International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

# Static and Dynamic Issues and Economic Theory. Part I. Models Based on Utility Functions

Aubin, J.-P.

IIASA Working Paper

WP-92-063

August 1992

# Working Paper

## Static and Dynamic Issues in Economic Theory Part I. Models Based on Utility Functions

*Jean-Pierre Aubin*

WP-92-63
August 1992

IIASA
International Institute for Applied Systems Analysis ◻ A-2361 Laxenburg ◻ Austria
Telephone: +43 2236 715210 ◻ Telex: 079 137 iiasa a ◻ Telefax: +43 2236 71313

# Static and Dynamic Issues in Economic Theory
# Part I. Models Based on Utility Functions

*Jean-Pierre Aubin*

WP-92-63
August 1992



IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Jean-Pierre Aubin

# Static and Dynamic Issues in Economic Theory

## Part I. Models Based on Utility Functions

# Contents


0 **The Issues** 5
- 0.1 The Main Issue: Decentralization . . . . . 5
- 0.2 Adam Smith's Invisible Hand . . . . . 6
- 0.3 Walras' Choice . . . . . 7
- 0.4 The Visible Consumers . . . . . 8
- 0.5 Selection Mechanisms . . . . . 9
- 0.6 The Inertia Principle . . . . . 11
- 0.7 Heavy Evolutions . . . . . 12
- 0.8 A Simple Economic Example. . . . . . 13
- 0.9 Outline . . . . . 20

1 **Allocations of Scarce Resources** 24
- 1.1 The Commodity Space . . . . . 24
- 1.2 The Value Space . . . . . 25
- 1.3 The Price Space . . . . . 26
- 1.4 The set of Resources . . . . . 28
- 1.5 Introducing the Consumers . . . . . 29
- 1.6 The Set of Allocations . . . . . 30

2 **The Optimal Allocation Theorem** 30
- 2.1 Demand and Change Maps . . . . . 31
- 2.2 Supply and Pricing Maps . . . . . 32
- 2.3 Optimal Allocations . . . . . 33
- 2.4 The Walras Tâtonnement . . . . . 34
- 2.5 The Nontâtonnement Process . . . . . 34

3 **Convex Functions** 37
- 3.1 Extended Functions and their Epigraphs . . . . . 37
- 3.2 Subdifferential of Convex Functions . . . . . 39
- 3.3 Support Functions and Conjugate Functions . . . . . 42
- 3.4 Fenchel's Theorem . . . . . 47
- 3.5 Properties of Conjugate Functions . . . . . 51
- 3.6 Subdifferential Calculus . . . . . 53
- 3.7 Moreau-Yosida Approximations . . . . . 56

4 **Subgradient Differential Inclusion** 64

5 **Subgradient Algorithms** 70

**6 Duality Theory** **75**
6.1 The Duality Theorem . . . . . . . . . . . . . . . . . . . . . . . . 75
6.2 Minimization Problems with Constraints . . . . . . . . . . . . 79
6.3 Optimal Allocations . . . . . . . . . . . . . . . . . . . . . . . . . 81

**7 Calculus of Tangent Cones** **87**

*"Il pourrait au contraire sembler à quelques uns qu'étant donnée cette complexité des phénomènes économiques, l'économie mathématique est justement beaucoup trop simple. Ceux-là n'auraient qu'à se rappeler que les premières propositions de la géométrie sont, elles aussi, très simples, sans que cette simplicité interdise en rien les complications ultérieures.*
*Enfin, si quelques uns, tout en reconnaissant que la méthode mathématique n'est pas superflue pour constituer la science de l'économie publique, et tout en reconnaissant sans doute que cette science n'atteint pas toute la complexité du réel vivant, mais que ces résultats cependant sont toujours sous-jacents, pour ainsi dire, à ce réel, se défiaient de cette science elle-même, et de son importance, et craignaient qu'on ne s'y enfermât un peu complaisamment, ceux-là n'auraient pas été frappés des admirables paroles, et décisives, où le savant fait lui-même la part de la science :* "la réforme sociale doit procéder à la fois du sentiment socialiste et de la science économique".

Charles Péguy

*un économiste socialiste : Léon Walras*
La Revue Socialiste, 1897, 146, 174-186

*"Finally, let us note a point at which the theory of social phenomena will presumably take a very definite turn away from the existing patterns of mathematical physics. This is, of course, only a surmise on a subject where much uncertainty and obscurity prevail... A dynamic theory, when one is found, will probably describe the changes in terms of simpler concepts".*

John von Neumann and Oscar Morgenstern

*Theory of Games and Economic Behavior*, (1944).

# Foreword

We shall devote these papers to the simplest economic problem we can think of:

**how to allocate scarce resources among consumers**

by complying to the basic economic constraint

**It is impossible to consume more physical goods than available**

In other words, let us introduce the set of **allocations** of these scarce resources among the consumers. This means that each consumer receives a commodity the sum of which is viable in the sense that the total consumptions is an available resource.

This problem looks at first glance somewhat silly and simple minded, since it amounts to pick up an element in this allocation set (i.e., to choose an allocation) in the case of static models, or to evolve in this set, regarded as a **viability set**, in the case of dynamical systems. However, it elucidates the basic difficulties characteristic of economic theory.

Static models assign one or several elements in the allocation set. But it may be time to answer the wish J. von Neumann and O. Morgenstern expressed in 1944 at the end of the first chapter of their monograph "Theory of Games and Economic Behavior":

*"Our theory is thoroughly static. A dynamic theory would unquestionably be more complete and therefore, preferable. But there is ample evidence from other branches of science that it is futile to try to build one as long as the static side is not thoroughly understood".*

We study here some mechanisms which govern the evolution of allocations of scarce resources[1].

In these dynamical models, the laws which govern the evolution of allocations are most often represented by differential equations (or differential inclusions) with or without memory.

Static models are particular cases of (time-independent) dynamical models yielding "constant evolutions", which are also called "equilibria". By the way, the concept of equilibrium often covers two different meanings in economics. The first one, the meaning we use in these lectures, is derived from

[1] By the way, in dynamical models, we can assume that the subset of allocations evolves with time, depends upon the history of the evolution.

mechanics, where an equilibrium is a constant function, or a "rest point". The second meaning is covered here by what we call the viability constraints, such as the total consumption must be less than or equal to the total supply, etc.

# 0 The Issues

## 0.1 The Main Issue: Decentralization

We begin by distinguishing between **centralized and decentralized models.** In the first category of models, consumers delegate their decision power to another "agent" who, knowing the behaviors of the consumers and the set of scarce resources, solves the problem at the global level.

For instance, consumers must agree to describe their behavior by a **collective utility function**

$$x := (x_1, \ldots, x_n) \mapsto U(x) = U(x_1, \ldots, x_n) \in \mathbf{R}$$

Then, this agent (planning bureau, big computers or big brothers, ...) knowing $U$ and the subset $M$, decides to maximize $U$ over the allocation set $K$. **The problem is then transferred to the question of choosing the collective utility function** $U$.

Or, in the dynamical version, they agree to represent their behavior by, say, a system of differential equations

$$x_i'(t) \;=\; f_i(x_1(t), \ldots, x_n(t)) \quad (i = 1, \ldots, n)$$

where the variations of the consumption of each agent depend upon the knowledge of both the whole set of scarce resources and the choices of every other agents.

In a **decentralized mechanism**, the information on the problem is split and mediated by, say, a "message" which summarizes part of the information. In our case, we use for message the "price" $p$. Knowing the price $p$, consumers are supposed to know how to choose their consumption bundle, without

- knowing the behavior of their fellow consumers
- knowing the set of scarce resources

Then the problem is to find what is the message which carries the relevant information.

Actually, we have to ask whether it is possible to find such a relevant message and then, how to find it. If it is possible to answer the first type of question, it is much more difficult to investigate the second, leaving such problems to mythical players such as the "market", Adam Smith's "invisible hand", etc. We shall bethink that these players are not really operating on the price system, which we shall propose to regard as a **regulatory control** (a "regulee") to help the consumers to respect the scarcity constraints by delivering them proper informations on the behavior of all consumers and the set of available resources.

## 0.2 Adam Smith's Invisible Hand

Indeed, there is no doubt that Adam Smith is at the origin of what we now call **decentralization**, i.e., the ability for a complex system moved by different actions in pursuit of different objectives to achieve an allocation of scarce resources. The difficulty to grasp such a disordered way of regulation of economic processes, contrary to apparently more logical (or simple minded?) attractive organizational processes based on several varieties of planning procedures[2] led him to express it in a poetic manner. Let us quote the celebrated citation of the WEALTH OF NATIONS published in 1796, two centuries ago:

*"Every individual endeavours to employ his capital so that its produce may be of greatest value. He generally neither intends to promote the public interest, nor knows how much he is promoting it. He intends only his own security, only his own gain. And he is in this led by an* **invisible hand** *to promote an end which was no part of his intention. By pursuing his own interest, he frequently thus promotes that of society more effectually that when he really intends to promote it"*

However, Adam Smith did not provide a careful statement of what the **invisible hand** "manipulates" nor, a fortiori, for its existence.

We had to wait a century more for Léon Walras, a former engineer, to recognize that this invisible hand "operates" on economic agents through prices, gaining enough information on the desires of the agents and the available commodities for guaranteeing their consistency, or the viability of the allocation system.

He presented in 1874 the general equilibrium concept in ÉLÉMENTS D'ÉCONOMIE POLITIQUE PURE as a solution to a system of nonlinear equations. At that time, when only linear systems were understood, the fact that the number of equations was equal to the number of unknowns led him and his immediate followers to make the optimistic assumption that a solution should necessary exist[3].

[2] in favor among military organizational schemes.

[3] But it took another century, until 1954, for Kenneth Arrow and Gérard Debreu to find a mathematical solution to this problem. This solution, however, could not have been obtained without the fundamental Brouwer Fixed Point Theorem in 1910, which in turn

## 0.3 Walras' Choice

In modern terms, the behavior of each consumer is described by a demand function $d_i$ allowing the consumer to choose a commodity $x_i = d_i(p)$ knowing only the price $p$. The problem is then to find a price $\bar{p}$ (the Walrasian equilibrium price) such that $(d_1(\bar{p}), \ldots, d_n(\bar{p}))$ forms an allocation. This is **a decentralized model** because consumers do not need to know neither the choices of other consumers nor the set $M$ of available commodities. The basic Arrow-Debreu Theorem states in this case that such an equilibrium exists whenever a budgetary rule known as **Walras law** — **it is forbidden to spend more monetary units than earned** — is obeyed by consumer's demand functions.

Furthermore, such a price $\bar{p}$ is an equilibrium of an underlying dynamical process, called the **Walrasian tâtonnement**[4]: in its continuous version, it is defined by the **differential inclusion**

$$p'(t) \in E(p(t))$$

where $E$ is the **excess demand map** given by

$$E(p) := \sum_{i=1}^{n} d_i(p) - M$$

Hence, according to this law of supply and demand, the price increases whenever the excess demand is positive and decreases in the opposite case.

We observe that if $p(t)$ is a price supplied by the Walras tâtonnement process and if it is not an equilibrium, **it cannot be implemented** because the associated demand is not necessarily available.

Hence, this model **forbids consumers to transact as long as the prices are not equilibria**. It is as if there was a super auctioneer calling prices and receiving offers from consumers. If the offers do not match, he calls another price according to the above dynamical process, but **does not allow transactions to take place** as long as the offers are not consistent, and this happens only at equilibria!.

**Tâtonnement is therefore not viable.**

---

required much modification to tailor it to this specific problem — by proving theorems whose assumptions could bear the same degree of economic interpretation as the conclusion.

[4]Tâtonnement means "tentative process", "trial and error" — literally, cumbersomely walking in obscurity by touch (tâter).

And it may be too much to ask the entity which regulates the price (the market, the invisible hand, the Gosplan, ...) to behave as a real decision-maker.

The concept of economic equilibrium and tâtonnement that we owe to Léon Walras is not his only claim to our gratitude: Léon Walras was one of the first persons (after Condorcet, Boda, Cournot, Canard, and few other) to suggest that mathematics could be useful in economic theory. Originality is often more a question of finding a new way of looking at the world than of making discoveries that attract the attention of one's peers. Walras introduced mathematical rigor into a domain which had never before been subjected to detailed analysis. He did it with disregard for – even in opposition to – the prevailing economic thinking of the times, despite tremendous difficulties, alone and without help, without the encouragement and moral support of his colleagues. He did it because, deep within him, he realized the far-reaching consequences of his bold vision.

However, the legitimate admiration that he deserves should not imply a dogmatic respect of his contribution by his followers: the equilibrium concept was a simplifying step in the attempt to grasp some essential economic feature in an otherwise complex maze of concepts. This concept had its use, as a first approximation, despite the fact that it rarely happened in economic history. So, its **dépassement**, as well as the observation that the Walrasian tâtonnement is not viable and should be replaced by a viable dynamical system, should not be regarded by the faithfuls as a crime of **lèse majesté**. On the other hand, smart — but superficial — minds should not use these shortcomings to claim that any decentralized mechanism using prices is merely a fantasy dreamed by mathematicians from their ivory towers — an empty box, as it has been written — and even, to reject the relevance of mathematical metaphors in economics. This is a typical instance of impatience and the totalitarian desire for monist explanations.

## 0.4 The Visible Consumers

It may be wise indeed to let the **real decision-makers**, the consumers in our case, to govern the evolution of their consumption through differential equations

$$x_i'(t) = c_i(x_i(t), p(t))$$

parametrized (or controlled) by the price $p(t)$, so that consumers change their consumptions knowing only the price $p(t)$ at each time $t$, **without taking into account neither the behavior of the other consumers nor the knowledge of** the set $M$ of scarce resources. Hence it shares with the Walras static model its decentralization property.

The problem is then to find a price function $p(t)$ such that the associated solutions $x_i(t)$ of the above differential equations do form an allocation at each time $t$. We prove that this viability property holds true under a dynamical version of the Walras law and even prove the existence of an equilibrium of this dynamical model.

Actually, we would like to know more than a time-dependent price function (which can be regarded as an open loop control). We wish to obtain "feedback prices", or, more generally, set-valued "regulation maps" associating with each allocation $x \in K$ one price, or more generally, the set $\Pi(x)$ of relevant messages, so that the evolution law of the relevant message is

$$\forall\, t,\;\; p(t) \;\in\; \Pi(x(t)) \;=\; \Pi(x_1(t),\ldots,x_n(t))$$

## 0.5 Selection Mechanisms

The set of viable prices (regarded as relevant messages) may contain more than one element. The question arises to select one of these prices, or, to shrink the set of viable prices by an adequate mechanism. This can be done by optimization techniques, or, more generally, by game theoretical methods.

In the dynamical case, this question splits in two: we have to distinguish between "intertemporal optimization" problems and "myopic or instantaneous optimization" problems.

In intertemporal optimization, we maximize intertemporal utility functions of the form

$$U(x(\cdot),p(\cdot)) \;:=\; \int_0^T u(t,x(t),p(t))dt + v(x(T),p(T))$$

under the constraint $(x(\cdot),p(\cdot)) \in \mathrm{Graph}(\Pi)$.

These are questions with which Calculus of Variations and Optimal Control Theory deal with.

But Optimal Control Theory does require the Market or Adam Smith's invisible hand to "guide" the system by optimizing such an intertemporal optimality criterion, the choice of which is open to question even in static models, even when multicriteria or several decision makers are involved in the model.

Furthermore, the choice (even conditional) of the controls is made once and for all at some initial time, so that they cannot be modified at each instant so as to take into account possible modifications of the environment of the system, forbidding therefore adaptation to scarcity constraints.

Finally, intertemporal optimization theory does require the knowledge of the future (even of a stochastic nature.) This requires the possibility of experimentation or the belief that the phenomenon under study is periodic. Experimentation, by assuming that the evolution of the state of the system starting from a given initial state for a same period of time will be the same whatever the initial time, allows one to translate the time interval back and forth, and, thus, to "know" the future evolution of the system.

But in economics, as well as in biological evolution, experimentation is not possible[5]. Furthermore, the dynamics of the system disappear and cannot be recreated. Most economic systems do involve myopic behavior; while they cannot take into account the future, they are certainly constrained by the past. Hence, forecasting or prediction of the future are not the issues which we shall address here. *La prévision est un rêve duquel l'événement nous tire,* wrote Paul Valéry.

We shall instead attempt to understand how the evolution of economic systems is governed.

Therefore, instead of using intertemporal optimization[6] that involves the future, we shall propose to use Viability Theory for providing selection procedures of viable evolutions obeying, at each instant, scarcity or more generally, viability constraints which depend upon the present or the past. (This does not exclude anticipations, which are extrapolations of past evolutions, constraining in the last analysis the evolution of the system to be a function of its history.)

However, the use of optimal control theory led to the popular theory of rational expectations. It shares with general equilibrium theory the feature of growing up from available mathematical theories and being transferred to economics. The pretty large consensus around these concepts make them "real" according to the following definition of the degree of reality for a social group at a given time: Reality is the consensus interpretations of the group member's perceptions of their physical, biological, social and cultural environments.

[5]The twentieth century Soviet type (or military type) economic experimentation showed experimentally the limits of centralized operation of complex systems.

[6]which can be traced back to Sumerian mythology which is at the origin of Genesis: one Decision-Maker, deciding what is good and bad and choosing the best (fortunately, on an intertemporal basis, thus wisely postponing to eternity the verification of optimality), knowing the future, and having taken the optimal decisions, well, during one week...

But it should be time for this consensus to evolve by looking for economic facts to motivate new mathematical theories and not the other way around.

In myopic optimization, we use the feedback relation and we select for each allocation $x \in K$ a price $p \in \Pi(x)$ by a static optimization technique (or any other kind of technique). For instance, we can choose the element $\pi^0(x) \in \Pi(x)$ of minimal norm. Despite the lack of continuity of such a selection, we still can prove that the system of differential equations

$$x_i'(t) = c_i(x_i(t), \pi^0(x(t)))$$

has viable solutions, which are called "slow allocations".

However, this type of selection may not enjoy economic meaning. We propose another one which may be closer in spirits to economic mechanisms.

## 0.6 The Inertia Principle

Actually, if the behavior of the consumers is well defined, what about either the market or the planning bureau, the task of which is to find the prices $p(t)$ in $\Pi(x(t))$? They do not behave as actual decision makers, knowing what is good or not (this is the case of even a planning bureau as soon as it involves more than three bureaucrats!). Hence, their role is only a **regulatory** one. If they are not able to optimize, we may assume that they only are able to correct the prices when the viability of the economic system is at stake, i.e., when the total consumption is no longer available.

Hence, we assume that the market (Adam Smith's "invisible hand") or the planning bureau are able to "pilot" or "act" on the system by choosing such controls according to the **inertia principle**:

> **Keep the price constant as long as the evolution provides allocations of available resources, and change them only when the viability is at stakes.**

Indeed, as long as the state of the system lies in the interior of the allocation set (the set of states satisfying scarcity constraints), any price will work. Therefore, the system can maintain the price inherited from the past. This happens if the system obeys the inertia principle. Since the allocations may evolve while the price remains constant, the total consumption may reach the boundary of the set of scarce resources with an "outward" velocity. This event corresponds to a period of *crisis*: To survive, the system must find another price such that the new associated velocity forces the solution back inside the allocation set. Alternatively, if the scarcity constraints can evolve, another way to resolve the crisis is to relax the constraints (by technological progress, for instance) so that the state of the system lies in the interior of the new allocation set. When this is not possible, **strategies for structural change fail:** by design, this means that the solution leaves the allocation set and "dies".

This management by crisis or bankruptcy has been observed in economic history, so that we suggest to take these phenomena into account in the framework of this Inertia Principle[7].

## 0.7 Heavy Evolutions

This inertia principle is not strong enough to select an evolution of a relevant price, since we have to provide rules for choosing prices when viability is at stakes.

The simplest one (and most often, the most reasonable one) is to assume that at each instant, the prices are changed as slowly as possible.

We called evolutions obeying this principle "heavy[8] evolutions", in the sense of heavy trends. Hence heavy evolution is obtained by requiring at each instant the (norm of the) velocity of the price to be as small as possible.

Therefore, for implementing this inertia principle, we have to provide conditions under which relevant prices $p(\cdot)$ are differentiable (almost everywhere), to built the differential inclusion which governs the evolution of differentiable relevant prices and then, select a differential equation in this differential inclusion (called a "dynamical closed loop") which will obey the inertia principle.

In summary, given the decentralized behavior of the consumers described by the differential equations $x_i' = c_i(x_i, p)$ and the set of scarce resources, we can built the dynamics $\varpi$ governing the behavior of the market, so that the

[7]This Inertia Principle provides an explanation of the concept punctuated equilibrium introduced in 1972 by Elredge and Gould in paleontology. Excavations at Kenya's Lake Turkana have provided clear evidence of evolution from one species to another. The rock strata there contain a series of fossils that show every small step of an evolution journey that seems to have proceeded in fits and starts. Examination of more than 3,000 fossils by P. Williamson showed how 13 species evolved. The record indicated that the animals stayed much the same for immensely long stretches of time. But twice, about two million years ago and then, 700,000 years ago, the pool of life seemed to explode — set off, apparently, by a drop in the lake's water level. Intermediate forms appeared very quickly, new species evolving in 5,000 to 50,000 years, after millions of years of constancy, leading paleontologists to challenge the accepted idea of continuous evolution.

[8]This is justified by the fact that the velocity of the price is related to the acceleration of the consumptions, and thus, the iverse of the mass.

evolution of the economic system is described by the system of differential equations

$$\begin{cases} i) \quad x_i'(t) = c_i(x_i(t), p(t)) \quad (i = 1, \ldots, n) \\ ii) \quad p'(t) = \varpi'(x(t), p(t)) \end{cases}$$

Contrary to other dynamical models, this law governing the evolution of prices is not a modeling assumption, but a **consequence** of the modeling data of this elementary model.

In summary, we assume implicitly that the "Market" follows an "opportunistic"m "conservative" and "lazy" behavior of the system: a behavior which enables the system to allocate scarce resources among consumers as long as any price makes possible its regulation and to keep this price as long as it is possible.

We shall attempt to explain the evolution of allocations and prices and to reveal the concealed feedbacks which allow the system to be regulated by prices.

We illustrate the concept of heavy solution by the simplest dynamical economic model (one commodity, one consumer.)

## 0.8 A Simple Economic Example.

Let $K := [0, b]$ the subset of a scarce commodity $x$. Assume that the consumption rate of our greedy consumer is equal to $a > 0$, so that, without any further restriction, her exponential consumption will leave the allocation subset $[0, b]$. Hence her consumption is slowed down by a price (which is regarded as a control). In summary, the evolution of its consumption is governed by the system

$$\text{for almost all } t \geq 0, \ \ x'(t) = ax(t) - p(t), \ \ \text{where } \ p(t) \geq 0$$

subjected to the constraints

$$\forall\, t \geq 0, \ \ x(t) \in [0, b]$$

(See figure 1) We see at once that the viable equilibria of the system range over the **equilibrium line** $p = ax$.

The regulation map is given by the formula

$$\Pi_K(0) = \{0\}, \ \Pi_K(x) = \mathbf{R}_+ \ \text{ when } \ x \in ]0, b[ \ \ \& \ \ \Pi_K(b) = [ab, +\infty[$$

Indeed, if $x = 0$, the velocity should be non negative, and the only price we can achieve it is $p := 0$. If $0 < x < b$, any velocity allows to keep the state between 0 and $b$ for a short period of time, so that any price can be used. If $x = b$, then the velocities $x' = ab - p$ should be non positive to keep the state in the interval $[0, b]$.

Viability is thus guaranteed each time that the price $p(t)$ is chosen in $\Pi(x(t))$, i.e., $p = 0$ when $x = 0$ (and thus, the system cannot leave the equilibrium because negative prices are not allowed "to start" the system) and $p \geq ab$ when $x = b$, so that the price is large enough to stop or decrease consumption.

Assume that the system obeys the inertia principle: **it keeps the price constant as long as it works.** Take for instance $x_0 > 0$ and $p_0 \in [0, ax_0[$. Then the consumption increases[9] and when it reaches the boundary $b$ of the interval, the system has to switch very quickly to a velocity large enough to slow down the consumption for the solution to remain in the interval $[0, b]$.

But there is a bound to the growth of prices (and inflation rates), so that we should set a bound on price velocities: $|p'(t)| \leq c$. **We shall associate with such a bound a "last warning" threshold to modify the price**: there is a level of consumption after which it will be impossible to slow down the consumption with a velocity smaller than or equal to $c$ to forbid it to increase beyond the boundary $b$.

We shall find this bound and introduce **heavy solutions** which will be studied in greater generality later for building this regulation law. They are the one whose controls evolve with the "smallest velocity".

We thus consider the solutions to the system

$$\begin{cases} i) & \text{for almost all } t \geq 0, \;\; x'(t) = ax(t) - p(t) \\ ii) & \text{and } \;\; -c \leq p'(t) \leq c \end{cases} \tag{0.1}$$

which are viable in $[0, b] \times \mathbf{R}_+$.

We introduce the functions $\rho_c^\sharp$ and $\rho_c^\flat$ defined on $[0, \infty[$ by

$$\begin{cases} i) & \rho_c^\flat(p) := \frac{c}{a^2}(e^{-ap/c} - 1 + \frac{a}{c}p) \approx \frac{p^2}{2c} \\ ii) & \rho_c^\sharp(p) := -ce^{a(p-ab)/c}/a^2 + p/a + c/a^2 \end{cases}$$

[9]it is equal to $\frac{e^{at}(ax_0 - p_0) + p_0}{a}$.

and the functions $r_c^\sharp$ and $r_c^\flat$ defined on $[0,b]$ by

$$\begin{cases} i) & r_c^\flat(x) = p \text{ if and only if } x = \rho_c^\flat(p) \\ ii) & r_c^\sharp(x) = 0 \text{ if } x \in [0, \rho_c^\sharp(0)] \ \ (\rho_c^\sharp(0) = \frac{c}{a^2}(1 - e^{-a^2 b/c})) \\ iii) & r_c^\sharp(x) = p \text{ if and only if } x = \rho_c^\sharp(p) \text{ when } x \in [\rho_c^\sharp(0), b] \end{cases}$$

We shall show that these maps $r_c^\sharp$ and $r_c^\flat$ are solutions to the nonlinear "first-order partial differential inclusion"

$$0 \in r'(x)(ax - r(x)) + [-c, +c] \tag{0.2}$$

and that they can be regarded as **planning procedures.**

We introduce now the set-valued map $R^c$ defined by[10]

$$\forall\, x \in [0,b],\ \ R^c(x) = [r_c^\sharp(x), r_c^\flat(x)] \tag{0.4}$$

**There exist solutions to (0.1) if and only if the initial state satisfies $p_0 \in R^c(x_0)$. In this case, prices and commodities are related by the regulation law:**

$$\forall\, t \geq 0,\ \ p(t) \in R^c(x(t))$$

Indeed, set $p^\sharp(t) := p_0 + ct$ and $p^\flat(t) := p_0 - ct$ and denote by $x^\sharp(\cdot)$ and $x^\flat(\cdot)$ the solutions starting at $x_0$ to differential equations $x' = ax - p^\sharp(t)$ and $x' = ax - p^t(\cdot)$ respectively. Then any solution $(x(\cdot), p(\cdot))$ to the system (0.1) satisfies $p^\flat(\cdot) \leq p(\cdot) \leq p^\sharp(\cdot)$ and thus, $x^\sharp(\cdot) \leq x(\cdot) \leq x^\flat(\cdot)$ because

$$x(t) = e^{at}x_0 - \int_0^t e^{a(t-s)}p(s)ds$$

We also observe that the equations of the curves $t \mapsto (x^\sharp(t), p^\sharp(t))$ and $t \mapsto (x^\flat(t), p^\flat(t))$ passing through $(x_0, u_0)$ are solutions to the differential equations

$$d\rho_c^\sharp = \frac{1}{c}(a\rho_c^\sharp - p)dp \ \ \& \ \ d\rho_c^\flat = -\frac{1}{c}(a\rho_c^\flat - p)dp$$

[10] By using tools of set-valued analysis, and in particular, the concept of contingent derivative $DR(x,u)$ of a set-valued map $R$, we shall see that $R^c$ is a set-valued solution to the first-order partial differential inclusion

$$\forall\, (x,p) \in \mathrm{Graph}(R),\ \ 0 \in DR(x,p)(ax - p) + [-c, +c] \tag{0.3}$$

and actually, the largest one with closed graph.

Figure 1: Evolution of a Heavy Solution

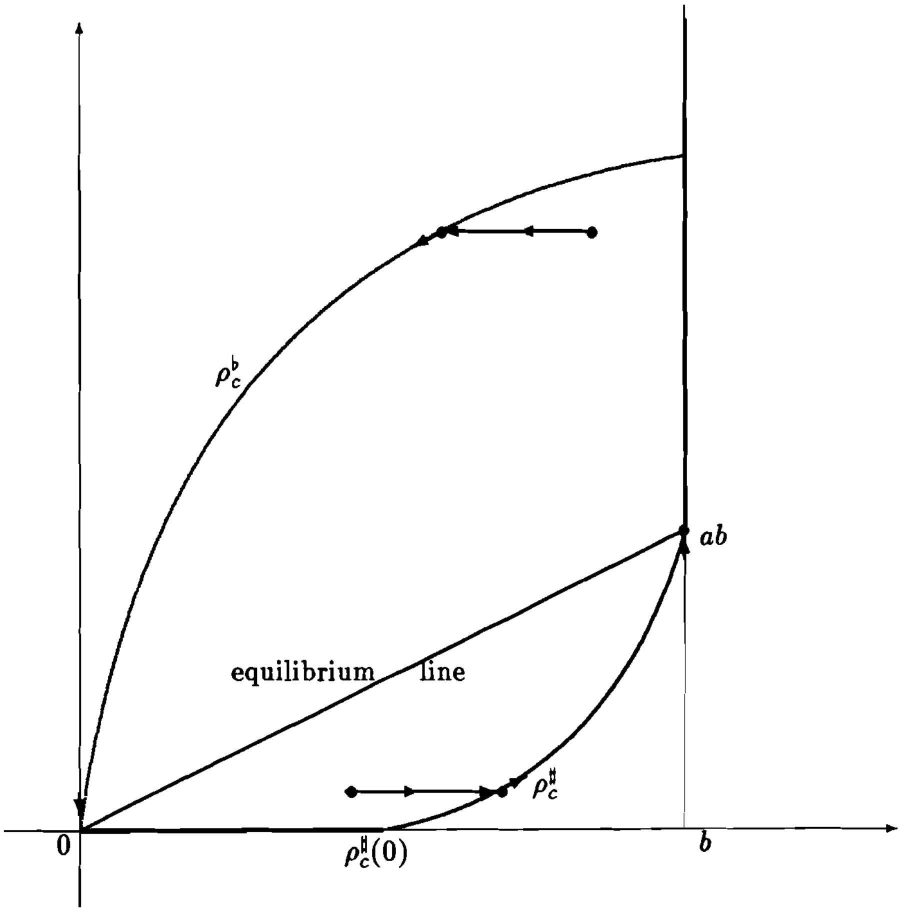
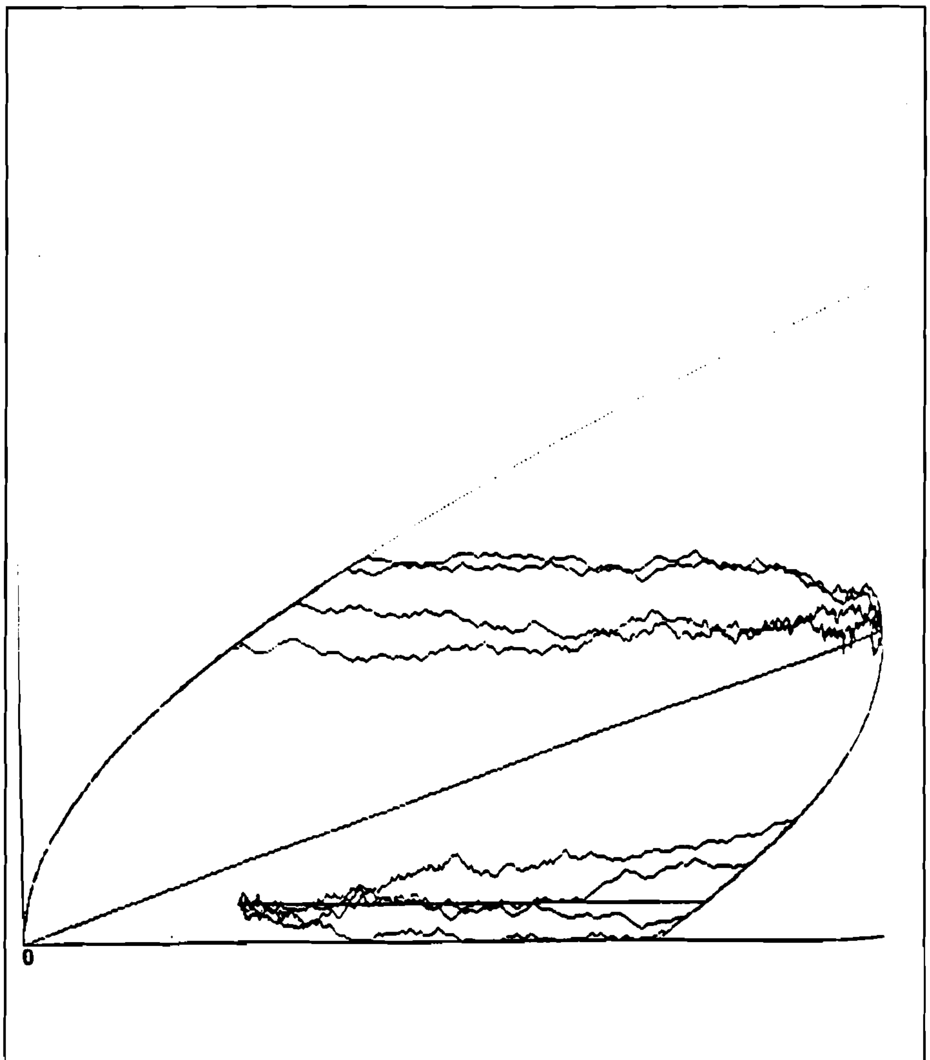


Figure 2: Other Solutions and Semipermeability of the Boundary




Examples of other solutions (0.1) where the velocities $u'$ of the controls are randomly generated. This computer simulation (due to Morin & Vandanjon) illustrates the Quincampoix Theorem (which is not prove in this lecture notes) on the semipermeability property of the part of the boundary of the "viability kernel" contained in the interior of $[0, b] \times \mathbf{R}_+$: The solutions which reach this boundary cannot come back to it, and have to remain on its boundary.

the solutions of which are

$$\begin{cases} i) \quad \rho_c^\sharp(p) = e^{a(p-p_0)/c}(x_0 - p_0/a - c/a^2) + p/a + c/a^2 \\ ii) \quad \rho_c^\flat(p) = e^{a(p_0-p)/c}(x_0 - p_0/a + c/a^2) + p/a - c/a^2 \end{cases}$$

Let $\rho_c^\flat$ be the solution passing through $(0,0)$, which is equal to $\rho_c^\flat(p) = \frac{c}{a^2}(e^{-ap/c} - 1 + \frac{a}{c}p)$ and $\rho_c^\sharp(p) = -ce^{a(p-ab)/c}/a^2 + p/a + c/a^2$ be the solution passing through the pair $(ab, b)$.

— IF $p_0 > r_c^\flat(x_0)$, THEN ANY SOLUTION $(x(\cdot), p(\cdot))$ STARTING FROM $(x_0, p_0)$ LEAVES Graph$(R^c)$: it satisfies

$$x(t) \leq x^\flat(t) = \rho_c^\flat(p^\flat(t)) \leq \rho_c^\flat(p(t))$$

because $\rho_c^\flat(\cdot)$ is nondecreasing. Hence, when $x(t_1) = 0$, we deduce that $p(t_1) > 0$, so that such solution is not viable.

— IF $0 \leq p_0 < r_c^\sharp(x_0)$, ANY SOLUTION $(x(\cdot), y(\cdot))$ STARTING FROM $(x_0, p_0)$ LEAVES Graph$(R^c)$: it satisfies inequalities

$$x(t) \geq x^\sharp(t) = \rho_c^\sharp(p^\sharp(t)) \geq \rho_c^\sharp(p(t))$$

Therefore, when $x(t_1) = b$ for some time $t_1$, its velocity $x'(t_1) = ab - p(t_1)$ is positive, so that the solution is not viable.

— It remains to show that starting from any point $(x_0, p_0)$ of the graph of $R^c$, there exist **heavy solutions**.

Naturally, if we start from an equilibrium, both the state and the controls can be kept constant.

We now investigate the cases when the initial control $p_0$ is below or above the equilibrium line.

Consider the case when $x_0 > 0$ and the price $p_0 \in [r_c^\sharp(x_0), ax_0[$. Since we want to choose the price velocity with minimal norm, we take $p'(t) = 0$ as long as the solution $x(\cdot)$ to the differential equation $x' = ax - p_0$ yields a consumption $x(t) < \rho_c^\sharp(p_0)$. When for some time $t_1$, the consumption $x(t_1) = \rho_c^\sharp(p_0)$, so the solution has to be slowed down. Indeed, otherwise $(x(t_1 + \varepsilon), p_0)$ will be below the curve $\rho_c^\sharp$ and we saw that in this case, any solution starting from this situation will eventually cease to be viable. Therefore, prices should increase to slow down the consumption growth. The idea is to take the smallest velocity $p'$ such that the vector $(x'(t_1), p')$ takes the state inside the graph of $R^c$: they are the velocities $p' \geq x'(t_1)/\rho_c^{\sharp'}(p_0)$.

By construction, it is achieved by the velocity of $x^\sharp(\cdot)$, which is the highest one allowed to increase prices. Therefore, by taking

$$x(t) := x^\sharp(t) := e^{a(t-t_1)}(x(t_1) - p_0/a - c/a^2) + c(t - t_1)/a + p_0/a + c/a^2$$

and $p(t) := p_0 + c(t - t_1)$ for $t \in [t_1, t_1 + (ab - p_0)/c]$, we get a solution which ranges over the curve $x^\sharp(t) = \rho_c^\sharp(p^\sharp(t))$. This a heavy solution because, for the same reason as above, the smallest velocity of the price (which is unique along this curve) is chosen. According to the above differential equation, we see that $x(t)$ increases to $b$ where it arrives with velocity 0 and the price increases linearly until it arrives at the equilibrium price $ab$. Since $(b, ab)$ is an equilibrium, the heavy solution stays there: we take $x(t) \equiv b$ and $p(t) \equiv ab$ when $t \geq t_1 + p_0/c$. So we have built a viable solution starting from $(x_0, p_0)$.

Consider now the case when $p_0 \in [ax_0, r_c^\flat(x_0)]$, where we follow the same construction of the heavy viable solution. We start by taking $p'(t) = 0$. Thus, $p(t) = p_0$, as long as the solution $x(\cdot)$ to the differential equation $x' = ax - p_0$, which decreases, satisfies $x(t) > \rho_c^\flat(p_0)$. Then, when $x(t_1) = \rho_c^\flat(p_0)$ for some $t_1$, we take

$$x(t) = x^\flat(t) := e^a(t - t_1)(x(t_1) - p_0/a + c/a^2) - c(t - t_1)/a + p_0/a - c/a^2$$

and $p(t) := p_0 - c(t - t_1)$ for $t \in [t_1, t_1 + p_0/c]$ in order to prevent the solution from leaving the graph of $R^c$. Finally, for $t \geq t_1 + p_0/c$, we take $x(t) \equiv 0$ and $p(t) \equiv 0$. □

**Remark** — We observe that for any $x \in ]0, b[$,

$$\lim_{c \to 0+} r_c^\flat(x) = \lim_{c \to 0+} r_c^\sharp(x) = ax, \quad \lim_{c \to \infty} r_c^\sharp(x) = 0 \;\&\; \lim_{c \to \infty} r_c^\flat(x) = +\infty$$

□

Quincampoix has proved that **the part of the boundary of the graph of $R^c$ which lies in the interior of the cylinder $[0, b] \times \mathbf{R}_+$ is a barrier**. This means that from any point $(x, p)$, **all viable solutions remain on the part of the boundary** contained in the interior of the cylinder. **They cannot enter the graph of $R^c$.** Once the solution bumps onto such a part of the boundary, its trajectory remains on it, and there is no way, in this example, for the price to evolve with a velocity smaller than $c$ in absolute value.

In a daily language, if one interprets situations where the pair $(x, p)$ lies in the boundary of the graph of $R^c$ as a crisis, there is no possibility to get out of this crisis situation as long the pair $(x, p)$ is in the interior of the cylinder.

This phenomenon is illustrated by computer simulations represented in figure 2. Velocities of the prices are generated randomly. As soon as the solution butts against the boundary of the graph, it continues to evolve on the boundary.

## 0.9 Outline

This first part, which we have tried to keep as self-contained as we could, cover the two main view points brought to this basic problem of allocation of scarce resources when the behavior of consumers is represented by utility functions.

It deals with the problem of optimal allocation of resources in the framework of Convex Analysis and its Duality Theory, of which we summarize the very basic facts[11]. Its main purpose is to show that this optimal allocation problem conceals the two main rival dynamical processes which compete in the economic literature: The Walras tâtonnement model and the Nontâtonnement model. Starting with utility functions which represent the behavior of consumers, one can derive

1. **demand and supply maps**, and then, the concept of excess demand on which the tâtonnement dynamics are built, and the associated equilibria, the Walras equilibrium prices,

2. **change and pricing maps**, and then, the nontâtonnement dynamical economy which can be built, and the associated equilibria, made of allocations which are not changed by the consumers.

We shall supply the proofs of the continuous and discrete versions of the gradient methods we shall provide in this context both the **tâtonnement** model and a **nontâtonnement** algorithm (at least, in the continuous case).

[11] A more comprehensive exposition can be found in [9, Aubin] and [24, Aubin & Ekeland].

# Optimal Allocations

## Introduction

The first type of selection mechanism of an allocation which comes up to the mind is an **optimization** mechanism permitting to select an allocation.

For doing so, we need to introduce a **collective utility function**

$$U : x := (x_1, \ldots, x_n) \in X := Y^n = \mathbf{R}^{ln} \mapsto U(x_1, \ldots, x_n)$$

and to look for an allocation

$$\hat{x} \in K := \left\{ x := (x_1, \ldots, x_n) \in \prod_{i=1}^{n} L_i \;\middle|\; \sum_{i=1}^{n} x_i \in M \right\}$$

maximizing the utility function $U$:

$$\hat{x} \in K \ \& \ U(\hat{x}) = \sup_{x \in K} U(x) \tag{0.5}$$

This concept of utility function played (and is still playing) a crucial role in economic theory, and has been at the origin and the them of many heated debates.

Among the first question which arises is the following: who will choose this collective utility function ?, the public interest (who knows it ?), a dictator ?, a planning bureau ?

Since the $n$ consumers are composing the "collectivity" whose behavior is described by the collective utility function $U$, one generally acquiesces to build $U$ from the utility functions $U_i$ of the consumers. For instance, the collective utility function is a weighted sum of the individual utility functions:

$$U(x) := \sum_{i=1}^{n} \lambda_i U_i(x_i)$$

In this case, the problem is shifted to the one of choosing the weights $\lambda_i$ attributed to each consumer. This is typically a game-theoretical issue. But even if we assume that this problem of allotting weights among consumers is solved, the question remains to know whether utility functions are the right metaphors for the behavior of consumers. Indeed, the concept of utility function has raised and still raises many issues.

The **cardinality versus ordinality** dispute is by now settled. Many economists did challenge the possibility for any economic agent to associate with any commodity a cardinal number measuring her utility or satisfaction and stressed the fact that utility functions played uniquely an "ordinal" role for comparing two commodities. What matter is the **preference preorder** $\preceq$ defined by

$$x \preceq y \text{ if and only if } U(x) \geq U(y)$$

Recall that a preorder is a reflexive and transitive binary relation, **complete** if any two elements are comparable, **partial** in the opposite case.

We can associate with any preorder the equivalence relation $\sim$ defined by

$$x \sim y \text{ if and only if } x \preceq y \;\&\; y \preceq x$$

The "projection" of the preorder to the factor space $X/\sim$ is then an **order relation**.

If $\varphi : \mathbf{R} \mapsto \mathbf{R}$ is an increasing function, then the utility functions $U$ and $\varphi \circ U$ generate the same preference preorder.

Hence the problem of representing any given preference preorder by a utility function was a real issue, because, in particular, the lexicographic order in $\mathbf{R}^n$ cannot be associated with a continuous utility function. Debreu ended this debate by giving reasonable sufficient conditions for a preorder on a finite dimensional vector-space to be represented by a continuous utility function.

Utility functions do not provide the more judicious representation of a consumer in a dynamical framework. In this case, concepts of change or transformations in a given direction $v$ are better embodied in the various concepts of **directional derivatives**. Starting from a commodity $x \in Y$ in a direction $v$, the **satisfaction caused by this move** can be described by the infinitesimal utility increment measured by adequate limits of the differential quotients

$$\frac{U(x+hv) - U(x)}{h}$$

when $h \mapsto 0+$ (the usual gradient is no longer sufficient, because in many cases, utility functions being built not only through standard algebraic operations, but also by supremum or infimum, are no longer differentiable in the classical sense. Nonsmooth analysis is then required.)

Nowadays, the concept of **rationality** became synonymous of the narrow notion of making optimal decisions. An individual, regarded as a decision-maker, is then reduced to an utility function postulated to summarize her behavior. Even the broader conceit of the ability of making transitive inferences is more a dream than a reality, as cognitive psychology acknowledges nowadays.

We shall see later that both in the static and dynamical frameworks, we can discard utility functions.

However, because of the historical importance of this point of view on one hand, and the importance of optimization theory on the other hand, we shall recall the main results of optimization theory in this framework.

Convexity will play a major role in this study, and, in particular, utility functions will be assumed to be concave. In order to avoid using both adjectives, convex and concave, we shall avoid maximizing concave functions and we simply shall minimize convex functions. This is the reason why we shall replace utility functions $U$ by ... **loss functions** $V := -U!$, asking economists to forsake their traditions for the comfort (or laziness) of the mathematicians.

We devote the first section to state the Optimal Allocation Theorem in the convex case. Indeed, convex analysis goes much beyond providing the mere existence of an optimal allocation. Duality Theory exhibits **prices** that emerge from the problem, which solve an associated **dual optimization problem**. Duality Theory of Convex Analysis reveals **demand and supply maps** on one hand, **change and pricing maps** on the other, which are concealed in this simple optimization problem.

It shows that **such a price clears the market**, in the sense that the optimal allocation is made of consumptions which belong to the demand maps of each consumer and that the total consumption is in the supply map.

It demonstrates in a dual way that for such a price, consumers never change tehir consumptions.

It exhibits also the **marginal property** of such a price, which measures the marginal variations of the collective utilities when the set of scarce resources is perturbed.

Last but not least, it conceals two dynamical algorithms which are the prototypes of both the Walras tâtonnement model (which is not viable) and the nontâtonnement model which we shall study in the third part.

The Walras tâtonnement model is nothing else than the (continuous) gradient method applied to the "dual minimization problem". It states that the variations of the prices are in the excess demand (demand minus supply) maps: It increases whenever demand increases. One can prove that it converges to an equilibrium price, which is a Walras equilibrium for this particular excess demand map. But, as it was already mentioned, the associated consumptions do not constitute an allocation whenever the price is not an equilibrium.

But thanks to the concept of subdifferentiability of nondifferentiable convex functions, one can show that the (continuous) gradient method applied to the initial optimization problem provides the evolution of consumptions which form at each instant allocations and which converges to an optimal allocation. Indeed, not only Convex Analysis reveals demand and supply maps, but also the **change and pricing** maps with which we shall build the general nontâtonnement models of Part 3 and beyond.

The next sections supply the minimum needed in Convex Analysis and Duality Theory to prove this Theorem. We shall not prove however the theorem stating the existence, uniqueness and convergence of solutions to gradient inclusions $x'(t) \in -\partial V(x(t))$.

# 1 Allocations of Scarce Resources

## 1.1 The Commodity Space

An **economic commodity** is by definition a good or a service supplied with a **measure unit**. Commodities can be dated, localized, contingent, etc. In this case, they are different. Two dated commodities with different dates and otherwise the same characteristics are different commodities.

Actually, one can characterize commodities by the services which they produce.

In summary, we start with $l$ commodities labeled $h = 1, \ldots, l$ and we denote by

$$e^h := (0, \ldots, 1, \ldots, 0) \in Y := \mathbf{R}^l$$

the **unit commodity** $h$ (where 1 is at the $h$ place).

We begin by assuming that **the commodities are indefinitely divisible**. This a quite rough approximation of economic reality, but an imperative one which allows to describe mathematically the space of commodities as a finite dimensional vector-space.

A **commodity bundle**, or, in a more descriptive way, a **commodity basket**, is a basket made of $x_1$ units of commodity 1, $x_2$ units of commodity 2, ... , $x_l$ units of commodity $l$. It is represented by the vector

$$x := \sum_{h=1}^{l} x_h e^h = (x_1, x_2, \ldots, x_l) \in Y$$

For the sake of simplicity, we shall speak from now on of commodities instead of commodity bundles or baskets.

So, the familiar finite dimensional vector-space $Y := \mathbf{R}^l$ is regarded as the **commodity space**, the canonical basis of which is made of the units of goods.

Naturally, some commodities will be eliminated, such as, for instance, commodities with negative units, which, at first glance, do not make sense.

Actually, it may be wise to accept negative goods if they are adequately interpreted. Later on, we will distinguish among **produced** commodities and **consumed** commodities. One may represent for instance consumed commodities with a positive sign and produced commodities with a negative sign. Or, as another example, **labor** can be regarded as a negative **leisure**.

Why do we represent the commodity space by the finite dimensional vector-space $\mathbf{R}^l$? The reason is that we can indeed add commodities and multiply them by scalars, i.e., perform linear combinations of commodities. Therefore, we shall be able to exploit the rich structure of linear spaces.

## 1.2 The Value Space

Very early economic activity, actually, trading activity, required the comparison of two commodities before an exchange, or a barter or swap. This is done by associating with each commodity its **value** expressed in **accounting or monetary units**, such as the ECU (European Currency Unit), the Franc, the Dollar, etc.

Even though barter is still used (in international trade), the idea to use a specific commodity, easy to handle and sufficiently divisible (liquid), as a unique mean of comparison appeared quite early. In order to compare two arbitrary commodities, each of them is compared with this specific one.

This specific commodity used to compare arbitrary commodities is called the **numéraire**. The **value** of a commodity can be expressed in amounts of units of numéraire judged equivalent to this commodity.

The choice of a numéraire requires a consensus among the economic agents trading the commodities, the faith or belief that everyone agrees on the common value of the unit of numéraire. This is why the numéraire is called a fiduciary good.

Economic history shows the evolution of numéraires, from specific and useful goods (camels, cows, etc.) to seldom employable goods (like shells, gold, etc.) to paper money and now, to abstract figures concealed in computer memories of some banks. Nowadays, the numéraire is made of an explicit, although abstract, commodity bundle used to make an **index**.

On should note confuse **money** with either **accounting units** or **numéraire**. Money has no definite meaning, conveys many different kinds of concepts and play different roles (storage of values, reward for risk taking, etc.).

Here, **sociopsychology**, in the sense of psychology of masses, plays an important role, since the consensus on the choice of a numéraire must be reached before we can use it in an economic model.

Hence, the choice of a numéraire and its value depend upon the set of economic agents which accept it and evolves with time. Hence, the space of values is one dimensional space, the unit of which is the unit of account, the Lira of May 15, 1992 for instance.

Actually, in complex economies of today, there are many different fiduciary goods, which add to the space of physical commodities (subject to inviolable scarcity constraints) a more and more complex space of fiduciary commodities (subject to psychological constraints, resulting from unknown psychological mechanisms governing the emergence of fashions, etc.). This aspect of things will not be taken into account in these lectures, naturally, but they should be kept in mind in order not to rely too much to the very unassuming and crude mathematical description of our humble economic problem.

## 1.3 The Price Space

How can two commodities be compared through a numéraire ? The simple idea is, as we have said, to express the value of each commodities in terms of units of numéraire.

The mechanism which associates with a commodity this amount of numéraire is what is called the **price system** or simply, the **price**. A price $p$ is then a map from the commodity space $Y$ to the values space $\mathbf{R}$, associating to each commodity its value.

Since we have represented the commodity space by a vector space, in which one can perform linear combination of commodities, it is natural to continue to accept the relevance of the linear structure and to assume that the price is linear: the value of the sum of two commodities is the sum of the values and the value of $\lambda$ times a commodity is $\lambda$ times its value. In other words,

$$p(\lambda_1 x_1 + \lambda_2 x_2) \;=\; \lambda_1 p(x_1) + \lambda_2 p(x_2)$$

The Price Space is then the **dual** $Y^\star := \mathcal{L}(Y, \mathbf{R})$ of the Commodity Space.

We shall supply the Price Space with the **dual basis**

$$(e_1^\star, e_2^\star, \ldots, e_l^\star)$$

where $e_h^\star$ associates with any commodity (basket) $x := (x_h)_{h=1,\ldots,l} \in Y$ the amount of units of the $h$th commodity: $e_h^\star(x) = x_h$.

We then deduce that

$$p(x) \;=\; \sum_{h=1}^{l} x_h p(e^h) \;=\; \sum_{h=1}^{l} e_h^\star(x) p(e^h) \;=\; \left(\sum_{h=1}^{l} e_h^\star(x) p(e^h)\right)(x)$$

This shows that $p := \left(\sum_{h=1}^{l} e_h^\star(x) p(e^h)\right)$ is a linear combination of the elements of the canonical basis. The components $p^h := p(e^h)$ of $p$ in this dual basis is the **value of the unit of commodity** $h$, what is meant in the day to day language by the **price of** $h$.

We the write

$$p(x) \;=\; \sum_{h=1}^{l} p^h x_h \;=:\; \langle p, x\rangle$$

This nondegenerate bilinear form

$$(p, x) \in Y^\star \times Y \;\mapsto\; \langle p, x\rangle \;:=\; p(x) \in \mathbf{R}$$

is called the **duality pairing**.

In general, we shall be led to choose **nonnegative prices**, i.e., prices in the positive cone

$$\mathbf{R}_+^{l\star} \;:=\; \left\{p \in \mathbf{R}^{l\star} \mid p^h \geq 0\right\}$$

This not always judicious. In instances when one consumer is forced to consume all available goods, it is sensible to accept negative prices. A glass of water in the desert may be attributed by someone a positive price, whereas in a dirty basement of a torture building, a victim is ready to attribute the last glass of water a negative price. Remember your childhood when a excessively caring mother forced you to finish your soup ...

So, to summarize, the first role played by a price is to compare two commodities $x$ and $y$ by comparing their value $\langle p, x\rangle$ and $\langle p, y\rangle$.

If a price $p$ plays this role, so are the prices $\lambda p$ for any positive scalar $\lambda$. Therefore, one can change the scale of prices (or price level) without altering this role (this is called **monetary illusion**).

So, we need a further condition to fix the price level. This is done by fixing the value of a numéraire $\omega \in Y$: the unit of gold (the "Bretton Woods gold exchange standard" until Nixon cancelled it on August 13, 1971), a commodity basket entering the composition of an **index**, etc.:

$$\langle p, \omega \rangle = \mu$$

There are no longer "real" numéraires nowadays, but **commodity indexes**, the value of which is observed and measured (rather than being fixed in an evolving — and not really controllable — world).

This is time for a warning that in evolving models, the value of the numéraire evolves (although it should remain constant to satisfy the expectations (or dreams) of economists and finance ministers).

Here, we shall take for numéraire the commodity

$$\omega := (1,1,\ldots,1) \ \& \ \mu := 1$$

We shall agree to take for price set the **price simplex** $S^l$ defined by the normalization rule:

$$S^l := \left\{ p \in \mathbf{R}_+^{l\star} \mid \sum_{h=1}^{l} p^h = 1 \right\}$$

## 1.4 The set of Resources

We denote by $M \subset Y$ the set of physical **scarce resources** to be allocated among $n$ consumers.

Scarcity is the key word, the basic requirement without which there would be no need of economics.

When producers are taken into account, the commodities of $M$ to be allocated among consumers are **produced** by producers. The set $M$ thus depends upon past and present decisions of producers. At least for the beginning, we shall assume only that there is a constant set $M$ of scarce resources available at each time to consumers.

The first law of economics we shall comply with states that **it is impossible to consume more (physical) resources than available** (by opposition to fiduciary goods).

Throughout this book, we assume mainly that $M$ is a **closed** and that it satisfies the **free disposal assumption**:

$$M = M - \mathbf{R}_+^l \tag{1.1}$$

This means that any commodity $y \leq x$ smaller than or equal to an available commodity $x$ is still available.

Since $M$ is a set of resources, it should be **bounded above** in the sense that

$$\exists\, \overline{y} \in Y \quad \text{such that } M \subset \overline{y} - \mathbf{R}^l_+ \tag{1.2}$$

(It cannot be bounded below because of the free disposal assumption).

We shall assume sometimes, for simplicity, that $M$ is **convex**: convex combinations of svarce resources are still available.

This is interpreted by economists by saying that **decreasing return to scale** prevails. If further more $M$ is a cone, they say that **constant return to scale** prevails.

Actually, we shall be able to bypass this assumption in the dynamical case[12].

Later on, when evolution will be taken into account, it will be possible to have $M$ depend upon the time and cumulated consequences of past allocations, in order to take into account investments, pollution, etc.).

## 1.5 Introducing the Consumers

We begin now the mathematical description of the $n$ consumers $i = 1, \ldots, n$.

It starts by her **consumption set** $L_i \subset Y$, which represents the set of potential consumptions. Actually, it is better to say that she will never accept a commodity outside her consumption set $L_i$. Most often, $L_i$ is chosen to be the orthant $\mathbf{R}^l_+$.

Throughout this book, we assume mainly that the consumption sets $L_i$ are **closed**.

Consumers are often assumed to have no **satiation**: this means that they is no limit to their desire to squander[13]. We describe it in mathematical terms by stating that $L_i = L_i + \mathbf{R}^l_+$.

We shall assume also that $L_i$ is **bounded below**, i.e., that

$$\exists\, \overline{x}_i \in Y \quad L_i \subset \overline{x}_i + \mathbf{R}^l_+$$

Again, for simplicity, we shall sometimes assume that the consumptions sets are **convex**.

[12] and replace it by a regularity assumption called **sleekness**.

[13] Despite its first glance appeal, this assumption is not always sensible.

### 1.6 The Set of Allocations

We translate now the first economic law: **it is impossible to consume more (physical) commodities than available** by introducing the set $K$ of allocations of scarce resources among $n$ consumers.

We denote by $X := Y^n = \mathbf{R}^{ln}$ the Consumption Space of the $n$ consumers. We set

$$x := (x_1, \ldots, x_n) \in X$$

where $x_i$ does no longer denote a component of a commodity bundle, but the commodity bundle of consumer $i$[14].

Therefore, the set of allocations is equal to

$$K := \left\{ x := (x_1, \ldots, x_n) \in \prod_{i=1}^{n} L_i \;\middle|\; \sum_{i=1}^{n} x_i \in M \right\} \tag{1.3}$$

Conforming to the first economic law amounts to evolving in the allocation set $K$ or to choosing elements (optimal ones or equilibria) in this allocation set.

Consequently, to proceed further, we need to make novel assumptions on the nature of the questions to answer and the behavior of consumers.

If everyone may easily agree on accepting the first economic law, the consensus about the behaviors of consumers and the way to describe them mathetically is far to be perfect and bound to evolve.

## 2 The Optimal Allocation Theorem

For simplicity, we shall incorporate the weights $\lambda_i$ and the consumption sets $L_i$ in the loss functions $V_i : Y \mapsto \mathbf{R} \cup \{+\infty\}$ of the consumers $i = 1, \ldots, n$ by setting

$$V_i(x) := \begin{cases} -\lambda_i U_i(x) & \text{if } x \in L_i \\ \\ +\infty & \text{if } x \notin L_i \end{cases}$$

Hence, an optimal allocation $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_n)$ is a solution to the minimization problem

$$v := \inf_{x \in K} \sum_{i=1}^{n} V_i(x_i) = \sum_{i=1}^{n} V_i(\overline{x_i})$$

[14]It is hoped that this slight abuse of notation is forgiven by the reader. The context should efface any ambiguity.

where

$$K := \left\{ x := (x_1, \ldots, x_n) \in \prod_{i=1}^{n} \mathrm{Dom}(V_i) \;\middle|\; \sum_{i=1}^{n} x_i \in M \right\}$$

We shall assume that

$$\forall\, q \in \mathbf{R}^l_+, \;\; \inf_{x \in L_i} (\langle q, x\rangle + V_i(x)) > -\infty \tag{2.1}$$

and that the set of scarce resources $M$ satisfy

$$\left\{ \begin{array}{ll} i) & M = M - \mathbf{R}^l_+ \text{ is a closed convex subset} \\ ii) & M \subset \underline{y} - \mathbf{R}^l_+ \end{array} \right. \tag{2.2}$$

From the knowledge of the loss functions and the set of scarce resources we shall extract concealed features on the behavior of the consumers.

## 2.1 Demand and Change Maps

We shall denote by

$$\widetilde{D}_i(q) := \left\{ x^i \in \mathbf{R}^l \mid \langle q, x_i\rangle + V_i(x_i)) = \inf_{x \in \mathbf{R}^l} (\langle q, x\rangle + V_i(x)) \right\} \tag{2.3}$$

the Walrasian demand of consumer $i$,

$$B_i(q, r) := \{x \in \mathrm{Dom}(V_i) \mid \langle q, x\rangle \leq r\}$$

the **budget set** of consumer $i$ and by

$$D_i(q, r) := \left\{ x_i \in B_i(q, r) \,|V_i(x_i) = \inf_{x \in B_i(q,r)} V_i(x) \right\}$$

her **demand set**. The **demand map** is the set-valued map $(q, r) \rightsquigarrow D_i(q, r)$. We observe at once that

$$\forall\, \overline{x}_i \in \widetilde{D}_i(q), \;\; \widetilde{D}_i(q) = D_i(q, \langle q, \overline{x}_i\rangle)$$

is independent of the choice of $\overline{x}_i \in D_i(q, \langle q, \overline{x}_i\rangle)$.

**Change maps** $C_i : L_i \times S^l \rightsquigarrow Y^\star$ which express the satisfaction of consumer $i$, are defined by:

$$C_i(x, q) := \left\{ p \in \mathbf{R}^{l\star} \mid V_i(x) + \langle p + q, x\rangle = \inf_{y \in Y} (V_i(y) + \langle p + q, y\rangle) \right\}$$

In other words, $p \in C_i(x,q)$ is the price for which the commodity $x$ minimizes the sum of the loss $V_i(y) + \langle q, y\rangle$ and the cost $\langle p, y\rangle$.

If the loss function $V_i$ is differentiable at $x$, then $C_i$ is single-valued and can be written

$$C_i(x,q) \;=\; -V_i'(x) - q$$

We observe that

$$0 \;\in\; C_i(x_i, q) \text{ if and only if } x_i \;\in\; \widetilde{D}_i(q)$$

In summary, we can associate with any consumer $i$ represented by a loss function $V_i$ a demand map $\widetilde{D}_i$ associating with any price $q$ a set of commodities minimizing her loss under budgetary constraints and a change map $C_i$ associating with any price $q$ the change of consumption.

## 2.2 Supply and Pricing Maps

We associate now with the set $M \subset Y$ of scarce resources the supply map $S_M$ associating with any $q \in Y^\star$ the supply set $S_M(q) \subset M$ defined by

$$S_M(q) \;:=\; \left\{ \overline{y} \in M \mid \langle q, \overline{y}\rangle \;=\; \sup_{y\in M}\langle q, y\rangle \right\}$$

of scarce resources which maximize the income

$$\sigma_M(q) \;:=\; \sup_{y\in M}\langle q, y\rangle$$

induced by the available resources.

We also introduce the inverse $N_M := S_M^{-1}$ of the supply map $S_M$:

$$q \;\in\; N_M(y) \text{ if and only if } y \;\in\; S_M(q)$$

which we regard as the pricing map.

Assumption (2.2) implies that

$$\sigma_M(q) \;<\; +\infty \text{ if and only if } q \;\in\; \mathbf{R}^l_+$$

We also observe that

$$\forall\, y \in \mathrm{Int}(M),\;\; N_M(y) \;=\; \{0\}$$

## 2.3 Optimal Allocations

**Theorem 2.1** *Let us assume the set $M$ of scarce resources satisfy assumptions (2.2), that the consumption set $L_i$ are closed and convex and that the loss functions $V_i$ are nontrivial, convex and lower semicontinuous and satisfy assumptions (2.3). Assume furthermore that*

$$0 \in \operatorname{Int}\left(\sum_{i=1}^{n} L_i - M\right) \tag{2.4}$$

*Then there exists an optimal allocation $\overline{x} \in K$ which is a solution to the optimal allocation problem*

$$v := \inf_{x \in K} \sum_{i=1}^{n} V_i(\overline{x_i})$$

*where*

$$K := \left\{ x := (x_1, \ldots, x_n) \in \prod_{i=1}^{n} \operatorname{Dom}(V_i) \;\middle|\; \sum_{i=1}^{n} x_i \in M \right\}$$

*Furthermore, there exists a price $\overline{q} \in Y^\star$ such that*

*a) the price $\overline{q}$ and the allocation $(\overline{x}_1, \ldots, \overline{x}_n)$ satisfy*

$$\begin{cases} i) & \forall\, i = 1, \ldots, n, \;\; \overline{x_i} \in \widetilde{D}_i(\overline{q}) \\ & \text{i.e., each } \overline{x}_i \text{ belongs to consumer } i\text{'s demand set} \\ ii) & \sum_{i=1}^{n} \overline{x}_i \in S_M(\overline{q}) \\ & \text{i.e., } \sum_{i=1}^{n} \overline{x}_i \text{ maximizes the available income } \langle \overline{q}, y \rangle \end{cases}$$

*b) the optimal price $\overline{q}$* clears the market *in the sense when it is a solution to the inclusion*

$$0 \in S_M(\overline{q}) - \sum_{i=1}^{n} \widetilde{D}_i(\overline{q})$$

*stating that the supply $S_M(\overline{q})$ is balanced by the total demand $\sum_{i=1}^{n} \widetilde{D}_i(\overline{q})$.*

*c) the optimal allocation $\overline{x}$ is an equilibrium of the associated nontâtonnement process in the sense that*

$$\begin{cases} i) & \forall\, i = 1, \ldots n, \;\; 0 \in C_i(\overline{x}_i, \overline{q}) \\ ii) & \overline{q} \in N_M\left(\sum_{i=1}^{n} \overline{x}_i\right) \end{cases}$$

We shall also prove that this price $\overline{q}$ enjoys a **marginal property**. We introduce perturbations on the resources and we define the **marginal function** $v$ associating with any resource $y$ the optimal value

$$v(y) := \inf_{x \in K(y)} \sum_{i=1}^{n} V_i(\overline{x_i})$$

where

$$K(y) := \left\{ x := (x_1, \ldots, x_n) \in \prod_{i=1}^{n} \mathrm{Dom}(V_i) \;\middle|\; \sum_{i=1}^{n} x_i \in M - y \right\}$$

We observe that $v(0) = v$. Naturally, the marginal function $v$ is not necessarily differentiable. But it is convex, and we shall extend the concept of differential to a concept of **subdifferential**. With this notion, we shall prove that $\overline{q}$ **belongs to the subdifferential of the marginal function at** $y := 0$.

## 2.4 The Walras Tâtonnement

Since we have associated with any consumer her demand map $\widetilde{D}_i$ and we have defined a supply map $S_M$, we can define the continuous **tâtonnement process** defined by

$$q'(t) \in \sum_{i=1}^{n} \widetilde{D}_i(q(t)) - S_M(q(t)) \tag{2.5}$$

which is a metaphor for the law of supply and demand. Then the following result hods true:

**Theorem 2.2** *We posit the assumptions of Theorem 2.1. For any positive initial price $q_0$, there exists a unique solution $q(\cdot)$ to the tâtonnement process (2.5) starting at $q_0$ with converges to an equilibrium $\overline{q}$ when $t \to +\infty$.*

Unfortunately, we have seen that tâtonnement processes are not viable.

## 2.5 The Nontâtonnement Process

However, we can design another dynamical process which is viable, through **change maps** $C_i : L_i \times S^l \rightsquigarrow Y$ which express the satisfaction of consumer $i$.

We then define the dynamical behavior of consumer $i$ by the **differential inclusion**

$$x_i'(t) \in C_i(x_i(t), q(t))$$

which is controlled by the price $q(t)$.

**Theorem 2.3** *We posit the assumptions of Theorem 2.1. From any initial allocation* $x_0 = (x_{01}, \ldots, x_{0n}) \in K$ *starts a* unique solution $x(\cdot) = (x_1(\cdot), \ldots, x_n(\cdot))$ *to the system of differential inclusions*

$$x_i'(t) \in C_i(x_i(t), q(t))$$

*where the price* $q(t)$ *satisfies*

$$\forall\, t \geq 0, \;\; q(t) \in N_M\left(\sum_{i=1}^{n} x_i(t)\right)$$

*and* which are viable in $K$ *in the sense that*

$$\forall\, t \geq 0, \;\; x_i(t) \in L_i \; (i = 1, \ldots, n) \;\; \& \;\; \sum_{i=1}^{n} x_i(t) \in M$$

*Furthermore, the total loss* $t \to \sum_{i=1}^{n} V_i(x_i(t))$ *decreases and the allocation* $x(t)$ *converges to an optimal allocation* $\overline{x}$ *when* $t \to +\infty$.

Observe that $q(t) = 0$ whenever the total consumption $\sum_{i=1}^{n} x_i(t)$ belongs to the interior of the set $M$ of scarce resources.

Therefore, under convexity assumptions, one can derive from the problem of optimal allocation many more informations than the mere existence of an optimal allocation. First, the concept of price emerges, and we can associate the concepts of demand maps and supply maps. We stated that there exists a price $\overline{q}$ which clears the marker: the total optimal consumption is in the supply set and each optimal consumption belongs to the demand set. This price has marginal properties.

The tâtonnement process can be defined, and, given an initial price, it has a unique solution converging to an equilibrium price.

We can also derive decentralized dynamical processes of each consumer, described by a differential inclusion controlled by prices. Starting from any initial allocation, there exists a unique allocation evolving according to this controlled dynamical process which converges to an optimal allocation when $t \mapsto +\infty$.

In the next parts, we shall answer the same type of questions (existence of an equilibrium, evolution of allocations) without grounding the theory on the assumptions of (convex) utility functions.

Convexity is indeed the main ingredient guaranteeing the above results. In order to prove them, we provide below the minimal exposition of convex analysis.

# 3 Convex Functions

## 3.1 Extended Functions and their Epigraphs

A function $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ is called an **extended (real-valued) function**. Its **domain** is the set of points at which $V$ is finite:

$$\text{Dom}(V) := \{x \in X \mid V(x) \neq \pm\infty\}$$

A function is said to be **nontrivial**[15] if its domain is not empty. Any function $V$ defined on a subset $K \subset X$ can be regarded as the extended function $V_K$ equal to $V$ on $K$ and to $+\infty$ outside of $K$, whose domain is $K$.

Since the order relation on the real numbers is involved in the definition of the Lyapunov property as well as in minimization problems, we no longer characterize a real-valued function by its graph, but rather by its **epigraph**

$$\mathcal{E}p(V) := \{(x, \lambda) \in X \times \mathbf{R} \mid V(x) \leq \lambda\}$$

or by its **hypograph** defined in a symmetric way by

$$\mathcal{H}p(V) := \{(x, \lambda) \in X \times \mathbf{R} \mid V(x) \geq \lambda\} = -\mathcal{E}p(-V)$$

**The graph of a real-valued function is then the intersection of its epigraph and its hypograph.**

We also remark that some properties of a function are actually properties of their epigraphs. For instance, **an extended function $V$ is convex (resp. positively homogeneous) if and only if its epigraph is convex (resp. a cone).**

The main examples of extended functions are the **indicators** $\psi_K$ **of subsets** $K$ defined by

$$\psi_K(x) := \begin{cases} 0 & \text{if } x \in K \\ +\infty & \text{if not} \end{cases}$$

It can be regarded as a **membership cost**[16] to $K$: it costs nothing to belong to $K$, and $+\infty$ to step outside of $K$.

The indicator $\psi_K$ is lower semicontinuous if and only if $K$ is closed and $\psi_K$ is convex if and only if $K$ is convex. One can regard the sum $V + \psi_K$ as

[15] Such a function is said to be proper in convex and non smooth analysis. We chose this terminology for avoiding confusion with proper maps.

[16] Functions $V : X \mapsto [0, +\infty]$ can be regarded as some kind of fuzzy sets, called toll sets. For "fuzzy differential inclusions" and "fuzzy viability", we refer to Chapter 10, Section 3 of [16, Aubin].

the restriction of $V$ to $K$. Therefore, a constrained minimization problem is equivalent to an unconstrained one for a new criterion function, which embodies the constraints so to speak:

$$\inf_{x\in K} V(x) \equiv \inf_{x\in X}(V(x)+\psi_K(x))$$

We recall the convention $\inf(\emptyset) := +\infty$.

**Lemma 3.1** *Consider a function $V : X \mapsto \mathbf{R}\cup\{\pm\infty\}$. Its epigraph is closed if and only if*

$$\forall\, v \in X,\ V(v) = \liminf_{v'\to v} V(v')$$

*For extended functions $V$ which never take the value $-\infty$, this is equivalent to the* lower semicontinuity *of $V$.*

*Assume that the epigraph of $V$ is a closed cone. Then the following conditions are equivalent:*

$$\begin{cases} i) & \forall\, v \in X,\ V(v) > -\infty \\ ii) & V(0) = 0 \\ iii) & (0,-1) \notin \mathcal{E}p(V) \end{cases}$$

**Proof** — Assume that the epigraph of $V$ is closed and pick $v \in X$. There exists a sequence of elements $v_n$ converging to $v$ such that

$$\lim_{n\to\infty} V(v_n) = \liminf_{v'\to v} V(v')$$

Hence, for any $\lambda > \liminf_{v'\to v} V(v')$, there exist $N$ such that, for all $n \geq N$, $V(v_n) \leq \lambda$, i.e., such that $(v_n,\lambda) \in \mathcal{E}p(V)$. By taking the limit, we infer that $V(v) \leq \lambda$, and thus, that $V(v) \leq \liminf_{v'\to v} V(v')$. The converse statement is obvious.

Suppose next that the epigraph of $V$ is a cone. Then it contains $(0,0)$ and $V(0) \leq 0$. The statements *ii)* and *iii)* are clearly equivalent.

If *i)* holds true and $V(0) < 0$, then

$$(0,-1) = \frac{1}{-V(0)}(0,V(0))$$

belongs to the epigraph of $V$, as well as all $(0,-\lambda)$, and (by letting $\lambda \to +\infty$) we deduce that $V(0) = -\infty$, so that *i)* implies *ii)*.

To end the proof, assume that $V(0) = 0$ and that for some $v$, $V(v) = -\infty$. Then, for any $\varepsilon > 0$, the pair $(v,-1/\varepsilon)$ belongs to the epigraph of $V$, as well as the pairs $(\varepsilon v,-1)$. By letting $\varepsilon$ converge to 0, we infer that $(0,-1)$ belongs also to the epigraph, since it is closed. Hence $V(0) < 0$, a contradiction. □

## 3.2 Subdifferential of Convex Functions

Convex functions enjoy further properties. We already mentioned that an extended function is convex (respectively lower semicontinuous) if and only if its epigraph is convex (respectively closed.)

Moreau and Rockafellar introduced the subdifferential of convex functions in the early 60's:

**Definition 3.2** *Consider a nontrivial function $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ and $x \in \mathrm{Dom}(V)$. The closed convex subset $\partial V(x)$ defined by*

$$\partial V(x) \; = \; \{p \; \in \; X^\star \mid \forall \, y \; \in \; X, \; < p, y - x > \leq \; V(y) - V(x)\}$$

*(which may be empty) is called the* subdifferential *of $V$ at $x$. We say that $V$ is* subdifferentiable at *$x$ if $\partial V(x) \neq \emptyset$.*

From the definition, we see that the Fermat Rule follows immediately:

**Theorem 3.3** *Let $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ be a non trivial function. Then the following conditions are equivalent:*

$$\left\{ \begin{array}{ll} i) & 0 \; \in \; \partial V(\overline{x}) \;\; \text{(the Fermat Rule)} \\ ii) & \overline{x} \;\; \text{minimizes } V \end{array} \right.$$

We also observe that the concept of subdifferential generalizes the concept of gradient in the following sense:

**Proposition 3.4** *If $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ is convex and differentiable at a point $x \in \mathrm{Int}(\mathrm{Dom}(V))$, then*

$$\partial V(x) \; = \; \{V'(x)\}$$

**Proof** — First, the gradient $V'(x)$ belongs to $\partial V(x)$, since, $V$ being convex, inequalities

$$\frac{V(x + h(y - x))}{h} \; \leq \; V(y) - V(x)$$

imply by letting $h$ converge to 0 that

$$\forall \, y \; \in \; X, \; < V'(x), y - x > \leq \; V(y) - V(x)$$

Conversely, if $p \in \partial V(x)$, we obtain, by taking $y = x + hu$ that

$$\langle p, u \rangle \; \leq \; \frac{V(x + hu) - V(x)}{h}$$

By letting $h$ converge to 0, we infer that for every $u \in X$, $\langle p, u \rangle \leq \langle V'(x), u \rangle$, so that $p = V'(x)$. □

**Proposition 3.5** *The subdifferential map* $x \rightsquigarrow \partial V(x)$ *is* monotone *in the sense that*

$$\forall\, x, y \in X,\ \ \forall\, p \in \partial V(x),\ q \in \partial V(y),\ \ \langle p - q, x - y \rangle \geq 0 \tag{3.1}$$

**Proof —** Indeed, since

$$\begin{cases} \forall\, p \in \partial V(x), \ \ \langle p, x - y \rangle \ \geq \ V(x) - V(y) \\ \\ \forall\, q \in \partial V(y), \ \ \langle q, y - x \rangle \ \geq \ V(y) - V(x) \end{cases}$$

we deduce the monotonicity of the subdifferential by adding those two inequalities. □

Monotone maps, and above all, maximal monotone maps, enjoy many of the properties of positive linear operators. We refer to [24, Aubin & Ekeland] for more details on monotone maps. □

We recall the following important property of convex functions defined on finite dimensional vector-spaces:

**Theorem 3.6** *A convex function defined on a finite dimensional vector-space is locally Lipschitz and subdifferentiable on the interior of its domain.*

(See for instance [24, Aubin & Ekeland] for a proof.)

Therefore, in order to apply Fermat Rule, we need a Subdifferentiable Calculus, for which we need the concept of conjugate functions. This is how prices will emerge when we shall apply the Fermat Rule to the optimization problem (0.5).

But we have to mention right now that the Fermat Rule replaces the minimization problem by an equilibrium problem: Indeed, the inclusion

$$0 \ \in \ \partial V(\overline{x})$$

shows that the constant function $x(t) \equiv \overline{x}$ is a solution to the differential inclusion

$$\text{for almost all } t \geq 0, \ \ x'(t) \ \in \ -\partial V(x)$$

(continuous descent method or subdifferential algorithm).

Actually, we shall show that such differential inclusions do have solutions in section 1.4.

**Theorem 3.7** *Assume that $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ is nontrivial, convex, lower semicontinuous and bounded below. Then, for any initial state $x_0 \in$ Dom$(V)$, there exists a unique solution to the differential inclusion*

$$\text{for almost all } t \geq 0, \ \ x'(t) \ \in \ -\partial V(x) \tag{3.2}$$

*starting at $x_0$.*

*Let $V_0'(x))$ denote the element of $\partial V(x)$ with the smallest norm. Then the solution $x(\cdot)$ is* slow *in the sense that for almost any $t$, the norm of the velocity $x'(t)$ is the smallest one:*

$$\text{for almost all } t \geq 0, \ \ x'(t) \ = \ -V_0'(x(t))$$

*Furthermore, if $V$ is inf-compact, then $x(t)$ converges when $t \to \infty$ to a limit $x_*$ which achieves the minimum of $V$:*

$$\lim_{t \to \infty} V(x(t)) \ = \ \inf_{x \in X} V(x) \ = \ V(x_*)$$

**Theorem 3.8** *Let us assume that a convex function $V : X \mapsto \mathbf{R}$ is bounded below.*

*Assume also that the steps of the subgradient algorithm*

$$x_{n+1} \ := \ x_n - \delta_n \frac{p_n}{\|p_n\|}$$

*where $p_n \in \partial V(x_n)$ satisfy*

$$\lim_{n \to \infty} \delta_n = 0 \ \ \& \ \ \sum_{n=0}^{\infty} \delta_n = +\infty$$

*Then the decreasing sequence of scalars*

$$\theta_k \ := \ \min_{n=0,\ldots,k} V(x_n)$$

*converges to the infimum $v := \inf_{x \in X} V(x)$ of $V$ when $k \to \infty$.*

We shall prove this theorem in Section 1.5, as well as a generalization to the case of lower semicontinuous extended convex functions.

## 3.3 Support Functions and Conjugate Functions

There is more to that: lower semicontinuous convex functions enjoy duality properties. In the same way that we associated with cones their polar cones, with closed convex processes their transposes, we can, following Fenchel, associate with lower semicontinuous convex functions conjugate functions for the same reasons, and with the same success.

**Definition 3.9** *Let $K$ be a nonempty subset of a finite dimensional vector-space $X$. We associate with any continuous linear form $p \in X^\star$*

$$\sigma_K(p) := \sigma(K,p) := \sup_{x\in K} < p,x > \in \mathbf{R} \cup \{+\infty\}$$

*The function $\sigma_K : X^\star \mapsto \mathbf{R} \cup \{+\infty\}$ is called the* support function of $K$. *We say that the subsets of $X^\star$ defined by*

$$\begin{cases} i) & K^- := \{p \in X^\star \mid \sigma_K(p) \leq 0\} \\ ii) & K^\perp := \{p \in X^\star \mid \forall\, x \in K, \; < p,x > = 0\} \end{cases}$$

*are the* (negative) polar cone, *and* orthogonal *of $K$ respectively.*

**Examples**

- When $K = \{x\}$, then $\sigma_K(p) = < p,x >$
- When $K = B_X$, then $\sigma_{B_X}(p) = \|p\|_\star$
- If $K$ is a cone, then

$$\sigma_K(p) = \begin{cases} 0 & \text{if} \quad p \in K^- \\ +\infty & \text{if} \quad p \notin K^- \end{cases} \quad \square$$

When $K = \emptyset$, we set $\sigma_\emptyset(p) = -\infty$ for every $p \in X^\star$.
We observe that

$$\forall\, \lambda,\, v > 0, \;\; \sigma_{\lambda L + \mu M}(p) = \lambda \sigma_L(p) + \mu \sigma_M(p)$$

and in particular, that if $P$ is a cone, then

$$\sigma_{M+P}(p) = \begin{cases} \sigma_M(p) & \text{if} \quad p \in P^- \\ +\infty & \text{if} \quad p \notin P^- \end{cases} \quad \square$$

The Separation Theorem[17] can be stated in the following way:

**Theorem 3.10 (Separation theorem)** *Let $K$ be a nonempty subset of a Banach space $X$. Its* closed convex hull *is characterized by linear constraint inequalities in the following way:*

$$\overline{co}(K) = \{x \in X \mid \forall\, p \in X^\star,\ \langle p, x\rangle \leq \sigma_K(p)\}$$

*Furthermore,* there is a bijective correspondence between nonempty closed convex subsets of $X$ and nontrivial lower semicontinuous positively homogeneous convex functions on $X^\star$.

Since the epigraph of a lower semicontinuous convex function is a closed convex subset, it is tempting to compute its support function, and in particular, to observe that

$$\sigma_{\mathcal{E}p(V)}(p, -1) = \sup_{x\in X,\, \lambda\geq V(x)} (\langle p, x\rangle - \lambda) = \sup_{x\in X}(\langle p, x\rangle - V(x))$$

**Definition 3.11** *Let $V : X \to \mathbf{R}\cup\{+\infty\}$ be any nontrivial extended function defined on a finite dimensional vector-space $X$. We associate with it its conjugate function $V^\star : X^\star \to \mathbf{R}\cup\{+\infty\}$ defined on the dual of $X$ by*

$$\forall\, p \in X^\star,\ V^\star(p) := \sup_{x\in X}(< p, x > -V(x))$$

*Its biconjugate $V^{\star\star} : X \mapsto \mathbf{R}\cup\{\pm\infty\}$ is defined by*

$$V^{\star\star}(x) := \sup_{p\,\in\, X^\star}(< p, x > -V^\star(p))$$

We see at once that **the conjugate function of the indicator $\psi_K$ of a subset $K$ is the support function $\sigma_K$.**

We deduce from the definition the following convenient inequality

$$\forall\, x \in X,\ p \in X^\star,\ < p, x > \leq V(x) + V^\star(p)$$

[17]This Separation Theorem is one corner stone of linear and convex functional analysis. It was discovered by the German mathematician Minkowski at the beginning of this century in finite dimensional spaces and extended in the 30's by Hahn, an Austrian mathematician, and Banach, the Polish founder of Linear Functional Analysis, in Banach spaces and in Hausdorff locally convex spaces, including weak topologies of Banach spaces. It is then known under the name of the Hahn-Banach Theorem.

known as **Fenchel's Inequality**. The epigraphs of the conjugate and biconjugate functions being closed convex subsets, the conjugate function is lower semicontinuous and convex and so is its biconjugate when it never takes the value $-\infty$. We observe that

$$\forall\, x \in X, \;\; V^{\star\star}(x) \leq V(x)$$

If equality holds, then $V$ is convex and lower semicontinuous. The converse statement, a consequence of the Hahn-Banach Separation Theorem, is the first basic theorem of convex analysis:

**Theorem 3.12** *A nontrivial extended function $V : X \to \mathbf{R} \cup \{+\infty\}$ is convex and lower semicontinuous if and only if it coincides with its biconjugate. In this case, the conjugate function $V^\star$ is nontrivial.*

So, the Fenchel correspondence associating with any function $V$ its conjugate $V^\star$ is a one to one correspondence between the sets of nontrivial lower semicontinuous convex functions defined on $X$ and its dual $X^\star$. This fact is at the root of duality theory in convex optimization.

**Proof**

a) Suppose that $a < V(x)$. Since the pair $(x, a)$ does not belong to $\mathrm{Ep}(V)$, which is convex and closed, there exist a continuous, linear form $(p, b) \in X^* \times \mathbf{R}$ and $\varepsilon > 0$ such that

$$\forall y \in \mathrm{Dom} V,\; \forall \lambda \geq V(y), \qquad \langle p, y\rangle - b\lambda \leq \langle p, x\rangle - ba - \varepsilon \tag{3.3}$$

by virtue of the **Separation Theorem** (Theorem 2.4).

b) We note that $b \geq 0$. If not, we take $y$ in the domain of $V$ and $\lambda = V(y) + \mu$. We would have

$$-b\mu \leq \langle p, x - y\rangle + b(V(y) - a) - \varepsilon < +\infty.$$

Then we obtain a contradiction if we let $\mu$ tend to $+\infty$.

c) We show that if $b > 0$, then $a < V^{**}(x)$. In fact, we may divide the inequality (3.3) by $b$; whence, setting $\bar{p} = p/b$ and taking $\lambda = V(y)$, we obtain

$$\forall y \in \mathrm{Dom} V, \qquad \langle \bar{p}, y\rangle - V(y) \leq \langle \bar{p}, x\rangle - a - \varepsilon/b.$$

Then, taking the supremum with respect to $y$, we have

$$V^*(\bar{p}) < \langle \bar{p}, x\rangle - a.$$

This implies that

$$\begin{array}{ll} i) & \bar{p} \text{ belongs to the domain of } V^* \\ ii) & a < \langle \bar{p}, x\rangle - V^*(\bar{p}) \leq V^{**}(x). \end{array} \tag{3.4}$$

d) We consider the case in which $x$ belongs to the domain of $V$. In this case, $b$ is always strictly positive. To see this, it is sufficient to take $y = x$ and $\lambda = V(x)$ in formula (3.3) to show that

$$b \geq \varepsilon/(V(x) - a)$$

since $V(x) - a$ is a strictly-positive, real number. Then, from part b), we deduce the existence of $\bar{p} \in \mathrm{Dom} V^*$ and that $a \leq V^{**}(x) \leq V(x)$ for all $a < V(x)$. Thus, $V^{**}(x)$ is equal to $V(x)$.

e) We consider the case in which $V(x) = +\infty$ and $a$ is an arbitrarily-large number. Either $b$ is strictly positive, in which case part b) implies that $a < V^{**}(x)$, or $b = 0$. In the latter case, (3.3) implies that

$$\forall y \in \mathrm{Dom}\, V, \qquad \langle p, y - x\rangle + \varepsilon \leq 0. \tag{3.5}$$

Let us take $\bar{p}$ in the domain of $V^*$ (we have shown that such an element exists, since $\mathrm{Dom}\, V$ is non-empty). Fenchel's inequality implies that

$$\langle \bar{p}, y\rangle - V^*(\bar{p}) - V(y) \leq 0. \tag{3.6}$$

We take $\mu > 0$, multiply the inequality (3.5) by $\mu$ and add it to the inequality (3.6) to obtain

$$\langle \bar{p} + \mu p, y\rangle - V(y) \leq V^*(\bar{p}) + \mu\langle p, x\rangle - \mu\varepsilon.$$

Taking the supremum with respect to $y$, we obtain:

$$V^*(\bar{p} + \mu p) \leq V^*(\bar{p}) + \mu\langle p, x\rangle - \mu\varepsilon$$

which may be written in the form

$$\langle p, x\rangle + \mu\varepsilon - V^*(\bar{p}) \leq \langle \bar{p} + \mu p, x\rangle - V^*(\bar{p} + \mu p) \leq V^{**}(x).$$

Taking $\mu = \dfrac{a + V^*(\bar{p}) - \langle \bar{p}, x\rangle}{\varepsilon}$, which is strictly positive whenever $a$ is large enough, we have again proved that $a \leq V^{**}(x)$. Thus, since $V^{**}(x)$ is greater than an arbitrary finite number $a$, we deduce that $V^{**}(x) = +\infty$. □

We deduce at once the following characterization of the subdifferential:

**Proposition 3.13** *Let $V : X \to \mathbf{R} \cup \{+\infty\}$ be a nontrivial extended convex function defined on a finite dimensional vector-space $X$. Then*

$$p \in \partial V(x) \iff <p, x> = V(x) + V^\star(p)$$

*If moreover the function $V$ is lower semicontinuous, then* the inverse of the subdifferential $\partial V(\cdot)$ is the subdifferential $\partial V^\star(\cdot)$ of the conjugate function*:*

$$p \in \partial V(x) \iff x \in \partial V^\star(p)$$

This result allows us to derive a subdifferential calculus form the calculus of conjugate functions, based on the following Fenchel Theorem.

Since $-V^\star(0) = \inf_{x \in X} V(x)$, the Fermat Rule becomes:

**Theorem 3.14** *Let $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ be a nontrivial lower semicontinuous convex extended function defined on a finite dimensional space $X$. Then $\partial V^\star(0)$ is the set of minimizers of $V$.*

As an example, we obtain

**Corollary 3.15** *Let $K \subset X$ be a closed convex subset. Then*

$$\begin{cases} i) & \partial \psi_K(x) = \{p \in X^\star \quad \text{such that } \langle p, x \rangle = \sup_{y \in K} \langle p, y \rangle\} \\ ii) & \partial \sigma_K(p) = \{x \in K \quad \text{such that } \langle p, x \rangle = \sup_{y \in K} \langle p, y \rangle\} \end{cases}$$

**Definition 3.16** *The first subset is the* normal cone to $K$ at $x$ *and the second one is called the* support zone of $K$ at $p$.

*The negative polar cone of the normal cone $N_K(x)$ to a convex subset is called the* tangent cone to $K$ *at $x$ and is denoted by*

$$T_K(x) := N_K(x)^-$$

It can be easily characterized by:

$$T_K(x) = \overline{S_K(x)}$$

where

$$S_K(x) := \bigcup_{h > 0} \frac{K - x}{h} \quad \square$$

The problem of finding an optimal allocation

$$\hat{x} \in K \;\&\; \sum_{i=1}^{n} V_i(\hat{x}_i) = \inf_{x\in K} \sum_{i=1}^{n} V_i(\hat{x}_i)$$

where

$$K := \left\{ x := (x_1, \ldots, x_n) \in \prod_{i=1}^{n} L_i \;\middle|\; \sum_{i=1}^{n} x_i \in M \right\}$$

can be embedded in the problem of the form

$$v := \inf_{x\in X} [V(x) + W(Ax)].$$

where $X := Y^n$,

$$V(x) := \sum_{i=1}^{n} V_i|_{L_i}(x), \;\; W(y) := \psi_M(y) \;\&\; Ax := \sum_{i=1}^{n} x_i$$

which we shall now study in this simpler and more general framework.

## 3.4 Fenchel's Theorem

Suppose we have two finite dimensional vector-spaces $X$ and $Y$, together with

$$\begin{cases} i) & \text{a continuous, linear operator } A \in L(X,Y) \\ ii) & \text{two nontrivial, convex, lower semi-continuous functions} \\ & V : X \to \mathbf{R} \cup \{+\infty\} \text{ and } W : X \to \mathbf{R} \cup \{+\infty\} \end{cases}$$

We shall study the minimization problem

$$v := \inf_{x\in X} [V(x) + W(Ax)]. \tag{3.7}$$

Note that the function $V + W \circ A$ which we propose to minimize is only nontrivial if $A\,\mathrm{Dom}V \cap \mathrm{Dom}\,W \neq \emptyset$, that is to say, if

$$0 \in A\,(\mathrm{Dom}\,V) - \mathrm{Dom}\,W. \tag{3.8}$$

In this case, we have $v < +\infty$.

Now we introduce the **dual minimization problem**

$$v^* := \inf_{q\in Y^*} [V^*(-A^*q) + W^*(q)] \tag{3.9}$$

where $A^* \in L(Y^*, X^*)$ is the transpose of $A$, $V^* : X^* \to \mathbf{R} \cup \{+\infty\}$ is the conjugate of $V$ and $W^* : Y^* \to \mathbf{R} \cup \{+\infty\}$ is the conjugate of $W$. This only makes sense if we assume that

$$0 \in A^* \operatorname{Dom} W^* + \operatorname{Dom} V^* \tag{3.10}$$

and in this case, $v^* < +\infty$.

Note that we still have the inequality

$$v + v^* \geq 0 \tag{3.11}$$

since, by virtue of Fenchel's inequality,

$$V(x) + W(Ax) + V^*(-Aq) + W^*(q) \geq \langle -A^*q, x \rangle + \langle q, Ax \rangle = 0.$$

Consequently, conditions (3.8) and (3.10) imply that $v$ and $v^*$ are finite.

**Theorem 3.17** *Suppose that $A \in L(X,Y)$ is a linear operator from $X$ to $Y$ and that $V : X \to \mathbf{R} \cup \{+\infty\}$ and $W : Y \to \mathbf{R} \cup \{+\infty\}$ are nontrivial, convex, lower semi-continuous functions. We consider the case in which $0 \in A(\operatorname{Dom} V) - \operatorname{Dom} W$ and $0 \in A^*(\operatorname{Dom} W^*) + \operatorname{Dom} V^*$ (which is equivalent to the assumption that $v$ and $v^*$ are finite).*

*If we suppose that*

$$0 \in \operatorname{Int}(A^* \operatorname{Dom} W^* + \operatorname{Dom} V^*), \tag{3.12}$$

*then*

$$\begin{array}{ll} i) & v + v^* = 0 \\ ii) & \exists\, \bar{x} \in X \ \textit{ such that } \ V(\bar{x}) + W(A\bar{x}) = v. \end{array} \tag{3.13}$$

*If we suppose that*

$$0 \in \operatorname{Int}(A \operatorname{Dom} V - \operatorname{Dom} W), \tag{3.14}$$

*then*

$$\begin{array}{ll} i) & v + v^* = 0 \\ ii) & \exists\, \bar{q} \in Y^* \ \text{ such that } \ V^*(-A^*\bar{q}) + W^*(\bar{q}) = v^* \end{array} \tag{3.15}$$

**Proof —** We introduce the map $\Phi$ from $\mathrm{Dom}\, V \times \mathrm{Dom}\, W$ to $Y \times \mathbf{R}$ defined by

$$\Phi(x,y) = \{Ax - y, V(x) + W(y)\} \tag{3.16}$$

together with

$$\begin{array}{ll} i) & \text{the vector } (0, -v^\star) \in Y \times \mathbf{R} \\ ii) & \text{the cone } Q = \{0\} \times ]0, \infty[ \subset Y \times \mathbf{R} \end{array} \tag{3.17}$$

It is easy to show that the linearity of $A$ and the convexity of the functions $V$ and $W$ imply that

$$\mathcal{K} \ := \ \Phi(\mathrm{Dom}\, V \times \mathrm{Dom}\, W) + Q \ \text{ is a convex subset of } Y \times \mathbf{R}. \tag{3.18}$$

It is enough to prove that

$$(0, -v^\star) \in \Phi(\mathrm{Dom}\, V \times \mathrm{Dom}\, W) + Q. \tag{3.19}$$

because this inclusion implies the existence of a pair $(\overline{x}, \overline{y})$ satisfying $A\overline{x} = \overline{y}$ and

$$-v^\star \ \geq \ V(\overline{x}) + W(\overline{y}) \ = \ V(\overline{x}) + W(A\overline{x}) \ \geq \ v \ \geq \ -v^\star$$

so that $\overline{x}$ is a solution to our problem.

Assume for the time that $\mathcal{K}$ is also closed. Then we infer that $(0, -v^\star) \in \mathcal{K}$.

If not, the Separation Theorem implies the existence of a continuous linear form $(p, a) \in X^* \times \mathbf{R}$ such that

$$\left\{ \begin{array}{l} \sup_{\substack{x \in \mathrm{Dom} V \\ y \in \mathrm{Dom} W}} [-a(V(x) + W(y)) + \langle -p, Ax - y \rangle] + \sup_{\theta > 0}(-a\theta) \\ \leq \ \langle (-p, -a), (0, -v^\star) \rangle - \varepsilon \ = \ av^\star - \varepsilon \end{array} \right. \tag{3.20}$$

Since the number $\sup_{\theta > 0}(-a\theta)$ is bounded above, we deduce that it is zero and that $a$ is positive or zero. We cannot have $a = 0$, since in that case, the inequality (3.20)ii) would imply the contradiction

$$0 \ = \ \langle p, 0 \rangle \ \leq \ \sup_{\substack{x \in \mathrm{Dom} V \\ y \in \mathrm{Dom} W}} \langle p, Ax - y \rangle \ \leq \ -\varepsilon \tag{3.21}$$

since $0 \in A\, \mathrm{Dom} V - \mathrm{Dom}\, W$.

Consequently, $a$ is strictly positive. Dividing the inequality (3.20)ii) by $a$ and taking $\bar{p} = p/a$, we obtain

$$\left\{ \begin{array}{l} \sup_{\substack{x \in \mathrm{Dom} V \\ y \in \mathrm{Dom} W}} [(\langle -A^\star \bar{p}, x \rangle - V(x)) + + \langle \bar{p}, y \rangle - W(y))] + 0 \\ = \ V^\star(-A^\star \bar{p}) + W^\star(\bar{p}) \ \leq \ v^\star - \frac{\varepsilon}{a} \end{array} \right.$$

which implies the contradiction $v^\star \leq v^\star - \dfrac{\varepsilon}{a}$.

It remains to prove that $\mathcal{K}$ is closed. For that purpose, we consider sequences $x_n \in \text{Dom}(V)$ and $y_n \in \text{Dom}(W)$ such that $v_n \geq V(x_n) + W(y_n)$ converges to some $v$ and $z_n := Ax_n - y_n$ converges to some $z$.

The idea is to deduce from assumption (3.12) that the sequence $x_n$ is bounded, because, in this case, it will remain in a compact subset and we will be able to extract a converging subsequence.

So, by assumption, there exists $\eta > 0$ such that

$$\eta B \subset A^\star \text{Dom}(W^\star) + \text{Dom}(V^\star)$$

so that we can associate with any $p \in X^\star$ elements $q \in \text{Dom}(W^\star)$and $r \in \text{Dom}(V^\star)$ such that $\eta \frac{p}{\|p\|} = A^\star q + r$.

Hence,

$$\begin{cases} \left\langle \eta \frac{p}{\|p\|}, x_n \right\rangle = \langle q, Ax_n \rangle + \langle r, x_n \rangle = \langle q, z_n \rangle + \langle q, y_n \rangle + \langle r, x_n \rangle \\ \leq \langle q, z_n \rangle + V(x_n) + V^\star(r) + W(y_n) + W^\star(q) \\ \leq \langle q, z_n \rangle + v_n + V^\star(r) + W^\star(q) < +\infty \end{cases}$$

Therefore, a subsequence (again denoted by) $x_n$ converges to some $x$ and $y_n := z_n - Ax_n$ converges to $z - Ax$. Since $V$ and $W$ are lower semicontinuous, we infer that

$$V(x) + W(y) \leq v \;\&\; z = Ax - y$$

which shows that the limit $(z, v)$ belongs to $\mathcal{K}$. □

**Remark —** This proof shows that the Fenchel Theorem remains true when $X$ is a reflexive Banach space (suplied with the weak topology) and $Y$ is a Banach space. Indeed, we proved that the convex set $\mathcal{K}$ is closed, i.e., weakly closed because

1. the sequence $x_n$ is weakly bounded, and thus, weakly compact, so that a subsequence (again denoted by) $x_n$ converges weakly to some $x$
2. every lower semicontinuous convex function is weakly lower semicontinuous. □

Actually, in finite dimensional vector-space, we do not really need to assume that the functions $V$ and $W$ are lower semicontinuous.

**Corollary 3.18** *Let $L \subset X$ and $M \subset Y$ to closed convex subsets and $A \in \mathcal{L}(X, Y)$ a linear operator linked by the* constraint qualification condition

$$0 \in \text{Int}(AL - M)$$

*Then the normal cone to* $L \cap A^{-1}(M)$

$$N_{L\cap A^{-1}(M)}(x) \;=\; N_L(x) + A^\star N_M(Ax)$$

*and the tangent cone by*

$$T_{L\cap A^{-1}(M)}(x) \;=\; T_L(x) \cap A^{-1} T_M(Ax)$$

**Proof** — Since $\psi_{L\cap A^{-1}(M)}(x) = \psi_L(x) + \psi_M(Ax)$ and since

$$N_{L\cap A^{-1}(M)}(x) = \partial\psi_{L\cap A^{-1}(M)}(x)$$

we deduce the formula for the normal cones. The one for tangent cones is obtained by polarity and transposition. □

**Remark** — Without the constraint qualification condition, the above Corollary can be false. Take for instance $X = Y := \mathbf{R}^2$, $A = \mathbf{1}$ and two balls $L$ and $M$ tangent at a point $x$. The tangent cone to the intersection $\{x\}$ is reduced to $\{0\}$, whereas the intersection of the tangent cones is a hyperplane. □

## 3.5 Properties of Conjugate Functions

Firstly, we note the following elementary propositions.

**Proposition 3.19** *a) If* $V \leq W$, *then* $W^\star \leq V^\star$.
*b) If* $A \in L(X,X)$ *is an isomorphism, then*

$$(V \circ A)^\star = V^\star \circ A^{\star -1}.$$

*c) If* $W(x) := V(x - x_0) + \langle p_0, x\rangle + a$, *then*

$$W^\star(p) = V^\star(p - p_0) + \langle p, x_0\rangle - (a + \langle p_0, x_0\rangle)$$

*d) If* $W(x) := V(\lambda x)$, *then* $W^\star(p) = V^\star\left(\frac{p}{\lambda}\right)$ *and if* $U(x) := \lambda V(x)$, *then* $Z^\star(p) = \lambda V^\star\left(\frac{p}{\lambda}\right)$

**Proof** — The first assertion is evident. The second assertion may be proved by showing that

$$\sup_{x\in X}[\langle p,x\rangle - V(Ax)] = \sup_{y\in X}[\langle A^{*-1}p,y\rangle - V(y)] = V^*(A^{*-1}p).$$

For the third assertion, we observe that

$$\begin{aligned}\sup_{x\in X}[\langle p,x\rangle - W(x)] &= \sup_{x\in X}[\langle p-p_0,x\rangle - V(x-x_0)] - a\\ &= \sup_{x\in X}[\langle p-p_0,y\rangle - V(y)] - a + \langle p-p_0,x_0\rangle\\ &= V^*(p-p_0) + \langle p,x_0\rangle - a - \langle p_0,x_0\rangle \ \square\end{aligned}$$

**Proposition 3.20** *Suppose that $X$ and $Y$ are two finite dimensional vector-spaces and that $V$ is a nontrivial, convex function from $X\times Y$ to $\mathbf{R}\cup\{+\infty\}$. Set $W(y) := \inf_{x\in X} V(x,y)$. Then*

$$W^*(q) = V^*(0,q) \tag{3.22}$$

**Proof**

$$\begin{aligned}W^*(q) &= \sup_{y\in Y}[\langle q,y\rangle - \inf_{x\in X} V(x,y)]\\ &= \sup_{y\in Y}\sup_{x\in X}[\langle 0,x\rangle + \langle q,y\rangle - V(x,y)] = V^*(0,q) \ \square\end{aligned}$$

**Proposition 3.4.** *Suppose that $A \in L(X,Y)$ is a linear operator and that $V : X \to \mathbf{R}\cup\{+\infty\}$ and $W : Y \to \mathbf{R}\cup\{+\infty\}$ are two nontrivial, lower semi-continuous functions. Suppose further that*

$$0 \in \mathrm{Int}\,(A\,\mathrm{Dom}\,V - \mathrm{Dom}\,W). \tag{3.23}$$

*Then, for all $p \in A^*\,\mathrm{Dom}\,W^* + \mathrm{Dom}\,V^*$, there exists $\bar{q} \in Y^*$ such that*

$$\begin{aligned}(V + W\circ A)^*(p) &= V^*(p - A^*\bar{q}) + W^*(\bar{q})\\ &= \inf_{q\in Y^*}(V^*(p - A^*q) + W^*(q))\end{aligned} \tag{3.24}$$

**Proof** — We may write

$$\sup_{x\in X}[\langle p,x\rangle - V(x) - W(Ax)] = -\inf[V(x) - \langle p,x\rangle + W(Ax)]$$

We apply Fenchel's theorem with $V$ replaced by $V(\cdot) - \langle p, \cdot\rangle$, the domain of which coincides with that of $V$ and the conjugate function of which is equal to $q \to V^*(q + p)$. Thus, there exists $\bar{q} \in Y^*$ such that

$$\begin{aligned}\sup_{x\in X}[\langle p, x\rangle - V(x) - W(Ax)] &= V^*(p - A^*\bar{q}) + W^*(\bar{q}) \\ &= \inf_{q\in Y}[V^*(p - A^*q) + W^*(q)] \quad \square\end{aligned}$$

It is useful to state the following consequence explicitly:

**Proposition 3.5.** *Suppose that $A \in L(X, Y)$ is a linear operator from $X$ to $Y$ and that $W : Y \to \mathbf{R} \cup \{+\infty\}$ is a nontrivial, convex, lower semi-continuous function. We suppose further that*

$$0 \in \operatorname{Int}(\operatorname{Im} A - \operatorname{Dom} W) \tag{3.25}$$

*Then, for all $p \in A^* \operatorname{Dom} W^*$, there exists $\bar{q} \in \operatorname{Dom} W^*$ satisfying*

$$A^*\bar{q} = p \quad \text{and} \quad (W \circ A)^*(p) = W^*(\bar{q}) = \min_{A^*q=p} W^*(q)$$

**Proof**— We apply the previous proposition with $V = 0$, where the domain is the whole space $X$. Its conjugate function $V^*$ is defined by $V^*(p) = \{0\}$ if $p = 0$ and $V^*(p) = +\infty$ otherwise. Consequently, $V^*(p - A^*q)$ is finite (and equal to 0) if and only if $p = A^*q$. $\square$

## 3.6 Subdifferential Calculus

We can deduce easily from the calculus of conjugate functions a subdifferential calculus.

**Theorem 3.21** *We consider a linear operator $A \in L(X, Y)$ and two non-trivial, convex, lower semi-continuous functions $V : X \to \mathbf{R} \cup \{+\infty\}$ and $W : Y \to \mathbf{R} \cup \{+\infty\}$.*

*We assume further that*

$$0 \in \operatorname{Int}(A \operatorname{Dom} V - \operatorname{Dom} W). \tag{3.26}$$

*Then,*

$$\partial(V + W \circ A)(x) = \partial V(x) + A^*\partial W(Ax) \tag{3.27}$$

**Proof** — It is easy to check that $\partial V(x) + A^*\partial W(Ax)$ is always contained in $\partial(V + W \circ A)(x)$. The inverse inclusion follows from Proposition 3.4. We take $p \in (\partial(V + W \circ A)(x)$. There exists $\bar{q} \in Y^*$ such that $(V + W \circ A)^*(p) = V^*(p - A^*\bar{q}) + W^*(\bar{q})$. Thus, from equation (3.24),

$$\begin{array}{rcl} \langle p, x \rangle & = & V(x) + W(Ax) + (V + W \circ A)^*(p) \\ & = & (V(x) + V^*(p - A^*\bar{q})) + (W(Ax) + W^*(\bar{q})). \end{array}$$

Consequently,

$$0 = (\langle p - A^*\bar{q}, x \rangle - V(x) - V^*(p - A^*\bar{q})) + (\langle \bar{q}, Ax \rangle - W(Ax) - W^*(\bar{q})).$$

Since each of these two expressions is negative or zero, it follows that they are both zero, whence that $\bar{q} \in \partial W(Ax)$ and $p - A^*\bar{q} \in \partial V(x)$. Thus, we have shown that $p = p - A^*\bar{q} + A^*\bar{q} \in \partial V(x) + A^*\partial W(Ax)$. □

**Corollary 3.22** *If $V$ and $W$ are two nontrivial, convex, lower semi-continuous functions from $X$ to $\mathbf{R} \cup \{+\infty\}$ and if*

$$0 \in \mathrm{Int}\,(\mathrm{Dom}\,V - \mathrm{Dom}\,W) \tag{3.28}$$

*then*

$$\partial(V + W)(x) = \partial V(x) + \partial W(x). \tag{3.29}$$

*If $W$ is a nontrivial, convex, lower semi-continuous function from $Y$ to $\mathbf{R} \cup \{+\infty\}$ and if $A \in L(X, Y)$ satisfies*

$$0 \in \mathrm{Int}\,(\mathrm{Im}\,A - \mathrm{Dom}\,W) \tag{3.30}$$

*then*

$$\partial(W \circ A)(x) = A^*\partial W(Ax). \tag{3.31}$$

**Proposition 3.23** *Let $W$ be a nontrivial, convex function from $X \times Y$ to $\mathbf{R} \cup \{+\infty\}$. Consider the function $U : Y \to \mathbf{R} \cup \{+\infty\}$ defined by*

$$U(y) := \inf_{x \in X} W(x, y). \tag{3.32}$$

*If $\bar{x} \in X$ satisfies $U(y) = W(\bar{x}, y)$, then the following conditions are equivalent:*

$$\begin{array}{ll} (a) & q \in \partial U(y) \\ (b) & (0, q) \in \partial W(\bar{x}, y) \end{array} \tag{3.33}$$

**Proof** — Since $U^*(q) = W^*(0,q)$, following Proposition 3.20, we deduce that $q$ belongs to $\partial U(y)$ if and only if $\langle q,y\rangle = U(y) + h^*(q) = W(\bar{x},y) + W^*(0,q)$, that is, if and only if $(0,q) \in \partial W(\bar{x},y)$. □

**Proposition 3.24** *We consider a family of convex functions $x \to V(x,p)$ indexed by a parameter $p$ running over a set $P$. We assume that*

$$\left\{\begin{array}{ll} i) & P \text{ is compact} \\ ii) & \text{There exists a neighborhood } \mathcal{N} \text{ of } x \text{ such that,} \\ & \text{for all } y \text{ in } \mathcal{N},\ p \to V(y,p) \text{ is upper semi-continuous.} \\ iii) & \forall p \in P, y \to V(y,p) \text{ is continuous at } x. \end{array}\right. \tag{3.34}$$

*Consider the upper envelope $U$ of the functions $V(\cdot,p)$, defined by $U(y) = \sup_{p\in P} V(y,p)$. Set*

$$P(x) := \{p \in P \mid U(x) = V(x,p)\}. \tag{3.35}$$

*Then*

$$DU(x)(v) = \sup_{p\in P(x)} DV(x,p)(v) \tag{3.36}$$

*and*

$$\partial U(x) = \overline{\mathrm{co}}\left(\bigcup_{p\in P(x)} \partial V(x,p)\right) \tag{3.37}$$

**Proof** — Since when $p$ belongs to $P(x)$, we may write

$$\frac{V(x+hv,p) - V(x,p)}{h} \leq \frac{U(x+hv) - U(x)}{h},$$

letting $h$ tend to 0 we obtain

$$\sup_{p\in P(x)} DV(x,p)(v) \leq DU(x)(v). \tag{3.38}$$

We must establish the inverse inequality. Fix $\varepsilon > 0$; we shall show that there exists $p \in P(x)$ such that $DU(x)(v) - \varepsilon \leq DV(x,p)(v)$. Since the function $U$ is convex, we know that

$$DU(x)(v) = \inf_{h>0} \frac{U(x+hv) - U(x)}{h}$$

Then, for all $h > 0$, the set

$$B_h := \left\{p \in P \,\middle|\, \frac{V(x+hv,p) - U(x)}{h} \geq DU(x)(v) - \varepsilon\right\} \tag{3.39}$$

is non-empty. Consider the neighborhood $\mathcal{N}$ mentioned in assumption (3.34)ii). There exists $h_0 > 0$ such that $x + hv$ belongs to $\mathcal{N}$ for all $h \leq h_0$. Since $p \to V(x + hv, p)$ is upper semi-continuous, the set $B_h$ is closed. On the other hand, if $h_1 \leq h_2$, then $B_{h_1} \subset B_{h_2}$; if $p$ belongs to $B_{h_1}$, the convexity of $V$ with respect to $x$ implies that

$$\begin{cases} DU(x)(v) - \varepsilon \\ \leq \frac{1}{h_1}\left[\left(1 - \frac{h_1}{h_2}\right)(V(x,p) - U(x)) + \frac{h_1}{h_2}(V(x + h_2 v, p) - U(x))\right] \\ \leq \frac{1}{h_2}(V(x + h_2 v, p) - U(x)) \end{cases}$$

since $x + h_1 v = \left(1 - \dfrac{h_1}{h_2}\right)x + \dfrac{h_1}{h_2}(x + h_2 v)$ and since $V(x,p) - U(x) \leq 0$ for all $p$. Consequently, since $P$ is compact, the intersection $\bigcap_{0<h\leq h_0} B_h$ is non-empty and all elements $\bar{p}$ of this intersection satisfy

$$h(DU(x)(v) - \varepsilon) \leq V(x + hv, \bar{p}) - U(x). \tag{3.40}$$

Letting $h$ tend to 0, we deduce that $V(x,\bar{p}) - U(x) \geq 0$, whence $\bar{p}$ belongs to $P(x)$. Dividing by $h > 0$, we obtain the inequality

$$DU(x)(v) - \varepsilon \leq DV(x,\bar{p})(v) \leq \sup_{p\in P(x)} DV(x,p)(v)$$

Thus, it is sufficient to let $\varepsilon$ tend to 0.

Since $y \to V(y,p)$ is continuous at $x$, we know that $DV(x,p)(\cdot)$ is continuous for each $p$, whence that $DU(x)(\cdot)$ is lower semi-continuous. Equation (3.36) may be written as

$$\sigma(\partial U(x,v)) = \sup_{p\in P(x)} \sigma(\partial V(x,p), v)$$

which implies equation (3.37) □

**Corollary 3.25** *Consider $n$ convex functions $V_i$ continuous at a point $x$. Then*

$$\partial\left(\sup_{i=1,\ldots,n} V_i\right)(x) = \overline{\text{co}}\left(\bigcup_{i\in I(x)} \partial V_i(x)\right) \tag{3.41}$$

*where* $I(x) = \{i = 1, \ldots, n | V_i(x) = \sup_{j=1,\ldots,n} V_j(x)\}$.

## 3.7 Moreau-Yosida Approximations

Consider a nontrivial, convex, lower semi-continuous function $V$ from a Hilbert space $X$ to $\mathbf{R} \cup \{+\infty\}$. With any $\lambda > 0$ we associate the function $V_\lambda$ defined by

$$V_\lambda(x) := \inf_{y\in X}\left[V(y) + \frac{1}{2\lambda}\|y - x\|^2\right] \tag{3.42}$$

We shall show that the functions $V_\lambda$ are convex, differentiable functions which are simply convergent to the function $V$ as $\lambda$ tends to 0. This provides us with a **regularization** procedure which enables us to approximate $V$ by a more regular function.

**Theorem 2.2.** *Suppose that $V : X \to \mathbf{R}\cup\{+\infty\}$ is a nontrivial, convex, lower semi-continuous function from a Hilbert space $X$ to $\mathbf{R}\cup\{+\infty\}$. There exists a* **unique** *solution (denoted by $J_\lambda(x)$) of the problem $V^\lambda(x)$:*

$$V_\lambda(x) = V(J_\lambda x) + \frac{1}{2\lambda}\|J_\lambda x - x\|^2.$$

By applying it to the case where $V = \psi_K$ is the indicator function of a set $K$, we obtain the projection theorem, since in this case

$$V_\lambda(x) = \frac{1}{2\lambda}d(x,K)^2$$

where $d(x,K) := \inf_{y\in K}\|x-y\|$ is the distance from $x$ to $K$.

**Proof**

Since the Cauchy–Schwartz inequality implies that

$$\langle p, x-y\rangle \le \frac{1}{\lambda}\|\lambda p\|\,\|x-y\| \le \frac{\lambda}{2}\|p\|^2 + \frac{1}{2\lambda}\|y-x\|^2,$$

this inequality implies that

$$\begin{aligned} V(y) + \frac{1}{2\lambda}\|y-x\|^2 &\ge \langle p, y-x\rangle + a + \langle p, x\rangle + \frac{1}{2\lambda}\|y-x\|^2 \\ &\ge a + \langle p, x\rangle - \frac{\lambda}{2}\|p\|^2 \end{aligned}$$

and thus that

$$V_\lambda(x) \ge a + \langle p, x\rangle - \frac{\lambda}{2}\|p\|^2 > -\infty.$$

b) There exists a solution $\bar{x}$ of the problem $V_\lambda(x)$. To prove this, we consider a minimising sequence of elements $x_n \in X$ satisfying

$$V(x_n) + \frac{1}{2\lambda}\|x_n - x\|^2 \le V_\lambda(x) + \frac{1}{n}.$$

We shall show that this is a Cauchy sequence. In fact, the so-called median formula implies that

$$\|x_n - x_m\|^2 = 2\|x_n - x\|^2 + 2\|x_m - x\|^2 - 4\left\|\frac{x_n + x_m}{2} - x\right\|^2.$$

Consequently, we have

$$\begin{aligned}
\|x_n - x_m\|^2 &\leq 4\lambda\left(\frac{1}{n} + \frac{1}{m} + 2V_\lambda(x) - V(x_m) - V(x_n)\right) \\
&\quad +8\lambda\left(V\left(\frac{x_n + x_m}{2}\right) - V_\lambda(x)\right) \\
&= 4\lambda\left(\frac{1}{n} + \frac{1}{m} + 2V\left(\frac{x_n + x_m}{2}\right) - V(x_n) - V(x_m)\right) \\
&\leq 4\lambda\left(\frac{1}{n} + \frac{1}{m}\right)
\end{aligned}$$

since $V$ is convex.

Thus, $x_n$ converges to an element $\bar{x}$ of $X$, since $X$ is complete.

The lower semi-continuity of $V$ implies that

$$\begin{aligned}
V(\bar{x}) + \frac{1}{2\lambda}\|\bar{x} - x\|^2 &\leq \liminf_{x_n \to \bar{x}}\left(V(x_n) + \frac{1}{2\lambda}\|x_n - \bar{x}\|^2\right) \\
&\leq V_\lambda(x).
\end{aligned}$$

Whence $V_\lambda(x) = V(\bar{x}) + \frac{1}{2\lambda}\|\bar{x} - x\|^2$.

Since the Hilbertian norm is strictly convex, this solution is unique. □

The Fermat rule and the subdifferential calculus imply that

$$x \in J_\lambda x + \lambda\partial V(J_\lambda x) = (1 + \lambda\partial V(\cdot))(J_\lambda x) \tag{3.43}$$

which can be written

$$\forall y \in X, \qquad \frac{1}{\lambda}\langle J_\lambda x - x, J_\lambda x - y\rangle + V(J_\lambda x) - V(y) \leq 0 \tag{3.44}$$

We set:

$$A_\lambda(x) := \frac{1}{\lambda}(x - J_\lambda x) \in \partial V(J_\lambda x)$$

Thus, $J_\lambda$ **is the inverse of the set-valued map** $1 + \lambda\partial V(\cdot)$. The map $A_\lambda$ is called the **Moreau-Yosida approximation** of the subdifferential $\partial V(\cdot)$.

We note that the maps $J_\lambda$ and $1 - J_\lambda$ are both continuous, indeed Lipschitz with constant 1.

**Proposition 3.26** *The maps $J_\lambda$ and $1 - J_\lambda$ are Lipschitz with constant 1 (independent of $\lambda$) and "monotone":*

$$\left\{\begin{array}{ll} i) & \langle J_\lambda x - J_\lambda y, x - y\rangle \geq \|J_\lambda x - J_\lambda y\|^2 \\ ii) & \langle (1 - J_\lambda)x - (1 - J_\lambda)y, x - y\rangle \geq \|(1 - J_\lambda)x - (1 - J_\lambda)y\|^2 \end{array}\right. \tag{3.45}$$

**Proof** — The variational inequality which characterizes $J_\lambda x$ implies that

$$V(J_\lambda x) - V(J_\lambda y) + \frac{1}{\lambda}\langle J_\lambda x - x, J_\lambda x - J_\lambda y\rangle \leq 0$$

Switching the roles of $x$ and $y$, we have

$$V(J_\lambda y) - V(J_\lambda x) + \frac{1}{\lambda}\langle J_\lambda y - y, J_\lambda y - J_\lambda x\rangle \leq 0.$$

Adding these two inequalities, we find that

$$\langle J_\lambda x - J_\lambda y - (x - y), J_\lambda x - J_\lambda y\rangle \leq 0 \tag{3.46}$$

The inequalities (3.45)i) and ii) follow from this inequality.

This being so, we write

$$\begin{array}{rcl} \|x - y\|^2 & = & \|x - J_\lambda x - (y - J_\lambda y) + (J_\lambda x - J_\lambda y)\|^2 \\ & = & \|(1 - J_\lambda)x - (1 - J_\lambda)y\|^2 + \|J_\lambda x - J_\lambda y\|^2 \\ & & +2\langle (1 - J_\lambda)x - (1 - J_\lambda)y, J_\lambda x - J_\lambda y\rangle \end{array}$$

Following (3.46), we deduce that

$$\|x - y\|^2 \geq \|(1 - J_\lambda)x - (1 - J_\lambda)y\|^2 + \|J_\lambda x - J_\lambda y\|^2$$

This completes the proof. □

**Theorem 3.27** *Suppose that $V : X \to \mathbf{R} \cup \{+\infty\}$ is a nontrivial, convex, lower semi-continuous function. Then the functions $V_\lambda$ are convex and differentiable and*

$$A_\lambda(x) = \nabla V_\lambda(x) \tag{3.47}$$

*Moreover, when $\lambda$ tends to 0,*

$$\forall x \in \mathrm{Dom}\, V,\ \ V_\lambda(x) \to V(x) \ \text{ and } \ J_\lambda x \to x \tag{3.48}$$

**Proof**

a) For $x$ belonging to the domain of $V$, we shall show that $J_\lambda x$ converges to $x$. We take any $p$ in the domain of $V^*$ (which is non-empty). Since

$$\frac{1}{2\lambda}\|J_\lambda x - x\|^2 + V(J_\lambda x) = V_\lambda(x) \leq V(x)$$

and since

$$-V(J_\lambda x) \leq V^*(p) - \langle p, J_\lambda x\rangle$$

we deduce that

$$\begin{aligned}\frac{1}{2\lambda}\|J_\lambda x - x\|^2 &\leq V(x) + V^*(p) - \langle p, x\rangle + \langle p, x - J_\lambda x\rangle \\ &\leq \frac{1}{4\lambda}\|J_\lambda x - x\|^2 + V(x) + V^*(p) - \langle p, x\rangle + \lambda\|p\|^2\end{aligned}$$

(since $ab \leq a^2/4\lambda + b^2\lambda$). Thus, since $\lambda$ converges to 0,

$$\|J_\lambda x - x\|^2 \leq 4\lambda(V(x) + V^*(p) - \langle p, x\rangle + \lambda\|p\|^2) \to 0$$

b) Moreover, $V_\lambda(x) \leq V(x) + \frac{1}{2\lambda}\|x - x\|^2 = V(x)$. Since $V(x) \leq \liminf_{\lambda\to 0} V(J_\lambda x)$ (because $V$ is lower semi-continuous) and since

$$V(J_\lambda x) = V_\lambda(x) - \frac{1}{2\lambda}\|J_\lambda x - x\|^2 \leq V_\lambda(x),$$

it follows that $V(x) \leq \liminf_{\lambda\to 0} V_\lambda(x)$. Thus, $V(x) = \lim_{\lambda\to 0} V_\lambda(x)$.

c) We observed that $A_\lambda(x)$ belongs to $\partial V(J_\lambda x)$. Thus,

$$\begin{aligned}V_\lambda(x) - V_\lambda(y) &= V(J_\lambda x) - V(J_\lambda y) + \frac{\lambda}{2}\|A_\lambda(x)\|^2 - \frac{\lambda}{2}\|A_\lambda(y)\|^2 \\ &\leq \langle A_\lambda(x), J_\lambda x - J_\lambda y\rangle + \frac{\lambda}{2}\|A_\lambda(x)\|^2 - \frac{\lambda}{2}\|A_\lambda(y)\|^2\end{aligned}$$

(because $A_\lambda(x) \in \partial V(J_\lambda x)$)

$$\leq \langle A_\lambda(x), x - y\rangle - \lambda\langle A_\lambda(x), A_\lambda(x) - A_\lambda(y)\rangle + \frac{\lambda}{2}\|A_\lambda(x)\|^2 - \frac{\lambda}{2}\|A_\lambda(y)\|^2$$

(because $J_\lambda = 1 - \lambda A_\lambda$)

$$\begin{aligned}
&= \langle A_\lambda(x), x-y\rangle - \lambda\left(\frac{1}{2}\|A_\lambda(x)\|^2 + \frac{1}{2}\|A_\lambda(y)\|^2 - \langle A_\lambda(x), A_\lambda(y)\rangle\right)\\
&= \langle A_\lambda(x), x-y\rangle - \frac{\lambda}{2}\|A_\lambda(x) - A_\lambda(y)\|^2\\
&\leq \langle A_\lambda(x), x-y\rangle
\end{aligned}$$

Thus, we have shown that

$$A_\lambda(x) \in \partial V_\lambda(x). \tag{3.49}$$

Moreover, since $A_\lambda(y)$ belongs to $\partial V_\lambda(y)$ for all $y$, we obtain the inequalities

$$\begin{aligned}
V_\lambda(x) - V_\lambda(y) &\geq \langle A_\lambda(y), x-y\rangle\\
&= \langle A_\lambda(x), x-y\rangle + \langle A_\lambda(y) - A_\lambda(x), x-y\rangle\\
&\geq \langle A_\lambda(x), x-y\rangle - \|A_\lambda(y) - A_\lambda(x)\|\|x-y\|\\
&\geq \langle A_\lambda(x), x-y\rangle - \frac{1}{\lambda}\|x-y\|^2
\end{aligned}$$

since $\|A_\lambda(x) - A_\lambda(y)\| \leq \frac{1}{\lambda}\|x-y\|$ (see Proposition 3.26, since $A_\lambda = \frac{1}{\lambda}(1 - J_\lambda)$). Thus,

$$\left|\frac{V_\lambda(x) - V_\lambda(y) - \langle A_\lambda(x), x-y\rangle}{\|x-y\|}\right| \leq \frac{1}{\lambda}\|x-y\|,$$

whence $A_\lambda(x) = \nabla V_\lambda(x)$.

□

**Corollary 3.28** *Let $V : \mathbf{R} \cup \{+\infty\}$ be a nontrivial, convex lower semicontinuous function. Then $V$ is subdifferentiable on a dense subset of the domain of $V$.*

**Proposition 3.29** *Let $V_0'(x))$ denote the element of $\partial V(x)$ with the smallest norm. We also have*

$$\forall\, x \in \mathrm{Dom}(\partial V),\ \ \|A_\lambda(x) - V_0'(x))\|^2 \ \leq\ \|V_0'(x))\|^2 - \|A_\lambda(x)\|^2$$

*and for all $x \in \mathrm{Dom}(\partial V)$,*

$$A_\lambda(x) \text{ converges to } V_0'(x)) \text{ when } \lambda \to 0+$$

**Proof**

1. — Let $x \in \text{Dom}(\partial V)$. Then

$$\begin{aligned} & \|A_\lambda(x) - V_0'(x))\|^2 \\ = \;& \|A_\lambda(x)\|^2 + \|V_0'(x))\|^2 - 2 < A_\lambda(x), V_0'(x)) > \\ = \;& \|V_0'(x))\|^2 - \|A_\lambda(x)\|^2 - 2 < A_\lambda(x), V_0'(x)) - A_\lambda(x) > \end{aligned}$$

Using that $\partial V$ is monotone, that $V_0'(x)) \in \partial V(x)$ and $A_\lambda(x) \in \partial V(J_\lambda(x))$, we obtain

$$< A_\lambda(x), V_0'(x)) - A_\lambda(x) >= \frac{1}{\lambda} < x - J_\lambda(x), V_0'(x)) - A_\lambda(x) > \geq 0$$

Therefore, we have proved inequality

$$\|A_\lambda(x) - V_0'(x))\|^2 \leq \|V_0'(x))\|^2 - \|A_\lambda(x)\|^2 \tag{3.50}$$

2. — We deduce that $y = A_\lambda(x)$ is a solution to the equation $y \in \partial V(x - \lambda y)$. Indeed, setting $z = x - \lambda y$, this equation becomes $x \in z + \lambda \partial V(z)$. Hence,

$$z = J_\lambda(x) \;\; \& \;\; y = (x - J_\lambda(x))/\lambda = A_\lambda(x)$$

This remark implies that

$$A_{\mu+\lambda}(x) = (A_\mu)_\lambda(x)$$

Indeed, $y = A_{\mu+\lambda}(x)$ is a solution to the equation $y \in \partial V(x - \lambda y - \mu y)$; then $y \in A_\mu(x - \lambda y)$. Applying again the preceding remark to the Yosida approximation $A_\mu$, which is maximal monotone, we deduce that $y = (A_\mu)_\lambda(x)$.

3. — Now we use inequality (3.50), replacing $\partial V$ by $A_\mu$. Since $V_0'{}_\mu(x)) = A_\mu(x)$, we obtain

$$\|A_{\mu+\lambda}(x) - A_\mu(x)\|^2 \leq \|A_\mu(x)\|^2 - \|A_{\lambda+\mu}(x)\|^2$$

Then the sequence $\|A_\mu(x)\|^2$ is monotone and bounded from above by $\|V_0'(x))\|^2$, so that it converges to some real number $\alpha$ when $\lambda \to 0+$. This implies that

$$\lim_{\lambda,\mu \to 0} \|A_{\mu+\lambda}(x) - A_\mu(x)\|^2 \leq \alpha - \alpha = 0$$

Hence, $A_\lambda(x)$ satisfies the Cauchy criterion and converges to some element $v$ in $X$. Since $A_\lambda(x) \in \partial V(J_\lambda(x))$ and the graph of $\partial V$ is closed, we deduce that $v \in \partial V(x)$. Also

$$\|v\| \;=\; \lim_{\lambda \to 0+} \|A_\lambda(x)\| \;\leq\; \|V_0'(x))\|$$

Since $\partial V(x)$ is closed and convex, the projection of zero onto $\partial V(x)$ is unique and consequently, $v = V_0'(x))$. Therefore, $A_\lambda(x)$ converges to $V_0'(x))$ for all $x \in \mathrm{Dom}(\partial V)$. □

**Remark** — When $\lambda$ tends to infinity, we may interpret the minimisation problem (3.42) as a **penalization** of the minimisation problem

$$-V^*(0) = \inf_{x \in X} V(x). \tag{3.51}$$

We observe that the Fenchel Theorem implies that

$$V_\lambda(x) + \inf_{p \in X^*} \left( V^*(p) - \langle p, x\rangle + \frac{\lambda}{2}\|p\|^2 \right) \;=\; 0.$$

that the minimisation problem $V_\lambda(x)$ has a solution denoted by $J_\lambda x$ and that its dual problem

$$\inf_{p \in X^*} \left( V^*(p) - \langle p, x\rangle + \frac{\lambda}{2}\|p\|^2 \right)$$

has also a solution dented by $A_\lambda(x)$.

When $\lambda \to \infty$,

$$V_\lambda(x) \;\text{ tends to }\; -V^*(0) = \inf_{x \in X} V(x). \tag{3.52}$$

From Fenchel's Theorem, we know that

$$V_\lambda(x) + \inf_{q \in X^*} \left( V^*(-q) + \frac{\lambda}{2}\|q\|^2 - \langle q, x\rangle \right) = 0.$$

In other words, we may write

$$V_\lambda(x) + (V^* - x)_{1/\lambda}(0) = 0. \tag{3.53}$$

Consequently, when $\lambda \to \infty$, $(V^* - x)_{1/\lambda}(0)$ tends to $(V^* - x)(0) = V^*(0) = -\inf_{y \in X} V(y)$ from the above.

Assume that 0 belongs to the domain of $\partial V^*$ (in other words, if there exists a minimum of $V$). Since $\nabla V_\lambda(x) = A_\lambda(x)$ is the unique solution of the problem $(V^* - x)_{1/\lambda}$, then $\nabla V_\lambda(x)$ converges to 0 as $\lambda$ tends to infinity. Consequently, if the limit of $J_\lambda(x)$ as $\lambda$ tends to infinity exists, it belongs to $\partial V^*(0)$, in other words, it achieves the minimum of $V$. □

## 4 Subgradient Differential Inclusion

We prove in this section the Existence Theorem 3.7 of a unique solution to the Cauchy problem of a subgradient differential inclusion.

For proving this theorem, we shall first approximate the lower semicontinuous convex function by its differentiable convex Moreau-Yosida approximation $V_\lambda$ defined on the whole space and prove that the solutions $x_\lambda(\cdot)$ of the **gradient equation**

$$x'_\lambda(t) = -\nabla V_\lambda(x_\lambda(t))$$

converge to a solution to the subgradient differential inclusion (4.3).

We first assume that a solution $x(\cdot)$ to the differential inclusion (4.3) exists and derive its properties. The solutions $x_\lambda(\cdot)$ to the approximate gradient differential equation enjoy naturally the same properties, which shall be used in the proof of the convergence.

**Lemma 4.1** *Assume that $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ is nontrivial, convex, lower semicontinuous and bounded below. Let $V'_0(x))$ denote the element of $\partial V(x)$ with the smallest norm.*

*Let $x(\cdot)$ and $y(\cdot)$ be two solutions of the differential inclusion (4.3) starting at $x_0$ and $y_0$ respectively. Then*

$$\sup_{t\geq 0} \|x(t) - y(t)\| \leq \|x_0 - y_0\|$$

*Therefore, from any initial state $x_0 \in \mathrm{Dom}(V)$ starts at most a* **unique solution** *to the differential inclusion (4.3) satisfying $t \mapsto \|x'(t)\|$ is not increasing.*

*If $x_\star$ achieves the minimum of $V$, then*

$$\forall\, t \geq s,\ \ \|x(t) - x_\star\| \leq \|x(s) - x_\star\| \tag{4.1}$$

**Proof—** Let $x(\cdot)$ and $y(\cdot)$ be two solutions of the differential inclusion (4.3) starting at $x_0$ and $y_0$ respectively. The monotonicity property of the subdifferential implies that

$$\frac{d}{dt}\|x(t) - y(t)\|^2 = 2\langle x'(t) - y'(t), x(t) - y(t)\rangle \leq 0$$

so that, by integrating this inequality, we obtain

$$\sup_{t\geq 0} \|x(t) - y(t)\| \leq \|x_0 - y_0\|$$

By taking $x_0 = y_0$, we infer the uniqueness of the solution. By taking $x(t) := x(t+s)$ and $y(t) := x(t+s+h)$, which are solutions to (4.3) starting at $x(s)$ and $x(s+h)$ respectively, we deduce that

$$\|x(t+s+h) - x(t+s)\| \;\leq\; \|x(s+h) - x(s)\|$$

Assume that $x(\cdot)$ is differentiable at $s$ and $t+s$. Then, dividing by $h > 0$ and letting $h$ converge to 0, we infer that

$$\forall\, t \geq 0, \;\; \|x'(t+s)\| \;\leq\; \|x'(s)\| \tag{4.2}$$

so tat the function $t \mapsto \|x'(t)\|$ is not increasing.

If $x_\star$ achieves the minimum of $V$, then the Fermat rule implies that $0 \in \partial V(x_\star)$ and we deduce from the above inequality that

$$\forall\, t \geq s, \;\; \|x(t) - x_\star\| \;\leq\; \|x(s) - x_\star\|$$

**Lemma 4.2** *Assume that $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ is nontrivial, convex, lower semicontinuous and bounded below. Then, for any initial state $x_0 \in \mathrm{Dom}(V)$, there exists a* unique solution *to the differential inclusion*

$$\text{for almost all } t \geq 0, \;\; x'(t) \;\in\; -\partial V(x) \tag{4.3}$$

*starting at $x_0$.*

**Proof** — Let us consider the sequence of solutions $x_\lambda(\cdot)$ to the gradient equation $x'_\lambda(t) = -V'_\lambda(x_\lambda(t))$ starting at $x_0$, which exist by the Cauchy-Lipschitz Theorem since $V'_\lambda =: A_\lambda$ is Lipschitz.

Let us set

$$\left\{\begin{array}{l} a(t) \;:=\; \frac{1}{2}\|x_\lambda(t) - x_\mu(t)\|^2 \;=\; \frac{1}{2}\int_0^t \frac{d}{d\tau}\|x_\lambda(\tau) - x_\mu(\tau)\|^2 d\tau \\[2ex] = \; -\int_0^t \left\langle V'_\lambda(x_\lambda(\tau)) - V'_\mu(x_\mu(\tau)), x_\lambda(\tau) - x_\mu(\tau)\right\rangle d\tau \\[2ex] = \; -\int_0^t \langle A_\lambda(x_\lambda(\tau)) - A_\mu(x_\mu(\tau)), x_\lambda(\tau) - x_\mu(\tau)\rangle\, d\tau \end{array}\right. \tag{4.4}$$

We use now the relation $\lambda A_\lambda = \mathbf{1} - J_\lambda$, the fact that $A_\lambda(x) \in \partial V(J_\lambda(x))$

and the monotonicity property of $\partial V(\cdot)$.

$$\left\{\begin{array}{l} a(t) := -\int_0^t \langle A_\lambda(x_\lambda(\tau)) - A_\mu(x_\mu(\tau)), \lambda A_\lambda(x_\lambda(\tau)) - \mu A_\mu(x_\mu(\tau))\rangle \, d\tau \\ -\int_0^t \langle A_\lambda(x_\lambda(\tau)) - A_\mu(x_\mu(\tau)), J_\lambda x_\lambda(\tau) - J_\mu x_\mu(\tau)\rangle \, d\tau \\ \leq -\int_0^t \langle A_\lambda(x_\lambda(\tau)) - A_\mu(x_\mu(\tau)), \lambda A_\lambda(x_\lambda(\tau)) - \mu A_\mu(x_\mu(\tau))\rangle \, d\tau \\ = \int_0^t \langle \lambda A_\lambda(x_\lambda(\tau)), A_\mu(x_\mu(\tau))\rangle \, d\tau + \int_0^t \langle \mu A_\mu(x_\mu(\tau)), A_\lambda(x_\lambda(\tau))\rangle \, d\tau \\ - \int_0^t \lambda \|A_\lambda(x_\lambda(\tau))\|^2 \, d\tau - \int_0^t \mu \|A_\mu(x_\mu(\tau))\|^2 \, d\tau \end{array}\right. \tag{4.5}$$

We note that

$$\left\{\begin{array}{l} \langle \lambda A_\lambda(x_\lambda(\tau)), A_\mu(x_\mu(\tau))\rangle \leq \lambda \|A_\lambda(x_\lambda(\tau))\| \|A_\mu(x_\mu(\tau))\| \\ \leq \lambda \|A_\lambda(x_\lambda(\tau))\|^2 + \frac{\lambda}{4} \|A_\mu(x_\mu(\tau))\|^2 \end{array}\right.$$

and, in the same way, that

$$\langle \lambda A_\lambda(x_\lambda(\tau)), A_\mu(x_\mu(\tau))\rangle \leq \mu \|A_\mu(x_\mu(\tau))\|^2 + \frac{\mu}{4} \|A_\lambda(x_\lambda(\tau))\|^2$$

Therefore, we deduce from these remarks and from (4.5) that

$$a(t) \leq \frac{1}{4}\int_0^t \left(\mu \|A_\lambda(x_\lambda(\tau))\|^2 + \lambda \|A_\mu(x_\mu(\tau))\|^2\right) d\tau \leq \frac{\lambda+\mu}{4} t \|V_0'(x_0)\|^2$$

since

$$\|A_\lambda(x(t))\| = \|x_\lambda'(t)\| \leq \|x_\lambda'(0)\| = \|A_\lambda(x_0)\| \leq \|V_0'(x_0)\|$$

Hence $x_\lambda(\cdot)$ is a Cauchy sequence of the space $C(0,1;X)$, which thus converges uniformly to a continuous function $x(\cdot)$.

The inequality

$$\|x_\lambda(t) - J_\lambda(x_\lambda(t))\| = \lambda \|A_\lambda(x_\lambda(t))\| \leq \lambda \|V_0'(x_0)\|$$

implies that $J_\lambda(x_\lambda(t))$ also converges uniformly to $x(t)$.

The sequence of derivatives $x'_\lambda(\cdot)$ being bounded in $L^2(0,1;X)$, a subsequence converges weakly to $x'(\cdot)$. Inequalities

$$\int_\tau^s (V(x_\lambda(t)) - V(y))\,dt \;\leq\; \int_\tau^s \langle -x'_\lambda(t), x_\lambda(t) - y\rangle dt$$

and Fatou's Lemma imply by going to the limit

$$\int_\tau^s (V(x(t)) - V(y))\,dt \;\leq\; \int_\tau^s \langle -x'(t), x(t) - y\rangle dt$$

Therefore, we deduce that

$$\forall\, y \in X, \text{ for almost all } t \geq 0,\; V(x(t)) - V(y) \;\leq\; \langle -x'(t), x(t) - y\rangle$$

i.e.,

$$\text{for almost all } t \geq 0,\; -x'(t) \;\in\; \partial V(x(t)) \;\square$$

**Lemma 4.3** *Assume that $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ is nontrivial, convex, lower semicontinuous and bounded below. Let $V_0'(x))$ denote the element of $\partial V(x)$ with the smallest norm. Then the solution $x(\cdot)$ is* slow *in the sense that for almost any $t$, the norm of the velocity $x'(t)$ is the smallest one:*

$$\text{for almost all } t \geq 0,\; x'(t) \;=\; -V_0'(x(t))$$

*Furthermore, $t \mapsto V_0'(x(t))$ is nonincreasing and continuous from the right and*

$$\forall\, t \geq 0,\; \lim_{h \mapsto 0+} \frac{x(t+h) - x(t)}{h} \;=\; -V_0'(x(t))$$

**Proof** — We have seen that for $t \geq 0$,

$$\|x'_\lambda(t)\| \leq \|x'_\lambda(0)\| = \|A_\lambda(x_\lambda(0))\| \leq \|V_0'(x_\lambda(0))\| = \|V_0'(x_0)\|$$

On the other hand, every weak cluster $v(t)$ of the bounded sequence $x'_\lambda(t) \in \partial V(J_\lambda(x_\lambda)(t))$ belongs to $-\partial V(x(t))$. Since $\|x'_\lambda(t)\| \leq \|V_0'(x_0)\|$, we also deduce that the solutions are uniformly Lipschitz.

On the other hand, since

$$\|v(t)\| \;\leq\; \liminf_{\lambda \mapsto 0+} \|x'_\lambda(t)\| \;\leq\; \|V_0'(x_0)\|$$

we infer that

$$\|V_0'(x(t))\| \;\leq\; \|v(t)\| \;\leq\; \|V_0'(x(0))\|$$

This implies that the function $t \mapsto \|V_0'(x(t))\|$ is not increasing.

Let $t_0$ be a point where $x(\cdot)$ is differentiable. Since

$$\frac{d}{dt}\|x(t)\| \leq \|x'(t)\| \leq V_0'(x(t_0))$$

we infer that

$$\left\|\frac{x(t_0+h)-x(t_0)}{h}\right\| \leq \|V_0'(x(t_0))\|$$

Since $x'(t_0) \in -\partial V(x(t_0))$, we infer that $x'(t_0) = -V_0'(x(t_0))$.

Therefore, we deduce that for any $t$, the solution is differentiable from the right. Indeed,

$$\frac{x(t+h)-x(t)}{h} = -\frac{1}{h}\int_t^{t+h} V_0'(x(s))ds$$

It is then sufficient to prove that that $V_0'(x(t))$ is continuous from the right. Indeed, let us consider a sequence $t_n \geq t$ converging to $t$. Since $\|V_0'(x(t_n))\| \leq \|V_0'(x(t))\|$, a subsequence (again denoted by) $V_0'(x(t_n))$ converges weakly to some $p$.

Inequalities

$$V(x(t_n)) - V(x(t)) \leq \langle V_0'(x(t_n)), x(t_n) - x(t)\rangle$$

imply that $p$ belongs to $\partial V(x(t))$. Hence

$$\|p\| \leq \liminf_{n\to\infty} \|V_0'(x(t_n))\| \leq \|V_0'(x(t))\|$$

Hence $p = V_0'(x(t))$ is the weak limit of $V_0'(x(t_n))$ and $\|V_0'(x(t))\|$ the limit of $\|V_0'(x(t_n))\|$. Therefore $V_0'(x(t))$ is the strong limit of $V_0'(x(t_n))$.

Therefore, equation

$$\frac{x(t+h)-x(t)}{h} = -\frac{1}{h}\int_t^{t+h} V_0'(x(\tau))d\tau$$

implies that the for every $t \geq 0$, $x(\cdot)$ has a derivative from the right which is equal to $-V_0'(x(\cdot))$. □

**Lemma 4.4** *Assume that $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ is nontrivial, convex, lower semicontinuous and bounded below. Then the solution $x(\cdot)$ satisfies*

$$\text{for almost all } t \geq 0,\ \frac{d}{dt}V(x(t)) + \|x'(t)\|^2 = 0$$

*Furthermore, if $V$ is inf-compact, then $x(t)$ converges when $t \to \infty$ to a limit $x_*$ which achieves the minimum of $V$:*

$$\lim_{t\to\infty} V(x(t)) = \inf_{x\in X} V(x) = V(x_*)$$

**Proof** — Since $x'(t) = -V_0'(x(t))$ is continuous from the right, inequalities

$$\begin{cases} V(x(t)) - V(x(t+h)) \leq \langle -x'(t), x(t) - x(t+h)\rangle \\ V(x(t+h)) - V(x(t)) \leq \langle -x'(t+h), x(t+h) - x(t)\rangle \end{cases}$$

imply that

$$\text{for almost all } t \geq 0, \ \frac{d}{dt}V(x(t)) + \|x'(t)\|^2 = 0$$

This implies in particular that the function $t \mapsto V(x(t))$ is decreasing and thus, that it converges to some $\alpha \geq -V^\star(0)$.

Integrating this inequality from $r$ to $s$, we deduce that

$$\int_r^s \|x'(t)\|^2 dt \leq V(x(r)) - V(x(s))$$

so that, thanks to the Cauchy criterion, we infer that

$$\int_0^\infty \|x'(t)\|^2 dt < +\infty$$

This implies that for any $\varepsilon > 0$, the measure of the set

$$T_\varepsilon := \{t \in \mathbf{R}_+ \mid \|x'(t)\| < \varepsilon\}$$

is infinite. Otherwise, its measure would be finite, so that the measure of $\mathbf{R}_+ \backslash T_\varepsilon$ would be infinite and

$$\infty = \text{meas}(\mathbf{R}_+ \backslash T_\varepsilon) \leq \frac{1}{\varepsilon^2}\int_{\mathbf{R}_+ \backslash T_\varepsilon} \|x'(t)\|^2 dt < +\infty$$

which is impossible. So, for any $\varepsilon > 0$, there exists one $t \geq 0$ such that $\|x'(t)\| \leq \varepsilon$. For such a $t$ and for any $x_\star$ achieving the minimum of $V$, we obtain

$$\begin{cases} \alpha + V^\star(0) \leq V(x(t)) - V(x_\star) \leq \langle -x'(t), x(t) - x_\star\rangle \\ \leq \|x'(t)\|\|x(t) - x_\star\| \leq \|x'(t)\|\|x(0) - x_\star\| \leq \varepsilon\|x(0) - x_\star\| \end{cases}$$

by formula (4.1). By letting $\varepsilon$ converging to 0, we infer that

$$\inf_{t\geq 0} V(x(t)) = \inf_{y\in X} V(y)$$

Since $V$ is inf-compact and $V(x(t)) \leq V(x_0)$, the trajectory remains in a compact subset, so that there exists a cluster point $x_*$, limit of a subsequence $x(t_n)$. Therefore

$$V(x_*) \leq \liminf_{n\to\infty} V(x(t_n)) = \lim_{t\to\infty} V(x(t)) = \inf_{y\in X} V(y)$$

# 5 Subgradient Algorithms

We prove in this section Theorem 3.8 on the convergence of the subgradient algorithm

$$x_{n+1} := x_n - \delta_n \frac{p_n}{\|p_n\|}$$

where $p_n \in \partial V(x_n)$ satisfy

$$\lim_{n\to\infty} \delta_n = 0 \ \& \ \sum_{n=0}^{\infty} \delta_n = +\infty$$

for convex finite functions defined on a finite dimensional vector-space $X$.

We recall that we have assumed that

$$\lim_{n\to\infty} \delta_n = 0 \ \& \ \sum_{n=0}^{\infty} \delta_n = +\infty \tag{5.1}$$

We have to prove that the decreasing sequence of scalars

$$\theta_k := \min_{n=0,\dots,k} V(x_n)$$

converges to the infimum $v := \inf_{x\in X} V(x)$ of $V$ when $k \to \infty$.

**Proof of Theorem 3.8** — We prove this theorem by contradiction. If the conclusion is false, there exists $\eta > 0$ such that $v + 2\eta \leq \theta_k \leq V(x_k)$. Let $\bar{x} \in X$ such that $V(\bar{x}) < v + \eta \leq \theta_k - \eta$. Hence

$$\forall\, k > 0, \ \ V(\bar{x}) + \eta < \theta_k \leq V(x_k)$$

We shall contradict this assumption by constructing a subsequence $x_{n_k}$ such that $\lim_{k\to\infty} V(x_{n_k}) \leq V(\bar{x})$.

First, we observe that

$$\|x_{n+1} - \bar{x}\|^2 = \|x_n - \bar{x}\|^2 - 2\langle x_n - x_{n+1}, x_n - \bar{x}\rangle + \|x_{n+1} - x_n\|^2$$

so that, by recalling that $\|x_{n+1} - x_n\| = \delta_n$ and that $\dfrac{x_n - x_{n+1}}{\delta_n} = \dfrac{p_n}{\|p_n\|}$, we have

$$\|x_{n+1} - \bar{x}\|^2 = \|x_n - \bar{x}\|^2 - 2\delta_n \left\langle \frac{p_n}{\|p_n\|}, x_n - \bar{x} \right\rangle + \delta_n^2$$

Let us set

$$\alpha_k := \min_{n=0,\dots,k} \left\langle \frac{p_n}{\|p_n\|}, x_n - \bar{x} \right\rangle$$

By the definition of the subdifferential and the choice of $\bar{x}$, we deduce that

$$\eta < V(x_k) - V(\bar{x}) \le \langle p_k, x_k - \bar{x} \rangle$$

so that $\alpha_k > 0$. By summing up the above inequalities from $n = 0$ to $k$, we obtain:

$$\|x_{k+1} - \bar{x}\|^2 \le \|x_0 - \bar{x}\|^2 - 2\alpha_k \sum_{n=0}^{k} \delta_n + \sum_{n=0}^{k} \delta_n^2 \tag{5.2}$$

On the other hand, we check easily that under assumption (5.1),

$$\frac{\sum_{n=0}^{k} \delta_n^2}{\sum_{n=0}^{k} \delta_n} \text{ converges to } 0 \tag{5.3}$$

Indeed, set $\gamma_k := \sum_{n=0}^{k} \delta_n^2$, $\tau_k := \sum_{n=0}^{k} \delta_n$ and $K(\varepsilon)$ the integer such that $\delta_k \le \varepsilon$ whenever $k \ge K(\varepsilon)$. Then

$$\gamma_k = \gamma_{K(\varepsilon)-1} + \sum_{k=K(\varepsilon)}^{k} \delta_n^2 \le \gamma_{K(\varepsilon)-1} + \varepsilon \sum_{k=K(\varepsilon)}^{k} \delta_n = \gamma_{K(\varepsilon)-1} + \varepsilon\tau_k$$

so that

$$\forall\, k \ge K(\varepsilon),\ \frac{\gamma_k}{\tau_k} \le \frac{\gamma_{K(\varepsilon)-1}}{\tau_k} + \varepsilon$$

Since $\tau_k \to \infty$, we infer that

$$\limsup_{k\to\infty} \frac{\gamma_k}{\tau_k} \le \varepsilon$$

By letting $\varepsilon$ converge to 0, we have checked (5.3).

Properties (5.2) and (5.3) imply

$$\lim_{k\to\infty} \alpha_k = 0 \tag{5.4}$$

Let $n_k$ be the index such that

$$\left\langle \frac{p_{n_k}}{\|p_{n_k}\|}, x_{n_k} - \bar{x} \right\rangle := \alpha_k := \min_{n=0,\ldots,k} \left\langle \frac{p_n}{\|p_n\|}, x_n - \bar{x} \right\rangle$$

Let us set

$$\bar{x}_{n_k} := \bar{x} + \frac{\langle p_{n_k}, x_{n_k} - \bar{x}\rangle}{\|p_{n_k}\|^2} p_{n_k}$$

We see at once that

$$\begin{cases} \langle p_{n_k}, \bar{x}_{n_k}\rangle = \langle p_{n_k}, x_{n_k}\rangle \\ \|\bar{x}_{n_k} - \bar{x}\| = \left\langle \frac{p_{n_k}}{\|p_{n_k}\|}, x_{n_k} - \bar{x} \right\rangle = \alpha_k \end{cases}$$

The first inequality implies that

$$V(x_{n_k}) - V(\bar{x}_{n_k}) \le \langle p_{n_k}, x_{n_k} - \bar{x}_{n_k}\rangle = 0$$

by the definition of the subdifferential. The second implies that there exists $l > 0$ such that, for $k$ large enough

$$V(\bar{x}_{n_k}) - V(\bar{x}) \le l\|\bar{x}_{n_k} - \bar{x}\| \le l\alpha_k$$

since a convex function defined on a finite dimensional vector-space is locally Lipschitz on the interior of its domain.

Therefore $V(x_{n_k}) \le V(\bar{x}) + l\alpha_k$, so that, passing to the limit, we obtain the contradiction $\lim_{k\to\infty} V(x_{n_k}) \le V(\bar{x})$ we were looking for. □

When $V$ is a lower semicontinuous convex extended function, the subgradient algorithm makes no longer sense since we do not know whether $x_{n+1} := x_n - \delta_n \frac{p_n}{\|p_n\|}$ belongs to the domain of $V$. Hence the idea is to approximate $V$ by its Moreau-Yosida approximation $V_\lambda$ defined by

$$V_\lambda(x) := \inf_{y\in X} \left[ V(y) + \frac{1}{2\lambda}\|y - x\|^2 \right]$$

and to use the gradient method for the Moreau-Yosida approximation. Hence, we have a sequence with two indices, the step of the approximation and the parameter $\lambda$.

Recall that $V_\lambda$ is convex and differentiable. If $J_\lambda x$ denotes the unique point which achieves the minimum of $V_\lambda$, then

$$V_\lambda'(x) = A_\lambda(x) := \frac{1}{\lambda}(x - J_\lambda x) \in \partial V(J_\lambda x)$$

**Theorem 5.1** *Let us consider the Moreau-Yosida approximations $V_\lambda$ of a nontrivial lower semicontinuous convex function $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ is bounded below.*

*We consider the* **regularized gradient method**

$$x_{n+1}^\lambda := x_n^\lambda - \delta_n \frac{p_n^\lambda}{\|p_n^\lambda\|}$$

*where*

$$p_n^\lambda := V_\lambda'(x_n^\lambda) = \frac{1}{\lambda}(x_n^\lambda - J_\lambda x_n^\lambda)$$

*Assume that*

$$\lim_{n\to\infty} \delta_n = 0 \ \& \ \sum_{n=0}^{\infty} \delta_n = +\infty$$

*Then there exists a subsequence of $V_\lambda(x_k^\lambda)$ which converges to the infimum $v := \inf_{x\in X} V(x)$ of $V$ when $k \to \infty$ and $\lambda \mapsto 0+$.*

**Proof** — We prove this theorem by contradiction. If the conclusion is false, there exist $\eta > 0$, $N > 0$ and $\rho > 0$ such that

$$\forall\, n \geq N, \ \forall\, \lambda \leq \rho, \ v + 2\eta \leq V_\lambda(x_n^\lambda)$$

Let $\bar{x} \in X$ such that $V(\bar{x}) < v + \eta \leq V_\lambda(x_n^\lambda) - \eta$. Hence

$$\forall\, n \geq N, \ \forall\, \lambda \leq \rho, \ V(\bar{x}) + \eta \leq V_\lambda(x_k^\lambda) \tag{5.5}$$

First, we observe that

$$\|x_{n+1}^\lambda - \bar{x}\|^2 = \|x_n^\lambda - \bar{x}\|^2 - 2\left\langle x_n^\lambda - x_{n+1}^\lambda, x_n^\lambda - \bar{x}\right\rangle + \|x_{n+1}^\lambda - x_n^\lambda\|^2$$

so that, by recalling that $\|x_{n+1}^\lambda - x_n^\lambda\| = \delta_n$ and that $\dfrac{x_n^\lambda - x_{n+1}^\lambda}{\delta_n} = \dfrac{p_n^\lambda}{\|p_n^\lambda\|}$, we have

$$\|x_{n+1}^\lambda - \bar{x}\|^2 = \|x_n^\lambda - \bar{x}\|^2 - 2\delta_n \left\langle \frac{p_n^\lambda}{\|p_n^\lambda\|}, x_n^\lambda - \bar{x}\right\rangle + \delta_n^2$$

Let us set for any $k \geq N$

$$\alpha_k^\lambda := \min_{n=N,\ldots,k} \left\langle \frac{p_n^\lambda}{\|p_n^\lambda\|}, x_n^\lambda - \bar{x}\right\rangle$$

Since $V_\lambda(\bar{x}) \leq V(\bar{x})$, we deduce that from the definition of the subdifferential and the choice of $\bar{x}$ that

$$\eta \;\leq\; V_\lambda(x_n^\lambda) - V(\bar{x}) \;\leq\; V_\lambda(x_n^\lambda) - V_\lambda(\bar{x}) \;\leq\; \left\langle p_n^\lambda, x_n^\lambda - \bar{x} \right\rangle$$

so that $\alpha_k^\lambda > 0$. By summing up the above inequalities from $n = N$ to $k > N$, we obtain:

$$\|x_{k+1}^\lambda - \bar{x}\|^2 \leq \|x_N - \bar{x}\|^2 - 2\alpha_k \sum_{n=N}^{k} \delta_n + \sum_{n=N}^{k} \delta_n^2 \tag{5.6}$$

On the other hand, we check easily that under assumption (5.1),

$$\frac{\sum_{n=N}^k \delta_n^2}{\sum_{n=N}^k \delta_n} \text{ converges to } 0 \tag{5.7}$$

Indeed, set $\gamma_k := \sum_{n=N}^k \delta_n^2$, $\tau_k := \sum_{n=N}^k \delta_n$ and $K(\varepsilon)$ the integer such that $\delta_k \leq \varepsilon$ whenever $k \geq K(\varepsilon)$. Then

$$\gamma_k = \gamma_{K(\varepsilon)-1} + \sum_{k=K(\varepsilon)}^{k} \delta_n^2 \leq \gamma_{K(\varepsilon)-1} + \varepsilon \sum_{k=K(\varepsilon)}^{k} \delta_n = \gamma_{K(\varepsilon)-1} + \varepsilon\tau_k$$

so that

$$\forall\, k \geq K(\varepsilon),\; \frac{\gamma_k}{\tau_k} \;\leq\; \frac{\gamma_{K(\varepsilon)-1}}{\tau_k} + \varepsilon$$

Since $\tau_k \to \infty$, we infer that

$$\limsup_{k\to\infty} \frac{\gamma_k}{\tau_k} \;\leq\; \varepsilon$$

By letting $\varepsilon$ converge to 0, we have checked (5.7).

Properties (5.6) and (5.7) imply

$$\alpha_k^\lambda \leq \beta_k := \frac{\sum_{n=N}^k \delta_n^2}{2\sum_{n=N}^k \delta_n} + \frac{\|x_N - \bar{x}\|^2}{2\sum_{n=N}^k \delta_n} \text{ converges to } 0 \tag{5.8}$$

Let us take $\lambda := \beta_k$ and $n_k$ be the index such that

$$\left\langle \frac{p_{n_k}^{\beta_k}}{\|p_{n_k}^{\beta_k}\|}, x_{n_k}^{\beta_k} - \bar{x} \right\rangle \;:=\; \alpha_k^{\beta_k} \;:=\; \min_{n=N,\ldots,k} \left\langle \frac{p_n^{\beta_k}}{\|p_n^{\beta_k}\|}, x_n^{\beta_k} - \bar{x} \right\rangle$$

Let us set

$$y_{n_k}^{\beta_k} := \bar{x} + \frac{\left\langle p_{n_k}^{\beta_k}, x_{n_k}^{\beta_k} - \bar{x} \right\rangle}{\|p_{n_k}^{\beta_k}\|^2} p_{n_k}^{\beta_k}$$

We see at once that

$$\begin{cases} \left\langle p_{n_k}^{\beta_k}, y_{n_k}^{\beta_k} \right\rangle = \left\langle p_{n_k}^{\beta_k}, x_{n_k}^{\beta_k} \right\rangle \\ \\ \|y_{n_k}^{\beta_k} - \bar{x}\| = \left\langle \frac{p_{n_k}^{\beta_k}}{\|p_{n_k}^{\beta_k}\|}, x_{n_k}^{\beta_k} - \bar{x} \right\rangle = \alpha_k^{\beta_k} \end{cases}$$

The first inequality implies that

$$V_{\beta_k}(x_{n_k}^{\beta_k}) - V_{\beta_k}(y_{n_k}^{\beta_k}) \leq \left\langle p_{n_k}^{\beta_k}, x_{n_k}^{\beta_k} - y_{n_k}^{\beta_k} \right\rangle = 0$$

by the definition of the subdifferential.

We thus deduce from (5.5) that

$$\begin{cases} V(\bar{x}) + \eta \leq V_{\beta_k}(x_{n_k}^{\beta_k}) \leq V_{\beta_k}(y_{n_k}^{\beta_k}) \\ \\ \leq V(\bar{x}) + \dfrac{1}{2\beta_k}\|y_{n_k}^{\beta_k} - \bar{x}\|^2 \leq V(\bar{x}) + \dfrac{\alpha_k^{\beta_k^2}}{2\beta_k} \leq V(\bar{x}) + \dfrac{\beta_k^2}{2\beta_k} \end{cases} \tag{5.9}$$

so that we obtain the contradiction $\eta \leq \frac{\beta_k}{2}$ which converges to 0.

# 6 Duality Theory

## 6.1 The Duality Theorem

We use the above subdifferential calculus for implementing the Fermat Rules and duality theory for the following general class of convex minimization problems: We consider

1. two finite dimensional spaces $X$ and $Y$;

2. two nontrivial, convex, lower semi-continuous functions

$$\begin{cases} i) & V : X \to \mathbf{R} \cup \{+\infty\} \\ ii) & W : Y \to \mathbf{R} \cup \{+\infty\} \\ iii) & \text{a continuous, linear operator } A \in L(X,Y) \end{cases} \tag{6.1}$$

We shall choose elements $y \in Y$ and $p \in X^*$ as parameters of the optimization problems

$$v := \inf_{x \in X} (V(x) - \langle p, x\rangle + W(Ax + y)) \tag{6.2}$$

and

$$v_* := \inf_{q \in Y^*} (V^*(p - A^*q) + W^*(q) - \langle q, y\rangle) \tag{6.3}$$

which we shall solve at the same time.

We shall say the minimization problems $v$ and $v_*$ are dual.

**Theorem 6.1** *a) We suppose that the conditions (6.1) are satisfied. If*

$$p \in \operatorname{Int}(\operatorname{Dom} V^* + A^* \operatorname{Dom} W^*), \tag{6.4}$$

*then there exists a solution $\overline{x}$ of the problem $v$ and*

$$v + v_* = 0. \tag{6.5}$$

*b) If we suppose further that*

$$y \in \operatorname{Int}(\operatorname{Dom} W - A \operatorname{Dom} V) \tag{6.6}$$

*then the following conditions are equivalent*

$$\begin{cases} i) & \overline{x} \text{ is a solution of the "primal" problem } v \\ ii) & \overline{x} \text{ is a solution of the inclusion } p \in \partial V(\overline{x}) + A^*\partial W(A\overline{x} + y). \end{cases} \tag{6.7}$$

*c) Similarly, assumption (6.6) implies that there exists a solution $\overline{q}$ of the dual problem $v_*$ and the two assumptions imply that the following conditions are equivalent:*

$$\begin{cases} i) & \overline{q} \text{ is a solution of the problem } v_* \\ ii) & \overline{q} \text{ is a solution of the inclusion } y \in \partial W^*(\overline{q}) - A\partial V^*(p - A^*\overline{q}). \end{cases}$$

*d) The two assumptions imply that the solutions $\overline{x}$ and $\overline{q}$ of the problems $v$ and $v_*$ are solutions of the system of inclusions*

$$\begin{aligned} i) \quad & p \in \partial V(\overline{x}) + A^*(\overline{q}) \\ ii) \quad & y \in -A\overline{x} + \partial W^*(\overline{q}). \end{aligned} \tag{6.8}$$

**Remark** — An optimal solution of the dual minimization problem $v_*$ is usually called a *Lagrange (or Kuhn–Tucker) multiplier*, the inclusion (6.7)iii) is usually called the *Euler–Lagrange inclusion* and the inclusion (6.1)iii) is the *Euler–Lagrange dual inclusion.* The system of inclusions (6.8) is usually called the *Hamiltonian system.* □

The set-valued map $(x,q) \to (\partial V(x) + A^*q) \times (-Ax + \partial W^*(q))$ from $X \times Y^*$ to its dual $X^* \times Y$ may be written symbolically in matrix form by

$$\begin{pmatrix} \partial V & A^* \\ -A & \partial W^* \end{pmatrix} \tag{6.9}$$

The set of solutions $(\overline{x}, \overline{q})$ of the minimization problems $v$ and $v_*$ may then be written in the suggestive form

$$\begin{pmatrix} \partial V & A^* \\ -A & \partial W^* \end{pmatrix}^{-1} \begin{pmatrix} p \\ y \end{pmatrix}$$

This notation highlights the variation of the set of solutions as a function of the parameters $p \in X^*$ and $y \in Y$.

**Remark** — When assumptions (6.4) and (6.6) of Theorem 6.1 are satisfied, solution of the problem $v$ is equivalent to solution of the inclusion (set-valued equation)

$$p \in \partial V(\overline{x}) + A^*\partial W(A\overline{x} + y). \tag{6.10}$$

Theorem 6.1 indicates another way of solving this problem. This involves first solving the inclusion

$$y \in \partial W^*(\overline{q}) - A\partial V^*(p - A^*\overline{q}) \tag{6.11}$$

and then choosing $\overline{x}$ in the set

$$\partial V^*(p - A^*\overline{q}) \cap A^{-1}(\partial W^*(\overline{q}) - y). \tag{6.12}$$

This procedure is only sensible if the second inclusion is easier to solve than the first. This clearly depends on the functions $V$ and $W$. If $W$ is differentiable, it may be better to solve the inclusion (16). If, instead, $V^*$ is differentiable, it may be easier to solve the inclusion (17), which in this case may be written as

$$A\nabla V^{\star}(p - A^{\star}\overline{q}) + y \in \partial W^{\star}(\overline{q}) \tag{6.13}$$

or as

$$\forall q \in Y, \qquad \langle -A\nabla V^{\star}(p - A^{\star}\overline{q}) - y, \overline{q} - q\rangle + W^{\star}(\overline{q}) - W^{\star}(q) \leq 0. \tag{6.14}$$

**Remark: Gradient Methods —** For solving the minimization problem $v$ (or $v_{\star}$), we can use the gradient method to either problem $v$ or $v_{\star}$.

We deduce from Theorem 3.7 the following result:

**Theorem 6.2** *We posit the assumptions of Theorem 6.1. Then, for any initial state* $x_0 \in \mathrm{Dom}(V)$, *there exists a* unique solution *to the differential inclusion*

$$-x'(t) \in \partial V(x(t)) + A^{\star}\partial W(Ax(t))$$

*starting at* $x_0$, *converging to an optimal solution* $\overline{x}$ *and satisfying*

$$\lim_{t\to\infty}(V(x(t)) + W(Ax(t))) = \inf_{x\in X}(V(x) + W(Ax))$$

*For any initial state* $q_0 \in \mathrm{Dom}(V)$, *there exists a* unique solution *to the differential inclusion*

$$-q'(t) \in \partial W^{\star}(q(t)) - A^{\star}\partial V^{\star}(-A^{\star}q(t))$$

*starting at* $q_0$, *converging to a solution* $\overline{q}$ *of the dual problem and satisfying*

$$\lim_{t\to\infty}(W^{\star}(q(t)) + V^{\star}(-A^{\star}q(t))) = \inf_{q\in Y^{\star}}(W^{\star}(q) + V^{\star}(-A^{\star}q))$$

The discrete subgradient algorithms can be used when either the function $V + W \circ A$ or $V^{\star} \circ (-A^{\star}) + W^{\star}$ are finite and continuous. In this case. they yield

$$\begin{cases} x_{n+1} - x_n \in -\delta_n \dfrac{p_n}{\|p_n\|} \text{ where} \\ \\ p_n \in \partial V(x_n) + A^{\star}\partial W(Ax_n + y) \end{cases}$$

and

$$\begin{cases} q_{n+1} - q_n \in -\delta_n \dfrac{y_n}{\|y_n\|} \text{ where} \\ \\ y_n \in \partial W^{\star}(q_n) - A\partial V^{\star}(p - A^{\star}q_n) \ \square \end{cases}$$

Otherwise, we have to take their Moreau-Yosida approximations.

## 6.2 Minimization Problems with Constraints

Let us consider
i) two finite dimensional spaces $X$ and $Y$,
ii) a continuous, linear operator $A \in L(X,Y)$,
iii) a convex, closed subset $M \subset Y$,
iv) a nontrivial, convex, lower semi-continuous function $V : X \to \mathbf{R} \cup \{+\infty\}$ and two elements $y \in Y$ and $p \in X^*$.

We consider the minimization problem

$$v := \inf_{Ax \in M - y} (V(x) - \langle p, x \rangle) \tag{6.15}$$

with its associated dual problem

$$v_* := \inf_{q \in Y^*} (V^*(p - A^* q) + \sigma_M(q) - \langle q, y \rangle). \tag{6.16}$$

**Corollary 6.3** *If we suppose that*

$$p \in \mathrm{Int}\,(\mathrm{Dom}\, V^* + A^* \mathrm{Dom}(\psi_M)) \tag{6.17}$$

*then there exists a solution $\overline{x}$ (satisfying $A\overline{x} \in M - y$) of the problem $v$. If we suppose further that*

$$y \in \mathrm{Int}\,(M - A\,\mathrm{Dom}\, V) \tag{6.18}$$

*then the solutions $\overline{x}$ of the problem $v$ are the solutions of the inclusion*

$$p \in \partial V(\overline{x}) + A^* N_M(A\overline{x} + y) \tag{6.19}$$

*The following conditions are then equivalent:*
*i) $\overline{q}$ is a solution of the inclusion $y \in \partial \sigma_M(\overline{q}) - A \partial V^*(p - A^* \overline{q})$.*
*ii) The optimal solutions $\overline{x}$ and $\overline{q}$ of the problems $v$ and $v_*$ are related by*

$$p \in \partial V(\overline{x}) + A^* \overline{q} \quad \text{and} \quad \overline{q} \in N_M(A\overline{x} + y). \tag{6.20}$$

The minimization problem

$$v := \inf_{Ax + y = 0} (V(x) - \langle p, x \rangle) \tag{6.21}$$

which is a minimization problem with 'constraints of equality' is obtained as the particular case in which $M = \{0\}$. Its dual problem is

$$v_* := \inf_{q \in Y^*} (V^*(p - A^* q) - \langle q, y \rangle). \tag{6.22}$$

**Corollary 6.4** *If we suppose that*

$$p \in \operatorname{Int}(\operatorname{Dom} V^* + \operatorname{Im} A^*) \tag{6.23}$$

*then there exists a solution $\overline{x}$ of the problem $v$.*

*If we suppose further that*

$$-y \in \operatorname{Int}(A \operatorname{Dom} V) \tag{6.24}$$

*then the solutions $\overline{x}$ of the problem $v$ are the solutions of the inclusion*

$$p \in \partial V(\overline{x}) + \operatorname{Im} A^*, \qquad A\overline{x} + y = 0 \tag{6.25}$$

*The following conditions are equivalent*

*i) $\overline{q}$ is a solution of the problem $v_*$;*

*ii) $\overline{q}$ is a solution of the inclusion $y \in -A\partial V^*(p - A^*\overline{q})$.* (34)

*The optimal solutions $\overline{x}$ and $\overline{q}$ of the problems $v$ and $v_*$ are related by*

$$p \in \partial V(\overline{x}) + A^*\overline{q}. \tag{6.26}$$

Suppose that $P \subset Y$ is a convex, closed cone and denote its negative polar cone by $P^-$. The cone $P$ defines an order relation $\geq$ by

$$y_1 \geq y_2 \text{ if and only if } y_1 - y_2 \in P \tag{6.27}$$

and the cone $P^-$ defines the order relation

$$q_1 \leq q_2 \text{ if and only if } q_1 - q_2 \in P^-. \tag{6.28}$$

The minimization problem

$$v := \inf_{Ax+y\geq 0}(V(x) - \langle p, x\rangle) \tag{6.29}$$

which is a minimization problem with "inequality constraints" is obtained in the special case in which $M = P$. Its dual problem is

$$v_* := \inf_{q\in P^-}(V^*(p - Aq) - \langle q, y\rangle). \tag{6.30}$$

**Corollary 6.5** *If we suppose that*

$$p \in \mathrm{Int}\,(\mathrm{Dom}\, V^* + A^* P^-) \tag{6.31}$$

*then there exists a solution $\overline{x}$ of the problem $v$.*

*If we suppose further that*

$$y \in \mathrm{Int}\,(P - A\,\mathrm{Dom}\, V) \tag{6.32}$$

*then the solutions $\overline{x}$ of the problem $v$ are the solutions of the inclusion*

$$p \in \partial V(\overline{x}) + A^* N_P(A\overline{x} + y) \tag{6.33}$$

*The following conditions are equivalent*
*i) $\overline{q}$ is a solution of $v_*$*
*ii) $\overline{q}$ is a solution of the inclusion $y \in N_{P^-}(\overline{q}) - A\partial V^*(p - A^*\overline{q})$.*

*The solutions $\overline{x}$ and $\overline{q}$ of the problems $v$ and $v_*$ are related by*

$$\begin{array}{ll} i) & p \in \partial V(\overline{x}) - A^*\overline{q} \\ ii) & A\overline{x} + y \geq 0, \;\; \overline{q} \leq 0 \;\; \textit{and} \;\; \langle \overline{q}, A\overline{x} + y \rangle = 0. \end{array} \tag{6.34}$$

## 6.3 Optimal Allocations

We shall denote by

$$B_i(q, r) \;:=\; \{x \in \mathrm{Dom}(V_i) \mid \langle q, x \rangle \;\leq\; r\}$$

the **budget set** of consumer $i$ and by

$$D_i(q, r) \;:=\; \left\{ x_i \in B_i(q, r) \,|V_i(x_i) \;=\; \inf_{x \in B_i(q,r)} V_i(x) \right\}$$

her **demand set**. The **demand map** is the set-valued map $(q, r) \rightsquigarrow D_i(q, r)$.

We observe that

$$\forall\, \overline{x}_i \in \partial(V_i^\star)(-q), \;\; D_i(q, \langle q, \overline{x}_i \rangle) \;:=\; \widetilde{D_i}(q) \;=\; \partial(V_i^\star)(-q)$$

Indeed, to say that $x_i \in \partial(V_i^\star)(-q)$ amounts to saying that

$$0 \;\in\; \partial V_i(x_i) + q$$

or, equivalently, that

$$\forall\, y \in Y, \;\; V_i(x_i) + \langle q, x_i \rangle \;\leq\; V_i(x) + \langle q, x \rangle$$

This can be written

$$(V_i(x_i) - V_i(x)) \leq \langle q, x - x_i \rangle$$

We define **Change maps** $C_i : L_i \times S^l \rightsquigarrow Y$ by:

$$C_i(x,q) := \left\{ p \in \mathbf{R}^l \mid V_i(x) + \langle p+q, x\rangle = \inf_{y \in Y} (V_i(y) + \langle p+q, y\rangle) \right\}$$

On the other hand, we observe that the **supply map** associating with any $q \in Y^\star$ the subset $S_M(q) \subset M$ defined by

$$S_M(q) := \left\{ \overline{y} \in M \mid \langle q, \overline{y}\rangle = \sigma_M(q) := \sup_{y \in M} \langle q, y\rangle \right\}$$

is equal to the support zone of $M$:

$$S_M(q) = \partial\sigma_M(q)$$

so that

$$y \in S_M(q) \text{ if and only if } q \in N_M(y)$$

We observe that assumption

$$M = M - \mathbf{R}^l_+$$

implies that

$$\forall\, y \in M,\ N_M(y) \subset \mathbf{R}^l_+$$

because for any price $q \in N_M(y)$, we have

$$\langle q, y\rangle = \sigma_M(q) = \sigma_M(q) + \sigma_{-\mathbf{R}^l_+}(q) = \begin{cases} \sigma_M(q) & \text{if } q \in \mathbf{R}^l_+ \\ +\infty & \text{if } q \notin \mathbf{R}^l_+ \end{cases}$$

so that $q$ can only ne nonnegative. Furthermore, if $M$ is bounded above, we infer that

$$\mathrm{Dom}(\sigma_M) \supset \mathbf{R}^l_+$$

because, for any nonnegative price $q \in \mathbf{R}^l_+$,

$$\sigma_M(q) \leq \langle q, \overline{y}\rangle + \sigma_{-\mathbf{R}^l_+}(q) < +\infty$$

Hence assumption (2.2):

$$\begin{cases} i) & M = M - \mathbf{R}^l_+ \text{ is a closed convex subset} \\ ii) & M \subset \underline{y} - \mathbf{R}^l_+ \end{cases}$$

implies that

$$\mathrm{Dom}(\sigma_M) \;=\; \mathbf{R}^l_+$$

Assumption (2.3):

$$\forall\, q \in \mathbf{R}^l_+, \;\; \inf_{x \in L_i} (\langle q, x\rangle + V_i(x)) \;>\; -\infty$$

states that the negative cone $\mathbf{R}^l_-$ is contained in the domain of each $V_i^\star$.

Since the conjugate function of

$$V(x) \;:=\; \sum_{i=1}^{n} V_i(x_i)$$

is the function defined by

$$V^\star(p) \;=\; \sum_{i=1}^{n} V_i^\star(p_i)$$

we see that

$$\mathrm{Dom}(V^\star) \;=\; \prod_{i=1}^{n} \mathrm{Dom}(V_i^\star) \;=\; \mathbf{R}^{ln}_-$$

We take for operator $A$ the sum:

$$Ax \;:=\; \sum_{i=1}^{n} x_i$$

the transpose of which is equal to

$$A^\star q \;=\; (q, \ldots, q)$$

Therefore, property

$$0 \;\in\; \mathrm{Int}(\mathrm{Dom}(V^\star) + A^\star \mathrm{Dom}(W^\star))$$

is satisfied because $\mathrm{Dom}(V^\star) + A^\star\mathrm{Dom}(W^\star) = \mathbf{R}^{ln}_- + A^\star\mathbf{R}^{ln}_+ = \mathbf{R}^{ln}$.

**Theorem 6.6** *Let us assume the set $M$ of scarce resources is closed and convex and satisfies assumption (2.3) and that the loss functions $V_i : Y \mapsto \mathbf{R}\cup\{+\infty\}$ are nontrivial, convex and lower semicontinuous and satisfy (2.3). Assume furthermore that*

$$0 \in \operatorname{Int}\left(\sum_{i=1}^{n} \operatorname{Dom}(V_i) - M\right)$$

*Then there exists an optimal allocation $\overline{x_i} \in \operatorname{Dom}(V_i)$ and a price $\overline{q} \in Y^\star$ which are solutions to the optimal allocation problem*

$$v := \inf_{x \in K} \sum_{i=1}^{n} V_i(\overline{x_i})$$

*where*

$$K := \left\{ x := (x_1, \ldots, x_n) \in \prod_{i=1}^{n} \operatorname{Dom}(V_i) \;\middle|\; \sum_{i=1}^{n} x_i \in M \right\}$$

*and to its dual problem*

$$v_\star := \inf_{q \in Y} \left(\sigma_M(q) + \sum_{i=1}^{n} V_i^\star(-q)\right)$$

*The following conditions are equivalent:*

*a) the price $\overline{q}$ and the allocation $(\overline{x}_1, \ldots, \overline{x}_n)$ satisfy*

$$\begin{cases} i) & \forall\, i = 1, \ldots, n, \;\; \overline{x}_i \in \widetilde{D}_i(\overline{q}) \\ & \text{each } \overline{x}_i \text{ belongs to consumer } i\text{'s demand set} \\ ii) & \displaystyle\sum_{i=1}^{n} \overline{x}_i \in S_M(\overline{q}) \\ & \text{i.e., } \displaystyle\sum_{i=1}^{n} \overline{x}_i \text{ maximizes the available income } \langle \overline{q}, y \rangle \end{cases}$$

*b) the optimal price $\overline{q}$* **clears the market** *in the sense when it is a solution to the inclusion*

$$0 \in \partial\sigma_M(\overline{q}) - \sum_{i=1}^{n} \widetilde{D}_i(\overline{q})$$

*stating that the supply $\partial\sigma_M(\overline{q})$ is balanced by the total demand $\sum_{i=1}^{n} \widetilde{D}_i(\overline{q})$. c) the price $\overline{q}$ and the allocation $(\overline{x}_1, \ldots, \overline{x}_n)$ satisfy*

$$\begin{cases} i) & \forall\, i = 1, \ldots, n, \;\; 0 \;\in\; C_i(\overline{x}_i, \overline{q}) \\ & \text{each } \overline{x}_i \text{ is an equilibrium of the consumer } i\text{'s change map} \\ ii) & \overline{q} \;\in\; N_N\left(\sum_{i=1}^{n} \overline{x}_i\right) \end{cases}$$

By taking $V_i := \psi_{L_i}$, we deduce the following characterization of the tangent and normal cones to the sets of allocations

**Corollary 6.7** *Assume that the Resource Set and the Consumption Sets are closed and convex and satisfy*

$$0 \;\in\; \mathrm{Int}\left(\sum_{i=1}^{n} L_i - M\right)$$

*Then*

$$T_K(x) \;:=\; \left\{ v \;:=\; (v_1, \ldots, v_n) \;\in\; \prod_{i=1}^{n} T_{L_i}(x) \;\middle|\; \sum_{i=1}^{n} v_i \;\in\; T_M\left(\sum_{i=1}^{n} x_i\right) \right\} \tag{6.35}$$

*and*

$$N_K(x) := \left\{ (p + q_1, \ldots, p + q_n) \text{ where } q_i \in N_{L_i}(x) \;\&\; p \in N_M\left(\sum_{i=1}^{n} x_i\right) \right\} \tag{6.36}$$

**Proposition 6.8** *We posit the assumptions of Theorem 6.6. Then $\overline{q}$ is a solution of the dual problem if and only if it belongs to the subdifferential $\partial v(y)$ of the marginal function $v$ defined by*

$$v(y) \;:=\; \inf_{x \in K(y)} \sum_{i=1}^{n} V_i(\overline{x_i})$$

*where*

$$K(y) \;:=\; \left\{ x \;:=\; (x_1, \ldots, x_n) \;\in\; \prod_{i=1}^{n} \mathrm{Dom}(V_i) \;\middle|\; \sum_{i=1}^{n} x_i \;\in\; M - y \right\}$$

**Proof —** The marginal function $v$ can be written $v(y) = \inf_{U(x,y)}$ where

$$U(x,y) := \sum_{i=1}^{n} V_i(\overline{x_i}) + \psi_M\left(\sum_{i=1}^{n} x_i + y\right)$$

By Proposition 3.23, we know that $\overline{p} \in \partial v(0)$ if and only if $(0,\overline{q})$ belongs to $\partial U(\overline{x},0)$ since $v(0) = U(\overline{x},0)$. The latter inclusion can be written

$$\begin{cases} \forall\, i = 1,\ldots,n, \;\; 0 \in \partial V_i(\overline{x}) + N_M\left(\sum_{i=1}^{n} \overline{x}_i\right) \\ \overline{q} \;\in\; N_M\left(\sum_{i=1}^{n} \overline{x}_i\right) \end{cases}$$

Therefore, $\overline{q}$ is the solution to the dual problem. □

We know that the differential inclusions

$$-x'(t) \in \partial\left(\sum_{i=1}^{n} V_i(x_i(t)) + \psi_M\left(\sum_{i=1}^{n} x_i(t)\right)\right)$$

and

$$-q'(t) \in \partial\left(\sigma_M(q) + \sum_{i=1}^{n} V_i^\star(-q)\right)$$

have unique solutions converging to solutions of the optimal allocation problem and its dual respectively. Under the assumption of Theorem 6.6, the first differential inclusion can be written in the form

$$\begin{cases} \forall\, i = 1,\ldots,n, \;\; x_i'(t) \;\in\; -\partial V_i(x_i(t)) - q(t) \;=:\; C_i(x_i(t), q(t)) \\ \text{where} \\ q(t) \;\in\; N_M\left(\sum_{i=1}^{n} x_i(t)\right) \end{cases}$$

and the second one in the form

$$-q'(t) \in \partial\sigma_M(q(t)) - \sum_{i=1}^{n} \partial V_i^\star(-q(t))$$

This proves Theorems 2.2 and2.3. □

**Remark —** We can also derive from the duality a third system of diferential inclusions of a Hamiltonian flavor

$$\begin{cases} i) & -x'(t) \in \partial \left(\sum_{i=1}^{n} V_i(x_i(t))\right) + q(t) \\ ii) & -q'(t) \in \partial\sigma_M(q) - \sum_{i=1}^{n} x_i(t) \end{cases}$$

This an algorithm of the form of a dynamical system proposed by S. Smale in [150], which unfortunately shares with the tâtonnement process the flaw of not being necessarily viable. □

# 7 Calculus of Tangent Cones

It may be useful to recall the characterization of the interior of the tangent cone to a convex subset.

**Proposition 7.1 (Interior of a Tangent Cone)** *Assume that the interior of $K \subset X$ is not empty. Then*

$$\forall\, x \in K, \quad \mathrm{Int}(T_K(x)) \;=\; \bigcup_{h>0} \left(\frac{\mathrm{Int}(K) - x}{h}\right)$$

*Furthermore, the graph of the set-valued map $K \ni x \rightsquigarrow \mathrm{Int}(T_K(x))$ is open.*

For the convenience of the reader, we list in the Table 1 some useful formulas of the calculus of tangent cones to convex subsets which will be proved later, in which the subsets $K$, $K_i$, $L$, $M$, …are assumed to be convex.)

We shall need the following characterization of the normal cone to a convex cone:

**Lemma 7.2** *Let $K \subset X$ be a convex cone of a normed space $X$ and $x \in K$. Then*

$$p \in N_K(x) \Longleftrightarrow x \in K,\ p \in K^- \ \&\ <p,x> = 0 \Longleftrightarrow x \in N_{K^-}(p)$$

*where* $N_{K^-}(p) := \{x \in K \mid \forall\, q \in K^-,\ <q-p,x> \leq 0\}$.

**Proof —** To say that $p \in N_K(x)$ means that $<p,x>= \sigma_K(p)$, which is equal to 0 if and only if $p \in K^-$, and the first statement of the lemma follows. □

Table 1: **Properties of Tangent Cones to Convex Sets.**

| | | |
|---|---|---|
| (1) | ▷ | If $x \in K \subset L \subset X$, then<br>$T_K(x) \subset T_L(x)$ & $N_L(x) \subset N_K(x)$ |
| (3) | ▷ | If $x_i \in K_i \subset X_i, \quad (i = 1, \cdots, n)$, then<br>$T_{\prod_{i=1}^n K_i}(x_1, \ldots, x_n) = \prod_{i=1}^n T_{K_i}(x_i)$<br>$N_{\prod_{i=1}^n K_i}(x_1, \ldots, x_n) = \prod_{i=1}^n N_{K_i}(x_i)$ |
| (4)a) | ▷ | If $A \in \mathcal{L}(X,Y)$ and $x \in K \subset X$, then<br>$T_{A(K)}(Ax) = \overline{A(T_K(x))}$<br>$N_{A(K)}(Ax) = {A^\star}^{-1} N_K(x)$ |
| (4)b) | ▷ | If $K_1,\ K_2 \subset X$, $x_i \in K_i$, $i = 1, 2$, then<br>$T_{K_1+K_2}(x_1 + x_2) = \overline{T_{K_1}(x_1) + T_{K_2}(x_2)}$<br>$N_{K_1+K_2}(x_1 + x_2) = N_{K_1}(x_1) \cap N_{K_2}(x_2)$<br>In particular, if $x_1 \in K$ and $x_2$ belongs to a closed subspace $P$ of $X$, then<br>$T_{K+P}(x_1 + x_2) = \overline{T_{K_1}(x_1) + P}$<br>$N_{K+P}(x_1 + x_2) = N_K(x_1) \cap P^\perp$ |
| (5) | ▷ | If $L \subset X$ and $M \subset Y$ are **closed convex** subsets and $A \in \mathcal{L}(X,Y)$ satisfies the<br>**constraint qualification assumption**<br>$0 \in \mathrm{Int}(M - A(L))$, then, for every $x \in L \cap A^{-1}(M)$,<br>$T_{L \cap A^{-1}(M)} = T_L(x) \cap A^{-1} T_M(Ax)$<br>$N_{L \cap A^{-1}(M)} = N_L(x) + A^\star N_M(Ax)$ |
| (5)a) | ▷ | If $M \subset Y$ is *closed convex* and if $A \in \mathcal{L}(X,Y)$ satisfies $0 \in \mathrm{Int}(\mathrm{Im}(A) - M)$,<br>then, for any $x \in A^{-1}(M)$,<br>$T_{A^{-1}(M)}(x) = A^{-1} T_M(Ax)$<br>$N_{A^{-1}(M)}(x) = A^\star N_M(Ax)$ |
| (5)b) | ▷ | If $K_1,\ K_2 \subset X$ are closed convex and satisfy $0 \in \mathrm{Int}(K_1 - K_2)$, then, for any $x \in K_1 \cap K_2$<br>$T_{K_1 \cap K_2}(x) = T_{K_1}(x) \cap T_{K_2}(x)$<br>$N_{K_1 \cap K_2}(x) = N_{K_1}(x) + N_{K_2}(x)$ |
| (5)c) | ▷ | If $K_i \subset X, \quad (i = 1, \ldots, n)$, are closed and convex, $x \in \bigcap_{i=1}^n K_i$ and if there exists $\gamma > 0$ satisfying $\forall x_i$ such that $\|x_i\| \leq \gamma$, $\bigcap_{i=1}^n (K_i - x_i) \neq \emptyset$, then<br>$T_{\bigcap_{i=1}^n K_i}(x) = \bigcap_{i=1}^n T_{K_i}(x)$<br>$N_{\bigcap_{i=1}^n K_i}(x) = \sum_{i=1}^n N_{K_i}(x)$ |

# References


[1] AUBIN J.-P. (1979) MATHEMATICAL METHODS OF GAME AND ECONOMIC THEORY, North-Holland (Studies in Mathematics and its applications, Vol. 7, 1-619)

[2] AUBIN J.-P. & CLARKE (1979) *Shadow prices and duality for a class of optimal control problems*, SIAM J. Opt. and Control, 17, 567-586

[3] AUBIN J.-P. (1980) *Monotone Trajectories of Differential Inclusions : a Darwinian Approach*, Methods of Operations research, 37, 19-40

[4] AUBIN J.-P. (1981) *A dynamical, pure exchange economy with feedback pricing*, J. Economic Behavior and Organizations, 2, 95-127

[5] AUBIN J.-P. (1981) *Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions*, Advances in Mathematics, Supplementary studies, Ed. Nachbin L., 160-232

[6] AUBIN J.-P. (1981) *Cooperative fuzzy games*, Math. Op. Res., 6, 1-13

[7] AUBIN J.-P. (1981) *Locally lipchitz cooperative games*, J. Math. Economics, 8, 241-262

[8] AUBIN J.-P. (1982) MÉTHODES EXPLICITES DE L'OPTIMISATION, Dunod (Traduction Anglaise: Explicit Methods of Optimization, Dunod, 1985)

[9] AUBIN J.-P. (1983) L'ANALYSE NON LINÉAIRE ET SES MOTIVATIONS ÉCONOMIQUES, Masson

[10] AUBIN J.-P. (1987) *Smooth and Heavy Solutions to Control Problems*, in NONLINEAR AND CONVEX ANALYSIS, Eds. B-L. Lin & Simons S., Proceedings in honor of Ky Fan, Lecture Notes in pure and applied mathematics, June 24-26, 1985,

[11] AUBIN J.-P. (1987) EXERCICES D'ANALYSE NON LINÉAIRE, Masson

[12] AUBIN J.-P. (1988) *L'évolution des systèmes historiques vu à travers la théorie de la viabilité*, Stratégique, 1, 129-142

[13] AUBIN J.-P. (1990) *A survey of viability theory*, SIAM J. on Control and Optimization, 28, 749-788

[14] AUBIN J.-P. (1990) *Differential games: a viability approach*, SIAM J. on Control and Optimization, 28, 1294-1320

[15] AUBIN J.-P. (1990) *Fuzzy differential inclusions*, Problems of Control and Information Theory, 19, 55-67

[16] AUBIN J.-P. (1991) VIABILITY THEORY

[17] AUBIN J.-P. (to appear) NEURAL NETWORKS AND QUALITATIVE PHYSICS: A VIABILITY APPROACH

[18] AUBIN J.-P. (in preparation)

[19] AUBIN J.-P. & CELLINA A. (1984) DIFFERENTIAL INCLUSIONS, Springer-Verlag, Grundlehren der math. Wiss.

[20] AUBIN J.-P. & DA PRATO G. (1990) *Stochastic Viability and Invariance*, Annali Scuola Normale di Pisa, 27, 595-694

[21] AUBIN J.-P. & DA PRATO G. ( 1990) *Solutions contingentes de l'équation de la variété centrale*, Comptes-Rendus de l'Académie des Sciences, Paris, 311, 295-300

[22] AUBIN J.-P. & DA PRATO G. (1992) *Contingent Solutions to the Center Manifold Equation*, Annales de l'Institut Henri Poincaré, Analyse Non Linéaire, 9, 13-28

[23] AUBIN J.-P. & DAY R. H. (1980) *Homeostatic trajectories for a class of adaptive economic systems*, J. Economic Dyn. Control, 2, 185-203

[24] AUBIN J.-P. & EKELAND I. (1984) APPLIED NONLINEAR ANALYSIS, Wiley-Interscience

[25] AUBIN J.-P. & FRANKOWSKA H. (1984) *Trajectoires lourdes de systèmes contrôlés*, Comptes-Rendus de l'Académie des Sciences, PARIS ,298, 521-524

[26] AUBIN J.-P. & FRANKOWSKA H. (1985) *Heavy viable trajectories of controlled systems*, Annales de l'Institut Heanri-Poincaré, Analyse Non Linéaire, 2, 371-395

[27] AUBIN J.-P. & FRANKOWSKA H. (1990) SET-VALUED ANALYSIS, Systems and Control: Foundations and Applications, Birkhäuser, Boston, Basel, Berlin

[28] AUBIN J.-P. & FRANKOWSKA H. (1990) *Inclusions aux dérivées partielles gouvernant des contrôles de rétroaction* , Comptes-Rendus de l'Académie des Sciences, Paris, 311, 851-856

[29] AUBIN J.-P. & FRANKOWSKA H. (1991) *Systèmes hyperboliques d'inclusions aux dérivées partielles*, Comptes-Rendus de l'Académie des Sciences, Paris, 312, 271-276

[30] AUBIN J.-P. & FRANKOWSKA H. (1991) *Feedback control for uncertain systems*, in MODELING, ESTIMATION AND CONTROL OF SYSTEMS WITH UNCERTAINTY, Di Masi, Gombani & Kurzhanski Eds, 1-21

[31] AUBIN J.-P. & FRANKOWSKA H. (to appear) *Hyperbolic systems of partial differential inclusions*, Annali Scuola Normale di Pisa

[32] AUBIN J.-P. & FRANKOWSKA H. (to appear) *Partial differential inclusions governing feedback controls*, IIASA WP-90-028

[33] AUBIN J.-P. & FRANKOWSKA H. (in preparation) *Contingent partial differential inclusions governing feedback controls*,

[34] BALASKO Y. (1988) FOUNDATIONS OF THE THEORY OF GENERAL EUILIBRIUM, Academic Press

[35] BEBERNES J. W. & SCHUUR J. D. (1970) *The Ważewski topological method for contingent equations*, Annali di Mat. Pura ed Appli., 87, 271-280

[36] BENSOUSSAN A. & LIONS J.L (1982) CONTRÔLE IMPULSIONNEL ET INÉQUATIONS QUASI-VARIATIONNELLES, Dunod, Paris

[37] BENSOUSSAN A. (1982) STOCHATIC CONTROL BY FUNCTIONAL ANALYSIS METHODS, North-Holland

[38] BOULIGAND G. (1932) INTRODUCTION À LA GÉOMÉTRIE INFINITÉSIMALE DIRECTE, Gauthier-Villars

[39] BRESSAN A., CELLINA A. & COLOMBO G. (1988) *Upper semicontinuous differential inclusions without convexity*, SISSA

[40] BRESSAN A., CELLINA A. & FRYSZKOWSKI A. (to appear) *A class of absolute retracts in spaces of integrable functions*, Trans. AMS

[41] BRESSAN A. & COLOMBO G. (1986) *Extensions and selections of maps with decomposable values*, Preprint SISSA

[42] BRESSAN A. & COLOMBO G. (1987) *Generalized Baire category and differential inclusions in Banach spaces*, Preprint SISSA

[43] BRESSAN A. & COLOMBO G. (to appear) *Existence and continuous dependence for discontinuous O.D.E.'s*, Boll. Un. Mat. Ital. submitted

[44] BRESSAN A. & CORTESI A. (to appear) *Directionally continuous selections in Banach spaces*, Preprint

[45] BRESSAN A. (1983) *Solutions of lower semicontinuous differential inclusions on closed sets*, Rend. Sem. Mat. Univ. Padova, 69, 99-107

[46] BRESSAN A. (1987) *Directional convexity and finite optimality conditions*, Journal of Mathematical Analysis and Applications, 125, 1, 234-246

[47] BRESSAN A. (to appear) *Directionally continuous selections and differential inclusions*, Funkc. Ekvac.

[48] BRESSAN A. (to appear) *On the qualitative theory of lower semicontinuous differential inclusions*, J. Diff. Equations

[49] BRESSAN A. (to appear) *Upper and lower semicontinuous differential inclusions: a unified approach*, Controllability & optimal control

[50] CASTAING CH. & VALADIER M. (1977) CONVEX ANALYSIS AND MEASURABLE MULTIFUNCTIONS, Springer-Velag, Lecture Notes in Math. # 580

[51] CELLINA A., COLOMBO G. & FONDA A. (1986) *A continuous version of Liapunov's theorem on vector measures* , SISSA

[52] CELLINA A. & COLOMBO G. (1988) *On a classical problem of the calculus of variations without convexity assumptions*, SISSA

[53] CELLINA A. & COLOMBO R.M. (1989) *On the representation of measurable set valued maps through selections*, Preprint SISSA

[54] CELLINA A. & COLOMBO R.M. (1990) *Qualitative and quantitative results on a differential inclusion*, SISSA

[55] CELLINA A. & MARICONDA C. (to appear) *The kuratowski index of a decomposable set*, Bull. Pol. Acad. Sci.

[56] CELLINA A. & ORNELAS A. (1988) *Convexity and the closure of the solution set to differential inclusions*, SISSA

[57] CELLINA A. & ORNELAS A. (1988) *Representation of the attainable set for lipschitzian differential inclusions*, SISSA

[58] CELLINA A., FRYSZHOWSKI A. & RZEZUCHOWSKI T. (to appear) *Upper semicontinuity of Nemytskij operators*, SISSA, 1989

[59] CELLINA A. & STAICU V. (1989) *On evolution equations having monotonicities of opposite sign*, SISSA

[60] CELLINA A. (1969) *A theorem on the approximation of set-valued mappings*, Rend. Ac, Na. Lincei, 47, 429-433

[61] CELLINA A. (1969) *Approximation of set-valued functions and fixed point theorems*, Annali Mat. Pura. Appl., 82, 17-24

[62] CELLINA A. (to appear) *On the set solutions to lipschitzean differential inclusions*, Preprint SISSA

[63] CLARKE F.H. (1975) *Generalized gradients and applications*, Trans. Am. Math. Soc., 205, 247-262

[64] CLARKE F.H. (1983) OPTIMIZATION AND NONSMOOTH ANALYSIS, Wiley-Interscience

[65] CORNET B. & HADDAD G. (1985) *Viability theorems for second order differential inclusions*, Israel J. Math.

[66] CORNET B. (1976) *On planning procedures defined by multivalued differential equations*, Système dynamiques et modèles économiques (C.N.R.S)

[67] CORNET B. (1981) *Regular properties of tangent and normal cones*, Cahiers de Maths. de la Décision #8130

[68] CORNET B. (1983) *An existence theorem of slow solutions for a class of differential inclusions*,

[69] CORNET B. (1983) *Monotone dynamic processes in a pure exchange economy*,

[70] CORNET B. (1984) *Regular properties of open tangent cones*, Center for Operations Research & Econometrics - Université Catholique de Louvain

[71] CRANDALL M.G. & LIONS P.L. (1983) *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 277, 1-42

[72] CRANDALL M.G. & PAZY A. (1972) *Nonlinear evolution equations in Banach spaces*, Israel J. Math., 11, 57-94

[73] DA PRATO G. & LUNARDI A. (1988) *Solvability on the real line of a class of linear Volterra integrodifferential equations of parabolic type*, Annali di Mat. Pura ed Applic., 150, 67-118

[74] DORDAN O. (1989) *Simulation qualitative des equations différentielles sur le simplexe*, C.R. Acad. Sc. Paris 302, Série I

[75] DORDAN O. (1990) *Algorithme de simulation qualitative d'une équation différentielle sur le simplexe*, Comptes-Rendus de l'Académie des Sciences, Paris, 310, 479-482

[76] DORDAN O. (1990) *Analyse qualitative* , Thèse Université de Paris-Dauphine

[77] DORDAN O. (to appear) *Mathematical problems arising in qualitative simulation of a differential equation*, Artificial Intelligence

[78] DULUC R. & VIGNERON C. (1990) *Linear functional viability constraints*, Proceedings of the 9th International Conference on Analysis and Optimization of Systems, Nice, June 1990, Lecture Notes in Control and Information Sciences, Springer-Verlag,

[79] FALCONE M. & SAINT-PIERRE P. (1987) *Slow and quasi-slow solutions of differential inclusions*, J. Nonlinear Anal.,T.,M.,A., 3, 367-377

[80] FAN KY (1972) *A minimax inequality and applications*, Inequalities III, 103-113, Shisha Ed., Academic Press

[81] FILIPPOV A.F. (1958) *On some problems of optimal control theory*, Vestnik Moskowskogo Universiteta, Math. no.2, 25-32

[82] FILIPPOV A.F. (1962) *On certain questions in the theory of optimal control*, SIAM J. Control, 1, 76-84

[83] FILIPPOV A.F. (1967) *Classical solutions of differential equations with multivalued right hand side*, SIAM J. on Control, 5, 609-621

[84] FILIPPOV A.F. (1985) DIFFERENTIAL EQUATIONS WITH DISCONTINUOUS RIGHT-HAND SIDE, Nauka

[85] FILIPPOV A.F. (1988) DIFFERENTIAL EQUATIONS WITH DISCONTINUOUS RIGHTHAND SIDE, Kluwer Academic Publishers

[86] FLAM S. D. & ZOWE J. (to appear) *A primal-dual differential method for convex programming*, in OPTIMIZATION AND NONLINEAR ANALYSIS, IOFFE A., MARCUS M. & REICH S. EDS.. Longman

[87] FLAM S. D. (to appear) *Paths to constrained Nash equilibria*, Applied Math. Opt.

[88] FRANKOWSKA H. (1987) *Local controllability and infinitesimal generators of semigroups of set-valued maps*, SIAM J. on Control and Optimization, 25, 412-432

[89] FRANKOWSKA H. (1989) *Contingent cones to reachable sets of control systems*, SIAM J. on Control and Optimization, 27, 170-198

[90] FRANKOWSKA H. (1989) *Estimations a priori pour les inclusions différentielles opérationelles*, Comptes-Rendus de l'Académie des Sciences, PARIS, Série 1, 308, 47-55

[91] FRANKOWSKA H. (1989) *Hamilton-Jacobi equation: viscosity solutions and generalized gradients*, J. of Math. Analysis and Appl. 141, 21-26 )

[92] FRANKOWSKA H. (1989) *High order inverse function theorems*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 283-304

[93] FRANKOWSKA H. (1989) *Local controllability of control systems with feedbacks*, Journal of Optimization Theory and Applications, 60, 277-296

[94] FRANKOWSKA H. (1989) *Non smooth solutions to an Hamilton-Jacobi equation*, Proceedings of the International Conference Bellman Continuum, Antibes, France, June 13-14, 1988, Lecture Notes in Control and Information Sciences, Springer Verlag

[95] FRANKOWSKA H. (1989) *Optimal trajectories associated to a solution of contingent Hamilton-Jacobi equations*, Applied Mathematics and Optimization, 19, 291-311

[96] FRANKOWSKA H. (1989) *Set-valued analysis and some control problems*, Proceedings of the International Conference 30 YEARS OF MODERN CONTROL THEORY, Kingston, June 3-6, 1988 E. Roxin Editor, Marcel Dekker

[97] FRANKOWSKA H. (1990) *A priori estimates for operational differential inclusions* , J. Diff. Eqs.

[98] FRANKOWSKA H. (1990) *On controllability and observability of implicit systems* , Systems and Control Letters, 14, 219-225

[99] FRANKOWSKA H. (1990) *Some inverse mapping theorems*, Ann. Inst. Henri Poincaré, Analyse Non Linéaire, 3, 183-234

[100] FRANKOWSKA H. (1992) SET-VALUED ANALYSIS AND CONTROL THEORY, Systems and Control: Foundations and Applications, Birkhäuser, Boston, Basel, Berlin

[101] FRANKOWSKA H. & QUINCAMPOIX M. (1991) *Viability kernels of differential inclusions with constraints*, Mathematics of Systems, Estimation and Control,

[102] FRANKOWSKA H. & QUINCAMPOIX M. *Un algorithme déterminant les noyaux de viabilité pour des inclusions différentielles avec contraintes*, Comptes-Rendus de l'Académie des Sciences, PARIS, Série 1, 312, 31-36

[103] FRANKOWSKA H. & QUINCAMPOIX M. (to appear) *Value-functions for differential games with feedback strategies*

[104] GOMEZ G. L. (1992) DYNAMIC PROBABILISTIC MODELS AND SOCIAL STRUCTURES, Reidel

[105] HADDAD G. & LASRY J.-M. (1983) *Periodic solutions of functional differential inclusions and fixed points of s-selectionable correspondences*, J. Math. Anal. Appl.

[106] HADDAD G. (1981) *Monotone trajectories of differential inclusions with memory*, Isr. J. Math., 39, 83-100

[107] HADDAD G. (1981) *Monotone viable trajectories for functional differential inclusions*, J. Diff. Eq., 42, 1-24

[108] HADDAD G. (1981) *Topological properties of the set of solutions for functional differential differential inclusions*, Nonlinear Anal. Theory, Meth. Appl., 5, 1349-1366

[109] HOFBAUER J. & SIGMUND K. (1988) THE THEORY OF EVOLUTION AND DYNAMICAL SYSTEMS, Cambridge University Press, London Math. Soc. # 7

[110] KURATOWSKI K. (1958) TOPOLOGIE, VOLS. 1 AND 2 4TH. ED. CORRECTED, Panstowowe Wyd Nauk, Warszawa. (Academic Press, New York, 1966)

[111] LYAPUNOV A.M. (1907) *Problème general de la stabilite du mouvement*, Ann. Fac. Toulouse, vol. 9, 203-474 ; also in : Annals of mathematics study, # 17; Princeton University Press, 1949.

[112] MADERNER N. (to appear) *Regulation of control systems under inequality viability constraints*,

[113] MARCHAUD H. (1934) *Sur les champs de demi-cônes et les équations différentielles du premier ordre*, Bull. Sc. Math., 62, 1-38

[114] MARCHAUD H. (1936) *Sur les champs continus de demi-cônes convexes et leurs intégrales*, Compos. Math. 1, 89-127

[115] MICHAEL E. (1956) *Continuous selections I*, Annals of Math., 63, 361-381

[116] MICHAEL E. (1956) *Continuous selections II*, Annals of Math., 64, 562-580

[117] MICHAEL E. (1957) *Continuous selections III*, Annals of Math., 65, 375-390

[118] MILLER R.K. & MICHEL A.N. (1982) ORDINARY DIFFERENTIAL EQUATIONS, New York : Academic Press

[119] NESTEROV Y. (1984) *Minimization methods for nonsmooth convex and quasiconvex functions*, Matekon 20, 519-531

[120] PAVEL N. H. (1984) DIFFERENTIAL EQUATIONS, FLOW INVARIANCE ANS APLLICATIONS, Pitman Research Notes in Mathematics # 113

[121] PAZY A. (1983) *Semi-groups of linear operators and applications*, Springer Verlag, Berlin

[122] QUINCAMPOIX M. (1990) *Frontières de domaines d'invariance et de viabilité pour des inclusions différentielles avec contraintes*, Comptes-Rendus de l'Académie des Sciences, Paris, 311, 411-416

[123] QUINCAMPOIX M. (1990) *Playable differentiable games*, J. Math. Anal. Appl.

[124] QUINCAMPOIX M. (1991) *Differential inclusions and target problems*, SIAM J. Control & Optimization, IIASA WP-90

[125] QUINCAMPOIX M. (to appear) *Domaines de victoire et de défaite d'un jeu différentiel à une cible et deux joueurs*, Cahiers de Mathématiques de la décision

[126] QUINCAMPOIX M. (to appear) *Enveloppes d'invariance*, Cahiers de Mathématiques de la décision

[127] ROBINSON S.M. (1976) *Stability theory for systems of inequalities, part ii: differentiable nonlinear systems*, SIAM J. on Numerical Analysis,13, 497-513

[128] ROCKAFELLAR R. T. & WETS R. (to appear) ,

[129] ROCKAFELLAR R.T. (1967) MONOTONE PROCESSES OF CONVEX AND CONCAVE TYPE, Mem. of AMS # 77

[130] ROCKAFELLAR R.T. (1970) CONVEX ANALYSIS, Princeton University Press

[131] ROCKAFELLAR R.T. (1981) THE THEORY OF SUBGRADIENTS AND ITS APPLICATIONS TO PROBLEMS OF OPTIMIZATION CONVEX AND NONCONVEX FUNCTIONS , Helderman Verlag, W. Berlin 1981

[132] ROCKAFELLAR R.T. (1985) *Lipschitzian properties of multifunctions*, J. Nonlinear Anal. T.M.A., 9, 867-885

[133] SAARI D. G. (1986) *The source of some paradoxes from social choice and probability*, IIASA CP-8625

[134] SAARI D.G. (1984) *The ultimate of chaos resulting from weighted voting systems*, Advances in Applied Mathematics, 5, 286-308

[135] SAARI D.G. (1989) *Erratic behavior in economic models*, Working paper # 225

[136] SAINT-PIERRE P. (1990) *Approximation of slow solutions to differential inclusions*, Applied Mathematics & Optimisation

[137] SAINT-PIERRE P. (to appear) *Newton's method for set-valued maps*, Cahiers de Mathématiques de la Décision

[138] SAINT-PIERRE P. (to appear) *Viability property of the boundary of the viability kernel*,

[139] SEUBE N. (1991) *Apprentisage de lois de contrôle régulant des contraintes sur l'état par réseaux de neurones*, Comptes-Rendus de l'Académie des Sciences, Paris, 312, 446-450

[140] SEUBE N. (1992)Régulation de systèmes contrôlés avec contraintes sur l'état par réseaux de neurones, Thèse Université de Paris-Dauphine

[141] SEUBE N. (to appear) *A Neural Network Approach For Autonomous Underwater Vehicle Control Based On Viability Theory*,

[142] SEUBE N. (to appear) *A neural network learning scheme for viable feedback control laws*,

[143] SEUBE N. (to appear) *Neural Network Learning Rules For Control: Application To Auv Tracking Control*,

[144] SEUBE N. (to appear) *Neural Networks Learning Rules for Control: Uniform Dynamic Backpropagation, Heavy Adaptive Learning Rule,*

[145] SHI SHUZHONG (1986) *Théorèmes de viabilité pour les inclusions aux dérivées partielles,* C. R. Acad. Sc. Paris, 303, Série I, 11-14

[146] SHI SHUZHONG (1988) *Nagumo type condition for partial differential inclusions,* Nonlinear Analysis, T.M.A., 12, 951-967

[147] SHI SHUZHONG (1988) *Optimal control of strongly monotone variational inequalities,* SIAM J. Control & Optimization, 26, 274-290

[148] SHI SHUZHONG (1989) *Viability theorems for a class of differential-operator inclusions,* J. of Diff. Eq., 79, 232-257

[149] SHI SHUZHONG (to appear) *Viability theory for partial differential inclusions,* Cahier de MD # 8601 - Université Paris-Dauphine

[150] SMALE S. (1976) *Exchange processes with price adjustement,* Journal Mathematical Economics 3, 211-226

[151] STACCHETTI (1985) *Analysis of a dynamic, decentralized exchange economy,* Journal of Mathematical Economics, 14, 241-259

[152] TCHOU N. A. (1988) *Existence of slow monotone solutions to a differential inclusion,* J. Math. Anal. Appl.

[153] WAŻEWSKI T. (1990) SELECTED PAPERS, PWN (Polish Scientific Publishers)

[154] YORKE J.A. (1967) *Invariance for ordinary differential équations,* Math. Systems Theory, 1, 353-372

[155] YORKE J.A. (1969) *Invariance for contingent equations,* Lecture Notes in Operations Research and Mathematical Economics, 12, Springer Verlag, Berlin, 379-381

[156] YORKE J.A. (1970) *Differential inequalities and non-lipschitz scalar functions,* Math. Systems Theory, 4, 140-153

[157] YORKE J.A. (1971) *Another proof of the Liapounov convexity theorem,* SIAM J. Control Optim. 9, 351-353

[158] YORKE J.A. (1972) *The maximum principle and controllability of nonlinear differential equations,* SIAM J. of Control and Optimization, 10, 334-338

[159] YOSHIZAWA T. (1966) STABILITY THEORY BY LYAPUNOV SECOND METHOD, The Mathematical Society of Japan

[160] YOSHIZAWA T. (1966) STABILITY THEORY BY LYAPUNOV'S SECOND METHOD, Tokyo : The mathematical society of Japan

[161] YOUNG L.C. (1969) LECTURES ON THE CALCULUS OF VARIATIONS AND OPTIMAL CONTROL THEORY, Saunders, Philadelphia

[162] ZADEH L.A. (1965) *Fuzzy sets*, Information and Control, 8, 338-353

[163] ZAREMBA S.C. (1934) *Sur une extension de la notion d'équation diférentielle*, COMPTES RENDUS ACAD. SCIENCES 199, 545-548

[164] ZAREMBA S.C. (1936) *Sur les équations au paratingent*, Bull. Sc. Math., 60, 139-160