Università di Padova – Dipartimento di Matematica Pura ed Applicata

Scuole di Dottorato in Matematica Pura e Matematica Computazionale

Seminario Dottorato 2007/08

Presentazione **2**

Sunti dei seminari (tratti dalla pagina web del Seminario Dottorato) **3**

Note dei seminari **10**

Marco Di Summa, *Extended formulations for some mixed-integer sets* 10
Gabriele Terrone, *Viscosity methods for the order reduction of singularly perturbed control systems* 17
Agnese Cadel, *Spin glasses and directed polymers* 23
Olga Bernardi, *Sistemi dinamici e insiemi di Aubry-Mather* 29
Matteo Dalla Riva, *A functional analytic approach for the analysis of singularly perturbed boundary value problems* 36
Cristina Vagnoni, *Construction of balanced complex polytopes in* $\mathbb{C}^2$ 42
Giovanni Cerulli Irelli, *Cluster algebras: some motivating examples for their introduction* 55
Giorgia Callegaro, *Computing VaR and CVaR for energy derivatives* 67
Luca Prelli, *Sheaves on subanalytic sites and* $\mathcal{D}$*-modules* 79
Manolo Venturin, *Numerical modeling for convection–dominated problems* 87
Paola Toto, *The Basic Picture on sets evaluated over an overlap algebra* 92
Luigi Manca, *Chaotic phenomena described by stochastic equations* 103
Roberto Monti, *Lunghezza di curve e misure di superficie* 108
David Barbato, *An introduction to Stochastic Fluid Dynamic Models* 115

Presentazione

Questo documento offre un resoconto dell'attività del Seminario Dottorato 2007/08. Giunto al suo secondo anno di svolgimento regolare, il Seminario Dottorato fornisce un'opportunità a dottorandi, assegnisti e borsisti (ma, a volte, anche a ricercatori affermati) in Matematica Pura e Computazionale di comunicare le proprie ricerche in modo interessante per un pubblico matematicamente ben istruito ma non specialista.
Un grazie va a tutti coloro che hanno contribuito alla riuscita dell'iniziativa, in primo luogo naturalmente a chi ha accettato di svolgere i seminari e di stendere queste note.

Padova, 20 giugno 2008

Corrado Marastoni, Tiziano Vargiolu

Sunti dei seminari (tratti dalla pagina web del Seminario Dottorato)

26 settembre 2007

Algebra and Topology

Dan Segal (professor at Oxford-All Souls College)

Actually the subject begins with number theory. In the 1930s Wolfgang Krull extended the Fundamental Theorem of Galois Theory from finite Galois extensions to infinite Galois extensions. In order to obtain a bijective correspondence between intermediate fields and subgroups of the Galois group, Krull realized that it is necessary to consider the latter as a topological group: each field corresponds to a subgroup and conversely. The topology is defined by taking as neighbourhoods of the identity the Galois groups of the big field over (larger and larger) finite sub-extensions of the small field. In this way, the Galois group appears as the inverse limit of a system of finite (Galois) groups. A group that is the inverse limit of an inverse system of finite groups is called a profinite group. It is in a natural way a compact, totally disconnected topological group (inheriting these properties from the finite groups considered as finite spaces). An infinite abstract group may have many different structures as a profinite group (i.e. different topologies) (or of course none). But it was discovered by J-P. Serre in the 1970s that for certain kinds of profinite group, the topology is uniquely determined by the underlying group. These are the so-called finitely generated pro-p groups. Serre wondered whether the same might be true for finitely generated profinite groups in general; after about 30 years of partial results by several mathematicians, we have recently shown that the answer is "yes". In fact, what the proof does is to show that many closed subgroups can be constructed in a purely algebraic way. In the talk I will try to sketch some of the mathematics involved in the proof, and mention other related results and open problems.

---

17 ottobre 2007

Formulazioni estese per problemi di programmazione intera mista

Marco Di Summa (dottorato in Matematica Computazionale)

In certi problemi di ottimizzazione, detti problemi di programmazione intera mista, è necessario studiare regioni dello spazio definite da disequazioni lineari, con la condizione aggiuntiva che alcune delle coordinate possono assumere solo valori interi. L'analisi di queste regioni nel loro spazio naturale di definizione è resa complessa proprio dai vincoli di interezza. Tuttavia in certi casi l'introduzione di variabili aggiuntive permette di descrivere in modo molto più semplice la regione in esame. Tali formulazioni, date in uno spazio di dimensione superiore, sono dette "formulazioni estese" e sono di fondamentale importanza per la soluzione di problemi di questo tipo. In questo seminario, dopo un'ampia panoramica introduttiva, illustrerò una tecnica che consente di ottenere semplici formulazioni estese per una vasta classe di problemi. Metterò in evidenza potenzialità e limiti di questo approccio. (Lavoro in collaborazione con M. Conforti, F. Eisenbrand e L. Wolsey)

---

31 ottobre 2007

Metodi di viscosità per la riduzione dell'ordine di sistemi di controllo singolarmente perturbati

Gabriele Terrone (dottorato in Matematica Pura)

Si considera un sistema di due equazioni differenziali ordinarie per la coppia di variabili $(x(t), y(t))$. L'evoluzione di $x(t)$ e $y(t)$ avviene su due scale temporali differenti: la velocità delle variabili "veloci" $y(t)$ è proporzionale ad un parametro positivo $(\epsilon)^{-1}$. Si determina una dinamica "limite", per le sole variabili "lente" $x(t)$, che rappresenta il comportamento del sistema originario quando ϵ tende a zero. Si mostra anche che il sistema limite è in grado di fornire informazioni sulla stabilità del sistema originario.

---

14 novembre 2007

Sistemi disordinati: vetri di spin e polimeri diretti

Agnese Cadel (dottorato in Matematica Computazionale)

Si parla di sistemi disordinati (o complessi) quando sono presenti eterogeneità a livello microscopico e per questo manifestano una ricca varietà di comportamenti. Dopo una breve introduzione alla meccanica statistica dei sistemi complessi, parleremo dei due più famosi esempi di questo tipo di sistemi: i vetri di spin e i polimeri diretti.

---

28 novembre 2007

Sistemi dinamici e insiemi di Aubry-Mather

Olga Bernardi (assegnista in Matematica Pura)

Un sistema dinamico consiste di uno spazio delle fasi che descrive gli stati permessi ad un sistema e di una legge che definisce l'evoluzione temporale di questi stati. L'evoluzione può essere continua, come per le equazioni differenziali, o discreta, come per le mappe. Nello studio dei sistemi dinamici un ruolo fondamentale è svolto dagli insiemi invarianti per la dinamica. Dopo una introduzione per non-esperti ai sistemi dinamici, si definiscono gli insiemi invarianti di Aubry-Mather per una classe di mappe quasi-integrabili in 2 dimensioni e si discute la loro localizzazione tramite tecniche di regolarizzazione ispirate alle teorie di viscosità.

---

12 dicembre 2007

Un calcolo logico per la computazione quantistica

Paola Zizzi (dottorato in Matematica Pura)

Il calcolo dei sequenti (LK), un sistema di deduzione logica introdotto da Gentzen inizialmente per la logica Classica, e in seguito esteso alla logica Intuizionistica (LJ), esiste oggi anche per le logiche sub-strutturali, come la logica Lineare di Girard, e la logica di Base di Sambin. In questo seminario, dopo una prima parte introduttiva, ci proponiamo di introdurre un adeguato calcolo dei sequenti per la computazione quantistica (finora descritta solo in termini di reti di cancelli logici quantistici). Dai risultati finora ottenuti, sembra che il calcolo dei sequenti della logica di Base, possa, con le opportune modifiche, servire a tale scopo.

---

16 gennaio 2008

Metodi di teoria del potenziale per l'analisi di problemi col dato al bordo singolarmente perturbati

Matteo Dalla Riva (dottorato in Matematica Pura)

Si considererà un problema con dato al bordo definito su un aperto limitato dello spazio Euclideo 3-dimensionale. Tale aperto avrà un buco al suo interno. Il nostro scopo è di descrivere il comportamento della soluzione del problema con dato al bordo quando il buco collassa ad un punto. Problemi di questo genere sono stati lungamente studiati tramite le tecniche dell' "analisi asintotica" (si vedano ad esempio i lavori di Keller, Kozlov, Movchan, Maz'ya, Nazarov, Plamenewskii, Ozawa e Ward). Illustreremo in un facile esempio quale tipo di risultato possiamo attenderci applicando tali tecniche. Poi mostreremo il risultato che si ottiene tramite l'approccio alternativo proposto da Lanza de Cristoforis in alcuni lavori a partire dal 2001 e metteremo in luce le principali differenze tra i due risultati.

---

30 gennaio 2008

Algorithms for the computation of the joint spectral radius

Cristina Vagnoni (dottorato in Matematica Computazionale)

The asymptotic behaviour of the solutions of a discrete linear dynamical system is related to the spectral radius R of its associated family F; in particular, a system is stable if R ¡= 1 and there exists an extremal norm for F. In the last decades some algorithms have been proposed in order to find real extremal norms of polytope type in the case of finite families. However, recently it has been observed that it is more useful to consider complex polytope norms. In this talk we show an approach based on the notion of "balanced complex polytopes"; due to the strong increase in complexity of the geometry of such objects, the exposition will be confined to the two-dimensional case. In particular, we give original theoretical results on the geometry of two-dimensional balanced complex polytopes in order to present two efficient algorithms, one for the geometric representation of a balanced complex polytope and the other the computation of the corresponding complex polytope norm of a vector.

---

13 febbraio 2008

Cluster Algebras: an overview

Giovanni Cerulli Irelli (dottorato in Matematica Pura)

Cluster algebras were introduced in 2001 by S. Fomin and A. Zelevinsky with the aim of studying total positivity and canonical basis in semi-simple algebraic groups. After its introduction, the theory has been developed in several unexpected fields of mathematics, e.g. quiver representations, Grassmannians and projective configurations, a new family of convex polytopes (generalized associahedra) including as a special case Stasheff's associahedron, Al. Zamolodchikov's Y-systems in thermodynamic Bethe Ansatz, discrete dynamical system, Teichmuller spaces and Poisson geometry,etc.. In this talk I will recall the definition of such algebraic structure and I will give some motivating examples arising from algebraic geometry.

---

27 febbraio 2008

Computing VaR and CVaR for energy derivatives

Giorgia Callegaro (Matematica Computazionale, S.N.S. di Pisa)

The aim of the talk is to give an idea of the possible applications of mathematics to energy derivatives markets, when computing the risk related to an investment in such a market. First of all we will introduce the notion of derivative asset, starting with an analysis of the basic cases of Call and Put options and arriving to the more complicated swing option case, that are all financial products generally traded on option markets all over the world, with an "underlying" that can be anything, from foreign currencies to stocks, oranges, gas or timber. We will explain how the underlying price dynamics are modeled in energy markets, in basic cases and we will present the problems of "pricing" a derivative and computing the risk related to an investment. In particular, focusing on the gas market, we will explain how the fair price of swing options can be (numerically) computed, by applying the Dynamic Programming Principle and the vectorial quantization. In the same setting, we will also obtain numerical estimates, by means of stochastic recursive algorithms of the Robbins-Monro type, for two different risk measures, namely the "Value at Risk" (VaR) and the "Conditional Value at Risk" (CVaR). (Keywords: swing option, dynamic programming, quantization, risk measure, Robbins-Monro algorithms.)

---

12 marzo 2008

Subanalytic sheaves and D-modules

Luca Prelli (assegnista in Matematica Pura)

This seminar is divided in two parts. We start with an introduction to sheaf theory and then we define sheaves on the subanalytic site. Thanks to these objects we can describe functional spaces which are not defined by local properties (as tempered distributions). In the second part we introduce the notion of D-modules to apply the preceding constructions.

---

2 aprile 2008

Numerical modeling for convection-dominated problems

Manolo Venturin (borsista in Matematica Computazionale)

During the last years, there has been a great interest in the development of sophisticated mathematical models for the simulation of real life applications which involves convection-dominated phenomena. For example, these problems concern the solution of scalar advection-diffusion equations, the Navier-Stokes equations and the Shallow Water equations. The main goal of this seminar is to review the most important difficulties that arise in the numerical approximation of this kind of problems when convection dominates the transport process. Moreover, we present a method for the treatment of this equations with the use of the finite element discretization on the domain.

---

16 aprile 2008

The Basic Picture on sets evaluated over an overlap algebra

Paola Toto (dottorato in Matematica Pura - Università del Salento)

In his forthcoming book, G. Sambin introduces a new topological theory, called "The Basic Picture". In this theory both the notion of topological space and its point-free version are generalized. The concept of overlap algebra is also introduced in order to put in algebraic form the properties needed to define the new topological structures. In this seminar we shall give a tutorial introduction to our work, whose ultimate goal is to generalize such topological notions in the context of many-valued sets. In many-valued set theory sets are built by using propositions evaluated in an algebraic structure. To reach our goal a key point is to check whether the original algebrization of Sambin's topological notions can be considered also as the algebrization of their many-valued version. We prove that this is the case if and only if we take an overlap algebra as the underlying structure of truth values.

---

29 aprile 2008

Quiver mutation and derived equivalence

Bernhard Keller (professor at Paris 7)

[The seminar will be divided in two parts of 40' each, the first of which, of introductory type, will be suitable for a large public] 1. In the first part, we will define and study quiver mutation. This is an elementary operation on quivers (=oriented graphs) which was introduced by Fomin and Zelevinsky in the definition of cluster algebras at the beginning of this decade. The combinatorics behind quiver mutation are rich and varied. We will illustrate them on numerous examples using computer animations. 2. In the second part of the talk, we will "categorify" quiver mutation using representation theory. More precisely, by combining recent work of Derksen-Weyman-Zelevinsky

and Ginzburg, we will show how quiver mutations give rise to equivalences between derived categories of certain differential graded algebras. These derived categories are closely related to cluster categories and thus to cluster algebras. This is joint work with Dong Yang.

---

30 aprile 2008

Chaotic phenomena described by stochastic equations

Luigi Manca (grant holder, Dip. Mat.)

It is well known that many natural phenomena such as population dynamics, stock exchange, diffusion of particles, can be seen as 'chaotic'. To give a mathematical description of these 'chaotic' phenomena has been developed the theory of stochastic processes and of the related stochastic differential equations. Starting by the fundamental concept of Brownian motion, I shall introduce the main ideas and the basic tools in order to understand some easy models driven by stochastic equations. Moreover, I shall describe how stochastic equations can be used to study some deterministic model.

---

14 maggio 2008

Length of curves and surface measures

Roberto Monti (researcher, Dip. Mat.)

We discuss different definitions for the length of a curve and for the area of a hypersurface in the Euclidean space and in more general metric spaces. The talk has an expository character and is an introduction to Geometric Measure Theory.

---

28 maggio 2008

An Introduction to Stochastic Fluid Dynamic Models

David Barbato (researcher, Dip. Mat.)

The Navier-Stokes problem, still unsolved by more than 150 years, represents the starting point for lots of mathematical research topics. The aim of the talk is to present selected fluidodynamic models, in the deterministic and stochastic case, developed from Navier-Stokes equations. In particular the GOY shell model, a Fourier system simplified with respect to the Navier-Stokes one, will be described, and some recent rigorous results discussed. Finally open questions and conjectures on turbolence flows will be presented.

---

11 giugno 2008

An overview on low degree non-abelian cohomology

Pietro Polesello (researcher, Dip. Mat.)

In the first part of the seminar, which will be of introductory level, I will recall some basic facts about the first cohomology set $H^1(X;\underline{G})$, for a given (not necessarily abelian) topological group G, such as the classification of principal G-bundles, the Hurwitz formula and the classification of G-coverings. The second part of the seminar will be devoted to the generalization of some of these results to the "second non abelian cohomology" of G.

Extended formulations for some mixed-integer sets

Marco Di Summa (*)

Abstract. In a mixed-integer programming problem one is required to optimize a linear function over a subset of $\mathbb{R}^n$ defined by a system of linear inequalities, with the additional restriction that some variables must take an integer value. Since subsets of this type are usually very complicated to describe, mixed-integer programming is a hard problem. The introduction of additional variables sometimes leads to a simpler description of the problem (extended formulation) in a higher dimensional space. We present a technique to construct extended formulations for mixed-integer programs with special structure.

Sunto. In programmazione intera mista si deve ottimizzare una funzione lineare su un sottoinsieme di $\mathbb{R}^n$ definito da un sistema di disequazioni lineari, con la condizione aggiuntiva che alcune variabili devono assumere valore intero. A causa della complessità di tali sottoinsiemi di $\mathbb{R}^n$, la programmazione intera mista è un problema di difficile trattazione. A volte l'introduzione di nuove variabili consente di ottenere una descrizione più semplice del problema (formulazione estesa). Illustriamo qui una tecnica per costruire formulazioni estese per problemi di programmazione intera mista aventi una struttura speciale.

1 Introduction

A *mixed-integer (linear) program* is an optimization problem where one is required to minimize (or maximize) a linear function over a subset of $\mathbb{R}^n$ defined by a system of linear inequalities, with the additional restriction that some of the variables must take an integer value. Any mixed-integer program can then be formulated as

$$\min \quad c^\top x \tag{1}$$

$$\text{subject to} \quad Ax \geq b, \tag{2}$$

$$x_i \in \mathbb{Z} \quad \text{for } i \in I, \tag{3}$$

(*)Ph.D. School in Applied Mathematics. Università di Padova, Dipartimento di Matematica Pura ed Applicata, Via Trieste 63, 35121 Padova, Italy. E-mail: mdsumma@math.unipd.it. Seminar held on 17 October 2007.

where A is an $m \times n$ matrix, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ and I is a nonempty subset of $\{1, \ldots, n\}$. In the above problem, $c^\top x$ is the *objective function*, while the set defined by conditions (2)–(3) is the *feasible region*. Variables x_i for $i \in I$ are called the *integer variables*, x_i for $i \notin I$ are the *continuous variables*. A subset of $\mathbb{R}^n$ that is the feasible region of a mixed-integer program is called a *mixed-integer set*.

When $I = \{1, \ldots, n\}$, problem (1)–(3) is a *pure integer (linear) program* (or simply *integer program*). Thus we view integer programs as special types of mixed-integer programs. A problem of the form (1)–(2), with no integrality restrictions, is a *linear program*.

Linear and mixed-integer programming are fundamental areas of operations research. A large number of real-world problems can be formulated as linear or mixed-integer programs, such as problems arising in transportation, manufacturing, scheduling, production planning and many other fields (see e.g. [8,13,14]).

While linear programming is a tractable problem, mixed-integer programming is difficult in general, as the region defined by conditions (2)–(3) is usually very complicated to describe. In some special cases, the introduction of new variables in the problem allows one to give a simpler description of a mixed-integer set. A description of this type, which is given in a higher dimensional space, is called an *extended formulation* of the set (a more precise definition is given in Section 2.1).

In this note we consider mixed-integer sets (2)–(3) whose constraint matrix A has some special structure that we will specify later. We present a technique that allows one to construct extended formulations for an arbitrary set having such a structure.

The rest of the note is organized as follows. In Section 2 we recall some basic results linking mixed-integer programming to linear programming and we also introduce the concept of extended formulation. In Section 3 we define the family of sets that is the object of this study and present our technique to construct an extended formulation for any set in this class. We conclude in Section 4 with some final comments.

2 Links between mixed-integer programming and linear programming

Recall that a *polyhedron* is the set of solutions to a linear system of the type $Ax \leq b$ for some matrix A and some vector b. Thus the feasible region of any linear programming problem is a polyhedron.

Consider now a mixed-integer program (1)–(3) and let X denote its feasible region, i.e. X is the mixed-integer set defined by conditions (2)–(3). We denote by $\mathrm{conv}(X)$ the *convex hull* of X, i.e. the set of points that are convex combinations of points in X.

It can be easily shown that problem (1)–(3) is equivalent to the optimization problem

$$\min \left\{ c^\top x : x \in \mathrm{conv}(X) \right\}. \tag{4}$$

The following result gives some useful information about the set $\mathrm{conv}(X)$.

Theorem 1 [10] *If all entries of A and b are rational numbers,(*) the convex hull of (2)–(3) is a polyhedron.*

(*)The hypothesis of rationality is standard in optimization and complexity theory.

Thus, under the assumption of rationality, (4) is the problem of optimizing a linear function over a polyhedron. If the linear inequalities defining such a polyhedron are explicitly known, problem (4) is a linear program and thus the original mixed-integer programming problem (1)-(3) can be solved by means of linear programming algorithms.

Unfortunately the polyhedron $\text{conv}(X)$ may be defined by a number of inequalities which is exponential in the size of the original system (2), and it is usually very hard to characterize them. Thus the approach described above does not result (in general) in an algorithm for solving mixed-integer programming problems.

2.1 Extended formulations

The above discussion indicates that characterizing the convex hull of a mixed-integer set by means of linear inequalities is a major task in mixed-integer programming. We point out here how such a characterization can be given in a higher dimensional space by using additional variables. To do this, we need the concepts of projection and extended formulation.

Given a set Q in the space $\mathbb{R}^{n+p}$ (with variables $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^p$), the *projection* of Q onto the space of the x-variables is the set of points $x \in \mathbb{R}^n$ that can be completed to a vector (x, y) of Q:

$$\text{proj}_x(Q) := \{x \in \mathbb{R}^n : \text{there exists } y \in \mathbb{R}^p \text{ such that } (x, y) \in Q\}.$$

The projection of a polyhedron is always a polyhedron (see e.g. [18]).

Given a polyhedron P in the space $\mathbb{R}^n$ (with variables x), an *extended formulation* of P is a system of linear inequalities $A'x + B'y \leq d'$ defining a polyhedron Q in a space $\mathbb{R}^{n+p}$ (with variables x and y) such that $P = \text{proj}_x(Q)$.

Extended formulations are useful because a polyhedron that is defined by a huge number of inequalities in its original space may admit a smaller description in an extended space.

We pointed out that given a mixed-integer set X defined by conditions (2)–(3), the structure of the polyhedron $\text{conv}(X)$ is usually very complicated to describe in its original space. However the introduction of additional variables sometimes lead to a simpler description of $\text{conv}(X)$ in an extended space. If such an extended formulation of $\text{conv}(X)$ is known and is compact (i.e. it uses a number of variables and inequalities that is polynomial in the size of the original system $Ax \leq b$), then the mixed-integer program (1)–(3) can be solved in polynomial time by solving the equivalent linear program (4) in the extended space.

Theorem 2 *Let X be the mixed-integer set defined by (2)–(3). If a compact extended formulation of* $\text{conv}(X)$ *is known, the mixed-integer program (1)–(3) can be solved in polynomial time by means of linear programming algorithms.*

3 A technique to construct extended formulations

We define the family MIX^{2TU} as the class of mixed-integer sets (2)–(3) such that the matrix A satisfies the following conditions:

- A is *totally unimodular* (TU), i.e. the determinant of every square submatrix of A is 0, 1 or -1;
- each row of A contains at most two nonzero entries.

In this section we illustrate how to construct an extended formulation for the convex hull of any set in this family.

Our interest in the former of the above two conditions comes form the fact that TU matrices play an important role in pure integer programming. Specifically, in the pure integer case, if A is TU and b has integer components, then the convex hull of the set is obtained by just dropping the integrality requirements [9]. It is then natural to investigate this class of matrices in the mixed-integer case.

Apart from the above theoretical motivation, sets of type MIX^{2TU} are interesting because they include several models appearing in practical problems such as production planning problems [2, 3, 5, 7, 16].

Given any set X in the family MIX^{2TU}, the construction of the extended formulation is based on the enumeration of all the possible fractional parts that the continuous variables take over the vertices of $\text{conv}(X)$. Our approach is an abstraction of an idea used recently by some authors to tackle some specific mixed-integer sets [11, 12, 15, 17]. The details of the process sketched below can be found in [1] and [6].

We use the following notation: given a real number α, we denote by $f(\alpha)$ the fractional part of α, i.e. $f(\alpha) := \alpha - \lfloor \alpha \rfloor$. Define $N := \{1, \dots, n\}$. It can be shown that (possibly after changing the sign of some variables) every set X of the type MIX^{2TU} can be written in this form:

$$x_i - x_j \geq l_{ij}, \quad (i,j) \in N^e, \tag{5}$$

$$x_i \geq l_i, \quad i \in N^l, \tag{6}$$

$$x_i \leq u_i, \quad i \in N^u, \tag{7}$$

$$x_i \in \mathbb{Z}, \quad i \in I, \tag{8}$$

where $N^e \subseteq N \times N$ and $N^l, N^u, I \subseteq N$. The values l_{ij}, l_i, u_i are arbitrary real numbers. We remark that the above system may also include constraints of the type $x_i - x_j \leq u_{ij}$, as this inequality is equivalent to $x_j - x_i \geq l_{ij}$ for $l_{ij} := -u_{ij}$.

Suppose we are given a list of decreasing *fractional parts* $\mathcal{L} = \{f_1, \dots, f_k\}$, i.e. $1 > f_1 > \cdots > f_k \geq 0$, satisfying the following property:

$$\text{If } \bar{x} \text{ is a vertex of } \text{conv}(X), \text{ then } f(\bar{x}_i) \in \mathcal{L} \text{ for each } i \in N \setminus I. \tag{9}$$

A list $\mathcal{L}$ satisfying the above property is called a *complete* list for X. (We will discuss in Section 3.1 how such a list can be found.)

Let $X^{\mathcal{L}}$ be the set of points x satisfying (5)–(8) along with the additional condition that each continuous variable takes a fractional part in $\mathcal{L}$:

$$X^{\mathcal{L}} := \{x \in \mathbb{R}^n : x \text{ satisfies (5)–(8)}, \ f(x_i) \in \mathcal{L} \text{ for } i \in N \setminus I\}.$$

Define $K := \{1, \ldots, k\}$. It is easy to see that, by introducing additional variables μ^i, δ^i_ℓ for $i \in N \setminus I$ and $\ell \in K$, the set $X^{\mathcal{L}}$ can be modeled by the following constraints:

$$x_i = \mu^i + \textstyle\sum_{\ell=1}^k f_\ell \delta^i_\ell, \quad i \in N \setminus I, \tag{10}$$

$$\textstyle\sum_{\ell=1}^k \delta^i_\ell = 1,\ \delta^i_\ell \geq 0, \quad i \in N,\ \ell \in K, \tag{11}$$

$$x_i - x_j \geq l_{ij}, \qquad (i,j) \in N^e, \tag{12}$$

$$x_i \geq l_i, \qquad i \in N^l, \tag{13}$$

$$x_i \leq u_i, \qquad i \in N^u, \tag{14}$$

$$x_i \in \mathbb{Z}, \qquad i \in I, \tag{15}$$

$$\mu^i, \delta^i_\ell \in \mathbb{Z}, \qquad i \in N \setminus I,\ \ell \in K. \tag{16}$$

In other words, $X^{\mathcal{L}}$ is the projection of the mixed-integer set (10)–(16) onto the space of the x-variables.

By using property (9), one shows that $\text{conv}(X) = \text{conv}(X^{\mathcal{L}})$. Thus a linear inequality description of the convex hull of (10)–(16) is an extended formulation of $\text{conv}(X)$. To find such a linear inequality description, we rewrite constraints (12)–(14) as a linear system with TU matrix involving only integer variables. We then use a standard result on total unimodularity [9] to drop the integrality requirements.

Consider an inequality of type (13) for some $i \in N^l$. If $i \in I$ then the inequality can be tightened to $x_i \geq \lceil l_i \rceil$, as x_i is an integer variable. If $i \notin I$, one can verify that under conditions (10)-(11) and (16), inequality (13) is equivalent to the following:

$$\mu^i + \sum_{\ell : f_\ell \geq f(l_i)} \delta^i_\ell \geq \lfloor l_i \rfloor + 1.$$

Consider now an inequality of type (12) with both $i, j \notin I$. Define $k_{ij} := \max\{\ell : f_\ell + f(l_{ij}) \geq 1\}$. It can be proven that under conditions (10)–(11) and (16), inequality (12) is equivalent to the following linear system:

$$\mu^i + \sum_{\ell : f_\ell \geq f(f_t + f(l_{ij}))} \delta^i_\ell - \mu^j - \sum_{\ell : f_\ell \geq f_t} \delta^i_\ell \geq \lfloor l_{ij} \rfloor + 1, \quad 1 \leq t \leq k_{ij},$$

$$\mu^i + \sum_{\ell : f_\ell \geq f(f_t + f(l_{ij}))} \delta^i_\ell - \mu^j - \sum_{\ell : f_\ell \geq f_t} \delta^i_\ell \geq \lfloor l_{ij} \rfloor, \qquad k_{ij} < t \leq k.$$

The other inequalities of the system (12)–(14) can be modeled similarly.

Using these results, the mixed-integer set (10)–(16) can be equivalently rewritten in the form

$$x_i = \mu^i + \textstyle\sum_{\ell=1}^k f_\ell \delta^i_\ell, \quad i \in N \setminus I, \tag{17}$$

$$Cz \geq d, \tag{18}$$

$$z \text{ integral}, \tag{19}$$

where C is a TU matrix, d is an integral vector and z stands for the vector of variables μ, δ and x_i for $i \in I$. The system $Cz \geq d$ includes constraints (11) as well as the reformulation of inequalities (12)-(14) described above.

Except for equations (17), the above is a *pure* integer set. Since equations (17) just define the value of variables x_i for $i \in N \setminus I$ (and these variable do not appear in the remainder of the system), the well-known result of Hoffman and Kruskal [9] implies that the convex hull of the above set is obtained by dropping the integrality restriction (19). Thus (17)–(18) is an extended formulation of $\text{conv}(MIX^{2TU})$.

3.1 How to find a complete list

The construction of the extended formulation described above relies upon the knowledge of a complete list of fractional parts for a set X of the type MIX^{2TU}. We briefly discuss how such a list can be computed.

Proposition 3 *Let $X := \{x \in \mathbb{R}^n : Ax \geq b, x_i \in \mathbb{Z} \text{ for } i \in I\}$ be a mixed-integer set, where A is an $m \times n$ TU matrix, $b \in \mathbb{R}^m$ and $I \subseteq \{1, \ldots, n\}$. The following list is complete for X:*

$$\mathcal{L} := \left\{ f\left(\textstyle\sum_{j=1}^m \sigma_j b_j\right) : \sigma_1, \ldots, \sigma_m \in \{0, \pm 1\} \right\}. \tag{20}$$

The above result provides a complete list for any set of the type MIX^{2TU}, thus formulation (17)–(18) can be always written explicitly, though the above list is (in general) very long and originates a non-compact extended formulation. Note however that Proposition 3 only exploits the total unimodularity of matrix A. By also using the information that every row of A contains at most two nonzero entries, the above list can be refined. This process is based on the construction of a graph associated with the set X and in many interesting cases (e.g. in mixed-integer sets arising from production planning problems) generates a complete list with only a polynomial number of elements. Thus in these cases formulation (17)–(18) is compact and can be used to optimize in polynomial time.

4 Concluding remarks

The previous sections shows how a formulation of the convex hull of a set X of the type MIX^{2TU} can be given in an extended space. Nonetheless the formulation of $\text{conv}(X)$ in its original space of definition is important as well. Such a formulation can be computed (in principle) by projecting our extended formulation onto the original space of the x-variables. Though the calculation of the projection seems to be a prohibitive task in general, the formulation in the original space was obtained for some specific (classes of) sets of the type MIX^{2TU} ([3,5,6,7,16]).

Finally, a natural question one could ask is whether the above technique can be extended to mixed-integer sets that do not belong to the class MIX^{2TU}. Some results addressing this problem can be found in [4].

References

[1] M. Conforti, M. Di Summa, F. Eisenbrand, and L.A. Wolsey, *Network formulations of mixed-integer programs*. CORE Discussion Paper 2006/117, Université catholique de Louvain, Belgium, 2006. Accepted by Mathematics of Operations Research.

[2] M. Conforti, M. Di Summa, and L.A. Wolsey, *The intersection of continuous mixing polyhedra and the continuous mixing polyhedron with flows*. In Integer Programming and Combinatorial Optimization, M. Fischetti and D.P. Williamson, eds., vol. 4513 of Lecture Notes in Computer Science, Springer, 2007, pp. 352–366.

[3] M. Conforti, M. Di Summa, and L.A. Wolsey, *The mixing set with flows*. SIAM Journal on Discrete Mathematics, 21 (2007), pp. 396–407.

[4] M. Conforti, M. Di Summa, and L.A. Wolsey, *The mixing set with divisible capacities*. In Integer Programming and Combinatorial Optimization, A. Lodi, A. Panconesi and G. Rinaldi, eds., vol. 5035 of Lecture Notes in Computer Science, Springer, 2008, pp. 435-449.

[5] M. Conforti, B. Gerards, and G. Zambelli, *Mixed-integer vertex covers on bipartite graphs*. In Integer Programming and Combinatorial Optimization, M. Fischetti and D.P. Williamson, eds., vol. 4513 of Lecture Notes in Computer Science, Springer, 2007, pp. 324–336.

[6] M. Di Summa, *Formulations of Mixed-Integer Sets Defined by Totally Unimodular Constraint Matrices*. PhD thesis, Università degli Studi di Padova, Italy, 2008.

[7] O. Günlük and Y. Pochet, *Mixing mixed-integer inequalities*. Mathematical Programming, 90 (2001), pp. 429–457.

[8] F.S. Hillier and G.J. Lieberman, "Introduction to Operations Research". McGraw-Hill School Education Group, seventh ed., 2000.

[9] A.J. Hoffman and J.B. Kruskal, *Integral boundary points of convex polyhedra*. In Linear Inequalities and Related Systems, H.W. Kuhn and A.W. Tucker, eds., no. 38 in Annals of Mathematical Study, Princeton Unversity Press, 1956, pp. 223–246.

[10] R.R. Meyer, *On the existence of optimal solutions to integer and mixed-integer programming problems*. Mathematical Programming, 7 (1974), pp. 223–235.

[11] A.J. Miller and L.A. Wolsey, *Tight formulations for some simple mixed integer programs and convex objective integer programs*. Mathematical Programming, 98 (2003), pp. 73–88.

[12] A.J. Miller and L.A. Wolsey, *Tight mip formulation for multi-item discrete lot-sizing problems*. Operations Research, 51 (2003), pp. 557–565.

[13] G.L. Nemhauser and L.A. Wolsey, "Integer and Combinatorial Optimization". Wiley-Interscience, New York, 1988.

[14] Y. Pochet and L.A. Wolsey, "Production Planning by Mixed Integer Programming". Springer, 2006.

[15] M. Van Vyve, *A Solution Approach of Production Planning Problems Based on Compact Formulations for Single-Item Lot-Sizing Models*. PhD thesis, Faculté des Sciences Appliquées, Université catholique de Louvain, Belgium, 2003.

[16] M. Van Vyve, *The continuous mixing polyhedron*. Mathematics of Operations Research, 30 (2005), pp. 441–452.

[17] M. Van Vyve, *Linear-programming extended formulations for the single-item lot-sizing problem with backlogging and constant capacity*. Mathematical Programming, 108 (2006), pp. 53–77.

[18] G.M. Ziegler, "Lectures on Polytopes". Springer, 1995.

Viscosity methods for the order reduction of singularly perturbed control systems

Gabriele Terrone (*)

Abstract. In a singularly perturbed system the state variables evolves along two different time scales: a positive parameter ε appears in front of the time derivative of the *fast* variables. We define a *limit* dynamics, just for the slow states, describing the asymptotic behavior of a singularly perturbed control system, as the parameter ε vanishes. Then, we show that it is possible to infer the asymptotic stabilizability of the two-scale system from the one of the limit dynamics. More precisely, using viscosity theoretical methods, we exhibit a Lyapunov function for the original dynamics, obtained as an ε perturbation of a given Lyapunov function for the limit dynamics.

Sunto. In un sistema singolarmente perturbato le variabili di stato evolvono lungo due differenti scale temporali: un parametro ε positivo appare davanti alla derivata temporale delle variabili veloci. Si vuole determinare una dinamica *limite*, solo per le variabili lente, che descriva il comportamento asintotico del sistema singolarmente perturbato quando il parametro ε svanisce. Quindi si vuole mostrare che è possibile inferire la stabilizzabilità asintotica della dinamica singolarmente perturbata da quella della dinamica limite. Più precisamente, usando metodi della teoria di viscosità, si esibisce una funzione di Lyapunov per la dinamica originaria ottenuta come perturbazione in ε di una funzione di Lyapunov della dinamica limite.

Let us consider the following singularly perturbed control system

$$(S_\varepsilon) \qquad \begin{cases} \dot{x}(t) = f(x(t), y(t), a(t)), & x(0) = x \\ \varepsilon \dot{y}(t) = g(x(t), y(t), a(t)), & y(0) = y \end{cases}$$

where

- $x \in \mathbb{R}^N$ and y belongs to the flat torus $\mathbb{T}^M \simeq \mathbb{R}^M/\mathbb{Z}^M$;
- the functions $a(\cdot)$, the *controls*, are measurable functions from $\mathbb{R}^+$ to a compact metric space A. The set of these functions is denoted by $\mathcal{A}$. We will also write a to denote the elements of A, when no ambiguities can arise;

(*)Ph.D. School in Pure Mathematics. Università di Padova, Dipartimento di Matematica Pura ed Applicata, Via Trieste 63, 35121 Padova, Italy. E-mail: gabter@math.unipd.it. Seminar held on 31 October 2007.

– the functions f and g are continuous from $\mathbb{R}^N \times \mathbb{T}^M \times A$ to $\mathbb{R}^N$ and $\mathbb{T}^M$ respectively, and Lipschitz–continuous in (x, y) uniformly with respect to a.

The role of the positive parameter ε in front of the time derivative of the y variables is apparent: the state variables are divided in two groups, a group of N *slow* variables, evolving on a macroscopic time scale, and a group of M *fast* variables, evolving along a microscopic time scale. This is the reason why the study of multiscale problems is motivated by many phenomena arising in engineering, chemistry and physics.

We shall try to reply to the following questions:

(Q1). Is it possible to find a dynamics $(\bar{S})$ involving only the slow variables, and describing the asymptotic behavior of (S_ε) as $\varepsilon \to 0$?

(Q2). If yes, is it possible to use it to derive any qualitative information for (S_ε)? In particular, if we know that $(\bar{S})$ is stable (in some sense) what can we say about the stability of (S_ε)?

Q1. At the beginning, the Levinson-Tikhonov method has been used to detect, under suitable assumption, the limit dynamics of a singularly perturbed system. This approach consists in considering, as the natural candidate for the limit, the system obtained by setting $\varepsilon = 0$ in (S_ε). The result is an ordinary differential equation combined with an algebraic equation. This approach gives the appropriate limit when the stationary points of the fast dynamics are attractive, a condition that may fail to be satisfied by systems with more general asymptotic behavior of the fast variables. Other averaging approaches have been proposed by Artstein in the context of invariant measure theory (see [3]), and by Gaitsgory and Leizarowitz, using limit occupational measures (see [6]).

Our way to detect the limit system combines the theory of occupational measure and the theory of homogenization of partial differential equations. It consists in considering the *fast subsystem* associated to (S_ε), that is the M–dimensional system

$$\dot{y}(t) = g(x, y(t), a(t)), \quad y(0) = y; \tag{1}$$

here x is frozen and considered as a parameter. We assume that the fast subsystem is *bounded time controllable*, *i.e.* that there any pair of points on $\mathbb{T}^M$ can be connected by an admissible trajectory of (1) in a uniformly bounded time.

An *occupational measure* μ_s for (1) is a Radon probability measure on $\mathcal{B}(\mathbb{T}^M \times A)$ (the Borel σ–algebra of $\mathbb{T}^M \times A$) defined by:

$$\mu_s := \frac{1}{s} \int_0^s \delta_{(y(t),a(t))} dt$$

where $\delta_{(y(t),a(t))}$ is the Dirac mass concentrated on $(y(t), a(t))$. For any Borel set Q, $\mu_s(Q)$ gives the proportion of time spent by the trajectory on Q, with respect to the time interval $[0, s]$. We say that a measure μ on $\mathcal{B}(\mathbb{T}^M \times A)$ is a *limiting relaxed control* if there exist a control function $a \in \mathcal{A}$ and a diverging sequence t_n such that the occupational measure μ_{t_n} of the corresponding solution of (1) converges weak star to μ. We will denote by $M(x)$ the set of all limiting relaxed measures obtained by solutions of (1). For any x the set

$M(x)$ constitutes a convex compact subset of the probability measures on $\mathcal{B}(\mathbb{T}^M \times A)$, and the measures in $M(x)$ are used to replace the dependence of the data by y and a with a dependence by the relaxed control μ:

$$\bar{f}(x,\mu) := \int f(x,y,a)\mu(dy,da), \qquad \mu \in M(x)$$

The limit system is then detected as the differential inclusion

$$(\bar{S}) \qquad \begin{cases} x(t) \in F(x(t)) \\ x(0) = x \end{cases}$$

where

$$F(x) := \left\{ \bar{f}(x,\mu) \,\middle|\, \mu \in M(x) \right\}.$$

Even if the way to obtain the limiting system proposed in [3] and in [6] is actually close to the one described here, the arguments used are different: instead of techniques connected to the theory of invariant measures and occupational measures we will exploit only viscosity theoretical techniques. This is indeed one of the main peculiarity of our contributions. The fact that $(\bar{S})$ provides the correct approximation of (S_ε) for small ε, has been proved in [10] by studying the partial differential equations and the value functions associated to these systems.

It is well known that the *value function* of (S_ε), *i.e.* the function

$$u^\varepsilon(t,x,y) := \inf \left\{ \int_0^t l(x(s),y(s),a(s))ds + h(x(t),y(t)) \right\},$$

where l and h are given functions and the inf is taken among all admissible trajectories of (S_ε), satisfies

$$\partial_t u^\varepsilon + H\left(x, y, D_x u^\varepsilon, \tfrac{1}{\varepsilon} D_y u^\varepsilon\right) = 0$$

$$u^\varepsilon(0,x,y) = h(x,y)$$

where $D_x u^\varepsilon$ and $D_y u^\varepsilon$ stand for the gradient of u^ε with respect to the slow and fast variables respectively, and

$$H(x,y,p,q) = \max_{a \in A} \{-p \cdot f(x,y,a) - q \cdot g(x,y,a) - l(x,y,a)\}.$$

To reply to question **Q1**, instead of passing to the limit in the dynamics, we pass to the limit in the PDE, using the machinery of the homogenization theory. Under suitable conditions, it is possible to define an *effective Hamiltonian* $\bar{H}(x,p)$, and an *effective initial data* $\bar{h}(x)$ such that the u^ε converges locally uniformly, as $\varepsilon \to 0$ to a solution of

$$\partial_t u + \bar{H}(x, D_x u) = 0$$

$$u(0,x) = \bar{h}(x)$$

The proof of the existence of such an operator, and the analysis of some property of it, constitute a wide line of research, going back to the firsts pioneering works on homogenization of PDEs, in particular to the famous unpublished preprint by Lions, Papanicolaou,

and Varadhan. Recently, two crucial properties about the convergence of the u^ε have been singled out by Alvarez and Bardi in [2]. The first is an *ergodicity* property of the operator, and pertains with the definition of the effective Hamiltonian; the second property regards the possibility to define an effective initial datum for the effective Cauchy problem. These two properties also permit to establish the uniform convergence of u^ε to the solution of the effective equation.

It turns out that the effective Hamiltonian is the partial differential operator associated to the limit control problem $(\bar{S})$, in fact it can be proved that, for any x, p

$$\bar{H}(x,p) = \max_{\mu \in M(x)} \left\{ -p \cdot \bar{f}(x,\mu) - \bar{l}(x,\mu) \right\}.$$

Using this representation formula for $\bar{H}$, has been proved that the value function of the effective control problem, *i.e.* the function

$$\bar{u}(t,x) := \inf \left\{ \int_0^t \bar{l}(x(s),\mu(s))ds + \bar{h}(x(t)) \right\},$$

where the inf is taken among all admissible trajectories of $(\bar{S})$, satisfies the effective Cauchy problem. Moreover, under uniqueness of solutions for the effective Cauchy problem, $\bar{u}$ coincides with the limit of the value functions u^ε. Then we derive also an additional interesting information: the local uniform limit of value function of optimal control problems is a value function of another optimal control problem.

Q2. Now we want to establish if stability properties can be deducted for the singularly perturbed system, assuming that the limit system is stable. In this direction few results are available in the literature, essentially in the context of the Levinson-Tikhonov theory; see [7] and [9]. One of the most relevant results has been established, for *non-controlled* systems, by Artstein in [3]. The main result in [3] asserts that if the equilibrium is asymptotically stable for the limit differential inclusion, then the slow part of the singularly perturbed system is asymptotically stable *near* the origin.

We concentrate on *asymptotic stabilizability*, a condition implying the existence of *at least* a trajectory of the the dynamics driving asymptotically the system to a certain target. It is the classical notion of asymptotic stability adapted to control systems. Stability properties are studied by means of Lyapunov pairs, *i.e.* pairs constituted by a Lyapunov function and another function estimating its infinitesimal decrease along integral trajectories driving to a certain target, that we suppose for simplicity to be the origin. We characterize such a monotonicity property with a suitable Hamilton–Jacobi differential inequality, interpreted in viscosity sense. More precisely we say that $(V(x), W(x))$ is a *Control Lyapunov pair* for $(\bar{S})$ with respect to the target $\{x = 0\}$ if V and W are proper and positive definite, and satisfy

$$\bar{H}(x, D_x V) \geq W(x);$$

analogously we say that $(V^\varepsilon(x,y), W^\varepsilon(x,y))$ is a *Control Lyapunov pair* for (S_ε) with respect to the target $\{0\} \times \mathbb{T}^M$ if

$$H\left(x, y, D_x V^\varepsilon, \frac{1}{\varepsilon} D_y V^\varepsilon\right) \geq W^\varepsilon.$$

We want to construct a Lyapunov function V^ε for the singularly perturbed system as a first order perturbation in ε of a given Lyapunov function for the limiting system.

In order to explain our ideas in the simpler way, let us try a formal expansion. Suppose to posses a Control Lyapunov pair (V, W) for the effective dynamics, with V *smooth*, say C^1. For any fixed point $\bar{x}$, let $\chi(y)$ be the solution[(*)] of the equation

$$H(\bar{x}, y, DV(\bar{x}), D\chi) = \bar{H}(\bar{x}, DV(\bar{x}))$$

and set

$$V^\varepsilon(x, y) := V(x) + \varepsilon\chi(y).$$

This function should satisfy

$$H\left(x, y, D_x V^\varepsilon, \frac{1}{\varepsilon} D_y V^\varepsilon\right) = H(x, y, DV(x), D_y\chi(y)) = \bar{H}(x, DV(x)) \geq W(x)$$

at $x = \bar{x}$... Such computation is evidently incorrect, because the function χ depends not only on y but also on x! Therefore in the previous computation, in place of $D_x V^\varepsilon$ should appear also a contribution of χ. Unfortunately, the dependence of χ on x is not clear, and still remains an open question.

The problem is eliminated by considering an alternative corrector. For any fixed a positive radius ρ, we obtain - as a (continuous) viscosity supersolution of a suitable auxiliary problem - a function χ_ρ which is uniform in x in the complement of the ball $B(0, \rho)$. Consequently, the former expansion $V^\varepsilon(x, y) := V(x) + \varepsilon\chi_\rho(y)$ can be used to prove rigorously the desired inequality outside the ball of radius ρ. Since the function V^ε does not satisfy the required monotonicity property in a whole neighborhood of the origin, but in the complement of any arbitrarily small ball, no dynamical consequence can be directly derived by Lyapunov theorems. Nevertheless, exploiting comparison and superoptimality principles for viscosity supersolutions of Hamilton-Jacobi equations (see [8]) we are able to establish the asymptotic stabilizability of the singularly pertutbed system.

Theorem. *If the limiting differential inclusion $(\bar{S})$ is asymptotically stable to the origin, then the slow part of the singularly perturbed system (S_ε) is asymptotically stable to the same state, if ε is small enough.*

References

[1] O. Alvarez, M. Bardi, *Viscosity solutions methods for singular perturbations in deterministic and stochastic control.* SIAM J. Control Optim. Vol. 40, No. 4 (2001), 1159–1188.

[2] O. Alvarez, M. Bardi, *Singular Perturbations of Nonlinear Degenerate Parabolic PDEs: a General Convergence Result.* Arch. Rational Mech. Anal. Vol. 170 (2003), 17–61.

[3] Z. Artstein, *Stability in presence of singular perturbations.* Nonlinear Analysis Vol.33 (1998), 817–827.

[(*)] The function χ is usually called *first corrector*, and the equation solved by χ is called *cell problem.*

[4] M. Bardi, I. Capuzzo Dolcetta, "Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellmann equations". Birkhäuser, Boston, 1997.

[5] L. C. Evans, *The perturbed test function method for viscosity solutions of nonlinear PDE.* Proceedings of the Royal Society of Edinburgh Vol.111A (1989), 359–375.

[6] V. Gaitsgory, A. Leizarowitz, *Limit Occupational Measures Set for a Control System and Averaging of Singularly Perturbed Control System.* Journal of Mathematical Analysis and Applications Vol. 233 (1999), 461–475.

[7] P. Kokotović, H. Khalil, J. O'Reilly, "Singular Perturbation Methods in Control: Analysis and Design". Academic Press, London, 1986.

[8] P. Soravia, *Optimality principles and representation formulas for viscosity solutions of Hamilton-Jacobi equations. II. Equations of control problems with state constraints.* Differential and Integral Equations Vol. 12 no. 2 (1999), 275–293.

[9] A.R. Teel, L. Moreau and D. Nešić, *A unified framework for input-to-state stability in systems with two time scales.* IEEE Trans. Automat. Control Vol. 48 no. 9 (2003), 1526–1544.

[10] G. Terrone, "Singular Perturbation and Homogenization Problems in Control Theory, Differential Games and fully nonlinear Partial Differential equations". Ph.D. Thesis. Università degli Studi di Padova, 2007.

Spin glasses and directed polymers

Agnese Cadel [(*)]

1 Introduction: statistical mechanics

Statistical mechanics is a theory that attempts to explain the behaviour of systems that are composed of many individual components, like gases, liquids or crystalline solids. According to this theory a system in equilibrium is described by an energy functional, the *Hamiltonian*, which associates a macroscopic energy to each microscopic configuration of the system. The principal aim of this theory is to obtain the general laws of thermodynamics and the thermodynamic functions associates to the system.
In particular we are interested in the study of the *free energy*, that is the amount of thermodynamic energy that can be converted into work at a constant temperature and pressure. Two facts are known about free energy: in a system with constant temperature and pressure the free energy does not increase, and in such a system the minimum of the free energy represents the equilibrium.

One speaks of disordered (or complex) system when the dynamics, or the structures that appears within the system, exhibits a rich variety of behaviours, while the microscopic entities the system is made of, and the interactions among these entities, are a priori simple.
In the last years a lot of examples of systems characterised by disordered molecular aggregation have been found. Two typical examples are **spin glasses**, *i.e.* metals containing random magnetic impurities, like the blue gold (an alloy of gold and iron), and **polymers**, *i.e.* chemical compounds consisting of repeating units called monomers, like DNA or RNA.

Suppose that we want to study the blue gold (AuFe). We are in a lattice and some sites will be occupied by the iron atoms, while others by the golden ones. If, for example, we heat the system, the atoms will become mobile and will change places, but at low temperatures this motion of the atoms is suppressed, and if we wait, even for a large time, the macroscopic realization of the material will not change. In this case we say that the positions of the atoms are 'frozen' and the system will not be in thermal equilibrium.

However, if we are interested in the *magnetic properties* of the system, we don't have

[(*)]Ph.D. School in Applied Mathematics. Università di Padova, Dipartimento di Matematica Pura ed Applicata, Via Trieste 63, 35121 Padova, Italy. E-mail: agnese@math.unipd.it . Seminar held on 14 November 2007.

to look at the position of the atoms, but at their magnetic moments, to be more precise, we have to look at the orientation of these moments. On the other hand we have to remember that to make a good description of the system we can not forget the position of the iron atoms because the interactions among them depend on their distance. The problem is that it is very difficult to study all possible systems for all possible placements of these atoms. Hence we have to suppose that only certain properties are important and that the microscopic details of these arrangements are negligible, *i.e.* we want to model the disordered system as a random model, by introducing a probability measure on the space of the possible realizations of the positions of the iron atoms.

2 Spin glasses

The movement of the electrons around the nucleus generates a microscopic magnetic moment called *spin*. This spin can be seen has a vector in a 3-dimensional space and for it only two directions are allowed: *up* or *down*. In ferromagnets each spin has a tendency to align with the one in its proximity. At high temperature the motion of the spins is so erratic that almost half of them are pointing up and the other half down, thus the net macroscopic magnetization is zero (the microscopic magnetic field generated by each spin cancel each other out). At low temperature, however, the spins become more sensible to their mutual interaction because their erratic movement is reduced. The fundamental feature of ferromagnets is that there exists a critical temperature T_c below which the spins exhibit a collective behaviour in that a majority of them point in the same direction (up or down). This is called *spontaneous magnetization*.
Let Λ be lattice, then at each site i we assign a variable σ_i (the spin) such that $\sigma_i \in \{-1, 1\}$ and we can define the Hamiltonian as

$$-H_\Lambda(\sigma) = \sum_{\substack{i,j\in\Lambda \\ i\sim j}} \sigma_i\sigma_j,$$

where $i \sim j$ means that i and j are neighbouring sites. At a given temperature T the state of the system is described by the *Gibbs' measure* associated with the Hamiltonian

$$G_\Lambda(\sigma) = \frac{1}{Z_\Lambda}\exp\left(-\frac{H_\Lambda(\sigma)}{\kappa T}\right)$$

where k is the Boltzmann constant and Z_Λ is the normalizing factor that makes G_Λ a probability measure: it is the probability to find the system at the given configuration σ. At low temperature the Gibbs' measure is peaked around the configuration of minimal energy of the system (the so called ground states). In the case of ferromagnets these configurations are those in which the spins have the same value.

In the case of *spin glasses*, instead, some pair of neighbouring spins want to be aligned, but some others prefer to be anti-aligned. In the first case we'll say that the interaction is ferromagnetic, while for the latter we talk about anti-ferromagnetic interaction. For any pair of spins the type of interaction is chosen randomly with the same probability. As a consequence, in the Hamiltonian we can't simply consider the interaction among pairs of

spins (like we did for the pure ferromagnets) but we have to consider a random variable to show the type of the interaction (ferromagnetic or anti-ferromagnetic) among them. Thus the Hamiltonian becomes

$$-H_\Lambda(\sigma) = \sum_{\substack{i,j\in\Lambda \\ i\sim j}} J_{i,j}\sigma_i\sigma_j \tag{1}$$

where $J_{i,j}$ is a random variable that takes values 1 or -1 with the same probability. Besides the presence of the randomness in the Hamiltonian, in order to have a spin glass we need another feature: there must be *frustration*. We say that a system is frustrated when the Hamiltonian cannot be written as the sum of many terms, all of which can be minimized by a single ground state configuration.
To see this let us make an example: consider a group of N people and suppose that each person knows each other. We also assume that any couple of individuals can be either friends or enemies and that the friendship-enmity relations are assigned randomly and independently for each couple. Now, one wants to divide the N individuals into two parties, so as to minimize social discomfort, *i.e.* putting as much as possible friends together and enemies apart. The system is obviously frustrated because if A is a friend of B and B a friend of C, we can say nothing about the relationship between A and C. Thus there are 2^N possible ways the N people can be divided. Let us assign at each person the variable

$$\begin{cases} \sigma_i = 1 \quad \text{if } i \text{ is in the first group} \\ \\ \sigma_i = -1 \quad \text{if } i \text{ is in the second group} \end{cases} \tag{2}$$

Furthermore, given a pair of people i and j set $J_{i,j} = 1$ if they are friends or $J_{i,j} = -1$ if they are enemies. Thus the problem to find the optimal division of the group is equivalent to find the minimum of the 'cost function'

$$-H_N(\sigma) = \sum_{1\le i<j\le N} J_{i,j}\sigma_i\sigma_j,$$

that is almost our Hamiltonian, except that in this case we consider the sum over *all* the 2^N configurations. Notice that in this example the role of disorder is played by the random choice of the relation, friendship or hostility.

The Hamiltonian (1) was introduced by Edwards and Anderson. The Edwards-Anderson model is one of the most difficult models to analyze, from both the analytical and numerical point of view. This is due to frustration: it is non-trivial to say something about its *ground states*. The reason is that the couplings take both signs, favouring alignment or non-alignment of the spins and it is clearly impossible to satisfy the demands of all couplings. It seems quite natural, then, to use approximations. *Mean fields* theory is widely used in such situations. In other words, we start from the study of a simplified model maintaining the two fundamental features of disorder and frustration: the geometry of the lattice is broken so that every magnetic moment interacts with *all* others (and not only with the neighbouring ones, like in the previous example).

The Edwards-Anderson model was modified using the mean field approximation by Sherrington and Kirkpatrick. The Hamiltonian, that for sake of simplicity we consider dependent also on the inverse of the temperature, is

$$-H_N(\sigma) = \frac{\beta}{\sqrt{N}} \sum_{1 \le i < j \le N} g_{i,j}\sigma_i\sigma_j,$$

where the $g_{i,j}$ are standard Gaussian random variables and $\beta = (\kappa T)^{-1}$. We can point out that $H_N(\sigma)$ is a family of centered Gaussian random variables, characterized by the covariance function

$$\mathbf{Cov}(H_N(\sigma^1)H_N(\sigma^2)) = \frac{1}{N} \sum_{1 \le i,j \le N} \sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2.$$

In my thesis we studied a new version of the SK-model, where we considered an additional ferromagnetic interaction. Then the Hamiltonian becomes

$$-H_N(\sigma) = \frac{\beta_1}{2N}\left(\sum_{i \le N}\sigma_i\right)^2 + \frac{\beta_2}{\sqrt{N}} \sum_{1 \le i < j \le N} g_{i,j}\sigma_i\sigma_j + h\sum_{i \le N}\sigma_i,$$

where we introduced the interaction of an external magnetic field $h > 0$ (that favours the + spins over the - ones) and where β_1 and β_2 represent the inverse of two different temperatures. We studied the behaviour of the system in the thermodynamical limit, *i.e.* $N \to \infty$, in particular we found a limit for the free energy, that in this model is defined by

$$p_N(\beta) = \frac{1}{N}\mathbf{E}\left[\log Z_N\right].$$

3 Polymers

A *polymer* is a chain of monomers linked by chemical bonds and these monomers can be either single atoms or molecules. As we said, typical examples are DNA, RNA or lipids. According to this definition, an important and natural example of a polymer model is given by a d-dimensional random walk where the monomers are thought of as increments. In order to simplify, we want to suppress entanglement and U-turns of the polymer, so we would like to deal with *self-avoiding walks*. In the lattice case, by this we mean a random walk that cannot visit again the sites it has already visited. The problem is that it is very difficult to deal with this kind of walks. To avoid this, we impose a simpler constraint: we work with *directed walks*, *i.e.* we force one of the coordinates to be strictly increasing. Thus the polymer is supposed to live in $(1+d)$-dimensional lattice and to stretch in the direction of the first coordinate.

In my thesis we focused on directed polymers in a random environment, which can be thought of as paths of stochastic processes interacting with a quenched disorder, depending on both time and space. Each path is weighted not only according to its length, but also according to the random impurities (disorder) that it meets along its route.

We first considered a Brownian polymer in a Gaussian environment: the polymer itself is modeled by a Brownian motion $b = \{b_t; t \geq 0\}$ and the random environment is represented by a centered Gaussian random field W, defined on another independent complete probability space. Once b and W are defined, the polymer measure itself can be described as follows: for any $t > 0$, the energy of a given path (or configuration) b on $[0, t]$ is given by the *Hamiltonian*

$$-H_t(b) = \int_0^t W(ds, b_s).$$

Based on this Hamiltonian, for any $x \in \mathbb{R}^d$, and a given constant β (that is again the inverse of the temperature of the system), we define our (random) polymer measure G_t^x as follows:

$$dG_t^x(b) = \frac{e^{-\beta H_t(b)}}{Z_t^x} dP_b^x(b), \quad \text{with} \quad Z_t^x = E_b^x \left[e^{-\beta H_t(b)} \right].$$

The second model we considered is the continuous time random walk on $\mathbb{Z}^d$ in a white noise potential, which can be defined similarly to the Brownian polymer above: the polymer is modeled by a continuous time random walk $\hat{b} = \{\hat{b}_t; t \geq 0\}$ on $\mathbb{Z}^d$, defined on a complete filtered probability space and the random environment $\hat{W}$ will be defined as a sequence $\{\hat{W}(., z); z \in \mathbb{Z}^d\}$ of Brownian motions, defined on another independent complete probability space.

Also in the case of polymers we are interested in the *free energy*, that in our models can be defined similarly as

$$p(\beta) = \lim_{t \to \infty} \frac{1}{t} \mathbf{E} \left[\log(Z_t) \right], \quad \text{and} \quad \hat{p}(\beta) = \lim_{t \to \infty} \frac{1}{t} \mathbf{E} \left[\log(\hat{Z}_t) \right].$$

The most important question in these models is:

how does the disorder affect the global shape of the polymer?

The answer for the random walk-type models can be find for example in [5] and it is

- If $d \geq 3$ and β small enough, the impurities do not affect the global shape of the polymer and we say that the polymer is in the weak disorder phase.
- If either
  (i) $d \leq 2$ and $\beta \neq 0$ or
  (ii) $d \geq 3$ and β large enough
  then the impurities change drastically the global shape of the polymer and we say that the polymer is in the strong disorder phase.

It is known that the free energy $p(\beta)$ is bounded from above by $Q(0)\beta^2/2$. It is then possible to separate the regions of *strong* and *weak* disorder according to the value of $p(\beta)$: we will say that the polymer is in the *weak* disorder regime if $p(\beta) = Q(0)\beta^2/2$ while the *strong* disorder phase is defined by $p(\beta) < Q(0)\beta^2/2$. Besides it is expected for any model of polymer in a random media that the strong disorder regime is attained whenever β is large enough. In my thesis we obtain some sharp estimates on the free energies $p(\beta)$ and $\hat{p}(\beta)$ of the two systems described above, for large β.

References

[1] X. Bardina, D. Márquez-Carreras, C. Rovira, S. Tindel, *The p-spin interaction model with external field*. Potential Analysis 21 (2004), 311–362.

[2] A. Bovier, "Statistical Mechanics of Disordered Systems". Cambridge Series in Statistical and Probabilistic Mathematics, 2006.

[3] P. Carmona, Y. Hu, *On the partition function of a directed polymer in a Gaussian random environment*. Probab. Theory Relat. Fields, 124 (2002), 431–457.

[4] P. Carmona, Y. Hu, *Strong disorder implies strong localization for directed polymers in a random environment*. ALEA Lat. Am. J. Probab. Math. Stat. 2 (2006), 217–229.

[5] F. Comets, T. Shiga, N. Yoshida, *Probabilistic analysis of directed polymers in a random environment: a review*. Adv. Stud. Pure Math. 39 (2003), 115–142.

[6] S. F. Edwards, P. W. Anderson, *Theory of spin glasses*. J. Phys. F. no 5 (1965).

[7] G. Giacomin, "Random Polymer models". Imperial College Press, World Scientific, 2007.

[8] H. Nishimori, "Statistical Physics of Spin Glasses and Information Processing". Oxford University Press, 2001.

[9] C. Rovira, S. Tindel, *On the Brownian-directed polymer in a Gaussian random environment*. J. Funct. Anal. 222 no. 1 (2005), 178–201.

[10] D. Sherrington, S. Kirkpatrick, *A solvable model of a spin glass*. Phys. Rev. Lett. 35 (1975), 1792–1796.

[11] M. Talagrand, "Spin Glasses: A Challenge for Mathematicians". Springer-Verlag, 2003.

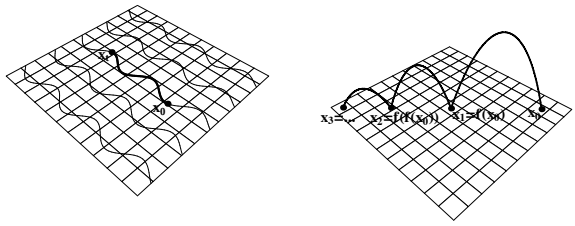
Sistemi dinamici e insiemi di Aubry-Mather

Olga Bernardi (*)

Sunto. Dopo una breve introduzione per non-esperti ai sistemi dinamici, si definiscono gli insiemi invarianti di Aubry-Mather per una classe di mappe quasi-integrabili in 2 dimensioni (mappa standard) e si discute la loro localizzazione tramite tecniche di regolarizzazione ispirate alle teorie di viscosità.

1 Introduzione

Un sistema dinamico consiste di uno spazio delle fasi che descrive gli stati permessi ad un sistema e di una legge che definisce l'evoluzione temporale di questi stati. L'evoluzione può essere continua, come per le equazioni differenziali, oppure discreta, e in tal caso è fornita dall'iterazione di una mappa:



In linea di principio, ogni modello fisico è un sistema dinamico; tuttavia, molti di essi sono sistemi dinamici Hamiltoniani che forniscono mappe simplettiche.

Questioni fondamentali di interesse fisico nello studio di un sistema dinamico includono la stabilità delle soluzioni (intesa come dipendenza dai dati iniziali), la stabilità strutturale del sistema dinamico (ovvero, come e quanto cambiano le soluzioni per piccole variazioni della legge che regola l'evoluzione temporale degli stati) e infine questioni riguardanti l'esistenza e la determinazione di insiemi invarianti per la dinamica. Le precedenti questioni necessitano di risultati teorici, infatti le similazioni dirette di sistemi dinamici per tempi molto lunghi non sono sempre possibili e –quando possibili– sono suscettibili di errori numerici. Lo scopo di questa breve nota è capire la teoria e la struttura degli insiemi invarianti per la mappa standard, che è un caso particolare di mappa twist.

(*)Assegnista in Matematica Pura - Dipartimento di Matematica Pura ed Applicata, Via Belzoni 7 - 35131 Padova, Italy - obern@math.unipd.it - Seminario tenuto il 28 novembre 2007.

2 Mappa twist

Restringiamo le nostre considerazioni a mappe 2-dimensionali e assumiamo che lo spazio delle fasi (x, y) sia un cilindro, con x la coordinata angolo: tale spazio delle fasi si ottiene in molti esempi in cui il momento y rappresenta la velocità, e quindi è illimitato, e la configurazione x rappresenta la coordinata angolo (si pensi all'oscillatore armonico). Sia ora

$$\mathbb{S}^1 \times \mathbb{R} \to \mathbb{S}^1 \times \mathbb{R}, \qquad (x, y) \mapsto (\bar{x}, \bar{y}) \tag{1}$$

una mappa dal cilindro in sè. La condizione di twist è comune (e naturale) nelle applicazioni fisiche: traduce il fatto che per punti con velocità y più grande la variabile configurazione x cresce più rapidamente.

Definizione 1 *La mappa (1) è una mappa twist se esiste una costante $K > 0$ tale che:*

$$\frac{d\bar{x}}{dy}|_x \geq K,$$

ovvero $\bar{x}$ è funzione monotona crescente della y.

Un caso particolare è quello in cui la legge ha la seguente struttura ($K = 1$):

$$\begin{cases} \bar{x} = x + y \pmod{2\pi} \\ \bar{y} = ... \end{cases}$$

3 Mappa standard

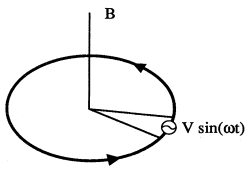
Sia f una funzione analitica, periodica e a media nulla. La mappa standard ϕ è un caso particolare di mappa twist e risulta così definita:

$$\phi : \begin{cases} \bar{x} = x + y \pmod{2\pi} \\ \bar{y} = y + \varepsilon f(\bar{x}) = y + \varepsilon f(x + y) \end{cases} \tag{2}$$

La mappa standard si presenta in vari esempi di interesse fisico: descriveremo il ciclotrone e il modello Frenkel-Kontorova.

3.1 Alcuni esempi

Il ciclotrone è un acceleratore di particelle consistente di un campo magnetico costante $\mathbf{B} = B_0\mathbf{e}_z$ e di un salto di potenziale dipendente dal tempo $V \sin(\omega t)$:

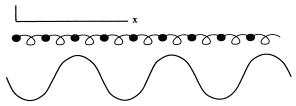


Supponiamo che ci sia un elettrone orbitante all'interno del ciclotrone. Il tempo per l'elettrone di compiere un giro attorno al circuito è $T = 2\pi\frac{E}{eBc}$ (dove E rappresenta l'energia della particella), mentre la sua variazione di energia è $\Delta E = -eV\sin(\omega t)$. Conseguentemente, nel piano energia-tempo (E, t), la dinamica dell'elettrone è data da:

$$\begin{cases} \bar{E} = E - eV\sin(\omega t) \\ \bar{t} = t + 2\pi\frac{\bar{E}}{eBc} \end{cases}$$

Un opportuno cambio di coordinate mostra che la mappa ottenuta è la mappa standard, per una specifica scelta della funzione f.

Il modello Frenkel-Kontorova consiste invece di una catena 1-dimensionale di particelle connesse da molle: per semplicità, prendiamo le costanti elastiche delle molle uguali a 1. Supponiamo inoltre che la catena sia posta su una superficie di un cristallo, rappresentato dal potenziale periodico $V(x) = \frac{k}{4\pi^2}\cos(2\pi x)$:



L'interazione tra il potenziale e le forze inter-atomiche risulta in equilibrio allorchè il seguente bilancio tra forze è soddisfatto:

$$(x_{j+1} - x_j) - (x_j - x_{j-1}) + \frac{k}{2\pi}\sin(2\pi x_j) = 0.$$

Definendo $y_j = x_j - x_{j-1}$ e interpretando l'indice j della particella come tempo, si ottiene:

$$\begin{cases} x_j = x_{j-1} + y_j \\ y_j = y_{j-1} + \frac{k}{2\pi}\sin(2\pi x_{j-1}) \end{cases}$$

Anche quest'ultima mappa è –a meno di un cambio di coordinate– la mappa standard.

3.2 Insiemi invarianti, frequenza

Indicheremo nel seguito con $\tilde{\phi}$ e $\tilde{x}$ la mappa ϕ e la variabile angolo x pensate nei rivestimenti di $\mathbb{S}^1 \times \mathbb{R}$ e di $\mathbb{S}^1$ rispettivamente. Iniziamo con l'introdurre la nozione di frequenza.

Definizione 2 *Un moto* $(x_t, y_t) = \phi^t(x_0, y_0),\ t \in \mathbb{Z}$ *ha frequenza* α *se*

$$\lim_{t\to\infty} \frac{\tilde{x}_t}{t} = \alpha.$$

La mappa standard (2) dipende dal parametro perturbativo $\varepsilon > 0$: in particolare, nel caso (imperturbato) in cui $\varepsilon = 0$, ogni insieme $y = \text{cost}$ è invariante per ϕ e su di esso ϕ è una traslazione di y: $x_t = x_0 + ty_0 (\text{mod } 2\pi),\quad y_t = y_0$. Inoltre, sempre per $\varepsilon = 0$: se $\alpha/2\pi$ è razionale, le orbite della mappa standard risultano periodiche, mentre se $\alpha/2\pi$ è irrazionale, le orbite della mappa standard risultano quasi-periodiche, ovvero ogni

x_t ($t \in \mathbb{Z}$) è denso in $y = \text{cost}$.
Conseguentemente, è naturale porsi i seguenti quesiti: al crescere del parametro perturbativo $\varepsilon > 0$, persistono insiemi invarianti? E se si, di che tipo? Sono curve oppure insiemi dalla topologia più raffinata? È quest'ultima una domanda tutt'altro che banale, alla quale risponde –nel caso di frequenze diofantine e piccole perturbazioni– la teoria KAM.

3.3 Insiemi di Aubry-Mather

Un punto chiave della teoria di Aubry-Mather è quindi la ricerca di tali strutture invarianti e il risultato che Aubry e Mather hanno indipendentemente dimostrato negli anni '80 può essere riassunto come segue.

Teorema 1 *Per ogni $\varepsilon > 0$ e per ogni $\alpha/2\pi$ irrazionale, la mappa standard:*

$$\phi : \begin{cases} \bar{x} = x + y \ (mod\ 2\pi) \\ \bar{y} = y + \varepsilon f(\bar{x}) = y + \varepsilon f(x + y) \end{cases}$$

ammette o una curva o un insieme di Cantor, che si proietta iniettivamente su $y = 0$, che è invariante e che supporta solo moti con numero di rotazione α: è l'insieme di Aubry-Mather M_α.

Tale problema –di natura dinamica– può essere tradotto in un problema di risoluzione di un'equazione nel modo che segue. Si cerca una riparametrizzazione di M_α:

$$\sigma_\alpha : \vartheta \to (\tilde{x} = u(\vartheta), y = v(\vartheta))$$

con le seguenti due proprietà:

(a) $(u(\vartheta + 2\pi), v(\vartheta + 2\pi)) = (u(\vartheta) + 2\pi, v(\vartheta))$,

(b) $\sigma_\alpha(\vartheta + \alpha) = \tilde{\phi} \circ \sigma_\alpha(\vartheta)$.

Le precedenti condizioni equivalgono a richiedere rispettivamente che:

(a) l'immagine di σ_α rappresenti un insieme invariante per $\tilde{\phi}$,

(b) la mappa indotta su tale insieme da $\tilde{\phi}$ sia una traslazione di α.

In maggior dettaglio, utilizzando la definizione della mappa standard, si ottiene che le componenti della riparametrizzazione $\sigma_\alpha(\vartheta) = (u(\vartheta), v(\vartheta))$ devono soddisfare alle seguenti equazioni:

$$(3) \qquad \begin{cases} u(\vartheta) - \frac{u(\vartheta+\alpha)+u(\vartheta-\alpha)}{2} + \frac{\varepsilon}{2} f \circ u(\vartheta) = 0 \\ \\ v(\vartheta) = u(\vartheta + \alpha) - u(\vartheta) \end{cases}$$

e che l'insieme di Aubry-Mather M_α risulta:

$$M_\alpha = \{(\tilde{x}, y) : \ \tilde{x} = u(\vartheta), \ y = u(\vartheta + \alpha) - u(\vartheta), \ \vartheta \in \mathbb{S}^1\}.$$

3.4 Risultati di esistenza, unicità e regolarizzazione

Il problema dell'esistenza per gli insiemi di Aubry-Mather è stato brillantemente risolto da Mather [2] con una tecnica variazionale. Definita la funzione (generatrice) $h(\tilde{x}_0, \tilde{x}_1) = \frac{(\tilde{x}_1-\tilde{x}_0)^2}{2} + \varepsilon F(\tilde{x}_1)$, dove $F' = f$, egli ha dimostrato che il minimo del funzionale

$$I_\alpha[u] = \int_0^1 h(u(\vartheta), u(\vartheta+\alpha))d\vartheta$$

nella classe $\{u : \ u(\vartheta) \leq u(\vartheta') \text{ per } \vartheta < \vartheta', \ u(\vartheta+2\pi) = u(\vartheta)+2\pi\}$ risolve l'equazione (3) per $u(\vartheta)$. Conseguentemente, essendo tale minimo una funzione non necessariamente differenziabile, essa può avere al più un insieme numerabile di discontinuità e risulta unica a meno di traslazioni e a meno dei valori alle discontinuità.

Un'ulteriore dimostrazione dell'esistenza e dell'unicità di tali insiemi invarianti M_α è stata poi messa a punto da Moser ([3] e [4]): la sua dimostrazione risulta importante perchè è supportata da metodi di viscosità. Egli ha dimostrato che, per ogni $\nu > 0$, il seguente funzionale regolarizzato:

$$I^\nu_\alpha[u] = \int_0^1 \Big(-\nu \log \frac{\partial u}{\partial \vartheta} + h(u(\vartheta), u(\vartheta+\alpha))\Big) d\theta$$

ammette un minimo $u^\nu(\vartheta)$ di classe C^2 e che le curve $u^\nu(\vartheta)$ convergono puntualmente a $u(\vartheta)$, soluzione per (3), $\nu \to 0$ in tutti i punti di continuità di $u(\vartheta)$. In maggior dettaglio, le curve $u^\nu(\vartheta)$, al variare di $\nu > 0$, formano una foliazione di $u(\vartheta)$.

3.5 Tecnica di localizzazione

Se da un punto di vista teorico, come abbiamo visto nelle precedenti sezioni, gli insiemi di Aubry-Mather sono stati ampiamente studiati e le fondamentali questioni riguardo alla loro esistenza ed unicità risolte, tali strutture invarianti M_α non vengono rilevate da tecniche numeriche standard: l'idea alla base dell'articolo [1] è infatti quella di proporre e mettere in pratica una nuova tecnica per la loro localizzazione. Nel seguito tratteremo le linee di tale procedimento iterativo.

1. Poniamo $u(\vartheta) = \vartheta + U(\vartheta)$, con $U(\vartheta)$ periodica in $[0, 2\pi]$:

$$U(\vartheta) = \sum_{k\in\mathbb{Z}} u_k e^{ik\vartheta}.$$

L'equazione (3) per $u(\vartheta)$ si riscrive per $U(\vartheta)$ nel modo seguente:

$$F[U](\vartheta) := U(\vartheta) - \frac{1}{2}\Big(U(\vartheta+\alpha) + U(\vartheta-\alpha) - \varepsilon f(\vartheta + U(\vartheta))\Big) = 0.$$

A partire a da un guess iniziale U^1, si definisce una successione di soluzioni approssimate U^k, $k \leq K$:

$$U^k = U^{k-1} + \delta U^k,$$

richiedendo che (qui $U = U^{k-1}$ e $\delta U = \delta U^k$) per $\nu_0 > 0$ e $\nu_1 > 0$:

$$\nu_0\delta U - \nu_1\delta U'' + \delta U - \frac{\delta U(\vartheta+\alpha) + \delta U(\vartheta-\alpha)}{2} + U - \frac{U(\vartheta+\alpha) + U(\vartheta-\alpha)}{2} + \frac{\varepsilon}{2} f(\vartheta+U) = 0.$$

Il seguente risultato analitico ci garantisce che il precedente procedimento definisce effettivamente una successione di funzioni approssimante la vera soluzione $U(\vartheta)$ di $F[U](\vartheta) = 0$.

Proposizione 1 *Sulla striscia di altezza σ denotiamo con $|\cdot|_\sigma$ la sup-norma e con $\lambda(\sigma)$ la costante di Lipschitz di f. Allora:*

(a)

$$\frac{|F[U+\delta U]|_{\sigma-\delta}}{|F[U]|_\sigma} \leq \sum_{n\in\mathbb{Z}\setminus 0} \frac{\nu_0 + \nu_1 n^2 + \frac{\epsilon}{2}\lambda(\sigma)}{\nu_0 + \nu_1 n^2 + 1 - \cos n\alpha} e^{-|n|\delta}.$$

(b) Esistono $\nu_0^* > 0,\ \nu_1^* > 0$ tali che

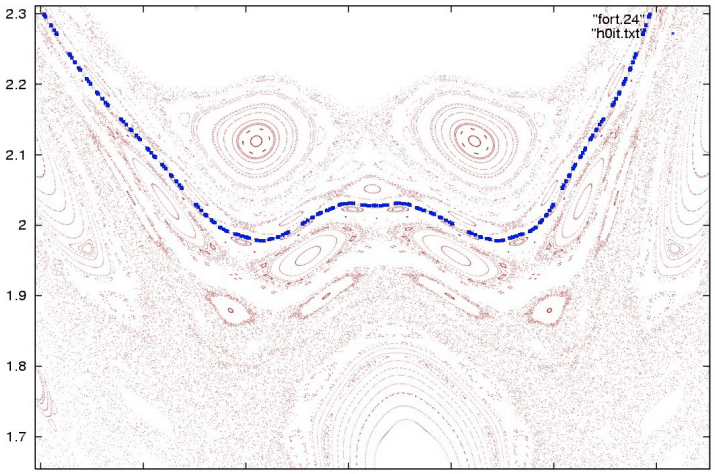
$$\frac{|F[U+\delta U]|_{\sigma-\delta}}{|F[U]|_\sigma} < 1.$$

Poiché i parametri $\nu_0,\ \nu_1 > 0$ agiscono sulla differenza tra due approssimazioni successive, ad ogni passo possono essere scelti $\nu_0^* > 0,\ \nu_1^* > 0$ in modo da garantire che $\frac{|F[U+\delta U]|_{\sigma-\delta}}{|F[U]|_\sigma} < 1$. Conseguentemente, non si lavora al limite per $\nu_0,\ \nu_1 \to 0$ e in questa costruzione ciò risulta fondamentale per ovviare alla presenza dei piccoli divisori.
2. Ricordiamo che $u(\vartheta) = \vartheta + U(\vartheta)$. Avendo usato tecniche di regolarizzazione, la soluzione approssimata $u(\vartheta)$ è regolare e l'insieme di Aubry-Mather risulta localizzato vicino alla curva:

$$\tilde{M}_\alpha = \{\big(u(\vartheta), u(\vartheta+\alpha) - u(\vartheta)\big),\ \vartheta \in \mathbb{S}^1\}.$$

3. Come localizzare infine i gaps dell'insieme di Aubry-Mather sulle curve approssimanti $\tilde{M}_\alpha$? Abbiamo utilizzato una misura su $\mathbb{S}^1$ corrispondente alla densità dei punti sull'insieme di Aubry-Mather.

Nella figura che segue, riportiamo in rilievo uno degli insieme di Aubry-Mather cosí ottenuti.



Bibliografia

[1] Guzzo, M., Bernardi, O., Cardin, F., *The experimental localization of Aubry-Mather sets using regularization techniques inspired by viscosity theory.* Chaos 17 (2007), no. 3, 033107, 9 pp.

[2] Mather, J. N., *Existence of quasiperiodic orbits for twist homeomorphisms of the annulus.* Topology 21 (1982), no. 4, 457–467.

[3] Moser, J., *An unusual variational problem connected with Mather's theory for monotone twist mappings.* Seminar on Dynamical Systems (St. Petersburg, 1991); Progr. Nonlinear Differential Equations Appl. 12 (1994), 81–89.

[4] Moser, J., *Smooth approximation of Mather sets of monotone twist mappings.* Comm. Pure Appl. Math. 47 (1994) no. 5, 625–652.

A functional analytic approach for the analysis of singularly perturbed boundary value problems

Matteo Dalla Riva (*)

Abstract. We consider a boundary value problem defined on an open and bounded subset of the 3-dimensional Euclidean space. Such an open set presents a hole in its interior. Our aim is to describe the behavior of the solution of the boundary value problem as the hole shrinks to a point. This kind of problem is not new at all. Indeed it has been long investigated by the techniques of the *asymptotic analysis* (see, e.g., the works of Keller, Kozlov, Movchan, Maz'ya, Nazarov, Plamenewskii, Ozawa and Ward.) By a simple example, we illustrate the kind of result that we can expect by the approach of the asymptotic analysis. Then, we show the result obtained by the alternative approach proposed by Lanza de Cristoforis in some papers from 2001. We point out the main differences between the two results.

Sunto. Si considererà un problema con dato al bordo definito su un aperto limitato dello spazio Euclideo 3-dimensionale. Tale aperto avrà un buco al suo interno. Il nostro scopo è di descrivere il comportamento della soluzione del problema con dato al bordo quando il buco collassa ad un punto. Problemi di questo genere sono stati lungamente studiati tramite le tecniche dell' *analisi asintotica* (si vedano ad esempio i lavori di Keller, Kozlov, Movchan, Maz'ya, Nazarov, Plamenewskii, Ozawa e Ward). Illustreremo in un facile esempio quale tipo di risultato possiamo attenderci applicando tali tecniche. Poi mostreremo il risultato che si ottiene tramite l'approccio alternativo proposto da Lanza de Cristoforis in alcuni lavori a partire dal 2001 e metteremo in luce le principali differenze tra i due risultati.

1 Some notation

Let λ be a real number in the open interval $]0,1[$. Let m,n be natural numbers, $m,n \geq 1$. Let Ω be an open subset of the n-dimensional Euclidean space $\mathbb{R}^n$. We denote by $C^{1,\lambda}(\mathrm{cl}\Omega,\mathbb{R}^m)$ the set of the bounded continuous and differentiable functions $f : \mathrm{cl}\Omega \to \mathbb{R}^m$ whose partial derivatives of order one extend to a bounded continuous function of $\mathrm{cl}\Omega$ to

(*)Ph.D. School in Pure Mathematics. Università di Padova, Dipartimento di Matematica Pura ed Applicata, Via Trieste 63, 35121 Padova, Italy. E-mail: mdallari@math.unipd.it. Seminar held on 16 January 2008.

$\mathbb{R}^n$ which is Hölder continuous with exponent λ. If $m = 1$ we write $C^{1,\lambda}(\mathrm{cl}\Omega)$ instead of $C^{1,\lambda}(\mathrm{cl}\Omega, \mathbb{R})$. Moreover, we say that Ω is of class $C^{1,\lambda}$ if its boundary $\partial\Omega$ is a sub-manifold of $\mathbb{R}^n$ of class $C^{1,\lambda}$. If Ω is a bounded open subset of $\mathbb{R}^n$ of class $C^{1,\lambda}$ we define the space $C^{1,\lambda}(\partial\Omega, \mathbb{R}^m)$ by exploiting the local parametrizations.

2 The perforated domain A_ϵ

We now fix two open bounded subsets A and B of the 3-dimensional Euclidean space $\mathbb{R}^3$, A and B of class $C^{1,\lambda}$. We assume that A and B are connected and that their exteriors, $\mathbb{R}^3 \setminus \mathrm{cl}A$ and $\mathbb{R}^3 \setminus \mathrm{cl}B$, are also connected (condition which means that A and B have no holes.) We further assume that both A and B contains the origin $0 \equiv (0, 0, 0)$ of $\mathbb{R}^3$. Then, one immediately verifies that there exists a real number $\epsilon_0 > 0$ such that the closure $\mathrm{cl}(\epsilon B)$ of the set ϵB is contained in A for all $\epsilon \in]0, \epsilon_0[$. We denote by A_ϵ the set $A \setminus (\mathrm{cl}\,\epsilon B)$, for all $\epsilon \in]0, \epsilon_0[$. We will introduce in the following paragraph a boundary value problem for the Laplace operator Δ in the perforated domain domain A_ϵ, with $\epsilon \in]0, \epsilon_0[$.

3 A boundary value problem in A_ϵ

Let $\epsilon \in]0, \epsilon_0[$. Let $f \in C^{1,\lambda}(\partial A)$ and $g \in C^{1,\lambda}(\partial B)$. We consider the following boundary value problem. Find $u \in C^{1,\lambda}(\mathrm{cl}A_\epsilon) \cap C^2(A_\epsilon)$ such that

$$(1) \qquad \begin{cases} \Delta u(x) = 0 & \text{if } x \in A_\epsilon, \\ u(x) = f(x) & \text{if } x \in \partial A, \\ u(x) = g(x/\epsilon) & \text{if } x \in \partial(\epsilon B). \end{cases}$$

As it is well known, problem (1) admits a unique solution $u \in C^{1,\lambda}(\mathrm{cl}A_\epsilon) \cap C^2(\mathrm{cl}A_\epsilon)$. We denote by u_ϵ such a solution. Our aim is to describe the behavior of u_ϵ as ϵ goes to 0.

To do so, we fix a point x_0 of A. We assume that x_0 is different from the origin 0. Then, it is easily verify that there exists a positive real number ϵ_1, smaller than the number ϵ_0 introduced above, such that, the point x_0 belongs to A_ϵ for all $\epsilon \in]0, \epsilon_1[$. In particular, the solution u_ϵ of problem (1) is defined on x_0. Therefore, we can consider the map of $]0, \epsilon_1[$ to $\mathbb{R}$ which takes ϵ to $u_\epsilon(x_0)$ and it makes sense to investigate the limit

$$\lim_{\epsilon \to 0^+} u_\epsilon(x_0).$$

4 Asymptotic analysis

The question introduced in the previous paragraph is not new at all. Indeed it has been long investigated by the techniques of *asymptotic analysis*. It is perhaps difficult to provide a complete list of contributions. Here we mention the work of Kozlov, Maz'ya and Movchan [2], Maz'ya, Nazarov and Plamenewskii [8,9], Ozawa [10], Ward and Keller [11].

The following Theorem 1 due to Maz'ya, Nazarov and Plamenewskii [8, Theorem 2.1.1] explains the kind of results that we can expect by asymptotic analysis.

Theorem 1 *There exist functions $v_j \in C^{1,\lambda}(\mathrm{cl}A) \cap C^2(A)$ and $w_j \in C^{1,\lambda}(\mathbb{R}^n \setminus B) \cap C^2(\mathbb{R}^n \setminus \mathrm{cl}B)$, $j \in \mathbb{N}$, with $\Delta v_j = 0$ in A and $\Delta w_j = 0$ in $\mathbb{R}^n \setminus \mathrm{cl}B$, such that, the asymptotic behavior of the solution u_ϵ as $\epsilon \to 0^+$ is delivered by the following equation,*

$$u_\epsilon(x_0) = v_0(x_0) + \sum_{j=1}^{N} \epsilon^j \left(v_j(x_0) + w_j(x_0/\epsilon)\right) + O(\epsilon^{N+1}), \quad \text{as } \epsilon \to 0^+, \tag{2}$$

for all $x_0 \in A \setminus \{0\}$ and all $N \in \mathbb{N}$, where we understand the sum on the right hand side is 0 if $N = 0$. In particular,

$$\lim_{\epsilon \to 0^+} u_\epsilon(x_0) = v_0(x_0), \quad \forall\ x_0 \in A \setminus \{0\},$$

and v_0 is the unique solution of

$$\begin{cases} \Delta v_0 = 0 & \text{in } A, \\ v_0 = f & \text{on } \partial A. \end{cases}$$

In the sequel we shall present a result which is obtained by an approach alternative to that of the asymptotic analysis. Such approach exploits Potential Theoretic Methods and methods of nonlinear functional analysis for real analytic functions (cf. Lanza de Cristoforis [3,4,5,6,7] and [1].) The aim is in some sense different from the one of the asymptotic analysis. Indeed we wish to express the dependence of u_ϵ upon ϵ in terms of a real analytic function defined in a whole open neighborhood of $\epsilon = 0$. In particular, we want to get rid of the reminder term in expression (2). Moreover, we do not confine our attention to the dependence of the solution u_ϵ upon ϵ. We consider also the dependence upon perturbation of the point where the hole is situated, and of the shape of the hole, and of the shape of the outer domain, and of the boundary data on the boundary of the hole and on the outer boundary. To do so we introduce some more notation.

5 The perforated domain $\mathbb{A}[\omega, \epsilon, \phi^A, \phi^B]$

Let A and B be the domains introduced above. We denote by $\mathcal{A}_{\partial A}$ the set of all functions ϕ^A from ∂A to $\mathbb{R}^3$ which we retain as admissible. These are the functions $\phi^A \in C^{1,\lambda}(\partial A, \mathbb{R}^3)$ which are injective and whose differential $d\phi^A(x)$ is injective for all $x \in \partial A$.

The reason for which these functions are told *admissible* is the following. If $\phi^A \in \mathcal{A}_{\partial A}$, then its image $\phi^A(\partial A)$ separates $\mathbb{R}^3$ in two connected components, a bounded one, which we denote by $\mathbb{I}[\phi^A]$, and an unbounded one, which we denote by $\mathbb{E}[\phi^A]$. Thus, $\mathbb{I}[\phi^A]$ is an open bounded connected subset of $\mathbb{R}^3$ with no holes, whose boundary $\phi^A(\partial A)$ is a sub-manifold of $\mathbb{R}^3$ of class $C^{1,\lambda}$ parametrized by the function ϕ^A.

Similarly, we denote by $\mathcal{A}_{\partial B}$ the set of all functions $\phi \in C^{1,\lambda}(\partial B, \mathbb{R}^3)$ which are injective and whose differential $d\phi(x)$ is injective for all $x \in \partial B$. We denote by $\mathbb{I}[\phi^B]$ the bounded open domain with boundary $\phi^B(\partial B)$.

Now, let $\phi^A \in \mathcal{A}_{\partial A}$ and $\phi^B \in \mathcal{A}_{\partial B}$. Let ω be a point of $\mathbb{I}[\phi^A]$. Let $\epsilon \in \mathbb{R}$. Clearly, if ϵ is small enough, the closure of the set $\omega + \epsilon\mathbb{I}[\phi^B]$ is contained in $\mathbb{I}[\phi^A]$. If this is the case, $\omega + \epsilon\mathbb{I}[\phi^B]$ is our hole, and we obtain the perforated domain $\mathbb{A}[\omega, \epsilon, \phi^A, \phi^B]$ by removing the closure of the hole $\omega + \epsilon\mathbb{I}[\phi^B]$ from the domain $\mathbb{I}[\phi^A]$. We note that $\mathbb{A}[\omega, \epsilon, \phi^A, \phi^B]$ is a bounded open and connected subset of $\mathbb{R}^3$ with boundary made of two connected components, $\omega + \epsilon\phi^B(\partial B)$ and $\phi^A(\partial A)$.

We denote by $\mathcal{E}^{1,\lambda}$ the set of all the admissible quadruples $(\omega, \epsilon, \phi^A, \phi^B)$ which give rise to a perforated domain $\mathbb{A}[\omega, \epsilon, \phi^A, \phi^B]$ and we point out that $\mathcal{E}^{1,\lambda}$ is an open subset of the Banach space $\mathbb{R}^3 \times \mathbb{R} \times C^{1,\lambda}(\partial A, \mathbb{R}^3) \times C^{1,\lambda}(\partial B, \mathbb{R}^3)$. In particular, $\mathbb{A}[\omega, 0, \phi^A, \phi^B] = \mathbb{I}[\phi^A] \setminus \{\omega\}$.

6 A boundary value problem in $\mathbb{A}[\omega, \epsilon, \phi^A, \phi^B]$

Let $(\omega, \epsilon, \phi^A, \phi^B)$ belong to $\mathcal{E}^{1,\lambda}$. Let $\epsilon > 0$. Let f belong to $C^{1,\lambda}(\partial A)$ and g belong to $C^{1,\lambda}(\partial B)$. We consider the following boundary value problem in $\mathbb{A}[\omega, \epsilon, \phi^A, \phi^B]$. Find $u \in C^{1,\lambda}(\mathrm{cl}\mathbb{A}[\omega, \epsilon, \phi^A, \phi^B]) \cap C^2(\mathbb{A}[\omega, \epsilon, \phi^A, \phi^B])$ such that

$$\text{(3)} \qquad \begin{cases} \Delta u = 0 & \text{in } \mathbb{A}[\omega, \epsilon, \phi^A, \phi^B], \\ u = f \circ (\phi^A)^{(-1)} & \text{on } \phi^A(\partial A), \\ u = g \circ (\omega + \epsilon\phi^B)^{(-1)} & \text{on } \omega + \epsilon\phi^B(\partial\Omega_B). \end{cases}$$

As it is well known, problem (3) admits a unique solution u for each given $(\omega, \epsilon, \phi^A, \phi^B) \in \mathcal{E}^{1,\lambda}$ and $(f, g) \in C^{1,\lambda}(\partial A) \times C^{1,\lambda}(\partial B)$. So it makes sense to consider such a solution as a function of the variables $(\omega, \epsilon, \phi^A, \phi^B, f, g)$ and to write $u[\omega, \epsilon, \phi^A, \phi^B, f, g]$ to denote it. Our purpose is to investigate the dependence of $u[\omega, \epsilon, \phi^A, \phi^B, f, g]$ upon $(\omega, \epsilon, \phi^A, \phi^B, f, g)$ in a neighborhood of a *degenerate* 7-tuple $(\omega_0, 0, \phi_0^A, \phi_0^B, f_0, g_0)$.

7 A functional analytic approach

We show in the following Theorem 2 a real analytic continuation result in a neighborhood of a *degenerate* 7-tuple $(\omega_0, 0, \phi_0^A, \phi_0^B, f_0, g_0)$ for the function which takes $(\omega, \epsilon, \phi^A, \phi^B, f, g)$ to $u[\omega, \epsilon, \phi^A, \phi^B, f, g]$. We omit to present here a proof. We just remark that Theorem 2 is an immediate corollary of Theorem 5.7 in Lanza de Cristoforis [5].

Theorem 2 *Let $(\omega_0, 0, \phi_0^A, \phi_0^B, f_0, g_0) \in \mathcal{E}^{1,\lambda} \times C^{1,\lambda}(\partial A) \times C^{1,\lambda}(\partial B)$. Let Ω be an open subset of $\mathbb{R}^3$ such that $\mathrm{cl}\Omega \subset \mathbb{I}[\phi_0^A] \setminus \{\omega_0\}$. Then there exist an open neighborhood $\mathcal{U}$ of $(\omega_0, 0, \phi_0^A, \phi_0^B, f_0, g_0)$ in $\mathcal{E}^{1,\lambda} \times C^{1,\lambda}(\partial A) \times C^{1,\lambda}(\partial B)$ and a real analytic operator $U[\cdot]$ of $\mathcal{U}$ to $C(\mathrm{cl}\Omega)$, endowed with the norm of the uniform convergence, such that the following conditions hold.*

(i) $\mathrm{cl}\Omega \subset \mathbb{A}[\omega, \epsilon, \phi^A, \phi^B]$ *for all* $(\omega, \epsilon, \phi^A, \phi^B, f, g) \in \mathcal{U}$.

(ii) We have

$$u[\omega, \epsilon, \phi^A, \phi^B, f, g](x) = U[\omega, \epsilon, \phi^A, \phi^B, f, g](x), \quad \forall\, x \in \mathrm{cl}\Omega,$$

for all $(\omega, \epsilon, \phi^A, \phi^B, f, g) \in \mathcal{U}$ *with* $\epsilon > 0$.

(iii)

$$U[\omega, 0, \phi^A, \phi^B, f, g](x) = v_0[\phi^A, f](x)$$

for all $x \in \mathrm{cl}\Omega$ *and for all* $(\omega, 0, \phi^A, \phi^B, f, g) \in \mathcal{U}$, *where* $v_0[\phi^A, f] \in C^{1,\lambda}(\mathrm{cl}\mathbb{I}[\phi^A]) \cap C^2(\mathbb{I}[\phi^A])$ *is the unique solution of*

$$\begin{cases} \Delta v_0[\phi^A, f] = 0 & \text{in } \mathbb{I}[\phi^A], \\ v_0[\phi^A, f] = f \circ (\phi^A)^{(-1)} & \text{on } \phi^A(\partial A). \end{cases}$$

We observe that the result stated in Theorem 2 is in accordance with the behavior one would expect by looking at the asymptotic expansion (2). Actually, equation (2) may *suggest* the validity of the above result, at least for fixed values of ω, ϕ^A, ϕ^B, f, g. Perhaps, one could also try to prove such result for fixed values of ω, ϕ^A, ϕ^B, f, g, by showing that the series on the right hand side of (2) converge to the corresponding function. Such approach however may be non trivial, and for its possible feasibility we take no credit and refer to some expert in asymptotic analysis. We also mention that one could think of proving Theorem 2 by considering a real analytic curve of sextuples $(\omega, \epsilon, \phi^A, \phi^B, f, g)$ through a degenerate sextuple $(\omega_0, 0, \phi_0^A, \phi_0^B, f_0, g_0)$ depending on a real parameter and then by showing the appropriate continuation properties of the solution of (3) as a function of the parameter of the curve. Such method, also known as "parameter method" at least since the thirties would anyway yield a weaker form of results. Indeed it is well known that for an operator in a Banach space, even in the finite dimensional case, real analyticity on all real analytic curves does not imply real analyticity.

References

[1] M. Dalla Riva, "Potential theoretic methods for the analysis of singularly perturbed problems in linearized elasticity". Doctoral dissertation, Advisor M. Lanza de Cristoforis; Università degli Studi di Padova, 2008.

[2] Vladimir Kozlov, Vladimir Maz'ya, and Alexander Movchan, "Asymptotic analysis of fields in multi-structures". Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 1999.

[3] Massimo Lanza de Cristoforis, *Asymptotic behaviour of the conformal representation of a Jordan domain with a small hole, and relative capacity.* Agranovsky, Mark (ed.) et al., Complex analysis and dynamical systems. Proceedings of the 3rd international conference, Karmiel, Israel, June 19–22, 2001. Providence, RI: American Mathematical Society (AMS). Contemporary Mathematics 364. Israel Mathematical Conference Proceedings.

[4] Massimo Lanza de Cristoforis, *Asymptotic behaviour of the conformal representation of a Jordan domain with a small hole in Schauder spaces.* Comput. Methods Funct. Theory, 2(1) (2002), 1–27.

[5] Massimo Lanza de Cristoforis, *Asymptotic behavior of the solutions of the Dirichlet problem for the Laplace operator in a domain with a small hole. A functional analytic approach.* To appear, 2005.

[6] Massimo Lanza de Cristoforis, *A domain perturbation problem for the Poisson equation.* Complex Var. Theory Appl., 50(7-11) (2005), 851–867.

[7] Massimo Lanza de Cristoforis, *Perturbation problems in potential theory, a functional analytic approach.* J. Appl. Funct. Anal., 2(3) (2007), 197–222.

[8] Vladimir Maz′ya, Serguei Nazarov, and Boris Plamenevskij, *Asymptotic theory of elliptic boundary value problems in singularly perturbed domains. Vol. I, volume 111 of* Operator Theory: Advances and Applications. Birkhäuser Verlag, Basel (2000). Translated from the German by Georg Heinig and Christian Posthoff.

[9] Vladimir Maz′ya, Serguei Nazarov, and Boris Plamenevskij, *Asymptotic theory of elliptic boundary value problems in singularly perturbed domains. Vol. II, volume 112 of* Operator Theory: Advances and Applications. Birkhäuser Verlag, Basel (2000). Translated from the German by Plamenevskij.

[10] Shin Ozawa, *Electrostatic capacity and eigenvalues of the Laplacian.* J. Fac. Sci. Univ. Tokyo Sect. IA Math., 30(1) (1983), 53–62.

[11] Michael J. Ward and Joseph B. Keller, *Strong localized perturbations of eigenvalue problems.* SIAM J. Appl. Math., 53(3) (1993), 770–798.

Construction of balanced complex polytopes in $\mathbb{C}^2$

Cristina Vagnoni (*)

Abstract. The asymptotic behaviour of the solutions of a discrete linear dynamical system may be related to the spectral radius ρ of its associated family $\mathcal{F}$; in particular, a system is stable if $\rho \leq 1$ and there exists an extremal norm for $\mathcal{F}$. Since the extremal norms play an important role, in the last decades some algorithms have been proposed in order to find real extremal norms of polytope type in the case of finite families. However, recently it has been observed that it is more useful to consider *complex polytope norms*. Such norms require the notion of *balanced complex polytopes* and so it is interesting to analyze the geometric representation of such objects. However, due to the strong increase in complexity of their geometry, the work will be confined to the two-dimensional case. In particular, we give original theoretical results on the geometry of two-dimensional balanced complex polytopes in order to present two efficient algorithms, one for the geometric representation of a balanced complex polytope and the other for the computation of the corresponding complex polytope norm of a vector.

Sunto. La stabilità di un sistema dinamico lineare discreto è un importante problema che può essere affrontato mediante il calcolo del raggio spettrale ρ della famiglia di matrici $\mathcal{F}$ che definisce il sistema stesso; in particolare, il sistema è stabile se $\rho \leq 1$ ed esiste una norma estremale per la famiglia $\mathcal{F}$. Dato il ruolo importante rivestito dalle norme estremali nel presente problema, negli ultimi anni sono stati proposti alcuni algoritmi per il calcolo di norme estremali di tipo politopico reale. Tuttavia, è stato recentemente osservato che può essere necessario considerare *norme di tipo politopico complesso*, le quali richiedono la nozione di *politopo complesso bilanciato*. Per questo motivo, il lavoro si è incentrato sullo studio della geometria di politopi complessi e, vista la notevole complessità della loro geometria, ci si è limitati al caso complesso bidimensionale. In particolare, verrano presentati sia nuovi risultati teorici sulla geometria dei politopi complessi bilanciati in $\mathbb{C}^2$ sia i primi algoritmi efficienti che forniscono una completa caratterizzazione della loro geometria e permettono di calcolare la corrispondente norma politopica complessa di un arbitrario vettore in $\mathbb{C}^2$.

(*)Ph.D. School in Applied Mathematics. Università di Padova, Dipartimento di Matematica Pura ed Applicata, Via Trieste 63, 35121 Padova, Italy. E-mail: vagnoni@math.unipd.it. Seminar held on 30 January 2008.

1 Introduction

The knowledge of discrete linear dynamical systems of the kind

$$x^{(i+1)} = A^{(i)}x^{(i)},\ i = 0, 1, 2, \dots \tag{1}$$

where $x^{(0)}$ is given and $A^{(i)} \in \mathbb{C}^{n\times n}$, $i = 0, 1, 2, \dots$, are given matrices, is of great importance in many fields of applied mathematics, such as Engineering, Physics, Biology, Chemistry, etc. We note that sometimes, it may be more convenient to perform the analysis of continuous processes on the corresponding discretized ones which are, substantially, of the form (1). A crucial aspect related to a discrete linear dynamical system is its stability properties, that is, the asymptotic behaviour of its solutions. For a given starting point $x^{(0)}$, they are given by

$$x^{(i+1)} = P^{(i)}x^{(0)}\,,\ \ P^{(i)} = A^{(i)} \cdots A^{(0)}.$$

Therefore, it is clear that the products of the matrices of the family $\mathcal{F} = \{A^{(i)}\}_{i\geq 0}$ play an essential role in the behaviour of the linear system. Indeed, by defining (see [7]) the *spectral radius* of the family $\mathcal{F}$ as

$$\rho(\mathcal{F}) = \limsup_{k\to+\infty} \overline{\rho}_k(\mathcal{F}), \quad \overline{\rho}_k(\mathcal{F}) = \sup_{P\in\Sigma_k(\mathcal{F})} \rho(P),$$

where $\Sigma_k(\mathcal{F}) = \{P \in \mathbb{C}^{n\times n} : P = \prod_{l=1}^{k} A^{(i_l)}\}$, we have that the asymptotic stability is guaranteed if $\rho(\mathcal{F}) < 1$. So, we are naturally led to the computation of the spectral radius of $\mathcal{F}$. This is not an easy task at all, but it suggests an important way to approach the stability question. The computation of $\rho(\mathcal{F})$ may be simplified by introducing the concept of extremal norm, which is a norm $\|\cdot\|_*$ satisfying the condition $\|\mathcal{F}\|_* = \rho(\mathcal{F})$, where the norm of the family $\mathcal{F}$ is defined, for any induced matrix norm $\|\cdot\|$, as $\|\mathcal{F}\| = \sup_{i\geq 0} \|A^{(i)}\|$.

In literature some algorithms have been proposed for the construction of extremal norms of polytope type and, in particular, norms whose unit ball a symmetric real polytope (see [1]).

However, how we can see in the next section, even if the dynamical system is real, the real polytope norms may be unsatisfactory. Therefore, it may be useful to enlarge the search of extremal norms to the class of the complex polytope norms (see [3], [4] and [6]). To this end we recall, in the second section, some theoretical results on the b.c.p. in order to give an algorithm for the computation of an extremal norm in the class of the complex polytope norms. Then, in the third section we give our new theoretical results on the geometry of a b.c.p. in $\mathbb{C}^2$, and present the first algorithm for the geometric representation of a b.c.p. and the related algorithm for the computation of a complex polytope norm. Finally, in the last section, we sketch some computational improvements of the previous algorithms.

2 An algorithm for the computation of an extremal norm

In order to present an algorithm for the computation of an extremal norm in the class of the complex polytope norms, we first need to define a balanced complex polytope (b.c.p.),

to give a first characterization of its boundary and finally to introduce the concept of complex polytope norm (see [5] and [4]).

Definition 2.1 *A bounded set $\mathcal{P} \subset \mathbb{C}^n$ is a* balanced complex polytope *if there exists a finite set of vectors $\mathcal{X} = \{x^{(i)}\}_{1 \le i \le m}$ such that $span(\mathcal{X}) = \mathbb{C}^n$ and*

$$\mathcal{P} = absco(\mathcal{X}) = \left\{ x \in \mathbb{C}^n \;\middle|\; x = \sum_{i=1}^{m} \lambda_i x^{(i)} \quad with\ \lambda_i \in \mathbb{C},\ \sum_{i=1}^{m} |\lambda_i| \le 1 \right\}.$$

Moreover, if $absco(\mathcal{X}') \subsetneq absco(\mathcal{X})$ for all $\mathcal{X}' \subsetneq \mathcal{X}$, then $\mathcal{X}$ is called an essential system of vertices (e.s.v.) for $\mathcal{P}$, whereas any vector $ux^{(i)}$ with $u \in \mathbb{C}$, $|u| = 1$, is called a vertex of $\mathcal{P}$.

Note that, from a geometrical point of view, if we identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$, a b.c.p. $\mathcal{P}$ and, in general, even the intersection $\mathcal{P} \cap \mathbb{R}^n$ are not classical polytopes because they are not bounded by hyperplanes. However, if the b.c.p. $\mathcal{P}$ admits an essential system of real vertices, then $\mathcal{P} \cap \mathbb{R}^n$ is a classical polytope.

In order to give a first characterization of the boundary of a b.c.p. we introduce the concept of b.c.p. of adjoint type as follow.

Definition 2.2 *A bounded set $\mathcal{P}^* \subset \mathbb{C}^n$ is a* b.c.p. of adjoint type *if there exist a finite set of vectors $\mathcal{X} = \{x^{(i)}\}_{1 \le i \le m}$ such that $span(\mathcal{X}) = \mathbb{C}^n$ and*

$$\mathcal{P}^* = \mathrm{adj}(\mathcal{X}) = \left\{ y \in \mathbb{C}^n \;\middle|\; |\langle y, x^{(i)} \rangle| \le 1,\ i = 1, \dots, m \right\}.$$

Moreover, if $\mathrm{adj}(\mathcal{X}') \supsetneq \mathrm{adj}(\mathcal{X})$ for all $\mathcal{X}' \subsetneq \mathcal{X}$, then $\mathcal{X}$ is called an essential system of facets *for $\mathcal{P}^*$, whereas any vector $ux^{(i)}$ with $u \in \mathbb{C}$, $|u| = 1$, is called a* facet *of $\mathcal{P}^*$.*

A b.c.p. $\mathcal{P}$ and its adjoint $\mathcal{P}^*$ may be related with the following theorem.

Theorem 2.1 *Let $\mathcal{P}$ be a b.c.p. and let $\mathcal{P}^* = \mathrm{adj}(\mathcal{P})$. Then $\mathcal{P} = adj(\mathcal{P}^*)$. Conversely, let $\mathcal{P}^*$ be a b.c.p. of adjoint type and $\mathcal{P} = \mathrm{adj}(\mathcal{P}^*)$. Then $\mathcal{P}^* = \mathrm{adj}(\mathcal{P})$.*

The boundary of a b.c.p. may be characterized through its *face* and *facets*, which are defined as follow.

Definition 2.3 *Let $\mathcal{P}$ be a b.c.p., $\mathcal{P}^* = \mathrm{adj}(\mathcal{P})$ and $y \in \partial\mathcal{P}^*$. Then the convex set*

$$F_y = \left\{ x \in \mathcal{P} \;\middle|\; \langle y, x \rangle = 1 \right\}$$

is called a (geometric) face *of $\mathcal{P}$, whereas y is called the* functional *associated to F_y. Moreover, any vertex of $\mathcal{P}$ belonging to F_y is called a* vertex *of F_y and the* dimension *of F_y is*

$$dim(F_y) = dim(span(F_y)) - 1.$$

In particular, if a face F_y has n linearly independent vertices (i.e. $dim(F_y) = n - 1$), then it is called a (geometric) facet *of $\mathcal{P}$.*

Moreover, it can be proved that any face is the convex hull of its vertices $\mathcal{X}_y$, that is,

$$F_y = \mathrm{co}(\mathcal{X}_y).$$

Finally, we can state the following theorem which characterize the boundary of a b.c.p.

Theorem 2.2

$$\partial\mathcal{P} = \bigcup_{y \in F^*(\mathcal{P})} F_y,$$

where $F^*(\mathcal{P})$ is the set of all the facets of $\mathcal{P}$. In other words, $\partial\mathcal{P}$ is the union of all the (geometric) facets of $\mathcal{P}$.

Now, the concept of polytope norm is extended to the complex case with the following Lemma.

Lemma 2.1 *Any b.c.p. $\mathcal{P}$ is the unit ball of a norm $\|\cdot\|_{\mathcal{P}}$ on $\mathbb{C}^n$ such that, for any $z \in \mathbb{C}^n$,*

$$\|z\|_{\mathcal{P}} = inf\left\{\rho > 0 \;\middle|\; z \in \rho\mathcal{P}\right\}.$$

Definition 2.4 *A* complex polytope norm *is any norm $\|\cdot\|_{\mathcal{P}}$ whose unit ball is a b.c.p.*

Moreover, denoting with $\mathcal{X} = \{x^{(i)}\}_{1 \le i \le m}$ an essential system of vertices of $\mathcal{P}$, it can be proved that for any $z \in \mathbb{C}^n$, it holds that

$$\|z\|_{\mathcal{P}} = min\left\{\sum_{i=1}^{m} |\lambda_i| \;\middle|\; z = \sum_{i=1}^{m} \lambda_i x^{(i)}\right\}.$$

We conclude this section giving a sketch of the algorithm for the computation of an extremal norm for a bounded family $\mathcal{F} = \{A^{(i)}\}_{i \in \mathcal{I}}$ of $n \times n$-matrices.

Algorithm 2.1

1. Choose a suitable product $Q \in \Sigma_k(\mathcal{F})$ (for some k).

2. Set $\rho = \rho(Q)^{1/k}$ and define the scaled family $\mathcal{F}^* = \{\rho^{-1}A^{(i)}\}_{i \in \mathcal{I}}$.

3. Compute a leading eigenvector v of Q and set $x^{(1)} = v$.

4. Set $s = 1$, $\mathcal{X}^{(1)} = \{x^{(1)}\}$ and $\mathcal{P}^{(1)} = absco(\mathcal{X}^{(1)})$.

5. Compute the set of vectors $\mathcal{V}^{(s)} = \mathcal{F}^*(\mathcal{X}^{(s)})$.

6. **If** $\mathcal{V}^{(s)} \subseteq \mathcal{P}^{(s)}$ **STOP** ($\mathcal{P}^{(s)}$ is the unit ball of an estremal norm).

7. Set $\mathcal{P}^{(s+1)} = absco(\mathcal{V}^{(s)} \cup \mathcal{X}^{(s)})$ and compute an e.s.v. $\mathcal{X}^{(s+1)}$ of $\mathcal{P}^{(s+1)}$.

8. Set $s = s + 1$ and go to **5.**

Note that, even if the family $\mathcal{F}$ is real, the eigenvector v may be complex, an this giustify the search of an extremal norm in the class of the complex polytope norms.

So the geometric representation of a b.c.p. may be an important tool, but since we have proved that the complexity of the boundary of a b.c.p. increase with the dimension n (see [8]), in the next section we limit ourselves to analyze the two-dimensional complex case.

3 The geometric representation of a b.c.p. in $\mathbb{C}^2$

In this section we give our original theoretical results on the geometry of a b.c.p. in $\mathbb{C}^2$, in order to present the first algorithm for the geometric representation of a b.c.p. and the related algorithm for the computation of a complex polytope norm. For a deeper knowledge of the incoming results see [8] and [9].

3.1 Theoretical results

The construction of a b.c.p. is done in an iterative manner, starting from a two vertex polytope and then by adding one of the remaining points at a time. Following this idea, we assume, without loss of generality, that the first two points $x^{(1)}, x^{(2)}$ of $\mathcal{X}$ are linearly independent and, starting from $\mathcal{P}^{(2)} = \text{absco}(\{x^{(1)}, x^{(2)}\})$, we construct $\mathcal{P}^{(k)} = \text{absco}(\mathcal{P}^{(k-1)} \cup \{x^{(k)}\})$, for $3 \leq k \leq m$, by adding $x^{(k)}$ to $\mathcal{P}^{(k-1)}$. In this incremental construction of the b.c.p. $\mathcal{P}$, we distinguish the construction of a b.c.p. with two essential vertices from the case of a b.c.p. with three essential vertices and the general case of a b.c.p. with $m \geq 4$ essential vertices, because these three cases are substantially different from one another.

We begin by analyzing the boundary of a b.c.p. with two essential vertices, which is $\mathcal{P} = absco(\{x^{(1)}, x^{(2)}\}) \subseteq \mathbb{C}^2$, with $x^{(1)}, x^{(2)}$ linearly independent. In the light of Theorem 2.2 we have that $\partial\mathcal{P}$ is the union of all the facets of $\mathcal{P}$, which are given by the segments which joint all the possible pairs of non-proportional vertices, that are

$$F_{y(\theta_1,\theta_2)} = e^{\mathrm{i}\theta_1}x^{(1)} \bullet\!\!-\!\!\bullet\, e^{\mathrm{i}\theta_2}x^{(2)} = co(\{e^{\mathrm{i}\theta_1}x^{(1)}, e^{\mathrm{i}\theta_2}x^{(2)}\}), \quad \theta_1, \theta_2 \in (-\pi, \pi].$$

Moreover, if we denote with $\|\cdot\|_{\mathcal{P}}$ the corresponding polytope norm, it can be proved that the boundary of $\mathcal{P}$, that is $\partial\mathcal{P} = \{x \in \mathbb{C}^2 \,:\, \|x\|_{\mathcal{P}} = 1\}$, is a branch of a fourth order surface in $\mathbb{C}^2$ (identified with $\mathbb{R}^4$) and so it is not a classical polytope.

Now we examine the addition of a third point $x^{(3)}$ to the already available two-vertex b.c.p. $\mathcal{P}^{(2)} = \text{absco}(\{x^{(1)}, x^{(2)}\}) \subseteq \mathbb{C}^2$ in order to construct $\mathcal{P}^{(3)} = \text{absco}(\{x^{(1)}, x^{(2)}, x^{(3)}\})$ in the case that $x^{(3)} \notin \mathcal{P}^{(2)}$ and does not delete none of the vertices of $\mathcal{P}^{(2)}$.

This construction may be done, substantially, with the following three steps.

- Find the *deleting interval* $\mathcal{D}_{12} \subseteq (-\pi, \pi]$ representing all the facets (apart from scalar factors of unitary modulus)

$$F_{y_{12}(\theta)} : x^{(1)} \bullet\!\!-\!\!\bullet\, e^{\mathrm{i}\theta}x^{(2)}$$

of $\mathcal{P}_{12}$ to delete, that is (using Theorem 2.1), s.t.

$$|\langle y_{12}(\theta), x^{(3)}\rangle| > 1,$$

where $y_{12}(\theta) = ([x^{(1)}, x^{(2)}]^*)^{-1} \begin{bmatrix} 1 \\ e^{\mathrm{i}\theta} \end{bmatrix}$.

- Define the *existence interval* of surviving facets $F_{y_{12}(\theta)}$ as $\mathcal{E}_{12} = (-\pi, \pi] \setminus \mathcal{D}_{12}$.
- Compute the *existence intervals* $\mathcal{E}_{13}, \mathcal{E}_{23} \subseteq (-\pi, \pi]$ of the facets to add, which are respectively

$$F_{y_{13}(\theta)} : x^{(1)} \bullet\!\!-\!\!\bullet\, e^{\mathrm{i}\theta} x^{(3)} \quad \text{for } \theta \text{ s.t.} \quad |\langle y_{13}(\theta), x^{(2)}\rangle| \leq 1,$$

$$F_{y_{23}(\theta)} : x^{(2)} \bullet\!\!-\!\!\bullet\, e^{\mathrm{i}\theta} x^{(3)} \quad \text{for } \theta \text{ s.t.} \quad |\langle y_{23}(\theta), x^{(1)}\rangle| \leq 1.$$

Before ending the characterization of a b.c.p. with three essential vertices, we first need to classify the facets of a b.c.p. with three or more essential vertices as stated in the following definitions.

Definition 3.1 *A facet F_y of a b.c.p. $\mathcal{P}$ is called* regular *if it contains exactly two vertices.*

Definition 3.2 *A facet F_y of a b.c.p. $\mathcal{P}$ is called* special *if it contains three or more vertices, that is if there exist $x^{(i_1)}, \cdots, x^{(i_s)} \in \mathcal{X}$, with $3 \leq s \leq m$, such that the associated functional y satisfies*

$$\langle y, \mathrm{e}^{\mathrm{i}\theta_j} x^{(i_j)}\rangle = 1 \qquad \textit{for suitable} \quad \theta_j,\ j = 1, \cdots, s.$$

It is clear that, a special facet F_y is the union of all possible *triangles* of the type $e^{\mathrm{i}\theta_{j_1}} x^{(i_{j_1})} \blacktriangle e^{\mathrm{i}\theta_{j_2}} x^{(i_{j_2})} \blacktriangle e^{\mathrm{i}\theta_{j_3}} x^{(i_{j_3})} = \mathrm{co}\left(\left\{e^{\mathrm{i}\theta_{j_1}} x^{(i_{j_1})}, e^{\mathrm{i}\theta_{j_2}} x^{(i_{j_2})}, e^{\mathrm{i}\theta_{j_3}} x^{(i_{j_3})}\right)\right\}$, i.e.

$$F_y = \bigcup_{1 \leq j_1 < j_2 < j_3 \leq s} e^{\mathrm{i}\theta_{j_1}} x^{(i_{j_1})} \blacktriangle e^{\mathrm{i}\theta_{j_2}} x^{(i_{j_2})} \blacktriangle e^{\mathrm{i}\theta_{j_3}} x^{(i_{j_3})}.$$

Note that, unlike the real case, the dimension on $\mathbb{R}$, thought as a part of an affine subspace of $\mathbb{R}^4$, is equal to 2, whereas the dimension on $\mathbb{R}$ of a regular facet is obviously equal to 1.

Moreover, it can be proved that the intersection with $\mathbb{R}^2$ of the triangles contained in special facets of $\mathcal{P}$ is given by two pairs of symmetric arcs of ellipsis, where one or even both may degenerate into a pair of symmetric straight segments.

Definition 3.3 *A segment $x^{(i)} \bullet\!\!-\!\!\bullet\, \mathrm{e}^{\mathrm{i}\theta_j} x^{(j)}$ containing two vertices of a b.c.p. $\mathcal{P}$ is called* isolated facet *if it is contained in a facet of $\mathcal{P}$ and if all the segments $x^{(i)} \bullet\!\!-\!\!\bullet\, \mathrm{e}^{\mathrm{i}(\theta+\theta_j)} x^{(j)}$ are not facets of $\mathcal{P}$ for all $\theta \in (-\theta_0, \theta_0) \setminus \{0\}$ for some $\theta_0 > 0$.*

We have proved the following Theorem related to isolated facets, which, as we can see at the end of this section, will play an important role in our constructive algorithm of a b.c.p.

Theorem 3.1 *None of the vertices of a b.c.p. $\mathcal{P}$ may belong only to isolated facets.*

Now, we conclude the construction of a b.c.p. with three essential vertices with the next Theorem. To this aim, we first note that its existence intervals are of the form:

$$\mathcal{E}_{12} \quad : \quad [\theta_{12}^-, \theta_{12}^+] \quad \text{or} \quad (-\pi, \theta_{12}^+] \cup [\theta_{12}^-, \pi]$$

$$\mathcal{E}_{i3} \quad : \quad [\theta_{i3}^-, \theta_{i3}^+] \quad \text{or} \quad (-\pi, \theta_{i3}^+] \cup [\theta_{i3}^-, \pi], \quad i = 1, 2.$$

Thus, we have

Theorem 3.2 *Let $\mathcal{P} = absco(\{x^{(1)}, x^{(2)}, x^{(3)}\})$, where $x^{(1)}, x^{(2)}, x^{(3)} \in \mathbb{C}^2$ represent an e.s.v. Then $\mathcal{P}$ has exactly two special facets with three vertices (modulo scalar factors of unitary modulus) given by*

$$x^{(1)} \blacktriangle e^{\mathrm{i}\theta_{12}^+} x^{(2)} \blacktriangle e^{\mathrm{i}\theta_{13}^-} x^{(3)} \qquad \textit{and} \qquad x^{(1)} \blacktriangle e^{\mathrm{i}\theta_{12}^-} x^{(2)} \blacktriangle e^{\mathrm{i}\theta_{13}^+} x^{(3)},$$

where $\theta_{12}^{\pm}$ and $\theta_{13}^{\pm}$ satisfy

$$\theta_{12}^+ + \theta_{23}^+ = \theta_{13}^- \mathrm{Mod} 2\pi \qquad \textit{and} \qquad \theta_{12}^- + \theta_{23}^- = \theta_{13}^+ \mathrm{Mod} 2\pi.$$

We conclude the analysis of a three vertex b.c.p. with the following Example which shows its intersection with the real plane.

Example 3.1 Let $\mathcal{P}$ be b.c.p. with essential vertices $x^{(1)} = [1; i], x^{(2)} = [1; 1-i], x^{(3)} = [1; 1]$. In order to compute $\partial\mathcal{P} \cap \mathbb{R}^2$, we first determine the intersection with $\mathbb{R}^2$ of each of the three 2-vertex subpolytopes $\partial\mathcal{P}_{ij}$ for $i, j = 1, 2, 3$ with $i < j$, which is reported on the left of Figure 1 and then, the intersection of the two special facets which is reported on the right of the same figure by the blue ellipse and by the pair of symmetric straight green lines. So, we can conclude that $\partial\mathcal{P} \cap \mathbb{R}^2$ is given by the black curve on the right of the figure.

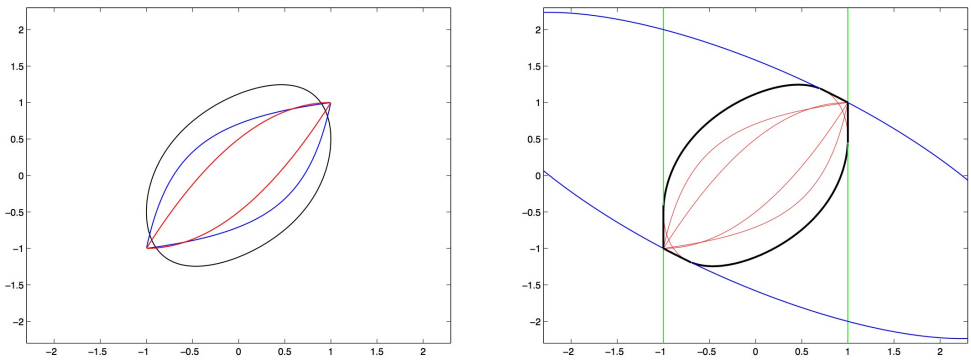


Figure 1

We conclude this subsection by examining the general case of a b.c.p. with $m \geq 4$ essential vertices, starting with the generalization of the definition of existence interval given for a 3-vertex b.c.p.

Definition 3.4 *Let $\mathcal{P} = absco(\mathcal{X})$, where $\mathcal{X} = \{x^{(i)}\}_{1\leq i\leq m}$ is an e.s.v. with $m \geq 4$. The* existence pluri-interval *of the facets $F_{y_{ij}(\theta)} = x^{(i)} \bullet\!\!-\!\!\bullet\, \mathrm{e}^{\mathrm{i}\theta}x^{(j)}$ of $\mathcal{P}$ is the set $\mathcal{E}_{ij} \subseteq (-\pi, \pi]$ s.t. $\theta \in \mathcal{E}_{ij}$ if and only if the segment $x^{(i)} \bullet\!\!-\!\!\bullet\, \mathrm{e}^{\mathrm{i}\theta}x^{(j)}$ is a regular facet or included in a special facet of $\mathcal{P}$.*

It can be proved that $\mathcal{E}_{ij}$ is the union of p_{ij} disjoint intervals, $0 \leq p_{ij} \leq m-1$, that is, for appropriate $\theta^-_{ij,l}, \theta^+_{ij,l} \in [-\pi, \pi]$, $\mathcal{E}_{ij}$ is of the form

$$\mathcal{E}_{ij} = \bigcup_{l=1}^{p_{ij}} [\theta^-_{ij,l}, \theta^+_{ij,l}] \setminus \{-\pi\}.$$

We conclude the characterization of the boundary of a b.c.p. $\mathcal{P}$ with $m \geq 4$ essential vertices with the following Theorem which allow us to compute the (triangles contained in) special facets of $\mathcal{P}$, from the knowledge of its existence pluri-intervals.

Theorem 3.3 *For each triplet (i, j, k) with $1 \leq i < j < k \leq m$, there exist at most two triangles of the type $x^{(i)} \blacktriangle \mathrm{e}^{\mathrm{i}\theta_{\mathrm{ij}}} x^{(j)} \blacktriangle \mathrm{e}^{\mathrm{i}\theta_{\mathrm{ik}}} x^{(k)}$ contained in special facets of $\mathcal{P}$. Moreover, such triangles, if any, are of the type*

$$x^{(i)} \blacktriangle \mathrm{e}^{\mathrm{i}\theta^+_{\mathrm{ij}}} x^{(j)} \blacktriangle \mathrm{e}^{\mathrm{i}\theta^-_{\mathrm{ik}}} x^{(k)} \qquad with \quad \theta^+_{ij} + \theta^+_{jk} = \theta^-_{ik} \quad \mathrm{Mod}2\pi$$

and/or

$$x^{(i)} \blacktriangle \mathrm{e}^{\mathrm{i}\theta^-_{\mathrm{ij}}} x^{(j)} \blacktriangle \mathrm{e}^{\mathrm{i}\theta^+_{\mathrm{ik}}} x^{(k)} \qquad with \quad \theta^-_{ij} + \theta^-_{jk} = \theta^+_{ik} \quad \mathrm{Mod}2\pi,$$

where $[\theta^-_{rs}, \theta^+_{rs}]$ stays for one of the intervals of the existence pluri-interval $\mathcal{E}_{rs}$, $r, s \in \{i, j, k\}$.

3.2 Algorithms for the construction of a b.c.p. $\mathcal{P}$ and for the computation of $\|\cdot\|_{\mathcal{P}}$

We begin this subsection by presenting our algorithm for the construction of a b.c.p. $\mathcal{P} = absco(\mathcal{V})$, where $\mathcal{V} = \{x^{(i)}\}_{1\leq i\leq m}$ is a set of vectors of $\mathbb{C}^2$ where we assume $x^{(1)}$ and $x^{(2)}$ linearly independent.

The algorithm works in an iterative fashion starting from

$$\mathcal{P}^{(2)} = absco(\mathcal{X}^{(2)}), \quad \text{where} \quad \mathcal{X}^{(2)} = \{x^{(1)}, x^{(2)}\}.$$

Then, from the already available b.c.p. $\mathcal{P}^{(k-1)} = absco(\mathcal{X}^{(k-1)})$, $k \geq 3$, we construct $\mathcal{P}^{(k)}$ by adding $x^{(k)}$ to $\mathcal{P}^{(k-1)}$. So, the k-th step of the constructive algorithm may be summarized as follows.

Algorithm 3.1 (k-th step)

- For $1 \leq i < j \leq k-1$, consider the 2-vertex subpolytopes $\mathcal{P}_{ij} = absco(\{x^{(i)}, x^{(j)}\})$ of $\mathcal{P}^{(k-1)}$ and write $x^{(k)} = \lambda_{ij}^{(k)} x^{(i)} + \mu_{ij}^{(k)} x^{(j)}$ with $\lambda_{ij}^{(k)}, \mu_{ij}^{(k)} \in \mathbb{C}$.

  If $\|x^{(k)}\|_{\mathcal{P}_{ij}} = |\lambda_{ij}^{(k)}| + |\mu_{ij}^{(k)}| \leq 1$ for some (i,j)
  $x^{(k)} \in \mathcal{P}^{(k-1)}$ is deleted
  else
  - $\mathcal{X}^{(k)} = \mathcal{X}^{(k-1)} \cup \{x^{(k)}\}$
  - Remove from $\mathcal{X}^{(k)}$ all those vectors, if any, which are deleted by $x^{(k)}$ as vertices of a 2-vertex subpolytope $\mathcal{P}_{ij}$ and detect those facets of $\mathcal{P}^{(k)}$ which were already facets of $\mathcal{P}^{(k-1)}$ and survive the addition of $x^{(k)}$
  - Add the new facets of $\mathcal{P}^{(k)}$ whose second vertex is $x^{(k)}$
  - Remove from $\mathcal{X}^{(k)}$ all those vectors which belong only to isolated facets or to no facets at all (in the light of Theorem 3.1)

  end

Once we have processed the last vector of $\mathcal{V}$, we have found the existence pluri-intervals of all the regular facets of $\mathcal{P}$ and an essential system of vertices $\mathcal{X} \subseteq \mathcal{V}$ for $\mathcal{P}$.

Finally, in the light of Theorem 3.3, the algorithm ends with the detection of all the (triangles included in) special facets of $\mathcal{P}$ by analyzing its existence pluri-intervals. Remark that any triangle $x^{(i)} \blacktriangle e^{i\theta} x^{(j)} \blacktriangle e^{i\phi} x^{(k)}$ included in the special facets of $\mathcal{P}$ is characterized by its *existence pair* $S_{ijk} = \{\theta, \phi\}$.

We conclude this subsection by giving our algorithm for the computation of the complex polytope norm $\|z\|_{\mathcal{P}}$, for any $z \in \mathbb{C}^2$.

We recall that the input of the algorithm is the b.c.p. $\mathcal{P} = absco(\mathcal{X})$ characterized by:

- its essential system of vertices $\mathcal{X} = \{x^{(1)}, x^{(2)}, \ldots, x^{(m)}\}$,
- its regular facets, described by existence pluri-intervals $\mathcal{E}_{ij}$,
- the triangles contained in special facets, described by existence pairs S_{ijk}.

Algorithm 3.2 (computation of $\|z\|_{\mathcal{P}}$)

If z is proportional to $x^{(i)}$
$\|z\|_{\mathcal{P}} = \|z\| / \|x^{(i)}\|$
else
compute $n_{rs} = \min_{1 \leq i < j \leq m} \|z\|_{\mathcal{P}_{ij}} = \min_{1 \leq i < j \leq m} (|\lambda_{ij}| + |\mu_{ij}|)$, where the indexes r, s reach the minimum and

$$\begin{bmatrix} \lambda_{ij} \\ \mu_{ij} \end{bmatrix} = [x^{(i)}\, x^{(j)}]^{-1} z$$

if $\arg(\lambda_{rs}) - \arg(\mu_{rs}) \in \mathcal{E}_{rs}$

% z projects on $e^{\mathrm{i}arg(\lambda_{rs})}x^{(r)}$ •–• $e^{\mathrm{i}arg(\mu_{rs})}x^{(s)}$
$\|z\|_{\mathcal{P}} = |\lambda_{rs}| + |\mu_{rs}|$
else
% z projects on $x^{(i)} \blacktriangle e^{\mathrm{i}\theta}x^{(j)} \blacktriangle e^{\mathrm{i}\phi}x^{(k)}$
find the existence pair $S_{ijk} = \{\theta, \phi\}$, s. t. there exist
$\lambda_i, \lambda_j, \lambda_k > 0 : z = \lambda_i x^{(i)} + \lambda_j e^{\mathrm{i}\theta}x^{(j)} + \lambda_k e^{\mathrm{i}\phi}x^{(k)}$,
$\|z\|_{\mathcal{P}} = \lambda_i + \lambda_j + \lambda_k$
end
end

4 Computational aspects

From our numerical experiments, made using MATLAB, on the previous algorithms, we have seen that they are computationally expensive. Indeed, the k-th step of Algorithm 3.1 consider all the facets of all the two-vertex subpolytopes $\mathcal{P}_{ij} = \mathrm{absco}(\{x^{(i)}, x^{(j)}\})$ of $\mathcal{P}^{(k-1)}$, which may uselessly be too time consuming. For this reason, in this section, we present an improvement, where, in general, only a subset of the regular facets of $\mathcal{P}^{(k-1)}$ is involved. This improvement is based on the extension to $\mathbb{C}^2$ of the *limit cone* idea used in the Beneath-Beyond (B–B) method to construct of real polytopes (see [2]). In order to show this idea we recall the k-th step of the method in $\mathbb{R}^2$, that is the addition of $x^{(k)}$ to $\mathcal{P}^{(k-1)}$, whose steps are also shown in Figure 2 in two different cases.

Algorithm 4.1 (k-ht step)

- Determine the facet $x^{(i)}$ •–• $x^{(j)}$ on which $x^{(k)}$ projects, which is the one intersected by the line $\lambda x^{(k)}$, $\lambda > 0$;

If if $x^{(k)} \in \mathcal{P}_{ij} = absco(\{x^{(i)}, x^{(j)}\})$
delete $x^{(k)}$
else
- delete the facets of $\mathcal{P}^{(k-1)}$ which are seen by $x^{(k)}$, that is, the facets inside the limit cone of apex $x^{(k)}$
- add the facets whose second vertex is $x^{(k)}$, by only using the two vertices belonging to the boundary of the limit cone

end

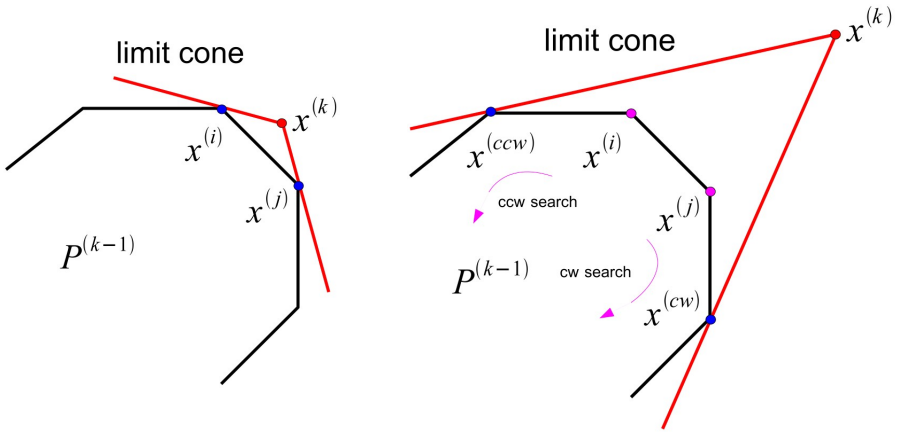


Figure 2. The point $x^{(k)}$ projects on the facet $F_{i,j} : x^{(i)} \bullet\!\!-\!\!\bullet\, x^{(j)}$. On the left, none of the existing vertices have to be removed since the limit cone does not contain any of them (i.e. $x^{(ccw)} = x^{(i)}$ and $x^{(cw)} = x^{(j)}$). On the right, the two magenta vertices $x^{(i)}$ and $x^{(j)}$ have to be removed since they lie inside the limit cone. These two vertices no longer belong to the set of vertices of $\mathcal{P}^{(k)}$.

Our aim is to use the limit cone idea also in $\mathbb{C}^2$, so as to be allowed to check only a minimal subset of the regular facets of $\mathcal{P}^{(k-1)}$.

The *balanced limit cone* in $\mathbb{C}^2$ of apex $x^{(k)}$, which is tangent to $\mathcal{P}^{(k-1)}$, delimits the set D of the regular facets of $\mathcal{P}^{(k-1)}$ which are *seen* (and so deleted) by the *circle* generated by $x^{(k)} \notin \mathcal{P}^{(k-1)}$. Since $\mathcal{P}^{(k-1)}$ is convex, D is connected and thus, in order to update the regular facets of $\mathcal{P}^{(k-1)}$ due to the addition of $x^{(k)}$, we can start from any seen facet and then find, moving by connection, all the other regular facets which are seen by the circle generated by $x^{(k)}$. Consequently, we have to add only the facets $x^{(l)} \bullet\!\!-\!\!\bullet\, e^{\mathrm{i}\theta} x^{(k)}$, where $x^{(l)}$ belongs to the set of the non-deleted vertices of seen facets.

To perform the search of a seen facet, if any, we propose a first criterion in order to guess those regular facets of $\mathcal{P}$ that have the greatest chances to be seen by $x^{(k)}$. This criterion is based on the reasonable assumption that, in most cases, a facet which is seen by $x^{(k)}$ includes vertices that are among the closest, in the Euclidean distance, to

$$R_{x^{(k)}} = \{\rho x^{(k)} \mid \rho > 0\}.$$

Therefore, $\forall i = 1, \ldots, k-1$, we compute the Euclidean distances

$$\delta_i = \min_{-\pi < \theta \leq \pi,\, \rho > 0} \|\mathrm{e}^{\mathrm{i}\theta} x^{(i)} - \rho x^{(k)}\|_2 = \sqrt{\|x^{(i)}\|_2^2 - |\langle x^{(i)}, x^{(k)} \rangle|^2 / \|x^{(k)}\|_2^2}$$

of $R_{x^{(k)}}$ from the circle generated by $x^{(i)}$, and we reorder the indexes in non decreasing order with respect to the distances δ_i. Subsequently, we define a total order relation "$\prec$" on the set of the index pairs (i, j) reordered in reversed lexicographical way, that is,

$$(i, j) \prec (h, k) \iff j < k \quad \text{o} \quad (j = k \ \ \& \ \ i < h);$$

next following this total order relation, we find, if any, the first seen facet and then moving by connection, also all the other seen facets of $\mathcal{P}^{(k-1)}$.

Also the algorithm for the computation of the polytope norm may unnecessarily involve all the two-vertex subpolytopes $\mathcal{P}_{ij}$ of $\mathcal{P}$. So, we use the same criterion already used to improve the procedure for the computation of $\|z\|_{\mathcal{P}}$, $z \in \mathbb{C}^2$. First, we compute for $i = 1, \ldots, m$ the Euclidean distances δ_i of $R_z = \{\rho z \mid \rho > 0\}$ from the circle generated by the m essential vertices of $\mathcal{P}$; next, we reorder the indexes of these m vertices in non-decreasing order with respect to the distances δ_i; then we define, as before, a total order relation "$\prec$" on the set of the index pairs (i, j), and we extend the total order relation "$\prec$" to the triplets (i, j, k), $1 \leq i < j < k \leq m$, by inserting a triplet (r, s, t) soon after the last of the three pairs (r, s), (r, t), (s, t). In this way, following this total order relation, we can process, one after the other, pairs and triplet of indexes, until we find the facet, regular or special, on which z projects.

The speed-up obtained by these improvements is confirmed by the numerical tests; for more details see [8].

5 Conclusions

To summarise, in this work we have deepened the theoretical study of the geometry of a b.c.p. in $\mathbb{C}^2$ presenting the first efficient algorithm for the construction of a balanced complex polytope $\mathcal{P}$ in $\mathbb{C}^2$ which completely describes the geometry of $\mathcal{P}$. Furthermore, we have also presented the first efficient algorithm to compute the complex polytope norm of a vector $z \in \mathbb{C}^2$ starting from the knowledge of the boundary of the corresponding unit ball.

References

[1] R. K. Brayton and C. H. Tong, *Stability of dynamical systems: a constructive algorithm.* IEEE Trans. Circuits Systems 26 (1979), 224–234.

[2] H. Edelsbrunner, "Algorithms in combinatorial geometry". EATCS Monographs on Theoretical Computer Science, Spring-Verlag, Heidelberg, 1987.

[3] N. Guglielmi and F. Wirth and M. Zennaro, *Complex polytope extremality results for families of matrices.* SIAM J. Matrix Anal. Appl. 27 (2005), 721–743.

[4] N. Guglielmi and M. Zennaro, *An algorithm for finding extremal polytope norms of matrix families.* Linear Algebra Appl. (to appear).

[5] N. Guglielmi and M. Zennaro, *Balanced complex polytopes and related vector and matrix norms.* J. Convex Anal. 14 (2007), 729–766.

[6] S. Miani and C. Savorgnan, *Complex polytopic control Lyapunov functions.* 45th IEEE Conference on Decision and Control (San Diego, CA, USA 13–15 December 2006), 13–15 December.

[7] G. C. Rota and G. Strang, *A note on the joint spectral radius.* Indag. Math. 22 (1960), 379–381.

[8] C. Vagnoni, "Algorithms for the computation of the joint spectral radius". PhD Thesis, Università degli Studi di Padova, 2008.

[9] C. Vagnoni and M. Zennaro, *The analysis and the construction of balanced complex polytopes in 2d.* Foundation of Computational Mathematics (to appear).

Cluster algebras: some motivating examples for their introduction

Giovanni Cerulli Irelli (*)

Abstract. In this note we give some of the ideas behind the introduction of the theory of cluster algebras. This theory was first introduced in [CAI] by S. Fomin and A. Zelevinsky in 2001. From their introduction, cluster algebras have found place in several different fields of mathematics. After its introduction the theory by itself has been developed by now in three papers, [CAII], [CAIII] and [CAIV]. We follow the very good survey of the subject given by [FZNotes].

Sunto. Questa nota illustra a grandi linee alcune delle motivazioni che hanno spinto S. Fomin e A. Zelevinsky a far nascere e sviluppare la teoria delle algebre cluster. Questa teoria nasce nel 2001 allo scopo di creare una struttura algebrica nella quale studiare i concetti di positivitá totale e base canonica in gruppi algebrici semi–semplici. Qui si richiamano brevemente alcuni dei risultati preliminari alla teoria che sono molto ben descritti in [FZNotes] la quale rimane la principale referenza di queste note.

1 Double–Bruhat cells and total positivity

Let $G = SL_n(\mathbb{C}) = \{X = (x_{ij}) \in M_{n\times n}(\mathbb{C}) | \, det(X) = 1\}$ be the group of complex $n \times n$ matrices with determinant 1. Let $W = Sym_n$ be the symmetric group generated by the $n-1$ simple transpositions $s_1, \cdots, s_{n-1}$ where $s_i = (i, i+1)$ denotes the transposition of i and $i+1$. We see W as a subgroup of G in the natural way.

Inside G we consider the (Borel) subgroups B and B_- respectively of the upper and lower–triangular matrices. A *Bruhat cell* in G is the double coset BwB or B_-wB_- for $w \in W$. It is known that G is disjoint union of Bruhat cells: $G = \bigsqcup_{w\in W} BwB = \bigsqcup_{w\in W} B_-wB_-$. A *double Bruhat cell* is the intersection of two Bruhat cells:

$$G^{(u,v)} \doteq BuB \cap B_-vB_-$$

for $u, v \in W$. In particular $G = \sqcup_{(u,v)\in W\times W} G^{(u,v)}$. Let us describe a double Bruhat cell geometrically. Recall that given an element w of the symmetric group W, the *lenght* of w,

(*)Ph.D. School in Pure Mathematics. Università di Padova, Dipartimento di Matematica Pura ed Applicata, Via Trieste 63, 35121 Padova, Italy. E-mail: gcerulli@math.unipd.it . Seminar held on 13 February 2008.

denoted by $l(w)$, is the shortest lenght of a sequence of indices $\mathbf{i} = (i_1, \cdots, i_l)$ such that $w = s_{i_1} \cdots s_{i_l}$. Such a sequence is called a *reduced word* for w.

Theorem 1.1 [1, Theorem 1.1] *For every $u, v \in W$ the double Bruhat cell $G^{(u,v)}$ is isomorphic to a Zarisky open subset of an affine space of dimension $n + l(u) + l(v)$.*

Roughly speaking $G^{(u,v)}$ is isomorphic as an algebraic variety to a subset of $\mathbb{C}^{n+l(u)+l(v)}$ obtained by excluding common zeroes of a finite set of polynomials.

For two subsets I and J of $[1, n] \doteq \{1, \cdots, n\}$ of the *same* size, we denote by $\Delta_{I,J}$ the minor with row set I and column set J. We now give a description of a Bruhat cell BwB. We use the notations: $[1, i\,] \doteq \{1, \cdots, i\}$ and for $w \in W$, $w([1, i\,]) \doteq \{w(1), \cdots, w(i)\}$.

Proposition 1.2 [1, Proposition 4.1] *Let $x \in G$, $w \in W$. Then $x \in BwB$ if and only if*

(a) $\Delta_{w([1,i\,]),[1,i\,]} \neq 0$ *for* $i = 1, 2, \cdots, n-1$;

(b) $\Delta_{w([1,i-1\,]\cup\{j\}),[1,i\,]} = 0$ *for every* $1 \leq i < j \leq n$ *such that* $w(i) < w(j)$.

To illustrate the previous result let us consider the case in which $n = 3$ and $w = w_0 = (13)$ is the longest element of Sym_3. The first condition says that $\Delta_{3,1} \neq 0$ and $\Delta_{[2,3],[1,2]} \neq 0$; the second condition is not applied in this case (this is a general fact: if $w = w_0$ is the longest element of Sym_n then there are not $i < j$ such that $w_0(i) < w_0(j)$). We hence conclude that the Bruhat cell Bw_0B is given by

$$\begin{aligned} Bw_0B &= \{x \in SL_3(\mathbb{C}) |\, \Delta_{3,1}(x) \neq 0;\, \Delta_{[2,3],[1,2]} \neq 0\} \\ &= \{x \in SL_3(\mathbb{C}) |\, x_{31} \neq 0;\, x_{21}x_{32} - x_{22}x_{31} \neq 0\}. \end{aligned}$$

The transposition morphism $x \mapsto x^t$ sends $\Delta_{I,J}$ to $\Delta_{J,I}$ and the Bruhat cell BwB onto the Bruhat cell $B_-w^{-1}B_-$. One can hence use Proposition 1.2 in order to find the analogous result for the Bruhat cell $B_-w^{-1}B_-$. In particular we get the following description of the double Bruhat cell $G^{(u,v)}$.

Proposition 1.3 *Let $x \in G$, $u, v \in W$. Then $x \in G^{(u,v)}$ if and only if*

(a) $\Delta_{u([1,i\,]),[1,i\,]} \neq 0$ *for* $i = 1, 2, \cdots, n-1$;

(b) $\Delta_{u([1,i-1\,]\cup\{j\}),[1,i\,]} = 0$ *for every* $1 \leq i < j \leq n$ *such that* $u(i) < u(j)$;

(c) $\Delta_{[1,i\,],v^{-1}([1,i\,])} \neq 0$ *for* $i = 1, 2, \cdots, n-1$;

(d) $\Delta_{[1,i\,],v^{-1}([1,i-1\,]\cup\{j\})} = 0$ *for every* $1 \leq i < j \leq n$ *such that* $v^{-1}(i) < v^{-1}(j)$.

To illustrate the previous Proposition we continue our running example $n = 3$, $w = w_0 = (13)$: we get

$$\begin{aligned} G^{(w_0,w_0)} &= \{x \in SL_3(\mathbb{C}) |\, \Delta_{3,1}(x) \neq 0;\, \Delta_{[2,3],[1,2]} \neq 0,\, \Delta_{1,3}(x) \neq 0;\, \Delta_{[1,2],[2,3]} \neq 0\} \\ (1.1) &= \{x \in SL_3(\mathbb{C}) |\, x_{31} \neq 0;\, x_{21}x_{32} - x_{22}x_{31} \neq 0,\, x_{13} \neq 0;\, x_{12}x_{23} - x_{22}x_{13} \neq 0\}. \end{aligned}$$

This is a general result: if w_0 is the longest element of G then $G^{(w_0,w_0)}$ is the *open* double Bruhat cell

$$G^{(w_0,w_0)} = \{x \in G |\ \Delta_{w_0([1,i]),[1,i]}(x) \neq 0;\ \Delta_{[1,i],w_0([1,i])}(x) \neq 0\,,\ \text{for all } i \in [1,n]\}.$$

We denote by $G_{\geq 0}$ and $G_{>0}$ respectively the varieties of *totally nonnegative* and *totally positive* matrices, i.e. matrices whose minors are all non–negative (resp. positive) real numbers. The *totally positive* part of a double Bruhat cell $G^{(u,v)}$ is $G^{(u,v)}_{>0} \doteq G^{(u,v)} \cap G_{\geq 0}$.

Proposition 1.4 [1] *The totally positive part of the open double Bruhat cell $G^{(w_0,w_0)}$ is the totally positive variety.*

To illustrate the previous Proposition we study in detail our running example, $n = 3$, $w = w_0$: we consider the following regular functions on G:

$$\begin{array}{c} f_1 = \Delta_{1,3} = x_{13},\ \ f_2 = \Delta_{12,23}, \\ f_3 = \Delta_{1,2} = x_{12},\ f_4 = \Delta_{12,12},\ f_5 = \Delta_{1,1} = x_{11},\ f_6 = \Delta_{2,1} = x_{21}, \\ f_7 = \Delta_{23,12},\ f_8 = \Delta_{3,1} = x_{31}. \end{array} \tag{1.2}$$

Let $F = \{f_i |\ i = 1, \cdots, 8\}$ be the set of such minors. In particular the open double Bruhat cell is given by:

$$G^{(w_0,w_0)} = \{x \in G |\ f_k(x) \neq 0 \text{ for every } k \in \{1,2,7,8\}\}. \tag{1.3}$$

The following Lemma illustrate the previous Proposition.

Lemma 1.5 *Given an element x of $G = SL_3(\mathbb{C})$, every minor of x is a Laurent polynomial with non–negative integer coefficients in the variables $f_1, \cdots, f_8$. If all minors of $x \in SL_3(\mathbb{C})$ are nonnegative and x is in the open double Bruhat cell, i.e. $x \in G^{(w_0,w_0)}_{>0}$, then x is totally positive.*

Dimostrazione. The proof is by direct check. We give the Laurent expansion in $f_1, \cdots, f_8$ of every minor.

Minors of order 1

$$\begin{array}{lll} x_{11} = f_5, & x_{12} = f_3, & x_{13} = f_1, \\ x_{21} = f_6, & x_{22} = \frac{\Delta_{12,12}+x_{21}x_{12}}{x_{11}} = \frac{f_4+f_6f_3}{f_5}, & x_{23} = \frac{\Delta_{12,23}+x_{13}x_{22}}{x_{12}} = \frac{f_2f_5+f_1f_4+f_1f_3f_6}{f_3f_5}, \\ x_{31} = f_8, & x_{32} = \frac{\Delta_{23,12}+x_{22}x_{31}}{x_{21}} = \frac{f_5f_7+f_4f_8+f_3f_6f_8}{f_5f_6}. & \end{array}$$

Since the determinant of x is 1 we have that

$$\begin{aligned} 1 &= x_{31}\Delta_{12,23} - x_{32}\Delta_{12,13} + x_{33}\Delta_{12,12} = \cdots \\ &= f_8f_2 - \frac{(f_5f_7 + f_4f_8 + f_3f_6f_8)(f_2f_5 + f_1f_4)}{f_3f_5f_6} + x_{33}f_4 \end{aligned}$$

from which we get:

$$x_{33} = \frac{f_3f_5f_6 + (f_5f_7 + f_4f_8)(f_2f_5 + f_1f_4) + f_1f_3f_4f_6f_8}{f_3f_4f_5f_6}.$$

Minors of order two By using minors of order one it is not difficult to get the following equalities:

$$\begin{aligned}
\Delta_{12,12} &= f_4,\\
\Delta_{12,13} &= x_{11}x_{23} - x_{13}x_{21} = \frac{f_2f_5 + f_1f_4}{f_3},\\
\Delta_{12,23} &= f_2,\\
\Delta_{13,12} &= x_{11}x_{32} - x_{12}x_{31} = \frac{f_5f_7 + f_4f_8}{f_6},\\
\Delta_{13,13} &= x_{11}x_{33} - x_{13}x_{31} = \frac{f_3f_5f_6 + (f_5f_7 + f_4f_8)(f_2f_5 + f_1f_4)}{f_3f_4f_6},\\
\Delta_{13,23} &= x_{12}x_{33} - x_{13}x_{32} = \frac{f_3f_6 + f_2f_5f_7 + f_2f_4f_8}{f_4f_6},\\
\Delta_{23,12} &= f_7,\\
\Delta_{23,13} &= x_{21}x_{33} - x_{23}x_{31} = \frac{f_3f_6 + f_2f_5f_7 + f_1f_4f_7}{f_3f_4},\\
\Delta_{23,23} &= x_{22}x_{33} - x_{23}x_{32} = \frac{1 + f_3f_6 + f_2f_5f_7}{f_4f_5}
\end{aligned}$$

Suppose now that all the minors of a matrix $x \in SL_3(\mathbb{C})$ are nonnegative and $f_k(x) > 0$ for $k = 1, 2, 7, 8$; it then follows from the previous formulas that $f_k(x) > 0$ also for $k = 3, 4, 5, 6$ and hence x is totally positive. □

Definition 1.6 *A* Total Positive basis *(TP–basis) for $G^{(u,v)}$ is a collection of regular functions $F = \{f_1, \cdots, f_m\} \subset \mathbb{C}[G^{(u,v)}]$ such that*

(a) $f_1, \cdots, f_m$ are algebraically independent and generate the field $\mathbb{C}(G^{(u,v)})$. In particular $m = n + l(u) + l(v)$;

(b) $(f_1, \cdots, f_m) : G^{(u,v)} \to \mathbb{C}^m$ restricts to a bi–regular isomorphism $U(F) \to \mathbb{C}^m_{\neq 0}$ where

$$U(F) \doteq \{x \in G^{(u,v)} |\, f_k(x) \neq 0 \ \forall k = 1, \cdots, m\}; \tag{1.4}$$

(c) $(f_1, \cdots, f_m) : G^{(u,v)} \to \mathbb{C}^m$ restricts to an isomorphism $G^{(u,v)}_{>0} \to \mathbb{R}^m_{>0}$.

The third condition in Definition 1.6 gives rise to a *total positivity criterion in $G^{(u,v)}$*: a matrix $x \in G^{(u,v)}$ is totally nonnegative if and only if $f_k(x) > 0$ for every $k = 1, \cdots, m$.

A reduced word for $(u, v) \in W \times W$ is a sequence $\mathbf{i} = (i_1, \cdots, i_{l(u)+l(v)})$ of indices $i_k \in [1, n]$. Following the construction of [1], we add the indices $1, \cdots, n$ at the beginning of $\mathbf{i}$ and we highlight indices denoting u with the sign minus. The new sequence will be again called a reduced word of (u, v), denoted with $\mathbf{i}$ but it becomes the sequence $(1, \cdots, n, i_1, \cdots, i_{l(u)+l(v)})$ of length $m = n + l(u) + l(v)$ and $i_k \in -[1, n] \cup [1, n]$. For example if $u = s_1s_2s_1$ and $v = s_1s_2$, $\mathbf{i} = (1, 2, -1, -2, -1, 1, 2)$.

Definition 1.7 *Let* $\mathbf{i} = (i_1, \cdots, i_m) = (1, \cdots, n, i_{n+1}, \cdots, i_m)$ *be a reduced word for* $(u, v) \in W \times W$ *(in the previous sense). For* $k \in [1, m]$ *we define the the multi–indices* γ_k *and* δ_k *as* $\gamma_k = s_{-i_1} \cdots s_{-i_k}[1, |i_k|\,]$ *and* $\delta_k = s_{i_m} \cdots s_{i_{k+1}}[1, |i_k|\,]$ *with the convention* $s_{-i} = 1$ *for* $i \in [1, n]$.

Theorem 1.8 *Each reduced word* $\mathbf{i}$ *of* $(u, v) \in W \times W$ *gives rise to a TP–basis*

$$F_{\mathbf{i}} = \{\Delta_{\gamma_k, \delta_k} : k \in [1, m]\}$$

where the multi–indices γ_k *and* δ_k *are given in Definition 1.7.*

To illustrate the previous theorem, in $G = SL_3(\mathbb{C})$ we choose $u = v = (1, 3) = s_1 s_2 s_1$ and $\mathbf{i} = (1, 2, 1, 2, 1, -1, -2, -1)$. Then

$$\begin{aligned}
\gamma_1 &\doteq 1 & \delta_1 &= s_1 s_2 s_1 s_2[1] = 3 \\
\gamma_2 &\doteq [12] & \delta_2 &= s_1 s_2 s_1[12] = [23] \\
\gamma_3 &\doteq 1 & \delta_3 &= s_1 s_2[1] = 2 \\
\gamma_4 &\doteq [12] & \delta_4 &= s_1[12] = [12] \\
\gamma_5 &\doteq 1 & \delta_5 &= 1 \\
\gamma_6 &\doteq s_1[1] = 2 & \delta_6 &= 1 \\
\gamma_7 &\doteq s_1 s_2[12] = [23] & \delta_7 &= [12] \\
\gamma_8 &\doteq s_1 s_2 s_3[1] = 3 & \delta_8 &= 1
\end{aligned}$$

and we get $f_k = \Delta_{\gamma_k, \delta_k}$, $k = 1, \cdots, 8$, defined in (1.2).

In [1] they discovered that TP–bases can be "mutated" one into another by a combinatorial mechanism. This is the main idea behind the theory of cluster algebras.

Definition 1.9 *Let* $\mathbf{i} = (i_1, \cdots, i_m)$ *be a reduced word for* $(u, v) \in W \times W$. *We say that an index* $k \in [1, m]$ *is* $\mathbf{i}$–exchangeable *if*

(a) $n < k \leq m$;

(b) $|i_p| = |i_k|$ *for some* $p > k$.

We denote by $ex_{\mathbf{i}}$ *the set of* $\mathbf{i}$*–exchangeable indices.*

In the previous example $\mathbf{i} = (1, 2, 1, 2, 1, -1, -2, -1)$ and hence $ex_{\mathbf{i}} = [3, 6]$.

Lemma 1.10 *The subset* $\mathbf{c} = \{f_k : k \in [1, m] \setminus \mathbf{ex_i}\} \subset F_{\mathbf{i}}$ *depends only on* u *and* v*, not on the particular choice of a reduced word* $\mathbf{i}$. *In particular the cardinality* n *of* $\mathbf{ex_i}$ *depends only on* u *and* v. *Moreover each* $f_k \in \mathbf{c}$ *vanishes nowhere in* $G^{(u,v)}$.

In the previous example $\mathbf{c} = \{f_1, f_2, f_7, f_8\}$. An immediate corollary of this result is that the ring of Laurent polynomials $\mathbb{C}[\mathbf{c}^{\pm 1}]$ in the elements of $\mathbf{c}$ is contained in $\mathbb{C}[G^{(u,v)}]$.

Theorem 1.11 *There exists an* $m \times n$ *integer matrix* $\tilde{B}(\mathbf{i}) = (b_{ij})$ *such that*

- *for every* $k \in \mathbf{ex_i}$ *the function*

$$f'_k = \frac{\prod_{i:b_{ik}>0} f_i^{b_{ik}} + \prod_{i:b_{ik}<0} f_i^{-b_{ik}}}{f_k} \tag{1.5}$$

is regular on $G^{(u,v)}$ *(i.e.* $f'_k \in \mathbb{C}[G^{(u,v)}]$*) and the collection*

$$F_{\mathbf{i},k} \doteq F_{\mathbf{i}} \setminus \{f_k\} \cup \{f'_k\} \subset \mathbb{C}[G^{(u,v)}]$$

is a TP–basis for $G^{(u,v)}$*;*

- $\mathbb{C}[G^{(u,v)}] = \mathbb{C}[U]$ *where* U *is the Zariski open subset*

$$U = U(F_{\mathbf{i}}) \bigcup_{k \in \mathbf{ex_i}} U(F_{\mathbf{i},k})$$

and $U(F_{\mathbf{i}})$ *is defined in (1.4).*

Corollary 1.12 *The subalgebra* $\mathbb{C}[G^{(u,v)}]$ *of the field of rational functions* $\mathbb{C}(G^{(u,v)})$ *is the intersection of* $n+1$ *Laurent polynomial rings:*

$$\mathbb{C}[G^{(u,v)}] = \mathbb{C}[F_{\mathbf{i}}^{\pm 1}] \cap \bigcap_{k \in \mathbf{ex}} \mathbb{C}[F_{\mathbf{i},k}]$$

This corollary says that $\mathbb{C}[G^{(u,v)}]$ has a structure of *upper cluster algebra*. Every cluster algebra is contained in an upper cluster algebra. Sometimes it happens that they coincide. This is the case for example when $u = v = c = s_{i_1} \cdots s_{i_n}$ is a Coxeter element of W, i.e. the indices i_k are all distinct.

The definition of the matrix $\tilde{B}(\mathbf{i})$ is purely combinatorial. A particular importance for our point of view has the submatrix $B(\mathbf{i})$ of $\tilde{B}(\mathbf{i})$ with rows and columns parameterized by $ex_{\mathbf{i}}$. The submatrix $B(\mathbf{i})$ is called the *exchange matrix* of $F(\mathbf{i})$. In our example: $u = v = (1,3)$ and $\mathbf{i} = (1,2,1,2,1,-1,-2,-1)$ the two matrices are given by:

$$\tilde{B}(\mathbf{i}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & -1 & 1 \\ -1 & 1 & 0 & -1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad B(\mathbf{i}) = \begin{bmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & -1 & 1 \\ -1 & 1 & 0 & -1 \\ 0 & -1 & 1 & 0 \end{bmatrix}. \tag{1.6}$$

Note that the matrix $B(\mathbf{i})$ is skew–symmetric. In general $B(\mathbf{i})$ has a weaker property:

Proposition 1.13 [CAIII, Proposition 2.6] *The matrix* $\tilde{B}(\mathbf{i})$ *has full rank* n*. The exchange matrix* $B(\mathbf{i})$ *is* skew–symmetrizable*, i.e.* $d_i b_{ij} = -d_j b_{ji}$ *for some positive integers* $d_1, \cdots, d_n$*.*

We conclude this section by saying that all the previous results hold in a semi–simple complex, connected, simply–connected algebraic group G by substituting minors with *generalized* minors, that are particular regular functions.

2 The algebra of diagonals

We consider two classical combinatorial problems:

(a) Counting the number of products (bracketings) of $n+2$ elements with respect to a non–associative operation;

(b) counting the number of triangulations by non–crossing diagonals of a regular $(n+3)$–agon.

There is a natural bijection between bracketings of $n+2$ elements $a_1, \cdots, a_{n+2}$ and triangulations by (non–crossing) diagonals of the regular $n+3$–agon P_n: it consists in labeling all the edges but one of P_n by $a_1, \cdots, a_{n+2}$ and in associating to a triangulation T of P_n the bracketing in which a_i and a_j are in the same product if they label two edges of the same triangle of T. In Figure 1 we enumerate the five triangulations of a pentagon by $T_1, \cdots, T_5$ and we show the bijection with the five possible bracketings of four elements.

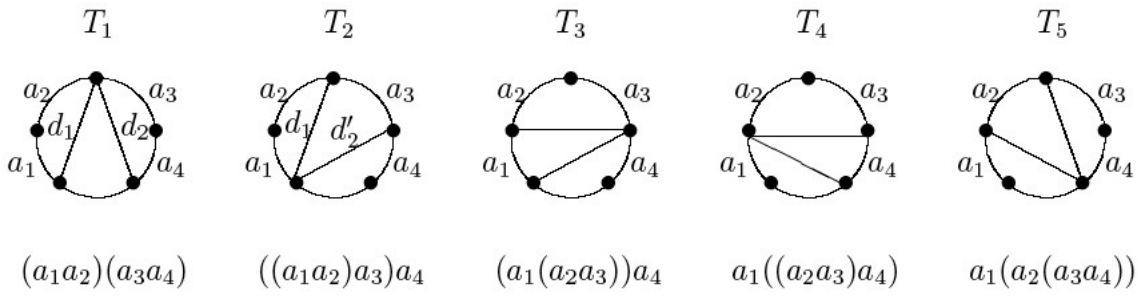


Figure 1: Triangulations of a pentagon and bracketings of four elements

The solution of the two problems is given by the $n+1$–th Catalan number (see e.g. [FR] and its bibliography):

$$c_{n+1} \doteq \frac{1}{n+2}\binom{2n+2}{n+1}.$$

In the first problem n is the number of (couples of) parenthesis while in the second problem n is the number of diagonals of every triangulation of P_n. We denote by $m = n+(n+3)$ the number of diagonals and edges of every triangulation of P_n. The *flip* mutation transforms a triangulation T of the $n+3$–agon P_n into the triangulation T' that have all the diagonals of T but one diagonal, say d, that is replaced by the unique diagonal d' of P_n such that d and d' are diagonals of a quadrilateral of elements of T (or T').

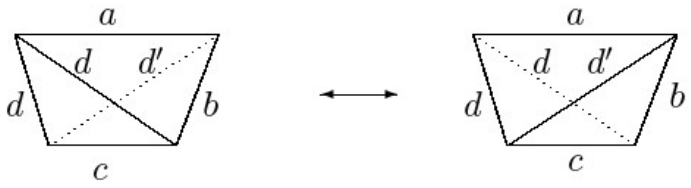


Figure 2: Diagonal Flip

We associate to P_n a graph called *exchange graph* whose vertices are triangulations by non–crossing diagonals of P_n and whose edges are diagonal flips. Such graph is well known in literature and it is proved that it is the 1–skeleton of a convex polytope called *n–dimensional associahedron* or Stasheff's polytope (see [FR] for a good survey on associahedra and their connection with cluster algebras). Figure 3 shows the exchange graph of a pentagon that coincide with the 2–dimensional associahedron.

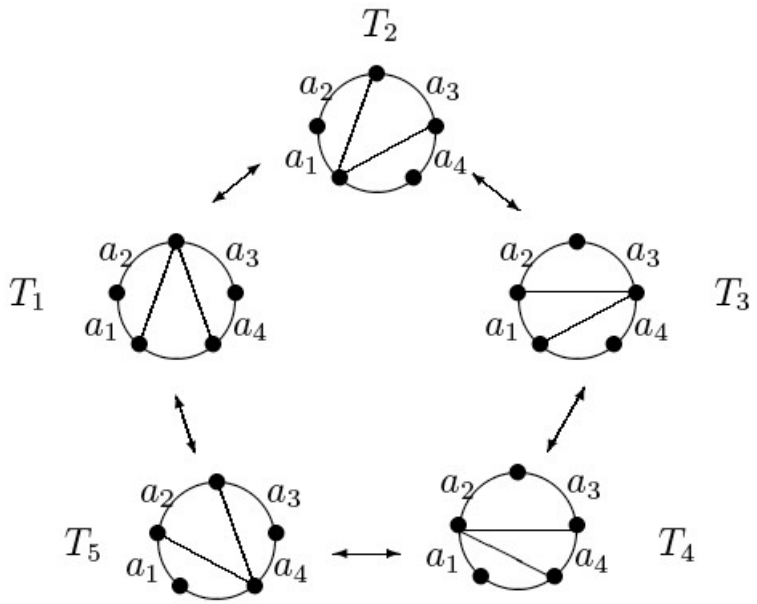


Figure 3: Exchange graph of a pentagon and two dimensional associahedron

To every triangulation T of P_n we associate an $m \times n$ integer matrix $\tilde{B}(T)$ whose columns are parameterized by edges and diagonals of T and whose rows only by diagonals and whose entries are given by:

$$b_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are sides in some triangle of } T \\ & \text{and } j \text{ follows } i \text{ in the clockwise order} \\ -1 & \text{if the same holds, in the counter–clockwise order} \\ 0 & \text{otherwise} \end{cases}$$

To illustrate: the matrix associated with the triangulation T_1 of Figure 1 has the columns parameterized by the diagonals d_1 and d_2 and the rows parameterized by both the diagonals d_1 and d_2 and by the edges $a_1, \cdots, a_4$. It is the following

	d_1	d_2
d_1	0	1
d_2	-1	0
a_1	-1	0
a_2	1	0
a_3	0	-1
a_4	0	1

Diagonal flips are hence translated into *matrix mutations* as follows: if a triangulation T' is obtained to the triangulation T by "flipping" the diagonal d, the respective matrices $\tilde{B}(T) = (b_{ij})$ and $\tilde{B}(T') = (b'_{ij})$ are related to each other by

$$b'_{ij} = \begin{cases} -b_{ij} & \text{if } i = d \text{ or } j = d \\ b_{ij} + sg(b_{id})[b_{id}b_{dj}]_+ & \text{otherwise} \end{cases} \tag{2.1}$$

where $[x]_+ \doteq \max(x, 0)$ and $sg(x)$ is the sign function with $sg(0) = 0$. In general given an $m \times n$ matrix B, we denote by $\mu_k(B)$ the matrix obtained from B by a matrix mutation (2.1) in direction $k \in [1, n]$, i.e. where k plays the role of d.

To illustrate we mutate in direction 1 the matrix associated with the triangulation T_1 of Figure 1. We get $\tilde{B}(T_2)$:

$$\tilde{B}(T_1) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \\ -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} \overset{\mu_1}{\longleftrightarrow} \begin{bmatrix} 0 & -1 \\ 1 & 0 \\ 1 & 0 \\ -1 & 1 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} = \tilde{B}(T_2)$$

We are now able to introduce the *algebra of diagonals* $\mathcal{A}_n$ of P_n: it has $\binom{n+3}{2}$ generators (associated to every diagonal and edge of P_n) with the following relations: if a diagonal d' is obtained from a diagonal d by a diagonal flip shown in Figure 2 then the corresponding variables x_d and $x_{d'}$ satisfy

$$x_d x_{d'} = x_a x_c + x_b x_d. \tag{2.2}$$

The $\binom{n+3}{4}$ relations (2.2) are called *Ptolemy relations* since of the classical Ptolemy theorem which asserts that the product of the lengths of two diagonals of an inscribed quadrilateral equals the sum of the product of the lengths of opposite sides.

A model for the algebra $\mathcal{A}_n$ is the coordinate ring $\mathbb{C}[X_n]$ of the variety X_n of all decomposable bi–vectors in $\bigwedge^2 \mathbb{C}^{n+3}$. An element x of X_n is written as $x = \sum_{i<j} P_{ij} e_i \wedge e_j$ and P_{ij} are called the Plücker coordinates of x in the standard basis $e_i \wedge e_j$ of $\bigwedge^2 \mathbb{C}^{n+3}$. To identify $\mathcal{A}_n$ with $\mathbb{C}[X_n]$ we associate to a chord $[i, j]$ (i.e. either a diagonal or an edge

between the distinct vertices i and j) of P_n the Plücker coordinate P_{ij}. The Ptolemy relations (2.2) translate into the *Grassmann–Plücker relations* $P_{ik}P_{jl} = P_{ij}P_{kl} + P_{il}P_{jk}$ for all $1 \leq i < j < k < l \leq n+3$.

One can develop a total positivity theory in X_n completely parallel to that one in Double–Bruhat cells mentioned in Section 1. We say that an element x of X_n is *totally positive* if all its Plücker coordinates are positive real numbers. We then say that a collection of elements of $\mathcal{A}_n = \mathbb{C}[X_n]$ is a *Total positive basis* (TP–basis) for X_n, if it satisfies hypothesis of Definition 1.6 where $m = 2n+3$.

Theorem 2.1 [CAII] *Every triangulation T of P_n gives rise to a TP–basis $\tilde{x}(T)$ for X_n which consists of the $2n+3$ generators x_a corresponding to the sides and diagonals of T.*

If a triangulation T' of P_n is obtained from the triangulation T by the flip of a diagonal d, then one can easily check that the Ptolemy relation (2.2) between the corresponding variables $x_k \doteq x_d$ and $x_{k'} \doteq x_{d'}$ is completely analogous to (1.5) i.e.

$$x'_k = \frac{\prod_{i:b_{ik}>0} x_i^{b_{ik}} + \prod_{i:b_{ik}<0} x_i^{-b_{ik}}}{x_k} \tag{2.3}$$

where $\tilde{B}(T) = (b_{ij})$ is the matrix associate with the triangulation T.

We saw in Section 1 that the TP–basis $F_{\mathbf{i}}$ can be mutated in n directions giving rise to other TP–bases. Here we can also continue mutating. The first concept is related with the concept of upper cluster algebra while the second one with that one of *cluster algebra*.

3 Definition of a cluster algebra

Let $m \geq n > 0$ two positive integers. Let $\mathcal{F}$ be the field of rational functions in m commuting variables. A *seed* in $\mathcal{F}$ is a couple $\Sigma = (\tilde{B}, \tilde{\mathbf{x}})$ where

- the first element $\tilde{B} = (b_{ij})$ is an $m \times n$ integer matrix whose principal part $B = (b_{ij})_{i,j=1,\cdots,n}$ is *skew–symmetrizable*, i.e. $d_i b_{ij} = -d_j b_{ji}$ for some positive integers $d_1, \cdots, d_n$. B is called the *exchange matrix* of Σ.
- The second element $\tilde{\mathbf{x}} = (x_1, \cdots, x_m)$ is an m–tuple of elements of $\mathcal{F}$ forming a free generating system for $\mathcal{F}$, i.e. $\mathcal{F} \simeq \mathbb{Q}(x_1, \cdots, x_m)$. The n–tuple $\mathbf{x} = (x_1, \cdots, x_n)$ is called the *cluster* of the seed Σ and its elements are the *cluster variables* of Σ. The set $\mathbf{c} = \{x_{n+1}, \cdots, x_m\}$ is called the set of coefficients.

In the previous sections we have seen that only *some* elements of a TP–basis need to be mutated in order to get another TP–basis while the others can remain the same. "Mutating" elements correspond to cluster variables while the others correspond to coefficients (which explains the notation), as we see in the following definition.

Definition 3.1 [Seed mutations] *Let $\Sigma = (\tilde{B}, \tilde{\mathbf{x}})$ be a seed in $\mathcal{F}$ and let $k \in I = [1, n]$. The* seed mutation μ_k in direction k *transforms the seed Σ into the seed $\mu_k(\Sigma) = (\tilde{B}', \tilde{\mathbf{x}}')$ defined as follows:*

- *The entries of* $\tilde{B}' = (b'_{ij})$ *are given by (2.1);*
- *The set* $\tilde{\mathbf{x}}' = \{x'_1, \cdots, x'_m\}$ *is given by* $x'_j = x_j$ *for* $j \neq k$*, while* $x'_k \in \mathcal{F}$ *is determined by the* exchange relation *(2.3).*

It is easy to see that μ_k is involutive and hence mutations define an equivalence relation in the class of seeds of $\mathcal{F}$: two seeds Σ and Σ' are equivalent if there exists a sequence $\{\mu_{k_1}, \mu_{k_2}, \cdots, \mu_{k_s}\}$ of mutations such that $\Sigma' = \mu_{k_s} \cdots \mu_{k_1} \Sigma$. We denote by $\mathcal{O}(\Sigma)$ the equivalence class of Σ and with $\chi(\Sigma)$ the set of cluster variables in $\mathcal{O}(\Sigma)$, that is the set of cluster variables of every seed in $\mathcal{O}(\Sigma)$.

The *cluster algebra* $\mathcal{A} = \mathcal{A}(\Sigma)$ with initial seed Σ is the $\mathbb{Z}[\mathbf{c}^{\pm 1}]$-subalgebra of $\mathcal{F}$ generated by cluster variables in $\mathcal{O}(\Sigma)$, in symbols: $\mathcal{A}(\Sigma) \doteq \mathbb{Z}[\mathbf{c}^{\pm 1}][\chi(\Sigma)]$.

A cluster algebra is called of *finite type* if it has finitely many clusters. In [CAII] the authors show that cluster algebras of finite type are classified by Dynkin diagrams of finite type, i.e. of type $A, B, \cdots, G$. Some coordinate rings of classical algebraic varieties have a structure of cluster algebras of finite type. This is a list of some of them:

	cluster type
$\mathbb{Q}[Gr_{2;n+3}]$	A_n
$\mathbb{Q}[Gr_{3;6}]$	D_4
$\mathbb{Q}[Gr_{3;7}]$	E_6
$\mathbb{Q}[Gr_{3;8}]$	E_8
$\mathbb{Q}[SL_2]$	A_1
$\mathbb{Q}[SL_3]$	D_4

where $Gr_{k;n}$ denotes the Grassmannians of k–vector spaces in an n–dimensional one. On the other hand in [CAII] and more recently in [2], authors give a geometric realization of every cluster algebra of finite type. In [3] the authors study in details cluster algebras of rank two, i.e. in which every cluster has cardinality two. In my phd thesis I have studied cluster algebras of rank three.

References

[CAIII] Arkady Berenstein, Sergey Fomin and Andrei Zelevinsky, *Cluster Algebras III:Upper bounds and double Bruhat cells.* Duke Math. J. 126 (2005), 1–52.

[FR] S. Fomin, Nathan Reading, "Root system and generalized associahedra". Arxiv: CO/0505518, 31 May 2005.

[FZ] S. Fomin, A. Zelevinsky, "Cluster algebras: Notes for the CDM-03 conference". math.RT/0311493.

[CAI] S. Fomin and A. Zelevinsky, *Cluster Algebras I: Foundations.* J. Amer. Math. Soc. 15 (2002), 497–529.

[CAII] S. Fomin and A. Zelevinsky, *Cluster Algebras II: Finite type classification.* Invent. Math. 154 (2003), 63–121.

[CAIV] S. Fomin and A. Zelevinsky, *Cluster Algebras IV: Coefficients.* Compositio Math. 143 (2007), 112–164.

[1] S. Fomin and A. Zelevinsky, *Double Bruhat cells and total positivity.* J. Amer. Math. Soc. 12 (1999), 335–380.

[2] S. Yang and A. Zelevinsky, *Cluster algebras of finite type via Coxeter elements and principal minors.* Preprint 2008, `arXiv:math.RA/0804.3403v1`.

[3] P. Sherman and A. Zelevinsky, *Positivity and canonical bases in rank 2 cluster algebras of finite and affine types.* Moscow Math. J. 4 (2004), no. 4, 947–974.

Computing VaR and CVaR for energy derivatives

Giorgia Callegaro (*)

1 Introduction

Due to the peculiarity of energy markets (think for example of the spot price seasonality or of transport and storability problems in the case of electricity and gas) many "variable volume" options have been introduced.
They are purchase or sale contracts which provide flexibility about the timing to delivery and about the overall minimum and maximum take amounts, usually called "swing" of "take or pay" options.

We focused our attention on the computation of risk measures related to this kind of options, namely we considered the "Value at Risk" (VaR) and the "Conditional Value at Risk" (CVaR) and we numerically computed them relatively to an investment in swing options.

2 From the spot price model to swing options

We will be interested in Swing options referring to *gas*. Since the spot price of the gas is not a traded quantity, usually one refers to prices of forward contracts with different maturities, denoted $(F_{s,t})_{s\in[0,T]}$ (where t is the maturity), so that the spot price at time t is equal to

$$S_t = F_{t,t} = \lim_{T\to t} F_{t,T}.$$

For simplicity we will work on the following log-normal model for the forward curve $F_{t,T}$, known as the "one-factor model", by L. Clewlow and C. Strickland (see [3]):

$$\frac{\mathrm{d}F_{t,T}}{F_{t,T}} = \sigma(t,T)\mathrm{d}W_t = \sigma e^{-\theta(T-t)}\mathrm{d}W_t, \quad t\in[0,T],$$

(*)Scuola Normale Superiore, I-56100 Pisa, Italy; Université d'Évry Val d'Essonne, F-91025 Évry Cedex; g.callegaro@sns.it. Seminar held on 28 February 2008.
This work was done during a five-months internship at "Gaz the France", Research and Development Division, supervised by M. Olivier Bardou (Gaz de France) and Prof. Gilles Pagès (Université Paris VI).

where $(W_t)_{0\le t\le T}$ is a standard Brownian motion on a (completed) filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P})$. By using Itô's Formula we find

$$S_t = F_{0,t} \exp\left\{ Y_{t,t} - \frac{1}{2}\Sigma_t^2 \right\}, \tag{1}$$

where the initial conditions are given and specified by the forward curve $F(0,t)_{0\le t\le T}$ and

$$Y_{t,T} = \sigma e^{-\theta T} \int_0^t e^{\theta s} \mathrm{d}W_s, \qquad \Sigma_t^2 = \frac{\sigma^2}{2\theta}(1 - e^{-2\theta t}). \tag{2}$$

Let us now describe a general Swing contract, following [2]. We fix a finite time horizon T and some intermediate dates $0 = t_0 < t_1 < \cdots < t_N = T$. We denote by q_{t_i} the volume of gas purchased at time t_i by the buyer of the contract, which is subject to the constraints

$$q_{\min} < q_{t_i} < q_{\max}, \qquad i = 0, \cdots, N-1, \tag{3}$$

and, introducing the cumulative volume of gas bought up to time t_i, $Q_{t_i} = \sum_{j=0}^{i-1} q_{t_j}$, we also want the following relation to be satisfied

$$Q_{\min} < Q_T < Q_{\max}.$$

Assuming that the purchase price is fixed and equal to K, the buyer of the contract at time t_i gets $q_{t_i}(S_{t_i} - K)$ and so, given the strategy $q = (q_{t_i})_{0\le i<N}$ we find that the value at time 0 of such an option is given by

$$\sum_{i=0}^{N-1} e^{-rt_i} q_{t_i}(S_{t_i} - K) + e^{-rT} P_T(S_T, Q_T),$$

where r is the interest rate and $P_T(S,Q)$ is the penalty function at maturity T for the case in which global purchase constraints are violated.
Its fair price at time 0 for a given $q = (q_{t_i})_{0\le i<N}$, is

$$\mathcal{P}(0,q) = \mathbb{E}\left\{ \sum_{i=0}^{N-1} e^{-rt_i} q_{t_i}(S_{t_i} - K) + e^{-rT} P_T(S_T, Q_T) \right\}, \tag{4}$$

where the expectation is taken under the risk neutral measure, and so to value such an option we have to compute

$$\sup_{q\in Adm} \mathcal{P}(0,q). \tag{5}$$

We have obtained a stochastic control problem, that can be solved using the Dynamic Programming (DP) Principle, under some additional hypotheses.
In [2] the evaluation problem is solved in a more general setting using the DP together with the *vectorial quantization* method (for an introduction to the method see [9] and for further results see [6] and [7]).

We will use the instruments developed in [2] to determine the optimal strategy q^* to compute risk measures associated to the random "investment"

$$X = \sum_{i=0}^{N} q_{t_i}^*(S_{t_i} - K). \tag{6}$$

3 Risk measures: VaR vs CVaR

If our gain is given by a real random variable X (as for example (6) and F is its distribution function ($F : \mathbb{R} \to [0,1]$, $F(x) = \mathbb{P}(X \leq x)$) we have

Definition 3.1 *Given $\alpha \in]0,1[$, the "Value at Risk" of level α is the $\alpha-$quantile of the X distibution, which is the real value*

$$\mathrm{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}. \tag{7}$$

Since VaR is not able to quantify the amount of gains that are in the right tail of the distribution, in recent years the attention has moved to new risk measures, such as the "Conditional Value at Risk" (CVaR).

Definition 3.2 *The CVaR is defined as*

$$\mathrm{CVaR}_\alpha(X) = \mathbb{E}(X | X \geq \mathrm{VaR}_\alpha(X)).$$

It quantifies the gains (or losses) that might be encountered in the $\alpha-$tail of the distribution.

The following result is the starting point of our work, it is due to R.T. Rockafellar and S. Uryasev and it can be found in "Optimization of conditional value-at-risk", Journal of Risk 2 (2000), 21-41 (see also [8]).

Proposition 3.3 *As a function of $\xi \in \mathbb{R}$*

$$F_\alpha(X,\xi) := \xi + \frac{1}{1-\alpha}\mathbb{E}\left[(X-\xi)_+\right] \tag{8}$$

is finite and convex (hence continuous) and we have

$$\mathrm{CVaR}_\alpha(X) = \min_\xi F_\alpha(X,\xi), \qquad \mathrm{VaR}_\alpha(X) = \xi^* \in \mathit{argmin}_\xi\, F_\alpha(X,\xi). \tag{9}$$

Remark 3.4 The minimization problem above can be transformed into a stochastic-root finding one by writing the first order optimality conditions

$$\min_\xi F_\alpha(X,\xi) = F_\alpha(X,\xi^*) \iff \frac{\partial F_\alpha}{\partial \xi}(X,\xi^*) = 0.$$

The solution ξ^* to the former problem can be asymptotically estimated using stochastic algorithms, if no explicit formula is available.

4 The Robbins-Monro algorithm(s)

In order to approximate the zero of a certain function h which is not known directly, but can be estimated via noisy observations, Robbins and Monro, in 1951, proposed the following general algorithm:

$$\xi_0 \text{ given}, \quad \xi_{n+1} = \xi_n - \pi_{n+1}\bar{h}(\xi_n),\ n \geq 0, \tag{10}$$

where

- $(\pi_n)_{n\geq 1}$ is an sequence of deterministic positive weights satisfying

$$\pi_n \to 0, \quad \sum_{n\geq 1} \pi_n = \infty, \quad \sum_{n\geq 1} \pi_n^2 < \infty,$$

  for example $\pi_n = \frac{C}{n}, \quad \forall n \geq 1$.

- $\bar{h}(\xi_n)$ is a "noisy" estimate of the value $h(\xi_n)$.

In order to solve problem (9) when X is given by (6), since in this case the distribution of X is not easily determined (and so is $F_\alpha(X,\xi)$), we can apply R-M stochastic recursive algorithms. If we define (see Remark 3.4)

$$H(\xi, X) := \xi + \frac{1}{1-\alpha}(X-\xi)_+, \quad \frac{\partial H(\xi,X)}{\partial \xi} = 1 - \frac{1}{1-\alpha}\mathbb{I}_{\{X>\xi\}} < +\infty \quad \mathbb{P}\text{–a.s.} \tag{11}$$

in order to determine $\mathrm{VaR}_\alpha(X) = \xi^*$ we have to find the zero of

$$h(\xi) := \mathbb{E}\left[\frac{\partial H(\xi,X)}{\partial \xi}\right]. \tag{12}$$

Adapting equation (10), the R-M algorithm consists in simulating a sequence $(X_n)_{n\geq 0}$ of copies i.i.d. of X, and then in applying the recursion

$$\xi_{n+1} = \xi_n - \pi_{n+1}\frac{\partial H(\xi_n, X_{n+1})}{\partial \xi}, \tag{13}$$

given the initial point ξ_0.

Remark 4.1 [The choice of ξ_0 and $(\pi_n)_n$] When trying to determine $\mathrm{VaR}_\alpha(X)$, we can for example choose $\xi_0 = \mathbb{E}(X)$ and $\pi_n = \frac{C}{n} \quad \forall n \geq 1$. What about C? As we will see in the examples, it has to be chosen depending on the problem and it has to depend on the order of magnitude of the quantities we deal with.

Remark 4.2 [Convergence] In order to prove the convergence of algorithm (13) it suffices to adapt more general theorems of convergence of stochastic recursive algorithms, such as for example [4, Ths. 1.4.26 and 2.2.12] and [1, Ths. 5 and 6].

5 Importance sampling as a variance reduction technique

The idea is to change the probability measure under which the paths (of X) are generated, in order to obtain a representation that gives more weight to "interesting" outcomes. This is possible thanks to Girsanov Theorem (for this we refer to any book treating stochastic calculus).
If we denote by $Z_{t,\mu}$ the Radon-Nikodym density process

$$Z_{t,\mu} = \frac{\mathrm{d}\mathbb{Q}_\mu}{\mathrm{d}\mathbb{P}}_{|\mathcal{F}_t} = \exp\left(\int_0^t \mu_s \mathrm{d}W_s - \frac{1}{2}\int_0^t ||\mu_s||^2 \mathrm{d}s\right) \tag{14}$$

(where process μ has to satisfy some suitable integrability conditions) and by $\mathbb{E}_\mu$ the expectation under $\mathbb{Q}_\mu$, if Girsanov Theorem's hypotheses apply we can write

$$\mathbb{E}[X] = \mathbb{E}_\mu\left[X\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}_\mu}\right] =: \mathbb{E}_\mu\left[XL_{T,\mu}\right]. \tag{15}$$

This means that for each "admissible" process (μ_t), we can simulate i.i.d. copies of $XL_{T,\mu}$ under measure $\mathbb{Q}_\mu$ and obtain unbiased estimators of the original value.
The importance sampling paradigm requires to choose the process (μ_t) which retains the initial expected value and minimizes the variance of the new estimator

$$\mathrm{Var}_\mu\left[X\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}_\mu}\right] = \mathbb{E}_\mu\left\{\left[XL_{T,\mu}\right]^2\right\} - \mathbb{E}_\mu\left[XL_{T,\mu}\right]^2 = \mathbb{E}\left[X^2L_{T,\mu}\right] - \mathbb{E}[X]^2.$$

The optimal importance sampling estimator has then to solve

$$\min_{\mu\in Adm}\mathbb{E}\left[X^2L_{T,\mu}\right] \Longleftrightarrow \min_{\mu\in Adm}\mathbb{E}\left[X^2e^{-\int_0^T\mu_s\mathrm{d}W_s+\frac{1}{2}\int_0^T||\mu_s||^2\mathrm{d}s}\right]. \tag{16}$$

As for (9), also in this case the function to minimize with respect to μ is not always easily accessible because of the form of X and we will resort to a second stochastic optimization algorithm.

6 VaR and CVaR with IS: a "parallel" R-M

For simplicity we will restrict our attention to the following case:

$$\mu_t \equiv \mu \tag{17}$$

and we will try to find an optimal "constant" version of μ^*, denoted $\bar{\mu}^*$.

Proposition 6.1 *In the case of swing contracts (see (6)) or sequences of call options in energy markets, with the spot derived in the one-factor model, problem (16) with $\mu_t \equiv \mu$ admits a unique solution and it is equivalent to the stochastic root finding one*

$$\frac{\partial}{\partial\mu}\left\{\mathbb{E}\left[X^2e^{-\mu W_T+\frac{1}{2}\mu^2T}\right]\right\} = 0. \tag{18}$$

To determine $\bar{\mu}^*$ that solves (18) our idea was to use a second R-M algorithm, made parallel to the previous, to approximate it with a sequence $(\mu_n)_{n\geq 0}$. This is possible introducing parameter μ_n in the first recursion.
The two optimization problems made parallel are (see (9), (11) and (16)

$$\text{(19)} \qquad \begin{cases} \min_{\xi\in\mathbb{R}} \left\{\xi + \frac{1}{1-\alpha}\mathbb{E}_\mu\left[(X-\xi)_+ L_{T,\mu}\right]\right\} & \Leftrightarrow \quad \min_{\xi\in\mathbb{R}} \mathbb{E}_\mu\left[H_1(\xi, X, \mu)\right] \\ \min_{\mu\in\mathbb{R}} \mathbb{E}\left[X^2 e^{-\mu W_T + \frac{1}{2}\mu^2 T}\right] & \Leftrightarrow \quad \min_{\mu\in\mathbb{R}} \mathbb{E}\left[H_2(\mu, X, W)\right] \end{cases}$$

and the corresponding two R-M procedures, obtained as in Section 4 (see (13), are

$$\text{(20)} \qquad \begin{cases} (\mathbb{Q}_n) \qquad \xi_{n+1} = \xi_n - \pi^1_{n+1}\dfrac{\partial H_1(\xi_n, X_{n+1}, \mu_n)}{\partial \xi}, \qquad \xi_0 \text{ given}, \\ (\mathbb{P}) \qquad \mu_{n+1} = \mu_n - \pi^2_{n+1}\dfrac{\partial H_2(\mu_n, X_{n+1}, W_{n+1})}{\partial \mu}, \quad \mu_0 \text{ given}, \end{cases}$$

where

$$\frac{\partial H_1(\xi, X, \mu)}{\partial \xi} = 1 - \frac{1}{1-\alpha} L_{T,\mu} \mathbb{1}_{\{X>\xi\}}, \quad \frac{\partial H_2(\mu, X, W)}{\partial \mu} = X^2 L_{T,\mu}(-W + \mu T)$$

and W denotes for simplicity W_T. The initial values are given (for example we choose $\mu_0 = 0$) and the weights are$\pi^1_n = \frac{C_1}{n}$, $\pi^2_n = \frac{C_2}{n}$, $\forall n \geq 1$, with constants C_1 and C_2 to be chosen wisely.

Remark 6.2 [Convergence] Because of the presence, in the first recursion, of an exponential of μ we numerically observe the explosion of the approximating sequence $(\mu_n)_n$ and so to achieve convergence we have to recur to projection methods. To show that the projected version of our variance reduction algorithm (the second one in (19)) converges it suffices to adapt the proof of Proposition 3.4.3 in [1].

7 An alternative: variance reduction as entropy minimization

As we have seen in Section 5, the optimal importance sampling process μ^* is given by the solution to problem (16), so that, since the minimum is attained at zero, we have

$$\text{(21)} \qquad \mathrm{Var}_{\mu^*}\left[X\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}_{\mu^*}}\right] = 0 \quad \text{i.e.} \quad \frac{\mathrm{d}\mathbb{Q}_{\mu^*}}{\mathrm{d}\mathbb{P}} = \frac{X}{\mathbb{E}(X)}.$$

Problem (16) can be rewritten as

$$\min_{\mu\in\mathrm{Adm}} \mathrm{Var}_\mu\left[X\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}_\mu}\right] = \min_{\mu\in\mathrm{Adm}} \mathbb{E}\left[X^2\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}_{\mu^*}}\frac{\mathrm{d}\mathbb{Q}_{\mu^*}}{\mathrm{d}\mathbb{Q}_\mu}\right],$$

so that we have transformed our original problem into the problem of determining μ that minimizes the "distance" between $\mathbb{Q}_{\mu^*}$ and $\mathbb{Q}_\mu$. As a measure of this "distance" between

two general probability distributions $\mathbb{P}$ and $\mathbb{Q}$ we used the relative entropy, which is defined as

$$\mathcal{D}_{KL}(\mathbb{Q},\mathbb{P}) = \begin{cases} \mathbb{E}_{\mathbb{Q}}\left[\ln\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right)\right] = \mathbb{E}_{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\ln\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right)\right] & \text{if } \mathbb{Q} << \mathbb{P}, \\ +\infty & \text{otherwise.} \end{cases}$$

After some manipulations, in the case when $\mu_t \equiv \mu$, the problem of minimizing the Kullback-Leibler divergence between $\mathbb{Q}_{\mu^*}$ and $\mathbb{Q}_\mu$ reduces to finding the

$$\min_{\mu\in\mathbb{R}} \mathbb{E}\left[X\left(-\mu W_T + \frac{1}{2}\mu^2 T\right)\right] \tag{22}$$

and this last problem can be solved at the same time with the determination of the VaR and CVaR (recall (19) in of Section 6, i.e. we will consider the two convex optimization problems

$$\begin{cases} \min_{\xi\in\mathbb{R}} \left\{\xi + \frac{1}{1-\alpha}\mathbb{E}_\mu\left[(X-\xi)_+ L_{T,\mu}\right]\right\} & \Leftrightarrow \quad \min_{\xi\in\mathbb{R}} \mathbb{E}_\mu\left[H_1(\xi,X,\mu)\right] \\ \min_{\mu\in\mathbb{R}} \mathbb{E}\left[X(-\mu W_T + \frac{1}{2}\mu^2 T)\right] & \Leftrightarrow \quad \min_{\mu\in\mathbb{R}} \mathbb{E}\left[H_3(\mu,X,W)\right] \end{cases} \tag{23}$$

The two corresponding R-M algorithms are then

$$\begin{cases} (\mathbb{Q}_n) \qquad \xi_{n+1} = \xi_n - \pi^1_{n+1}\dfrac{\partial H_1(\xi_n, X_{n+1}, \mu_n)}{\partial \xi}, \qquad \xi_0 \text{ given}, \\ (\mathbb{P}) \qquad \mu_{n+1} = \mu_n - \pi^3_{n+1}\dfrac{\partial H_3(\mu_n, X_{n+1}, W_{n+1})}{\partial \mu}, \quad \mu_0 \text{ given}, \end{cases} \tag{24}$$

where

$$\frac{\partial H_1(\xi,X,\mu)}{\partial \xi} = 1 - \frac{1}{1-\alpha}L_{T,\mu}\mathbb{1}_{\{X>\xi\}}, \quad \frac{\partial H_3(\mu,X,W)}{\partial \mu} = X(-W+\mu T),$$

W denotes for simplicity W_T, μ_0 can be chosen equal to zero and π^3_n is of the form $C_3/n, \forall n \geq 1$.

Remark 7.1 [Convergence] For the proof of the convergence of the sequence μ_n to $\bar{\mu}^*$ it suffices to adapt the result of Theorem 4.4.2.1 in [1].

8 Numerical results

We tested our recursive algorithms in the following three cases:

(C) $X = (S_T - K)_+$ (call option),

(SC) $X = \sum_{i=0}^{365} (S_{t_i} - K)_+$ (a sequence of call options),

(SW) $X = \sum_{i=0}^{365} q^*_{t_i}(S_{t_i} - K)$ (swing contract),

where process S satisfies equation (1), with the initial condition given by the forward curve of [2] (i.e. the time horizon is one year and $N = 365$).
For the sake of brevity, we will comment on results in the first two cases and show the most relevant ones in the swing option case. The common parameters are

$$\sigma = 0.36, \theta = 0.21, r = 0.00, T = 1.00, K = 10.0, \alpha = 0.95.$$

In all recursive procedures to obtain our simulations we used an Euler scheme with time-step equal to $\Delta t/2$ (instead of $\Delta t = 1/365$, to have more precise values).

For what concerns the computation of VaR and CVaR, i.e. the application of recursion (13), inspired by the definition of VaR and having in mind the gaussian case, we tested the following cases (recall Remark 4.1)

$$\xi_0 = p_0^{\mathrm{BS}}(X) + l \cdot \sigma(X), \qquad l \in \{1, 2, 3\} \tag{25}$$

$$\pi_n = \frac{C}{n} \quad \text{with} \quad C = \frac{m \cdot \sigma(X)}{\frac{1}{1-\alpha} - 1}, \quad m \in \{1, 2, 3\} \tag{26}$$

where p_0^{BS} is the Black-Scholes price at time zero and $\sigma(X)$ denotes the standard deviation of X. To indicate that we worked with a specific choice of l and m we used the notation "(l, m)".
Our "recipe" after n iterations is: simulate X_{n+1} and

- if $X_{n+1} > \xi_n$, then $\xi_{n+1} = \xi_n + \frac{m \cdot \sigma(X)}{n+1}$ (in this situation, with respect to our estimate ξ_n, a rare event has happened and so ξ_{n+1} will be on the right of ξ_n, shifted by a quantity depending on n);
- else if $X_{n+1} < \xi_n$, then $\xi_{n+1} = \xi_n - \frac{m \cdot \sigma(X)}{(n+1) \cdot \frac{\alpha}{1-\alpha}}$ (in this case the new estimate will be on the left of the previous one, since the simulation has not sampled in the "critical" zone).

In the R-M procedures (20) and (24) we worked with the same ξ_0 and with $\mu_0 = 0$. Furthermore, we have made the following choices

$$C_1 = C, \qquad C_2 = C_4 = \frac{1}{p_0^{BS}(X)^2}, \qquad C_3 = \frac{1}{p_0^{BS}(X)}, \tag{27}$$

in order for the three algorithms to be stable.
In the case of the standard variance reduction technique, we projected the sequence μ_n on the compact set $[-2, 2] \subset \mathbb{R}$, because some preliminary tests showed that $\bar{\mu}^*$ was very close to zero.
All the numerical results are obtained using a computer with the following technical characteristics:

OPERATING SYSTEM: Windows 2000,
PROCESSOR: Celeron,
CPU: 2.40 Ghz, RAM: 515568 KB.

8.1 The call option

The interest in treating case (C) is the presence of closed formulas, namely

$p_0^{\mathrm{BS}}(X)$	$\mathrm{var}(X)$	$\sigma(X)$	$\mathrm{VaR}_{95\%}(X)$
14.46012068	66.62334369	8.162312398	29.60797250

First of all we observed that in all the cases, $\bar{\mu}^* \in [0.535, 0.545]$ already after 75000 Monte-Carlo simulations (nMC), meaning that the R-M procedure applied to determine the optimal constant importance sampling parameter is early efficacious.
Furthermore, it was evident that the convergence to 29.60797250 was better achieved by applying a variance reduction technique on X, as it should be, and that the best results are obtained in the case $(2, 2)$.
Encouraged by these facts, we tested the algorithms in the more interesting cases (SC) and (SW), focusing our attention also to the CVaR, for which we do not have any reference value.

8.2 The sequence of call options

Since in order to run the simulations we need good approximations for $p_0^{\mathrm{BS}}(X)$ and $\sigma(X)$, we made the following choices

$$p_0^{\mathrm{BS}}(X) = \sum_{i=0}^{365} \mathrm{Call}(0, S_0; t_i, K) = 6950.923384, \tag{28}$$

$$\sigma(X) \cong \sum_{i=0}^{365} \sqrt{\mathrm{var}[(S_{t_i} - K)_+]} \cong 2950, \tag{29}$$

where the first one is the exact formula and the second one was tested to be a good approximation for the standard deviation.
For what concerns (20), as done with (C), we projected the sequence $(\mu_n)_n$ on the set $\mathcal{K} = [-2, 2]$.
First of all, the improvement given by the variance reduction was evident: the convergence speeded up, since especially in the cases $(2, 2)$ and $(2, 3)$ and already after 100000 simulations the values of VaR and CVaR became stable around approximatively 11100 and 12500 respectively.
Furthermore for different choices of m we achieved convergence to the same two values and in the case of CVaR the convergence is evident, especially under recursions (24).
Finally, $\bar{\mu}^* = 0.28$ and effectively the standard deviation of $XL_{T,\bar{\mu}^*}$ reduced from 2950 to approximatively 1030 (this result is based on a sample of 10000 simulations).
As a conclusion we can say that the variance reduction techniques we have tested enable a good convergence of $(\xi_n)_n$ to VaR and we find

$$\mathrm{VaR}_{95\%}(X) \cong 11100, \quad \mathrm{CVaR}_{95\%}(X) \cong 12500.$$

8.3 The swing option

In order to determine VaR and CVaR for the investment (SW) we based our work on the results provided in [2] relative to $p_0^{\mathrm{BS}}(X)$ (recall equation (5)), computed by means of DP and *vectorial quantization* (we can not enter here into details).
Since the computational cost increases rapidly with the number of Monte-Carlo simulations when $N = 365$, we decided to test the case of a contract lasting three months.
Furthermore, to reduce further on the simulation costs, we decided to use an Euler scheme with time step equal to Δt instead of $\Delta t/2$.
In all the cases $p_0^{\mathrm{BS}}(X)$ is computed as done in [2] with DP and quantization, while $\sigma(X)$ is estimated empirically from a sample of 10000 simulations.
The case we analyzed is the following

$$q_{\min} = 3.00, q_{\max} = 6.00, Q_{\min} = 300, Q_{\max} = 480, K = 20.0, T = 0.25$$

and we found

$$p_0^{\mathrm{BS}}(X) = 945.538498 \quad \text{and} \quad \sigma(X) \cong 900.$$

After some preliminary tests on $\bar{\mu}^*$ we chose to project the sequence $(\mu_n)_n$ in (20) on the compact set $\mathcal{K}' = [-3, 3]$.
In the following figures we present numerical results for the case $(2, 2)$, that is the best one: also in this case the convergence is evident already after 100000 Monte-Carlo iterations, especially in the CVaR case.

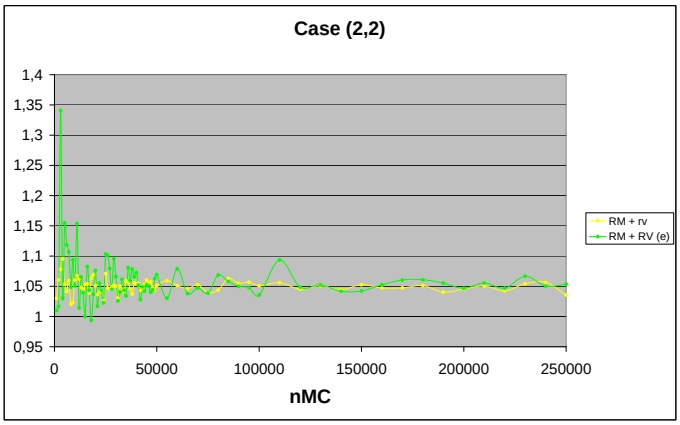


Figure 1: $\bar{\mu}^*$ as a function of nMC, obtained using the different variance reduction techniques in the case (2,2).

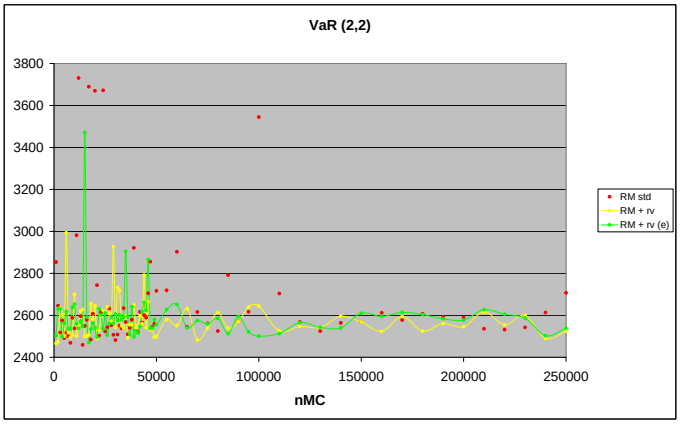


Figure 2: ξ^* as a function of nMC, obtained using different R-M procedures in the case $(2, 2)$.

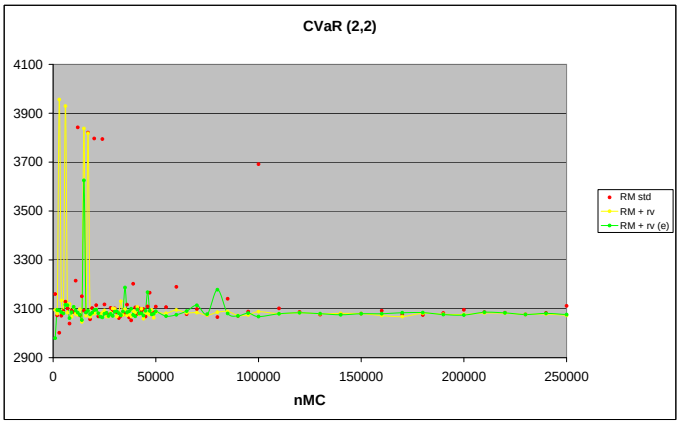


Figure 3: Conditional Value at Risk obtained using different R-M procedures in the case $(2, 2)$.

Remark 8.1 [Swing without constraints] A swing option without constraints (for example if $q_{\min} = 0.00, q_{\max} = 1.00, Q_{\min} = 0.00, Q_{\max} = 365, K = 10.0, T = 1.00$) becomes a sequence of Call options. For completeness we checked that in this case the results are comparable to the ones obtained in Section 8.2. Unfortunately we cannot enter here into details.

9 Conclusions

The most important result is that the three procedures we have used to determine VaR and CVaR, especially the ones with variance reduction are efficient, i.e. the algorithms tested converge and, furthermore, for different choices of l they converge to the same value.

In the case of swing options, even if the computational costs are higher because of the quantization method used to obtain the price simulations, the results seem to be precise, especially for the CVaR.
In this case we also know that VaR can be directly computed at the same time with the price, on the "Dynamic Programming quantized tree", but there is no possibility to compute the CVaR in the same way and so we did not concentrate on this possibility.

Since we know that VaR can be computed by means of a standard Monte-Carlo method, the main question is: is the use of Robbins-Monro type algorithms really convenient?
I think it is, since in the cases we analyzed (especially when using variance reduction techniques) convergence to our final estimate for the VaR starts with a number of Monte-Carlo simulations equal to 120000, which is not so high. Furthermore, for what concerns CVaR, which cannot be obtained by standard Monte-Carlo, we obtain estimates that converge very quickly: our algorithms provide good approximations also of this quantity.

The difficult in implementing the three variance reduction methods is almost the same, but the computational cost is not, for example because the sequence obtained when applying IS needs to be projected on a compact set, simulations are done under different measures and quantization is heavier than Euler scheme.

References

[1] B. Arouna, "Algorithmes Stochastiques et Méthodes de Monte Carlo". Thèse de doctorat de l'École Nationale des Ponts et Chaussées, 2004.

[2] S. Bouthemy, *Valorisation d'Option Swing par Quantification Vectorielle*. Rapport de stage (2006).

[3] L. Clewlow, C. Strickland, *Valuing Energy Options in a One Factor Model Fitted to Forward Prices*. Working paper, School of Finance and Economics, University of Technology, Sydney (1999).

[4] M. Duflo, "Raldom Iterative Models". Applications of Mathematics 34, Springer-Verlag, 1997.

[5] H. J. Kushner, G. G. Yin, "Stochastic Approximation and Recursive Algorithms and Applications". Springer, 2003.

[6] G. Pages and J. Printems, *Functional quantization for numerics with an application to option pricing*. Monte Carlo Methods & Applications, 11(4) (2005), 407–446.

[7] G. Pages and J. Printems, *Functional quantization for pricing derivatives*. Preprint of Laboratoire de Probabilités et Modèles Aléatoires PMA-930 (2004).

[8] R. T. Rockafellar and S. Uryasev, *Conditional Value-at-Risk for general loss distributions*. Research report #2001-5 (2001).

[9] A. Sellami, "Introduction à la quantification finidimensionelle". Février 2006, available at the web location `http://quantize.maths-fi.com/`.

Sheaves on subanalytic sites and $\mathcal{D}$-modules

Luca Prelli (*)

Abstract. We start with an introduction to sheaf theory with some examples and then we define sheaves on the subanalytic site. Thanks to these objects we can describe functional spaces which are not defined by local properties (as tempered distributions). Then we introduce the notion of $\mathcal{D}$-module to apply the preceding constructions.

Sunto. Cominceremo con un'introduzione alla teoria dei fasci con alcuni esempi e poi definiremo i fasci sul sito sottoanalitico. Grazie a questi oggetti è possibile descrivere spazi funzionali che non sono definiti da proprietà locali (come, ad esempio, le distribuzioni temperate). Introdurremo poi la nozione di $\mathcal{D}$-modulo per dare un'applicazione alle costruzioni fatte in precedenza.

1 Introduction

Let X be a real analytic manifold and k a field. The spaces of functions which are not defined by local properties, such as tempered distributions, tempered and Whitney $\mathcal{C}^\infty$ functions, etc., are very useful in the study of systems of linear partial differential equations (Laplace transform, tempered holomorphic solutions of $\mathcal{D}$-modules etc.). Although these spaces do not define sheaves on X, they define sheaves on a site associated to X, the subanalytic site X_{sa}, where one just considers open subanalytic sets and locally finite coverings.

In **Section 2** we define sheaves on topological spaces and we give some examples (continuous functions, locally constant functions, etc.). Then we show that if consider "less open" subsets and "less coverings", there are objects, which are not sheaves with the usual topology, that become sheaves (as continuous bounded functions when we consider only finite coverings). In order to do that we have to introduce the notion of topological site (a site whose objects are open subsets).

In **Section 3** we introduce the subanalytic site and we give a method to construct subanalytic sheaves. Then we see that tempered distributions define a sheaf on the subanalytic site associated to a real analytic manifold.

In **Section 4** we introduce the notion of $\mathcal{D}$-module, characteristic variety and complex of solutions of a $\mathcal{D}$-module to apply the preceding constructions.

(*)Grant holder in Pure Mathematics - Dipartimento di Matematica Pura ed Applicata, Via Belzoni 7 - 35131 Padova, Italy - lprelli@math.unipd.it - Seminar held on 12 March 2008.

References are made to [1, 2, 4] for the classical sheaf theory and to [8] for sheaves on Grothendieck topologies. The subanalytic site is defined in [5], and in [6] there is a detailed study of sheaves on subanalytic sites with the construction of the six Grothendieck operations. We refer to [3, 7] for an introduction to $\mathcal{D}$-modules.

2 Sheaves

2.1 Sheaves on topological spaces

Let X be a topological space and let k be a field. Let $\mathrm{Op}(X)$ be the category of open subsets of X, where the arrows are given by the inclusions (i.e. $U \to V$ means $U \subseteq V$).

Definition 2.1.1 *A presheaf of k-vector spaces is a contravariant functor from $\mathrm{Op}(X)$ to the category of k-vector spaces.*

In other words a presheaf associates:

(i) to each open subset U of X a k-vector space

$$\begin{array}{rcl} \mathrm{Op}(X) & \to & k\text{-vector spaces} \\ U & \mapsto & \Gamma(U;F) \quad (\text{or } F(U)), \end{array}$$

(ii) to each open inclusion $V \subseteq U$ a k-linear map (the restriction morphism)

$$\begin{array}{rcl} (V \subset U) & \mapsto & \big(F(U) \to F(V)\big) \quad \text{(restriction)} \\ & & s \mapsto s|_V. \end{array}$$

Definition 2.1.2 *Given two presheaves F e G, a morphism of presheaves $\phi : F \to G$ consists of a family $\phi(U) : F(U) \to G(U)$, (U open subset of X), commuting with restrictions, i.e. if $V \subseteq U$ the diagram*

$$\begin{array}{ccc} F(U) & \xrightarrow{\phi(U)} & G(U) \\ {\scriptstyle |_V}\downarrow & & \downarrow{\scriptstyle |_V} \\ F(V) & \xrightarrow{\phi(V)} & G(V) \end{array}$$

is commutative.

Definition 2.1.3 *A sheaf F on X is a presheaf satisfying the following gluing conditions. Let U be open and let $\{U_j\}_{j\in J}$ be a covering of U. We have the exact sequence*

$$0 \to F(U) \to \prod_{j\in J} F(U_j) \to \prod_{j,k\in J} F(U_j \cap U_k).$$

It means that

(i) if $s \in \Gamma(U;F)$ and $s|_{U_j} = 0$ for each j then $s = 0$,

(ii) if $s_j \in \Gamma(U_j;F)$ such that $s_j = s_k$ on $U_j \cap U_k$ then they glue to $s \in \Gamma(U;F)$ (i.e. $\exists\, s \in \Gamma(U;F)$ such that $s|_{U_j} = s_j$).

A morphism of sheaves is a morphism of the underlying presheaves.

Example 2.1.4 Let us consider

$$\begin{array}{rcl} \mathbb{R}_X : \text{Open sets of } X & \to & \mathbb{R}\text{-vector spaces} \\ U & \mapsto & \Gamma(U;\mathbb{R}_X) = \{\text{constant functions on } U\} \\ (V \subset U) & \mapsto & \big(\mathbb{R}_X(U) \to \mathbb{R}_X(V)\big) \quad \text{(restriction)} \\ & & s \mapsto s|_V. \end{array}$$

- If s is zero on a covering of U then $s = 0$.
- For example, let $X = \mathbb{R}$, $U_1 = (1,2)$, $U_2 = (2,3)$. We have $U_1 \cap U_2 = \varnothing$. The constant functions $s_1 = 0$ on U_1 and $s_2 = 1$ on U_2 do not glue on a constant function on $U_1 \cup U_2$.

Then the correspondence $U \mapsto \Gamma(U;\mathbb{R}_X) = \{\text{constant functions on } X\}$ does not define a sheaf on X. We obtain a sheaf if we consider locally constant functions.

Example 2.1.5 Let us consider

$$\begin{array}{rcl} \mathcal{C}_X : \text{Open sets of } X & \to & \mathbb{R}\text{-vector spaces} \\ U & \mapsto & \{\text{continuous real valued functions on } U\} \end{array}$$

- If s is a continuous function and s is zero on a covering of U then $s = 0$.
- If $\{s_i\}$ are continuous functions on a covering $\{U_i\}$ of U, such that $s_i = s_j$ on $U_i \cap U_j$, then there exists s continuous on U with $s = s_i$ on each U_i.

Then the correspondence $U \mapsto \Gamma(U;\mathcal{C}_X) = \{\text{continuous real valued functions on } U\}$ defines a sheaf on X.

Example 2.1.6 Let us consider

$$\begin{array}{rcl} \mathcal{C}^b_X : \text{Open sets of } X & \to & \mathbb{R}\text{-vector spaces} \\ U & \mapsto & \{\text{continuous bounded functions on } U\} \end{array}$$

- For example, let $X = \mathbb{R}$, $U_n = (-n,n)$, $n \in \mathbb{N}$, and $s_n \colon U_n \to \mathbb{R}$, $x \mapsto x^2$. Then s_n is bounded on U_n for each $n \in \mathbb{N}$, but $x \mapsto x^2$ is not bounded on $\mathbb{R}$.

Then the correspondence $U \mapsto \Gamma(U;\mathcal{C}^b_X) = \{\text{continuous bounded real valued functions on } U\}$ does not define a sheaf on X.

Other examples of sheaves are: holomorphic functions, $\mathcal{C}^\infty$ functions , distributions.
Other examples of presheaves which are not sheaves are: L^2 functions, tempered distributions. In fact they do not satisfy gluing conditions.

If we consider "less open subsets" and "less coverings" they may become sheaves. We need the notion of site.

2.2 Sheaves on topological sites

Remark that the definition of sheaf depends only on

- open subsets
- coverings

One can generalize this notion by choosing a subfamily $\mathcal{T}$ of $\mathrm{Op}(X)$ and for each $U \in \mathrm{Op}(X)$ a subfamily $\mathrm{Cov}(U)$ of coverings of U satisfying the following hypothesis (defining a site $X_{\mathcal{T}}$).

(i) $U \in \mathrm{Cov}(U)$,

(ii) if $S_1 \in \mathrm{Cov}(U)$ is a refinement of $S_2 \subset \mathrm{Op}(U)$, then $S_2 \in \mathrm{Cov}(U)$,

(iii) if $S \in \mathrm{Cov}(U)$, then for each $V \subseteq U$, $V \cap S \in \mathrm{Cov}(V)$,

(iv) if $S_1, S_2 \subset \mathrm{Op}(U)$, $S_1 \in \mathrm{Cov}(U)$ and $V \cap S_2 \in \mathrm{Cov}(V)$, then $S_2 \in \mathrm{Cov}(U)$.

Here $S \cap V = \{U_j \cap V\}_{j\in J}$ if $S = \{U_j\}_{j\in J}$.

Hence we can generalize the definitions of presheaf and sheaf to $X_{\mathcal{T}}$.

Definition 2.2.1 *Following the notations of Definitions 2.1.1 and 2.1.3:*

(i) a presheaf on $X_{\mathcal{T}}$ is a contravariant functor from $\mathcal{T}$ to the category of k-vector spaces,

(ii) a presheaf of k-vector spaces F is a sheaf on $X_{\mathcal{T}}$ if for each $U \in \mathcal{T}$ and each $\{U_j\}_{j\in J} \in \mathrm{Cov}(U)$ we have the exact sequence

$$0 \to F(U) \to \prod_{j\in J} F(U_j) \to \prod_{j,k\in J} F(U_j \cap U_k).$$

The definitions of morphisms of presheaves (resp. sheaves) on $X_{\mathcal{T}}$ are similar to those of presheaves (resp. sheaves) on X.

Example 2.2.2 Let us consider the site $X_{\mathcal{T}}$ where

- $\mathcal{T}$=open subsets of X
- $\mathrm{Cov}(U) = \{$coverings of U admitting a finite subcovering$\}$

and consider the correspondence $U \mapsto \Gamma(U; \mathcal{C}^b_X)$ (continuous bounded functions).

- If $\{s_i\}$ are bounded on a finite covering $\{U_i\}$ of U, such that $s_i = s_j$ on $U_i \cap U_j$, then there exists s bounded on U with $s = s_i$ on each U_i.

Then the correspondence $U \mapsto \Gamma(U; \mathcal{C}^b_X)$ defines a sheaf on $X_{\mathcal{T}}$.

3 Sheaves on subanalytic sites

3.1 The general case

Let X be a topological space and consider a family of open subsets $\mathcal{T}$ satisfying:

$$\begin{cases} \text{(i) } U, V \in \mathcal{T} \Leftrightarrow U \cap V, U \cup V \in \mathcal{T}, \\ \text{(ii) } U \setminus V \text{ has finite numbers of connected components } \forall U, V \in \mathcal{T}, \\ \text{(iii) } \mathcal{T} \text{ is a basis for the topology of } X. \end{cases}$$

Definition 3.1.1 *The site $X_{\mathcal{T}}$ is defined by:*

- *open subsets: elements of $\mathcal{T}$*
- $\mathrm{Cov}(U)$ *(coverings of $U \in \mathrm{Op}(X_{\mathcal{T}})$): coverings admitting a finite subcovering*

Example 3.1.2 Exemples of families $\mathcal{T}$ satisfying $(i) - (iii)$ are:

(i) $\mathcal{T}$=$\{$open semialgebraic subsets of $\mathbb{R}^n\}$

(ii) $\mathcal{T} = \{$open relatively compact subanalytic subsets of a real analytic manifold$\}$, the subanalytic site X_{sa}.

(iii) $\mathcal{T}$=$\{$open definable subsets of $N^n\}$, given an O-minimal structure $(N, <, \ldots)$, the site DTOP.

There is an easy method to verify if a presheaf on $X_{\mathcal{T}}$ is a sheaf.

Proposition 3.1.3 *Let F be a presheaf on $X_{\mathcal{T}}$. Assume that*

(i) $F(\varnothing) = 0$

(ii) $\forall U, V \in \mathcal{T}$ *the sequence*

$$0 \to F(U \cup V) \to F(U) \oplus F(V) \to F(U \cap V)$$

is exact.

Then F is a sheaf on $X_{\mathcal{T}}$.

3.2 Subanalytic sheaves

From now on we will consider the subanalytic site X_{sa}:

- open subsets: relatively compact subanalytic open subsets,
- $\mathrm{Cov}(U)$ (coverings of $U \in \mathrm{Op}(X_{sa})$): locally finite coverings of U.

Let us consider as an example the presheaf

$$U \mapsto \mathcal{D}b^t_X(U)$$

of tempered distribution over a real analytic manifold X. This is not a sheaf with the usual topology.

- For example, if $X = \mathbb{R}$, we can find tempered distributions s_n on $\left\{\frac{1}{n} < x < 1\right\}$, $n \in \mathbb{N}$ which do not glue to a tempered distribution s on $\{0 < x < 1\}$.

Anyway for U, V open subanalytic relatively compact subsets of X we have the exact sequence

$$0 \to \mathcal{D}b^t_X(U \cup V) \to \mathcal{D}b^t_X(U) \oplus \mathcal{D}b^t_X(V) \to \mathcal{D}b^t_X(U \cap V)$$

This implies that $U \mapsto \mathcal{D}b^t_X(U)$ is a sheaf on the subanalytic site X_{sa}.

One can also define the six Grothendieck operations for subanalytic sheaves. In fact we have the following result of [6]

Theorem 3.2.1 *Let $f : X \to Y$ be a morphism of real analytic manifolds. The six Grothendieck operations $\mathcal{H}om$, $\otimes$, f_*, f^{-1}, $f_{!!}$, $f^!$ are well defined for subanalytic sheaves.*

4 $\mathcal{D}$-modules

4.1 The ring of differential operators

Let X be a complex analytic manifold. We denote by $\mathcal{D}_X$ the sheaf of rings of differential operators. Locally, a section of $\Gamma(U; \mathcal{D}_X)$ may be written as $P = \sum_{|\alpha| \le m} a_\alpha(z)\partial_z^\alpha$ with $a_\alpha(z)$ holomorphic on U. We denote by $\mathrm{Mod}(\mathcal{D}_X)$ the sheaf of $\mathcal{D}_X$-modules (i.e. a sheaf F belongs to $\mathrm{Mod}(\mathcal{D}_X)$ if $F(U)$ is a $\mathcal{D}_X(U)$-module for each $U \in \mathrm{Op}(X)$).

Let $T^*X \xrightarrow{\pi} X$ be the cotangent bundle and let $\mathcal{S} \subset \pi^{-1}\mathcal{D}_X$

$$\mathcal{S}_{(x,\xi)} = \{P \in \pi^{-1}(\mathcal{D}_x)\ ;\ \sigma^{-1}(P)(x,\xi) \neq 0\}$$

where $\sigma(P)$ is the principal symbol of P. Set $\mathcal{E}_X = \mathcal{S}^{-1}(\pi^{-1}\mathcal{D}_X)$. It means that in $\mathcal{E}_X$ every P with $\sigma(P)(x,\xi) \neq 0$ is (locally) invertible.

Definition 4.1.1 *The characteristic variety* $\mathrm{Char}(\mathcal{M})$ *of a $\mathcal{D}_X$-module $\mathcal{M}$ is the support of $\mathcal{E}_X \otimes_{\pi^{-1}\mathcal{D}_X} \pi^{-1}\mathcal{M}$.*

Example 4.1.2 If P is a differential operator and $\mathcal{M} = \mathcal{D}_X/\mathcal{D}_X P$ (i.e. $\mathcal{M} = \mathrm{coker}(\mathcal{D}_X \xrightarrow{P} \mathcal{D}_X)$), $\mathrm{Char}(\mathcal{M})$ is the zero locus of the principal symbol $\sigma(P)$ of P. This is because if $\sigma(P) \neq 0$ then P (*locally*) has an inverse in $\mathcal{E}_X$.

Let $f : X \to Y$ be a morphism of complex analytic manifolds and let $f_\pi : X \times_Y T^*Y \to T^*Y$ be the base change map.

Definition 4.1.3 *f is non characteristic for $\mathcal{M}$ if*

$$f_\pi^{-1}(\mathrm{Char}(\mathcal{M})) \cap T^*_X Y \subseteq X \times_Y T^*_Y Y$$

Example 4.1.4 If $Y = \mathbb{C}^n$, $X = \{z_1 = 0\}$ and $f : X \hookrightarrow Y$, $\mathcal{M} = \mathcal{D}_X/\mathcal{D}_X P$ with $P = \sum_{|\alpha| \leq m} a_\alpha(z)\partial_z^\alpha$ then f is non characteristic if the coefficient of $\partial_{z_1}^m$ is $\neq 0$ on X.

The sheaf $\mathcal{O}_X$ of holomorphic functions has a structure of $\mathcal{D}_X$-module. Let $\mathcal{M}$ be a coherent $\mathcal{D}_X$-module (i.e. there is an exact sequence $\mathcal{D}_X^M \to \mathcal{D}_X^N \to \mathcal{M} \to 0$). We denote by $\mathcal{S}ol(\mathcal{M})$ the sheaf $R\mathcal{H}om_{\mathcal{D}_X}(\mathcal{M}, \mathcal{O}_X)$.

Example 4.1.5 If U is convex and $\mathcal{M} = \mathcal{D}_X/\mathcal{D}_X P$, then $\mathcal{S}ol(\mathcal{M})$ on U is the complex

$$\Gamma(U; \mathcal{O}_X) \xrightarrow{P} \Gamma(U; \mathcal{O}_X),$$

$$\begin{aligned} H^0(U; \mathcal{S}ol(\mathcal{M})) &= \{s \in \Gamma(U; \mathcal{O}_X),\ Ps = 0\} = \ker P, \\ H^1(U; \mathcal{S}ol(\mathcal{M})) &= \Gamma(U; \mathcal{O}_X)/P\Gamma(U; \mathcal{O}_X) = \mathrm{coker} P. \end{aligned}$$

4.2 Cauchy-Kowaleskaya-Kashiwara Theorem

The following theorem is known as the Cauchy-Kowaleskaya-Kashiwara Theorem.

Theorem 4.2.1 *Let $\mathcal{M}$ be a coherent $\mathcal{D}_Y$-module and suppose that f is non characteristic for $\mathcal{M}$. Then $f^{-1}\mathcal{S}ol(\mathcal{M}) \simeq \mathcal{S}ol(\underline{f}^{-1}\mathcal{M})$.*

Suppose that U is a convex neighborhood of a point of $X = \{z_1 = 0\}$, f is the embedding and $\mathcal{M} = \mathcal{D}_X/\mathcal{D}_X P$, with $P = \sum_{|\alpha| \leq m} a_\alpha \partial_z^\alpha$. In this case $\underline{f}^{-1}\mathcal{M} \simeq \mathcal{D}_X^m$. We are reduced to the isomorphism

$$\begin{aligned} \Gamma(U; f^{-1}\mathcal{S}ol(\mathcal{M})) &\xrightarrow{\sim} (\Gamma(U; \mathcal{O}_X))^m \\ s &\mapsto (s|_X, \partial_{z_1} s|_X, \ldots, \partial_{z_1}^{m-1} s|_X) \end{aligned}$$

i.e. to the existence and uniqueness of the solution of

$$\begin{cases} Ps = 0 \\ \partial^k s|_X = g_k \quad k = 0, \ldots, m-1 \end{cases}$$

for any $(g_k)_{k=0}^{m-1} \in (\Gamma(U;\mathcal{O}_X))^m$. Moreover $H^1(U;\mathcal{S}ol(\mathcal{M})) = 0$ means that P is surjective, i.e. the existence of the solution of $Ps = g$ for any $g \in \Gamma(U;\mathcal{O}_X)$.

Definition 4.2.2 *One denotes by $\mathcal{O}_X^t$ the sheaf of tempered holomorphic functions defined by the Dolbeault complex:*

$$0 \to \mathcal{D}b_X^t \xrightarrow{\overline{\partial}} \mathcal{D}b_X^{t\,(0,1)} \xrightarrow{\overline{\partial}} \cdots \xrightarrow{\overline{\partial}} \mathcal{D}b_X^{t\,(0,n)} \to 0.$$

The sheaf of tempered holomorphic functions has a structure of $\rho_!\mathcal{D}_X$-module. ($\Gamma(U;\rho_!\mathcal{D}_X)$ are differential operators $\sum_{|\alpha|\leq m} a_\alpha \partial_z^\alpha$ with a_α holomorphic in $\overline{U}$).

Remark 4.2.3 One shall be aware that if the dimension of X is > 1 the Dolbeaut complex is not concentrated in degree zero and $\mathcal{O}_X^t$ belongs to derived category of $\mathrm{Mod}(\mathcal{D}_X)$.

Denote $\mathcal{S}ol^t(\mathcal{M}) = R\mathcal{H}om_{\rho_!\mathcal{D}_X}(\rho_!\mathcal{M}, \mathcal{O}_X^t)$. Thanks to the theory of subanalytic sheaves one can obtain a tempered version of the Cauchy-Kowaleskaya-Kashiwara Theorem.

Theorem 4.2.4 *Let $\mathcal{M}$ be a coherent $\mathcal{D}_Y$-module and suppose that f is non characteristic for $\mathcal{M}$. Then $f^{-1}\mathcal{S}ol^t(\mathcal{M}) \simeq \mathcal{S}ol^t(\underline{f}^{-1}\mathcal{M})$.*

References

[1] R. Godement, "Topologie algébrique et théorie des faisceaux". Hermann, Paris, 1958.

[2] B. Iversen, "Cohomology of sheaves". Universitext Springer-Verlag, Berlin, 1986.

[3] M. Kashiwara, "$\mathcal{D}$-modules and microlocal calculus". Translations of Math. Monog. 217, Iwanami Series in Modern Math., American Math. Soc., Providence, 2003.

[4] M. Kashiwara, P. Schapira, "Sheaves on manifolds". Grundlehren der Math. 292, Springer-Verlag, Berlin, 1990.

[5] M. Kashiwara, P. Schapira, "Ind-sheaves". Astérisque 271, 2001.

[6] L. Prelli, *Sheaves on subanalytic sites.* To appear, Rend. Sem. Mat. Univ. Padova, available in arXiv: `math.AG/0505498`.

[7] J.-P. Schneiders, *An introduction to $\mathcal{D}$-modules.* Bull. Soc. Royale des Sciences de Liège 63, pp. 223-295 (1994).

[8] G. Tamme, "Introduction to étale cohomology". Universitext Springer-Verlag, Berlin, 1994.

Numerical modeling for convection–dominated problems

Manolo Venturin (*)

Abstract. During the last years, there has been a great interest in the development of sophisticated mathematical models for the simulation of real life applications which involves convection–dominated phenomena. For example, these problems concern the solution of scalar advection–diffusion equations, the Navier–Stokes equations and the Shallow Water equations.
The main goal of this seminar is to review the most important difficulties that arise in the numerical approximation of this kind of problems when convection dominates the transport process.
Moreover, we present a method for the treatment of this equations with the use of the finite element discretization on the domain.

Sunto. Nel corso degli ultimi anni, vi è stato un grande interesse nello sviluppo di modelli matematici sofisticati per la simulazione di problemi reali che coinvolgono fenomeni a convezione dominante. Ad esempio, riguardano la soluzione di equazioni di convezione–diffusione scalari, le equazioni di Navier–Stokes e le equazioni delle Acque Basse.
L'obiettivo principale di questo seminario è la revisione delle più importanti difficoltà che insorgono in ambito numerico per questo tipo di problemi quando la convezione domina il processo di trasporto.
Inoltre, viene presentato il trattamento di queste equazioni mediante l'uso del metodo degli elementi finiti.

A number of important phenomena encountered in coastal and environmental engineering, are described by the nonlinear shallow water equations. Their interests is largely motivated by environmental considerations, such as the study of tidal currents and water elevations for flood control, and the need for prediction of man–made alterations to the environment by the construction of harbours, barrages, etc.

In the above applications, it is increasingly important to be able to deal with complex geometries with several physical parameters and different boundary conditions. It is also desired to preserve accuracy of solutions in the computational steps and to save simulation time. Hence, the development of sound and flexible numerical tools are important in the investigations of such phenomena in order to have a "true" and "close" description of the reality. Moreover, transport models, governed by advection–diffusion equations, can also

(*)Grant holder in Applied Mathematics - Università di Padova, Dipartimento di Matematica Pura ed Applicata, Via Trieste 63, 35121 Padova, Italy. E-mail: mventuri@math.unipd.it. Seminar held on 2 April 2008.

be coupled to the shallow water hydrodynamic model making it possible to study pollutant dispersion or temperature distribution.

During the last years, considerable effort has been focused towards the development of two–dimensional models for the numerical approximation of the shallow water equations both in conservative and non–conservative forms, and many numerical schemes are now available for that purpose. In the past, the most popular methods used for this discretization are based on finite differences and finite volume methods. Recently, alternative approaches such as spectral element methods, and finite element schemes have been proposed. For details see [1, 2].

In this work, advection–diffusion and shallow water problems are considered. In particular, for the shallow water problems, the following gorverning equations (hydrodynamical model) are taken into account:

Continuity equation

$$\frac{\partial h}{\partial t} = \frac{1}{c^2}\frac{\partial p}{\partial t} = -\frac{\partial U_i}{\partial x_i} \quad \text{in } \Omega \times (0,T)$$

Momentum equations

$$\frac{\partial U_i}{\partial t} = -\frac{\partial}{\partial x_j}\left(U_i u_j\right) - \frac{\partial p}{\partial x_i} - Q_i \quad \text{in } \Omega \times (0,T)$$

where h or p and u_i are the unknowns; see Chapter 2 for a detailed description of the equations.

In the case of pollutant dispersion of a passive tracer (scalar variable) T ("passive" meaning the distribution of the tracer does not affect the fluid flow), the following advection–diffusion equations should be solve

Transport equation

$$\frac{\partial (hT)}{\partial t} + \frac{\partial}{\partial x_i}\left(hu_i T\right) - \frac{\partial}{\partial x_i}\left(hk\frac{\partial T}{\partial x_i}\right) + R = 0$$

in which T is the depth–averaged pollutant dispersion, k is the depth–averaged diffusion coefficient, h and u_i are the depth and fluid velocity previously computed, and R is a depth–averaged source ($R < 0$) or sink ($R > 0$) term.

In the case of the advection–diffusion equations, the main difficulty is due to the presence of non–symmetric convection operators that appear in formulations based on kinematic description other than Lagrangian.

When convection dominates the transport processes, the best approximation property in the energy norm of the Galerkin method — which is the basis of success in symmetric cases — is lost, and solutions are corrupted by spurious node to node oscillations. These can only be removed by mesh and time step refinements which destroy the practical utility of the method, and, this has motivated, the development of the Galerkin formulations, called stabilization techniques, which precludes oscillations without requiring mesh or time step refinements.

The stabilization methods that are the most known are the Streamline Upwinding Petrov–Galerkin (SUPG) method, the Galerkin Least Square (GLS) method, Residual–Free Bubble, Wavelet functions, the Taylor–Galerkin (TG) method and the Finite Incremental Calculus (FIC) method.

The numerical difficulties, encountered in shallow water problem, in the use of the standard Galerkin finite element method are mainly of three different kinds:

- the mixed type of the equations, which is due to the coupling of the momentum equation with the incompressibility condition, and subsequently, the treatment of the pressure or water elevation term;
- the advective–diffusive character of the equations, which have a viscous and a convective term;
- and finally, the nonlinearity of the problem.

The first is related to the incompressibility of the fluid and exhibits itself when an incorrect combination of element interpolation functions for the velocity and pressure or water elevation is employed. It consists of a constraint on the velocity field which must be divergence free. Then, the pressure has to be considered as a variable not related to any constitutive equation. Its presence in the momentum equation has the purpose of introducing an additional degree of freedom needed to satisfy the incompressibility constraint. The role of the pressure variable is thus to adjust itself instantaneously in order to satisfy the condition of divergence–free velocity. That is, the pressure is acting as a Lagrangian multiplier of the incompressibility constraint and thus there is a coupling between the velocity and the pressure unknowns.

Another source of numerical difficulty is due to the presence of nonlinear and non–symmetric convective terms in the momentum equations. As it is well known, the standard Galerkin formulation typically lacks stability when convective effects dominate and alternative spatial discretization procedure must be used to restore stability without compromising the accuracy.

As a starting point for the development of the finite element models for the shallow water equations, we consider the CBS algorithm proposed by *Zienkiewicz* and co – workers (a detailed description is available in [3]).

From its introduction the Characteristic Based Split (CBS) method has been used by a certain number of researchers to solve fluid dynamics problems. Such method does enjoy interesting stability and consistency properties and are nowadays widely used by the finite element community for solving convection–dominated problems. Its basis is the fractional step procedure introduced by Chorin and Temam for incompressible Navier–Stokes equations in the finite difference context.

The main idea of the Chorin–Temam method consists in the decomposition of the time advancement into a sequence of two or more steps that split the numerical treatment of the equation operators into relatively easier subproblems. The principle of the projection method is to compute the velocity and pressure fields separately through the computation of an intermediate velocity, which is then projected onto the subspace of solenoidal

vector functions. Basic to the derivation of projection methods is a theorem of orthogonal decomposition due to *Ladyzhenskaya*, which is based on the Helmholtz decomposition principle.

Several implementations have been proposed to perform such splitting and therefore a variety of fractional–step methods exists: fractional steps or splitting methods for evolution equations, methods based on a projection onto a subspace of the solenoidal vector functions; algebraic splitting methods and methods based on pressure or velocity correction. A detailed exposition of fractional–step methods can be found in [4] and references therein.

The CBS scheme combines the Characteristic–Galerkin method to deal with convection dominated flows with a splitting technique based on velocity correction. The velocity field is computed into two stages with the Characteristic–Galerkin method. In the first step, the pressure term (or elevations of the free surface) is retained from the momentum equations and an intermediate velocity field is estimated. Then, the continuity equation is solved using the intermediate vector field value and the pressure is carried out, by means of a Laplacian–type equation, whose self–adjoint form makes the Galerkin space discretization optimal. Finally, the velocity field is corrected using the new computed pressure term.

This leads to the following time–discretization formulae:

Intermediate momentum

$$\Delta U_i^* = \Delta t \left[-\frac{\partial}{\partial x_j} \left(U_i u_j \right)^n - Q_i^n \right] + \frac{\Delta t^2}{2} u_k^n \frac{\partial}{\partial x_k} \left(\frac{\partial}{\partial x_j} \left(U_i u_j \right)^n + Q_i^n \right)$$

Pressure equation

$$\left(\frac{1}{c^2} \right)^n \Delta p - \Delta t^2 \theta_1 \theta_2 \frac{\partial}{\partial x_i} \left(\frac{\partial (\Delta p)}{\partial x_i} \right) = -\Delta t \frac{\partial}{\partial x_i} \left(U_i^n + \theta_1 \Delta U_i^* \right) + \Delta t^2 \theta_1 \frac{\partial}{\partial x_i} \left(\frac{\partial p^n}{\partial x_i} \right)$$

Momentum correction

$$\Delta U_i = \Delta U_i^* - \Delta t \frac{\partial p^{n+\theta_2}}{\partial x_i} + \frac{\Delta t^2}{2} u_k^n \frac{\partial}{\partial x_k} \left(\frac{\partial p^n}{\partial x_i} \right)$$

where higher order terms have been neglected.

The approximation of the scalar transport equation, that can be added to the hydrodynamic model, is straightforward, and it requires only the application of the Characteristic–Galerkin method. Hence, the characteristic time discretization gives

Transport equation

$$\Delta T = \Delta t \left[-\frac{\partial}{\partial x_j} (u_j T)^n + \frac{\partial}{\partial x_j} \left(k \frac{\partial T}{\partial x_j} \right)^{n+\theta_3} - R^n \right] + \frac{\Delta t^2}{2} u_k^n \frac{\partial}{\partial x_k} \left(\frac{\partial}{\partial x_j} (u_j T)^n + R^n \right)$$

where T is multiplied by h, $\Delta T = T^{n+1} - T^n$.

The procedure has some interesting and useful advantages. The first is that dropping the pressure term, each momentum equation is similar to an advection–diffusion equation and so the Characteristic–Galerkin procedure can be applied. The idea of the

Characteristic–Galerkin scheme is to stabilize advection–diffusion equations using a finite difference discretization of the total derivative along the characteristic. Then, if the discretization of the space is done, a consistent artificial diffusion, which stabilized convective terms, appears. The splitting operation, being self–adjoint, can then be solved optimally using the Galerkin procedure. The second advantage is that removing the pressure from the momentum equations enhances the pressure stability and permits to avoid any restrictions on the nature of the interpolation functions for both velocity and pressure, *i.e.* the Babuška–Brezzi condition is satisfied. Finally, in the semi–implicit form the algorithm provides a critical time step dependent only on the current velocity instead of the wave celerity.

References

[1] M. Morandi Cecchi and M. Venturin, *Characteristic-based split (CBS) algorithm finite element modelling for shallow waters in the Venice lagoon.* Internat. J. Numer. Methods Engrg. 66 (2006), no. 10, 1641–1657.

[2] M. Venturin, "A finite element stabilization system for advection–diffusion problems". Tesi di Dottorato di Ricerca in Matematica Computazionale, Università degli Studi di Padova, dicembre 2005.

[3] O. C. Zienkiewicz and R. L. Taylor, "The finite element method. Vol. 3: Fluid dynamics". Butterworth–Heinemann, Oxford, fifth edition, 2000.

[4] L. Quartapelle, "Numerical solution of the incompressible Navier–Stokes equations". Volume 113 of *International Series of Numerical Mathematics*, Birkhäuser Verlag, Basel, 1993.

The Basic Picture on sets evaluated over an overlap algebra

Paola Toto (*)

Abstract. In his forthcoming book [4], G. Sambin introduces a new topological theory, called "The Basic Picture". In this theory both the notion of topological space and its point-free version are generalized. The concept of overlap algebra is also introduced in order to put in algebraic form the properties needed to define the new topological structures. The ultimate goal of our work is to generalize such topological notions in the context of many-valued sets. In many-valued set theory sets are built by using propositions evaluated in an algebraic structure. To reach our goal a key point is to check whether the original algebrization of Sambin's topological notions can be considered also as the algebrization of their many-valued version. We prove that this is the case if and only if we take an overlap algebra as the underlying structure of truth values. This is a joint work with Maria Emilia Maietti and Giovanni Sambin.

Sunto. Nel suo libro [4] di prossima uscita, G. Sambin introduce una nuova teoria, detta "The Basic Picture" in cui generalizza sia la nozione di spazio topologico sia la sua versione senza punti. Inoltre, egli introduce il concetto di overlap algebra, al fine di tradurre in forma algebrica le proprietà necessarie a definire queste nuove strutture topologiche. L'obiettivo finale del nostro lavoro è di generalizzare queste nozioni topologiche nel contesto degli insiemi a più valori. Nella teoria degli insiemi valutati, gli insiemi sono costruiti usando proposizioni valutate su una struttura algebrica fissata. Per raggiungere l'obiettivo prefissato, un passo fondamentale è stato quello di controllare quanto l'originaria algebrizzazione della nozione topologica di G. Sambin possa anche essere come l'algebrizzazione della sua versione a più valori. In questo lavoro si è provato che questo accade se e solo se si considera come sottostante struttura dei valori di verità un'overlap algebra. Quanto viene di seguito presentato è estratto da un lavoro fatto in collaborazione con Maria Emilia Maietti e Giovanni Sambin.

1 Preliminaries

All this work is developed in a predicative constructive set theory, using intuitionistic logic. As starting point we will give a tutorial introduction of our work, explaining the meaning of the previous words. Intuitionistic Logic can be succinctly described as Classical Logic

(*)Dottorato in Matematica Pura, XX ciclo. Università del Salento. Advisors: Giovanni Sambin (Università di Padova - Dipartimento di Matematica pura ed applicata), Cosimo Guido (Università del Salento - Dipartimento di Matematica). E-mail: paola.toto@unile.it. Seminar held on 16 April 2008.

without the Aristotelian law of excluded middle:

$$\text{(LEM): } (A \vee \neg A)$$

The rejection of (LEM) has far-reaching consequences. On the one hand, intuitionistically, *Reductio ad absurdum* can work only on negative statements, since $\neg\neg A \to A$ does not hold in general. Formalized intuitionistic logic is naturally motivated by the informal Brouwer-Heyting-Kolmogorov (BHK, for short) explication of intuitionistic truth. The constructive independence of the logical operations $\&, \vee, \to, \neg, \forall, \exists$ **contrasts with the classical situation**, where e.g. $(A \vee B)$ is equivalent to $\neg(\neg A \& \neg B)$ and $\exists x A(x)$ is equivalent to $\neg\forall x \neg A(x)$. In fact, in intuitionistic logic the following directions hold:

- $(A \vee B)$ implies $\neg(\neg A \& \neg B)$;
- $(\neg A \vee B)$ implies $(A \to B)$;
- $\exists x A(x)$ implies $\neg\forall x \neg A(x)$;
- A implies $\neg\neg A$;
- $\neg A$ is equivalent to $\neg\neg\neg A$;
- $\neg(A \vee B)$ is equivalent to $(\neg A \& \neg B)$;
- $\neg\exists x A(x)$ is equivalent to $\forall x \neg A(x)$.

From the BHK viewpoint, a sentence of the form $(A \vee B)$ asserts that either a proof of A, or a proof of B, has been constructed; while $\neg(\neg A \& \neg B)$ asserts that an algorithm has been constructed which would effectively convert any pair of constructions proving $\neg A$ and $\neg B$ respectively, into a proof of a known contradiction. Constructive mathematics is distinguished from its traditional counterpart, classical mathematics, by the strict interpretation of the phrase "there exists" as "we can construct" and the statement "$(A \vee B)$" is considered to be established only when one either can decide - that is, prove - that A or decide (prove) that B, namely if a constructive process can be presented that terminates with the indication that one of its components is true. In the intuitionistic propositional logic, this yields two properties:

(DP) If $(A \vee B)$ is a theorem, then A is a theorem or B is a theorem [Disjunction Property].

(EP) If $\exists x A(x)$ is a closed theorem, then for some closed term t, $A(t)$ is a theorem [Existence Property].

The disjunction and existence properties are special cases of a general phenomenon peculiar to nonclassical theories. In order to work constructively, we need to reinterpret not only the existential quantifier but all the logical connectives and quantifiers as instructions on how to construct a proof of the statement involving these logical expressions. Summarizing, in constructive mathematics, something is true if one can exhibit a proof of it, while in classical mathematics something is true if it is not contradictory. For example, when we pass from our initial, natural interpretation of $A \vee B$ to the unrestricted use of the

idealistic one, $\neg(\neg A \& \neg B)$, the resulting mathematics cannot enjoy the existence and the disjunction properties. This point is illustrated by a well-known example,

Proposition 1.1 *There exist irrational numbers a, b such that a^b is rational.*

A slick classical proof goes as follows. Either $\sqrt{2}^{\sqrt{2}}$ is rational, in which case we take $a = b = \sqrt{2}$, or else $\sqrt{2}^{\sqrt{2}}$ is irrational, in which case we take $a = \sqrt{2}^{\sqrt{2}}$ and $b = \sqrt{2}$. But as it stands, this proof does not enable us to pinpoint which of the two choices of the pair (a, b) has the required property. In order to determine the correct choice of (a, b), we would need to decide whether $\sqrt{2}^{\sqrt{2}}$ is rational or irrational, which is precisely to employ our initial interpretation of disjunction with P the statement "$\sqrt{2}^{\sqrt{2}}$ is rational".

A slick constructive proof goes as follows. Let us consider $a = \sqrt{2}$ and $b = log_2 9$. It is well known that $\sqrt{2} \notin \mathbb{Q}$. And $log_2 9 \notin \mathbb{Q}$, too. In fact, let us assume that $log_2 9 \in \mathbb{Q}$. Then $log_2 9 = \frac{m}{n}$, for some $m \in \mathbb{Z}$ and $n \in \mathbb{Z} \setminus 0$. Solving this equation, we get $9^n = 2^m$ and this is false, since an odd number is never equal an even one. Thus, $a^b = \sqrt{2}^{log_2 9} = 3 \in \mathbb{Q}$.

It should, by now, be clear that a full-blooded computational development of mathematics disallows the idealistic interpretations of disjunction and existence upon which most classical mathematics depends. In fact, in order to work constructively, we need to return from the classical interpretations back to the natural, constructive ones, as follows.

Connectives (Name)	**Formulas**	**Interpretation**
$\vee$ (or)	$P \vee Q$	to prove $P \vee Q$ we must have either a proof of P or a proof of Q
& (and)	$P \& Q$	to prove $P \& Q$ we must have a proof of P and a proof of Q
$\rightarrow$ (implies)	$P \rightarrow Q$	a proof of $P \rightarrow Q$ is an algorithm that converts a proof of P into a proof of Q
$\neg$ (not)	$\neg P$	to prove $\neg P$ we must show that P implies false (e.g. $0 = 1$)
$\exists$ (there exists)	$\exists x P(x)$	to prove $\exists x P(x)$ we must construct an object x and prove that $P(x)$ holds
$\forall$ (for each/all)	$\forall x P(x)$	a proof of $\forall x P(x)$ is an algorithm that, applied to any object x, proves that $P(x)$ holds

Why would we want to reinterpret logic in this way? First, there is the desire to retain, as far as possible, computational interpretations of our mathematics. Ideally, we are trying to develop mathematics in such a way that if a theorem asserts the existence of an object x with a property P, then the proof of the theorem embodies algorithms for constructing x and for demonstrating, by whatever calculations are necessary, that x has the property P.

From the algebraic point of view, Heyting algebras are models of intuitionistic logic, while Boolean algebras are models of classical logic. Heyting algebras are special partially ordered sets that constitute a generalization of Boolean algebras. A Heyting algebra H is a bounded lattice such that for all $a, b \in H$ there is a greatest element x of H such that $a \wedge x \leq b$. This element is the relative pseudo-complement of a with respect to b,

and is denoted $a \to b$. We write 1 and 0 for the greatest and the smallest element of H, respectively. In any Heyting algebra, one defines the pseudo-complement $\neg x$ of any element x by setting $\neg x = (x \to 0)$. By definition, $a \wedge \neg a = 0$, and $\neg a$ is the greatest element having this property. However, it is not in general true that $a \vee \neg a = 1$, thus $\neg$ is only a pseudo-complement, not a true complement, as would be the case in a Boolean algebra. A complete Heyting algebra is a Heyting algebra that is a complete lattice. Thus a complete Heyting algebra H is a complete lattice $(H, \leq, \wedge, \vee, \bigvee, \bigwedge, \to, 0, 1)$ such that the following properties hold: $a \to a = 1$; $a \wedge (a \to b) = a \wedge b$; $b \wedge (a \to b) = b$; $a \to (b \wedge c) = (a \to b) \wedge (a \to c)$, where $a \wedge b \leq c \Leftrightarrow b \leq a \to c$, for any $a, b, c \in H$.

Every topology $\mathcal{O}X$, of a topological space $(X, \mathcal{O}X)$, provides a complete Heyting algebra in the form of its open set lattice. In this case, the element $A \to B$ is the supremum of the collection $\{C \in \mathcal{O}X | A \cap C \subseteq B\}$. Not all complete Heyting algebras are of this form.

2 Sambin's Basic Picture

In topology an application of the constructive independence of the logical operation $\&, \vee, \to, \neg, \exists, \forall$ is for example related to closed subsets of a topological space. In fact, in constructive mathematics, the definition of a closed subset as a subset containing all its limit points is not equivalent to say that a closed subset is the complement to an open subset (as it happens in classical mathematics), since in intuitionistic logic, the existential quantifier $\exists$ is not equivalent to $\neg\forall\neg$. Thus, since from a constructive point of view, closed subsets are not necessarily defined as complements of these open ones, this leads to develop topology by considering open subsets and closed subsets primitively.

Moreover it often happens that the structure of open subsets can be given quite constructively, even when the corresponding points have an infinitary description. Typical is the case of real numbers: points are infinite sequences, while open subsets can be given starting from intervals with rational endpoints. Then the idea is that one can begin without real numbers: one defines the topology by giving open subsets and their coverings primitively, rather then quantifying on points.

Therefore Sambin's Basic picture is a generalization of both Topological Space and Point-free Topology. Topology begins when passing from one set to two sets linked by an arbitrary relation, a minimal structure here called a basic pair. In fact, one can see that the topological notions of interior and closure are the result of the dynamics between the two sets. This discovery allows one to see a clear structure underlying topology: logical duality between open and closed, symmetry between the traditional (pointwise) and the pointfree approach. The notion of continuity also has a structural characterization, since it turns out to be just a commutative square. The theory which follows from or which extends such structures, symmetries and dualities, Sambin has called the basic picture. It forms a structural basis for constructive topology and in the same time it generalizes both pointfree and pointwise topology.

Typical of the Basic Picture is a systematic use of the notion of "overlap" between two subsets (existence of a common element, which is intuitionistically different from nonempty intersection), which is logically dual of that of inclusion. The algebraic structure tradition-

ally associated with a topological space is that of its open subsets, which form a complete lattice satisfying infinite distributivity (of arbitrary joins over binary meets), also called frame or locale. Due to the topological interpretation of intuitionistic logic (propositions as open subsets), the structure of locale is also the intuitionistic algebraic counterpart of the structure of the power of a set; in this context, it is often called a complete Heyting algebra, to stress the presence of implication (which is anyway impredicatively definable in any locale). To be able to reflect into an algebraic definition also the presence of a primitive notion of closed subsets, Sambin introduced a new algebraic structure, the overlap algebra.

Definition 2.1 *An overlap algebra is:*

- *a collection $\mathcal{P}$, with objects $p, q, \cdots \in \mathcal{P}$;*
- *an order $\leq$ on $\mathcal{P}$ such that $(\mathcal{P}, \leq, 0, 1)$ is a complete Heyting algebra;*
- *a symmetric binary relation $\between$ on $\mathcal{P}$ such that:*
  - *$\between$ preserves infimum (if $p \between q$ then $p \between p \wedge q$);*
  - *$\between$ splits supremum ($p \between \bigvee_{i \in I} q_i$ iff $p \between q_i$ for some $i \in I$);*
  - *$\between$ satisfies density (for any $r : \mathcal{P}$, if $p \between r$ implies $q \between r$, then $p \leq q$).*

A main example of overlap algebra is the power collection $\mathcal{P}(X)$ of a set X, with the extra primitive $\between$ dual of that one of inclusion $\subseteq$. In fact comparing the definitions, if X is a set and $\mathcal{P}$ an overlap algebra, we get:

$\mathcal{P}$	$\mathcal{P}(X)$
$p \leq q$	$U \subseteq V \equiv \forall x(x \,\epsilon\, U \rightarrow x \,\epsilon\, V)$
$p = q$	$U = V \equiv \forall x(x \,\epsilon\, U \leftrightarrow x \,\epsilon\, V)$
$p \between q$	$U \between V \equiv \exists x(x \,\epsilon\, U \;\&\; x \,\epsilon\, V)$

Such structures can be organized into a category.

Definition 2.2 *An overlap relation from an overlap algebra $\mathcal{Q}_1$ into an overlap algebra $\mathcal{Q}_2$ is a quadruple of functions $F =< f, f^*, f^-, f^{-*} >$, where $f, f^{-*} : \mathcal{Q}_1 \rightarrow \mathcal{Q}_2$ and $f^-, f^* : \mathcal{Q}_2 \rightarrow \mathcal{Q}_1$, such that:*

- *$f \dashv f^*$, that is $fp \leq q$ iff $p \leq f^* q$;*
- *$f^- \dashv f^{-*}$, that is $f^- q \leq p$ iff $q \leq f^{-*} p$;*
- *$f \cdot|\cdot f^-$, that is $fp \between q$ iff $p \between f^- q$.*

In this way the definition of the category **OA** of overlap algebras as objects and overlap relations as arrows.

Now we show as a relation between two sets produce a morphism of overlap algebras.

Example 2.3 If we consider two sets X, Y, then every relation r between them, denoted by $r : X \rightarrow Y$, yields an overlap relation $\mathcal{P}(X) \rightarrow \mathcal{P}(Y)$, where, for any $D \subseteq X$ and $U \subseteq Y$:

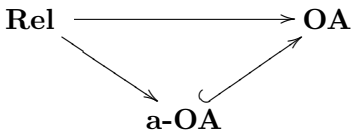
Existential operators	Universal operators
$rD \equiv \{a : \exists x \in X(xra \& x \,\epsilon\, D)\}$	$r^{-*}D \equiv \{a : \forall x \in X(xra \to x \,\epsilon\, D)\}$
$r^{-}U \equiv \{x : \exists a \in Y(xra \& a \,\epsilon\, U)\}$	$r^{*}U \equiv \{x : \forall a \in Y(xra \to a \,\epsilon\, U)\}$

To express in the language of overlap algebras the properties of singleton subsets in $\mathcal{P}(X)$, as the minimal inhabitated elements of the partially order collection $\mathcal{P}(X)$, we give the following

Definition 2.4 *In an overlap algebra $\mathcal{P}$, $p \in \mathcal{P}$ is atom if and only if*
$p \bowtie p$;
for every $q : \mathcal{P}$, if $q \bowtie p$ then $q \leq p$.
Intuitively, an overlap algebra $\mathcal{P}$ is atomic if the atoms are sufficient to determine all elements and their order, in the sense that for every $p, q \in \mathcal{P}$:

$$p \leq q \text{ if and only if } m \bowtie p \text{ implies } m \bowtie q \text{ for all atoms } m.$$

Denoting with **Rel** the category of relations between sets, it is possible to define the Power Functor $\mathcal{P} : \mathbf{Rel} \to \mathbf{OA}$ which is faithful, full and dense on atomic overlap algebras. This result can be expressed in the commutativity of the following diagram:



Our aim is to give a many-valued version of the previous diagram. In particular, to define the many-valued version of the Power Functor $\mathcal{P} : \mathbf{Rel} \to \mathbf{OA}$, we have to define in the many-valued structure of the power collection of a set X, $\mathcal{P}(X)$, in a suitable way.

3 Many-valued set theory

Many-valued set theory have been introduced by A. Zadeh as an extension of the classical notion of set theory. In classical set theory, the membership of elements in a set is assessed in binary terms according to a bivalent condition, an element either belongs or does not belong to the set. By contrast, many-valued (or fuzzy) set theory permits the gradual assessment of the membership of elements in a set; this is described with the aid of a membership function valued in a lattice. Fuzzy sets generalize classical sets, since the characteristic functions of classical sets are special cases of the membership functions of fuzzy sets, if the latter only take values 0 or 1. In our case, we have:

Set Theory	X	$a \in X$	$a = b \in X$	$C \subseteq X$
$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
Set Theory(H)	(X, E_X)	$E_X(a,a) \in H$	$E_X(a,b) \in H$	$C : X \to H$

And the idea is the following: since in the constructive set theory, a subset of a set X is a proposition $\phi(x)[x \in X]$ then in a many-valued setting it will be replaced by $\phi^*(x) \in H[x \in X]$, where $(H, \leq, \wedge, \vee, \rightarrow, 0, 1)$ is a complete Heyting algebra.

Definition 3.1 *An H-**valued model** or an H-**valuation** is a mapping from prop, the set of propositions of the Intuitionistic Set Theory, into a complete Heyting algebra H.*

The following table summarize the process of a H-**valuation**:

formula	H-**valuation**
$\bot$	0
$\top$	1
φ	φ^*
$\varphi \vee \psi$	$\varphi^* \vee \psi^*$
$\varphi \& \psi$	$\varphi^* \wedge \psi^*$
$\varphi \rightarrow \psi$	$\varphi^* \rightarrow \psi^*$
$\neg\varphi$	$(\neg\varphi)^* = \varphi^* \rightarrow 0$
$\forall x \varphi(x)$	$\bigwedge_{x \in X} \varphi^*(x)$
$\exists x \varphi(x)$	$\bigvee_{x \in X} \varphi^*(x)$

The H-valuation of the judgment "Γ is true implies that $\varphi(x)$ is true", is $\Gamma^* \leq \varphi^*(x)$. As corollary, the H-valuation of the judgment "$\varphi(x)$ is true" is $\varphi^*(x) = 1$.

3.1 H-sets and H-relations

Now, we are able to give some definitions.

Definition 3.2 *A H-**set** $\mathcal{X} \equiv (X, E)$ is a couple, where X is a set and $E_X : X \times X \rightarrow H$ is a map such that:*

(E1) $E_X(x, y) = E_X(y, x)$ *(symmetry)*

(E2) $E_X(x, y) \wedge E_X(y, z) \leq E_X(x, z)$ *(transitivity)*

The mapping E_X is the valuation in H of the equality of elements in X and it is called H-equality on X and $E_X(x, y)$ is interpreted as the *value* how much x and y *coincide*. In particular, $E_X(x, x)$ describes the domain or the extent of *existence* of x. It is easily seen that the From $(E2)$ and $(E1)$ it follows the *strictness axiom*:

(STR) $E_X(x, y) \leq E_X(x, x) \wedge E_X(y, y)$ (strictness).

We give some typical examples of H-sets.

Example 3.3

(a) Let X be a set. Then the *crisp equality* E_c on X determined by $E_c(x, y) = \vee\{1 | x = y\}$, for every $x, y \in X$, makes (X, E_c) a H-set.

(b) If we consider the binary meet operation $\wedge$ on the underlying Heyting algebra $(H, \leq, \wedge, \vee, \rightarrow, 0, 1)$, then $(H, \wedge)$ is a H-set.

(c) If we consider the bi-implication in $(H, \leq, \wedge, \vee, \rightarrow, 0, 1)$, that is $p \leftrightarrow q = (p \rightarrow q) \wedge (q \rightarrow p)$, for every $p, q \in H$, then $(H, \leftrightarrow)$ is a H-set.

(d) For every topological space $(X, \mathcal{O}X)$, let $H \equiv \mathcal{O}X$ be the complete Heyting algebra and let us consider the collection $\{f : X \rightarrow \mathbb{R} | f \text{ is a continuous function}\}$. Defining, for any continuous function $f, g : X \rightarrow \mathbb{R}$, $E_X(f, g) = int\{x \in X | f(x) = g(x)\}$, then (X, E_X) is a H-set.

Definition 3.4 *A* ***H-relation*** *r between two H-sets (X, E_X) and (Y, E_Y) is a map $r : X \times Y \rightarrow H$ such that, for all $x, y \in X$, $a, b \in Y$:*
$[Extensionality]$ $\quad r(x, a) \wedge E_X(x, y) \wedge E_Y(a, b) \leq r(y, b)$
$[Strictness]$ $\quad r(x, a) \leq E_X(x, x) \wedge E_Y(a, a)$

Definition 3.5 ***Rel(H)*** *is the category of H-sets and H-relations.*

Definition 3.6 *A* ***H-subset*** *D of a H-set $\mathcal{X} \equiv (X, E_X)$ is a map $D : X \rightarrow H$ such that, for every $x, y \in X$:*
$[Extensionality]$ $\quad D(x) \wedge E_X(x, y) \leq D(y)$;
$[Strictness]$ $\quad D(x) \leq E_X(x, x)$.

Denoting with $\mathcal{P}_H(\mathcal{X})$ the power collection of all H-subsets of a H-set $\mathcal{X} \equiv (X, E_X)$, we can define pointwise all operation between H-subsets, $\cup_H$, $\cap_H$, $\rightarrow_H$, $\bigcap_{i \in I}$, $\bigcup_{i \in I}$ and they satisfy Extensionality and Strictness axioms. Applying a H-valuation to the definition on $\subseteq$, $=$ between subsets appearing in $\mathcal{P}(X)$, we get:

	$\mathcal{P}(X)$		$\mathcal{P}_H(\mathcal{X})$
$\subseteq$	$\forall x(x \,\epsilon\, C \rightarrow x \,\epsilon\, D)$	$\Leftrightarrow$	$\bigwedge_{x \in X}(C(x) \rightarrow D(x)) = 1$
$=$	$\forall x(x \,\epsilon\, C \leftrightarrow x \,\epsilon\, D)$	$\Leftrightarrow$	$\bigwedge_{x \in X}(C(x) \leftrightarrow D(x)) = 1$

and therefore we have the definitions of H-inclusion and H-equality between H-subsets as follows:

Definition 3.7 *For any C and D H-subset of $\mathcal{X}$:*

(a) $C \subseteq_H D \equiv (\forall x \in X)(C(x) \leq D(x))$;

(b) $C =_H D \equiv (\forall x \in X)(C(x) = D(x))$.

Proposition 3.8 $(\mathcal{P}_H(\mathcal{X}), \subseteq_H, \cap_H, \cup_H, 1_\emptyset, 1_X)$ *is a complete Heyting algebra.*

We want to equip it with a structure of overlap algebra. Therefore, we have to define an overlap relation $\between_H$ on $\mathcal{P}_H(\mathcal{X})$. Applying a H-valuation also to the definition of $\between$ in $\mathcal{P}(X)$, we get:

	$\mathcal{P}(X)$		$\mathcal{P}_H(\mathcal{X})$
$\between$	$\exists x(x \in C \& x \in D)$	$\Leftrightarrow$	$\bigvee_{x\in X}(C(x) \wedge D(x)) = 1$

and therefore we have the definition of H-overlap between H-subsets, called **natural** as follows:

Definition 3.9 *For any C and D H-subset of $\mathcal{X}$:*

(a) $C \between_{H,n} D \equiv \bigvee_{x\in X}(C(x) \wedge D(x)) = 1$

The binary relation $\between_{H,n}$ on $\mathcal{P}_H(\mathcal{X})$ **does not work**, since, although it is symmetric and it preserves infima $\cap_H$:

- it does not satisfy density;
- it does not split suprema.

Remark 3.10 Thus, in this way, the power collection $\mathcal{P}_H(X)$ of a H-valued set does not in general enjoy all the algebraic properties of $\mathcal{P}(X)$ that Sambin expressed via the notion of overlap algebra. In more precise terms, the power collection of a H-set is not in general an overlap algebra, if we consider an overlap relation as exactly the many-valued version of that defined in $\mathcal{P}(X)$. Since H itself, seen as $\mathcal{P}(\mathtt{1})$, does not come necessarily equipped with a structure of an overlap algebra, we then conclude that H must be an overlap algebra as a starting point.

4 Sets evaluated over an overlap algebra

The above remark led us to place our project in the context of sets evaluated on an overlap algebra, that is the **$\mathcal{O}$-valued sets**, where $\mathcal{O} \equiv (\mathcal{O}, \leq, \wedge, \vee, \bigwedge, \bigvee, \rightarrow, \between, 0, 1)$ is an overlap algebra. On the complete Heyting algebra $(\mathcal{P}_\mathcal{O}(\mathcal{X}), \subseteq_\mathcal{O}, \cap_\mathcal{O}, \cup_\mathcal{O}, 1_\emptyset, 1_X)$, we define the **induced** overlap $\between_{\mathcal{O},i}$ as follows:

Definition 4.1 *For any $\mathcal{O}$-subset C and D of an $\mathcal{O}$-set $\mathcal{X} \equiv (X, E_X)$:*

$$C \between_{\mathcal{O},i} D \equiv (\exists x \in X)(C(x) \between D(x)).$$

And from now on we don't write the subscript i. The above definition is obtained from the following $\mathcal{O}$-valuation of the definition of $\between$ done in $\mathcal{P}(X)$:

	$\mathcal{P}(X)$	$\mathcal{P}_\mathcal{O}(\mathcal{X})$				
$\between$	$\exists x(x\epsilon C \ \&\ x\epsilon D)$	$\bigvee_{x\in X}(C(x) \wedge D(x)) \between 1$	$\Leftrightarrow$	$\exists x \in X(C(x) \wedge D(x)) \between 1$	$\Leftrightarrow$	$\exists x \in X(C(x) \between D(x))$

With this definition of overlap, we obtain the following result:

Theorem 4.2 *If $\mathcal{O}$ is an overlap algebra, then $\mathcal{P}_\mathcal{O}(\mathcal{X})$ is an overlap algebra with respect to the overlap relation induced from $\mathcal{O}$, that is*

$$C \between_\mathcal{O} D \equiv \exists x \in X(C(x) \between D(x)).$$

Moreover, we are able to prove that every $\mathcal{O}$-relation between $\mathcal{O}$-sets yields an overlap relation between the corresponding power collections of their $\mathcal{O}$-subsets. In fact, if $r : (X, E_X) \to (Y, E_Y)$ is an $\mathcal{O}$-relation between two $\mathcal{O}$-sets $\mathcal{X} \equiv (X, E_X)$ and $\mathcal{Y} \equiv (Y, E_Y)$, then there exist four functions $R, R^{-*} : \mathcal{P}_{\mathcal{O}}(\mathcal{X}) \to \mathcal{P}_{\mathcal{O}}(\mathcal{Y})$, $R^{-}, R^{*} : \mathcal{P}_{\mathcal{O}}(\mathcal{Y}) \to \mathcal{P}_{\mathcal{O}}(\mathcal{X})$, defined as follows, for every $D \subseteq_{\mathcal{O}} \mathcal{X}$, $U \subseteq_{\mathcal{O}} \mathcal{Y}$, $x \in X$ and $a \in Y$:

$$\begin{aligned} RD(a) &\equiv \bigvee_{x \in X} (r\ (x, a) \wedge D(x)) \\ R^{-*}D(a) &\equiv E_S(a, a) \wedge \bigwedge_{x \in X} (r\ (x, a) \to D(x)) \\ R^{-}U(x) &\equiv \bigvee_{a \in Y} (r\ (x, a) \wedge U(a)) \\ R^{*}U(x) &\equiv E_X(x, x) \wedge \bigwedge_{a \in Y} (r\ (x, a) \to U(a)), \end{aligned}$$

and $< R, R^{-}, R^{*}, R^{-*} >$ is an overlap relation between the overlap algebras $\mathcal{P}_{\mathcal{O}}(\mathcal{X})$ and $\mathcal{P}_{\mathcal{O}}(\mathcal{Y})$.

In this way, we have completely defined the Power Functor $\mathcal{P} : \mathbf{Rel}(\mathcal{O}) \to \mathbf{OA}$ as follows:

objects: $\mathcal{P}(X, E_X) \equiv (\mathcal{P}_{\mathcal{O}}(\mathcal{X}), \subseteq_{\mathcal{O}}, \cap_{\mathcal{O}}, \cup_{\mathcal{O}}, \to_{\mathcal{O}}, \lozenge_{\mathcal{O}}, 1_{\emptyset}, 1_X)$

morphisms: $\mathcal{P}(r : (X, E_X) \to (Y, E_Y)) \equiv < R, R^{-}, R^{*}, R^{-*} >$.

The Power Functor $\mathcal{P} : \mathbf{Rel}(\mathcal{O}) \to \mathbf{OA}$ is:

(a) well defined;

(b) faithful;

(c) not full, in general.

To give a complete many-valued version of what it happens in $\mathcal{P}(X)$, we have to define:

- an $\mathcal{O}$-valued overlap algebra, and then the category $\mathbf{oA}(\mathcal{O})$ of $\mathcal{O}$-valued overlap algebras;
- an atomic $\mathcal{O}$-valued overlap algebra, and then the category $\mathbf{a\text{-}oA}(\mathcal{O})$ of $\mathcal{O}$-valued overlap algebras;

in order to see whether it is possible to do a many-valued version of the diagram appearing at the end of Section 2. Since in particular, $\mathcal{P}_{\mathcal{O}}(\mathcal{X})$ will be the main example of an $\mathcal{O}$-valued overlap algebra, then we will give an $\mathcal{O}$-valuation of the definitions of inclusion, equality and overlap between $\mathcal{O}$-subsets of an $\mathcal{O}$-set $\mathcal{X}$, as follows:

Definition 4.3 *For every $\mathcal{O}$-subset C and D of an $\mathcal{O}$-set $\mathcal{X}$:*

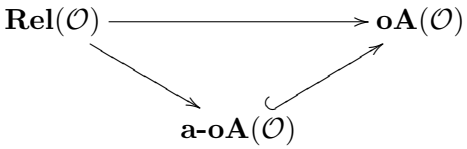
- $S_{\mathcal{P}_{\mathcal{O}}(\mathcal{X})}(C,D) \equiv \bigwedge_{x\in X}(C(x) \to D(x))$ *[subsethood degree];*
- $E_{\mathcal{P}_{\mathcal{O}}(\mathcal{X})}(C,D) \equiv \bigwedge_{x\in X}(C(x) \leftrightarrow D(x))$ *[equality degree];*
- $Ov_{\mathcal{P}_{\mathcal{O}}(\mathcal{X})}(C,D) \equiv \bigvee_{x\in X}(C(x) \wedge D(x))$ *[overlap degree].*

Remark 4.4 Let us observe that we obtain the previous definitions, in the following way:

- $S_{\mathcal{P}_{\mathcal{O}}(\mathcal{X})}(C,D) = 1 \Leftrightarrow C \subseteq_{\mathcal{O}} D$;
- $E_{\mathcal{P}_{\mathcal{O}}(\mathcal{X})}(C,D) = 1 \Leftrightarrow C =_{\mathcal{O}} D$;
- $Ov_{\mathcal{P}_{\mathcal{O}}(\mathcal{X})}(C,D) \gtrless 1 \Leftrightarrow C \between_{\mathcal{O}} D$;

for any $C, D \subseteq_{\mathcal{O}} \mathcal{X}$.

In such a way, since we achieved all our purposes, denoting by $\mathcal{P} : \mathbf{Rel}(\mathcal{O}) \to \mathbf{oA}(\mathcal{O})$, the $\mathcal{O}$-power functor, then we proved that $\mathcal{P}$ is well defined, faithful, full and dense on $\mathcal{O}$-valued atomic overlap algebras, and thus we obtain the commutativity of the following diagram:



References

[1] M. P. Fourman and D. S. Scott, *Notes on sheaves and logic.* In M. P. Fourman, C. J. Mulvey, and Dana S. Scott, Applications of Sheaves: Proceedings of the Research Symposium on Applications of Sheaf Theory to Logic, Algebra, and Analysis, Durham, July 9-21, 1977 (Lecture Notes in Mathematics Vol 753), Springer-Verlag (1979).

[2] M. E. Maietti and G. Sambin, *Toward a minimalist foundation for constructive mathematics.* In From sets and types to topology and analysis, 91–114, Oxford Logic Guides, 48, Oxford Univ. Press, Oxford (2005).

[3] P. Martin-Löf, "Intuitionistic type theory". Notes by G. Sambin of a series of lectures given in Padua, June 1980. Bibliopolis, Napoli, 1984.

[4] G. Sambin, "The Basic Picture - A structural basis for constructive topology". Oxford University Press, to appear.

[5] A. S. Troelstra, D. Van Dalen, "Constructivism in Mathematics. An introduction". Studies in logic and the foundations of mathematics, North Holland, 1988.

Chaotic phenomena described by stochastic equations

Luigi Manca (*)

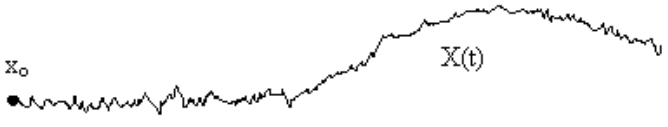
Abstract. It is well known that many natural phenomena such as population dynamics, stock exchange, diffusion of particles, can be seen as 'chaotic'. To give a mathematical description of these 'chaotic' phenomena the theory of stochastic processes and of the related stochastic differential equations has been developed. Starting by the fundamental concept of Brownian motion, I shall introduce the main ideas and the basic tools in order to understand some easy models driven by stochastic equations. Moreover, I shall describe how stochastic equations can be used to study some deterministic model.

Let us consider the ordinary differential equation

$$\text{(ODE)} \qquad \begin{cases} \frac{dx}{dt}(t) = b(x(t)), & t > 0 \\ x(0) = x_0 \end{cases}$$

where $b : \mathbb{R}^n \to \mathbb{R}^n$ is a given smooth vector field and $x_0 \in \mathbb{R}^n$. The solution is the trajectory $x(\cdot) : [0, \infty) \to \mathbb{R}^n$ which looks like

In many application, however, the experimentally measured trajectories of systems models by (ODE) behave as



(*)Grant holder in Pure Mathematics - Dipartimento di Matematica Pura ed Applicata, Via Belzoni 7 - 35131 Padova, Italy - manca@math.unipd.it - Seminar held on 30 April 2008.

Hence it seems reasonable to modify (ODE) in order to include random effects disturbing the system. *Formally* we write

$$\text{(1)} \qquad \begin{cases} \frac{dX}{dt}(t) = b(X(t)) + B(X(t))\xi(t), \quad t > 0 \\ X(0) = x_0 \end{cases}$$

where $B : \mathbb{R}^n \to \mathbb{M}^{n\times m}$ (the space of $n \times m$ matrices) and

$$\xi(\cdot) := m\text{-dimensional "white noise"}.$$

So we need

- Define the "white noise" $\xi(\cdot)$ in a rigorous way;
- Define what it means for $X(\cdot)$ to solve (1).

Some heuristics

Let us study (1) in the case $m = n$, $b \equiv 0$ and $B = I$. The solution of (1) turns to be the n-dimensional *Wiener process*, or *Brownian motion*, denoted by $W(\cdot)$. We write, formally,

$$\frac{dW}{dt}(\cdot) = \xi(\cdot),$$

asserting that *"white noise"* is the *time derivative of the Wiener process*.

With these notations, equation (1) becomes

$$\text{(SDE)} \qquad \begin{cases} dX(t) = b(X(t))dt + B(X(t))dW(t) \\ X(0) = x_0. \end{cases}$$

This expression is called a *stochastic differential equation*. Of course, we need to give it a precise meaning of what we call a *solution* of (SDE). However, this is not the purpose of these notes; so, the reader which wishes to go into this discussion can address to the below references.

We now try to give an idea of what is the *Wiener process* or Brownian motion.

The Brownian motion

Robert Brown in 1826-27 observed the irregular motion of a pollen particles suspended in water. He observed that

- the paht of a given particle is very irregular, having no tangent in any point;
- the motion of two distinct particles appear to be indipendent.

In 1905, Albert Einstein in [1] studied the Brownian phenomena in order to describe how an ink particle diffuses in water. He explained that the ink diffuses due to the collision with the particles of water. Starting with a probabilistic approach, he derived that the density $f(x,t)$ of the ink particles at point x and time t evolves accordingly with the *heat equation* (or *diffusion equation*)

$$f_t = \frac{D}{2} f_{xx},$$

where D is a constant depending by the gas constant, the absolute temperature, the friction coefficient and the Avogrado's number.

As well known, the above equation has solution

$$f(x,t) = \frac{1}{\sqrt{2\pi D t}} e^{-\frac{x^2}{2Dt}}.$$

This equation allowed J.Perrin to compute the Avogrado's number an help to confirm the atomic theory of matter.

The mathematical definition of Brownian motion: the Wiener process

Norbert Wiener in the 1920's put the basis for the mathematical theory of the Brownian motion.

Before giving the precise definition, we need some notation.

With $(\Omega, \mathcal{F}, \mathbb{P})$ we mean a probability space, where Ω is the space, $\mathcal{F}$ is the σ-algebra of measurable sets and $\mathbb{P}$ is the probability measure over $(\Omega, \mathcal{F})$. An elements $A \in \mathcal{F}$ is often called *event*.

We also recall that a *real random variable* is any measurable mapping $X : \Omega \to \mathbb{R}$, where $\mathbb{R}$ is equipped with the usual Borel sets $\mathcal{B}(\mathbb{R})$.

A *stochastic process* is a collection of random variable. We shall work with stochastic process of the form $\{X_t\}_{t\in I}$, where $X_t : \Omega \to \mathbb{R}$ is a real random variable for any $t \in I$, where $I \subset \mathbb{R}$ is the interval $I = [0, \infty)$ or $I = [0, T]$.

With these notations, the *trajectory* of the stochastic process $\{X_t\}_{t\in I}$ at $\omega \in \Omega$ is the real valued function

$$I \to \mathbb{R}, \quad t \mapsto X_t(\omega).$$

Given a real random variable $X : \Omega \to \mathbb{R}$ the *law* of X is the probability measure $\mathcal{L}_X$ on the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ given by

$$\mathcal{L}_X(A) := \mathbb{P}(X^{-1}(A)) \qquad (= \mathbb{P}\{\omega \in \Omega : X(\omega) \in A\})$$

We recall that two events $A, B \in \mathcal{F}$ are *indipendent* if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Finally, two random variables $X, Y : \Omega \to \mathbb{R}$ are indipendent if $\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) = \mathbb{P}(\{X \in A\})\mathbb{P}(\{Y \in B\})$, for any Borel sets $A, B \in \mathcal{B}(\mathbb{R})$. Equivalently, the random variables X, Y are indipendent if for any bounded Lebesgue measurable function $f : \mathbb{R}^2 \to \mathbb{R}$ it holds

$$\mathbb{E}[f(X,Y)] = \int_\Omega f(X(\omega), Y(\omega))\mathbb{P}(d\omega) = \int_{\mathbb{R}^2} f(x,y)\mathcal{L}_X \times \mathcal{L}_Y(dx, dy).$$

Definition 1 [of Wiener process] *A real valued stochastic process $\{B(t)\}_{t\geq 0}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a* Wiener process *if:*

- *The law of $B(0)$ is the Dirac measure at zero:*

$$\mathcal{L}_{B(0)} = \delta_0, \quad \text{where } \delta_0(A) = \begin{cases} = 0 \text{ if } 0 \notin A, \\ = 1 \text{ if } 0 \in A \end{cases} ;$$

  (the process "starts" at 0)

- *for any $0 \leq s < t$ the law of the random variable $B(t) - B(s)$ is gaussian of zero mean and variance $t - s$, that is*

$$\mathcal{L}_{B(t)-B(s)}(A) = \int_A \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{x^2}{2(t-s)}} dx, \quad \forall A \in \mathcal{B}(\mathbb{R});$$

- *for any $0 \leq t_0 < t_1 < \ldots < t_n$ the random variables $B(t_0)$, $B(t_1) - B(t_0)$, $\ldots$, $B(t_n) - B(t_{n-1})$ are indipendent (indipendent increments);*

- *the set $\{\omega \in \Omega :$ the trajectory of ω is continuous$\}$ has measure 1 (the trajectory are almost-surely continuous).*

There are many methods to construct a Wiener process which are in most of the cases quite technical. We refer to [2] for the Levy-Ciesielski construction and [3] for an infinite-dimensional analysis approach.

Properties of trajectories

The trajectories of the Wiener process enjoy many interesting properties. It is well known that they are nowhere differentiable and of infinite variation for each time interval. In particular, the sample path $t \mapsto B(t, \omega)$ is uniformly Hölder continuous for each exponent $\gamma < 1/2$, and it is nowhere Hölder continuous for each exponent $\gamma > 1/2$.

Wiener process and partial differential equations

Example 1 Let $U \subset \mathbb{R}^n$ be a bounded open set, with smooth boundary ∂U. According to stardard PDE theory, there exists a smooth solution u of the equation

$$\begin{cases} -\frac{1}{2}\Delta u = 1 \quad \text{in } U \\ u = 0 \quad \text{on } \partial U. \end{cases}$$

The solution u has a probabilistic representiation via the Wiener process.

For any $x \in U$ we consider the n-dimensional process $X(\cdot) = x + W(\cdot)$, where $W(\cdot)$ is a n-dimensional Wiener process. Define the random variable $\tau_x : \Omega \to [0, \infty]$ by

$$\tau_x = \inf \left\{ t \geq 0 \text{ such that } X(t, \omega) \notin \overline{U} \right\}.$$

It is possible to show that $\tau_x(\omega) < \infty$ for almost all $\omega \in \Omega$. This means that the Brownian sample paths starting at $x \in U$ will reach the boundary of U in a finite time with probability 1.

Such a random variables are often called *stopping time*.

Notice that $\tau_x \equiv 0$ if $x \in \partial U$.

Theorem 1 *We have*

$$u(x) = \mathbb{E}[\tau_x], \quad \text{for all } x \in U.$$

Example 2 Let $U \subset \mathbb{R}^n$ be a smooth, bounded domain and $g : \partial U \to \mathbb{R}$ a given continuous function. It is known that there exists a function $u \in C(\overline{U}) \cap C^2(U)$ satisfying the boundary value problem

$$\begin{cases} \Delta u = 0 & \text{in } U \\ u = g & \text{on } \partial U. \end{cases}$$

u is an *armonic* function.

Theorem 2 *For any $x \in U$ we have*

$$u(x) = \mathbb{E}[g(X(\tau_x))],$$

where $X(\cdot) = x + W(\cdot)$ is a Brownian motion starting at x.

References

[1] A. Einstein, "Investigations on the theory of the Brownian movement". Dover Publications Inc., New York, 1956. Edited with notes by R. Fürth; translated by A. D. Cowper.

[2] H. P. McKean, Jr., "Stochastic integrals". Probability and Mathematical Statistics, No. 5, Academic Press, New York, 1969.

[3] G. Da Prato, "An introduction to infinite-dimensional analysis". Universitext, Springer-Verlag, Berlin, 2006. Revised and extended from the 2001 original by Da Prato.

Lunghezza di curve e misure di superficie

Roberto Monti (*)

1 Lunghezza di curve in $\mathbb{R}^n$ e in spazi metrici

1.1 Lunghezza di curve C^1. La lunghezza di una curva $\gamma : [0,1] \to \mathbb{R}^n$ di classe C^1 si definisce come l'integrale della lunghezza (rispetto alla norma standard di $\mathbb{R}^n$) della sua derivata

$$L(\gamma) = \int_0^1 |\dot\gamma(t)| dt. \tag{1.1}$$

La definizione non dipende dalla parametrizzazione della curva ma solo dal supporto $\gamma([0,1])$. Infatti, se $\varphi : [0,1] \to [0,1]$ è un diffeomorfismo di classe C^1, la riparametrizzazione $\kappa = \gamma \circ \varphi$ ha stessa lunghezza di γ. Se γ è iniettiva, $L(\gamma)$ è la lunghezza del supporto $\gamma([0,1])$, in caso contrario può accadere che alcuni sottoinsiemi del supporto siano misurati con molteplicità maggiore di 1.

1.2 Variazione totale. Se $\gamma : [0,1] \to \mathbb{R}^n$ è una curva continua è possibile definire la sua variazione totale prendendo l'estremo superiore delle lunghezze delle curve poligonali inscritte sulla curva:

$$\mathrm{Var}(\gamma) = \sup \Big\{ \sum_{i=1}^N |\gamma(t_i) - \gamma(t_{i-1})| \,:\, 0 = t_0 < t_1 < ... < t_N = 1 \Big\}. \tag{1.2}$$

L'estremo superiore è calcolato su tutte le scomposizioni dell'intervallo $[0,1]$. Se $\mathrm{Var}(\gamma) < +\infty$ la curva γ si dice rettificabile. Lunghezza e variazione coincidono:

Proposizione 1.1 *Se $\gamma : [0,1] \to \mathbb{R}^n$ è di classe C^1 allora* $\mathrm{Var}(\gamma) = L(\gamma)$.

1.3 Misura di Hausdorff. Sia $K \subset \mathbb{R}^n$ un insieme di cui vogliamo definire la "lunghezza". Fissato un parametro $\delta > 0$, definiamo

$$\mathcal{H}^1_\delta(K) = \inf \Big\{ \sum_{i=1}^{+\infty} \mathrm{diam}(E_i) \,:\, E_i \subset \mathbb{R}^n,\, \mathrm{diam}(E_i) \le \delta,\, K \subset \bigcup_{i=1}^{+\infty} E_i \Big\}. \tag{1.3}$$

(*)Ricercatore presso il Dipartimento di Matematica Pura ed Applicata, Università di Padova, Via Trieste 63, I-35121 Padova, Italy. E-mail: monti@math.unipd.it. Seminario tenuto il 14 maggio 2008.

La funzione $\delta \mapsto \mathcal{H}^1_\delta(K)$ è decrescente e quindi esiste (finito o infinito) il limite

$$\mathcal{H}^1(K) = \lim_{\delta \to 0^+} \mathcal{H}^1_\delta(K). \tag{1.4}$$

La misura di lunghezza $\mathcal{H}^1$ è stata introdotta da Carathéodory nel 1914 [6].

Una curva $\gamma : [0,1] \to \mathbb{R}^n$ si dice Lipschitz se esiste $L > 0$ tale che $|\gamma(t) - \gamma(s)| \leq L|t-s|$ per ogni $s, t \in [0,1]$. Le curve Lipschitz sono derivabili in quasi ogni punto, essendo ciascuna coordinata della curva la differenza di due funzioni monotone.

Proposizione 1.2 *Sia $\gamma : [0,1] \to \mathbb{R}^n$ una curva Lipschitz e iniettiva. Allora:*

$$\mathcal{H}^1(\gamma([0,1])) = \mathrm{Var}(\gamma) = \int_0^1 |\dot\gamma(t)| dt. \tag{1.5}$$

La Proposizione 1.2 ha una naturale generalizzazione nel contesto degli spazi metrici.

1.4 Lunghezza di curve in spazi metrici. Sia (X, d) uno spazio metrico. Una curva $\gamma : [0,1] \to X$ si dice Lipschitz se esiste $L > 0$ tale che

$$d(\gamma(t), \gamma(s)) \leq L|t - s| \tag{1.6}$$

per ogni $s, t \in [0,1]$. La variazione totale di γ è

$$\mathrm{Var}(\gamma) = \sup \Big\{ \sum_{i=1}^N d(\gamma(t_i), \gamma(t_{i-1})) \,:\, 0 = t_0 < t_1 < ... < t_N = 1 \Big\}. \tag{1.7}$$

È chiaramente $\mathrm{Var}(\gamma) \leq L$.

Se $K \subset X$ è un insieme, si definisce la "lunghezza di K"

$$\mathcal{H}^1(K) = \lim_{\delta \to 0} \inf \Big\{ \sum_{i=1}^{+\infty} \mathrm{diam}(E_i) \,:\, E_i \subset X,\, \mathrm{diam}(E_i) \leq \delta,\, K \subset \bigcup_{i=1}^{+\infty} E_i \Big\} \tag{1.8}$$

Ora il diametro degli insiemi è calcolato rispetto alla distanza d di X.

Teorema 1.3 *Sia $\gamma : [0,1] \to (X, d)$ una curva Lipschitz. Allora:*

(i) *Il limite ("derivata metrica")*

$$|\dot\gamma|(t) := \lim_{\delta \to 0} \frac{d(\gamma(t+\delta), \gamma(t))}{|\delta|} \tag{1.9}$$

esiste per q.o. $t \in [0,1]$.

(ii) *Se, inoltre, γ è iniettiva, allora*

$$\mathcal{H}^1(\gamma([0,1])) = \mathrm{Var}(\gamma) = \int_0^1 |\dot\gamma|(t) dt. \tag{1.10}$$

Una prova del Teorema 1.3 si trova nel Capitolo 4 di [2].

2 Area di ipersuperfici e perimetri di insiemi

2.1 Formula dell'area per 2-superfici in $\mathbb{R}^3$. Siano $D \subset \mathbb{R}^2$ un insieme aperto e $\varphi : D \to \mathbb{R}^3$ una funzione di classe C^1 tale che le derivate parziali φ_x e φ_y siano linearmente indipendenti. L'insieme $\varphi(D) \subset \mathbb{R}^3$ è una superficie di classe C^1 che può avere autointersezioni ma se φ è iniettiva non ci sono autointersezioni. In questo caso, si definisce l'area di $\varphi(D)$ come

$$\text{Area}(\varphi(D)) = \int_D |\varphi_x \times \varphi_y| dxdy. \tag{2.11}$$

Questa definizione, in effetti, non dipende dalla parametrizzazione φ dell'insieme $\varphi(D)$. L'espressione $|\varphi_x \times \varphi_y|$ indica la lunghezza in $\mathbb{R}^3$ del vettore $\varphi_x \times \varphi_y$ e coincide con l'area del parallelogramma identificato dai vettori φ_x e φ_y.

2.2 Formula dell'area per grafici. Siano $D \subset \mathbb{R}^{n-1}$ un insieme aperto e $f : D \to \mathbb{R}$ una funzione di classe C^1. Il grafico della funzione f è l'insieme $\text{gr}(f) = \{(x, f(x)) \in \mathbb{R}^n : x \in D\}$ e la sua area è per definizione

$$\text{Area}(\text{gr}(f)) = \int_D \sqrt{1 + |\nabla f(x)|^2} dx. \tag{2.12}$$

Se la funzione φ in (2.11) è una parametrizzazione di tipo Cartesiano le definizioni di area (2.11) e (2.12) con $n = 3$ coincidono.

2.3 Misura di Hausdorff. Siano $K \subset \mathbb{R}^n$ un insieme e $\delta > 0$. Definiamo

$$\mathcal{H}^{n-1}_\delta(K) = \inf\Big\{\omega_{n-1} \sum_{i=1}^{+\infty} \Big(\frac{\text{diam}(E_i)}{2}\Big)^{n-1} \,:\, E_i \subset \mathbb{R}^n,\, \text{diam}(E_i) \le \delta,\, K \subset \bigcup_{i=1}^{+\infty} E_i\Big\}, \tag{2.13}$$

dove ω_{n-1} indica la misura di Lebesgue della palla unitaria in $\mathbb{R}^{n-1}$.

La funzione $\delta \mapsto \mathcal{H}^{n-1}_\delta(K)$ è decrescente e quindi esiste (finito o infinito) il limite

$$\mathcal{H}^{n-1}(K) = \lim_{\delta \to 0^+} \mathcal{H}^{n-1}_\delta(K). \tag{2.14}$$

Analoghe misure $\mathcal{H}^s$ per ogni reale $0 \le s \le n$ sono state introdotte nel 1918 da Hausdorff [12] che ha generalizzato la costruzione di Carathéodory.

Il legame fra misure di Hausdorff e area di grafici o immagini Lipschitziane è dato dalle formule dell'area.

Teorema 2.1 (Formula dell'area) *Siano $D \subset \mathbb{R}^{n-1}$ un compatto ed $f : D \to \mathbb{R}$ una funzione Lipschitz. Allora*

$$\mathcal{H}^{n-1}(\text{gr}(f)) = \int_D \sqrt{1 + |\nabla f(x)|^2} dx. \tag{2.15}$$

Formule analoghe sono note per superfici parametriche Lipschitz in $\mathbb{R}^n$ di dimensione $1 \leq m \leq n-1$ (Cfr. [10, Sez. 3.2] oppure [9, Capitolo 3]).

2.4 Contenuto di Minkowski. L'intorno tubolare di raggio $r > 0$ di un insieme $K \subset \mathbb{R}^n$ è l'insieme

$$K_r = \big\{x \in \mathbb{R}^n : \operatorname{dist}(x, K) < r\big\}. \tag{2.16}$$

Il contenuto di Minkowski $(n-1)$-dimensionale di K è - se il limite esiste -

$$\mathcal{M}^{n-1}(K) = \lim_{r \downarrow 0} \frac{\mathcal{L}^n(K_r)}{2r}, \tag{2.17}$$

dove $\mathcal{L}^n$ è la misura di Lebesgue in $\mathbb{R}^n$. Il contenuto di Minkowski $(n-s)$-dimensionale con $0 < s < n$ può essere definito in modo analogo, sostituendo $\omega_s r^{n-s}$ al posto di $2r$ nella (2.17).

Teorema 2.2 *Siano $D \subset \mathbb{R}^{n-1}$ un compatto, $\varphi : D \to \mathbb{R}^n$ una funzione Lipschitz e $K = \varphi(D)$. Allora*

$$\mathcal{M}^{n-1}(K) = \mathcal{H}^{n-1}(K). \tag{2.18}$$

La dimostrazione del teorema è di Federer [11, Sez. 3.2.37] (Cfr. anche [3, Sez. 2.13]).

2.5 Idea di Caccioppoli. Si vuole definire il "perimetro di E" ovvero la misura del "bordo di E" quando $E \subset \mathbb{R}^n$ è un insieme misurabile limitato. L'idea di Caccioppoli [5] è di approssimare E con insiemi regolari. Se $E, F \subset \mathbb{R}^n$ sono due insiemi misurabili, la funzione

$$\delta(E, F) = \mathcal{L}^n(E \setminus F) + \mathcal{L}^n(F \setminus E) \tag{2.19}$$

è una distanza. Si può allora definire:

$$P_C(E) = \inf\big\{ \liminf_{j \to \infty} P(F_j) : (F_j)_{j \in \mathbb{N}} \text{ poliedri tali che } F_j \to E \text{ in } \delta\big\}. \tag{2.20}$$

Se $F \subset \mathbb{R}^n$ è un poliedro il suo perimetro $P(F)$ è la somma delle aree delle facce di F. Al posto di poliedri è equivalente prendere insiemi con frontiera di classe C^∞ e definire il loro perimetro tramite una delle formule precedenti.

Il perimetro di Caccioppoli è l'inviluppo semicontinuo inferiore della misura di frontiere orientate regolari rispetto alla topologia L^1 di insiemi. La necessità di prendere l'estremo inferiore è dovuta al fatto che l'area delle triangolazioni di superfici con converge, in generale, all'area della superfice quando il diametro dei triangoli tende a zero (Schwarz ma mostrato questo fenomeno nel caso di un cilindro).

2.6 Idea di De Giorgi. La funzione caratteristica di un insieme $E \subset \mathbb{R}^n$ è la funzione χ_E tale che $\chi_E(x) = 1$ se $x \in E$ ed è 0 altrimenti. Dato un insieme misurabile e limitato

$E \subset \mathbb{R}^n$, sia $u \in C^\infty(\mathbb{R}^n \times (0,+\infty))$ la soluzione del problema di Cauchy per l'equazione del calore

$$\begin{cases} u_t = \Delta u & \text{in } \mathbb{R}^n \times (0,+\infty) \\ u(\cdot,0) = \chi_E, \end{cases} \tag{2.21}$$

dove il dato iniziale è preso in $L^1(\mathbb{R}^n)$. Usando le proprietà del semigruppo del calore, nel 1953 De Giorgi [8] ha osservato che la funzione della variabile $t > 0$

$$I(t;E) = \int_{\mathbb{R}^n} |\nabla_x u(x,t)| dx \tag{2.22}$$

è monotona decrescente. Quindi esiste, finito o infinito, il limite

$$P_{DG}(E) = \lim_{t\to 0} I(t;E). \tag{2.23}$$

Nello stesso lavoro De Giorgi prova che questa definizione coincide con quella di Caccioppoli: $P_C(E) = P_{DG}(E)$.

Dunque, il perimetro di un insieme può essere pensato come la norma L^1 del gradiente della funzione χ_E. Una combinazione delle idee di Caccioppoli e De Giorgi suggerisce di definire il perimetro di E come

$$\inf\Big\{ \liminf_{j\to\infty} \int_{\mathbb{R}^n} |\nabla u_j(x)| dx : u_j \in \mathrm{Lip}_0(\mathbb{R}^n),\ j \in \mathbb{N}, \text{ tali che } u_j \to \chi_E \text{ in } L^1(\mathbb{R}^n)\Big\}. \tag{2.24}$$

Questa definizione coincide con le precedenti e può essere generalizzata in spazi metrici con misura ([1] e [14]). Le funzioni $u_j \in \mathrm{Lip}_0(\mathbb{R}^n)$ sono nello spazio di Sobolev $W^{1,1}(\mathbb{R}^n)$, ma per la non riflessività della norma L^1 il loro limite esce da questo spazio.

2.7 Definizione variazionale. Sia $E \subset \mathbb{R}^n$ un aperto limitato di classe C^2. Data $\varphi \in C^1_0(\mathbb{R}^n)$, dal Teorema della divergenza segue che

$$\int_E \mathrm{div}\varphi \, dx = \int_{\partial E} \langle \varphi, \nu \rangle \, d\mathcal{H}^{n-1}, \tag{2.25}$$

dove ν è la normale esterna a ∂E. È possibile scegliere $\varphi \in C^1_0(\mathbb{R}^n)$ tale che $\varphi = \nu$ su ∂E. In questo caso si ottiene

$$\int_E \mathrm{div}\varphi \, dx = \mathcal{H}^{n-1}(\partial E). \tag{2.26}$$

Altre scelte con il vincolo $|\varphi| \le 1$ forniscono un numero minore. Dunque, per un insieme misurabile E si può definire:

$$P(E) = \sup\Big\{ \int_E \mathrm{div}\varphi \, dx : \varphi \in C^1_0(\mathbb{R}^n),\ |\varphi| \le 1 \Big\}. \tag{2.27}$$

Se $P(E) < +\infty$ l'insieme E si dice di perimetro finito. Mario Miranda [13] ha osservato nel 1964 che per gli insiemi di perimetro finito le derivate parziali distribuzionali della funzione

χ_E sono misure. Questa è la nozione di perimetro usata attualmente. Essa infatti produce in modo naturale una nozione di "normale" e di "frontiera" di E (Cfr. [3]). Ovviamente, questo perimetro coincide con quello di De Giorgi: $P(E) = P_{DG}(E)$.

2.8 Idea di Brezis. Una recente caratterizzazione del perimetro è stata suggerita dalla teoria degli spazi di Sobolev (Cfr. [4] e [7]):

Teorema 2.3 *Se $E \subset \mathbb{R}^n$ è un insieme limitato con perimetro finito, allora*

$$P(E) = \lim_{\varepsilon \to 0} \frac{1}{\omega_n \varepsilon^n} \int_{\mathbb{R}^n} \int_{|x-y|<\varepsilon} \frac{|\chi_E(x) - \chi_E(y)|}{|x-y|} dy\, dx. \tag{2.28}$$

Questa definizione si può riformulare in modo naturale anche in uno spazio metrico con misura. In tale contesto, tuttavia, non è chiaro il suo legame con altre nozioni di misura di ipersuperfice che è possibile introdurre.

Bibliografia

[1] L. Ambrosio, *Some fine properties of sets of finite perimeter in Ahlfors regular metric measure spaces*. Adv. Math. 159 (2001), no. 1, 51–67.

[2] L. Ambrosio & P. Tilli, "Selected topics on "analysis in metric spaces"". Appunti dei corsi tenuti da docenti della Scuola. Scuola Normale Superiore, Pisa, 2000.

[3] L. Ambrosio, N. Fusco, D. Pallara, "Functions of bounded variation and free discontinuity problems". Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[4] H. Brezis, *How to recognize constant functions. A connection with Sobolev spaces.* (Russian). Uspekhi Mat. Nauk 57 (2002), no. 4(346), 59–74; traduzione inglese in Russian Math. Surveys 57 (2002), no. 4, 693–708.

[5] R. Caccioppoli, *Misura e integrazione sugli insiemi dimensionalmente orientati. II.* (Italian). Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat. (8) 12, (1952), 3–11 e 137–146.

[6] C. Carathéodory, *Über das lineare Mass von Punktmengen - eine Veralgemeinerung des Längebegriffs*. Nachrichten K. Gesell. Wissensch. Gött. (1914), 404–426; traduzione inglese in A. Edgar, Classics on fractals, 2004.

[7] J. Dávila, *On an open question about functions of bounded variation*. Calc. Var. Partial Differential Equations 15 (2002), no. 4, 519–527.

[8] E. De Giorgi, *Definizione ed espressione analitica del perimetro di un insieme* (Italian). Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat. (8) 14, (1953).

[9] L. Evans and R. Gariepy, "Measure theory and fine properties of functions". Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[10] K. Falconer, "The geometry of fractal sets". Cambridge Tracts in Mathematics, 85. Cambridge University Press, Cambridge, 1986.

[11] H. Federer, "Geometric measure theory". Die Grundlehren der mathematischen Wissenschaften, Band 153 Springer-Verlag New York Inc., New York, 1969.

[12] F. Hausdorff, *Dimension und äusseres Mass*. Math. Annalen 79 (1918), 157–179; traduzione inglese in A. Edgar, Classics on fractals, 2004.

[13] Ma. Miranda, *Distribuzioni aventi derivate misure e insiemi di perimetro localmente finito*. Ann. Scuola Norm. Sup. Pisa (3) 18 (1964) 27–56.

[14] Mi. Miranda, *Functions of bounded variation on "good" metric spaces*. J. Math. Pures Appl. (9) 82 (2003), no. 8, 975–1004.

[15] Mi. Miranda, D. Pallara, F. Paronetto and M. Preunkert, *Heat semigroup and functions of bounded variation on Riemannian manifolds*. J. Reine Angew. Math. 613 (2007), 99–119.

An introduction to Stochastic Fluid Dynamic Models

David Barbato (*)

Abstract. The Navier-Stokes problem, still unsolved by more than 150 years, represents the starting point for lots of mathematical research topics. The aim of the talk is to present selected fluidodynamic models, in the deterministic and stochastic case, developed from Navier-Stokes equations. In particular the GOY shell model, a Fourier system simplified with respect to the Navier-Stokes one, will be described, and some recent rigorous results discussed.

Introduction

There are several difficult open issues in fluid dynamics related to energy cascade in Navier-Stokes equations. Due to the difficulty of such questions, it may be of help to analyse GOY model which captures some properties of the Navier-Stokes equations while simplifies others. The paper is organised as follows: in the first paragraph Navier-Stokes equations will be shown, in the second paragraph the Fourier Transform of Navier-Stokes equations will be derived, whereas in the third paragraph the stochastic GOY model is introduced.

1 Navier-Stokes equations

We consider the Navier-Stokes equations of a viscous, constant density, Newtonian fluid described by the system:

$$\begin{cases} \frac{\partial v}{\partial t} + (v \cdot \nabla)v = -\nabla p + \mu \Delta v + f \\ \operatorname{div} v = 0 \\ v(0,x) = v_0(x) \end{cases} \tag{1}$$

The unknows functions are $v(x,t) = (v_1(x,t), v_2(x,t), \dots, v_d(x,t))$ and $p(x,t)$ where $x = (x_1, \dots, x_d)$ and $t \geq 0$.
$v(x,t)$ is the velocity field of the moving gas or liquid, whereas $p(x,t)$ is the pressure. μ

(*)Researcher, Dipartimento di Matematica Pura ed Applicata, Università di Padova, Via Trieste 63, I-35121 Padova, Italy. E-mail: barbato@math.unipd.it . Seminar held on 28 May 2008.

is the kinematic viscosity and $f = (f_1(x,t), f_2(x,t), \dots, f_d(x,t))$ is a given force acting on the body.
Usually three cases are considered: I) $x \in \mathbb{R}^d$; II) $x \in \mathbb{T}^d$; III) $x \in Q \subseteq \mathbb{R}^d$ where Q is compact, have a smooth boundary and $u(x,t) = 0$ for $x \in \partial Q$. In this lecture we deal with the case II) and $f = 0$. About notations, recall that $\operatorname{div} v = \sum_{j=1}^d \frac{\partial v_j}{\partial x_j}$; we write $(v \cdot \nabla)v$ for the vector field $\sum_{k=1}^d v_k \frac{\partial v}{\partial x_k}$ and the Laplacian Δ is the differential operator $\sum_{k=1}^d \frac{\partial^2}{\partial x_k^2}$.
Meaning of the equations: the second equation of system (1) is the incompressibility condition; the third is the initial condition; the first equation is the momentum conservation, $(v\cdot\nabla)v$ is the convective acceleration the only non-linear term, ∇p is the pressure gradient and $\mu\Delta v$ is the viscosity term.

Energy balance

The basic property of the system (1) is the energy inequality. Let us suppose (v, p) be a suffinciently regular solution of the Navier-Stokes equations in a domain D.

$$\frac{\partial v_i}{\partial t} + \sum_{k=1}^d v_k \frac{\partial v_i}{\partial x_k} = -\frac{\partial p}{\partial x_i} + \mu \sum_{k=1}^d \frac{\partial^2 v_i}{\partial x_k^2} \tag{2}$$

Let $E(v)$ be the energy

$$E(v) = \frac{1}{2}\int_D \|v(x,t)\|^2 \, dx$$

$$\frac{dE}{dt} = \int_D \sum_{j=1}^d v_j(x,t) \frac{\partial v_j(x,t)}{\partial t} dx$$

$$= \int_D -\sum_{j=1}^d \sum_{k=1}^d v_j(x,t) \frac{\partial v_j(x,t)}{\partial x_k} v_k(x,t)$$

$$-\sum_{j=1}^d v_j(x,t) \frac{\partial p(x,t)}{\partial x_j}$$

$$+\mu \sum_{j=1}^d v_j(x,t) \Delta v_j(x,t) dx$$

The integration by parts of the first two terms gives zero in view of incompressibility condition. Whereas the integration by parts of the third term gives $-\mu \int_D \sum_{j=1}^d |\nabla v_j(x,t)|^2 \, dx$. Thus

$$\frac{dE}{dt} = -\mu \int_D \sum_{j=1}^d |\nabla v_j(x,t)|^2 \, dx \le 0$$

which is the main energy inequality.

2 The Navier-Stokes equations in Fourier coordinates

We derive the Navier-Stokes equations in the domain $\mathbb{T}^d = [0, 2\pi]^d$ with periodic boundary conditions, in the Fourier components. Consider the Fourier basis of $L^2([0, 2\pi]^d)$

$$e_{\mathbf{k}}(\mathbf{x}) = e^{i\langle \mathbf{k}, \mathbf{x}\rangle} \qquad \forall\ \mathbf{k} \in \mathbb{Z}^d\ ,\ \forall \mathbf{x} \in \Omega$$

$\{e_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d\}$ is an orthogonal basis of the Hilbert space $L^2([0, 2\pi]^d)$, so it is possible to write v in Fourier components

$$v(\mathbf{x}, t) = \sum_{\mathbf{k}\in\mathbb{Z}^d} C_{\mathbf{k}}(t) e_{\mathbf{k}}(\mathbf{x}) \qquad \forall\ t \geq 0\ ,\ \forall \mathbf{x} \in \Omega$$

with

$$C_{\mathbf{k}}(t) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} v(\mathbf{x}, t) e_{\mathbf{k}}(\mathbf{x}) d\mathbf{x} \qquad \forall\ \mathbf{k} \in \mathbb{Z}^d\ ,\ \forall\ t \geq 0$$

The real condition and the divergence-free condition become:

$$\left[\begin{array}{lcll} v \in \mathbb{R}^d & \Longleftrightarrow & C_{-k} = \overline{C_k} & \forall\ \mathbf{k} \in \mathbb{Z}^d \\ \operatorname{div} v = 0 & \Longleftrightarrow & \langle C_{\mathbf{k}}, \mathbf{k}\rangle = 0 & \forall\ \mathbf{k} \in \mathbb{Z}^d \end{array}\right.$$

Where $\overline{C_k}$ denote the conjugate of C_k and the second equality follows by

$$\frac{\partial e_{\mathbf{k}}}{\partial \mathbf{x}_j} = i k_j e_{\mathbf{k}} \qquad \text{for all} \qquad \mathbf{k} = (k_1, k_2, \ldots, k_d)$$

The Navier-Stokes equation (2) becomes:

$$\sum_{\mathbf{k}\in\mathbb{Z}^d} \frac{\partial}{\partial t} C_{\mathbf{k}} e_{\mathbf{k}} + \sum_{\substack{\mathbf{k},\tilde{\mathbf{k}},\tilde{\mathbf{h}}\in\mathbb{Z}^d \\ \tilde{\mathbf{k}}+\tilde{\mathbf{h}}=\mathbf{k}}} i\langle C_{\tilde{\mathbf{h}}}, \mathbf{k}\rangle \Big(C_{\tilde{\mathbf{k}}} - \frac{\mathbf{k}}{|\mathbf{k}|^2} \langle C_{\tilde{\mathbf{k}}}, \mathbf{k}\rangle \Big) e_{\mathbf{k}} = -\mu \sum_{\mathbf{k}\in\mathbb{Z}^d} |\mathbf{k}|^2 C_{\mathbf{k}} e_{\mathbf{k}}$$

$$\frac{\partial}{\partial t} C_{\mathbf{k}} = -\mu |\mathbf{k}|^2 C_{\mathbf{k}} - \sum_{\substack{\tilde{\mathbf{k}},\tilde{\mathbf{h}}\in\mathbb{Z}^d \\ \tilde{\mathbf{k}}+\tilde{\mathbf{h}}=\mathbf{k}}} i\langle C_{\tilde{\mathbf{h}}}, \mathbf{k}\rangle \Big(C_{\tilde{\mathbf{k}}} - \frac{\mathbf{k}}{|\mathbf{k}|^2} \langle C_{\tilde{\mathbf{k}}}, \mathbf{k}\rangle \Big) \quad \forall \mathbf{k} \in \mathbb{Z}^d \tag{3}$$

The projection onto the space of divergence-free vector fields is defined as

$$\mathcal{P}(a e_k) = \left(a - \frac{\langle a, \mathbf{k}\rangle}{|\mathbf{k}|^2} \mathbf{k} \right) e_k$$

Whereas P_k denote projection on the $k-$th component $P_k(a) = a - \frac{\langle a, \mathbf{k}\rangle}{|\mathbf{k}|^2}\mathbf{k}$ The equation (3) becomes:

$$\frac{\partial}{\partial t} C_{\mathbf{k}} = -\mu |\mathbf{k}|^2 C_{\mathbf{k}} - P_{\mathbf{k}} \sum_{\substack{\tilde{\mathbf{k}},\tilde{\mathbf{h}}\in\mathbb{Z}^d \\ \tilde{\mathbf{k}}+\tilde{\mathbf{h}}=\mathbf{k}}} i\langle C_{\tilde{\mathbf{h}}}, \mathbf{k}\rangle C_{\tilde{\mathbf{k}}} \qquad \forall \mathbf{k} \in \mathbb{Z}^d \tag{4}$$

Two elementary examples of flux in dimension d=2, showing the relationship between the wavenumber and the vortex radius, are reported (Fig. 1)
The first example is defined by:

$$v_1 = +e_{(1,1)} + e_{(1,-1)} + e_{(-1,1)} + e_{(-1,-1)}$$

$$v_2 = -e_{(1,1)} + e_{(1,-1)} + e_{(-1,1)} - e_{(-1,-1)}$$

Whereas the second example is defined by:

$$v_1 = +e_{(4,4)} + e_{(4,-4)} + e_{(-4,4)} + e_{(-4,-4)}$$

$$v_2 = -e_{(4,4)} + e_{(4,-4)} + e_{(-4,4)} - e_{(-4,-4)}$$

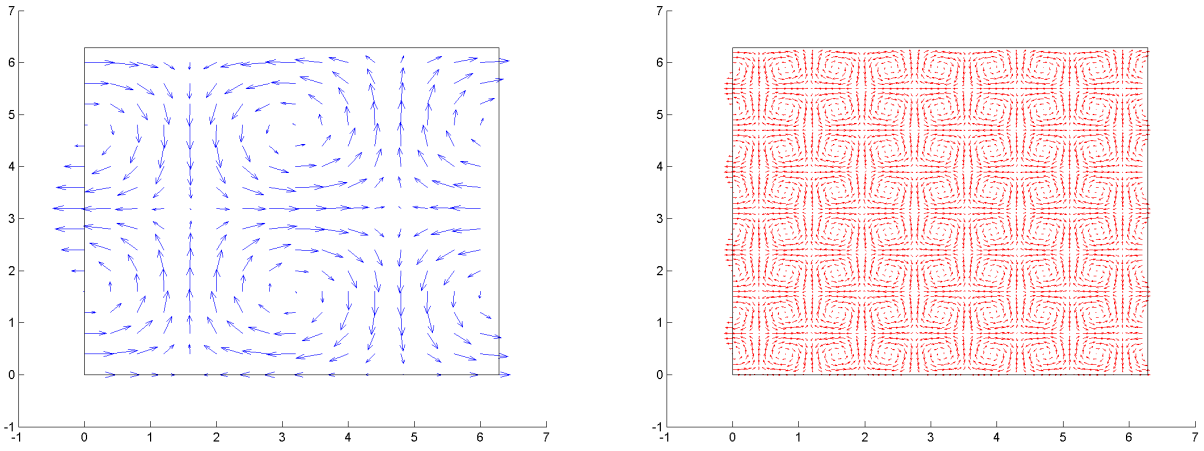


Figure 1. Examples of flux in dimension 2: a) flux with wavenumber 1 and b) flux with wavenumber 4.

Energy balance in Fourier components

We denote the Energy in the Fourier components

$$E(t) = \frac{1}{2} \sum_{k \in \mathbb{Z}^d} \|C_k(t)\|^2 = \frac{1}{2(2\pi)^d} \int_{\mathbb{T}^d} \|v(t)\|^2 \, dx$$

By a simple calculation it is possible to derive the following energy balance.

$$\frac{dE}{dt} = -\mu \sum_{k \in \mathbb{Z}^d} \|k\|^2 \, \|C_k(t)\|^2 \tag{5}$$

This equality shows that the bilinear term of the Navier-Stokes equation (4) does not cause direct dissipation of energy. Whereas the linear term of (4) $-\mu\|k\|^2 C_k(t)$ cause dissipation of energy with rate proportional to the square of the wavenumber k.

3 Goy model

The GOY model, from E. B. Gledzer, K. Ohkitani, M. Yamada, is a simplified Fourier system with respect to the Navier-Stokes one. It describes the dinamics of scalar velocity field in complex Fourier components u_n. $(u_n(t))_{n\geq -1}$ are complex value functions with $u_{-1}(t) = u_0(t) = 0$ for all $t \geq 0$ and satisfing for $n \geq 1$ the equation

$$\frac{d}{dt}u_n = -\nu k_n^2 u_n + ik_n \left(\frac{1}{4}\bar{u}_{n-1}\bar{u}_{n+1} - \bar{u}_{n+1}\bar{u}_{n+2} + \frac{1}{8}\bar{u}_{n-2}\bar{u}_{n-1}\right) \tag{6}$$

where $\nu \geq 0$ will be called viscosity, $\bar{u}_n$ denotes the complex conjugate, $k_n = k_0 2^n$ will be the wavenumbers associated to u_n. Heuristically, in the GOY model, the variable u_n^2 will correspond to the energy of the NavierStokes Fourier components of wavenumbers in a shell, namely the interval (k_n, k_{n+1}). $u_n^2 \sim \sum_{k_n<|k|<k_{n+1}} \|C_k\|^2$. As in the Navier-Stokes system the energy of GOY model is defined by

$$E(t) = \frac{1}{2}\sum_{n\geq 1} |u_n(t)|^2$$

With easy calculation it is possible to obtain the following energy balance equation

$$\frac{dE}{dt} = -\nu \sum_{n\geq 1} k_n^2 |u_n|^2$$

This equation justifies the analogy between the Navier-Stokes equations for Fourier components and the GOY model: the two model have the same energy dissipation rate, proportional to the wavenumber square.

Stochastic GOY model

The stochastic GOY model can be described by the following differential equation:

$$du_n = -\nu k_n^2 u_n dt + ik_n \left(\frac{1}{4}\bar{u}_{n-1}\bar{u}_{n+1} - \bar{u}_{n+1}\bar{u}_{n+2} + \frac{1}{8}\bar{u}_{n-2}\bar{u}_{n-1}\right) dt + \sigma_n d\beta_n \tag{7}$$

where $(\beta_n)_{n\geq 1}$ is a sequence of independent complex-valued Brownian motions on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ and σ_n are different from zero only for a finite values of n. The initial condition $(u_n(0))_{n\geq 1}$ is imposed a $\mathcal{F}_0$ measurable random variable with $\sum_{n\geq 1} |u_n(0)|^2 < \infty$. The stochastic goy model well posedness is proved in [1]. The following theorem holds:

Theorem 3.1 *Given an $\mathcal{F}_0$-measurable r.v. $u_0 : \Omega \to H$, there is a unique continuous in H adapted process $(u(t))_{t\geq 0}$ solution of Eq. (7).*

Where H is the space

$$H = \left\{u = (u_1, \dots) \in \mathbb{C}^\infty \ : \sum_{n=1}^{\infty} |u_n|^2 < \infty\right\}$$

with the norm given by $|u|_H^2 = \sum_{n=1}^{\infty} |u_n|^2$.

References

[1] D. Barbato, M. Barsanti, H. Bessaih, F. Flandoli, *Some Rigorous Results on a Stochastic GOY Model*. Journal of Statistical Physics 43-3 (2006), 673–712.

[2] P. Constantin, B. Levant, E. Titi, *Analytic study of shell models of turbulence*. Physica D, 219 (2006), 120–141.

[3] F. Flandoli, "An Introduction to 3D Stochastic Fluid Dynamics". CIME Lecture Notes, 2005.

[4] G. Gallavotti, "Foundations of Fluid Dynamics". Texts and Monographs in Physics, Springer-Verlag, Berlin, 2002. Translated from the Italian.

[5] A. N. Kolmogorov, *The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers*. Reprinted in Proc. Roy. Soc. London, Ser. A 434 (1991), 9–13. , ,