The equation (4.17) is famous Rankine-Hugoniot (RH) condition.*

Example. Consider the Burgers' equation

$$u_t + \left(\frac{u^2}{2}\right)_x \;=\; 0.$$

Given two different state $u^+ \neq u^-$, we have

$$\lambda \;=\; \frac{f(u^+) - f(u^-)}{u^+ - u^-} \;=\; \frac{u^+ + u^-}{2}.$$

The function

$$u(t,x) \;=\; \begin{cases} u^+ & \text{if } x > \frac{u^+ + u^-}{2}\cdot t \\ u^- & \text{if } x < \frac{u^+ + u^-}{2}\cdot t. \end{cases}$$

is a weak solution.

To derive (RH) condition for general weak solutions of (4.3), let us introduce the following:

Definition 4.9 (Approximate jump) *We say that a function $u \in \mathbf{L}^1_{\text{loc}}$ with value in $\mathbb{R}$ has an approximate jump discontinuity at a point $(\bar t, \bar x)$ if there exists $u^{\pm}, \lambda \in \mathbb{R}$ such that setting*

$$U(t,x) \;=\; \begin{cases} u^+ & \text{if } x > \lambda \cdot t \\ u^- & \text{if } x < \lambda \cdot t. \end{cases}$$

there holds

$$\lim_{r\to 0+} \frac{1}{r^2} \int\int_{[-r,r]^2} \big|u(\bar t + t, \bar x + x) - U(t,x)\big| dxdt \;=\; 0. \tag{4.18}$$

In this case, u^- and u^+ are the left and the right approximate limit of u at $(\bar t, \bar x)$, and λ is the jump speed. Moreover, we say that u is approximately continuous at $(\bar t, \bar x)$ if (4.18) holds for a function U defined as in (4.16) with $u^- = u^+$.

We will show that (RH) condition is satisfied at any point of approximate jump discontinuity of a weak solution of (4.3).

Proposition 4.9.1 *Let u be a bounded weak solution of (4.3) having an approximate jump discontinuity at a point $(\bar t, \bar x)$, i.e., (4.18) holds for some $u^{\pm}, \lambda \in R$. Then, the Rankine-Hugoniot equation (4.17) holds.*

Proof. **1.** For any fixed $\theta > 0$ sufficiently small, one can easily check that the rescaled function

$$u^{\theta}(t,x) \;=\; u(\bar t + \theta t, \bar x + \theta x)$$

is a weak solution of (4.3). Indeed, for any test function $\varphi \in \mathcal{C}^1(]0,\infty[\times\mathbb{R})$ with compact support, we find

$$\int_0^{\infty}\int_{-\infty}^{\infty} [u^{\theta}(t,x)\varphi_t(t,x) + f(u(t,x))\varphi_x(t,x)]dxdt$$
$$= \frac{1}{\theta^2}\cdot \int_{\bar t}^{\infty}\int_{-\infty}^{\infty} \left[u(\tau,z)\varphi_t\left(\frac{\tau-\bar t}{\theta}, \frac{z-\bar x}{\theta}\right) + f(u(\tau,z))\varphi_x\left(\frac{\tau-\bar t}{\theta}, \frac{z-\bar x}{\theta}\right)\right] d\tau dz.$$

Since the function $\varphi^{\theta}(\tau,z) = \varphi\left(\frac{\tau-\bar t}{\theta}, \frac{z-\bar x}{\theta}\right)$ is smooth with compact support, one has

$$\int_0^{\infty}\int_{-\infty}^{\infty} [u^{\theta}\varphi_t + f(u)\varphi_x]dxdt \;=\; \frac{1}{\theta^2}\cdot\int_{\bar t}^{\infty}\int_{-\infty}^{\infty} \left[u(\tau,z)\varphi^{\theta}_t + f(u(\tau,z))\varphi^{\theta}_x\right] d\tau dz \;=\; 0.$$

Hence, u^{θ} is a weak solution of (4.3).

2. We claim that u^{θ} converges to U in $\mathbf{L}^1_{\text{loc}}$ as $\theta \to 0^+$. Indeed, for any $R > 0$, one has that

$$\int_{-R}^{R}\int_{-R}^{R} \left|u^{\theta}(t,x) - U(t,x)\right| dxdt$$
$$= \frac{1}{\theta^2}\cdot\int_{\theta R}^{\theta R}\int_{-\theta R}^{\theta R} \left|u(\bar t + \tau, \bar x + z) - U\left(\frac{\tau}{\theta}, \frac{z}{\theta}\right)\right| d\tau dz$$
$$= R^2\cdot\left[\frac{1}{[\theta R]^2}\cdot\int_{\theta R}^{\theta R}\int_{-\theta R}^{\theta R} \left|u(\bar t + \tau, \bar x + z) - U(\tau,z)\right| d\tau dz\right]$$

Taking θ to $0+$, we then obtain from (4.18) that

$$\lim_{\theta\to 0+}\int_{-R}^{R}\int_{-R}^{R}\left|u^{\theta}(t,x)-U(t,x)\right|dxdt \;=\; 0.$$

3. Since u is bounded, the sequence of function u^{θ} is also bounded. Thus, one can apply Lemma 4.3 to derive that U is a weak solution of (4.3) and Lemma 4.28 implies the Rankine-Hugoniot condition. □

To conclude this part, we will provide necessary and sufficient conditions for a piecewise Lipschitz continuous function to be a weak solution. Here, we say that u is piecewise Lipschitz if u is in $\mathbf{L}^{\infty}$ and there exist a finite number of points $P_i=(t_i,x_i)$ and finitely many disjoint Lipschitz-continuous curves $\gamma_j:]a_j,b_j[\to\mathbb{R}$ such that

- For any point P outside the set $\left[(\bigcup P_i)\bigcup\left(\bigcup_j\gamma(]a_j,b_j[)\right)\right]$, u is Lipschitz continuous in $B(P,r)$ for some $r>0$;
- Given a point $Q\in\gamma_j(]a_j,b_j[)$ for some j, there exists a neighborhood V of Q such that u is Lipschitz continuous in

$$V^{+} \;=\; V\bigcap\{x>\gamma_j(t)\}\qquad\text{and}\qquad V^{-} \;=\; V\bigcap\{x<\gamma_j(t)\}$$

Assume that u is piecewise Lipscthiz, we can define

$$u_j^{+}(t) \;=\; \lim_{x\to\gamma_j(t)+}u(t,x)\qquad\text{and}\qquad u_j^{-}(t) \;=\; \lim_{x\to\gamma_j(t)-}u(t,x).$$

Proposition 4.9.2 *Let $u:\Omega\to\mathbb{R}$ be piecewise Lipschitz. Then, the followings are equivalent*

(i) u is a weak solution of (4.3);

(ii) u satisfies the quasilinear equation

$$u_t+f'(u)u_x \;=\; 0 \tag{4.19}$$

for almost every (t,x). Moreover, for every jump curve γ_j one has

$$\dot{\gamma}_j(t)\cdot\left[u_j^{+}(t)-u_j^{-}(t)\right]=f(u_j^{+}(t))-f(u_j^{-}(t)) \tag{4.20}$$

for almost every $t\in]a_j,b_j[$.

Sketch of Proof. 1. Assume that u is a weak solution of (4.3). For any point P out side the set $\left[(\bigcup P_i)\bigcup\left(\bigcup_j\gamma(]a_j,b_j[)\right)\right]$, u is Lipschitz continuous in $B(P,r)$

for some $r > 0$. For every given $\varphi \in \mathcal{C}^1$ with $supp(u) \subset B(P,r)$, one applies the divergence theorem to get

$$0 \ = \ \int\int_{B(P,r)} [u_t + [f(u)]_x] \cdot \varphi dtdx \ = \ 0.$$

Here, we used also Rademacher's theorem to say that $\operatorname{div}(u, f(u))$ is defined almost everywhere in $B(P,r)$ and is in $\mathbf{L}^\infty$. Thus, u satisfies (4.19).

On the other hand, one can use Proposition 4.9.1 to show that (4.20) satisfies at any time t where γ_j is differentiable.

2. Assume that u satisfies (ii). Let's consider the case where the set $(\bigcup P_i)$ is empty. For every $\varphi \in \mathcal{C}_c^1$, one apply the divergence theorem for $\mathbf{v} = (u\varphi), f(u)\varphi)$ to obtain that

$$\begin{aligned} 0 \ &= \ \int\int_\Omega u\varphi_t + f(u)\varphi_x dtdx \ = \ -\int\int_\Omega [u_t + f'(u)u_x] \cdot \varphi dtdx \\ &- \sum_j \int_{a_j}^{b_j} \left(\dot\gamma_j(t) \cdot \left[u_j^+(t) - u_j^-(t)\right] - [f(u_j^+(t)) - f(u_j^-(t))] \cdot \varphi(t, \gamma_j(t))\right) dt. \end{aligned}$$

Thus, u is a weak solution of (4.3).

3. For the general case where the set $(\bigcup P_i)$ is non-empty. For every given $\varphi \in \mathcal{C}_c^1$, one can construct a sequence of $(\varphi^n)_{n\geq 1} \subset \mathcal{C}_c^1$ such that

$$P_i \ \notin \ \operatorname{supp}(\varphi_n) \qquad \text{for all } i, n \geq 1$$

and $\lim_{n\to+\infty} \|\nabla\varphi_n - \nabla\varphi\|_{\mathbf{L}^1} = 0$. From the second step, one can obtain that

$$\begin{aligned} 0 \ &= \ \int\int_\Omega u\varphi_t + f(u)\varphi_x dtdx \\ &= \ \lim_{n\to+\infty} \left[\int\int_\Omega u\varphi_t^n + f(u)\varphi_x^n dtdx\right] \ = \ \lim_{n\to+\infty} \Big[-\int\int_\Omega [u_t + f'(u)u_x] \cdot \varphi dtdx \\ &- \sum_j \int_{a_j}^{b_j} \left(\dot\gamma_j(t) \cdot \left[u_j^+(t) - u_j^-(t)\right] - [f(u_j^+(t)) - f(u_j^-(t))]\right) \cdot \varphi^n(t, \gamma_j(t)) dt\Big] \ = \ 0. \end{aligned}$$

The proof is complete. □

4.2.3 Admissible conditions

Consider the Burgers' equation

$$u_t + \left(\frac{u^2}{2}\right)_x \ = \ 0 \quad \text{with} \quad u(0,x) \ = \ \begin{cases} 1 & x \geq 0 \\ 0 & x < 0. \end{cases}$$

For every $\alpha \in [0,1]$, the piecewise constant function

$$u_\alpha(t,x) \;=\; \begin{cases} 0 & \text{if} \quad x < \dfrac{\alpha}{2}t \\ \alpha & \text{if} \quad \dfrac{\alpha}{2}t \le x < \dfrac{\alpha+1}{2}\cdot t \\ 1 & \text{if} \quad x \ge \dfrac{1+\alpha}{2}\cdot \end{cases}$$

is a weak solution of (4.3). This shows that the concept of weak solution is not sufficient to single out a unique solution whenever a strong discontinuity appears in the solution.

Our goal is to supplement the notion of weak solution with further admissibility conditions, that can be present in a weak solution, in order to achieve uniqueness and continuous dependence on the initial data of the solutions.

1. Vanishing viscosity: We say that a weak solution $u : \Omega \to \mathbb{R}$ of (4.3) is admissible in the *vanishing viscosity sense* if there exists a sequence of smooth solutions of the viscous parabolic approximation

$$u_t^\varepsilon + f'(u^\varepsilon)\cdot u_x^\varepsilon \;=\; \varepsilon u_{xx}^\varepsilon \tag{4.21}$$

so that u^ε converges to u in $\mathbf{L}^1_{\mathrm{loc}}$ as $\varepsilon \to 0+$.

In general, it is very difficult to establish a-priori estimates on solutions to (4.3) that allow to prove the convergence as $\varepsilon \to 0^+$ and to characterize the corresponding limit. However, one can deduce from the vanishing viscosity condition other conditions that can be more easily verified in practice.

2. Entropy conditions: Motivated by the second principle of theorem dynamics for the Euler equation of gas, we introduce the concept of *entropy* which characterize *irreversible processes* (Kinetic energy dissipates when a shock appears: a part of it is transformed into heat).

Definition 4.10 (Entropy-Entropy flux) *We say that a pair of $\mathcal{C}^1$ (or locally Lipschitz) functions $(\eta, q) : \mathbb{R} \to \mathbb{R}$ is an entropy-entropy flux pair for (4.3) if*

$$q'(u) \;=\; \eta'(u)\cdot f'(u) \tag{4.22}$$

at every u where η, q and f are differentiable.

Notice that if u is a classical solution of (4.3) then u solves the equation

$$[\eta(u)]_t + [q(u)]_x \;=\; 0. \tag{4.23}$$

Remark 4.11 *In this case, $\eta(u)$ is conserved. However, when u is discontinuous, in general the quantity $\eta(u)$ is not conserved.*

Indeed, let's consider the Burgers' equation

$$u_t + \left(\frac{u^2}{2}\right)_x \;=\; 0$$

and a pair of entropy-entropy flux

$$(\eta, q) \;=\; \left(u^3, \frac{3}{4}\cdot u^4\right) \qquad \text{for all } u \in \mathbb{R}.$$

The following function

$$u(t,x) \;=\; \begin{cases} 1 & \text{if} \quad x < \dfrac{t}{2} \\ \\ 0 & \text{if} \quad x \geq \dfrac{t}{2} \end{cases}$$

satisfies Rankine-Hugoniot condition and thus is a weak solution of (4.3). However, u is not a weak solution of (4.23) since it does not satisfies Rankine-Hugoniot condition at 0 and 1, i.e.,

$$q(1) - q(0) \;=\; \frac{3}{4} \;\neq\; \frac{1}{2} \;=\; \frac{1}{2}\cdot(\eta(1) - \eta(0)).$$

$\square$

Entropy admissible solution. Let u^ε be the smooth solution of (4.21). It is easy to see that u^ε is also a solution of the following equation

$$[\eta(u^\varepsilon)]_t + [q(u^\varepsilon)]_x \;=\; \varepsilon\cdot\left([\eta(u^\varepsilon)]_{xx} - \eta''(u^\varepsilon)\cdot[u_x^\varepsilon]^2\right).$$

In particular, if η is convex and smooth then one has that

$$[\eta(u^\varepsilon)]_t + [q(u^\varepsilon)]_x \;\leq\; \varepsilon\cdot\eta(u^\varepsilon)]_{xx}$$

Thus, for every non-negative test functions $\varphi \in C_c^1$, it holds

$$\int\int_\Omega [\eta(u^\varepsilon)\varphi_t + q(u^\varepsilon)\varphi_x]\, dtdx \;\geq\; -\,\varepsilon\int\int_\Omega \eta(u^\varepsilon)\varphi_{xx} dtdx.$$

If u^ε converges to u in $\mathbf{L}^1_{\mathrm{loc}}$ then by taking ε to $0+$, we get

$$\int\int_\Omega [\eta(u)\varphi_t + q(u)\varphi_x]\, dtdx \;\geq\; 0. \tag{4.24}$$

This yields the following entropy admissible condition.

Definition 4.12 *A weak solution u of (4.3) is entropy admissible if it satisfies the inequality*

$$[\eta(u)]_t + [q(u)]_x \ \leq \ 0$$

in the distributional senses for every pair of convex entropy-entropy flux (η, q), i.e. (4.24) holds for every non-negative test functions $\varphi \in C^1_c$.

A particular class of entropy-entropy flux pairs which is quite useful in analyzing the behavior of entropy admissible weak solutions is given by the Kruzkhov's entropy: for each $k \in \mathbb{R}$, consider the functions

$$\eta_k(u) \ = \ |u-k|, \qquad q_k(u) \ = \ \mathrm{sign}(u-k)\cdot(f(u)-f(k)).$$

It is easy to check that $(\eta_k(u), q_k(u))$ is locally Lipschitz and satisfies (4.22) for every $u \neq k$.

Proposition 4.12.1 *A function $u \in \mathbf{L}^\infty(]0,T[\times\mathbb{R})$ is an entropy admissible weak solution of (4.3) iff for every $k \in \mathbb{R}$ it holds*

$$\int\int_{]0,T[\times\mathbb{R}} [\eta_k(u)\varphi_t + q_k(u)\varphi_x]\, dtdx \ \geq \ 0. \tag{4.25}$$

for all $\varphi \in \mathcal{C}^1_c(]0,T[\times\mathbb{R}, [0,+\infty[)$.

Sketch of proof. The proof is divided into several steps:

1. We show that if a function $u \in \mathbf{L}^\infty(]0,T[\times\mathbb{R})$ satisfies (4.25) for every $k \in \mathbb{R}$ then u is a weak solution of (4.3). Set $M = \|u\|_\infty + 1$, we have that

$$\eta_M(u) \ = \ M-u \qquad \text{and} \qquad q_M(u) \ = \ f(M)-f(u).$$

For every $\varphi \in \mathcal{C}^1_c(]0,T[\times\mathbb{R}, [0,+\infty[)$, we have

$$\int\int_{]0,T[\times\mathbb{R}} [\eta_M(u)\varphi_t + q_M(u)\varphi_x]\, dtdx \ \geq \ 0$$

and this implies that

$$\int\int_{]0,T[\times\mathbb{R}} [u\varphi_t + f(u)\varphi_x]\, dtdx \ \leq \ \int\int_{]0,T[\times\mathbb{R}} [M\varphi_t + f(M).\varphi_x]\, dtdx \ = \ 0$$

Similarly, since

$$\int\int_{]0,T[\times\mathbb{R}} [\eta_{-M}(u)\varphi_t + q_{-M}(u)\varphi_x]\, dtdx \ \geq \ 0,$$

one has

$$\int\int_{]0,T[\times\mathbb{R}} [u\varphi_t + f(u)\varphi_x]\, dtdx \ \geq \ -\int\int_{]0,T[\times\mathbb{R}} [M\varphi_t + f(M).\varphi_x]\, dtdx \ = \ 0.$$

Thus, for all $\varphi \in \mathcal{C}_c^1(]0,T[\times\mathbb{R},[0,+\infty[)$, it holds

$$\int\int_{]0,T[\times\mathbb{R}} [u\varphi_t + f(u)\varphi_x]\, dtdx \ = \ 0. \tag{4.26}$$

2. By an approximation argument, one can show that (4.26) also holds for all $\varphi \in \mathrm{Lip}_c(]0,T[\times\mathbb{R},[0,+\infty[)$. Therefore, for a general $\varphi \in \mathcal{C}_c^1(]0,T[\times\mathbb{R},\mathbb{R})$, it holds

$$\varphi \ = \ \varphi^+ - \varphi^- \qquad \text{with} \qquad \varphi^\pm \ = \ \frac{|\varphi| \pm \varphi}{2} \ \in \ \mathrm{Lip}_c(]0,T[\times\mathbb{R},[0,+\infty[).$$

Since (4.26) holds for both φ^+ and φ^-, one then has that (4.26) holds for φ.

3. To conclude the proof, we show that (4.24) is satisfied for every pair of convex entropy-entropy flux (η, q). It is divided in two main steps:

- Show that (4.24) every pair of convex entropy-entropy flux (η, q) where η is convex piecewise affine entropy of the form

$$\eta(u) \ = \ a_0 + a_i u + c_i \sum_{i=1}^{N} \frac{|u-k_i| + u - k_i}{2} \tag{4.27}$$

  for $a_0, a_1 \in \mathbb{R}$ and $c_i > 0$.

- For every pair of convex entropy-entropy flux (η, q), one can approximate uniformly (outside a set of measure zero) by a sequence of pair of convex entropy-entropy flux (η_n, q_n) with η_n talking the form (4.27).

□

Stability conditions. We wish to derive simple geometric conditions that can obtained purely from stability considerations, without any reference to physical models. Let us recall

$$U(t,x) \ = \ \begin{cases} u^+ & \text{if } x > \lambda \cdot t \\ u^- & \text{if } x < \lambda \cdot t. \end{cases}$$

a weak solution of (4.3), i.e.,

$$\lambda \ = \ \frac{f(u^+) - f(u^-)}{u^+ - u^-}.$$

Consider a slightly perturbed solution where the original shock joining two states $u^\pm$ is split into two separated smaller shocks that join u^+ and u^- with an intermediate state $u^\alpha = \alpha u^+ + (1-\alpha)u^+$ for some $\alpha \in]0,1[$. To ensure that the $\mathbf{L}^1$-distance between the original solution and the perturbed one does not increase in time, we need the following:

$$\text{[speed of jump behind]} \ \geq \ \text{[speed of jump ahead]}$$

By Rankine-Hugoniot condition, one has

$$\frac{f(u^\alpha) - f(u^-)}{u^\alpha - u^-} \;\geq\; \frac{f(u^+) - f(u^\alpha)}{u^+ - u^\alpha} \tag{4.28}$$

The above condition is equivalent to

$$\begin{cases} f(\alpha u^+ + (1-\alpha)u^-) \;\geq\; \alpha f(u^+) + (1-\alpha) f(u^-) & \text{if} \quad u^- < u^+ \\ f(\alpha u^+ + (1-\alpha)u^-) \;\leq\; \alpha f(u^+) + (1-\alpha) f(u^-) & \text{if} \quad u^- > u^+ \end{cases} \tag{4.29}$$

Proposition 4.12.2 *The function U is an entropy admissible solution of (4.3) iff (4.29) holds.*

Sketch of the proof. 1. One can show that U is entropy admissible weak solution of (4.3) if any only if

$$\lambda \cdot (\eta_k(u^+) - \eta_k(u^-)) \;\geq\; q_k(u^+) - q_k(u^-) \qquad \text{for all } k \in \mathbb{R}. \tag{4.30}$$

Thus, to prove the proposition, we need to show that (4.30) is verified if and only if

$$\lambda(u^+ - u^-) \;=\; f(u^+) - f(u^-)$$

and (4.29) holds for all $\alpha \in (0,1)$.

2. From the definition of (η_k, q_k), we can be rewriten (4.30) as

$$\begin{aligned} &\lambda \left[|u^+ - k| - |u^- - k|\right] \\ &\qquad \geq \; \left[(f(u^+) - f(k)) \cdot \text{sign}(u^+ - k) - (f(u^-) - f(k)) \cdot \text{sign}(u^- - k)\right]. \end{aligned} \tag{4.31}$$

Observe that

- If $k \geq \max\{u^+, u^-\}$ then (4.31) is equivalent to

$$\lambda \cdot (u^- - u^+) \;\geq\; f(u^-) - f(u^+).$$

- If $k \leq \min\{u^+, u^-\}$ then (4.31) is equivalent to

$$\lambda \cdot (u^+ - u^-) \;\geq\; f(u^+) - f(u^-).$$

Thus, (4.31) is verified for all $k \in]-\infty, \min\{u^+, u^-\}] \bigcup [\max\{u^+, u^-\}, +\infty[$ if any only if U satisfies the Rankine- Hugoniot condition.

3. To complete the proof, we need to show that (4.31) is verified for all $k \in]\min\{u^+, u^-\}, \max\{u^+, u^-\}[$ if any only if (4.29) is satisfied. Without loss of generality, we will assume that $u^- < u^+$. For every $k \in]-u^-, u^+[$,

$$\lambda \cdot (u^+ + u^- - 2k) \;\geq\; f(u^+) + f(u^-) - 2f(k)$$

By Rankine-Hugoniot condition, the above inequality is equivalent to

$$\big[f(u^+) - f(u^-)\big] \cdot (u^+ + u^- - 2k) \;\geq\; \big[f(u^+) + f(u^-) - 2f(k)\big] \cdot (u^+ - u^-)$$

Writing $k = \alpha u^+ + (1-\alpha)u^-$, we have

$$(1-2\alpha)\big[f(u^+)-f(u^-)\big][u^+-u^-] \;\geq\; \big[f(u^+)+f(u^-)-2f(\alpha u^+ +(1-\alpha)u^-)\big]\cdot(u^+-u^-)$$

and this implies that

$$f(\alpha u^+ + (1-\alpha)u^-) \;\geq\; \alpha f(u^+) + (1-\alpha)f(u^-).$$

The proof is complete. □

Rely on proposition 4.12.1 and the above proposition, one can show that

Theorem 4.13 *Let $u : [0,T[\times\mathbb{R} \to \mathbb{R}$ be piecewise Lipschitz. Then, the followings are equivalent:*

(i) u is entropy admissible weak solution.

(ii) u satisfies the quasilinear equation (4.19), and for every jump curve $\gamma_j :]a_j, b_j[\to \mathbb{R}$ the (RH) condition holds, i.e.,

$$\dot{\gamma}_j(t) \cdot \big[u_j^+(t) - u_j^-(t)\big] = f(u_j^+(t)) - f(u_j^-(t)),$$

together with the stability condition

$$\frac{f(u_j^\alpha(t)) - f(u_j^-(t))}{u_j^\alpha(t) - u_j^-(t)} \;\geq\; \frac{f(u_j^\alpha(t)) - f(u_j^+(t))}{u_j^\alpha(t) - u_j^+(t)} \qquad \text{for all } t \in]a_j, b_j[$$

with $u_j^\alpha = \alpha u_j^+ + (1-\alpha)u_j^-$.

Let us remark that if we take α go to $0+$ and $1-$ in (4.28) then we obtain the following condition

$$f'(u^-) \;\geq\; \frac{f(u^+) - f(u^-)}{u^+ - u^-} \;\geq\; f'(u^-) \tag{4.32}$$

which can be seen as another type of admissible condition :

Definition 4.14 *We say that a weak solution of (4.3) is admissible in the sense of Lax if at every point $(\bar{t}, \bar{x})$ of approximate jump discontinuity with the left and right states u^-, u^+, and speed λ, the Lax condition holds*

$$f'(u) \;\geq\; \lambda = \frac{f(u^+) - f(u^-)}{u^+ - u^-} \;\geq\; f'(u^+). \tag{4.33}$$

In the case of convex flux $f''(u) \geq 0$, the stability condition (4.29) and Lax condition (4.33) are equivalent. Moreover, if $f''(u) > 0$ then the Lax condition is equivalent to the condition:

$$u^- \ > \ u^+.$$

In the case of general flux, the Lax condition does not implies the stability condition.

Theorem 4.15 *Let $f : \mathbb{R} \to \mathbb{R}$ be locally Lipschitz continuous. Then there exists a continuous semigroup $S : [0, +\infty[\times \mathbf{L}^1 \to \mathbf{L}^1$ such that for each $\bar{u} \in \mathbf{L}^1 \bigcap \mathbf{L}^\infty$, the trajectory $t \mapsto S_t(\bar{u})$ yields a unique bounded, entropy-admissible weak solution of (4.3) with $u(0, \cdot) = \bar{u}$. Moreover, the following properties hold:*

(i) $S_0(\bar{u}) \ = \ \bar{u}, \qquad S_s \circ S_t(\bar{u}) = S_{s+t}(\bar{u})$;

(ii) $\|S_t(\bar{u}) - S_t(\bar{v})\|_{\mathbf{L}^1} \ \leq \ \|\bar{u} - \bar{v}\|_{\mathbf{L}^1}$;

(iii) If $\bar{u}(x) \leq \bar{v}$ for all $x \in \mathbb{R}$ then

$$S_t(\bar{u})(x) \ \leq \ S_t(\bar{v})(x) \qquad \textit{for all } (t, x) \in [0, \infty[\times \mathbb{R}.$$