Università di Padova – Dipartimento di Matematica

Scuole di Dottorato in Matematica Pura e Matematica Computazionale

Seminario Dottorato 2011/12

Preface

This document offers a large overview of the nine months' schedule of Seminario Dottorato 2011/12. Our "Seminario Dottorato" (Graduate Seminar) is a double-aimed activity. At one hand, the speakers (usually Ph.D. students or post-docs, but sometimes also senior researchers) are invited to think how to communicate their own researches to a public of mathematically well-educated but not specialist people, by preserving both understandability and the flavour of a research report. At the same time, people in the audience enjoy a rare opportunity to get an accessible but also precise idea of what's going on in some mathematical research area that they might not know very well.
Let us take this opportunity to warmly thank the speakers once again, in particular for their nice agreement to write down these notes to leave a concrete footstep of their participation. We are also grateful to the collegues who helped us, through their advices and suggestions, in building an interesting and culturally complete program.

Padova, 25 June 2012

Corrado Marastoni, Tiziano Vargiolu

Abstracts (from Seminario Dottorato's web page)

Wednesday 19 October 2011

Hopf Algebras. An introduction

Agustín GARCÍA IGLESIAS (Univ. Cordoba - Argentina)

In this introductory talk we will motivate and present the definition of a Hopf algebra and we will review some of its main properties. The talk will be illustrated with several examples. The concept of a Hopf algebra, in a slightly different version as we know it today, was introduced in the 50's and got its final shape towards the late 60's. They became popular in the 80's, with the appearance of quantum groups and their relation with phenomena of algebraic groups in positive characteristic. Soon after that, they spread over various fields of mathematics and mathematical physics.

Wednesday 2 November 2011

The explicit Laplace transform for the Wishart process

Alessandro GNOATTO (Padova, Dip. Mat.)

The first part of the talk will provide an introduction to mathematical finance. We start with a brief historical and philosophical perspective on the field and briefly review the main problems one tries to solve. We concentrate on the valuation of simple derivatives and review the famous Black-Scholes formula. After that we recall how the weaknesses of this standard approach motivated the introduction of more advanced models, in particular stochastic volatility (SV) models. In the context of SV models we review the role played by characteristic functions and the fast Fourier transform. This last point will serve as an introduction to the results of the paper, where we derive the explicit formula for the joint Laplace transform of the Wishart process and its time integral which extends the original approach of Bru (1991). We compare our methodology with the alternative results given by the variation of constants method, the linearization of the Matrix Riccati ODE's and the Runge-Kutta algorithm. The new formula turns out to be fast, accurate and very useful for applications when dealing with stochastic volatility and stochastic correlation modelling.

Wednesday 16 November 2011

L^2 theory and global regularity for $\overline{\partial}$ on pseudoconvex domains of $\mathbb{C}^n$

Stefano PINTON (Padova, Dip. Mat.)

This seminar is divided into two parts. The first one is an introduction to the first order partial differential operator $\overline{\partial}$ in a smooth bounded pseudoconvex domain D of $\mathbb{C}^n$. Only preliminary definitions are given and the basic estimate, due to Morrey, Kohn and Hormander, is established. It yields the existence of the d-bar Neumann operator, that is, the inverse to the complex Laplacian and the construction of the canonical solution to the equation $\overline{\partial}u = f$, for $f \in \ker \overline{\partial}$, that is the solution orthogonal to the kernel of $\overline{\partial}$. The second part is an introduction to the problem of the global regularity up to the boundary for the canonical solution of the $\overline{\partial}$ equation with data regular up to the boundary. In particular it is shown how compactness estimates are sufficient for global regularity as well as the existence of "good defining functions".

Wednesday 23 November 2011

Boundedness and compactness of matrix operators in weighted spaces of sequences

Zhanar TASPAGANBETOVA (University of Astana and Dip. Mat. Padova)

One of the main problems in the theory of matrices is to find necessary and sufficient conditions for the elements of a matrix so that the corresponding matrix operator maps continuously one normed space of sequences into another space of sequences. Thus it is very important to find the norm of a matrix operator, or at least, an upper or lower bound for the norm. However, in several spaces, which are very important both theoretically and in the applications, such problems have not been solved yet in full generality for operators corresponding to arbitrary matrices. Therefore, in such spaces researchers have considered some specific classes of matrix operators and have established criteria of boundedness and compactness for operators of such classes.
We prove a new discrete Hardy type inequality involving a kernel which has a more general form than those known in the literature. We obtain necessary and sufficient conditions for the boundedness and compactness of a matrix operator from the weighted $l_{p,v}$ space into the weighted $l_{q,u}$ space defined by $(Af)_j := \sum_{i=j}^{\infty} a_{i,j} f_i$, for all $f = \{f_i\}_{i=1}^{\infty} \in l_{p,v}$ in case $1 < p \leq q < \infty$ and $a_{i,j} \geq 0$. Then we deduce a corresponding dual statement.

Wednesday 30 November 2011

What does "Inverse Problems" mean ?

Giulia DEOLMI (Padova, Dip. Mat.)

Inverse Problems (IP) are described as situations where "the answer is known, but not the question, or where the results, or consequences are known, but not the cause" (Isakov, 2006). To better understand what this means, after an introductory overview, the talk will focus on two particular IP. The first one consists in the estimation of the quantity of pollutant released in a river, while the second one is about the estimation of the corrosion of an unobservable face of a metal slab. Both of them will be solved in a discrete context, using an adaptive parametrization.

Wednesday 14 December 2011

Regular biproduct decompositions of objects

NICOLA GIRARDI (Padova, Dip. Mat.)

Every vector space over a field K is the direct sum of a number of copies of the one-dimensional K-vector space K. Allowing scalars to be elements of a ring R instead of a field, we obtain a more general object called right (or left, depending on which side we write the scalars) R-module. Contrary to the trivial case of K-vector spaces, modules over R may or may not decompose into indecomposable submodules, and when they do, it is interesting to know whether their decompositions are unique in some sense or at least satisfy some sort of constraint. Beginning with the basics and with the classical results of the field we will end up giving some examples where modules have decompositions that satisfy a nice combinatorial condition. As a last step, we hint to a generalisation to the setting of biproduct decompositions in preadditive categories.

Wednesday 18 January 2012

Understanding Defaults

JUAN MIGUEL A. MONTES (Padova, Dip. Mat.)

The possibility that a debtor may default poses a big risk to investors. Such a risk, called credit risk, is one of the risks present in financial markets. It is traditionally modelled in isolation from other kinds of risk such as the risk due to volatility, known as equity risk. In fact equity risk is observed to be connected to credit risk. In this talk, we provide a discussion of the basic credit risk models and the pricing of credit risky derivatives. We also discuss the Fourier Transform approach to pricing, developed by Carr and Madan. Using the Fourier Transform approach, we can price options under a model of risky assets proposed by Fontana, that treats credit risk and equity risk jointly. Finally with time permitting we discuss the application of this approach to a defaultable Heston model.

Wednesday 1 February 2012

The Fermi Pasta Ulam Paradox: An Introduction to the problem and its recent developments

HELENI CHRISTODOULIDI (Padova, Dip. Mat.)

In 1950's Fermi, motivated by fundamental questions of statistical mechanics and by ergodic properties of nonlinear dynamical systems, started a numerical experiment on a nonlinear toy model,

in collaboration with Pasta and Ulam. The system was chosen to be a one dimensional chain of N nonlinear coupled oscillators, described by a quadratic potential of nearby particle interactions plus a cubic perturbation. Fermi's ergodic hypothesis states that a system under an arbitrarily small perturbing force becomes generically ergodic. Giving initial conditions to the so called FPU system on the longest wavelength normal modes, the system showed a contradicting and integrable-like behaviour. Many pioneer works followed for the explanation of this paradox, more prominent of them being the approximation of FPU dynamics in the thermodynamic limit by the integrable partial differential equation Korteweg-de Vries by Zabusky and Kruskal (1965), and the work of Ferguson et al. (1982), where the authors observed the vicinity of the integrable one dimensional Toda chain with the FPU model. Recent developments show a more complete picture of the problem and its explanation.

Wednesday 8 February 2012

Interpolation theorems for weighted spaces of smooth functions

Ademi OSPANOVA (University of Astana and Dip. Mat. Padova)

This talk is devoted to the description of the Petre interpolation spaces for pairs of multiweighted Sobolev spaces. Interpolation spaces for pairs of weighted Sobolev spaces are well known. This talk concerns interpolation spaces for pairs of Sobolev spaces with a more general class of weights. The problem of description of interpolation spaces for a pair function spaces is very important in the embedding and approximating theory. Descriptions of the above-mentioned interpolation spaces can be used in the research of multipliers for weighted spaces of differentiable functions and in the problems of description of traces, which are important in the analysis of boundary value problems of mathematical physics.

Wednesday 15 February 2012

Modeling and valuing make-up clauses in gas swing contracts

Enrico EDOLI (Padova, Dip. Mat.)

Europe is among the largest consumers of natural gas in the world, but has a very limited domestic production compared to its consumption. The excess demand is covered by massive natural gas imports from foreign producer countries like Russia and Algeria. Swing contracts, also known as take-or-pay, are long term supply contracts with a quite standard structure which, basically, permits flexibility of delivery. In the last 10 years, thanks to the worldwide energy liberalization process, the birth of competitive gas markets and the recent financial crisis, those kind of traditional long term swing contracts in Europe have been supplemented in a significant way by make-up clauses which allow postponing the withdrawal of gas to future years when it could be more profitable. This introduces more complexity in the pricing and optimal management of swing contracts.

This talk is devoted both to an introduction on swing contracts and to a proper quantitative modelization of make-up clause. More in detail, we succeed in building an algorithm, based on dynamic programming, to price and optimally manage a swing contract with make-up clause. After having described the dynamic programming equation for problem, we prove that this problem has a quadratic complexity with respect to the number of years. Then, as an example, we show the algorithm at work on a 3-year contract and we present a sensitivity analysis of the price and of the make-up policy with respect to various parameters relative both to the price dynamics and to the swing contract.

Wednesday 22 February 2012

Singular limits of reaction-diffusion equations and propagation of interfaces

Cecilia DE ZAN (Padova, Dip. Mat.)

Interfacial phenomena are commonplace in physics, chemistry, biology. They occur, for example, whenever a continuum that can exists in at least two different chemical or physical "states" is present, and there is some mechanism that generates or enforces a spatial separation between these states. The separation boundary is called an interface. In mathematics, interfaces appear in the study of the asymptotic limits of evolving systems, like reaction-diffusion equations. After a simple introduction about the connections between reaction-diffusion equations and the wavefronts they generates, we present some mathematical approaches to the study of evolving interfaces. We present the classical level set-approach and a geometrical approach introduced by Barles and Souganidis in 1998. Then we show how this second approach can be applied in the study of the asymptotic limits of reaction-diffusion equations. Finally we show a simple generalization we obtained for nonlinear and possibly degenerate diffusions.

Wednesday 7 March 2012

Smooth Asymptotics for a DIC option in a Binomial Tree Model

Jose Maria ESCANER IV (University of the Philippines Diliman)

The talk aims to compute for the smooth asymptotic expansion for a down-and-in call (DIC) barrier option that was modeled using the Cox-Ross-Rubenstein (CRR) binomial tree. For pricing option contracts, the most well-known model used both by practitioners and in the academe is the Black- Scholes continuous time model. Though less accurate, a simpler and easier understood way to model financial derivatives though would be to use discrete time models. Among the many different discrete time models, a simple and widely-used model is the CRR binomial tree model. It is well-known that the price of the Black-Scholes continuous time model is close to the price obtained with the CRR binomial tree model when the number of time steps is large, as the Black-Scholes price is the limit of the tree model price. As such, it is of interest to measure the

convergence of the CRR model using asymptotic expansion. We follow the framework used by Diener and Diener in measuring the asymptotic expansion for the convergence of barrier options. For the purpose of finding the asymptotics, we make use of Andres symmetry principle in order to find the exact pricing formula of the DIC barrier option. By the guidelines set by Joshi, we specify the parameters and define our CRR binomial tree in such a way as to make the pricing formula symmetric. This would allow us to formulate a complete and smooth asymptotic expansion for our DIC barrier option.

Wednesday 21 March 2012

Isoperimetric Inequality: an invitation

Francescopaolo MONTEFALCONE (Padova, Dip. Mat.)

In this talk, I will survey some classical topics concerning isoperimetric inequalities in the Euclidean space. In particular, I will discuss the main ideas behind an isoperimetric inequality for the case of compact smooth hypersurfaces, proved in the Seventies by Michael and Simon. Finally, in the last part of the talk I will quickly state a recent generalization of this inequality to the sub-Riemannian setting of Carnot groups.

Thursday 5 April 2012

Moment Problems and Spin Correlation Matrices

Neeraja SAHASRABUDHE (Padova, Dip. Mat.)

Moment problems are about realizability of a given pair correlation function or covariances (or higher moments), namely whether a probability distribution is determined by its moments. That is, given $m_0, m_1, ... \in \mathbb{R}$, one wants to find a probability measure μ such that $\int_{-\infty}^{+\infty} x^k d\mu(x) = m_k$ for $k = 0, 1, ...$ If the probability measure is determined uniquely by the given set of moments the problem is called a determinate moment problem. Generalized moment problems of this kind have been widely studied, mainly in the theoretical Engineering community, for continuous random variables (particularly in case of point processes).
?In this talk, I'll discuss moment problems in general, giving examples of some determinate and indeterminate moment problems. We will also look at some specific moment problems and the necessary and sufficient conditions for realizability. I'll also talk about my work which is about a moment problem on a system of spin random variables. I will discuss about the necessary and sufficient conditions for a correlation matrix of order $n \geq 2$ to be the correlation matrix of spin variables in the classical sense and finally try to give an algorithm to explicitly compute the probability measure that realizes the given correlations.

Wednesday 18 April 2012

A glimpse of categorical algebra

Isar GOYVAERTS (Vrije Universiteit Brussel)

Category theory occupies a central position in contemporary mathematics and theoretical computer science, and occurs as well in other areas such as mathematical physics and linguistics. Roughly speaking, it provides a general abstract theory of structures and of systems of structures. The study of categories is an attempt to axiomatically capture what is commonly found in various classes of related mathematical structures. A way of thinking about category theory is that it is a refinement (or "categorification") of ordinary algebra.
In this talk, we will focus on the notion of a monoidal category, i.e. the categorification of the notion of a monoid. Monoidal categories have numerous applications outside of category theory itself and provide a tool to connect a priori quite remote branches of mathematics. We start the talk with some basic notions and definitions and provide some examples. Next, we try to sketch the idea of doing "categorical algebra" by considering a concrete example in more detail. We finish the talk with mentioning some recent results and some of the type of problems we are interested in. No specific knowledge of category theory is required. The only prerequisites is some notion of the basic definitions in the theory of certain algebraic structures (such as monoids, groups and vector spaces).

Wednesday 2 May 2012

Numerical Solution of Richards' Equation

Mohammad Sayful ISLAM (Padova, Dip. Mat.)

As hydrological models become increasingly sophisticated (e.g., coupling with various meteorological, ecological, or biogeochemical components) and are applied in ever more computationally demanding contexts (e.g., the many realizations that are typically generated in parameter estimation, uncertainty analysis, data assimilation, or scenario studies), the need for robust, accurate, and efficient codes is greater than ever. The Richards equation for subsurface flow is highly nonlinear and requires iterative schemes for its solution. These schemes have been the subject of much research over the past two decades, but an effective all-purpose algorithm has thus far proven elusive. Ideally, rapid (quadratic as opposed to linear) and global (insensitive to initial guess) convergence is sought, in addition to applicability over a range of conditions (dry soils, storm-interstorm simulations, geological heterogeneity, 3D domains with complex boundary conditions, etc). Richards' equation can be mathematically formulated and numerically discretized in a variety of manners, and the specific form and scheme chosen will affect the mass balance behavior of the model. Using a mass conservative mixed formulation of Richards equation we implemented and tested a nested Newton-type algorithm (originally developed by Casulli and Zanolli, 2010) for solving Richards equation. Experimental results include a variable boundary condition 1D test case and a few multidimensional heterogeneous problems. The results show that judicious choice of the initial guess together with time-step adaptation ensure quadratic convergence for all tested flow regimes. We

will discuss future challenges and implications in the context of modern hydrological simulators for real world applications.

Wednesday 16 May 2012

Harmonic functions in a domain with a small hole

Paolo MUSOLINO (Padova, Dip. Mat. and IRMAR, Rennes)

The asymptotic behaviour of the solutions of boundary value problems in domains with small holes has been largely investigated by many authors with different approaches. In this seminar, we consider a Dirichlet problem for the Laplace operator in a bounded domain Ω^o of $\mathbb{R}^n$ ($n \geq 3$) containing the origin, where we remove a small set whose size is determined by a parameter ϵ and which collapses to 0 for $\epsilon = 0$. Then for $\epsilon \neq 0$ we denote the solution to such a problem by u_ϵ. If $p \in \Omega^o$ and $p \neq 0$, then it makes sense to consider for $\epsilon \neq 0$ and 'small' the value of the solution u_ϵ at the point p. It is natural to ask what can be said on the map which takes ϵ small and positive to $u_\epsilon(p)$ around the degenerate value $\epsilon = 0$. One can try to answer to this question in several ways. By the approach proposed by Lanza de Cristoforis, one can show that there exist $\epsilon_p > 0$ and a real analytic function U_p from $]-\epsilon_p, \epsilon_p[$ to $\mathbb{R}$ such that $u_\epsilon(p) = U_p[\epsilon]$ for all $\epsilon \in]0, \epsilon_p[$. Such an equality holds for ϵ positive, and thus one can ask what happens to it when ϵ is negative. After an introductory part, we answer to this question and we show that the answer depends on the parity of the dimension n. (Based on joint work with M. Dalla Riva, Universidade de Aveiro.)

Wednesday 23 May 2012

PAINT: PAreto front INTerpolation for solving wastewater treatment problems

Markus HARTIKAINEN (Jyvaskyla University, Finland)

In many real life problems, one seeks for an optimal solution in accordance with several criteria, which can be conflicting (for example, you want a fast car but you don't want to spend too much). These situations are approached mathematically by the theory of multiobjective optimization. In the solution of multiobjective optimization problems, interactive methods have been proven effective. In such methods, an iterative procedure is used to improve the solutions available to satisfy the preferences of a decision maker. Because of the lack of closed form equations for many real-life quantities to be optimized, simulation may have to be used to compute the values of objectives for various values of the decision variables. As a consequence, multiobjective optimization problems may be computationally expensive. We have developed a method called PAINT, which constructs an interpolation between a set of given Pareto optimal solutions. In this talk, we demonstrate how the PAINT method works and can be applied to a multiobjective wastewater treatment planning problem.

Wednesday 20 June 2012

Coherent states approximation and the semiclassical limit

Simone VAZZOLER (Padova, Dip. Mat.)

The problem of understanding the relation between Classical and Quantum Mechanics has a long history which dates back to Schrödinger in 1926. In this talk, after an introductory part on Classical and Quantum Mechanics, we start giving the definition of Coherent State and consider the problem of approximating the exact solution to the Schrodinger equation using this particular class of wave functions. In the last part we discuss about the error term of this approximation and introduce the so called "Ehrenfest time".

A short introduction to Hopf algebras

Agustín García Iglesias (*)

Abstract. We present the definition of Hopf algebras, together with several examples. We give some of the main properties of these objects, as well as a list of classification results.

1 Introduction

Hopf algebras, introduced in the 50's, have been studied in the first place in their relationship with algebraic groups, where they haven been particularly useful to describe phenomena in the case of positive characteristic. Soon after that, they began to be studied because of their intrinsic interest, see for example [25], and have attracted the interest of mathematicians from many areas since the introduction of Quantum Groups in 1986 by Drinfeld.

In particular, a great interest is in the classification of Hopf algebras. A special class of these objects, namely the class of pointed Hopf algebras, has been intensively studied with this aim.

In this short note, we give the definition of Hopf algebras, together with many examples. We present some of their most remarkable properties and comment on the state of the classification problem, in several of its directions.

2 Preliminaries

We fix $\Bbbk$ a field. If X is a set, we denote by $k\{x\,/x \in X\}$ the $\Bbbk$-vector space with a basis indexed by the elements of X. We denote by $\Bbbk\langle X\rangle$ the free algebra with the elements of X as generators. If $f_1, \dots, f_m \in k\langle X\rangle$, $m \in \mathbb{N}$, we denote by $\langle f_1, \dots, f_m\rangle$ the ideal generated by the f_i's and by $\Bbbk\langle X|f_1, \dots, f_m\rangle$ the quotient algebra $\Bbbk\langle X\rangle/\langle f_1, \dots, f_m\rangle$. All unadorned tensor products are assumed to be over the field $\Bbbk$.

We refer the reader to [25], [22], [26] for more information about Hopf algebras. Also see the rest of the bibliography at the end of the notes and the references therein.

(*)FaMAF-CIEM (CONICET), Universidad Nacional de Córdoba, Medina Allende s/n, Ciudad Universitaria, 5000 Córdoba, República Argentina. E-mail: aigarcia@famaf.unc.edu.ar . Seminar held on October 19th, 2011.

3 Definition and examples

The definition of a Hopf algebra involves the notion of several algebraic structures that we gradually introduce next.

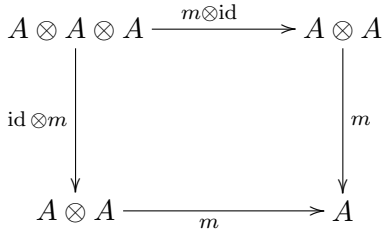
3.1 Algebras

We start by recalling, in a particular and convenient way, the definition of a $\Bbbk$-algebra.

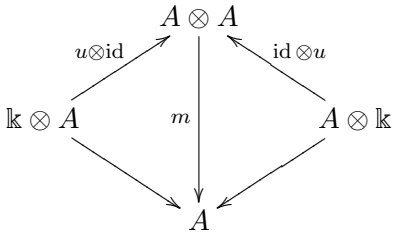
Definition 3.1 An associative algebra with unit is a vector space A with a bi-linear multiplication $m : A \times A \to A$, $m(a, b) := ab$, $a, b \in A$, and a distinguished element $1 = 1_A$, such that

$$a(bc) = (ab)c \quad \text{and} \quad 1a = a = a1 \qquad\qquad a, b, c \in A.$$

These axioms can be encoded in the commutativity of the following diagrams:



Associativity



Unit

where $u : k \to A$ is a ring homomorphism whose image is in the center of A. In this way, we get $1_A = u(1_\Bbbk)$.

From now on we will always consider associative algebras with 1, and we will just refer to this structures as algebras.

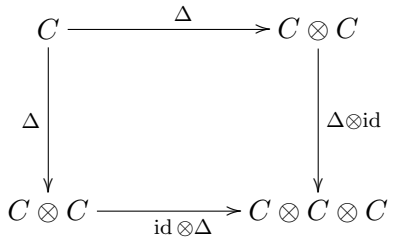
Example 3.2 Let Γ be a group, e its unit element. Let A be the vector space with basis $\{f_g, g \in \Gamma\}$. Then A is an algebra with multiplication $f_g f_h = f_{gh}$ and unit $1_A = f_e$ (notice that it is enough to define the multiplication on a basis, since it is a linear map). We denote this algebra by $\Bbbk\Gamma$.

3.2 Coalgebras

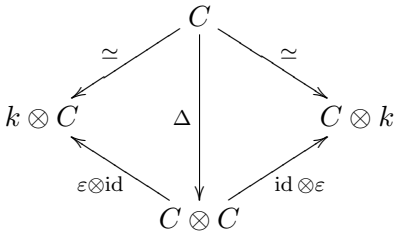
Dualizing the notion of an algebra, we arrive to the definition of a (coassociative, counital) coalgebra.

Definition 3.3 A *coalgebra* C over $\Bbbk$ is a $\Bbbk$-vector space together with linear morphisms

$\Delta : C \to C \otimes C$ and $\varepsilon : C \to \Bbbk$ such that the following diagrams commute:



Coassociativity



Counit

The map Δ is called the *comultiplication* and ε is called the *counit*. The axioms read as follows

$$(\Delta \otimes \mathrm{id})\Delta = (\mathrm{id} \otimes \Delta)\Delta \tag{1}$$

$$(\varepsilon \otimes \mathrm{id})\Delta = (\mathrm{id} \otimes \varepsilon)\Delta = \mathrm{id}\,. \tag{2}$$

Example 3.4 Let X be a set and consider the vector space $\Bbbk X = \Bbbk\{f_x, x \in X\}$. Let $\Delta : \Bbbk X \to \Bbbk X \otimes \Bbbk X$, $\varepsilon : \Bbbk X \to \Bbbk$ be the linear extensions of the maps $f_x \mapsto f_x \otimes f_x$ and $f_x \mapsto 1$. Then $\Bbbk X$ is a coalgebra with comultiplication Δ and counit ε. In fact, it is straightforward (and enough) to check the axioms in the basis:

$$(\Delta \otimes \mathrm{id})\Delta(f_x) = \Delta(f_x) \otimes f_x = f_x \otimes f_x \otimes f_x = f_x \otimes \Delta(f_x) = (\mathrm{id} \otimes \Delta)\Delta(f_x),$$
$$(\varepsilon \otimes \mathrm{id})\Delta(f_x) = \varepsilon(f_x)f_x = f_x = f_x\varepsilon(f_x) = (\mathrm{id} \otimes \varepsilon)\Delta(f_x),$$

for every $x \in X$.

Example 3.5 Let A be a finite dimensional algebra with multiplication m and unit map u. Then, identifying the spaces $(A \otimes A)^*$ and $A^* \otimes A^*$ the dual morphisms $m^* : A^* \to A^* \otimes A^*$ and $u^* : A^* \to \Bbbk$, given by $m^*(f)(a \otimes b) = f(ab)$ and $u^*(f) = f(1)$ give a structure of coalgebra to the vector space A^*. This follows from the fact that the axioms defining a coalgebra are precisely the dualization of the corresponding ones defining an algebra. In fact,

$$(m^* \otimes \mathrm{id})m^* = (m(m \otimes \mathrm{id}))^* = (m(\mathrm{id} \otimes m))^* = (\mathrm{id} \otimes m^*)m^*$$
$$(u^* \otimes \mathrm{id})m^* = (m(u \otimes \mathrm{id}))^* = \mathrm{id} = (m(\mathrm{id} \otimes u))^* = (\mathrm{id} \otimes u^*)m^*.$$

Example 3.6 In the context of the example above, let Γ be a finite group and consider the algebra $A = \Bbbk\Gamma$, presented in Example 3.2. Then A^* is identified with the function algebra $\Bbbk^\Gamma = \{f : \Gamma \to \Bbbk\}$, which has a basis $\{\delta_g\}_{g \in \Gamma}$ given by $\delta_g(f_h) = \delta_{g,h}$, $g, h \in \Gamma$. The

multiplication on this basis is $\delta_g\delta_h = \delta_{g,h}\delta_g$ and the unit is $1 = \sum_{g\in\Gamma}\delta_g$. Let Δ, ε be the duals of the multiplication and the unit map. Then

$$\begin{aligned} &\Delta(\delta_g(f_h \otimes f_t)) = \delta_g(f_{ht}) = \delta_{g,ht}, && g,h,t \in \Gamma,\\ &\varepsilon(\delta_g) = \delta_g(f_e) = \delta_{g,e}, && g \in \Gamma. \end{aligned}$$

In particular, we can write in this case an explicit formula for Δ:

$$\Delta(\delta_g) = \sum_{h\in\Gamma} \delta_g \otimes \delta_{g^{-1}h}.$$

Definition 3.7 A coalgebra C is said to be *cocommutative* if $\Delta = \tau \circ \Delta$, where $\tau : C \otimes C \to C \otimes C$ is the flip map: $x \otimes y \mapsto y \otimes x$.

Example 3.8 The coalgebra $\Bbbk X$ from Example 3.4 is cocommutative, while the function coalgebra k^Γ from Example 3.6 is not, whenever Γ is not commutative.

3.3 Bialgebras

We combine the concepts of algebra and coalgebra to reach to the definition of a bialgebra.

Definition 3.9 A *bialgebra* B is a $\Bbbk$-vector space with structures of algebra (B, m, u) and coalgebra (B, Δ, ε) such that $\Delta : B \to B \otimes B$ and $\varepsilon : B \to \Bbbk$ are algebra maps. Here, we consider $B \otimes B$ as an algebra with multiplication component-wise and the underlying algebra structure of the field $\Bbbk$.

Remark 3.10 It is equivalent to require $m : B \otimes B \to B$ and $u : k \to B$ to be coalgebra maps.

Example 3.11 Let B be the group algebra $\Bbbk\Gamma$ from Example 3.2 and consider the coalgebra structure on the vector space $\Bbbk\{f_g, g \in \Gamma\}$ as in Example 3.4. Then B is a bialgebra. In fact,

$$\begin{aligned} &\Delta(f_g f_h) = \Delta(f_{gh}) = f_{gh} \otimes f_{gh} = (f_g \otimes f_g)(f_h \otimes f_h) = \Delta(f_g)\Delta(f_h)\\ &\varepsilon(f_g f_h) = \varepsilon(f_{gh}) = 1 = \varepsilon(f_g)\varepsilon(f_h). \end{aligned}$$

The function algebra k^Γ is an algebra with pointwise multiplication, that is $\delta_g\delta_h = \delta_{g,h}\delta_g$, as seen in Example 3.4. This multiplication, together with the coalgebra structure introduced in that example, makes $\Bbbk^\Gamma$ a bialgebra.

Remark 3.12 If B is a bialgebra with multiplication m, unit u, comultiplication Δ and counit ε, then the vector space $A = \mathrm{Hom}_\Bbbk(B, B)$ becomes an algebra, with product given by the convolution $*$ and unit $u\varepsilon$. That is, if $f, g \in \mathrm{Hom}_\Bbbk(B, B)$, then $f * g \in \mathrm{Hom}_\Bbbk(B, B)$ via $(f * g)(a) = m(f \otimes g)\Delta(a)$.

3.4 Hopf algebras

Now we have all the elements needed to introduce the definition of a Hopf algebra.

Definition 3.13 A *Hopf algebra* is a $\Bbbk$-vector space H together with a linear map $\mathcal{S} : H \to H$ such that

$$m(\mathcal{S} \otimes \mathrm{id})\Delta = u\varepsilon = m(\mathrm{id} \otimes \mathcal{S})\Delta. \tag{3}$$

The map $\mathcal{S}$ is referred to as the *antipode*.

Remark 3.14 The map $\mathcal{S} : H \to H$ is an anti-algebra map, i.e. $\mathcal{S}(ab) = \mathcal{S}(b)\mathcal{S}(a)$ for every $a, b \in H$, see [22, Proposition 1.5.10].

Lemma 3.15 *Let H be a Hopf algebra with antipode $\mathcal{S}$ and consider the algebra $A = \mathrm{End}_{\Bbbk}(H, H)$, with the convolution product. Then $\mathcal{S}$ is the inverse of the identity map* id. *In particular, it is unique.*

Proof. Recall the the unit of the algebra A is given by the map $u\varepsilon$. Then the lemma follows from (3). □

Example 3.16 The group algebra $\Bbbk\Gamma$ is a Hopf algebra, with $\mathcal{S}(f_g) = f_{g^{-1}}$. Also, the function algebra $\Bbbk^{\Gamma}$ is a Hopf algebra, with $\mathcal{S}(\delta_g) = \delta_{g^{-1}}$.

Example 3.17 Let H be a finite-dimensional Hopf algebra, with algebra structure given by a multiplication m and a unit u, coalgebra structure given by maps Δ and ε and antipode $\mathcal{S}$. Then H^* is a Hopf algebra, with multiplication Δ^*, unit ε^*, comultiplication m^*, counit u^* and antipode given by $\mathcal{S}^*$.

Example 3.18 Let q be an Nth root of unity, $N \geq 2$. Then the algebra $T_q = \Bbbk\langle x, g | x^N = 0, g^N = 1, gx = q\, xg\rangle$ is a Hopf algebra, with comultiplication

$$\Delta(g) = g \otimes g, \qquad \Delta(x) = x \otimes 1 + g \otimes x$$

and antipode $\mathcal{S}(g) = g^{-1}$ and $\mathcal{S}(x) = -xg^{-1}$. This is the *Taft algebra*, introduced by Taft in 1971.

Example 3.19 Let $\mathfrak{g}$ be a Lie algebra and let $\mathcal{U}(g)$ be the universal enveloping algebra of $\mathfrak{g}$. Then $\mathcal{U}(\mathfrak{g})$ is a Hopf algebra with comultiplication and antipode defined on the set of generators $\mathfrak{g}$ by $\Delta(x) = x \otimes 1 + 1 \otimes x$ and $\mathcal{S}(x) = -x$, $x \in \mathfrak{g}$.

Example 3.20 Let $\mathfrak{g}$ be a Lie algebra and let Γ be a group acting by Lie algebra automorphisms on $\mathfrak{g}$. Then it is possible to consider the semidirect product $\mathcal{U}(\mathfrak{g})\#\Bbbk\Gamma$. This a Hopf algebra, with

$$\Delta(f_g) = f_g \otimes f_g, \qquad \Delta(x) = x \otimes 1 + 1 \otimes x,$$

for every $g \in \Gamma$, $x \in \mathfrak{g}$.

4 Invariants

When dealing with classification problems, it is important to determine invariants of the objects of study, that is to say properties or structures that are not altered by isomorphisms. A first invariant like this is of course the dimension. Two more refined invariants are defined as follows.

Definition 4.1 Let H be a Hopf algebra. Then the space of *group-like elements* of H is the set

$$G(H) = \{h \in H \,:\, \Delta(h) = h \otimes h \text{ and } \varepsilon(h) \neq 0\}.$$

The space of *primitive elements* of H is the set

$$\mathcal{P}(H) = \{h \in H \,:\, \Delta(h) = h \otimes 1 + 1 \otimes h\}.$$

More generally, given $g, k \in G(H)$, the space of (g,k)-*skew primitive elements* is given by

$$\mathcal{P}_{g,k}(H) = \{h \in H \,:\, \Delta(h) = h \otimes g + k \otimes h\}.$$

In particular, $\mathcal{P}(H) = \mathcal{P}_{1,1}(H)$.

Proposition 4.2 *$G(H)$ is a group, with product given by the multiplication in H. On the other hand, the bracket given by the commutator in H makes $\mathcal{P}(H)$ a Lie algebra. Also, $G(H)$ acts on $\mathcal{P}(H)$ by conjugation in H. Moreover, this action is a Lie algebra automorphism.*

Proof. It is clear that if $g, h \in G(H)$ then $gh \in G(H)$ and that $1 \in G(H)$. Finally notice that, if $g \in G(H)$, we have that $\varepsilon(g) = 1$ by (2) and $\mathcal{S}(g) = g^{-1}$ by (3).

On the other hand, it is straightforward to check that if $x, y \in \mathcal{P}(H)$, then $[x,y] = xy - yx \in \mathcal{P}(H)$. Also, if $g \in G(H)$ and $z \in \mathcal{P}(H)$, then $g \cdot z := gzg^{-1} \in \mathcal{P}(H)$ and $g \cdot [x,y] = [g \cdot x, g \cdot y]$ for every $x, y \in \mathcal{P}(H)$. □

Examples 4.3 Let Γ be a group, $\mathfrak{g}$ a Lie algebra. We have $G(\Bbbk\Gamma) = \Gamma$, $\mathcal{P}(\Bbbk\Gamma) = \{0\}$, $G(\mathcal{U}(\mathfrak{g})) = \{1\}$, $\mathcal{P}(\mathcal{U}(\mathfrak{g})) = \mathfrak{g}$. Let $N \in \mathbb{N}$, q an Nth root of unity. Then $G(T_q) = \langle g \rangle \cong \mathbb{Z}/N\mathbb{Z}$, $\Bbbk\{x\} = \mathcal{P}_{1,g}$.

5 Some important theorems

Theorem 5.1 [24] *Let H be a finite-dimensional Hopf algebra, and let $R \subseteq H$ be a Hopf subalgebra. Then H is a free R-module.*

As a consequence of the above theorem, we have the following "Lagrange's Theorem for Hopf algebras".

Corollary 5.2 *If $R \subseteq H$ are finite-dimensional Hopf algebras, then the dimension of R divides the dimension of H.*

Definition 5.3 A Hopf algebra is said to be *semisimple* if it is semisimple as an algebra. It is said to be *cosemisimple* if H^* is semisimple.

There is a rather direct way to check the (co)semisimplicity of a Hopf algebra. The following theorem is due to Larson and Radford, see [25] for a complete proof.

Theorem 5.4 *Assume that the characteristic of $\Bbbk$ is zero and let H be a finite-dimensional Hopf algebra over $\Bbbk$. Then H is semisimple if and only if H^* is semisimple if and only if $\mathcal{S}^2 = \mathrm{id}$.*

Remark 5.5 Let Γ be a group, $\Bbbk$ a field of characteristic zero. Then $\Bbbk\Gamma$ and $\Bbbk^\Gamma$ are semisimple Hopf algebras. On the other hand, the Taft algebra T_q from Example 3.18 is not semisimple, in fact

$$\mathcal{S}^2(x) = -\mathcal{S}(xg^{-1}) = gxg^{-1} = qx.$$

6 Classification results

6.1 By dimension

Assume $\Bbbk$ is an algebraically closed field of characteristic zero.

Theorem 6.1 [28] *Let H be a Hopf algebra over $\Bbbk$ of prime dimension p. Then H is isomorphic to the Hopf group algebra of the cyclic group $\mathbb{Z}/p\mathbb{Z}$.*

Theorem 6.2 [7, 21, 23] *Let H be a Hopf algebra over $\Bbbk$ of dimension p^2, p a prime number. Then H is isomorphic to one and only one of the following: the groups algebras $\Bbbk\mathbb{Z}/p^2\mathbb{Z}$ or $\Bbbk(\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z})$ or the Taft algebras T_q, with $q^p = 1$, $q \neq \pm 1$.*

Remark 6.3 For $\dim H = p^3$ there are complete results for $p = 2$ [27] and $p = 3$ [14]. For $\dim H = p^4$ the list is complete only for $p = 2$ in [19]. There are also classification results for dimensions pq, pq^2, pqr, for p, q, r different primes. There are not many further results. In fact, the classification is not known even for dimension equal to 20.

6.2 Cocommutive & commutative Hopf algebras

Definition 6.4 A Hopf algebra H is said to be *cocommutative* if the underlying coalgebra is. Analogously, H is *commutative* if it is commutative as an algebra.

Cocommutive Hopf algebras are completely determined by the next theorem, due to Cartier and Kostant.

Theorem 6.5 *Let H be a cocommutative Hopf algebra over a field of characteristic zero $\Bbbk$. Then there exists a Lie algebra $\mathfrak{g}$ and a group Γ, acting by Lie automorphisms in $\mathfrak{g}$, such that $H \cong \mathcal{U}(\mathfrak{g})\#\Bbbk\Gamma$ as Hopf algebras. Moreover, $\mathfrak{g} = \mathcal{P}(H)$ and $\Gamma = G(H)$.*

Corollary 6.6 *Let H be a finite-dimensional cocommutative Hopf algebra over $\Bbbk$. Then $H \cong \Bbbk\Gamma$, for some group Γ.*

Proof. $\mathcal{U}(\mathfrak{g})$ is infinite-dimensional, for $\mathfrak{g}$ such that $\dim \mathfrak{g} > 1$ and $\mathcal{U}(\Bbbk) = \Bbbk$. □

Corollary 6.7 *Let H be a finite-dimensional commutative Hopf algebra over $\Bbbk$. Then $H \cong \Bbbk^\Gamma$, for some group Γ.*

Proof. H^* is cocommutative. □

6.3 Other classes: Pointed and Copointed Hopf algebras

Definition 6.8 Let H be a Hopf algebra. The *coradical* H_0 of H is the largest cosemisimple subcoalgebra of H. That is, $H_0 = \sum_{C \in S} C$, for S the set $\{C \subset H \text{ simple coalgebra}\}$.

There are many classification results for non-semisimple Hopf algebras according to the structure of H_0, particularly when H_0 is a Hopf subalgebra of H.

Remark 6.9 The group algebra $\Bbbk\Gamma$, for $\Gamma = G(H)$ is always contained in H_0 since it is a cosemisimple subcoalgebra of H.

6.3.1 Pointed Hopf algebras

Definition 6.10 A Hopf algebra is called *pointed* if H_0 coincides with the group algebra $\Bbbk G(H)$.

Example 6.11 The Hopf algebras $\Bbbk\Gamma$, $\mathcal{U}(\mathfrak{g})\#\Bbbk\Gamma$, T_q are pointed Hopf algebras.

Remark 6.12 Set $\Bbbk = \mathbb{C}$. Let H be a finite dimensional Hopf algebra, $\Gamma = G(H)$. A complete classification has been achieved in [8] in the case in which Γ is abelian and its order is coprime with 210 (i.e. not divisible by small primes). Important related results from [11, 12, 13] have prepared the ground for a complete classification [9].

When Γ is not abelian, a complete classification is known for the symmetric groups $\mathbb{S}_3$ and $\mathbb{S}_4$ [6], [18], respectively and the dihedral groups D_{4m}, $m \in \mathbb{N}$ [16].

The classification is also finished for the alternating groups A_m, $m \geq 6$ [1] and some sporadic simple groups [17]. These families share the property that the only pointed Hopf algebra over Γ is just the group algebra $\Bbbk\Gamma$. This seems to be the case for many simple groups, see [2] and [3], also [15].

6.3.2 Copointed Hopf algebras

Definition 6.13 A Hopf algebra is called *copointed* if its coradical is isomorphic to the function algebra over a group Γ.

Remark 6.14 A complete classification for the copointed case has been achieved in [10] for the symmetric group $\mathbb{S}_3$. The case of the group $\mathbb{S}_4$, together with some other non-abelian groups, is being developed in [20].

References

[1] Andruskiewitsch, N., Fantino, F., Graña, M., and Vendramín, L., *Finite-dimensional pointed Hopf algebras with alternating groups are trivial*. Ann. Mat. Pura Appl. (4) Volume **190**, Number 2 (2011), 225–245.

[2] Andruskiewitsch, N., Fantino, F., Graña, M., and Vendramin, L., *On twisted homogeneous racks of type D*. Revista de la Unión Matemática Argentina vol. **51-2** (2010), 1–16.

[3] Andruskiewitsch, N., Fantino, F., Graña, M., and Vendramin, L., *On Nichols algebras associated to simple racks*. Groups, Algebras and Applications, Contemporary Mathematics, vol. **537**, Amer. Math. Soc., Providence, RI (2011), 31-56.

[4] Andruskiewitsch, N. and Ferrer Santos, W., *The beginnings of the theory of Hopf algebras*. Acta Appl. Math. **108** (2009), 3–17.

[5] Andruskiewitsch, N. and F., Graña, *From racks to pointed Hopf algebras*. Adv. in Math. **178** (2) (2003), 177–243.

[6] Andruskiewitsch, N., Heckenberger, I. and Schneider, H.J., *The Nichols algebra of a semisimple Yetter-Drinfeld module*. Amer. J. Math., vol. **132**, no. 6 (2010), 1493–1547.

[7] Andruskiewitsch, N. and Schneider, H.J., *Hopf algebras of order p^2 and braided Hopf algebras of order p*. J. Algebra **199**, no. 2 (1998), 430–454.

[8] Andruskiewitsch, N. and Schneider, H.J., *On the classification of finite-dimensional pointed Hopf algebras*. Ann. Math. **171**, No. 1 (2010), 375–417.

[9] Andruskiewitsch, N., Angiono, I. and García Iglesias, A., *Pointed Hopf algebras with abelian group*. Work in progress.

[10] Andruskiewitsch, N. and Vay, C., *Finite dimensional Hopf algebras over the dual group algebra of the symmetric group in three letters*. Commun. Algebra, to appear.

[11] Angiono, I., *On Nichols algebras with standard braiding*. Algebra and Number Theory Vol. 3, No. **1** (2009), 35–106.

[12] Angiono, I., *A presentation by generators and relations of Nichols algebras of diagonal type and convex orders on root systems*. Submitted. Preprint: `arXiv:1008.4144`.

[13] Angiono, I., *On Nichols algebras of diagonal type*. To appear in J. Reine Angew. Math.; `arXiv:1104.0268`.

[14] Beattie, M. and García , G. A., *Techniques for classifying Hopf algebras and applications to dimension p^3*. Preprint: `arXiv:1108.6037v1`.

[15] Carnovale, G. and García Iglesias, A., *Twisted conjugacy classes of type D in finite groups of Lie type*. Work in progress..

[16] Fantino, F. and García, G. A., *On pointed Hopf algebras over dihedral groups*. Preprint: `arXiv:1007.0227v1`.

[17] Fantino, F. and Vendramín, L., *On twisted conjugacy classes of type D in sporadic simple groups*. Submitted. Preprint: `arXiv:1107.0310`.

[18] García, G. A. and García Iglesias, A., *Finite-dimensional pointed Hopf algebras over $\mathbb{S}_4$*. Israel Journal of Mathematics **183** (2011), 417–444.

[19] García, G. A. and Vay, C., *Hopf algebras of dimension 16*. Algebr. Represent. Theory **13**, no. 4 (2010), 383–405.

[20] García Iglesias, A. and Vay, C., *Liftings of finite-dimensional Nichols algebras over non-abelian groups and their duals*. Work in progress.

[21] Masuoka, A., *The p^n theorem for semisimple Hopf algebra*. Proc. Amer. Math. Soc. **124**, no. 3 (1996), 735–737.

[22] Montgomery, S., "Hopf algebras and their action on rings". CBMS Lecture Notes **82**, American Math Society, Providence, RI, 1993.

[23] Ng, S.-H., *Non-semisimple Hopf algebras of dimension p^2*. J. Algebra **255**, no. 1 (2002), 182–197.

[24] Nichols, W.D. and Zoeller, M. B., *A Hopf algebra freeness theorem*. American Journal of Math **111** (1989), no. 1, 381–385.

[25] Schneider, H.-J., "Lectures on Hopf algebras". Universidad de Córdoba Trabajos de Matemática, Serie "B", No. 31/95, Córdoba, Argentina, 1995.

[26] Sweedler, M., "Hopf algebras". Benjamin, New York, 1969.

[27] Williams, R.e, "Finite dimensional Hopf algebras". Ph. D. Thesis, Florida State University.

[28] Zhu, Y., *Hopf algebras of prime dimensio*. Internat. Math. Res. Notices **1** (1994), 53–59.

The Explicit Laplace Transform for the Wishart process

Alessandro Gnoatto (*)

This paper is based on joint work with Martino Grasselli(†)

Abstract. We derive the explicit formula for the joint Laplace transform of the Wishart process and its time integral which extends the original approach of [Bru, 1]. We compare our methodology with the alternative results given by the variation of constants method, the linearization of the Matrix Riccati ODE's and the Runge-Kutta algorithm. The new formula turns out to be fast, accurate and very useful for applications when dealing with stochastic volatility and stochastic correlation modelling.

1 Introduction

Before we introduce the results of the paper, we provide a short introduction to Mathematical finance, which is a branch of mathematics trying to solve problems like:

- Pricing and hedging of financial instruments,
- Computation of optimal investment strategies,
- Risk management.

From an historical point of view, contributions coming from different fields have determined the evolution of Mathematical Finance.

- Mathematics (Probability, Stochastic Calculus, Complex Analysis, Geometry)
- Physics (Many PDE analogies)
- Financial economics (No arbitrage, completeness)

(*)Mathematisches Institut der LMU, München and Dipartimento di Matematica Pura ed Applicata, Padova (Italy). E-mail: gnoatto@math.lmu.de . Seminar held on November 2nd, 2011.

(†) Dipartimento di Matematica Pura ed Applicata, Padova (Italy) and ESILV Ecole Supérieure d'Ingénieurs Léonard de Vinci, Département Mathématiques et Ingénierie Financière, Paris La Défense (France). E-mail: grassell@math.unipd.it .

- Econometrics/Statistics

This emphasizes the fact that Mathematical Finance is a subject where different paradigms coexist. Physics in particular has played an important role in the early age, around the seventies/eighties. The universally recognized birth of Mathematical Finance coincides with the appearence of the PhD thesis of L. Bachelier, in 1900. In his PhD Thesis Bachelier proposed the first financial model, where the complexity of financial markets is captured by a stochastic process (Brownian motion). Unfortunately, the importance of his work was not recognized for many years. It was around the sixties that P. Samuelson redescovered the work of Bachelier and started to use it as a basis for financial modelling. Samuelson introduced a new kind of stochastic process, called Geometric Brownian Motion, which constitutes the standard example for a process describing asset prices. In the mean time, another important stream of research, focused on the search for optimal allocations among risky investment opportunities, was introduced by H. Markowitz in his PhD thesis. Finally, in 1973, F. Black and M. Scholes published their famous paper on option pricing which was given full recognition in 1997 with the Nobel prize.

2 Motivating example

2.1 Introduction to Option pricing

We will introduce Mathematical Finance by illustrating one of the most important problems in financial engineering: the pricing of derivatives securities.

Let $S = (S_t)_{0 \le t \le T}$ be the price of a stock, say Unicredit. Suppose you bought at time $t = 0$ a Unicredit stock at a price $S_0 = 0.80$. The payoff, i.e. what you earn or lose at a future time is simply given by:

$$p\&l_T = S_T - S_0 \tag{2.1}$$

which is a trivial linear function of S_T.

- Of course, the value of S_T is unknown at time $t = 0$.
- We model this lack of knowledge with randomness, so S_T is a random variable.
- Is it possible to annihilate the risk arising from the fluctuations of S?

Let us write $K = S_0$. The profit and loss may be decomposed as:

$$S_T - K = \max\{S_T - K, 0\} - \max\{K - S_T, 0\} \tag{2.2}$$

What we have introduced are European options. More precisely:

- A call option gives to its holder the right, but not the obligation, to buy a stock at T at a fixed price (exercise or strike price) K which is determined at t.
- A put option gives to its holder the right, but not the obligation, to sell a stock at T at a fixed price (exercise or strike price) K which is determined at t.

These are the simplest examples of derivatives. We would like to compute the fair price at time 0 of these products.

If you asked an economist about the price of such an instrument, he/she may tell you that prices are determined by demand and supply. However, financial markets are driven by many factors, which are not easy to describe theoretically and estimate empirically. The perspective of Mathematical Finance is different. The stock price is assumed to follow a path, which is similar to that of a particle of pollen, when it is hit by many moleculae of water. This kind of wild trajectories are described mathematically by a stochastic process which is known as Brownian Motion.

2.2 The Black-Scholes model

The assumptions behind the model that we will employ to price the derivative are the following.

- A filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$.
- $\mathcal{F}_t$ satisfies the usual assumptions.
- we consider two processes modeling two different financial assets
- a money-market account driven by the deterministic ODE

$$dB_t = rB_t dt \quad B_0 = 1 \tag{2.3}$$

- an SDE for the stock price

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t^{\mathbb{P}} \quad S_0 = s_0 \tag{2.4}$$

which admits a solution which is pathwise unique

$$S_T = S_0 \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma W_T\right\} \tag{2.5}$$

Under these assumptions, the evolution of a derivative will be a function of time and of the price of the stock. Given the SDE (2.4) we know how to compute the dynamics of $f(t, S_t)$ thanks to Ito's lemma:

$$df(t, S_t) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S_t}dS_t + \frac{\partial^2 f}{\partial S_t^2}d\langle S_t\rangle \tag{2.6}$$

the idea of Black Scholes is to construct a portfolio using the money-market account B and the stock S such that the random term that we have above vanishes. In this way we obtain a pricing PDE of the form:

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial S}rS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} - rf = 0 \tag{2.7}$$

$$f(T, S_T) = \max\{S_T - K, 0\} \tag{2.8}$$

We can solve the previous PDE (this was the approach in Black Scholes) or we can apply Feynman-Kac theorem, which tells us that the PDE is equivalent to an expectation computed w.r.t. an artificial probability measure, which is called Risk-neutral measure. Hence we write:

$$f(t, S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\max\left\{S_T - K, 0\right\}|\mathcal{F}_t\right] \tag{2.9}$$

In the case of this simple model it is not difficult to show that the solution is:

$$f(t, S_t) = S_t N(d_1) - K e^{-r(T-t)} N(d_2) \tag{2.10}$$

$$N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \tag{2.11}$$

$$d_1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \tag{2.12}$$

$$d_2 = d_1 - \sigma\sqrt{T-t} \tag{2.13}$$

We can proceed to make some observations.

- Notice that the terms $N(d_i)$ are probabilities, in particular $N(d_2) = \mathbb{Q}\left(S_T > K\right)$
- From (2.5) we realize that S_T is a lognormal random variable.
- Hence the distribution of the *log-asset price* $X_t := \log S_t$ is Gaussian
- So in this context we can easily write the conditional probability density of the log-asset price as

  $$\frac{1}{\sqrt{2\pi v}} e^{-\frac{1}{2}\frac{x-h}{v^2}} dx \tag{2.14}$$

  where:

  $$h = \log S_t + \left(r - \frac{\sigma^2}{2}\right)(T-t) \tag{2.15}$$

  $$v = \sigma^2(T-t) \tag{2.16}$$

Unfortunately, the Black and Scholes model is not realistic.

- The distribution of $X_T = \log S_T$ is not Gaussian and features asymmetry and fat tails
- If we observe a surface of prices for different values of K and the maturity T and then try to find the value for σ such that the Black Scholes price is equal to the market price then we do not observe a flat surface.
- Smiles and term structures of volatilities.
- Volatility is itself a stochastic process.

2.3 Characteristic Functions and more general models

Before we talk about more complex models, we would like to recall some well-known results which will be useful in the sequel. Let X be a random variable (X will be the log-price). Then its characteristic function is defined as:

$$\varphi(\theta) = \mathbb{E}\left[e^{i\theta X}\right] \tag{2.17}$$

Levy's inversion theorem gives a relationship between the characteristic function and the probability density of a random variable. If we know the characteristic function of a random variable, then we can recover its density:

Theorem 1 [Levy's inversion Theorem] *If* $\int_{\mathbb{R}} |\varphi(\theta)| d\theta < \infty$ *then* X *has a continuous probability density function* f*, and*

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\theta x} \varphi(\theta) d\theta \tag{2.18}$$

The CF for a Gaussian random variable with mean h and variance v^2 is given by:

$$\mathbb{E}\left[e^{i\theta X}\right] = e^{i\theta h - \frac{1}{2}\theta^2 v^2} \tag{2.19}$$

Using the relation between the CF and the density, we can rewrite the probabilities in the B&S formula, e.g.:

$$N(d_2) = \mathbb{Q}(S_T > K) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty Re\left[\frac{e^{-i\theta \log(K)\varphi(\theta)}}{i\theta}\right] d\theta \tag{2.20}$$

this means that if we know the characteristic function of $X_T = \log(S_T)$, then we can compute option prices, even for models which are more complicated w.r.t the standard Black Scholes case.

2.3.1 [Heston, 3]

$$\frac{dS_t}{S_t} = rdt + \sigma\sqrt{V_t} dW_t \tag{2.21}$$

$$dV_t = \kappa(\theta - V_t) + \eta\sqrt{V_t} dZ_t \tag{2.22}$$

2.3.2 [Da Fonseca et al., 2]

$$\frac{dS_t}{S_t} = rdt + Tr\left[\sqrt{X_t} dW_t\right] \tag{2.23}$$

$$dX_t = \left(\Omega\Omega^\top + MX_t + X_t M^\top\right) dt + \sqrt{X_t} dZ_t Q + Q^\top dZ_t^\top \sqrt{X_t} \tag{2.24}$$

The last model motivates the need for an explicit solution for the Laplace/Fourier transform of the process X, which is called Wishart process.

3 Our results on the Wishart process

We start from the following probabilistic setup.

- Filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$+usual assumptions.
- B matrix Brownian motion
- S_d^+ the cone of symmetric $d \times d$ matrices with scalar product given by the trace operator
- A Wishart process is governed by the following SDE

$$dS_t = \sqrt{S_t} dB_t Q + Q^\top dB_t^\top \sqrt{S_t} + \left(M S_t + S_t M^\top + \alpha Q^\top Q \right) dt$$
$$S_0 \in S_d^+, \quad t \geq 0 \tag{3.1}$$

- $Q \in GL_d$
- $M \in M_d$, $\lambda(M) < 0 \forall \lambda \in \sigma(M)$
- $\alpha > d - 1$
- $\sqrt{S_t}$ square root in matrix sense

Bru proved many interesting properties, like non collision of the eigenvalues. Most importantly, she characterized the conditional probability densities of S_t.

- calculation of the Laplace transform of the process
- first approach when $Q = I_d$, $M = 0$: $\rightarrow$ Kolmogorov type PDE
- second approach $\rightarrow$ matrix analogues of results from the theory of Bessel processes.

In particular, Bru successfully computed the quantity:

$$\mathbb{E}_{s_0}^{\mathbb{P}} \left[\exp \left\{ -Tr \left[w S_t + \int_0^t v S_s ds \right] \right\} \right], \tag{3.2}$$

where $w, v \in S_d^+$. The formula above is called, using Bru's terminology **Matrix Cameron-Martin Formula**. This formula is very useful for applications, unfortunately, Bru assumed that the matrices M and Q commute. We want to prove a more general version of the Cameron-Martin formula where we relax this commutativity assumption.

Our main result is the following:

Theorem 2 *Let $S \in WIS_d(S_0, \alpha, M, Q)$ be the Wishart process solving (3.1), assume*

$$M^\top \left(Q^\top Q \right)^{-1} = \left(Q^\top Q \right)^{-1} M, \tag{3.3}$$

let $\alpha \geq d+1$ and define the set of convergence of the Laplace transform

$$\mathcal{D}_t = \left\{ w, v \in S_d : \mathbb{E}^{\mathbb{P}}_{S_0} \left[\exp\left\{ -Tr\left[wS_t + \int_0^t vS_s ds \right] \right\} \right] < +\infty \right\}.$$

Then the joint moment generating function of the process and its integral is given by:

$$\begin{aligned} &\mathbb{E}^{\mathbb{P}}_{S_0} \left[\exp\left\{ -Tr\left[wS_t + \int_0^t vS_s ds \right] \right\} \right] \\ &= \det\left(e^{-Mt} \left(\cosh(\sqrt{\bar{v}}t) + \sinh(\sqrt{\bar{v}}t)k \right) \right)^{\frac{\alpha}{2}} \\ &\times \exp\left\{ Tr\left[\left(\frac{Q^{-1}\sqrt{\bar{v}}kQ^{\top^{-1}}}{2} - \frac{\left(Q^\top Q\right)^{-1} M}{2} \right) S_0 \right] \right\}, \end{aligned}$$

where the matrices $k, \bar{v}, \bar{w}$ are given by:

$$\begin{aligned} k &= -\left(\sqrt{\bar{v}} \cosh(\sqrt{\bar{v}}t) + \bar{w} \sinh(\sqrt{\bar{v}}t) \right)^{-1} \left(\sqrt{\bar{v}} \sinh(\sqrt{\bar{v}}t) + \bar{w} \cosh(\sqrt{\bar{v}}t) \right), \\ \bar{v} &= Q\left(2v + M^\top Q^{-1} Q^{\top^{-1}} M \right) Q^\top, \\ \bar{w} &= Q\left(2w - \left(Q^\top Q \right)^{-1} M \right) Q^\top. \end{aligned} \tag{3.4}$$

The proof of the result above proceeds along the following steps:

- Recall a matrix generalization of a result from the theory of Bessel processes,
- add the volatility matrix Q by means of an invariance result,
- add the drift matrix M by means of a measure change.

References

[1] Bru, M. F., *Wishart processes.* Journal of Theoretical Probability, 4 (1991), 725–751.

[2] Da Fonseca, J., Grasselli, M., and Tebaldi, C., *A multifactor volatility Heston model.* Quantitative Finance, 8(6) (2008), 591–604.

[3] Heston, S., *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options.* The Review of Financial Studies, 6(2) (1993), 327–343.

L^2 theory and global regularity for $\bar{\partial}$ on pseudoconvex domains of $\mathbb{C}^n$

Stefano Pinton (*)

Abstract. This seminar is divided into two parts. The first one is an introduction to the first order partial differential operator $\bar{\partial}$ in a smooth bounded pseudoconvex domain D of C^n. Only preliminary definitions are given and the basic estimate, due to Morrey, Kohn and Hormander, is established. It yields the existence of the $\bar{\partial}$ Neumann operator, that is, the inverse to the complex Laplacian and the construction of the canonical solution to the equation $\bar{\partial}u = f$, for $f \in \ker(\bar{\partial})$, that is the solution orthogonal to the kernel of $\bar{\partial}$. The second part is an introduction to the problem of the global regularity up to the boundary for the canonical solution of the $\bar{\partial}$ equation with data regular up to the boundary. In particular it is shown how compactness estimates are sufficient for global regularity as well as the existence of "good defining functions".

1 L^2 theory for $\bar{\partial}$ on pseudoconvex domains

In the first part of the seminar we will discuss the existence of the L^2 canonical solution for the equation $\bar{\partial}u = f$.

1.1 Preliminaries

Let $\mathbb{C}^n$, $n \geq 2$, be the euclidean complex $n-$dimensional space and $z \in \mathbb{C}^n$, $z = (z_1, \ldots, z_n)$, $z_j = x_j + iy_j$ be the complex coordinates. We define holomorphic and antiholomorphic derivatives by:

$$\partial_{\bar{z}_j} = \frac{1}{2}(\partial_{x_j} + i\partial_{y_j}) \quad \text{and} \quad \partial_{z_j} = \frac{1}{2}(\partial_{x_j} - i\partial_{y_j}).$$

Symmetrically we define the holomorphic and antiholomorphic forms by:

$$d\bar{z}_j = dx_j - idy_j \quad \text{and} \quad dz_j = dx_j + idy_j.$$

Let Ω be a compact subset of $\mathbb{C}^n$ with smooth boundary i.e. there exists a smooth real valued function $r : \mathbb{C}^n \to \mathbb{R}$ with $dr \neq 0$ such that $\Omega := \{z \in \mathbb{C}^n : r(z) < 0\}$.

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: pinton@math.unipd.it . Seminar held on November 16th, 2011.

Definition 1 A complex valued function $f \in C^1(\Omega)$ is said to be holomorphic when f satisfies the differential system $\partial_{\bar{z}_j} f = 0$, for $j = 1, \dots, n$.

The space $L^2(\Omega)^k$ of $k-$forms with L^2 coefficients is composed by $f = \sum_{|I|=k} f_I d\bar{z}_I$ where $f_I \in L^2(\Omega)$ and

$$\begin{aligned} I &= (i_1, \dots, i_k), \\ d\bar{z}_I &= d\bar{z}_{i_1} \wedge \cdots \wedge d\bar{z}_{i_k}, \\ f_{\sigma(I)} &= \operatorname{sign}(\sigma) f_I \quad (\sigma \text{ a permutation}), \end{aligned}$$

Usually $L^2(\Omega)^k$ is endowed with the scalar product

$$(f, g)_{L^2(\Omega)^k} = \sum_{|I|=k} \int_\Omega f_I \bar{g}_I \, dV.$$

We can consider also: $C^l(\bar{\Omega})^k$, $C^\infty(\bar{\Omega})^k$, $C^\infty(\Omega)^k$, $C_c^\infty(\Omega)^k$ and $H^s(\Omega)^k$.

1.2 $\bar{\partial}$ operator

Let $\bar{\partial} : C^\infty(\bar{\Omega})^k \to C^\infty(\bar{\Omega})^{k+1}$ for $k = 0, \dots, n-1$ be the first order differential operator defined by:

$$f = \sum_{|I|=k} f_I d\bar{z}_I \to \bar{\partial} f = \sum_{|I|=k} \sum_{i=1}^n \partial_{\bar{z}_i} f_I \, d\bar{z}_i \wedge d\bar{z}_I.$$

By the previous definition we have the following easy remarks:

Remark 2

- $\bar{\partial}\bar{\partial} f = 0$ (Schwartz Lemma);
- If $k = 0$, $\bar{\partial} f = \sum_{i=1}^n \partial_{\bar{z}_i} f \, d\bar{z}_i$.

Definition 3 The Levi form of a C^2 real valued function r is

$$L_r(z)(t) = \sum_{i,j} (\partial_{z_i} \partial_{\bar{z}_j} r(z)) \, t_i \bar{t}_j$$

for any $z \in \Omega$ and $t \in \mathbb{C}^n$. If the $L_r \geq 0$ then r is said plurisubharmonic.

Let $\vartheta : C^\infty(\bar{\Omega})^{k+1} \to C^\infty(\bar{\Omega})^k$ for $k = 1, \dots, n$ be the formal adjoint of $\bar{\partial}$ i.e. ϑ is defined by:

$$(\vartheta f, \varphi) = (f, \bar{\partial}\varphi) \quad \text{for any } \varphi \in C_c^\infty(\Omega)^k. \tag{1}$$

For any $f = \sum_{|I|=k+1} f_I d\bar{z}_I \in C^\infty(\bar{\Omega})^{k+1}$ integration by parts yields:

$$\vartheta f = -\sum_{|J|=k} \sum_{j=1}^n \partial_{z_j} f_{jJ} d\bar{z}_J.$$

By the previous definitions we have the following easy remarks:

Remark 4

- if $k = 1$ then $\vartheta(\sum_j f_j d\bar{z}_j) = -\sum_j \partial_{z_j} f_j$;
- we can think of both $\bar{\partial}$ and ϑ as (densily defined) operators on L^2-forms.

1.3 $\bar{\partial}^*$ operator and $\bar{\partial}$-Neumann condition

We denote by $\bar{\partial}^*$ the L^2-adjoint of $\bar{\partial}$ and again we have that $\bar{\partial}^*\bar{\partial}^* f = 0$. We observe that any $f \in C^\infty(\bar{\Omega})^k$ belongs to $\mathrm{Dom}(\bar{\partial})$ but not to $\mathrm{Dom}(\bar{\partial}^*)$; for this to hold the equality (1) must be satisfied $\forall \psi \in C^\infty(\bar{\Omega})^{k-1}$, not yet $C_c^\infty(\Omega)^{k-1}$:

$$(\bar{\partial}^* f, \psi) = (f, \bar{\partial}\psi) \qquad \forall \psi \in C^\infty(\bar{\Omega})^{k-1}. \tag{2}$$

But a form $\bar{\partial}^* f$ satisfying (2) exists in L^2 iff:

$$|(f, \bar{\partial}\psi)| \leq C\|\psi\| \qquad \forall \psi \in \mathrm{Dom}(\bar{\partial}) \quad \text{(Riesz)},$$

that is iff f satisfies Neumann condition:

$$(f \cdot \partial r)_{|b\Omega} = 0 \qquad \text{(integration by parts)}$$

here $\mathrm{Re}(\partial r) = dr$. When $f \in \mathrm{Dom}(\bar{\partial}^*)$ then $\bar{\partial}^* f = \vartheta f$. In the case of $1-$forms, i.e. $f = \sum_j^n f_j d\bar{z}_j$, we have $f \in \mathrm{Dom}(\bar{\partial}^*)$ iff $(\sum_{j=1}^n \partial_{z_j} r\, f_j)_{|b\Omega} = 0$. For a general form of degree k, the Neumann condition consists in:

$$\forall |J| = k-1 \qquad \big(\sum_{j=1}^n \partial_{z_j} r\, f_{jJ}\big)_{|b\Omega} = 0.$$

1.4 Basic estimate

When $f \in \mathrm{Dom}(\bar{\partial}) \cap \mathrm{Dom}(\bar{\partial}^*)$ then

$$\begin{aligned} \|\bar{\partial} f\|^2 &= -\sum_{i,j,|L|=k-1} \int_\Omega \partial_{\bar{z}_i} f_{jL} \overline{\partial_{\bar{z}_j} f_{iL}}\, dV + \sum_{i,|J|=k} \int_\Omega |\partial_{\bar{z}_i} f_J|^2\, dV \\ \|\bar{\partial}^* f\|^2 &= \sum_{i,j,|L|=k-1} \int_\Omega \partial_{z_i} f_{iL} \overline{\partial_{z_j} f_{jL}}\, dV. \end{aligned}$$

Integration by parts and the $\bar{\partial}-$Neumann conditions imply:

$$-\sum_{i,j,L} \int_\Omega \partial_{\bar{z}_i} f_{jL} \partial_{z_j} \bar{f}_{iL}\, dV + \sum_{i,j,L} \int_\Omega \partial_{z_i} f_{iL} \partial_{\bar{z}_j} \bar{f}_{jL}\, dV = \sum_{i,j,L} \int_{b\Omega} \partial_{z_i} \partial_{\bar{z}_j} r\, f_{iL} \bar{f}_{jL}\, dS$$

This yields:

$$\|\bar{\partial} f\|^2 + \|\bar{\partial}^* f\|^2 = \sum_{i,J} \|\partial_{\bar{z}_i} f_J\|^2 + \sum_{i,j,L} \int_{b\Omega} \partial_{z_i} \partial_{\bar{z}_j} r\, f_{iL} \bar{f}_{jL}\, dS.$$

Now if we consider a weighted scalar product:

$$(f,g)_\varphi = \sum_{|J|=k} \int_\Omega e^{-\varphi} f_J \bar{g}_J \, dV$$

we obtain:

$$\begin{aligned} \|\bar\partial f\|^2 + \|\bar\partial^* f\|^2 \;\gtrsim\; & \sum_{i,J} \|\partial_{\bar z_i} f_J\|^2 + \sum_{i,j,L} \int_{b\Omega} \partial_{z_i}\partial_{\bar z_j} r \, f_{iL} \bar f_{jL} \, dS \\ & + \sum_{i,j,L} \int_\Omega \partial_{z_i}\partial_{\bar z_j} \varphi \, f_{iL} \bar f_{jL} \, dV. \end{aligned} \tag{3}$$

Good weights, with big Levi form, always exist e.g. $\varphi = t|z|^2$, $t >> 1$; but excellent weights are those which have, in addition, a uniform bound in Ω near $b\Omega$:

$$\begin{cases} \partial_{z_i}\partial_{\bar z_j}\varphi_t \geq t \\ |\varphi_t| \leq 1 \qquad \text{near } b\Omega \end{cases} \tag{4}$$

When the $b\Omega$ is "'flat"' these weights do not exist. Now we introduce some domains for which the boundary integral in (3) is positive. They are called pseudoconvex domains.

Definition 5 Let Ω be bounded domain of $\mathbb{C}^n$ with smooth defining function r (i.e. $\Omega = \{ z \in \mathbb{C}^n : r(z) < 0 \}$, $dr \neq 0$). Ω is pseudoconvex iff

$$L_r(z)(t) = \sum_{i,j} (\partial_{z_i}\partial_{\bar z_j} r\,(z))\, t_i \bar t_j \geq 0 \tag{5}$$

for any $z \in b\Omega$ and for any $t \in \mathbb{C}^n \setminus \{0\}$ such that $\langle \partial r(z), t\rangle = 0$. When strict inequality holds in (5), Ω is said strictly pseudoconvex.

Thus for pseudoconvex domains we obtain the *Basic Estimates*:

$$\|\bar\partial f\| + \|\bar\partial^* f\| \gtrsim \sum_{i,J} \|\partial_{\bar z_i} f_J\|^2 + \int_\Omega \partial_{z_i}\partial_{\bar z_j}\varphi \, f_{iK} \bar f_{jK} \, dV \qquad \text{(basic estimate)} \tag{6}$$

for any $f \in \mathrm{Dom}(\bar\partial) \cap \mathrm{Dom}(\bar\partial^*)$ (for a complete derivation of the basic estimates we suggest [6, Chapter 4] or [5, Chapter 1.9]).

Here some examples of pseudoconvex domains:

- All convex domains are pseudoconvex. The ball $\mathcal{B} = \{ z \in \mathbb{C}^n : |z|^2 < 1 \}$ is strictly pseudoconvex.
- Let $f_1, \ldots, f_n$ be holomorphic functions of one variable. Then $\{ z \in \mathbb{C}^n : |f_1(z_1)|^2 + \cdots + |f_n(z_n)|^2 \leq C \}$ is pseudoconvex. In fact for $r = |f_1(z_1)|^2 + \cdots + |f_n(z_n)|^2 - C$, we have, for any $t \in \mathbb{C}^n$, that:

$$L_r(z)(t) = \sum_j |\partial_{z_j} f_j(z_j)|^2 \, |t|_j^2 \geq 0$$

.

1.5 Complex laplacian $\Box$ and Neumann operator N

Let $\Box = \bar\partial\bar\partial^* + \bar\partial^*\bar\partial : L^2(\Omega)^k \to L^2(\Omega)^k$ for $k = 1, \dots, n-1$, be a second order differential operator. This is called the complex laplacian. $\Box f$ should exist, then $\Box f = -\frac{1}{4}\sum_{|J|=k} \Delta f_J \, d\bar z_J$. If Ω is pseudoconvex, by the basic estimate (6), putting $\varphi = |z|^2$ and using the adjunction properties, we have:

$$\|f\|^2 \lesssim \|\bar\partial f\|^2 + \|\bar\partial^* f\|^2 = (\Box f, f) \le \|\Box f\|\|f\| \qquad \forall f \in \mathrm{Dom}(\Box)$$

where the last inequality is a Cauchy-Schwartz inequality. Thus

$$\|f\| \lesssim \|\Box f\|.$$

The last inequality tells us that $\Box$ is injective and it has closed range. Its inverse $N := \Box^{-1}$ is the **Neumann operator**:

$$N : L^2(\Omega)^k \to \mathrm{Dom}(\Box)^k \quad \text{for } k = 1, \dots, n-1.$$

It is a linear, selfadjoint and bounded operator.

We have the orthogonal decomposition:

$$\begin{aligned} L^2(\Omega)^k &= \mathrm{Range}(\bar\partial\bar\partial^* N) \oplus \mathrm{Range}(\bar\partial^*\bar\partial N) \\ &= \ker(\bar\partial) \oplus \ker(\bar\partial^*) \qquad \text{for any } k = 1, \dots, n-1. \end{aligned} \tag{7}$$

Thus we obtain the principal results of the first part.

Theorem 6 [L^2 canonical solution of $\bar\partial$ equation] *Let $\Omega \subset \mathbb{C}^n$ be a smooth bounded pseudoconvex domain and consider the equation $\bar\partial u = f$, where $f \in L^2(\Omega)^k$, for $k = 1, \dots, n$ and $\bar\partial f = 0$. Then $u = \bar\partial^* N f$ is a solution orthogonal to* $\ker(\bar\partial)$ (unique by (7)).

Proof. Since $\bar\partial f = 0$ and $\bar\partial N = N\bar\partial$ we have: $\bar\partial u = \bar\partial\bar\partial^* N f = \bar\partial\bar\partial^* N f + \bar\partial^* N \bar\partial f = (\bar\partial\bar\partial^* + \bar\partial^*\bar\partial) N f = \Box N f = f$. □

2 Regularity of the solution of $\bar\partial$ for pseudoconvex domains

In the second part of the seminar we will discuss the regularity up to the boundary of the canonical solution.

2.1 At the interior

Not all solutions are regular. For this we have to add $\bar\partial^*$ and consider the system $(\bar\partial, \bar\partial^*)$. Then we use the "'other half"' of the basic estimate:

$$\|\alpha\| + \sum_J \sum_j \|\partial_{\bar z_j} \alpha_J\| \lesssim \|\bar\partial\alpha\| + \|\bar\partial^*\alpha\|. \tag{8}$$

In particular, if $\alpha \in C_c^\infty(\Omega)^k$ converting each $\frac{1}{2}\|\partial_{\bar z_j}\alpha_J\|$ into $\frac{1}{2}\|\partial_{z_j}\alpha_J\|$ by integration by parts yields:

$$\frac{1}{2}\|\alpha\|_1 \lesssim \|\bar\partial\alpha\| + \|\bar\partial^*\alpha\| \quad \text{(elliptic estimate in the interior)}. \tag{9}$$

In particular if $f \in C^\infty(\Omega)^k \cap L^2(\Omega)^k$ then the canonical solution $\bar\partial^* N f \in C^\infty(\Omega)^{k-1}$. In other words, the system $\partial_{\bar z_j}\ j = 1, \dots, n$ is elliptic; to have its action over all coefficients α_J, the full action of $(\bar\partial, \bar\partial^*)$ over α is needed.

2.2 At the boundary

First, if $f \in C^\infty(\bar\Omega)$ there is a $C^\infty(\bar\Omega)$ solution [4]. Again, not all solutions are $C^\infty(\bar\Omega)$ but, differently from the interior, solutions of $(\bar\partial, \bar\partial^*)$ do not always inherit regularity from the data. In particular, for $C^\infty(\bar\Omega)$, $\bar\partial^* N f$ may happen not to be regular! At the boundary ($\alpha \in C^\infty(\bar\Omega)$ in (9) is needed instead of $\alpha \in C_c^\infty(\Omega)$) the elliptic estimate always fails: the best that can be expected is a $\frac{1}{2}-$subelliptic estimate (this is the case of strongly pseudoconvex domains, with $\|\alpha\|_1$ replaced by $\|\alpha\|_{\frac{1}{2}}$, $\alpha \in C^\infty(\bar\Omega)^k$); otherwise more general $\|\cdot\|_\delta-$norms with $\delta \le \frac{1}{2}$ or even weaker norms can be expected in the estimate:

$$\|\alpha\|_\delta \le \|\bar\partial\alpha\| + \|\bar\partial^*\alpha\| \qquad \alpha \in C^\infty(\bar\Omega)^k \cap \mathrm{Dom}(\bar\partial^*)^k$$

to obtain regularity up to the boundary for the canonical solutions.

The ultimate level of estimates for regularity at the boundary for the canonical solution are compactness estimates. They hold if $\forall \epsilon > 0$ and for suitable $C_\epsilon > 0$:

$$\frac{1}{\epsilon}\|\alpha\| \le \|\bar\partial\alpha\| + \|\bar\partial^*\alpha\| + C_\epsilon\|\alpha\|_{-1} \qquad \forall\, \alpha \in \mathrm{Dom}(\bar\partial) \cap \mathrm{Dom}(\bar\partial^*) \tag{10}$$

Sketch of the proof that (10) implies regularity.

$$\frac{1}{\epsilon}\|\alpha\|_s \lesssim \|\bar\partial\alpha\|_s + \|\bar\partial^*\alpha\|_s + s\|\alpha\|_s + c_\epsilon\|\alpha\|_0 \tag{11}$$

where $sD^s \sim [D^s, \bar\partial], [D^s, \bar\partial^*]$ and reduction from $\|\alpha\|_{s-1}$ to $\|\alpha\|_0$ is obtained by iteration. By choosing $\frac{1}{\epsilon} \gtrsim s$ one gets an estimate for H^s regularity: letting $s \nearrow +\infty$ we are done (again, since $(\bar\partial(\bar\partial^* N f) = f, \bar\partial^*(\bar\partial^* N f) = 0)$, (11) yield the $\bar\Omega-$regularity of $\bar\partial^* N f$ when f is $\bar\Omega-$regular). □

Condition for compactness: a family of bounded weights with arbitrarly large Levi-form, as in (4),suffices for the compactness estimate [7]. In fact by the weighted basic, (4) yields (10) $\forall \epsilon = \frac{1}{t}$. (10) is not necessary for $\bar\Omega-$regularity of $(\bar\partial, \bar\partial^*)$. Existence of defining functions which are fully plurisubharmonic (not only when restricted to vectors orthogonal to ∂r) is sufficient.

2.3 My contribution

I obtained some results in this field. Here I recall them briefly:

> Compactness estimates for the tangential $(\bar\partial, \bar\partial^*)-$system on an abstract pseudoconvex CR-manifold (not embedded nor orientable) [1].
>
> Regularity in absence of compactness for a more general r than plurisubharmonic. It is classical that if Ω is pseudoconvex, then there exists r and $0 < \eta \le 1$ such

that $-(-r)^\eta$ is plurisubharmonic in Ω. In my research the index s of regularity is related to (the inverse) of $(1-\eta)$. This was done by Kohn for functions [3] and is now generalized to any degree of forms [2].

References

[1] G. Zampieri, S. Pinton, *Compactness of $\Box_b$ in a CR manifold.* Proceedings of American Mathematical Society.

[2] G. Zampieri, S. Pinton, *The Diederich-Fornaess index and the global regularity of the $\bar\partial-$Neumann problem.* Preprint, `arXiv:1108.3598v2`.

[3] J. J. Kohn, *Quantitative estimates for global regularity.* Analysis and geometry in several complex variables (Katata, 1997), Trends Math., 97–128.

[4] J. J. Kohn, *Global regularity for $\bar\partial$ on weakly pseudoconvex manifolds.* Trans. Amer. Math. Soc., 181 (1973), 273–292.

[5] G. Zampieri, "Complex Analysis and CR Geometry". University Lecture Series, Amer. Math. Soc., volume 43.

[6] S. C. Chen, M. C. Shaw, "Partial Differential Equations in Several Complex Variables". Studies in Advanced Mathematics, Amer. Math. Soc..

[7] David W. Catlin, *Global regularity of the $\bar\partial$-Neumann.* Complex analysis of several variables (Madison, Wis., 1982), volume 41 of Proc. Sympos. Pure Math., Amer. Math. Soc., Providence, RI (1984), 39–49.

Boundedness and compactness of matrix operators in weighted spaces of sequences

Zhanar Taspaganbetova (*)

The present paper is based on joint work with Professor Ryskul Oinarov

Abstract. One of the main problems in the theory of matrices is to find necessary and sufficient conditions for the elements of a matrix so that the corresponding matrix operator maps continuously one normed space of sequences into another normed space of sequences. Thus it is very important to find the norm of a matrix operator, or at least, an upper or lower bound for the norm. However, in several spaces, which are very important both theoretically and in the applications, such problems have not been solved yet in full generality for operators corresponding to arbitrary matrices. Therefore, in such spaces researchers have considered some specific classes of matrix operators and have established criteria of boundedness and compactness for operators of such classes.

We prove a new discrete Hardy type inequality involving a kernel which has a more general form than those known in the literature. We obtain necessary and sufficient conditions for the boundedness and compactness of a matrix operator from the weighted $l_{p,v}$ space into the weighted $l_{q,u}$ space defined by $(A^+f)_i := \sum\limits_{j=1}^{i} a_{i,j}f_j$, for all $f = \{f_i\}_{i=1}^{\infty} \in l_{p,v}$ in case $1 < p \le q < \infty$ and $a_{i,j} \ge 0$. Then we deduce a corresponding dual statement.

1 Historical notes on the problem

In the second half of the last century reseachers singled out a class of integral operators, which is called the class of Hardy type operators, which is related to the work [4] of G. Hardy. In 1925 Hardy has established the boundedness of the operator H in $L_p(0,\infty)$ for $1 < p < \infty$ defined by

$$(Hf)(x) = \frac{1}{x}\int\limits_0^x f(s)ds \qquad \forall f \in L_p(0,\infty),$$

and has proved that $\|H\|_{p\to p} = \frac{p}{p-1}$.

(*)The L. N. Gumilyov Eurasian National University, Munaitpasov st., 5, 010008, Astana, Kazakhstan; E-mail: zhanara.t.a@yandex.kz . Seminar held on November 23rd, 2011.

However, in several applications in function theory, harmonic analysis and differential equations, one needs to consider weighted forms of Hardy operators. Namely, one needs to consider non-negative weights $u(x)$ and $v(x)$ in Lebesque spaces and operators K_0 of the form $(\mathcal{K}_0 f)(x) = u(x)\int\limits_0^x v(s)f(s)ds$. The problem was not easy. Only in 1969 the Italian mathematicians G. Talenti [5] and G. Tomaselli [6] have established, independently of each other, criteria of boundedness of the operator $\mathcal{K}_0$ in $L_p(0,\infty)$. During the next 11 years B. Muckenhoupt [7], J.S. Bradley [8], V.M. Kokilashvili [9], V.G. Maz'ja [10] have obtained criteria of boundedness for the operator $\mathcal{K}_0$ from $L_p(0,\infty)$ to $L_q(0,\infty)$ with $1 \le p, q \le \infty$. The initial results of G. Talenti, G. Tomaselli, B. Muckenhoupt gave a new impetus in the analysis of weighted embedding theorems and spectral problems for singular differential operators. Thus for example M. Otelbaev and his school have obtained important results in the 1970's concerning such topics (see e.g. [11], [12], [13]).

The next step was to study the operator

$$(\mathcal{K}f)(x) = \int\limits_0^x K(x,s)f(s)ds$$

with non-negative kernel $K(\cdot,\cdot)$. Such type of operators are called Hardy type operators. However, even in the space $L_2(0,\infty)$ finding a criterion of boundedness for such general form of operators in terms of the kernel $K(\cdot,\cdot)$ is very difficult and is still an unsolved problem. Therefore, many researchers have identified several classes of kernels, which satisfy some specific conditions and have proved boundedness criteria for the corresponding integral operators from $L_p(0,\infty)$ to $L_q(0,\infty)$, $1 < p, q < \infty$.

The first impulse in this direction was given by the works of F. Martin-Reyes, E. Sawyer [14], and V.D. Stepanov [15]-[18]. They have obtained criteria of boundedness of the Riemann-Liouville fractional integration operator from $L_{p,v}$ to $L_{q,u}$, $1 < p, q < \infty$, which has several applications in various fields of science. In [19] V.D. Stepanov has investigated the operator $\mathcal{K}$ with kernel $K(x,s) = k(x-s)$, where $k(\cdot)$ is nondecreasing and for which there exists $d \ge 1$ such that $k(x+s) \le d(k(x)+k(s))$, $x, s \in (0,\infty)$. In 1989-1990 R. Oinarov [20] and independently the American mathematicians S. Bloom and R. Kerman [21] in 1991 have studied the operator $\mathcal{K}$, when its kernel satisfies the following condition

$$\frac{1}{d}(K(x,t)+K(t,s)) \le K(x,s) \le d(K(x,t)+K(t,s)),$$

$x \ge t \ge s > 0$, $d \ge 1$. One of the important feature of this class of operators is that it includes almost all known operators of fractional integration. Nowadays, this condition imposed on the kernel $K(\cdot,\cdot)$ of the operator $\mathcal{K}$ is called the "Oinarov condition" in the mathematical literature. Operators $\mathcal{K}$ satisfying the Oinarov condition have been investigated by many authors (see e.g. [2], [22]).

In the twenties of the last century G. Hardy has considered the discrete analogue of the operator H in the form $(H^d f)_i = \frac{1}{i}\sum\limits_{j=1}^{i} f_j$ and has proved the boundedness of H^d in the space of sequences l_p and a formula for the norm $\|H^d\|_{p\to p} = \dfrac{p}{p-1}$, $1 < p < \infty$. As

in the continuous case, this result of Hardy had various applications in many problems. The discrete analogue $(K_0^d f)_i = u_i \sum\limits_{j=1}^{i} v_j f_j$ of the operator K_0 has been studied by many authors and the main results have been obtained in [23]-[28] only in 1987-1994 by K.F. Anderson, H.P. Heinig, G. Bennett, M.Sh. Braveman and V.D. Stepanov. Such delay of decades is related with the discrete changes of sequences $\{f_j\}$ and $\{(K_0^d f)_i\}$, which do not allow to transfer methods of the continuous case based on the continuity of $(K_0 f)(\cdot)$. The results which were obtained for the operator K_0^d have been successfully applied by mathematicians of different countries. For example in Kazakhstan M. Otelbaev [13], E.S. Smailov [29]-[31], A. Stikharnyi [32], R. Oinarov, A. Stikharnyi [33] and other authors have proved several applications in various problems of analysis.

An attempt to investigate more general matrix operator of Hardy type $(Af)_i = u_i \sum\limits_{j=1}^{i} a_{i,j} v_j f_j$, $a_{i,j} \geq 0$ has been done by K.F Andersen and H.P. Heinig [23], who have proved sufficient conditions for the boundedness of the operator A in the space l_p under some conditions on the matrix $(a_{i,j})$.

R. Oinarov, S.Kh. Shalgynbaeva [34] and R. Oinarov, C.A.Okpoti, L-E. Persson [35] have proved criteria of boundedness and compactness for the operator A from l_p to l_q, when the entries of the matrix $(a_{i,j})$ satisfy a discrete analogue of the "Oinarov condition".

The next problem was the following.

How to obtain criteria of boundedness and compactness for a wider class of matrix operators, or, in other words, how to prove a weaker condition on the matrix $(a_{i,j})$.

So we introduce the expanding classes of matrices, which are wider than the previously studied classes of matrices. When the matrices of operators $(A^+ f)_i := \sum\limits_{j=1}^{i} a_{i,j} f_j$, $i \geq 1$ and $(A^- f)_j := \sum\limits_{i=j}^{\infty} a_{i,j} f_i$, $j \geq 1$ belong to such classes, we obtain necessary and sufficient conditions for the boundedness and compactness of corresponding matrix operators from the weighted $l_{p,v}$ space into the weighted $l_{q,u}$ space $f = \{f_i\}_{i=1}^{\infty} \in l_{p,v}$ in case $1 < p \leq q < \infty$ and $a_{i,j} \geq 0$. Such operators have their own self interest and they are also a discrete analogue of integral operators, which play a very important role in functional analysis (see [2], [3]).

2 Introduction

Let $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{p'} = 1$ and $u = \{u_i\}_{i=1}^{\infty}$, $v = \{v_i\}_{i=1}^{\infty}$ be positive sequences of real numbers. Let $l_{p,v}$ be the space of sequences $f = \{f_i\}_{i=1}^{\infty}$ of real numbers such that

$$\|f\|_{p,v} := \left(\sum_{i=1}^{\infty} |v_i f_i|^p \right)^{\frac{1}{p}} < \infty, \qquad 1 < p < \infty.$$

We consider the problem of boundedness from the weighted $l_{p,v}$ space into the weighted

$l_{q,u}$ space of the matrix operators

$$\left(A^+f\right)_i := \sum_{j=1}^{i} a_{i,j}f_j, \qquad i \geq 1, \tag{2.1}$$

$$\left(A^-f\right)_j := \sum_{i=j}^{\infty} a_{i,j}f_i, \qquad j \geq 1, \tag{2.2}$$

which is equivalent to the validity of the following inequality

$$\|A^{\pm}f\|_{q,u} \leq C\|f\|_{p,v} \qquad \forall f \in l_{p,v}, \tag{2.3}$$

where C is a positive finite constant independent of f and $(a_{i,j})$ is a non-negative triangular matrix with entries $a_{i,j} \geq 0$ for $i \geq j \geq 1$ and $a_{i,j} = 0$ for $i < j$.

For $a_{i,j} = 1$, $i \geq j \geq 1$, the operators (2.1), (2.2) coincide with the discrete Hardy operators of the forms $\left(A_0^+f\right)_i := \sum\limits_{j=1}^{i} f_j$, $\left(A_0^-f\right)_j := \sum\limits_{i=j}^{\infty} f_i$, respectively. References about generalizations of the original forms of the discrete and continuous Hardy inequalities can be found in different books, see e.g. [4].

When one of parameters p or q is equal to 1 or ∞, necessary and sufficient conditions of the validity of (2.3) with the exact value of the best constant $C > 0$ have been obtained in [1]. In case $1 < p, q < \infty$ inequalities as (2.3) have not been established yet for arbitrary matrices $(a_{i,j})$. Instead inequality (2.3) has been established with certain restrictions on the matrix $(a_{i,j})$.

In [34], [35] necessary and sufficient conditions for the validity of (2.3) have been obtained for $1 < p, q < \infty$ under the assumption that there exists $d \geq 1$ such that the inequalities

$$\frac{1}{d}(a_{i,k} + a_{k,j}) \leq a_{i,j} \leq d(a_{i,k} + a_{k,j}), \qquad i \geq k \geq j \geq 1 \tag{2.4}$$

hold.

A sequence $\{a_i\}_{i=1}^{\infty}$ is called almost non-decreasing (non-increasing), if there exists $c > 0$ such that $ca_i \geq a_k$ ($a_k \leq ca_j$) for all $i \geq k \geq j \geq 1$.

In [36] estimate (2.3) has been studied under the assumption that there exist $d \geq 1$ and a sequence of positive numbers $\{\omega_k\}_{k=1}^{\infty}$, and a non-negative matrix $(b_{i,j})$, where $b_{i,j}$ is almost non-decreasing in i and almost non-increasing in j, such that the inequalities

$$\frac{1}{d}(b_{i,k}\omega_j + a_{k,j}) \leq a_{i,j} \leq d(b_{i,k}\omega_j + a_{k,j}) \tag{2.5}$$

hold for all $i \geq k \geq j \geq 1$.

In [37], [38] the authors have considered inequality (2.3) under the assumption that there exist $d \geq 1$, a sequence of positive numbers $\{\omega_k\}_{k=1}^{\infty}$, and a non-negative matrix

$(b_{i,j})$, whose entries $b_{i,j}$ are almost non-decreasing in i and almost non-increasing in j such that the inequalities

$$\frac{1}{d}(a_{i,k} + b_{k,j}\omega_i) \le a_{i,j} \le d(a_{i,k} + b_{k,j}\omega_i) \tag{2.6}$$

hold for all $i \ge k \ge j \ge 1$.

Conditions (2.5) and (2.6) include conditions (2.4), and complement each other.

Notation: If M and K are real valued functionals of sequences, then we understand that the symbol $M \ll K$ means that there exists $c > 0$ such that $M \le cK$, where c is a constant which may depend only on parameters such as p, q and r_n, but not on the sequences in the arguments of M and K. If $M \ll K \ll M$, then we write $M \approx K$.

3 Preliminaries and notation

We now introduce the classes $\mathcal{O}_n^+$ and $\mathcal{O}_n^-$ of matrices $(a_{i,j})$ for $n \ge 1$. We assume that $a_{i,j} \equiv a_{i,j}^{(n)}$ if $(a_{i,j}) \in \mathcal{O}_n^+$ or $(a_{i,j}) \in \mathcal{O}_n^-$.

We define the classes $\mathcal{O}_n^+$ for $n \ge 0$ by induction. Let $(a_{i,j})$ be non-negative and non-decreasing in the first index matrix for all $i \ge j \ge 1$. By definition matrices of the type $a_{i,j}^{(0)} = \alpha_j$, $\forall i \ge j \ge 1$ belong to the class $\mathcal{O}_0^+$. Let the classes $\mathcal{O}_\gamma^+$, $\gamma = 0, 1, ..., n-1$, $n \ge 1$ be defined. By definition, the matrix $(a_{i,j}) \equiv (a_{i,j}^{(n)})$ belongs to the class $\mathcal{O}_n^+$ if and only if there exist matrices $(a_{i,j}^{(\gamma)}) \in \mathcal{O}_\gamma^+$, $\gamma = 0, 1, ..., n-1$ and a number $r_n > 0$ such that

$$a_{i,j}^{(n)} \le r_n \sum_{\gamma=0}^{n} b_{i,k}^{n,\gamma} a_{k,j}^{(\gamma)} \tag{3.1}$$

for all $i \ge k \ge j \ge 1$, where $b_{i,k}^{n,n} \equiv 1$ and

$$b_{i,k}^{n,\gamma} = \inf_{1 \le j \le k} \frac{a_{i,j}^{(n)}}{a_{k,j}^{(\gamma)}}, \qquad \gamma = 0, 1, ..., n-1. \tag{3.2}$$

From (3.2) the following inequality follows

$$a_{i,j}^{(n)} \ge b_{i,k}^{n,\gamma} a_{k,j}^{(\gamma)} \tag{3.3}$$

$$i \ge k \ge j \ge 1, \gamma = 0, 1, ..., n, n = 0, 1, ...$$

Then for $(a_{i,j}^{(n)}) \in \mathcal{O}_n^+$ we have

$$a_{i,j}^{(n)} \approx \sum_{\gamma=0}^{n} b_{i,k}^{n,\gamma} a_{k,j}^{(\gamma)}, \qquad n \ge 0 \tag{3.4}$$

for all $i \geq k \geq j \geq 1$.

Remark 1 It is easy to show that if for the matrix $(a_{i,j}^{(n)})$, $n \geq 0$ there exist matrices $(a_{i,j}^{(\gamma)}) \in \mathcal{O}_\gamma^+$, $\gamma = 0, 1, ..., n-1$, and matrices $(\widetilde{b}_{i,k}^{n,\gamma})$, $\gamma = 0, 1, ..., n$ such that the equivalence (3.4) is valid for all $i \geq k \geq j \geq 1$, then $(a_{i,j}^{(n)}) \in \mathcal{O}_n^+$ and $\widetilde{b}_{i,k}^{n,\gamma} \approx b_{i,k}^{n,\gamma}$. Hence, we may assume that the matrices $(b_{i,k}^{n,\gamma})$ are arbitrary non-negative matrices which satisfy (3.4).

As above, we introduce the classes $\mathcal{O}_m^-$, $m \geq 0$. Let $(a_{i,j})$ be a non-negative matrix. Let $(a_{i,j})$ be non-increasing in the second index for all $i \geq j \geq 1$. By definition, a matrix $(a_{i,j}) = (a_{i,j}^{(0)})$ belongs to the class $\mathcal{O}_0^-$ if and only if it has the form $a_{i,j}^{(0)} = \beta_i$ for all $i \geq j \geq 1$. Let the classes $\mathcal{O}_\gamma^-$, $\gamma = 0, 1, ..., m-1$, $m \geq 1$ be defined. A matrix $(a_{i,j}) = (a_{i,j}^{(m)})$ belongs to the class $\mathcal{O}_m^-$ if and only if there exist matrices $(a_{i,j}^{(\gamma)}) \in \mathcal{O}_\gamma^-$, $\gamma = 0, 1, ..., m-1$ and matrices $(d_{k,j}^{\gamma,m})$, $\gamma = 0, 1, ..., m-1$, $m = 0, 1, ...$, such that

$$a_{i,j}^{(m)} \approx \sum_{\gamma=0}^{m} a_{i,k}^{(\gamma)} d_{k,j}^{\gamma,m}, \tag{3.5}$$

for all $i \geq k \geq j \geq 1$.

Remark 2 By the definitions of the classes $\mathcal{O}_n^\pm$, $n \geq 0$ we have $\mathcal{O}_0^\pm \subset \mathcal{O}_1^\pm \subset ... \subset \mathcal{O}_n^\pm \subset ....$

It is easy to see that the class $\mathcal{O}_1^+$ includes the matrices, whose entries satisfy conditions (2.4) and (2.5). Also it should be noted that the matrices satisfying conditions (2.4) and (2.6) belong to the class $\mathcal{O}_1^-$. Hence, the classes $\mathcal{O}_n^+$, $n \geq 1$ and $\mathcal{O}_m^-$, $m \geq 1$ are wider than the classes of matrices which have been studied before in the literature.

A continuous analogue of the classes $\mathcal{O}_n^+$ and $\mathcal{O}_n^-$, $n \geq 0$ has been considered by R. Oinarov in [39].

Next, we show properties of the classes of matrices $\mathcal{O}_n^+$ and $\mathcal{O}_n^-$ for $n \geq 0$.

We set

$$w_{i,k} = \sum_{j=k}^{i} a_{i,j} \sigma_{j,k}.$$

Then we have the following

Lemma 3.1 *Let* $(a_{i,j}) \in \mathcal{O}_n^+$, $(\sigma_{j,k}) \in \mathcal{O}_m^+$. *Then* $(w_{i,k}) \in \mathcal{O}_{m+n+1}^+$.

Now we set

$$\varphi_{k,j} = \sum_{i=j}^{k} \sigma_{k,i} a_{i,j}.$$

Then we have the following lemma.

Lemma 3.2 *Let* $(a_{i,j}) \in \mathcal{O}_n^-$, $(\sigma_{k,i}) \in \mathcal{O}_m^-$. *Then* $(\varphi_{k,j}) \in \mathcal{O}_{m+n+1}^-$.

4 Main results

We define

$$\left(\mathcal{B}^+_{p,q}\right)_k = \left(\sum_{j=1}^{k} v_j^{-p'} \left(\sum_{i=k}^{\infty} a_{i,j}^q u_i^q\right)^{\frac{p'}{q}}\right)^{\frac{1}{p'}},$$

$$\left(\mathcal{B}^-_{p,q}\right)_k = \left(\sum_{i=k}^{\infty} u_i^q \left(\sum_{j=1}^{k} a_{i,j}^{p'} v_j^{-p'}\right)^{\frac{q}{p'}}\right)^{\frac{1}{q}},$$

$$\left(\mathcal{A}^+_{p,q}\right)_k = \left(\sum_{j=1}^{k} u_j^q \left(\sum_{i=k}^{\infty} a_{i,j}^{p'} v_i^{-p'}\right)^{\frac{q}{p'}}\right)^{\frac{1}{q}},$$

$$\left(\mathcal{A}^-_{p,q}\right)_k = \left(\sum_{i=k}^{\infty} v_i^{-p'} \left(\sum_{j=1}^{k} a_{i,j}^q u_j^q\right)^{\frac{p'}{q}}\right)^{\frac{1}{p'}}.$$

We set $\mathcal{B}^+ = \sup_{k\geq 1} \left(\mathcal{B}^+_{p,q}\right)_k$, $\mathcal{B}^- = \sup_{k\geq 1} \left(\mathcal{B}^-_{p,q}\right)_k$, $\mathcal{A}^+ = \sup_{k\geq 1} \left(\mathcal{A}^+_{p,q}\right)_k$ and $\mathcal{A}^- = \sup_{k\geq 1} \left(\mathcal{A}^-_{p,q}\right)_k$.

Theorem 4.1 *Suppose that $1 < p \leq q < \infty$. Let the matrix $(a_{i,j})$ in (2.1) belong to the class $\mathcal{O}^+_m \cup \mathcal{O}^-_m$, $m \geq 0$. Let A^+ be the operator defined in (2.1). Then the following statements hold.*

(i) *A^+ is bounded from $l_{p,v}$ into $l_{q,u}$ if and only if at least one of the conditions $\mathcal{B}^+ < \infty$ and $\mathcal{B}^- < \infty$ holds. Moreover $\mathcal{B}^+ \approx \mathcal{B}^- \approx C$, where C is the best constant in (2.3).*

(ii) *A^+ is compact from $l_{p,v}$ into $l_{q,u}$ if and only if at least one of the conditions $\lim_{k\to\infty} \left(\mathcal{B}^+_{p,q}\right)_k = 0$ and $\lim_{k\to\infty} \left(\mathcal{B}^-_{p,q}\right)_k = 0$ holds.*

Theorem 4.2 *Suppose that $1 < p \leq q < \infty$. Let the matrix $(a_{i,j})$ in (2.2) belong to the class $\mathcal{O}^+_m \cup \mathcal{O}^-_m$, $m \geq 0$. Let A^- be the operator defined in (2.2). Then the following statements hold.*

(j) *A^- is bounded from $l_{p,v}$ into $l_{q,u}$ if and only if at least one of the conditions $\mathcal{A}^+ < \infty$ and $\mathcal{A}^- < \infty$ holds. Moreover $\mathcal{A}^+ \approx \mathcal{A}^- \approx C$, where C is the best constant in (2.3).*

(jj) *A^- is compact from $l_{p,v}$ into $l_{q,u}$ if and only if at least one of the conditions $\lim_{k\to\infty} \left(\mathcal{A}^+_{p,q}\right)_k = 0$ and $\lim_{k\to\infty} \left(\mathcal{A}^-_{p,q}\right)_k = 0$ holds.*

Remark 3 If we consider operator defined by (2.1) and operator of the following form

$$\left(\Sigma^+ g\right)_i = \sum_{j=1}^{i} \sigma_{i,j} g_j, \qquad i \geq 1.$$

Then

$$\left(A^+ \circ \Sigma^+\right)(g)_i \equiv \left(A^+ \left(\Sigma^+ g\right)\right)_i = \sum_{j=1}^{i} a_{i,j} \sum_{k=1}^{j} \sigma_{j,k} g_k$$

$$= \sum_{k=1}^{i} \left(\sum_{j=k}^{i} a_{i,j} \sigma_{j,k} \right) g_k = \sum_{k=1}^{i} w_{i,k} g_k.$$

Therefore if $(a_{i,j}) \in \mathcal{O}_n^+$, $(\sigma_{j,k}) \in \mathcal{O}_m^+$, then according to lemma 3.1 the matrix $(w_{i,k})$ of the operator $A^+ \circ \Sigma^+$ belongs to the class $\mathcal{O}_{m+n+1}^+$.

In general case, if matrices $(a_{i,j}^k)$ of operators $\left(A_k^+ f\right)_j = \sum\limits_{j=1}^{i} a_{i,j}^k f_j$ belong to the classes $\mathcal{O}_{m_k}^+$, $\quad k = 1, ..., n$, then the matrix of operator $\mathbb{A}_n^+ \equiv A_1^+ \circ A_2^+ \circ ... \circ A_n^+$ belongs to the class $\mathcal{O}_m^+$, where $m = \sum\limits_{k=1}^{n} m_k + n - 1$. So according to Theorem 4.1 we obtain criteria of boundedness and compactness of the matrix operator $\mathbb{A}_n^+$ from the weighted $l_{p,v}$ space into the weighted $l_{q,u}$ space, $1 < p \leq q < \infty$.

Similarly, if matrices $(a_{i,j}^k)$ of operators $\left(A_k^- g\right)_j = \sum\limits_{i=j}^{\infty} a_{i,j}^k g_i$ belong to the classes $\mathcal{O}_{m_k}^-$, $\quad k = 1, ..., n$, then based on lemma 3.2, Theorem 4.2 gives necessary and sufficient conditions for boundedness and compactness of operator $\mathbb{A}_n^- = A_1^- \circ A_2^- \circ ... \circ A_n^-$ from $l_{p,v}$ into $l_{q,u}$, $1 < p \leq q < \infty$.

Acknowledgements

The author thanks the University of Padova for the opportunity to have an internship, and expresses her gratitude to Professors Ryskul Oinarov and Massimo Lanza de Cristoforis for providing cooperative support and valuble assistance.

References

[1] Stieglitz M. and Tietz H., *Matrixtransformationen von Folgenräumen Eine Ergebnisübersicht.* Math. Z. 154 (1977), 1–16.

[2] Kufner L. and Persson L-E., "Weighted Inequalities of Hardy type". World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[3] Kufner A., Maligranda L. and Persson L-E., "The Hardy Inequality. About its History and Some Related Results". Vydavatelský Servis, Plzeň, 2007.

[4] Hardy G.H., *Notes on some points in the integral calculus - LX. An inequality between integrals.* Messenger of Math. 54 (1925), 150–156.

[5] Talenti G., *Osservazioni sopra una classe di disuguaglianze.* Rend. Mat. Fis. Milano 39 (1969), 171–185.

[6] Tomaselli G., *A class of inequalities.* Boll. Un. Mat. Ital. 2 (1969), 622–631.

[7] Muckenhoupt B., *Hardy's inequality with weights.* Studia Math. 44 (1972), 31–38.

[8] Bradley J.S., *Hardy inequalities with mixed norms.* Canada Math. Bull. 4 (1978), 405–408.

[9] Kokilashvili V.M.e, *On Hardy's inequalities in weighted spaces.* Soobshch. Akad. Nauk Gruzin. SSR. 96/1 (1979), 37–40.

[10] Maz'ja V.G., "Sobolev Spaces". Springer-Verlag, Berlin/Heidelberg, 1985.

[11] Otelbaev M., Apyshev O.D., *On the spectrum of a class of differential operators and some imbedding theorems.* Izvestiya Academy of Sciences of the USSR, serial Phys.-Mat. 43/4 (1979), 739–764.

[12] Otelbaev M., Mynbaev K.T., "Weighted functional spaces and the spectrum of differential operators". Moscow, Nauka., 1988.

[13] Otelbaev M., "Estimates of Spectrum of Sturm-Liouville Operator". Alma-Ata.-Gilim., 1990.

[14] Martin-Reyes F.J. and E. Sawyer, *Weighted inequalities for Riemann-Liouville fractional integrals of order one and greater.* Proc. Amer. Math. Soc. 3 (1989), 727–733.

[15] Stepanov V.D., *Weighted inequalities of Hardy type for higher derivatives, and their applications.* Doklady Academy of Sciences of the USSR 302/5 (1988), 1059–1062.

[16] Stepanov V.D., *Two-weighted estimates of Riemann-Liouville integrals.* Preprint, Computer centre of the Far East Scientific Centre of the Academy of Sciences of the USSR. Vladivostok (1988).

[17] Stepanov V.D., *Two-weighted estimates of Riemann-Liouville integrals.* Preprint, Computer centre of the Far East Scientific Centre of the Academy of Sciences of the USSR. Vladivostok (1988).

[18] Stepanov V.D., *On one weighted inequality of Hardy type for higher derivatives.* Trudy of Mathematical Institute of the Academy of Sciences of the USSR 187 (1989), 178–190.

[19] Stepanov V.D., *Two-weighted estimates for a class of convolution operators.* Preprint. Institute of Applied Mathematics of Far East Division of the Academy of Sciences of the USSR. Vladivostok (1989).

[20] Oinarov R., *Two-sided norm estimates for certain classes of integral operators.* Deposited in Kazakhstan Research Institute of Scientific and Technical Information. -27.033.91.-3354-K91.

[21] Bloom S., Kerman P., *Weighted norm inequalities for operators of Hardy type.* Proc. Amer. Math. Soc. 1 (1991), 135–141.

[22] Pankaj Jain, Pawan K. Jain and Babita Gupta, *Higher dimensional compactness of Hardy operators involving Oinarov-type kernels.* Math. Inequal. Appl. 4 (2006), 739–748.

[23] Andersen K. F., Heinig H.P., *Weighted norm inequalities for certain integral operators.* SIAM J. Math. 14 (1983), 834–844.

[24] Heinig H. P., *Weighted norm inequalities for certain integral operators. II.* Proc. Amer. Math. Soc. 95 (1985), 387–395.

[25] Bennett G., *Some elementary inequalities.* Quart. J. Math. Oxford Ser. 152 (1987), 401–425.

[26] Bennett G., *Some elementary inequalities II.* Quart. J. Math. Oxford Ser. 156 (1988), 385–400.

[27] Bennett G., *Some elementary inequalities III.* Quart. J. Math. Oxford Ser. 166 (1991), 149–174.

[28] Braverman M. Sh. and Stepanov V.D., *On the discrete Hardy inequality.* Bull. London Math. Soc. 3 (1994), 283–287.

[29] Smailov E.S., *Difference imbedding theorems for Sobolev spaces with weight and their applications.* Doklady Academy of Sciences of the USSR 270/1 (1983), 52–55.

[30] Smailov E.S., *Difference imbedding theorems for Sobolev spaces with weight.* Some applications of functional analysis to mathematical physics. Trudy seminara S.L. Soboleva. -Novosibirsk (1984), 122–143.

[31] Smailov E.S., "Fourier multipliers, embedding theorems, and related problems". Dr. Sci.(Phys.-Math.) Dissertation. Karaganda, 1997.

[32] Stikharnyi A.P., "Difference imbedding theorems and some properties of second order difference operator". (Phys.-Math.) PhD thesis. -Alma-Ata., 1991.

[33] Oinarov R., Stikharnyi A.P., *Boundedness and compactness criteria for a certain difference inclusion.* Mathematical Notes 50/5 (1991), 54–60.

[34] Oinarov R., Shalgynbaeva S. Kh., *Weighted additive estimate of a class of matrix operators.* Izvestiya NAN RK, Serial Phys.-Mat. 1 (2004), 39–49.

[35] Oinarov R., Okpoti C.A., Persson L-E., *Weighted inequalities of Hardy type for matrix operators: the case $q < p$.* Math. Inequal. Appl. 4 (2007), 843–861.

[36] Oinarov R., Persson L-E., Temirkhanova A.M., *Weighted inequalities for a class of matrix operators: the case $p \leq q$.* Math. Inequal. Appl. 12 (2009), 891–903.

[37] Temirkhanova A.M., Taspaganbetova Zh.A., *Boundedness and compactness criteria of a certain class of matrix operators.* Mathematical Journal 11, No. 2 (40) (2011), 125–139.

[38] Temirkhanova A.M., Taspaganbetova Zh.A., *Criteria on Boundedness of matrix operators in weighted spaces of sequences and their applications.* Annals of Functional Analysis 2/1 (2011), 114–127.

[39] Oinarov R., *Boundedness and compactness of integral operators of Volterra type.* Sibirskii Matematicheskii Zhurnal 48/5 (2007), 1100–1115.

What does "Inverse Problems" mean?

Giulia Deolmi (*)

Abstract. Inverse Problems (IP) are described as situations where *the answer is known, but not the question, or where the results, or consequences are known, but not the cause* (Isakov, 2006). To better understand what this means, after an introductory overview, we will focus on two particular IP's. The first one consists in the estimation of the corrosion of an unobservable face of a metal slab, while the second one is about the estimation of the quantity of pollutant released in a river.

1 Introduction

As presented in [7], inverse problems are largely used in applications, for solving, among the others, medical (e.g. in tomographical methods), industrial (e.g. monitoring oil pipelines), image analysis and mine detection problems (e.g. ground penetration radar and electromagnetic induction). The first issue consists in understanding what *inverse problem* means. Following [8], to characterize them mathematically, we present some examples.

Example 1.1 Find a polynomial p of degree n with given zeros $x_1, \dots, x_n$. Inverse problem's solution is simply $p(x) = c(x - x_1) \dots (x - x_n)$, $c \in \mathbb{R}$. The corresponding direct problem reads: find the zeros $x_1, \dots, x_n$ of a given polynomial p.

Example 1.2 [Inverse scattering problem] Find the shape of a scattering object, given the intensity (and phase) of sound or electromagnetic waves scattered by this object. The corresponding direct problem is that of calculating the scattered wave for a given object. The rigorous mathematical description could be found in [8].

Example 1.3 [Backward heat equation] Consider the one-dimensional heat equation

$$\frac{\partial u(x,t)}{\partial t} = \frac{\partial^2 u(x,t)}{\partial x^2}$$

with boundary conditions

$$u(0,t) = u(\pi,t) = 0, \ t \geq 0$$

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: gdeolmi@math.unipd.it. Seminar held on November 30th, 2011.

and initial condition

$$u(x,0) = u_0(x),\ 0 \le x \le \pi.$$

In the inverse problem one measures the final temperature distribution $u(\cdot, T)$ and tries to determine the initial temperature $u(\cdot, 0)$.

Example 1.4 [Diffusion in an inhomogeneous medium] The equation of diffusion in an inhomogeneous medium is

$$\frac{\partial u(x,t)}{\partial t} = \frac{1}{c} div(k \nabla u(x,t)),\ x \in D,\ t > 0,$$

where c is a constant and $k = k(x)$ is a parameter describing the medium. In the inverse problem one measures u and the flux $\frac{\partial u}{\partial n}$ on the boundary ∂D and tries to determine the unknown function k in D.

As noted in [8], given two normed spaces X and Y, an operator $\mathcal{K} : X \to Y$ and a measurement y, in all of these examples we can formulate the inverse problem as the solution of the equation

$$\mathcal{K}(x) = y. \tag{1}$$

In order to formulate an inverse problem, the definition of the operator $\mathcal{K}$, including its domain and range, has to be given. In general the evaluation of $\mathcal{K}(x)$ means *solving a boundary value problem for a differential equation* or *evaluating an integral*.

Definition 1.1 Let X and Y be normed spaces, $\mathcal{K} : X \to Y$ a (linear or nonlinear) mapping. The equation $\mathcal{K}(x) = y$ is called **properly posed** or **well-posed** if the following holds:

(a) Existence: for every $y \in Y$ there is at least one $x \in X$ s.t. $\mathcal{K}(x) = y$.

(b) Uniqueness: for every $y \in Y$ there is at most one $x \in X$ with $\mathcal{K}(x) = y$.

(c) Stability: the solution x depends continuously on y, i.e. for every sequence (x_n) with $\mathcal{K}(x_n) \to \mathcal{K}(x)$ as $n \to \infty$, it follows that $x_n \to x$.

Equations for which at least one of these properties does not hold are called **improperly posed** or **ill-posed**.

Usually inverse problems are *ill-posed* or *improperly posed* in the sense of Hadamard [5]: to compensate the loss of information we use some additional *a priori knowledge* about the problem, i.e. we adopt a *regularization method*.

To solve an inverse problem two different approaches can be adopted:

(a) *first optimize than discretize* strategy: given (1), first an optimization problem is defined, and then it is discretized;

(b) *first discretize than optimize* strategy: first (1) is discretized and then a discrete optimization problem is solved.

1.1 First optimize then discretize strategy

Following [7], we denote the measured perturbed data by $y^{(\delta)}$ and assume that these noisy data satisfy

$$\left\| y^{(\delta)} - y \right\| \leq \delta,$$

$\delta > 0$.

The most well-known method for solving ill-posed problems is *Tikhonov regularization*: it consists in approximating a solution of (1) by a minimizer $x_\alpha^{(\delta)}$ of

$$J(x) := \|\mathcal{K}(x) - y\|^2 + \alpha \|x - x_0\|^2, \tag{2}$$

where $x_0 \in X$ typically unifies all available a priori information on the solution and $\alpha > 0$ is the *regularization parameter*.

An alternative consists in using *Iterative regularization methods* [7]:

$$x_{k+1}^{(\delta)} = x_k^{(\delta)} + G_k(x_k^{(\delta)}, y^{(\delta)}),\ k \in \mathbb{N},$$

for various choices of G_k.

1.2 First discretize then optimize strategy

This methodology will be adopted in the following paragraphs.

First of all, we assume that the unknown $x \in X$, solution of (1), can be described by a vector $\boldsymbol{\vartheta} \in \mathbb{R}^{n_\theta}$ of non negative parameters. The way this assumption is imposed is problem dependent: we are going to see how this can be done in the following sections.

Thus we can restate the continuous problem (1) as a discrete one: given the operator $\tilde{\mathcal{K}} :\ \mathbb{R}^{n_\theta} \to \mathbb{R}^{Nn_y}$ and a measurement $\mathbf{y} \in \mathbb{R}^{Nn_y}$, find $\boldsymbol{\vartheta} \in \mathbb{R}^{n_\theta}$ such that

$$\tilde{\mathcal{K}}(\boldsymbol{\vartheta}) = \mathbf{y}. \tag{3}$$

1.2.1 Least-squares approach

To solve (3), we minimize the following functional

$$\tilde{J}(\boldsymbol{\vartheta}) := \frac{1}{N} \|\mathbf{e}_\theta\|_2^2, \tag{4}$$

where the *residual* or *prediction error* $\mathbf{e}_\theta :\ \mathbb{R}^{n_\theta} \to \mathbb{R}^{Nn_y}$, $\mathbf{e}_\theta := \tilde{\mathcal{K}}(\boldsymbol{\vartheta}) - \mathbf{y}$. This methodology is called *least squares*.

Observe that in this context *regularization* is done using an *adaptive parametrization*, as mentioned in the following sections. More details can be found in [2, 3, 4].

To solve the least-squares minimization problem, in the sequel we will adopt a line search method: starting from $\boldsymbol{\vartheta}_0$, we compute a sequence

$$\boldsymbol{\vartheta}_{k+1} = \boldsymbol{\vartheta}_k + \alpha_k \mathbf{s}_k.$$

The search direction $\mathbf{s}_k$ is computed solving the *Gauss Newton* equation

$$\psi_{\vartheta_k}^T \psi_{\vartheta_k} \mathbf{s}_k = -\psi_{\vartheta_k}^T \mathbf{e}_{\vartheta_k}, \tag{5}$$

where the *sensitivity matrix* is defined as

$$\psi_\vartheta := \left(\frac{\partial e_\theta(j)}{\partial \vartheta_i} \right)_{j=1,\dots,Nn_y;\ i=1,\dots,n_\theta}.$$

In the following sections it will be explained how this solution strategy has been applied to solve two inverse problems coming from applications.

2 An inverse problem of corrosion estimation

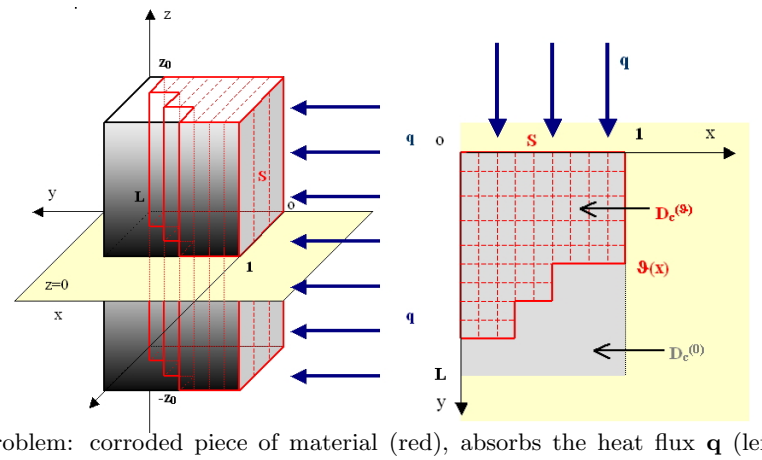


Figure 1: 3D problem: corroded piece of material (red), absorbs the heat flux $\mathbf{q}$ (left); 2D reduction, dealing with its section over $z = 0$ (right).

Suppose to deal with a metal slab, $D_c^{(0)}$, whose thickness and thermo physical properties are known, and to interact only with one face S, which is provided with n_y temperature sensors. A nondestructive test is used, consisting of an infrared thermographic inspection: in the time interval $[0, t_f]$, $t_f > 0$, S is heated with a thermal flash $\mathbf{q}(t)$ and experimental temperatures are collected. Suppose that the material surface, excluding S, is adiabatic: there is no heat exchange with the outside environment.

Consider the real *corroded domain* $D_c^{(\vartheta)}$ (cfr. the dashed domain in Figure 1), described by a scalar function $\vartheta \in \mathcal{L}^2(S)$. The corresponding PDE over $D_c^{(\vartheta)}$ is the following

$$\begin{cases} \rho C \frac{\partial}{\partial t} T^{(\vartheta)} &= k\, \Delta T^{(\vartheta)}, & in\ D_c^{(\vartheta)} \times [0, t_f] \\ k\, \nabla T^{(\vartheta)} \cdot \mathbf{n}_S &= q(t), & on\ S \times [0, t_f] \\ k\, \nabla T^{(\vartheta)} \cdot \mathbf{n} &= 0, & on\ \delta D_c^{(\vartheta)} / S \times [0, t_f] \\ T^{(\vartheta)}(0, \cdot) &= T_0(\cdot), & in\ D_c^{(\vartheta)}. \end{cases} \tag{6}$$

ρC is the heat capacity of the material, k is its thermal conductivity, and $\mathbf{n}_S$ and $\mathbf{n}$ are respectively the outward normal to S and $\delta D_c^{(0)}/S$. Suppose to know ρC, k and the heat flux $\mathbf{q}(t) = -q(t)\mathbf{n}_S$, which is assumed to be approximately a Dirac impulse in time, centered in $t = 0$, and constant over S. Consider a temporal discretization of $[0, t_f]$, $\{t_0, \dots, t_{N-1}\}$, $t_0 = 0$, $t_{N-1} = t_f$. The experimental data of the corroded model are denoted by $T_c^s \in \mathbb{R}^{n_y \times N}$, such that $(T_c^s)_{ij}$ represents the temperature in the i-th sensor at time t_{j-1}, $i = 1, \dots, n_y$, $j = 1, \dots, N$.

In the following, we assume that the corrosion does not vary along the z-axis, such that (6) can be restated as a 2D problem, considering $S = [0, 1]$ and $D_c^{(0)} = [0, 1] \times [0, L]$ (Figure 1 right). Thus we can describe analytically the corroded region in the following way:

$$D_c^{(\vartheta)} := \{(x, y) \; s.t. \; x \in [0, 1], \; 0 \le y \le L - \vartheta(x)\},$$

where $\vartheta(x) : \; [0, 1] \to [0, L]$ is a suitable smooth non negative function, such that $\vartheta(0) = 0 = \vartheta(1)$, which represents the *corrosion profile*.

First of all a particular approximation of the real corrosion profile $\vartheta(x)$ is introduced, choosing a *piecewise constant function*. This approach characterize the inverse problem from a geometrical point of view and it can be seen as additional a priori information about the problem. As explained below, under this hypothesis, instead of estimating a continuous unknown $\vartheta(x)$, we hand up with a vectorial parameter estimation problem, and thus with a *discrete inverse problem*.

To understand how this can be done, consider a subdivision of $[0, 1]$, coincident with a subset of the n_y temperature sensors' locations, with distinct spatial nodes $\{x_i\}_{i=1,\dots,n_\theta}$, $n_\theta \le n_y$, $x_0 = 0$, $x_{n_\theta} = 1$, and a uniform subdivision of $[0, L]$, with step h_y, $\{y_i\}_{i=0,\dots,n_L}$, $y_0 = 0$, $y_{n_L} = L$. Define

$$\theta_j := \frac{1}{h_c(j)} \int_{x_j}^{x_{j+1}} \vartheta(x) dx \approx L - y_k,$$

for a suitable $k \in \{0, \dots, n_L\}$, $h_c(j) := |x_{j+1} - x_j|$, $j = 1, \dots, n_\theta - 1$.

Consider now the set of functions

$$\mathcal{P} = \left\{ \tilde{\vartheta} \; s.t. \; \tilde{\vartheta} : \; [0, 1] \longrightarrow [0, L], \; \tilde{\vartheta}(x) = \sum_{j=1}^{n_\theta - 1} \theta_j \chi_{[x_i, x_{i+1})}(x) \right\},$$

where $\chi_{[x_i - x_{i+1})}(x) = \begin{cases} 1, \; x \in [x_i - x_{i+1}) \\ 0, \; \text{elsewhere} \end{cases}$ is the characteristic function of $[x_i, x_{i+1})$. The approximated corroded domain is defined as follows

$$D_c^{(\tilde{\theta})} := D_c^{(0)} \setminus \int_0^1 \tilde{\vartheta}(x) dx.$$

Thus $D_c^{(\tilde{\theta})}$ is identified by the vector of parameters $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta - 1}$.

Define now the *matrix of prediction errors* $E_\theta := T_c^s - T_h^{(\theta)} \in \mathbb{R}^{n_y \times N}$ where $T_h^{(\theta)} \in \mathbb{R}^{n_y \times N}$ denotes the Finite Element solution at every time discretization point in S' n_y nodes, solving (6) on the approximated corroded domain $D_c^{(\tilde{\theta})}$.

Consider the real valued function $\tilde{J}:\ \mathbb{R}^{n_\theta-1} \to \mathbb{R}$,

$$\tilde{J}(\boldsymbol{\theta}) := \frac{1}{N}\sum_{n=1}^{N} \|E_\theta(\cdot,n)\|_2^2 = \frac{1}{N}\sum_{k=1}^{n_y N}(e_\theta(k))^2, \tag{7}$$

where the prediction error $\mathbf{e}_\theta \in \mathbb{R}^{Nn_y}$ is such that $e_\theta((n-1)n_y+1 : nn_y) = E_\theta(\cdot,n)$, $n = 1,\dots,N$, i.e. the matrix E_θ is reshaped as a vector $\mathbf{e}_\theta$.

It corresponds to find the optimal $\tilde{\vartheta}^* \in \mathcal{P}$, or equivalently the optimal parameters θ_j^*, $j = 1,\dots,n_\theta - 1$ such that

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}\in\mathbb{R}^{n_\theta-1}} \tilde{J}(\boldsymbol{\theta}). \tag{8}$$

This least squares problem, equivalent to (4), is solved using the Gauss Newton approach. To deal with the intrinsic ill-posedness of the inverse problem of corrosion detection, an *adaptive formulation* is adopted, to reduce the computational cost. The idea is to determine the adaptive parametrization starting from an initial coarse subdivision. According to a suitable *a posteriori indicator*, the algorithm decides where eventually to refine locally the subdivision of the corrosion profile. The refinement operation corresponds to a bisection of the indicated segments, with a consequent increase in the number of segments and, therefore, of parameters of the model. More details can be found in [3, 2].

3 An inverse problem of pollution rate estimation

Let $[0,t_f) \subset \mathbb{R}$ and Ω be an open, limited and Lipschitz continuous boundary subset $\Omega \subset \mathbb{R}^2$, sufficiently regular. We denote with $\partial\Omega$ the boundary of Ω. Let $c:\ [0,t_f)\times\Omega \to \mathbb{R}$, $c = c(t,\mathbf{x})$ be the solution of the following (direct) parabolic convection-diffusion-reaction equation:

$$\begin{cases} \frac{\partial c}{\partial t} - \mu\Delta c + \nabla\cdot(\mathbf{u}c) + \sigma c &= 0, & in & (0,t_f)\times\Omega \\ c &= c_0, & on & \{0\}\times\Omega \\ c &= c_{in}, & on & (0,t_f)\times\Gamma_{in} \\ c &= c_{up}, & on & (0,t_f)\times\Gamma_{up} \\ \mu\frac{\partial c}{\partial n} &= 0, & on & (0,t_f)\times\Gamma_{down} \\ c &= 0, & on & (0,t_f)\times\Gamma_r \end{cases} \tag{9}$$

where Γ_{in}, Γ_{up}, Γ_{down} and Γ_r are given disjoint sets such that $\partial\Omega = \Gamma_{in}\cup\Gamma_{up}\cup\Gamma_{down}\cup\Gamma_r$.

Suppose that $c_{in} \in H^{\frac{1}{2}}(\Gamma_{in})$, $c_{up} \in H^{\frac{1}{2}}(\Gamma_{up})$, the initial condition $c_0 \in L^2(\Omega)$ and the coefficients are independent on time, moreover $\mu \in L^\infty(\Omega)$, $\mu(\mathbf{x}) \geq \mu_0 > 0$ for all $\mathbf{x} \in \Omega$, $\sigma \in L^\infty(\Omega)$, $\sigma(x) \geq 0$ a.e. in Ω, $\mathbf{u} \in [L^\infty(\Omega)]^2$, $div(\mathbf{u}) \in L^2(\Omega)$ are known. The *direct problem* consists in finding the concentration c over Ω at time t_f.

We assume that the physical properties of the fluid are constant and that the transported contaminant is considered as a passive scalar, which means that it does not affect the velocity field. Thus we suppose to know $\mathbf{u}$.

An example of the 2D domain Ω is illustrated in Figure 2.

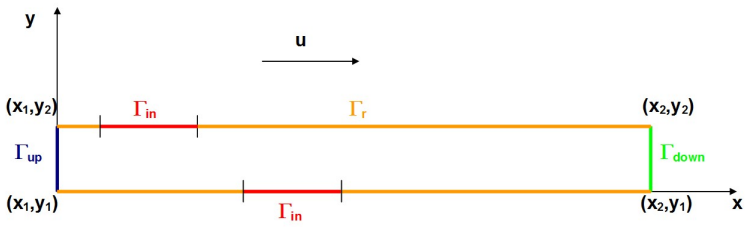


Figure 2: Example of problem's domain Ω.

Consider the set of time instants

$$\{t_j\}, \quad j = 0, \ldots, N-1. \tag{10}$$

In the *first discretize than optimize* context, we assume that $c_s(t, \mathbf{x})$ is known only in the n_y nodes of Γ_{down}, for every discrete time t_j. Let $\mathbf{C}_s(t_j) \in \mathbb{R}^{n_y}$ be the vector of measured concentration at $t = t_j$. For simplicity we suppose that

$$\Gamma_{in} = \bigcup_{l=1}^{n_\theta} \Gamma_{in}^{(l)},$$

being $\Gamma_{in}^{(l)}$ disjoint sets, such that c_{in} is constant on each $\Gamma_{in}^{(l)}$, for all $l = 1, \ldots, n_\theta$. Thus we have to estimate a vector $\boldsymbol{\vartheta}$ of n_θ non negative parameters: equivalently we assume that the function $c_{in} \in H^{\frac{1}{2}}(\Gamma_{in})$ is a piecewise constant function such that

$$c_{in}(\mathbf{x}) = \vartheta(l), \qquad \mathbf{x} \in \Gamma_{in}^{(l)}.$$

In this context the nodal vector solution is $\mathbf{C}(t_j) = \mathbf{C}(\boldsymbol{\vartheta}; t_j)$, where we have made explicit its dependence on $\boldsymbol{\vartheta}$.

Let $c_{down}(c_{in}; t, \mathbf{x}) := c(c_{in}; t, \mathbf{x})_{|\Gamma_{down}}$, be the *predicted concentration* on Γ_{down}, obtained by solving (9) imposing c_{in} on Γ_{in}, and $\boldsymbol{C_{down}}(\boldsymbol{\vartheta}; t_j)$ the corresponding nodal vector, computed at time $t = t_j$. The *discrete inverse problem* consists in finding

$$\hat{\boldsymbol{\vartheta}} = \arg \min_{\boldsymbol{\vartheta} \in \mathbb{R}_+^{n_\theta}} \tilde{J}(\boldsymbol{\vartheta}), \tag{11}$$

where the *discrete cost function* is defined as

$$\tilde{J}(\boldsymbol{\vartheta}) := \frac{1}{N} \sum_{j=1}^{N} \|\boldsymbol{C_{down}}(\boldsymbol{\vartheta}; t_j) - \mathbf{C}_s(t_j)\|_2^2 = \frac{1}{N} \sum_{k=1}^{n_y N} (e_\theta(k))^2, \tag{12}$$

where the prediction error $\mathbf{e}_\theta \in \mathbb{R}^{N n_y}$ is such that $e_\theta(j n_y + 1 : (j+1) n_y) = \boldsymbol{C_{down}}(\boldsymbol{\vartheta}; t_j) - \mathbf{C}_s(t_j)$, $j = 0, \ldots, N-1$. Observe that this is a *least squares problem* (cfr. Section 1.2.1).

If source location Γ_{in} is known then the problem can be solved applying the Gauss Newton method; whereas if Γ_{in} is unknown, also an adaptive parametrization must be adopted [4, 2].

References

[1] M. Bertero and P. Boccacci, "Introduction to Inverse problems in imaging". IOP Publishing Ltd, 1998.

[2] G. Deolmi, "Computational Parabolic Inverse Problems". Ph.D. Thesis, 2012 `http://www.math.unipd.it/~gdeolmi/PhD_thesis_Giulia_Deolmi.pdf`.

[3] G. Deolmi, F. Marcuzzi, S. Marinetti, S. Poles, *Numerical algorithms for an inverse problem of corrosion detectio*. Communications in Applied and Industrial Mathematics **1** (2010), 78–98.

[4] G. Deolmi, F. Marcuzzi, *Parabolic inverse convection-diffusion-reaction problem solved using an adaptive parametrization*. Preprint `http://arxiv.org/abs/1110.2376`.

[5] J. Hadamard, "Lectures on the Cauchy Problem in Linear Partial Differential Equations". Yale University Press, New Haven, 1923.

[6] V. Isakov, "Inverse Problems for Partial Differential Equations". Springer, 2006.

[7] B. Kaltenbacher, A. Neubauer, O. Scherzer, "Iterative Regularization Methods for Nonlinear Ill-Posed Problems". Walter de Gruyter, 2008.

[8] A. Kirsch, "An introduction to the mathematical theory of Inverse Problems". Springer, 1996.

[9] F. Marcuzzi and S. Marinetti, *Efficient reconstruction of corrosion profiles by infrared thermography*. Journal of Physics: Conference Series **124** (2008).

[10] F. Marcuzzi, *Space and time localization for the estimation of distributed parameters in a finite element model*. Computer Methods in Applied Mechanics and Engineering **198** (2009), 3020–3025.

[11] J. Nocedal, S. J. Wright, "Numerical optimization". Springer, 1999.

[12] A. N. Tikhonov, V. Y. Arsenin, "Solutions of Ill-Posed Problems". V HWinston, 1977.

Regular biproduct decompositions of objects

Nicola Girardi [(*)]

Abstract. Every vector space over a field K is the direct sum of a number of copies of the one-dimensional K-vector space K. Allowing scalars to be elements of a ring R instead of a field, we obtain a more general object called right (or left, depending on which side we write the scalars) R-module. Contrary to the trivial case of K-vector spaces, modules over R may or may not decompose into indecomposable submodules, and when they do, it is interesting to know whether their decompositions are unique in some sense or at least satisfy some sort of constraint. Beginning with the basics and with the classical results of the field we will end up giving some examples where modules have decompositions that satisfy a nice combinatorial condition. As a last step, we hint to a generalisation to the setting of biproduct decompositions in preadditive categories.

1 Introduction—The very basics

Every mathematical discourse begins with a few definitions. First let us recall the notion of a right module M over a ring R. This is an (additive) abelian group M with a mapping $M \times R \to M$ satisfying a number of properties. The image of (m, r) being denoted by $m.r$, we require that (1) $m.(r+r') = m.r + m.r'$, the former sum being in the ring R, the latter in the abelian group M, (2) $m.(rr') = (m.r).r'$, (3) $m.1 = m$, (4) $(m+m').r = m.r + m'.r$, where both sums are in the abelian group, for all $m, m' \in M$ and $r, r' \in R$. These rules probably ring a bell, in that this looks exactly like the definition of a vector space over a field K, which is certainly familiar to all readers, except for the fact that here the scalars form a ring instead of a field.[(1)] We recall that a ring has the same operations a field has—in fact, a field is a type of ring—but its properties are much less strict. For instance, elements of a ring need not be invertible (in the ring of integers $\mathbb{Z}$, only 1 and -1 are invertible), and there can be non-zero elements that multiply to zero (in the ring of integers modulo 6, the product of $\bar{2}$ and $\bar{3}$ is zero). Also, in a ring, the product of two elements generally depends on the order, that is, $a \cdot b$ and $b \cdot a$ may not yield the same result in a ring.

As it is common, after defining a mathematical structure, a notion of substructure

[(*)] Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: bruciasse@gmail.com . Seminar held on December 14th, 2011.

[(1)] Left modules are defined similarly; unless otherwise stated, all modules considered here are right modules; that is, scalars are applied on the right side of elements of the module in question.

naturally ensues.[2] Indeed, a submodule N of a module M is an abelian subgroup (that is, a non-empty subset of M closed by addition) which is closed by scalar multiples, that is, $m.r \in N$ for every $m \in N$ and every r in the ring.

Another common definition is that of "product structure." More generally, the product of finitely many structures is usually defined. In the case of a finite number of right R-modules $M_1, \ldots, M_n$, there is a right R-module $M_1 \oplus \cdots \oplus M_n$, called the *direct sum* of $M_1, \ldots, M_n$, which is the cartesian product $M_1 \times \cdots \times M_n$, endowed with pointwise operations, that is, $(m_1, \ldots, m_n) + (m'_1, \ldots, m'_n) = (m_1 + m'_1, \ldots, m_n + m'_n)$, and $(m_1, \ldots, m_n).r = (m_1.r, \ldots, m_n.r)$, where universal quantifiers are understood.

Morphisms of right R-modules are "structure-preserving mappings," as is generally the case. Precisely, in the case of modules, a mapping $g\colon M \to N$ is a *morphism* (of modules) if $g(m + m') = g(m) + g(m')$ and $g(m).r = g(m.r)$, where, again, universal quantifiers are understood. Of paramount importance, among all morphisms, are isomorphisms, as they are the mathematician's way of saying two objects are, essentially, the same. A morphism $g\colon M \to N$ is an *isomorphism* if there is another morphism $f\colon N \to M$ such that $g(f(n)) = n$ and $f(g(m)) = m$ for all $m \in M$ and $n \in N$. That is, $g \circ f$ and $f \circ g$ are the identity mappings of N and M respectively. In this situation we say that M and N are *isomorphic* modules.

2 Decompositions—The Krull-Schmidt Theorem

All of the above is quite standard in mathematics. Something which is quite standard all over the place in science is studying an entity through the study of its smaller parts, the possibly simpler parts that make up the bigger object. The study of modules is no exception to the rule. We say that M *decomposes*—but think of "separation" rather than "putrefaction"!—as the direct sum of the modules $M_1, \ldots, M_n$ if M and $M_1 \oplus \cdots \oplus M_n$ are isomorphic modules. M is said to be *indecomposable* if whenever $M \cong M_1 \oplus M_2$, either M_1 or M_2 is a singleton, that is, a zero module. Among the decompositions of a module M most important are those into indecomposable modules, that is, those where each M_i is indecomposable.

We may think of M as a molecule and view $M_1, \ldots, M_n$ as its atoms. Although, in general, modules do *not* break down into indecomposable modules as molecules break down into atoms. On the one hand we have the already mentioned vector spaces. A finite-dimensional vector space V over a field K is always of the form K^n. If we think of molecules, it's as if the entire universe were made of, say, hydrogen!

So we see that the situation for vector spaces is very simple. Modules, on the other hand, have a generally more complex behaviour. It may very well happen that, say, $H^2 \oplus O \cong M \oplus N$, where H, M, N, O are pairwise non-isomorphic. It's as if water could also be made of atoms other than two hydrogen and one oxygen atoms, which would result in a rather puzzling universe! In some cases though, modules can really be compared to molecules and atoms. The most common of such cases is that of modules of finite length. These are the modules M where all chains of submodules (read, families of submodules

[2] Here the term *structure* can in fact be intended in the technical sense of model theory.

of M totally ordered by set inclusion) are finite. For instance, the cyclic groups $\mathbb{Z}/n\mathbb{Z}$ ($n > 0$) are $\mathbb{Z}$-modules of finite length. A module of finite length has a decomposition into indecomposables (necessarily of finite length) and such decomposition is essentially unique. Precisely,

Theorem 2.1 [Krull-Schmidt, I] *Let M be a module of finite length. Then there is a decomposition $M \cong M_1 \oplus \cdots \oplus M_n$ where each M_i is an indecomposable module (of finite length), and if $M \cong N_1 \oplus \cdots \oplus N_m$, where each N_i is indecomposable, then $m = n$ and $M_i \cong N_{\sigma(i)}$ for a suitable permutation σ.*

So, to follow up on our metaphor, we have a periodic table consisting of the indecomposable modules of finite length, and every module of finite length (the molecule) uniquely determines its atoms $M_1, \ldots, M_n$, of which it is the direct sum.

The above result has been generalised in many ways. For instance, Theorem 2.1 holds if we replace every M_i and every N_i with modules whose endomorphism rings are local. A local ring S is one where for every $s \in S$, either s or $1 - s$ is invertible. Thus a local ring is not far from being a division ring. The set of endomorphisms of a module is canonically a ring; addition is carried out pointwise and multiplication is given by composition. For instance, consider the Prüfer group $\mathbb{Z}(p^\infty)$, which is the additive subgroup of $\mathbb{Q}/\mathbb{Z}$ consisting of those elements $\bar{x}$ such that $p^n\bar{x} = 0$ for some $n \geq 0$. The $\mathbb{Z}$-module $\mathbb{Z}(p^\infty)$ has a local ring of endomorphisms (in fact, isomorphic to the ring of p-adic integers $\mathbb{Z}_p$).

The last conclusion of Theorem 2.1 can be restated as follows: There is a permutation σ and a family of isomorphisms $g_i \colon M_i \to N_{\sigma(i)}$ indexed by $i = 1, \ldots, n$. Recall that every morphism $M_1 \oplus \cdots \oplus M_n \to N_1 \oplus \cdots \oplus N_m$ can be seen as a matrix of morphisms $g_{i,j} \colon M_j \to N_i$, much in the same way that we write morphisms between finite K-vector spaces as K-valued matrices. Precisely,

$$g(x_1, \ldots, x_n) = \left(\sum_{j=1}^{n} g_{i,j}(x_j) \right)_{i=1,\ldots,m}.$$

It is natural to ask whether, given an isomorphism $g \colon M_1 \oplus \cdots \oplus M_n \to N_1 \oplus \cdots \oplus N_m$ where all M_i and all N_i have a local ring of endomorphisms, it is possible to find the isomorphisms $g_i \colon M_i \to N_{\sigma(i)}$ of Theorem 2.1 in the "matrix" $(g_{i,j})_{i,j=1,\ldots,n}$. This is indeed the case [10]. In the following statement, the reader unfamiliar with additive categories can safely think of modules:

Theorem 2.2 [Krull-Schmidt, II] *Let $M_1, \ldots, M_n$ and $N_1, \ldots, N_m$ be objects with local endomorphism ring of an additive category, and let $g \colon M_1 \oplus \cdots \oplus M_n \to N_1 \oplus \cdots \oplus N_m$ be a module isomorphism. Then $n = m$ and there exists a permutation σ such that every $g_{\sigma(i),i} \colon M_i \to N_{\sigma(i)}$ is an isomorphism.*

Apart from some technical steps to embed the problem in a nice setting (an additive category where idempotents split), the proof goes by contradiction and employs Hall's Theorem from combinatorics. The details can be found in [10, Theorem 2.2]. The above result is the basis for other results that will follow on these pages.

3 Weak Krull-Schmidt Theorem

A class of modules for which there certainly is a decomposition into indecomposables is that of *artinian* modules. These are the modules whose strictly descending chains of submodules are finite. (The Prüfer group $\mathbb{Z}(p^\infty)$ is also an example of an Artinian module! In fact, the lattice of submodules of $\mathbb{Z}(p^\infty)$ is isomorphic to the natural numbers, in particular there are no infinite descending chains artinian.)

A module M is artinian if and only if every non-empty family of submodules of M has a minimal element. Using this property, the proof of the existence of a decomposition into indecomposables is rather easy: consider the family of submodules of M that do *not* have a decomposition into indecomposables; if the collection is not empty, it has a minimal element N; since N is not indecomposable, there is a decomposition $N = N_1 \oplus N_2$ where each $N_i \neq 0$. By minimality, each N_i has a decomposition into indecomposables, hence so does N, contradiction.

While proving the existence of a decomposition into indecomposables was easy, proving its uniqueness is impossible, because uniqueness doesn't actually hold in this case [7].

A module M is *uniserial* if for every pair of submodules X, Y of M, either $X \subseteq Y$ or $Y \subseteq X$. In other words, the submodules of M form a chain. (See again the Prüfer group for an example!) A uniserial module is, in particular, indecomposable. Finite direct sums of uniserial modules arise as the decompositions into indecomposables of "finitely presented modules over a serial ring" [11]. In the mentioned paper, Warfield asked whether such decomposition is unique as in Theorem 2.1, and in 1996 Facchini proved that that's not the case. Nevertheless, finite direct sums of uniserial modules enjoy an interesting property. First, suppose that M and N are uniserial modules; when there are surjective morphisms $M \to N$ and $N \to M$ we say that M and N have the same *epigeny* class, and write $[M]_e = [N]_e$. If there are injective morphisms $M \to N$ and $N \to M$, on the other hand, we say that M and N have the same *monogeny* class, and write $[M]_m = [N]_m$. It turns out that M and N are isomorphic if and only if they have the same epigeny class and the same monogeny class [3]. And here's the remarkable fact:

Theorem 3.1 [Weak Krull-Schmidt, I, [3]] *Let $M_1, \dots, M_n$ and $N_1, \dots, N_m$ be uniserial modules. Then $M_1 \oplus \cdots \oplus M_n \cong N_1 \oplus \cdots \oplus N_m$ if and only if $n = m$ and $[M_i]_e = [N_{\sigma(i)}]_e$ and $[M_i]_m = [N_{\tau(i)}]_m$ for suitable permutations σ and τ.*

So, for uniserial modules, "being isomorphic" breaks down into a conjunction of "having the same monogeny class" and "having the same epigeny class," and two direct sums are isomorphic if and only if the lists of monogeny classes and epigeny classes are the same up to order for both direct sums.

There are other classes of modules for which similar results hold, discovered relatively recently (years 2008–2010). One of these classes is that couniformly presented modules [6]. A module M is *couniformly presented* if there is a short exact sequence

$$0 \longrightarrow C \xrightarrow{f} P \xrightarrow{g} M \longrightarrow 0$$

where P is a couniform projective module and C is a couniform module. Exactness of the

sequence means that g is surjective, f is injective, and $g(y) = 0$ if and only if $y = f(x)$ for some $x \in C$. A module C is *couniform* if the sum of two proper submodules is a proper submodules.[3] We won't say here what projective modules are in general, but let's just say that the fact that P is projective allows us to find, for any given morphism $h\colon M \to M'$ between two couniformly presented modules M and M', a commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & C & \xrightarrow{f} & P & \xrightarrow{g} & M & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle h_1} & & \downarrow{\scriptstyle h_0} & & \downarrow{\scriptstyle h} & & \\ 0 & \longrightarrow & C' & \xrightarrow[f']{} & P' & \xrightarrow[g']{} & M' & \longrightarrow & 0 \end{array}$$

where it is understood that the bottom row is a couniform presentation of M'. ("Commutative diagram" here means that $hg = g'h_0$ and $h_0 f = f'h_1$.) Then one defines $[M]_\ell = [M']_\ell$ if there are morphisms $\alpha\colon M \to M'$ and $\beta\colon M' \to M$ such that α_1 and β_1 are surjective. When this happens, we say that M and M' have the same *lower part,* and write $[M]_\ell = [M']_\ell$. We've made a lot of choices here: we've chosen the two short exact sequences, and we've chosen all the morphisms α_0, α_1 and β_0, β_1, but it's not too hard to prove that the notion of lower-part is in fact well-defined. Then we have:

Theorem 3.2 [Weak Krull-Schmidt, II, [6]] *Let $M_1, \dots, M_n$ and $N_1, \dots, N_m$ be couniformly presented modules. Then $M_1 \oplus \cdots \oplus M_n \cong N_1 \oplus \cdots \oplus N_m$ if and only if $n = m$ and $[M_i]_e = [N_{\sigma(i)}]_e$ and $[M_i]_\ell = [N_{\tau(i)}]_\ell$ for suitable permutations σ and τ.*

In particular, of course, two couniformly presented modules are isomorphic if and only if have the same lower part and the same epigeny class.

Other classes of modules for which a result like Theorem 3.2 holds are biuniform modules, cyclically presented modules over a local ring, kernels of morphisms between indecomposable injective modules, for which we refer to [4], [1], and [5] respectively. All these classes of modules are related by categorical dualities and by the Auslander-Bridger transpose; for details see [6].

4 More invariants—The n-Krull-Schmidt Theorem

Let us now try to sketch how one can find a generalisation of the Weak Krull-Schmidt Theorem. Suppose to each module M that belongs to a specific class $\mathbf{C}$ we could associate a list of other modules $T_1(M), \dots, T_n(M)$, and to each morphisms $g\colon M \to N$ between two modules in $\mathbf{C}$ a list of morphisms $(T_i(g)\colon T_i(M) \to T_i(N))_{i=1,\dots,n}$ in a coherent way, that is,

(a) $T_i(g \circ f) = T_i(g) \circ T_i(f)$,

(b) $T_i(g + f) = T_i(g) + T_i(f)$, and

(c) $T_i(\text{identity of } M) = \text{identity of } T_i(M)$,

[3] Implicitly we're saying that submodules can be summed. If A and B are submodules of M, then $A+B$ is defined as the set $\{\, a + b : a \in A, b \in B \,\}$ and is the smallest submodule of M containing both A and B.

for all i and all g, f. (In other words, suppose we have a functor from a full subcategory of modules over a certain ring to a product of n full subcategories of modules over suitable rings.) Then whenever we have an isomorphism $g\colon M_1 \oplus \cdots \oplus M_s \to N_1 \oplus \cdots \oplus N_t$ we obtain isomorphisms $T_i(g)\colon T_i(M_1) \oplus \cdots \oplus T_i(M_s) \to T_i(N_1) \oplus \cdots \oplus T_i(N_t)$. If the modules $T_i(M_\nu)$ and $T_i(N_\nu)$ are nice enough, more precisely, if we assume that

(4) $T_i(M)$ has local endomorphism ring for all i and all M,

we can apply the Krull-Schmidt Theorem 2.1 and deduce that $s = t$ and there are permutations $(\sigma_i)_{i=1,\dots,n}$ such that $T_i(M_\nu) \cong T_i(N_{\sigma_i(\nu)})$. We actually know more: If we apply the refined version instead (Theorem 2.2), we draw a stronger conclusion: the permutations $(\sigma_i)_{i=1,\dots,n}$ can be chosen in such a way that $T_i(g_{\sigma_i(\nu),\nu})\colon T_i(M_\nu) \to T_i(N_{\sigma_i(\nu)})$ is an isomorphism, for all i and all ν.

After seeing pairs of equivalence relations like epigeny and monogeny, or epigeny and lower part, it is natural to come up with suitable invariants for this setting. Define equivalence relations $\{\equiv_i\}_{i<n}$ among the modules of $\mathbf{C}$ as follows: $M \equiv_i N$ if there are morphisms $g\colon M \to N$ and $f\colon N \to M$ such that $T_i(g)\colon M_i \to N_i$ and $T_i(f)\colon N_i \to M_i$ are isomorphisms. It is not hard to deduce from the above that, if $M_1, \dots, M_s$ and $N_1, \dots, N_t$ are modules in the class $\mathbf{C}$, and $M_1 \oplus \cdots \oplus M_s \cong N_1 \oplus \cdots \oplus N_t$, then $s = t$ and there are n permutations σ_i such that $M_\nu \equiv_i N_{\sigma_i(\nu)}$ (for all i, for all ν). In other words, if the two direct sums are isomorphic, the multiset (set with multiplicities) of invariants of the modules $M_1, \dots, M_s$ is the same as the multiset of invariants of $N_1, \dots, N_t$.

The converse is also true:

Theorem 4.1 [n-Krull-Schmidt, [10]] *Suppose $\mathbf{C}$ is a class of modules satisfying the four conditions above. For modules $X_1, \dots, X_s$ and $Y_1, \dots, Y_t$ in $\mathbf{C}$, the following are equivalent:*

(a) $X_1 \oplus \cdots \oplus X_s \cong Y_1 \oplus \cdots \oplus Y_t$

(b) *$s = t$ and there are permutations $(\sigma_i)_{i=1,\dots,n}$ such that $X_\nu \equiv_i Y_{\sigma_i(\nu)}$, for all i and ν.*

To prove the remaining implication some results by Facchini, Příhoda, and Perone are used [8, 9]. One of the ideas used in those papers can be extrapolated and it could be dubbed the "Chinese Remainder Theorem for preadditive categories," whose statement we include for those readers with enough background and for the most curious:

Theorem 4.2 *Let $\mathbf{C}$ be a preadditive category and $\{\mathbf{I}_i\}_{i<\kappa}$ a family of pairwise comaximal ideals of $\mathbf{C}$, and suppose that for every object X of $\mathbf{C}$ the set* $\operatorname{supp}(X) = \{i < \kappa : \mathbf{I}_i(X)$ *is proper* $\}$ *is finite. Then the canonical faithful additive functor*

$$F\colon \mathbf{C}/\bigcap_{i<\kappa} \mathbf{I}_i \to \prod_{i<\kappa} \mathbf{C}/\mathbf{I}_i$$

is also full. As a consequence, F reflects isomorphisms and retracts.

We close illustrating the n-Krull-Schmidt Theorem 4.1 with an example. We mentioned there is no uniqueness theorem for the decomposition into indecomposables of an artinian

module. Nevertheless, the n-Krull-Schmidt theorem holds for certain subclasses of artinian modules.

Let $S_1, \dots, S_n$ be pairwise non-isomorphic simple modules (= with no proper non-zero submodule) and $\mathbf{C}$ the category of artinian modules M such that $\mathrm{Soc}(M) \cong \bigoplus_{i=1}^{n} S_i$ (the socle $\mathrm{Soc}(M)$ is defined for every module, and it is characterised as the largest submodule which is a direct sum of simple modules). For $A, B \in \mathbf{C}$ let $A \equiv_i B$ (for $i = 1, \dots, n$) if there are morphisms $f\colon A \to B$ and $g\colon B \to A$ such that $\mathrm{Soc}_{S_i}(f)\colon \mathrm{Soc}_{S_i}(A) \to \mathrm{Soc}_{S_i}(B)$ and $\mathrm{Soc}_{S_i}(g)\colon \mathrm{Soc}_{S_i}(B) \to \mathrm{Soc}_{S_i}(A)$ are isomorphisms—in this case we might say that A and B are *S_i-isomorphic*. (The submodule $\mathrm{Soc}_{S_i}(M)$ is also defined for every module M, and it is characterised as the largest submodule of M which is a direct sum of copies of S_i. These are just particular cases of the "trace" of a class of modules in a certain module M, cf. [2].)

Theorem 4.3 *For modules $X_1, \dots, X_s$ and $Y_1, \dots, Y_t$ in $\mathbf{C}$, the following are equivalent:*

(a) $X_1 \oplus \cdots \oplus X_s \cong Y_1 \oplus \cdots \oplus Y_t$

(b) *$s = t$ and there are permutations $(\sigma_i)_{i=1,\dots,n}$ such that the modules X_ν and $Y_{\sigma_i(\nu)}$ are S_i-isomorphic.*

References

[1] Babak Amini, Afshin Amini, and Alberto Facchini, *Equivalence of diagonal matrices over local rings*. J. Algebra 320/3 (2008), 1288–1310.

[2] Frank W. Anderson and Kent R. Fuller, "Rings and categories of modules". Volume 13 of Graduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1992.

[3] Alberto Facchini, *Krull-Schmidt fails for serial modules*. Trans. Amer. Math. Soc. 348/11 (1996), 4561–4575.

[4] Alberto Facchini, "Module theory". Volume 167 of Progress in Mathematics. Birkhuser Verlag, Basel, 1998. Endomorphism rings and direct sum decompositions in some classes of modules.

[5] Alberto Facchini, Şule Ecevit, and M. Tamer Koşan, *Kernels of morphisms between indecomposable injective module*. Glasg. Math. J. 52/A (2010), 69–82.

[6] Alberto Facchini and Nicola Girardi, *Couniformly presented modules and dualities*. In "Advances in ring theory", Trends Math., Birkhuser/Springer Basel AG, Basel (2010) pp. 149–164.

[7] Alberto Facchini, Dolors Herbera, Lawrence S. Levy, and Peter Vámos, *Krull-Schmidt fails for Artinian modules*. Proc. Amer. Math. Soc. 123/12 (1995), 3587–3592.

[8] Alberto Facchini and Pavel Příhoda, *Endomorphism rings with finitely many maximal right ideals*. Comm. Algebra (2009).

[9] Alberto Facchini and Marco Perone, *Maximal ideals in preadditive categories and semilocal categories*. To appear (2010).

[10] Nicola Girardi, *Finite direct sums controlled by finitely many permutations*. Rocky Mountain J. Math. (2011).

[11] R. B. Warfield, Jr., *Serial rings and finitely presented modules*. J. Algebra 37/2 (1975), 187–222.

Understanding Defaults: An affine framework for the joint modelling of equity and credit risk

Juan Miguel A. Montes (*)

Based on joint work with Claudio Fontana(†)

Abstract. The possibility that a debtor may default poses a big risk to investors. Such a risk, called credit risk, is one of the risks present in financial markets. It is traditionally modelled in isolation from other kinds of risk such as the risk due to volatility, known as equity risk. In fact equity risk is observed to be connected to credit risk. In this talk, we provide a discussion of the basic credit risk models and the pricing of credit risky derivatives. We also discuss the Fourier Transform approach to pricing, developed by Carr and Madan. Using the Fourier Transform approach, we can price options under a model of risky assets proposed by Fontana, that treats credit risk and equity risk jointly. Finally we discuss the application of this approach to a defaultable Heston model. This joint work is based on [11]. (Keywords: Credit Risk Models, Stochastic Volatility, Hybrid Equity-Credit Risk Models, Heston Model with Jump-to-Default, Multi-factor Affine Models, FFT-based Option pricing.

1 Introduction

The last few years have witnessed an increasing popularity of hybrid equity/credit risk models. One of the most appealing features of such models is represented by their capability to link the stochastic behavior of the stock price (and of its volatility) with the random occurrence of the default event and, as a consequence, with the level of credit spreads. The relation between equity and credit risk is supported by strong empirical evidence (see the introductory sections of [3] and [6] for an overview of the related literature) and several studies document significant relationships between stock price volatility and credit spreads of corporate bonds and spreads of Credit Default Swaps.

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: montes@math.unipd.it. Seminar held on January 18th, 2012.

(†)University of Evry Val d'Essonne, Department of Mathematics, 23 Bd. de France, 91037, Evry (France); E-mail: fontana.claudio@libero.it.

In this paper we first provide a brief overview of a widely used framework for modelling Credit Risk, which is the framework of *Reduced Form Models* with an affine default intensity. We then propose a general framework for the joint modelling of equity and credit risk which is an extension of the basic reduced form model and which allows for a flexible correlation structure between stock price, stochastic volatility and default intensity. The proposed framework is fully analytically tractable, since it relies on the powerful technology of affine processes, and nests several stochastic volatility models which have been proposed in the literature, thereby extending their scope to a defaultable setting. Furthermore, unlike the models proposed in [3], [5] and [6], we jointly consider both the historical and the risk-neutral probability measures and, by relying on the results of [10], we explicitly solve risk-management as well as pricing problems.

2 Overview of Reduced Form Models with an Affine Intensity Process

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq \bar{T}}, P)$ be a filtered probability space over the time period $[0, \bar{T}]$, under the physical measure P. The filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq \bar{T}}$ shall be interpreted as the *investor filtration* representing the information available to investors[(4)]. We define $\Pi_{rf}(t,T)$ as the price at time t of a zero-coupon bond with unit payoff at maturity T that is *risk-free* (for example a bond issued by government that cannot default on its obligations). Under a certain P -equivalent, risk-neutral probability measure Q under which no-arbitrage prices are computed as Q-conditional expectations of future payoffs, we have that the general formula for $\Pi_{rf}(t,T)$ is given by

$$\Pi_{rf}(t,T) = E^Q\left[e^{-\int_t^T r_s ds} \mid \mathcal{F}_t\right]$$

where r_t represents the instantaneous risk-free interest rate, also called the *short rate*. In a model for default risk, we shall assume that the random variable $\tau : \Omega \to \mathbb{R}_+$ represents the random time of the default event, $\mathbb{G} = (\mathcal{G}_t)_{0\leq t\leq \bar{T}}$ is an enlarged, *global* filtration with $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$ where $\mathcal{H} = \sigma(\{\tau < u\} : u \leq t)$. We then have that $\Pi_{df}(t,T)$, the price at time t of a zero-coupon corporate (defaultable) bond with unit payoff at maturity T (for example a bond issued by a company that can possibly fail to pay its obligations at T if a default event occurs at $\tau \leq T$) is given by

$$\Pi_{df}(t,T) = \mathbb{1}_{\{\tau>t\}} E^Q\left[e^{-\int_t^T r_s ds}\mathbb{1}_{\{\tau>T\}} \mid \mathcal{G}_t\right]$$

To evaluate the above conditional expectation, one must provide further assumptions on how the default time τ is specified. The literature in credit risk is generally distinguised among two main approaches for specifying the default: *Structural models* and *Reduced - form (or Intensity-based) models*. An excellent overview of the various Credit Risk Models is found in the notes [1], the book [2] and Chapter 12 of the book [9]. In this note we shall be concerned only with *reduced-form model* specification of defaults. Unlike *structural models*, which attempt to model the default-event of a firm as the first passage time of the value process of a firm's assets with respect to a determenistic or stochastic barrier level,

[(4)] Typically in the literature, $\mathcal{F} = \mathcal{F}_\infty = \vee_{0\leq t\leq \bar{T}}\mathcal{F}_t$

reduced-form models instead provide a probabilistic meaning to the default time itself, which is seen as a stopping time under the global filtration $\mathcal{G}$ but that in general is not a stopping time under the sub-filtration $\mathcal{F} \subset \mathcal{G} = \mathcal{F} \vee \mathcal{H}$[5] representing the information available to investors. While the two modelling frameworks were originally treated as separate models in the literature, in reality they can be seen as the same model but having different assumptions about the information available to the modeller and the investor (this updated point of view , now generally accepted in the literature, was first expounded in detail in [14]).

Let the *default intensity* λ_t be defined as

$$\lim_{h \downarrow 0} \frac{Q\left[t < \tau \leq t + h \mid \mathcal{F}_t\right]}{h}. \tag{1}$$

A basic result used in reduced-form models is the following: if we assume that λ_t exists and that the process $M_t := \mathbb{1}_{\{\tau < t\}} - \int_0^{t \wedge \tau} \lambda_u du$ is a $(Q, \mathbb{G})$-martingale (i.e. it is a $(Q, \mathbb{G})$-intensity), then (see [2]) we have that

$$E^Q\left[1\tau > TZ \mid \mathcal{G}_t\right] = \mathbb{1}_{\{\tau > t\}} E^Q\left[Ze^{-\int_t^T \lambda_u du} \mid \mathcal{F}_t\right] \tag{2}$$

for every random variable Z. Hence

$$\begin{aligned} \Pi_{df}(t,T) =& \mathbb{1}_{\{\tau > t\}} E^Q\left[e^{-\int_t^T r_s ds} \mathbb{1}_{\{\tau > T\}} \mid \mathcal{G}_t\right] \\ \text{(by (2))} \qquad =& \mathbb{1}_{\{\tau > t\}} E^Q\left[e^{-\int_t^T (r_s + \lambda_s) ds} \mid \mathcal{F}_t\right]. \end{aligned}$$

The conditional expectation on the RHS of the last equation above can further be evaluated if we restrict ourself to the class of *affine models* for the intensity, that is, if $\lambda_t = \bar{\lambda} + \Lambda^\top X_t$ for $\bar{\lambda} \in \mathbb{R}_+$, and $\Lambda \in \mathbb{R}_+^m$ and X_t an m-dimensional affine diffusion process of the same class as in [8] whose drift and diffusion parameters satisfy certain non-degeneracy conditions as in [7] (see also Chapter 10 of [9]). In such a case, scaled conditional characteristic functions of the type $f(t,T,u,X_t) = E\left[e^{-\int_t^T \lambda_s ds} e^{u^\top X_t} \mid \mathcal{F}_t\right]$ are given in the form

$$f(t,T,u,X_t) = e^{A(t,T) + B(t,T)^\top X_t} \tag{3}$$

where $A(t,T)$ and $B(t,T)$ are deterministic functions (in $\mathbb{R}$ and $\mathbb{R}^m$ respectively) that are solutions of certain ODE's of Riccati type.

An important quantity observed in financial markets that is often considered as a reference quantity for understanding the credit risk present in markets is the so-called *Credit Spread* $CS(t,T)$ associated to a default-free and defaultable bond pair $\Pi_{rf}(t,T)$, $\Pi_{df}(t,T)$ which is defined as

$$CS(t,T) = -\frac{1}{T-t} \log \frac{\Pi_{rf}(t,T)}{\Pi_{rf}(t,T)}\ .$$

Thus an affine intensity-based model provides us with a probabilistic model for the credit spread where, as it can be easily seen, $CS(t,T)$ is given explicitly up to the evaluation of

[5] $\mathcal{H}$ represents the information due to the history of the defaults.

scaled characteristic function of the type (3). Such an approach is followed for example in [12] with applications to filtering under incomplete information and the inverse problem of model parameter estimation based on data observed on the financial markets.

3 Proposed Joint Modelling framework

Let $\left(\Omega, \mathcal{G}, (\mathcal{F}_t)_{0\le t\le T}, P\right)$ be a given filtered probability space, where P denotes the physical (or historical) probability measure and $(\mathcal{F}_t)_{0\le t\le T}$ is the filtration generated by an $\mathbb{R}^d$-valued Brownian motion $(W_t)_{0\le t\le T}$. Let $\tau : \Omega \to [0,T]\cup\{\infty\}$ be a random time which represents the *default time* of a given firm. We assume that τ is a *doubly stochastic random time* (in the sense of [9], Sect. 12.3.1) with stochastic P-*intensity* $\left(\lambda_t^P\right)_{0\le t\le T}$. Let the filtration $(\mathcal{G}_t)_{0\le t\le T}$ be the *progressive enlargement* of $(\mathcal{F}_t)_{0\le t\le T}$ with respect to τ and let $\mathcal{G}=\mathcal{G}_T$. Intuitively, the filtration $(\mathcal{F}_t)_{0\le t\le T}$ contains only the default-free market information, while the enlarged filtration $(\mathcal{G}_t)_{0\le t\le T}$ contains the full market information.

Let us denote by S_t the price at time $t\in[0,T]$ of one share issued by the defaultable firm and let $\tilde{S}_t$ be the corresponding *pre-default* value, i.e. $S_t = \mathbf{1}_{\{\tau>t\}}\tilde{S}_t$, for all $t\in[0,T]$. This corresponds to assuming that the stock price process jumps to zero as soon as the default event occurs and remains thereafter frozen at zero. Let $(v_t)_{0\le t\le T}$ denote the stochastic volatility of the stock and let $(X_t)_{0\le t\le T}$ be an $\mathbb{R}^{d-2}$-valued stochastic factor process which describes the evolution of the economy. Let also $L_t := \log\tilde{S}_t$ and $V_t := (v_t, X_t, L_t)$. We model the $\mathbb{R}^d$-valued process $(V_t)_{0\le t\le T}$ as the solution to the following SDE:

$$dV_t = (AV_t + b)\,dt + \Sigma\sqrt{R_t}\,dW_t \qquad V_0 = (v_0, X_0, \log S_0) \tag{4}$$

with parameters $(A,b,\Sigma)\in\mathbb{R}^{d\times d}\times\mathbb{R}^d\times\mathbb{R}^{d\times d}$ and where R_t is a diagonal $(d\times d)$-matrix with elements given by $R_t^{i,i} = \alpha_i + \beta_i^\top V_t$. Under suitable assumptions (see e.g. [9], Chpt. 10), there exists a unique strong solution to (4) on $\mathbb{R}^m_{++}\times\mathbb{R}^{d-m}$, for some $m\in\{1,\dots,d-1\}$. Furthermore, it can be easily shown that:

$$dS_t = S_{t-}\left(\bar{s}+\mu_1\log S_{t-}+\mu_2 v_t+\sum_{i=1}^{d-2}\gamma_i X_t^i\right)dt + S_{t-}\sum_{i=1}^{d}\Sigma_{d,i}\sqrt{R_t^{i,i}}dW_t^i - S_{t-}d\mathbf{1}_{\{\tau\le t\}} \tag{5}$$

for suitable parameters $\bar{s},\mu_1,\mu_2$ and γ, thus giving rich and flexible dynamics to the defaultable stock price process. For simplicity, we suppose that the risk-free interest rate is deterministic and equal to $r>0$. The P-intensity $\left(\lambda_t^P\right)_{0\le t\le T}$ of the default time τ is modeled as follows:

$$\lambda_t^P = \bar{\lambda}^P + \left(\Lambda^P\right)^\top V_t \qquad \text{for all } t\in[0,T] \tag{6}$$

with $\bar{\lambda}^P\in\mathbb{R}_+$ and $\Lambda^P\in\mathbb{R}^m_+\times\{0\}^{d-m}$. Observe that the specification (4)-(6) allows for both direct and indirect interactions between the stock price process, its stochastic volatility and the default intensity. Note also that so far everything has been specified with respect to the original (historical) probability measure P.

4 Risk-management applications

By relying on the affine structure of the general framework outlined in Section 3, we derive some simple results which can be of interest for risk-management purposes. As a preliminary, let us recall that for all $t \leq T' \leq T$ and $u \in \mathbb{C}^d$ (up to some technical conditions, see e.g. [9], Thm. 10.4):

$$E^P\left[e^{-\int_t^{T'}\lambda_s^P ds}\, e^{u^\top V_{T'}}\Big|\mathcal{F}_t\right] = e^{\Phi^P(T'-t,u)+\Psi^P(T'-t,u)^\top V_t} \tag{7}$$

where the functions $\Phi^P : [0,T]\times\mathbb{C}^d \to \mathbb{R}$ and $\Psi^P : [0,T]\times\mathbb{C}^d \to \mathbb{R}^d$ are given as solutions of a system of Riccati ODEs. As a first application, we can explicitly compute the $\mathcal{G}_t$-conditional probability of surviving until $T' \in [t,T]$ by noting that $P(\tau > T'|\mathcal{G}_t) = \mathbf{1}_{\{\tau>t\}} E^P\left[\exp\left(-\int_t^{T'}\lambda_s^P ds\right)|\mathcal{F}_t\right]$ and letting $u=0$ in (7).

We show that many quantities of interest in view of risk-management applications can be computed explicitly after changing the measure from P to $P^{T'}$, where $P^{T'}$ denotes the *T'-survival measure*, defined as follows:

$$dP^{T'}/dP := e^{-\int_0^{T'}\lambda_s^P ds}\left(E^P\left[e^{-\int_0^{T'}\lambda_s^P ds}\right]\right)^{-1} \qquad \text{for } T' \leq T \tag{8}$$

Note that, for $t \leq T' \leq T$, the $\mathcal{F}_t$-conditional characteristic function of $V_{T'}$ under the measure $P^{T'}$ can be easily obtained from (7). In particular, by relying on Fourier inversion techniques, we derive an explicit expression for the quantiles of the $\mathcal{G}_t$-conditional distribution (under the historical probability P) of the defaultable stock price at a given future date $T' \leq T$ in terms of the $\mathcal{F}_t$-conditional characteristic function of $V_{T'}$ under the T'-survival measure $P^{T'}$. This result will be important for the computation of *Value-at-Risk* and other risk measures.

5 Valuation of default-sensitive derivatives

Aiming at the valuation of default-sensitive financial derivatives, we need to shift our model from P to some *Equivalent Local Martingale Measure (ELMM) Q*. It can be shown that all densities dQ/dP admit the following representation:

$$\frac{dQ}{dP} = \exp\left(\sum_{i=1}^d \int_0^T \theta_t^i dW_t^i - \frac{1}{2}\sum_{i=1}^d \int_0^T \left(\theta_t^i\right)^2 dt - \int_0^{\tau\wedge T} \gamma_t \lambda_t^P dt\right)\left(1 + \mathbf{1}_{\{\tau\leq t\}}\gamma_\tau\right) \tag{9}$$

where, due to (5), the *risk-premia* processes $(\theta_t)_{0\leq t\leq T}$ and $(\gamma_t)_{0\leq t\leq T}$ satisfy P-a.s. the following condition, for all $t \in [0, T\wedge\tau]$:

$$\bar{s} + \mu_1 \log S_{t-} + \mu_2\, v_t + \sum_{i=1}^{d-2}\gamma_i X_t^i + \sum_{i=1}^d \Sigma_{d,i}\sqrt{R_t^{i,i}}\theta_t^i - \lambda_t^P\,(1+\gamma_t) = r \tag{10}$$

Note that the process $(\gamma_t)_{0\leq t\leq T}$ represents the risk-premium associated to the default event and accounts for the non-difersifiability of default risk, see e.g. [13].

By relying on [10], we provide a full characterisation of all ELMMs which preserve the affine structure (in the sense that the specification (4)-(6) holds after the change of measure) by giving necessary and sufficient conditions on the processes $(\theta_t)_{0\le t\le T}$ and $(\gamma_t)_{0\le t\le T}$ satisfying (9)-(10). Working under an *affine preserving* ELMM will ensure analytical tractability under both the physical and the risk-neutral probability measures. Then, similarly as in Section 4, we show that many pricing problems can be simplified by shifting the model to the T'*-survival risk-neutral measure* $Q^{T'}$, with:

$$dQ^{T'}/dQ := e^{-\int_0^{T'}\lambda_s^P(1+\gamma_s)ds}\left(E^Q\left[e^{-\int_0^{T'}\lambda_s^P(1+\gamma_s)ds}\right]\right)^{-1} \qquad \text{for } T'\le T \tag{11}$$

Let us consider an European defaultable derivative with maturity $T'\le T$ and payoff $F(V_{T'})$ in the case of survival. Then we have the following pricing formula:

$$\begin{aligned} &E^Q\left[e^{-r(T'-t)}F(V_{T'})\,\mathbf{1}_{\{\tau>T'\}}\middle|\mathcal{G}_t\right] \\ &\qquad = \mathbf{1}_{\{\tau>t\}}e^{-r(T'-t)+\Phi^Q(T'-t,0)+\Psi^Q(T'-t,0)^\top V_t}\times\times E^{Q^{T'}}[F(V_{T'})\,|\mathcal{F}_t] \end{aligned} \tag{12}$$

where the functions $\Phi^Q\colon[0,T]\times\mathbb{C}^d\to\mathbb{R}$ and $\Psi^Q\colon[0,T]\times\mathbb{C}^d\to\mathbb{R}^d$ are given as solutions of a system of Riccati ODEs. Since the $\mathcal{F}_t$-conditional characteristic function of $V_{T'}$ under the measure $Q^{T'}$ can be obtained explicitly, the quantity $E^{Q^{T'}}[F(V_{T'})\,|\mathcal{F}_t]$ can be computed in semi-closed form by relying on Fourier inversion techniques. In particular, we derive explicit expressions for the prices of corporate defaultable bonds and *Call* and *Put* options written on a defaultable stock together with a defaultable version of the classical *Put-Call parity* relation.

6 An example: the Heston with jump-to-default model

We illustrate the essential features of our general framework in the context of a simple example, which corresponds to a *jump-to-default* extension of the classical Heston stochastic volatility model. More specifically, with $d=3$, we consider the following specification of (4):

$$A=\begin{pmatrix}-k & 0 & 0\\ 0 & -k_0 & 0\\ -1/2 & 0 & 0\end{pmatrix}\quad b=\begin{pmatrix}k\bar v\\ k_0\,x\\ \mu\end{pmatrix}\quad \Sigma=\begin{pmatrix}\bar\sigma & 0 & 0\\ 0 & \sigma_0 & 0\\ \rho & 0 & \sqrt{1-\rho^2}\end{pmatrix}\quad R_t=\begin{pmatrix}v_t & 0 & 0\\ 0 & X_t & 0\\ 0 & 0 & v_t\end{pmatrix}$$

with $k\bar v\ge\bar\sigma^2/2$, $k_0x\ge\sigma_0^2/2$ and $\rho\in[-1,1]$. The default P-intensity $\left(\lambda_t^P\right)_{0\le t\le T}$ is specified as in (6). This specification extends the Heston jump-to-default model considered in [5] by allowing the default intensity to be a function of v_t and of an additional stochastic factor X_t.

By specialising the general results of Sections 4-5, we are able numerically investigate the following issues:

(i) the impact of stochastic volatility and default risk on the $\mathcal{G}_t$-conditional distribution of the defaultable stock price;

(ii) the impact of default risk and of different specifications of the default intensity on the implied volatility surface of European vanilla options written on the defaultable stock.

To this end, we computed Call prices written on the defaultable stock price S_t of this example, where the option is issued by the defaultable firm itself. We recall that such a Call price in our setting is given by a formula computed by a standard Fourier inversion technique. This was implemented in MATLAB, computing also the related *Implied Volatility surfaces*. The results were then compared to those of a Call option on an underlying stock with the same Heston-type dynamics but which is default-free, and issued by a default-free agent. The surface plots of the Call prices and the Implied Volatilites are shown in Figure 1 and Figure 2 respectively. Figure 3 shows the relative difference between the Call price on a defaultable underlying and the Call price on a default-free underlying.

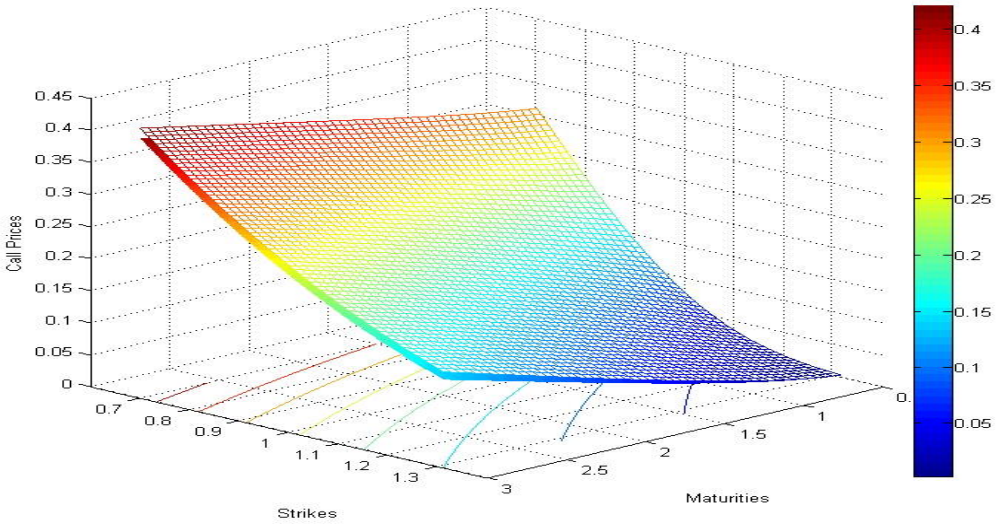


Figure 1: Call price surfaces under a Standard Heston Model (solid surface) and under a defaultable Heston Model (mesh surface). The fixed paramaters are: $k = 0.06$, $x = 1$, $\bar{\sigma} = 0.2$, $k_0 = 0.04$, $\bar{v} = 0.3$, $\sigma_0 = 0.1$, $\rho = -0.6$, $\bar{\mu} = 0$, $\bar{r} = 0.01$, $\Theta_{1,1} = \Theta_{2,2} = \hat{\theta} = 0$; For the defaultable underlying, $\Lambda_1^Q = 0.01$, $\Lambda_2^Q = 0.01$, $\bar{\lambda}^Q = 0.01$, while for the default-free underlying $\Lambda_1^Q = 0$, $\Lambda_2^Q = 0$, $\bar{\lambda}^Q = 0$.

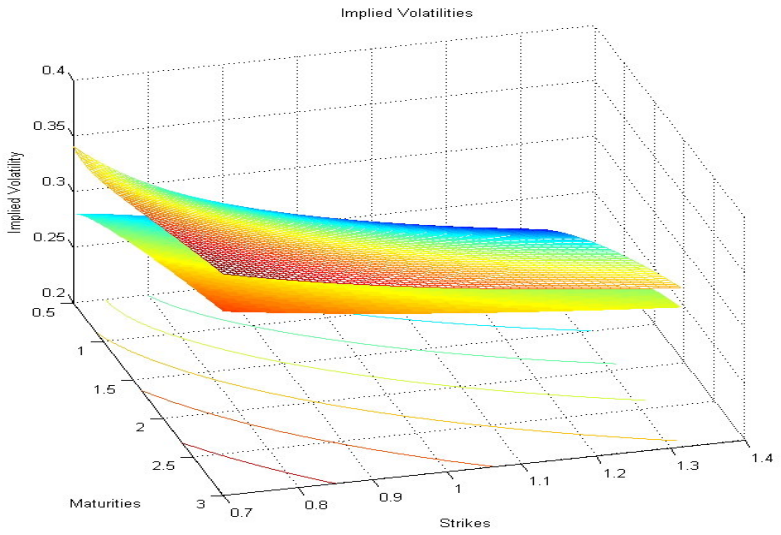


Figure 2: Implied Volatility surfaces corresponding, respectively, to the Call Prices under a standard Heston model (solid surface) and under a defaultable Heston model (mesh surface) from Figure 1. For the computation of the Implied Volatility, we used *blsimpv*, a program that is part of the Financial Toolbox package of MATLAB.

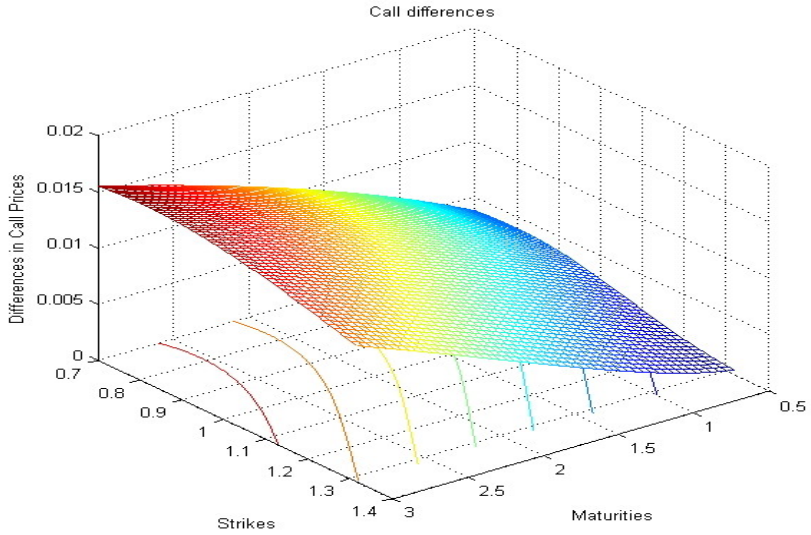


Figure 3: The surface plot of the relative differences $\tilde{C}(K,T) - C(K,T)$, where $\tilde{C}(K,T)$ are the Call Prices under a defaultable Heston model (mesh surface in Figure 1) and $C(K,T)$ are the Call Prices under a standard Heston model (solid surface in Figure 1) for each value K and T of strikes and maturities.

Based on Figures 1 - 3, we see that our model framework captures the expected behavior of a Call price when the underlying is made to be defaultable: the Call prices and the implied volatilities become higher than the analogous Call prices for options on default-free stocks.

An important future work is to establish, by means of a simple example, that our model is particularly well-suited to the analysis of problems which involve simultaneously both the historical and the risk-neutral probability measures. In particular, possible applications to the valuation of *mortality-linked* insurance products and credit rating trigger swaps are of interest.

References

[1] Bielecki, T. R., Jeanblanc, M. and Rutkowski M., "Credit risk modelling". Lecture Notes, Center for the Study of Finance and Insurance, Osaka University, Osaka, Japan, 2009.

[2] Bielecki, T. R. and Rutkowski M., "Credit Risk: Modeling, Valuation and Hedging". Springer, 2002.

[3] Carr, P. and Linetsky, V., *A jump to default extended CEV model: An application of Bessel processes.* Finance and Stochastics 10 (2006), 303–330.

[4] Carr, P. and Madan, D. B., *Option Valuation using the Fast Fourier Transform.* Journal of Computational Finance 2/4 (1999), 61–73.

[5] Carr, P. and Schoutens, W., *Hedging under the Heston model with jump-to-default.* International Journal of Theoretical and Applied Finance 11/4 (2008), 403–414.

[6] Chung, T. K. and Kwok, Y. K., *Equity-credit modeling under affine jump-diffusion models with jump-to-default.* Forthcoming in: Chen, S. and Kaboudan, M. (eds.), "Handbook on Computational Economics and Finance", Oxford University Press (2012).

[7] Duffie, D. and Kan, R., *A Yield-factor Model of Interest Rates.* Mathematical Finance 6/4 (1996), 379–406.

[8] Dai, Q. and Singleton, K.J., *Specification Analysis of Affine Term Structure Models.* Journal of Finance, LV/5 (2000), 1943–1978.

[9] Filipović, D., "Term-Structure Models: A Graduate Course". Springer Berlin-Heidelberg, 2009.

[10] Fontana, C., *Measure changes for reduced-form affine credit risk model.* Preprint (2011).

[11] Fontana, C. and Montes, J. M., *An affine framework for the joint modelling of equity and credit risk.* International Conference on Applied and Computational Mathematics 2012, Ankara, Turkey. Accepted for presentation.

[12] Fontana, C. and Runggaldier, W. J., *Credit Risk and Incomplete Information: Filtering and EM Parameter Estimation.* International Journal of Theoretical and Applied Finance 13/5 (2010), 683–715.

[13] Jarrow, R. A., Lando, D. and Yu, F., *Default risk and diversification: Theory and empirical implications.* Mathematical Finance 15/1 (2005), 1–26.

[14] Jarrow, R. A. and Protter, P., *Structural versus reduced form models: A new information based perspective.* Journal Of Investment Management 2 (2004), 1–10.

The Fermi Pasta Ulam Paradox: An Introduction to the problem and its recent developments

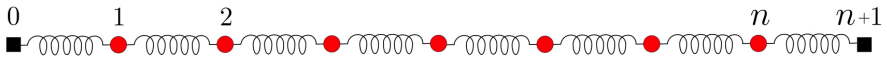
Heleni Christodoulidi (*)

Abstract. In 1950's Fermi, motivated by fundamental questions of statistical mechanics and by ergodic properties of nonlinear dynamical systems, started a numerical experiment on a nonlinear toy model, in collaboration with Pasta and Ulam. The system was chosen to be a one dimensional chain of N nonlinear coupled oscillators, described by a quadratic potential of nearby particle interactions plus a cubic perturbation. Fermi's ergodic hypothesis states that a system under an arbitrarily small perturbing force becomes generically ergodic. Giving initial conditions to the so called FPU system on the longest wavelength normal modes, the system showed a contradicting and integrable-like behaviour. Many pioneer works followed for the explanation of this paradox, more prominent of them being the approximation of FPU dynamics in the thermodynamic limit by the integrable partial differential equation Korteweg-de Vries by Zabusky and Kruskal (1965), and the work of Ferguson et al. (1982), where the authors observed the vicinity of the integrable one dimensional Toda chain with the FPU model. Recent developments show a more complete picture of the problem and its explanation.

1 Introduction

In 1954 Fermi, Pasta and Ulam [1] made a numerical experiment on a dynamical system that consists of N coupled oscillators in order to test the 'ergodic hypothesis'. Assuming that an arbitrary nonlinear system behaves as ergodic, especially for systems with many degrees of freedom, as statistical mechanics indicate, they chose a one dimensional chain of N nonlinear coupled oscillators, described by a quadratic potential of nearby particle interactions plus a cubic perturbation. Despite the ergodic expectations, the numerical integration of the system's dynamics showed an integrable behaviour, resulting to the famous Fermi – Pasta – Ulam (FPU) paradox.

(*)Università di Padova, Dipartimento di Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: echrist@math.unipd.it. Seminar held on February 1st, 2012.



In particular, the system that Fermi, Pasta and Ulam considered was the one dimensional lattice of N moving particles with fixed boundary conditions, described by the Hamiltonian

$$H = \frac{1}{2}\sum_{k=1}^{N} y_k^2 + \frac{1}{2}\sum_{k=0}^{N}(x_{k+1}-x_k)^2 + \frac{\alpha}{3}\sum_{k=0}^{N}(x_{k+1}-x_k)^3 \quad , \tag{1}$$

where x_k is the k–th particle's position with respect to equilibrium and y_k its canonically conjugate momentum and it is called the *FPU–α model*. Fixed boundary conditions are defined by setting $x_0 = x_{N+1} = 0$.

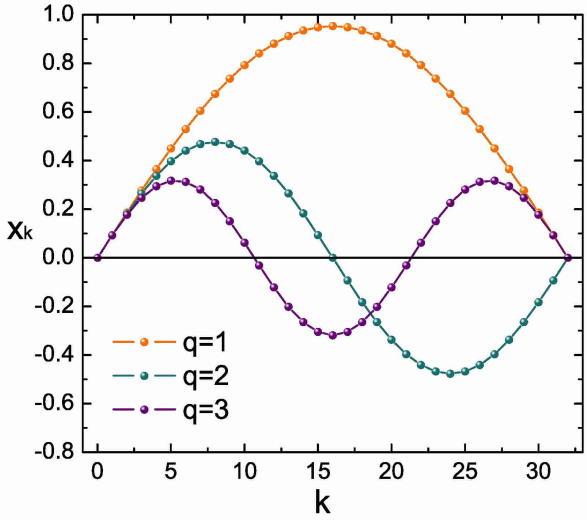


Figure 1: Initial conditions for x_k that correspond to the excitation of the first (orange), second (green) and third (purple) normal modes, when $y_k = 0$, $k = 1, ..., N$.

Before explaining the FPU experiment in detail, let us discuss the integrable system of N linearly coupled oscillators, given by Eq. (1) for $\alpha = 0$. The normal mode canonical variables (Q_q, P_q), $q = 1, \ldots, N$ are introduced by the linear canonical transformations

$$\begin{aligned} x_k &= \sqrt{\frac{2}{N+1}}\sum_{q=1}^{N} Q_q \sin\left(\frac{qk\pi}{N+1}\right), \\ y_k &= \sqrt{\frac{2}{N+1}}\sum_{q=1}^{N} P_q \sin\left(\frac{qk\pi}{N+1}\right), \quad k = 1, \ldots, N \end{aligned} \tag{2}$$

that uncouple the system into N independent harmonic oscillators

$$\ddot{Q}_q = -\Omega_q^2 Q_q, \quad q = 1, \ldots, N$$

where q is an integer identifying the q–th normal mode and Ω_q are the harmonic frequencies of the system

$$\Omega_q = 2 \sin \frac{q\pi}{2(N+1)}, \quad q = 1, \ldots, N \quad . \tag{3}$$

The constants of motion for Eq. (1) with $\alpha = 0$ are the harmonic energies $E_q = (P_q^2 + \Omega_q^2 Q_q^2)/2$ of each normal mode q. Let us point out here that *excitation of the q–th mode* can be obtained by any initial condition of the form $x_k = A \sin\left(\frac{qk\pi}{N+1}\right)$, $y_k = B \sin\left(\frac{qk\pi}{N+1}\right)$, $k = 1, \ldots, N$ and corresponds to a simple periodic orbit of the linear system $\alpha = 0$. In Fig. 1 are given 3 examples of mode excitations, for the modes $q = 1$ (orange), 2 (green) and 3 (purple).

'Switching on' the α parameter in the FPU system, the harmonic energies E_q for each normal mode q are not anymore constants of motion. If a system is ergodic, then any scalar function with domain the phase space of the system has an average over time, as $t \to \infty$, which is equal to its space average. If the FPU system is not integrable, the same property should hold for the modal energies $E_q(t)$, i.e.

$$\overline{E_q} = E_q^* \qquad \Leftrightarrow \qquad \lim_{T\to\infty} \frac{1}{T} \int_0^T E_q(t)dt = E/N, \quad q = 1, \ldots, N \quad . \tag{4}$$

where $\overline{E_q}$ is the time average of the harmonic energy function $E_q(t)$, E_q^* is the space average and $\varepsilon = E/N$ the specific energy.

In the classical experiment, the authors started with the initial condition on the first normal mode, as shown in Fig. 1 for $q = 1$, and expected to observe thermalization of the medium, i.e. that the energy of each normal mode q will tend to the value of the specific energy $\varepsilon = E/N$ of the system (relation (4)). This means that the energy of the first mode $E_1(0) = E$ is expected to diffuse to the rest ones, leading to *equipartition of energy* between all the normal modes of the system. Plotting the evolution of $E_q(t)$, $q = 1, \ldots, N$, as in Fig. 2, the experiment's results were very surprising. The first mode looses energy, which flows to only a few nearby ones and this evolution is reversed shortly after, with the energy flowing backwards to the first normal mode, yielding the famous *FPU recurrences* of the modal energies, a phenomenon that is periodically repeated until the available numerical integration time. This contradicting integrable–like behaviour consists the so–called *FPU paradox*, that although has been extensively studied, remains an open problem.

2 One problem, many approaches

In this section we are going to briefly outline the progress made in explaining the FPU paradox. Most approaches to the problem concern a comparison of the FPU model with a corresponding integrable one, that is achieved either numerically or by applying techniques

of perturbation theory. This comparison can be classified into three main categories: FPU can be regarded as (i) a perturbed system of harmonic coupled oscillators, (ii) a perturbed Toda system and (iii) a system that in the thermodynamic limit $E \to \infty$, $N \to \infty$ with $\varepsilon = E/N = const.$ is approximated by the integrable partial differential equation Korteweg–de Vries (KdV).

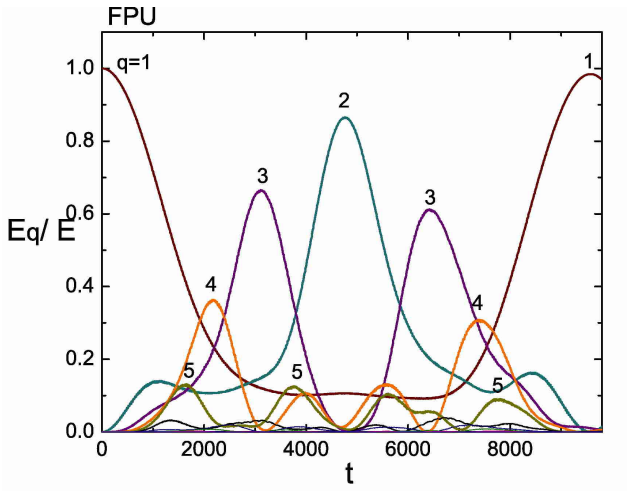


Figure 2: The evolution of the harmonic energies $E_q(t)/E$, $q = 1, \ldots, N$ for the 32–particle FPU chain, with $E = 0.07$ and $\alpha = 0.33$.

2.1 FPU as a perturbed linear system

This is the same case with the original one of Fermi, Pasta and Ulam, described extensively in the Introduction, where the Hamiltonian of the system is expressed by the Hamiltonian of the linear system plus a cubic perturbation, given by Eq. (1).

A very promising and challenging study was the implementation of KAM theorem, which apparently was discovered almost the same time with the FPU experiment. The most important work towards this direction was [3] by B. Rink and F. Verhulst, who prove the existence of KAM tori in the phase space of the FPU system, but for an energy extremely close to the ground state and with an N–dependence, implying that the range of existence of these tori tends to zero, as $N \to \infty$. Contrariwise, these results do not agree with numerical findings, as for example in [7], which indicate that the system, even towards the thermodynamic limit, depends on the specific energy ε and not on the energy E. Furthermore, new results [19] concerning the time T_{eq} the system needs to reach equipartition[(6)] show again a dependence of T_{eq} on ε.

Numerical integration of FPU dynamics indicates that when the energy is given initially to a low–frequency mode, the flow of energy from this mode to the unexcited ones may appear 'freezed' for quite long times, as observed in Fig. 3, resulting to the so–called

[(6)] The question of energy equipartition is one of the most fundamental in the FPU problem, therefore we suggest the reader to study [19], as well as the references therein.

"metastable state" [4], [5], [8], [9], in which the system exhibits important deviations from energy equipartition. Furthermore, by observing the averaged in time energy spectrum of the system is understood that it is exponentially localized in modal space, characterized by the energy concentration in one (the initially excited one) or more modes, accompanied by a tail of the remaining modes that is exponentially localized in normal mode space.

This metastable state, when concerns the excitation of a single low–frequency mode, can take qualitatively two different forms, depending on the system's parameters $\mu = \alpha^{1/2}\varepsilon^{1/4}$ and N: $\mu N < 1$ and $\mu N > 1$.

In the first case, which refers to lower energy values for an N–dimensional system, important studies [12]–[16] show that orbits rising by the excitation of one mode lie close to simple periodic orbits, named *q–breathers*, obtained by the continuation of the q–th mode excitation of the linear system. The spectra of both orbits are almost identical, with the energy of the system concentrated on the q–th mode, leaving the rest modes exponentially localized in normal mode space with a slope that depends logarithmically on μN. It immediately derives that this 'q–breather regime' vanishes in the thermodynamic limit. An example of a q–breather profile is shown in Fig. 3 (a) by the blue spheres, taken by the excitation of the first mode, for the system with $N = 32$, $E = 0.00017$ and $\alpha = 0.33$. The line in Fig. 3 (a) corresponds to the q–breather profile given in [12] with slope equal to $4\log(\mu N/\pi)$.

The main feature of the second case is the appearance of an approximate 'plateau' in the lower–frequency part of the energy spectrum, named as *natural packet* of modes [10], [11], leaving the rest ones exponentially localized, but for a slope that depends linearly on μ, i.e. that is independent of N. The percentage of modes entering the natural packet is approximately equal to μ. Fig. 2, which shows the FPU recurrences, belongs in this case and has an averaged 'freezed' profile that appears in Fig. 3 (a) by the orange triangles.

Another concept related to q–breathers is the concept of q–tori, introduced in [17], [18] as the low–dimensional tori, that rise from the continuation of a percentage of modes s/N of the linear system. Main feature of q–tori is the concentration of energy on these s modes, accompanied by an exponential tail which has a slope that depends logarithmically on $\mu N/s = \alpha^{1/2}\varepsilon^{1/4}N/s$. Eventhough the existence of these tori is not formally proved, numerical results show that consist a prominent direction. New expectations, in explaining system's behaviour in the thermodynamic limit, rise under the framework of q–tori, since the slope is invariant as long as the specific energy ε and the percentage s/N are kept fixed. Thus, it appears that this 'q–tori regime' is $\mu N/s << 1$, which is independent of N, but implies that these results hold true for rather small specific energies of the system.

2.2 FPU as a perturbed Toda system

Ferguson et al. [6] in 1982 consider the Toda Hamiltonian

$$H_T = \sum_{k=1}^{N} y_k^2 + \frac{1}{4\alpha^2}\sum_{k=0}^{N} e^{2\alpha(x_{k+1}-x_k)} - \frac{N+1}{4\alpha^2} \quad , \tag{5}$$

which can be regarded as an approximation of the FPU–α Hamiltonian of order α^2, since

$$H_T = H_{FPU} - \sum_{k=1}^{N} \sum_{n=4}^{\infty} \frac{2^{n-2}\alpha^{n-2}}{n!}(x_{k+1} - x_k)^n \quad . \tag{6}$$

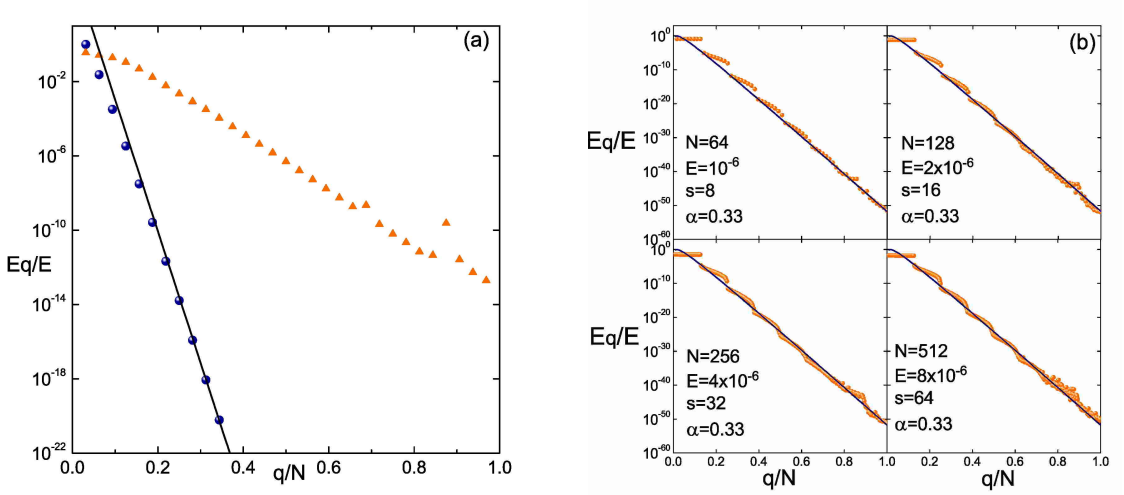


Figure 3: Normalized and averaged in time energy spectra versus q/N. (a) Excitation of the first mode for $\mu N = 0.86$ with blue spheres and for $\mu N = 3.38$ with orange triangles of the 32–particle system. The continuous line corresponds to the q–breather law $E_{q+1}/E = (\mu N/\pi)^{4q} q^2$. (b) 4 examples with $\varepsilon = E/N$ and percentage of modes initially excited s/N fixed, while E, s and N increase to double in each forward panel. The continuous line corresponds to the q–tori law $E_{k+1}/E = (\mu N/\pi s)^{2k} k^2$, $k = q/s$.

Numerical integration of the Toda system, when given same initial conditions with the FPU system, exhibits an impressively similar to the FPU behaviour. Such an example is shown in Fig. 4 (a), which is a repetition of the example in Fig. 2, but for the integrable Toda system. Both systems for the same initial conditions exhibit the same recurrences. Furthermore, the two systems have almost equal energy localization profiles, as can be seen from Fig. 4 (b), characterized by a natural packet of length $\mu = 3.38/N$ (3 to 4 modes share almost all the energy) and a tail with slope that depends linearly on μ.

2.3 Under KdV framework

Maybe the most pioneer work on the FPU paradox was the one by Zabusky and Kruskal [2] in 1965, in which a connection between the integrable KdV equation and FPU–α model for long wavelength excitations was for the first time observed. The authors point out that such kind of initial conditions lead eventually to the decomposition of the solution into a sequence of solitons, where under a Fourier transformation interpret the famous FPU recurrences. In particular, the authors in [2] state: 'solitons pass through one another without losing their identity'.

The partial differential equation Kortewegde Vries with dispersion is given by the equation

$$u_t + uu_x + \delta^2 u_{xxx} = 0, \quad x \in [0, 2] \tag{7}$$

where δ is the dispersion parameter. In the small dispersion limit $\delta \to 0$, Eq. (7) tends to Hopf equation $u_t + uu_x = 0$, which is also integrable and faces a shock in finite time.

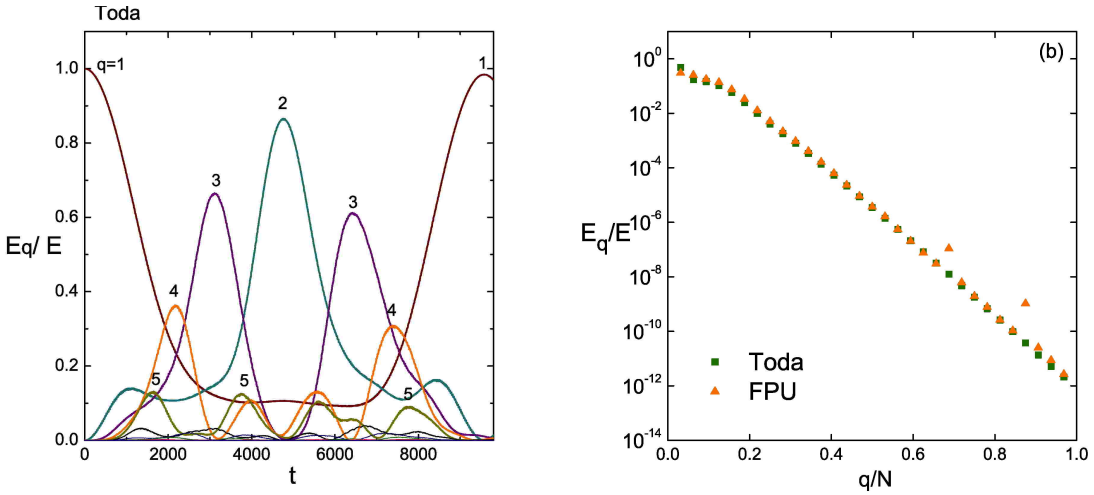


Figure 4: (a) The evolution of the harmonic energies $E_q(t)/E$, $q = 1, \ldots, N$ for the 32–particle Toda chain, with $E = 0.07$ and $\alpha = 0.33$, (b) Normalized and averaged in time energy spectra versus q/N for FPU with orange triangles and for Toda with green squares, that rise by the initial conditions of panel (a).

Only after 40 years, the KdV–direction in solving the FPU problem was revived by A. Ponno and D. Bambusi in [8], [9], where the two models are compared and connected through perturbation theory's techniques. By using a resonant normal form, that consists of two uncoupled KdV equations, provide a first order approximation of the FPU–α dynamics. Going back to the metastable scenario, which we described in the Introduction, the authors get rigorous estimates, explaining the natural packet's width and the localization law of the tail modes $E_k = c_k e^{-ck}$, where k represents the Fourier mode number.

In a recent work [20], we study numerically more general type of initial conditions for KdV

$$u(x, 0) = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} \cos(\pi k x + \varphi_k), \quad x \in [0, 2] \quad , \tag{8}$$

i.e. an initial excitation of m Fourier modes with random phases φ_k, $k = 1, ..., m$. By systematically measuring the slope σ of the KdV solution's Fourier profile, we find that there are two slope–regimes, depending on the dispersive parameter δ and on the number of modes excited m. In particular for $m > 1$, when $\delta \to 0$, so in the small dispersion limit, σ depends linearly on δ, while for larger δ values, where the linear term u_{xxx} of KdV in Eq. (7) is the dominant one, the slope σ depends logarithmically on δ.

Conclusions We can classify the approaches to the FPU–problem, when regarded as a perturbation of an integrable system, into three major cases: when the integrable system is i) the coupled harmonic oscillators (linear), ii) the Toda and iii) the KdV equation. It is evident that the FPU Hamiltonian expressed as a perturbation of the Toda system is a more prominent approach to the problem compared to linear–approach one, since in (ii) is regarded as a perturbation of higher order with respect to α, compared to (i). However, the implementation of perturbation techniques using Toda is still an open problem. Finally, the KdV approach completes both (i) and (ii) cases, where when more Fourier modes are excited, the slope of the energy spectra depends either logarithmically or linearly on system parameters, depending on the value of δ.

References

[1] E. Fermi, J. Pasta, and S. Ulam, *Studies of non linear problems.* Los Alamos report No LA-1940 (1955).

[2] N. J. Zabusky and M. D. Kruskal, *Interaction of 'Solitons' in a Collisionless Plasma and the Recurrence of Initial States.* Phys. Rev. Lett. 15 (1965), 240–243.

[3] B. Rink and F. Verhulst, *Near-integrability of periodic FPU-chains.* Physica A 285 (2000), 467–482.

[4] F. Fucito, F. Marchesoni, E. Marinari, G. Parisi, L. Politi, S. Ruffo, and A. Vulpiani, *Approach to equilibrium in a chain of nonlinear oscillators.* J. Physique 43 (1982), 707–713.

[5] R. Livi, M. Pettini, S. Ruffo, and A. Vulpiani, *Further results on the equipartition threshold in large nonlinear Hamiltonian systems.* Phys. Rev. A 31 (1985), 2740.

[6] W. E. Ferguson Jr., H. Flaschka and D. W. McLaughlin, *Nonlinear normal modes for the Toda chain.* J. Comp. Phys. 45 (1982), 157–209.

[7] P. Bocchieri, A. Scotti, B. Bearzi, A. Loinger, *Anharmonic chain with Lennard–Jones interaction.* Phys. Rev. A 2 (1970), 2013–2019.

[8] A. Ponno and D. Bambusi, *Korteweg-de Vries equation and energy sharing in Fermi-Pasta-Ulam.* Chaos 15 (2005), 015107.

[9] D. Bambusi and A. Ponno, *On Metastability in FPU.* Comm. Math. Phys. 264 (2006), 539–561.

[10] L. Berchialla, A. Giorgilli, and S. Paleari, *Exponentially long times to equipartition in the thermodynamic limit.* Phys. Lett. A 321 (2004), 167.

[11] L. Berchialla, L. Galgani, and A. Giorgilli, *Localization of energy in FPU chains.* Discr. Cont. Dyn. Sys. A 11 (2005), 855-866.

[12] S. Flach, M. V. Ivanchenko, and O. I. Kanakov, *q-Breathers and the Fermi-Pasta-Ulam Problem.* Phys. Rev. Lett. 95 (2005), 064102.

[13] S. Flach, M. V. Ivanchenko, and O. I. Kanakov, *q-Breathers in Fermi-Pasta-Ulam chains: Existence, localization, and stability.* Phys. Rev. E 73 (2006), 036618.

[14] S. Flach and A. Ponno, *The Fermi-Pasta-Ulam problem: Periodic orbits, normal forms and resonance overlap criteria.* Physica D 237 (2008), 908.

[15] S. Flach, O. Kanakov, M. Ivanchenko, and K. Mishagin, *q-Breathers in FPU lattices - scaling and properties for large system.* Int. J. Mod. Phys. B 21 (2007), 3925.

[16] O. Kanakov, S. Flach, M. Ivanchenko, and K. Mishagin, *Scaling properties of q-breathers in nonlinear acoustic lattices.* Phys. Lett. A 365 (2007), 416.

[17] H. Christodoulidi, Ch. Efthymiopoulos and T. Bountis, *Energy localization on q-tori, long-term stability, and the interpretation of Fermi-Pasta-Ulam recurrences.* Phys. Rev. E 81 (2010), 016210 .

[18] H. Christodoulidi, Ch. Efthymiopoulos. `http://arxiv.org/abs/1205.2573`.

[19] G. Benettin and A. Ponno, *Time-scales to equipartition in the Fermi-Pasta-Ulam problem: finite-size effects and thermodynamic limit.* J. Stat. Phys. 144 (2011), 793.

[20] H. Christodoulidi, A. Ponno and G. Benettin. In preparation.

Interpolation theorems for weighted spaces of smooth functions

Ademi Ospanova (*)

Introduction

Interpolation theory of function spaces is a new branch of functional analysis, which finds application in wide areas of theory of functions, as well as in other areas of mathematics. The most important fields of application of interpolation theory are: the theory of function spaces and differential operators, the theory of partial differential equations, the theory of Fourier series, approximation theory in Banach spaces, integral inequalities, singular integrals, theory of multipliers.

The problem of interpolation of Sobolev spaces with the weights of nonpowered character was considered for the first time by L. Kussainova [3, 4]. The work I present here is devoted to the problem of interpolation for Sobolev spaces with weights of more general form.

Now we will describe the purpose of interpolation theory in Banach spaces. Let $\mathcal{A}$ be a linear topological Hausdorff space. We say that a pair $\{A_0, A_1\}$ of Banach spaces is an interpolation pair in $\mathcal{A}$ provided that A_0 and A_1 are Banach spaces continuously embedded into $\mathcal{A}$. Let $\mathcal{B}$ be an other linear topological Hausdorff space. Let $\{B_0, B_1\}$ be an interpolation pair in $\mathcal{B}$. An interpolation theory aim at associating to each interpolation pair $\{A_0, A_1\}$ a Banach space A such that $A_0 \cap A_1 \subseteq A \subseteq A_0 + A_1$ and to each pair $\{B_0, B_1\}$ a Banach space B such that $B_0 \cap B_1 \subseteq B \subseteq B_0 + B_1$ in such a way that if T is a linear operator from $\mathcal{A}$ to $\mathcal{B}$ such that the restriction of T to A_i induces a linear and continuous map from A_i into B_i, $(i = 0, 1)$ that T induces necessarily a linear and continuous operator from A into B. The spaces A and B are called interpolation spaces with respect to the pairs $\{A_0, A_1\}$ and $\{B_0, B_1\}$. The main aim of interpolation is to obtain a construction of interpolation spaces in explicit form.

There are two interpolation methods of construction of interpolation spaces: the real method and the complex method. Interpolation spaces built by the real method are called the Peetre spaces. This method has many applications in function theory.

(*)The L. N. Gumilyov Eurasian National University, Munaitpasov st., 5, 010008, Astana, Kazakhstan; E-mail: ospanova@math.unipd.it . Seminar held on February 8th, 2012.

We will give the description of Peetre interpolation spaces of weighted Sobolev spaces defined on a domain G in a case, when $G = \mathbb{R}$ and in a case, when G is any domain in $\mathbb{R}^n$.

Interpolation theorems for pair of weighted Sobolev spaces

Let (A_0, A_1) be an interpolation pair. The real interpolation space or space Peetre is defined as follows. Let $0 < \theta < 1$ and $1 \le q < \infty$. By definition

$$(A_0, A_1)_{\theta,q} \;=\; \left\{ a : a \in A_0 + A_1,\ \|u\|_{(A_0,A_1)_{\theta,q}} = \left(\int_0^\infty (t^{-\theta} K(t,a))^q \frac{dt}{t} \right)^{1/q} < \infty \right\},$$

where

$$A_0 + A_1 = \{a :\ a = a_0 + a_1,\ a_0 \in A_0,\ a_1 \in A_1\},$$
$$K(t,a) = \inf_{a=a_0+a_1} (\|a_0\|_{A_0} + t\,\|a_1\|_{A_1}), \quad a \in A_0 + A_1 \quad (0 < t < \infty).$$

Example Let $0 < p_0 < p_1 \le \infty$. Let $\theta \in (0,1)$ such that $1/p = (1-\theta)/p_0 + \theta/p_1$, and $q = p$. Then one can prove that $(L_{p_0}, L_{p_1})_{\theta,p} = L_p$.

We consider the problem of the description of the Peetre interpolation space for pairs of weighted Sobolev spaces

$$(W^m_{p_0}(\upsilon), L_{p_1}),\ (W^m_{p_0}(\upsilon_0), L_{p_1}(\upsilon_1)),$$
$$(W^{m_0}_{p_0}(\upsilon_0), W^{m_1}_{p_1}(\upsilon_1)),\ (W^{m_0}_{p_0}(\rho_0, \upsilon_0), W^{m_1}_{p_1}(\rho_1, \upsilon_1)).$$

Now we introduce some notation. Let $Q_d(x)$ be an open n-dimensional cube centered at x:

$$Q = Q_d = Q_d(x) \equiv \{y \in \mathbb{R}^n : |y_i - x_i| < \tfrac{1}{2}d,\quad i = 1, \dots, n\} = \left(x_i - \tfrac{1}{2}d_i x_i + \tfrac{1}{2}d_i\right)^n.$$

In the one-dimensional case we have the open interval of length d.

Definition 1 We say that d is function of length of edge for a cube in $G \subset \mathbb{R}^n$ if d is a positive bounded function from G to $(0, +\infty)$ which satisfies conditions:

(a) immersion:

$$\forall x \in G \qquad Q(x) = Q_{d(x)}(x) \subset G\ ;$$

(b) slow oscillation:

$$(1) \qquad \exists\, \eta \in (0,1) :\quad \eta < \frac{d(y)}{d(x)} < \eta^{-1}, \quad \text{if}\ \ y \in (1-\varepsilon)Q(x)\ \ \text{and}\ x \in G\ ;$$

ε is some fixed number of $(0,1)$.

In a case, when $G = \mathbb{R}$, we have only condition b) and $\varepsilon = 1$.
Now we introduce some examples of weights satisfying the condition (1).

Examples The following functions satisfy the condition (1):

(a) $d(x) \equiv 1$ on $G = \mathbb{R}^n$, $n \geq 1$.

(b) $d(x) = \min\{x - a,\ b - x\}$ on $G = (a, b)$.

(c) $d(x) = (1 + |x|^{\mu})^{-1}$ on $G = \mathbb{R}^n$, $n \geq 1$.

(d) $d(x) = \min\{1,\ \mathrm{dist}(x, \partial G)\}$ on $G \subset \mathbb{R}^n$, where $\mathrm{dist}(x, \partial G) \stackrel{def}{=} \sup\limits_{h>0}\{h : Q_h(x) \subset G\}$ – distance from $x \in G$ to the boundary ∂G of bounded domain G.

Let $\rho(x)$, $\upsilon(x)$ be the weight functions given on the entire line $\mathbb{R} = (-\infty,\ \infty)$. Let $m \geq 1$ be an integer, $1 < p < \infty$. We define the Sobolev space $\mathcal{W}_p^m(\rho; \upsilon)$ as follows:

$$\mathcal{W}_p^m(\rho;\upsilon) = \{u \in L(\mathrm{loc}) :\ \left\|u;\ W_p^m(\rho,\ \upsilon)\right\| = \left\|u^{(m)};\ L_p(\rho)\right\| + \|u;\ L_p(\upsilon)\| \tag{2}$$
$$= \left(\int\limits_{-\infty}^{\infty} |u^{(m)}(x)|^p \rho(x) dx\right)^{1/p} + \left(\int\limits_{-\infty}^{\infty} |u(x)|^p \upsilon(x) dx\right)^{1/p} < \infty\}.$$

Let $W_p^m(\rho; \upsilon)$ be a two-weighted Sobolev space defined as completion of the space $C^{\infty}\mathcal{W}_p^m(\rho; \upsilon) = C^{\infty} \cap \mathcal{W}_p^m(\rho; \upsilon)$ with respect to the norm $\left\| \cdot\ ; \mathcal{W}_p^m(\rho, \upsilon)\right\|$ of (2). In particular, $W_p^m(1; \upsilon)$ will also be denoted by $W_p^m(\upsilon)$.

Class $(\Pi_{m,p})$**.** We say that a weight υ on $\mathbb{R}$ satisfies condition $(\Pi_{m,p})$ with respect to the function d, and write $\upsilon \in (\Pi_{m,p})$, if there are numbers $0 < \delta,\ \gamma < 1$ such that

$$d(x)^{mp-1} \inf_e \int_{Q \setminus e} \upsilon(y) dy \geq \gamma\ \forall Q = Q(x) \qquad \text{for almost all } x \in \mathbb{R}. \tag{3}$$

In (3) the inf is taken on all measurable subset e of Q such that $|e| \leq \delta\,|Q|$.

Now we introduce some examples of weights satisfying the condition $(\Pi_{m,p})$.

(E1) The weight $\upsilon = 1$ satisfies condition $(\Pi_{m,p})$ with respect to $d(x) = 1$. Indeed,

$$d(x)^{mp-1} \inf_e \upsilon\left(Q(x) \setminus e\right) \geq 1 - \delta = \gamma > 0, \qquad \text{for all } \delta \in (0, 1).$$

(E2) Let $\varepsilon \in (0, 1]$. Let $\upsilon(x) = |x - x_0|^{\mu}$, $0 < \mu < 1$, if $|x - x_0| < 1$ and $\upsilon(x) = \varepsilon$ if $|x - x_0| \geq 1$. Weight υ satisfies condition $(\Pi_{m,p})$ with respect to the function

$$d(x) = \begin{cases} (x - x_0)^{\gamma} & |x - x_0| < 1, \\ 1 & |x - x_0| \geq 1, \end{cases}$$

where $\gamma = \mu/mp$.

(E3) Let G be a domain with nonempty boundary,

$$d(x) = \begin{cases} 1/(2\sqrt{n}), \text{ if } \sigma(x) \geq 1, \\ \sigma(x)/(2\sqrt{n}), \text{ if } \sigma(x) < 1. \end{cases}$$

The function $\upsilon(x) \equiv e^{1/\sigma(x)}$ $(\sigma(x) = \operatorname{dist}(x, \partial G))$ satisfies condition $(\Pi_{m,p})$ with respect to the function $d(x)$.

Let $0 < s < \infty$, $1 \leq p < \infty$. We define a two-weighted space of Besov as follows:

$$B_p^s(\rho, \upsilon) \equiv \left\{ u \in L(\text{loc}) : \left\| u; B_p^s(\rho, \upsilon) \right\| < \infty \right\},$$

where

$$\left\| u; B_p^s(\rho, \upsilon) \right\| = \left\{ \int_0^\infty h^{-(s-k)p-1} \int_{\mathbb{R}} \left| \bigtriangleup^m(h) u^{(k)}(x) \right|^p \rho(x) dx dh \right\}^{1/p} + \left\| u; L_p(\upsilon) \right\|. \tag{4}$$

Here m, k – two integer non-negative numbers, satisfying the condition $m + k > s > k$. Let $\Delta^m(h)$ be the difference operator of order m $(m = 1, 2...)$ with step h defined as

$$\Delta^m(h) f(x) = \sum_{r=0}^{m} (-1)^{m-r} \frac{m!}{r!(m-r)!} f(x + rh).$$

One can prove that the above definition is independent of our specific choice of m and k as long as the above conditions are satisfied. We denote by $B_p^s(\upsilon)$ the space $B_p^s(1, \upsilon)$.

Theorem 1 *Let $0 < m_1 < m_0$ be integers, $1 < p_0 \leq p_1 < \infty$, $0 < \theta < 1$, $s = (1-\theta)m_0 + \theta m_1$, $1/p = (1-\theta)/p_0 + \theta/p_1$. Let $\upsilon_i(x)$ $(i = 0, 1)$ be weights on $\mathbb{R}$, satisfying the following conditions:*

1. *$\upsilon_i \in \Pi_{m_i, p_i}$ with respect to the function $d(x)$.*

2. *$K_i = \sup\limits_{x \in \mathbb{R}} d(x)^{m_i - 1/p_i} \left(\upsilon_i \left(Q(x) \right) \right)^{1/p_i} < \infty$.*

Then

$$\left(W_{p_0}^{m_0}(\upsilon_0),\ W_{p_1}^{m_1}(\upsilon_1) \right)_{\theta, p} = B_p^s(d(\cdot)^{-sp}).$$

Let $1 \leq p < \infty$. Let υ be a weight on $\mathbb{R}$. Let us assume that

$$\upsilon^*(x) = \sup_{d > 0} \left\{ d > 0 :\ d^{mp-1} \upsilon((Q_d(x)) \leqslant 1 \right\}, \tag{5}$$

$$Q^*(x) \equiv Q_d(x), \text{ where } d = \upsilon^*(x). \tag{6}$$

Note that the set in (5) is nonempty and

$$\text{for almost all } x \in \mathbb{R} \quad 0 < \upsilon^*(x) \le C < \infty \quad (mp > 1). \tag{7}$$

Example For the weight $\upsilon(x) = e^{|x|}$ we have $\upsilon^*(x) \equiv e^{-x/mp}$.

Remark 1 Let $mp > 1$. if weight υ satisfies condition (7) that $\upsilon \in \Pi_{m,p}$ with respect to function $d(x) = \upsilon^*(x)$.

Theorem 2 *Let $0 < m_1 < m_0$ be integers, $1 < p_0 \leqslant p_1 < \infty$, $0 < \theta < 1$, $s = (1-\theta)m_0 + \theta m_1$, $1/p = (1-\theta)/p_0 + \theta/p_1$. Let υ_i be weights on $\mathbb{R}$ satisfying the following conditions:*

1. *There exists $H > 0$ such that $\upsilon_i(Q_d(x)) \ge 1$ if $|x| > H$, $d > H$.*
2. *There exist $\delta \in (0,1)$ and $\Gamma > 1$ such that $\forall Q = Q_d$, $\upsilon_i(Q) \le \Gamma \upsilon_i(Q \backslash e)$ if $e \subset Q$, $|e| < \delta |Q|$.*

Then

$$\left(W_{p_0}^{m_0}(\upsilon_0),\ W_{p_1}^{m_1}(\upsilon_1)\right)_{\theta,p} = \left\{ u \in L(\mathrm{loc}) : \|u; b_p^s\| + \left(\int\limits_{\left\{\upsilon_0^* \ge \upsilon_1^*\right\}} |u(y)|^p \, \upsilon_0^{*-sp}(y) dy \right)^{1/p} + \left(\int\limits_{\left\{\upsilon_1^* > \upsilon_0^*\right\}} |u(y)|^p \, \upsilon_1^{*-sp}(y) dy \right)^{1/p} < \infty \right\} = B_p^s(\omega),$$

where $\omega(x) = \min\{(\upsilon_0^*(x))^{-sp}, (\upsilon_1^*(x))^{-sp}\}$.

Example 1 Let $\upsilon_0(x) = e^{|x|}$, $\upsilon_1(x) = 1$. Then from the above-mentioned example follows that $\upsilon_0^*(x) = e^{-x/(mp)}$. For $\upsilon_1(x)$ we have $\upsilon_1^*(x) = 1$. Hence,

$$\omega(x) = \min\{\upsilon_0^*(x)^{-sp}, \upsilon_1^*(x)^{-sp}\} \equiv \min\{e^{-\frac{|x|}{mp}\cdot(-sp)}, 1\} = 1. \tag{8}$$

Now from Theorem 2 we obtain the equality

$$\left(W_{p_0}^{m_0}(e^{|x|}), W_{p_1}^{m_1}(\mathbb{R})\right)_{\theta,p} = B_p^s(\omega) = B_p^s(\mathbb{R}).$$

Theorem 3 *Let $0 < m_1 < m_0$ be integers, $1 < p_0 \le p_1 < \infty$, $0 < \theta < 1$. Let $\upsilon_i = \rho_i d(\cdot)^{-m_i p_i}$ $(i = 0, 1)$, ρ_i be weights on $\mathbb{R}$ satisfying the following conditions: there exist constants $K > 0$, $b_i > 0$ $(i = 0, 1)$ such that:*

1. $b_i^{-1} \le \rho_i(y)/\rho_i(x) \le b_i$ *if* $y \in Q(x)$ *for all* $x \in \mathbb{R}$ (bounded oscillation condition)
2. $d(x)^{m_0-m_1}\rho_1(Q(x))^{1/p_1} \le K\rho_0(Q(x))^{1/p_0}$ *for all* $x \in \mathbb{R}$.

Then the following equality holds:

$$\left(W_{p_0}^{m_0}(\rho_0,\upsilon_0),\ W_{p_1}^{m_1}(\rho_1,\upsilon_1)\right)_{\theta,p} = B_p^s(\rho_\theta,\upsilon_\theta)$$

where $\rho_\theta^{1/p} = \rho_0^{(1-\theta)/p_0}\rho_1^{\theta/p_1}$, $\upsilon_\theta = \rho_\theta d(\cdot)^{-sp}$, $s = (1-\theta)m_0+\theta m_1$, $1/p = (1-\theta)/p_0+\theta/p_1$.

We will present some examples.

Example 2 From Theorems 1 and 3, in the case $\upsilon_0(x) = 1 = \upsilon_1(x)$ and $\rho_0(x) = 1 = \rho_1(x)$, then we obtain the description of the spaces $\left(W_{p_0}^{m_0}(\mathbb{R}), W_{p_1}^{m_1}(\mathbb{R})\right)_{\theta,p}$.

Example 3 If in Theorem 3 we assume that the weights are $\rho_i(x) = (1+|x|)^{\mu_i}$, $i = 0, 1$, and that $d(x) \equiv 1$, then we obtain a one-dimensional result of H. Triebel [2].

Example 4 If in the Example 3 we choose $\mu_0 = \mu_1 = 0$ then we obtain the known interpolation equality:

$$\left(W_{p_0}^{m_0}(\mathbb{R}),\ W_{p_1}^{m_1}(\mathbb{R})\right)_{\theta,p} = \left(W_{p_{0,0}}^{m_0}, W_{p_{1,0}}^{m_1}\right)_{\theta,p} = B_p^s(\mathbb{R}).$$

Let $W_{p,\mu}^m(\upsilon) = W_p^m(\upsilon^\mu, \upsilon^{\mu-mp})$, $B_{p,\nu}^s(\upsilon) = B_p^s(\upsilon^\nu, \upsilon^{\nu-sp})$, $-\infty < \mu, \nu < \infty$, $s > 0$.

Theorem 4 *Let* $1 < p_0 \le p_1 < \infty$, $-\infty < \mu_0, \mu_1 < \infty$, $\aleph = m_0 - m_1 - \frac{1+|\mu_0|}{p_0} + \frac{1+|\mu_1|}{p_1} \ge 0$, *let* υ *be a weight satisfying the condition* (7). *Then*

$$\left(W_{p_0,\mu_0}^{m_0}(\upsilon^*), W_{p_1,\mu_1}^{m_1}(\upsilon^*)\right)_{\theta,p} = \mathcal{B}_{p,\nu}^s(\upsilon^*), \tag{9}$$

where $0 < \theta < 1$, $s = (1-\theta)m_0+\theta m_1$, $1/p = (1-\theta)/p_0+\theta/p_1$, $\nu = (1-\theta)\mu_0/p_0+\theta\mu_1/p_1$.

The collection of intervals $I = \bigcup_{x\in\mathbb{R}} I_x$, where $I_x = \{Q = Q_d : Q_d \ni x\}$, is called differential basis. Let us denote by Mf the maximal operator of Peetre

$$Mf(x) = \sup_{(a,b)\ni x} \frac{1}{|b-a|}\int_a^b |f(y)|dy \quad f \in L(\mathrm{loc}).$$

Definition 2 Weight ω on R satisfies condition (A_1) ($\omega \in A_1$) (with respect to basis I) if there exists a constant $P_0 > 1$ such that

$$\frac{1}{|Q|}\int_Q \omega(y)dy \le P_0 \mathrm{vrai}\inf_{x\in Q}\omega(x) \quad \text{for all } Q \in I. \tag{10}$$

Definition 3 Weight ω on R satisfies condition (A^*) (record $\omega \in A^*$) (with respect to I) if there exist constants $P_1 > 1$ and $\tau \in (0,1)$ such that

$$\sup_Q M(\omega\chi_Q) \le P_1 \sup_{Q\setminus e} M(\omega\chi_Q), \quad \text{as } e \subset Q,\ |e| \le \tau|Q| \quad \text{for all } Q \in I. \tag{11}$$

Example If the weight υ "slowly oscillates" then $\upsilon \in A_1$ and $\upsilon \in A^*$.

Theorem 5 *Let $0 < m_1 < m_0$ be integers, $0 < \theta < 1$, $s = (1-\theta)m_0 + \theta m_1$, $r = 1 - m_1/m_0$, $1 < p < \infty$. Let υ be a continuous weight from class $A_1 \cap A^*$, satisfying condition $\upsilon(x) \ge c_0 > 0$. Then*

$$\left(W_p^{m_0}(\upsilon), W_p^{m_1}(\upsilon^{-r}, \upsilon)\right)_{\theta,p} = B_p^s(\upsilon^{-\theta r}, \upsilon).$$

Let G be an arbitrary domain of $\mathbb{R}^n$, let $\rho = \rho(x)$, $\upsilon = \upsilon(x)$ be non-negative locally summable functions (weights) in G.

We obtain embedding theorems and descriptions of real interpolation spaces for interpolation pairs of two-weighted Sobolev spaces with the different indexes of smoothness and summability.

Let $m > 0$ be an integer, $mp > n$, $1 \le p < \infty$. We define the two-weighted Sobolev space $W_p^m(G;\rho,\upsilon)$ as completion of the space $C^\infty \mathcal{W}_p^m(G;\rho,\upsilon) = C^\infty(G) \cap \mathcal{W}_p^m(G;\rho,\upsilon)$ with respect to the norm

$$\left\|u; \mathcal{W}_p^m(G;\rho,\upsilon)\right\| = \sum_{i=1}^n \left\|D_i^m u\right\|_{L_p(G;\rho)} + \left\|u\right\|_{L_p(G;\upsilon)}. \tag{12}$$

If $\rho = \upsilon = 1$ then the space $W_p^m(G;\rho,\upsilon)$ is well-known Sobolev space $W_p^m(G)$ with the corresponding norm.

Theorem 6 *Let $0 < m_1 < m_0$ be integers, $s = (1-\theta)m_0 + \theta m_1$, $0 < \theta < 1$, $1 < p_0 \le p_1 < \infty$, $1/p = (1-\theta)/p_0 + \theta/p_1$. Let weights $\upsilon_i(x) = \rho_i(x) d_i(x)^{-m_i p_i}$ $(i = 0,1)$, $m_i p_i > n$, and weights ρ_i on G satisfy conditions (with respect to the d):*

1. *there exist constants $b_i > 0$ $(i = 0,1)$ such that*

$$b_i^{-1} \le \frac{\rho_i(y)}{\rho_i(x)} \le b_i \text{ if } y \in Q_{(\varepsilon)}(x);$$

2. *there exists a constant $K > 0$ such that for almost all $x \in G$*

$$d(x)^{m_0 - m_1} \left(\int_{Q_{(\varepsilon)}(x)} \rho_1(y)dy\right)^{1/p_1} \le K \left(\int_{Q_{(\varepsilon)}(x)} \rho_0(y)dy\right)^{1/p_0}.$$

Then

$$\left\|u; W_{p_1}^{m_1}(G;\rho_1,\upsilon_1)\right\| \le cK \left\|u; W_{p_0}^{m_0}(G;\rho_0,\upsilon_0)\right\|, \quad u \in W_{p_0}^{m_0}(G;\rho_0,\upsilon_0).$$

Now the interpolation theorem for pair of spaces $\left(W_{p_0}^{m_0}(G;\rho_0,\upsilon_0), W_{p_1}^{m_1}(G;\rho_1,\upsilon_1)\right)$. Let $1 \leqslant p \leqslant \infty$, $s > 0$. We define a two-weighted Besov space as follows:

$$B_p^s(G;\rho,\upsilon) \equiv \left\{u \in L(G;\mathrm{loc}) : \left\|u; B_p^s(G;\rho,\upsilon)\right\| < \infty\right\},$$

where

$$\left\|u; B_p^s(G;\rho,\upsilon)\right\| = \|u; L_p(G;\upsilon)\| + \sum_{i=1}^{n} \left\{ \int_0^{h_0} h^{-(s-k)p-1} \left\|\Delta_i^m(h) f(x) D_i^k u(\xi)\rho(\xi)\right\|_p^p dh \right\}^{1/p}. \tag{13}$$

Here $h > 0$, $h_0 \in (0,1]$, $[x, x+mhe^i] \subset \{x : x + [-h,h]^n \subset G\}$; $m > 0$, $k \ge 0$ are integer numbers such that $m + k > s > k$,

$$\Delta_i^m(h) f(x) = \sum_{r=0}^{m} (-1)^{m-r} \frac{m!}{r!(m-r)!} f(x + rhe^i) \qquad (m = 1, 2...).$$

Theorem 7 *Let $0 < m_1 < m_0$ be integers, $m_i p_i > n$, $1 < p_0 \le p_1 < \infty$, $1/p = (1-\theta)/p_0 + \theta/p_1$, $s = (1-\theta)m_0 + \theta m_1$, $0 < \theta < 1$. Let weights ρ_i on G satisfy conditions 1 and 2 of Theorem 6; let $\upsilon_i(x) = \rho_i(x) d_i(x)^{-m_i p_i}$ $(i = 0, 1)$. Then interpolation equality*

$$\left(W_{p_0}^{m_0}(G;\rho_0,\upsilon_0), W_{p_1}^{m_1}(G;\rho_1,\upsilon_1)\right)_{\theta,p} = B_p^s(G;\rho_\theta,\upsilon_\theta),$$

where $\rho_\theta^{1/p} = \rho_0^{(1-\theta)/p_0} \rho_1^{\theta/p_1}$, $\upsilon_\theta = \rho_\theta d(\cdot)^{-sp}$, holds.

In these results the basic condition imposed on the weights is a bounded oscillation condition. This condition can be a consequence of any properties of behaviour of the gradient if the weight ρ is differentiable. But it is not such a rigid condition. Considering for simplicity the one-dimensional case, we can obtain such a bounded oscillation condition from the condition $\max\limits_{Q(x)} |\rho'| \le K \frac{\rho(x)}{d(x)}$, which comes up in the analysis of weighted spaces.

Example 5 If in the above theorems we assume that $d(x) = (1+|x|^\nu)^{-1}$, $\rho(x) = 1 + |x|^\nu$, $-\infty < \nu < \infty$, then we would obtain the interpolation theorem for two-weighted Sobolev spaces with the weights of power form given by H. Triebel [2].

Example 6 Let $\omega(x)$ be a weight of $\mathbb{R}$. If in Theorem 7 we assume that $\rho(x) = \omega^*(x)^\lambda$, $\upsilon = \rho(x)(\omega^*(x))^{-mp}$, $\lambda < mp$, then our theorems allow to obtain the interpolation theorems for two-weighted Sobolev spaces given by L. Kussainova. Here the weight $\omega^*(x)$ is defined by Definition (5).

Furthermore, if $m_0 = m_1 = m$ in the above statement then $s = m$ and one can see that the corresponding interpolation of Sobolev spaces with different weights yields the same space $B_p^s(G; d(\cdot)^{-mp})$, $G \subseteq \mathbb{R}^n$.

References

[1] Berg J., Lofstrom J., "Interpolation spaces. Introduction". Springer, 1976.

[2] Triebel H., "Interpolation theory. Function spaces. Differential operators". Berlin, 1978.

[3] Kussainova L. K., "Embedding theorems and interpolation of weighted Sobolev spaces". Doctoral thesis. Almaty, 1998.

[4] Kussainova L. K., *On interpolation of weighted Sobolev spaces.* Izvestiya MN–AN RK. Ser. Phis. Matem. 5 (1997), 33–51 (in Russian).

[5] Besov O. V., *Interpolation of spaces of differentiable functions on a domain.* Trudy Matem. Ins-ta Im. V. A. Steklova 214 (1997), 59–61 (in Russian).

Modeling and valuing *make-up* clauses in gas swing contracts

Enrico Edoli (*)

This presentation is based on the article [11].

Abstract. In the last ten years, thanks to the worldwile energy liberalization process, the birth of competitive gas markets and the recent financial crisis, traditional long term swing contracts in Europe have been supplemented in a significant way by make-up clauses which allow to postpone the withdrawal of gas to future years when it could be more profitable. This introduces more complexity in the pricing and optimal management of swing contracts. This paper is devoted to a proper quantitative modelization of a kind of make-up clause in a gas swing contract. More in detail, we succeed in building an algorithm to price and optimally manage the make-up gas allocation among the years and the gas taking in the swing subperiods within the years: we prove that this problem has a quadratic complexity with respect to the number of years. The algorithm can be adapted to different instances of make-up clauses as well as to some forms of carry-forward clauses. Then, as an example, we show the algorithm at work on a 3-year contract and we present a sensitivity analysis of the price and of the make-up policy with respect to various parameters relative both to the price dynamics as well as to the swing contract. To the authors' knowledge, this is the first time that such a quantitative treatment of make-up clauses appears in literature.

1 Introduction to long term supply contracts in European gas market

Europe is among the largest consumer of natural gas in the world, mainly used for heating and power generation. During the last thirty years natural gas has gradually replaced almost everywhere fuel oil for heating purposes and is actually competing with coal as main fuel source for electric power generation. Hence, long term trend of natural gas demand has been historically upward sloping. The economic crisis of 2008 has strongly impacted this tendency: global gas demand fell sharply by 3% between 2008 and 2009; however, the International Energy Agency (IEA) forecasts that OECD[7] gas demand would recover

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: edoli@math.unipd.it. Seminar held on February 15th, 2012.

[7]Current membership of OECD: Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States.

slowly with consumption returning to the 2008 levels by 2012 or 2013, depending on the region. In addition, recent events concerning nuclear power generation, post Fukushima's accident, are expected to provide new strength to the long term up-growing tendency of natural gas global demand. In fact, in the medium to long term, many countries are expected to reduce their nuclear ambitions and the fuel of choice to compensate for lower nuclear will reasonably be natural gas. Despite its significant consumption, Europe, meant either as OECD or European Union (EU), has only a limited inner production compared to its consumption and the excess demand is covered by massive natural gas imports from producer countries like Russia, Norway, Algeria

Natural gas imports are physically delivered via pipelines or via LNG (Liquified Natural Gas) cargoes and were traditionally based on long term oil-linked swing contracts (10-30 years duration) in order to guarantee the security of supply of such an important energy commodity. In the last ten years, thanks to the worldwide energy liberalization process and the birth of competitive gas markets, in almost all European countries these long term contracts have been supplemented by spot transactions (short term transactions) even if long term deals still represent the pillar of European gas system (see [21]).

The structure of long term gas agreements is pretty standardized in Europe. As already said, these long term contracts are swing (also known as *take-or-pay*, see [20] for details) in nature, with the peculiarity that the strike price typically depends upon a basket of crude and refined oil products, which is averaged through time in order to smooth undesired volatility effects; for more details we refer the interested reader to [1, Section 3.1].

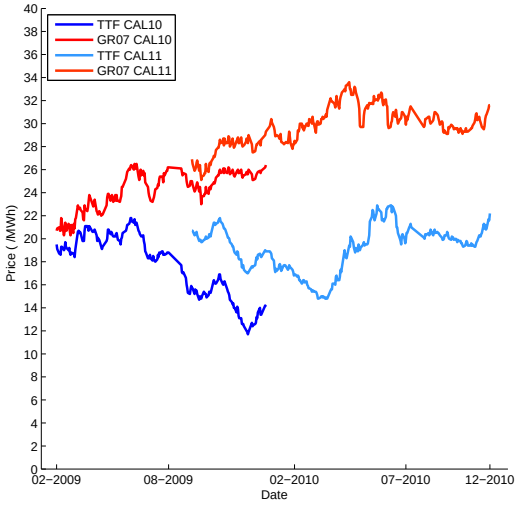


Figure 1: CAL10 and CAL11 (calendar forward) of TTF and GR07 (ENI Gas Release 2007, gas price formula) contracts. [2]

These long term contract structure and oil indexation have their origins in the early European gas market of the 1970s. Since that time sources of gas have increased, making gas markets and infrastructure much denser and open to competion. From 2008 onwards this traditional market framework has significantly changed especially for what concern

the oil-to-gas price relationship. Actually, the demand drop following the financial crisis with the subsequent economic recession associated to the significant increase in LNG (Liquefied Natural Gas) and unconventional gas supply sources flowing to Europe generated a consistent and pretty persistent oil decoupling of European gas market prices: since 2008, European gas markets are pricing systematically and significantly below indexes usually used for the strike price I above (see Figure 1), so the spread $P - I$ has became negative.

Obviously, this market phenomenon has determined a panic situation for all the owners of classical long term gas supply contracts. Significant losses have been faced at present by pipeline importers due to this kind of oil-to-gas price decoupling; moreover, the structural market change determined an increased sense of uncertainty about European gas market future development. Interested readers may refer to [21] for a detailed and updated analysis of oil-to-gas decoupling.

This new market scenario has induced many long term importers to engage a renegotiation process with their suppliers together with a more focused attention towards optimization and hedging possibilities which are naturally embedded in the current contracts. In fact, traditionally long term forward gas contracts are often equipped with some volumetric flexibilities in order to intertemporally manage gas demand fluctuations year by year. Among those, in this market situation a new and particular importance arose for the so-called *make-up* and *carry forward* clauses, which flank traditional constraints as minimum and maximum withdrawal quantity established for every contract year and every contract sub-period (day or month). Basically, these clauses allow the buyer of the contract to delay or anticipate respectively the withdrawal of gas from one year to another within the full respect of sub-period capacity constraints. In particular, the introduction of make-up clauses has become very important for European long term contracts holders: in fact, in the recent oil-to-gas price decoupling situation, contracts holders were induced to delay as much as possible the gas delivery for the sake of loss minimization. With a make-up clause contract holders can effectively postpone the delivery of gas when it is too expensive with respect to market prices, hoping that future gas prices will rise up and the exercise of the contract rights becomes again profitable.

The optimization/valuation problem of standard swing contracts is not a trivial problem *per se*, as sub-period decisions typically impact the possibility of exercising the option in the future due to annual volume constraints. Thus, in the recent years swing options received vast treatment in the literature (see for instance [5, 20] and references therein for what concerns gas markets, and references in [3, 4, 6, 14] for swing options in more general markets). The presence of make-up clauses further complicate things and introduces more complexity. Surprisingly, the quantitative literature appears scarce: in the authors' opinion, this is due to the fact that a make-up clause is worth more in a market where price decoupling is high, and the need to study such markets arose only in the last years. At a qualitative level, for instance, the make-up clause is described in [22, 24]. An algorithm to evaluate a swing contract with the carry-forward clause using the least square Monte Carlo approach is presented in [14], where the authors claim that the make-up can be evaluated similarly: this is true only for make-up clauses with a single final installment, i.e. when the make-up gas is to be paid only when it is called back (which corresponds to letting $\alpha = 0$ in Rule 4 of Section 2.3). However, typical make-up clauses have a

double-installment mechanism, i.e. the make-up gas has to be paid when nominated with a price proportion $\alpha > 0$ and when called back with a price proportion $1 - \alpha$ (as an example, [22] report $\alpha \in (0.85, 1]$ as typical proportions). To the authors' knowledge, an algorithm to properly price (and find out optimal policy) a swing contract with such a general double-installment make-up clause and where both the market and strike price are stochastic variables has never been presented so far. The aim of this paper is exactly to fill this literature gap. In particular, we will describe, frame and solve the optimization issues related to the presence of make-up clauses in long term swing contracts. Finally, we use the algorithm in order to explore the value of the contract with respect to the peculiar constraints introduced by the make-up.

2 The structure of swing contracts and *make-up* clause

2.1 Time structure and admissible strategies

Ordinary swing contract schemes are normally defined dividing each one of the D yearly delivery periods $\{[T_{j-1}, T_j)\}_{j=1,\dots,D}$, into N sub-periods $\{[t_{j,i-1}, t_{j,i})\}_{i=1,\dots,N}^{j=1,\dots,D}$ obtaining the sequence $\{t_{j,i}\}$ such that

$$\begin{aligned} 0 = T_0 = t_{1,0} < t_{1,2} < \ldots < t_{1,N} = T_1 = t_{2,0} < t_{2,1} < \ldots \\ \ldots < t_{j,i} < \ldots < t_{j,N} = T_j = t_{j+1,0} < \ldots < t_{D,N} = T_D \end{aligned}$$

In particular, in every year $[T_{j-1}, T_j]$ we have the $N + 1$ points $(t_{j,i})_{i=0,\dots,N}$ such that $t_{j,0} = T_j$ and $t_{j,N} = T_{j+1}$.

We are also assuming that N is also the number of exercise swing rights the holder has in every year, which can be exercised exactly at the points $t_{j,i}$, for $i = 0, \dots, N-1$ i.e. at the beginning of every sub-period. For example if the decisions are taken month by month, at the beginning of evey month, $N = 12$, if day by day $N = 365$.

Denote by $u_{j,i}$ the quantity of gas the holder decides to buy in the sub-period $[t_{j,i}, t_{j,i+1})$, $i = 0, \dots, N-1$, and by $z_{j,j}$ the cumulated gas quantity at time $t_{j,i}$. In particular we set $z_{j,0} = 0$ for all $j = 1, \dots, D$ and

$$z_{j,i+1} = \sum_{k=0}^{i} u_{j,k} = z_{j,i} + u_{j,i} \quad \forall i \in \{0, \dots, N-1\} \tag{1}$$

Over each one of the N sub-periods, minimum (mDQ) and maximum (MDQ) delivery quantities are established in the contract, which usually reflect physical effective transportation capacity limitations: thus, the quantities $u_{j,i}$ are constrained by

$$\mathsf{mDQ} \leqslant u_{j,i} \leqslant \mathsf{MDQ} \qquad \forall i = 0, \dots, (N-1), \quad \forall j = 1, \dots, D \tag{2}$$

For every contractual year, minimum and maximum quantities are also established, called respectively minimum annual quantity (mAQ) and annual contract quantity (ACQ). The difference between the maximum gas that the holder could physically take and his contract right is thus given by

$$\overline{\mathcal{M}} := N \cdot \mathsf{MDQ} - \mathsf{ACQ} \tag{3}$$

while the difference between the minimum gas that the holder must take by contract and the minimum which he could physically take is given by

$$\underline{\mathcal{M}} := \mathsf{mAQ} - \mathsf{mDQ} \cdot N \tag{4}$$

Often we have non-trivial volume constraints, in the sense that

$$\overline{\mathcal{M}} > 0, \quad \underline{\mathcal{M}} > 0 \tag{5}$$

Thus, in the light of the discussion above, without any additional clauses and with non-trivial constraints we have

$$N \cdot \mathsf{mDQ} < \mathsf{mAQ} \leqslant z_{j,N} \leqslant \mathsf{ACQ} < N \cdot \mathsf{MDQ} \qquad \forall j = 1, \dots, D$$

Penalty payments can be imposed if the volume constraints are exceeded in order to stimulate the buyer to respect the volumetric limits imposed (see for example [5]), but in this paper we do not take into account these penalties.

The difference between swing contracts with trivial and non-trivial volume constraints is extremely important in the pricing and hedging of the contract itself. In fact, with non-trivial volume constraints the holder must take into account, at time $t_{j,i}$, not only the quantity $u_{j,i}$ which would be optimal for that period, but also the effects of this quantity on the future decisions that he will be allowed to take after. This brings to model the so-called *space of controls*, i.e. the set where $u_{j,i}$ is allowed to take values, in the following way. For a given year $j = 1, \dots, D$, assume that we have a final constraint $z_{j,N} \in [\underline{z}, \overline{z}]$ for some $0 \leqslant \underline{z} < \overline{z}$. ($[\underline{z}, \overline{z}] = [\mathsf{mAQ}, \mathsf{ACQ}]$ in the absence of make-up or other clauses). Then, for a given time $t_{j,i}$, the space of controls $\mathcal{A}(t_{j,i}, z_{j,i}, [\underline{z}, \overline{z}])$ will in general depend on time $t_{j,i}$, cumulated quantity $z_{j,i}$ and $[\underline{z}, \overline{z}]$.

By the constraints (2) and construction of $z_{j,i}$, at time $t_{j,i}$ we can restrict our attention to the case when $z_{j,i}$ satisfies the constraints

$$\mathsf{mDQ} \cdot i \leqslant z_{j,i} \leqslant \mathsf{MDQ} \cdot i \quad \forall i = 0, \dots, N$$

and

$$N \cdot \mathsf{mDQ} \leqslant \underline{z} \leqslant \overline{z} \leqslant N \cdot \mathsf{MDQ}$$

The problem of determining the set $\mathcal{A}_{j,i}$ is non-trivial when Eq. (5) holds, which translates in

$$N \cdot \mathsf{mDQ} < \underline{z} \leqslant \overline{z} < N \cdot \mathsf{MDQ}$$

(otherwise we can always reach the values in $[N \cdot \mathsf{mDQ}, N \cdot \mathsf{MDQ}]$). In this non trivial case, we are not allowed to take $u_{j,i} = \mathsf{mDQ}$ for all $i = 0, \dots, N-1$: in fact, there exists a time τ_1 such that, if we have always took this minimum for $t \leqslant \tau_1$, then for $t > \tau_1$ we have to switch to $u_{j,i} = \mathsf{MDQ}$ in order to reach $\underline{z}$. This point τ_1 is the common point between the two lines $z = \mathsf{mDQ}(t - t_{j,0})$ and $z = \mathsf{MDQ}(t - t_{j,N}) + \underline{z}$, $\forall t \in [t_{j,0}, t_{j,N}]$. A simple calculation leads to

$$z_{j,i} \geqslant r_{\min}(t_{j,i}, \underline{z}) = \max\left\{\mathsf{mDQ}(t_{j,i} - t_{j,0}), \mathsf{MDQ}(t_{j,i} - t_{j,N}) + \underline{z}\right\}$$

Similarly, we are not allowed to take always $u_{j,i} = \mathsf{MDQ}$ either: in fact, there exists a time τ_2 such that, if we have always took this maximum for $t \leqslant \tau_2$, then for $t > \tau_2$ we have to switch to $u_{j,i} = \mathsf{mDQ}$ in order to reach, and not exceed, $\overline{z}$. The boundary for $z_{j,i}$ in this case is

$$z_{j,i} \leqslant r_{\max}(t_{j,i}, \overline{z}) = \min\left\{\mathsf{MDQ}\left(t_{j,i} - t_{j,0}\right), \mathsf{mDQ}\left(t_{j,i} - t_{j,N}\right) + \overline{z}\right\}$$

Figure 2 shows an example of the admissible area.

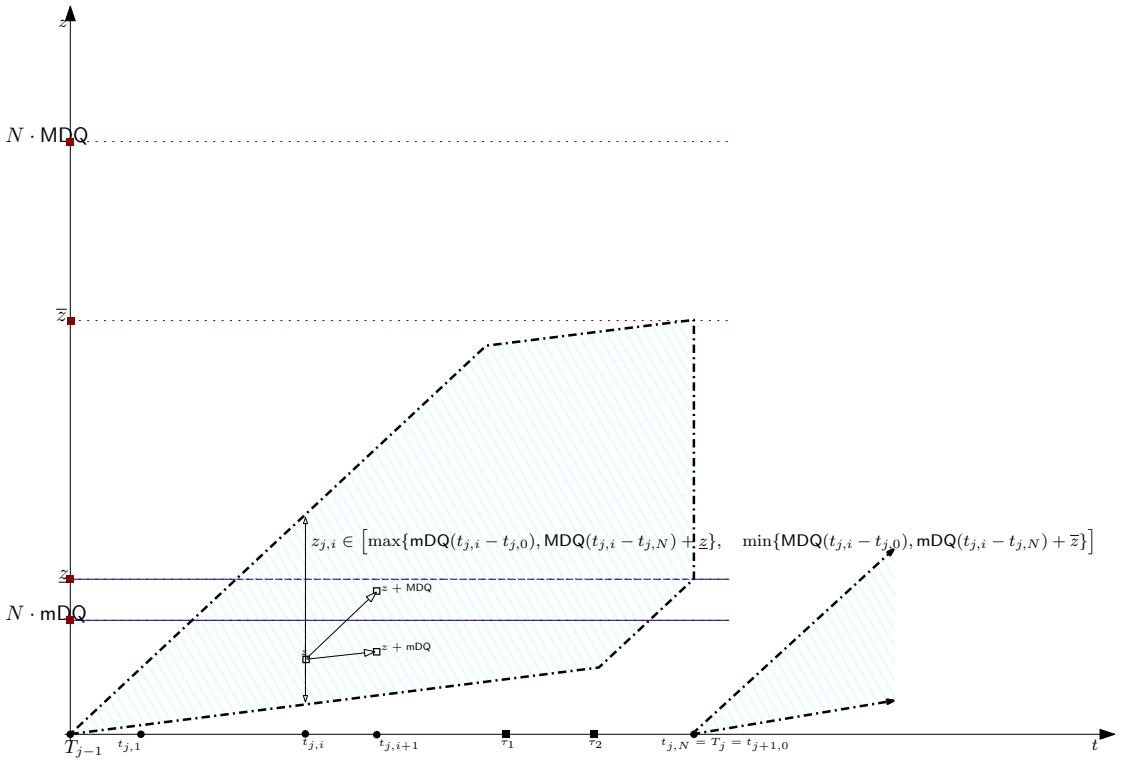


Figure 2: Typical admissible area for one year. Here $\underline{z} < \overline{z}$, leaving some optionality for the total intake $z_{j,N}$. If $\underline{z} = \overline{z}$ (typical of years when some make-up gas is nominated or called back), we have the constraint $z_{j,N} = \underline{z} = \overline{z}$ and the admissible region is like those in Figure 4.

In conclusion, the correct form of the space of controls $\mathcal{A}(t_{j,i}, z, [\underline{z}, \overline{z}])$ at time $t_{j,i}$, given the constraint $z_{j,N} \in [\underline{z}, \overline{z}]$ and the cumulated quantity $z_{j,i} = z$, is given by

$$\mathcal{A}(t_{j,i}, z, [\underline{z}, \overline{z}]) := \{u_{j,i} \in [\mathsf{mDQ}, \mathsf{MDQ}] \mid z + u_{j,i} \in [r_{\min}(t_{j,i+1}, \underline{z}), r_{\max}(t_{j,i+1}, \overline{z})]\} \tag{6}$$

which appears implicitly in [3, Equation 7] and is also a discretized version of the one in [6].

2.2 The price of a standard swing contract

We now present a standard procedure to price a swing option without the presence of additional clauses (such as make-up).

Let $P_{j,i}$ and $I_{j,i}$ be respectively the prices of gas and index in year $j = 1, \dots, D$, subperiod $[t_{j,i}, t_{j,i+1})$, $i = 0, \dots, N-1$: the contract holder has to buy the gas at the price $I_{j,i}$ and can sell it at the price $P_{j,i}$: of course with this notation we have $(P_{j,N}, I_{j,N}) = (P_{j+1,0}, I_{j+1,0})$ for each year $j = 1, \dots, D-1$. Even if $I = (I_{j,i})_{j=1,\dots,D, i=0,\dots,N}$ is a time average on several subperiod and thus has relevant memory effects, as mentioned in the Introduction, in the following we make the simplifying assumption that $(P_{j,i}, I_{j,i})_{j=1,\dots,D, i=0,\dots,N}$, evolve as a two-dimensional Markov process under a pricing measure $\mathbb{P}$, which is used in all the mathematical expectations that follow, while the numerical implementation that we use for the analysis of the next section will make use of the particular specification that we describe in Appendix A of [11]. We also assume that $\mathbb{E}[e^{-rt_{j,i}} P_{j,i}] = F^P_{j,i} < +\infty$ and $\mathbb{E}[e^{-rt_{j,i}} I_{j,i}] = F^I_{j,i} < +\infty$, where $F^P_{j,i}$, $F^I_{j,i}$ represent the forward prices of P and of I, respectively, for the delivery time $t_{j,i}$.

The objective of contract's holder is to maximize the discounted global margin of the contract (i.e. minimize the total loss), i.e., (s)he wants to calculate the value of

$$\begin{aligned} V(0, p_{1,0}, \iota_{1,0}, 0) &= \sup_{u \in \mathcal{A}} \mathbb{E}\left[\sum_{j=1}^{D} \sum_{i=0}^{N-1} e^{-rt_{j,i}} u_{j,i} \left(P_{j,i} - I_{j,i} \right) \right] \\ &= \sum_{j=1}^{D} \sup_{u \in \mathcal{A}} \mathbb{E}\left[\sum_{i=0}^{N-1} e^{-rt_{j,i}} u_{j,i} \left(P_{j,i} - I_{j,i} \right) \right] \end{aligned} \tag{7}$$

where the set $\mathcal{A}$ of *admissible controls* is defined by

$$\mathcal{A} := \{ (u_{j,i})_{j,i} \text{ adapted to } (P_{j,i}, I_{j,i})_{j,i} \text{ and s.t. } u_{j,i} \in \mathcal{A}(t_{j,i}, z_{j,i}, [\mathsf{mAQ}, \mathsf{ACQ}]) \}$$

in the absence of a make-up clause, and $r \geqslant 0$ is the risk-free annual interest rate. Equation (7) follows from the fact that $z_{j,0} = 0$ for all $j = 1, \dots, D$, i.e. in the absence of a make-up clause the swing contract can be valued independently year by year.

It is a standard result (see e.g. [3, 5, 20]), and it will also follow as a particular case of our results in Section 2.4, that this maximisation problem can be solved by the use of the Dynamic Programming: for each year $j = 1, \dots, D$, define the deterministic functions

$$V_j(N, p, \iota, z) := 0, \tag{8}$$

$$\begin{aligned} V_j(i, p, \iota, z) := \max_{u \in \mathcal{A}(t_{j,i}, z, [\mathsf{mAQ}, \mathsf{ACQ}])} & \mathbb{E}^{p,\iota}_{j,i} \left[e^{-rt_{j,i}} u \left(p - \iota \right) + \right. \\ & \left. + V_j(i+1, P_{j,i+1}, I_{j,i+1}, z + u) \right] \qquad \forall i < N \end{aligned} \tag{9}$$

where $\mathbb{E}^{p,\iota}_{j,i}$ indicates the expectation conditional to $P_{j,i} = p$ and $I_{j,i} = \iota$ (recall that, as these are Markov processes, these values are a sufficient statistics for the whole information up to subperiod i of year j). Then the original problem in Equation (7) is brought back to calculating

$$V(0, p_{1,0}, \iota_{1,0}, 0) = \mathbb{E}\left[\sum_{j=1}^{D} V_j(0, P_{j,0}, I_{j,0}, 0) \right]$$

2.3 Modeling the make-up clause

This subsection is devoted to the analytical representation of the make-up clause and its constraints. While long term contracts may have a length of 10-30 years, make-up clauses are typically written on a limited period of the contract life, often from 3 to 5 years. Given the fact that, as explained in Section 2.2, a contract without make-up clause can be evaluated as the sum of some yearly contract one independent from the other, we can split a contract with make-up written only on a subperiod of the whole contract life in two parts: the first part is a swing contract with a make-up clause with length equal to the original make-up clause, while the other part covers all the years when the make-up is not written. Thus, without loss of generality we can assume that the make-up clause is written on the whole contract's length, D years.

For each year $j = 1, \dots, D$, call M_j the make-up gas nominated and U_j the make-up gas called back in year j. With this notation, we assume that the precise structure of the make-up clause follows these rules.

(a) *For each year $j = 1, \dots, D-1$, the contract holder is allowed to take $z_{j,N} < \mathsf{mAQ}$, provided $u_{j,i} \geqslant \mathsf{mDQ}$ for all $i = 0, \dots, N-1$.*

Thus, the make-up gas nominated in year j is

$$M_j := (\mathsf{mAQ} - z_{j,N})^+ \quad \text{and must satisfy} \quad M_j \in [0, \underline{\mathcal{M}}] \tag{10}$$

where $x^+ := \max(x, 0)$ and $\underline{\mathcal{M}}$, defined in Equation (4), is also the maximum quantity of make-up gas that can be physically nominated in a given year.

(b) *The make-up M_j nominated in year j can be called back in one or more subsequent years (the quantity M_j can be splitted and called back in more than one year). This is possible only if the ACQ quantity has been reached in that year, and of course in that year we still have to satisfy $u_{j,i} \leqslant \mathsf{MDQ}$ for all $i = 1, \dots, N$.*

Thus, the make-up gas called back in year $j = 2, \dots, D$ is

$$U_j := (z_{j,N} - \mathsf{ACQ})^+ \quad \text{and is such that} \quad U_j \in [0, \overline{\mathcal{M}}] \tag{11}$$

where $\overline{\mathcal{M}}$, defined in Equation (3), is also the maximum quantity of make-up gas that can physically called back in a given year.

(c) *It is not possible to call back make-up gas before having nominated it, and at year D all the nominated make-up gas must have been called back.*

Thus, if we define the cumulated gas debt at year j, i.e. the make-up gas not yet called back, as

$$\boldsymbol{M}_j = \sum_{k=1}^{j} M_k - \sum_{k=2}^{j} U_k = \sum_{k=1}^{j} (M_k - U_k), \tag{12}$$

then $U_1 = M_D = 0$, $\boldsymbol{M}_j \geqslant 0$ for all $j = 1, \dots, D-1$ and $\boldsymbol{M}_D = 0$. Moreover,

$$\boldsymbol{M}_{j+1} = \boldsymbol{M}_j + M_j - U_j = \boldsymbol{M}_j + (\mathsf{mAQ} - z_{j,N})^+ - (z_{j,N} - \mathsf{ACQ})^+$$

Notice that conditions 2. and 3. imply, for example, that if at the beginning of the last contract year of the make-up clause we have some make-up gas not called back, i.e. $\boldsymbol{M}_{D-1} > 0$, in year d we necessarily have to reach the quantity $\mathsf{ACQ} + \boldsymbol{M}_{D-1}$.

Remark 1 More in general, for all years $j = 1, \dots, D$, the definition of $\boldsymbol{M}_j$ implies that $\boldsymbol{M}_j \leqslant j \cdot \underline{\mathcal{M}}$ and $\boldsymbol{M}_j \leqslant (D-j) \cdot \overline{\mathcal{M}}$. By combining these two constraints, the maximum gas debt is possible at year

$$\bar{j} := \frac{D\overline{\mathcal{M}}}{\underline{\mathcal{M}} + \overline{\mathcal{M}}} \tag{13}$$

if $\bar{j}$ is integer, and at one of the two nearest years if $\bar{j}$ is not integer. In particular, the gas debt $\boldsymbol{M}_j$ can increase without constraints for $j < \bar{j}$ and must possibly be decreased for $j > \bar{j}$.

(d) *The price of the make-up quantity nominated in year j and called back in year k, subperiod i, is defined as the weighted sum of two components respectively paid at two different times:*

a) at time $t_{j,N}$ (i.e. at the end of year j when M_j becomes known) the buyer pays the make-up gas at the price $\alpha\overline{\Gamma}_j$ for some $\alpha \in (0,1)$ defined in the contract, where $\overline{\Gamma}_j$ is the average index price observed in year j;

b) at time of withdrawal $t_{k,i}$, the price paid is $(1-\alpha)I_{k,i}$.

The price of make-up gas, as defined above, is associated to the gas volume $u_{k,i}$ physically delivered at time $t_{k,i}$. This means that the part $\alpha\overline{\Gamma}_j$ in (a) of the price needs to be capitalized from time $T_j = t_{j,N}$ up to time $t_{k,i}$: thus, the price $\mathcal{I}_{j,k,i}$ at time $t_{k,i}$ of the make-up gas nominated in year j and called back in year k, subperiod i is

$$\mathcal{I}_{j,k,i} = \alpha\overline{\Gamma}_j e^{r\left(t_{k,i} - t_{j,N}\right)} + (1-\alpha)I_{k,i} \tag{14}$$

By discounting at time $T_0 = 0$ the price of make-up gas called back at time $t_{k,i}$, we have

$$e^{-rt_{k,i}}\mathcal{I}_{j,k,i} = \alpha\overline{\Gamma}_j e^{-rt_{j,N}} + (1-\alpha)I_{k,i}e^{-rt_{k,i}}$$

It follows that, in a year $j = 1, \dots, D$ where the make-up clause is exercised to nominate or call back gas, the residual value of the swing contract at subperiod $i = 0, \dots, N-1$ for that year with the control policy $u_j := (u_{j,i})_{i=0,\dots,N-1}$ is given by

$$J_j(i, p, \iota, z_{j,i}; u_j) := \mathbb{E}^{p,\iota}_{j,i}\left[\sum_{k=i}^{N-1} e^{-rt_{j,k}} u_{j,k}\left(P_{j,k} - A_\alpha(u_{j,k}, z_{j,k})I_{j,k}\right) - e^{-rt_{j,N}}\alpha\overline{\Gamma}_j M_j\right] \tag{15}$$

where

$$A_\alpha(u,z) := 1-\alpha\left(1-\frac{\mathsf{ACQ}-z}{u}\right)^+ \mathbf{1}_{\{\mathsf{ACQ}-\mathsf{MDQ}<z\leqslant\mathsf{ACQ}\}} - \alpha\mathbf{1}_{\{z>\mathsf{ACQ}\}}$$

is a pricing coefficient in the interval $[1-\alpha,1]$ for $I_{j,k}$ to accomodate the gas quality (ordinary below ACQ, called back from previous years above). Here, the main apparent difference with respect to the case when no make-up is exercised is that we can end up a year with a non-null position in the make-up gas, i.e. with $M_j - U_j \neq 0$ (notice that this notation is not ambiguous as M_j and U_j cannot be both different from zero), where the quantity $M_j - U_j$ is by definition a deterministic function of $z_{j,N}$, thus of u_j. Notice that this generalizes the payoff to be maximised in Equation (7), which is reobtained for $z_{j,N} \in [\mathsf{mAQ},\mathsf{ACQ}]$, i.e. $M_j = U_j = 0$, and setting $z_{j,0} = 0$.

As a result, the total value of the swing option with the make-up clause described above is given by

$$\sup_{u\in\mathcal{A}} \mathbb{E}\left[\sum_{j=1}^{D} J_j(0,p_{j,0},\iota_{j,0},0;u_j)\right] \tag{16}$$

where $u=(u_{j,i})_{j=1,\dots,D,i=0,\dots,N-1}$ must now belong to the set

$$\mathcal{A} := \left\{ \begin{array}{l} (u_{j,i})_{j,i} \text{ adapted to } (P_{j,i},I_{j,i})_{j,i} \mid u_{j,i}\in[\mathsf{mDQ},\mathsf{MDQ}], \\ \\ \qquad U_1 = M_D = 0, \boldsymbol{M}_j \geqslant 0 \ \forall j=1,\dots,D, \boldsymbol{M}_D = 0 \end{array} \right\}$$

and the functions J_j are given by Equation (15). Here, the constraints on U and M induce constraints on $(u_{j,i})_{j,i}$, which will be treated in detail in the next subsection.

2.4 The price of swing contracts with make-up clauses

We just saw that the set of admissible strategies for $(u_{ji})_{ji}$ for a given year $j=1,\dots,D$ now depend on the gas debt $\boldsymbol{M}_{j-1}$ arriving from the previous years. For this reason, this quantity has to be explicitly taken into account in the evaluation of the swing contract for that year. More in detail, now we define the *value function* as

$$V_j(i,p,\iota,z,\boldsymbol{M}_{j-1}) := \sup_{u\in\mathcal{A}} \mathbb{E}^{p,\iota}_{j,i}\left[J_j(i,p,\iota,z;u_j) + \sum_{k=j+1}^{D} J_k(0,P_{k,0},I_{k,0},0;u_k)\right] \tag{17}$$

We now build a Dynamic Programming algorithm as in [5, 20]: for each year $j=1,\dots,D$, define the deterministic functions:

- if $j=D$, then $M_D=0$ and $U_D=\boldsymbol{M}_{D-1}$ (recall that $\boldsymbol{M}_D\equiv 0$), so we let

$$\underline{z} := \mathsf{mAQ}\mathbf{1}_{\{\boldsymbol{M}_{D-1}=0\}} + (\mathsf{ACQ}+\boldsymbol{M}_{D-1})\mathbf{1}_{\{\boldsymbol{M}_{D-1}>0\}}, \qquad \overline{z} := \mathsf{ACQ}+\boldsymbol{M}_{D-1} \tag{18}$$

and define

$$V_j(N,p,\iota,z,\boldsymbol{M}_{j-1}) := 0, \tag{19}$$

$$V_j(i,p,\iota,z,\boldsymbol{M}_{j-1}) := \max_{u\in\mathcal{A}(t_{j,i},z,[\underline{z},\overline{z}])} \mathbb{E}^{p,\iota}_{j,i}\Big[e^{-rt_{j,i}}u\,(p-A_\alpha(u,z)\iota)+ \\ +V_j(i+1,P_{j,i+1},I_{j,i+1},z+u,\boldsymbol{M}_{j-1})\Big] \tag{20}$$

Notice that the functions V_j depend on $\boldsymbol{M}_{j-1}$ through $[\underline{z},\overline{z}]$.

- for $j=1,\dots,D-1$ the key quantity is $\boldsymbol{M}_{j-1}$ which is known at the beginning of the year. Assume that $\boldsymbol{M}_{j-1}\leqslant(D-j+1)\overline{\mathcal{M}}$. For the lower bound we have two cases.
  - If $\boldsymbol{M}_{j-1}$ is admissible also for year j, i.e. if $\boldsymbol{M}_{j-1}\leqslant(D-j)\overline{\mathcal{M}}$, then we can nominate some other make-up gas M_j as long as

    $$\boldsymbol{M}_j=\boldsymbol{M}_{j-1}+M_j\leqslant(D-j)\overline{\mathcal{M}}\Rightarrow M_j\leqslant(D-j)\overline{\mathcal{M}}-\boldsymbol{M}_{j-1}$$

    Taking into account the mDQ constraints, the lower bound for $z_{j,N}$ is:

    $$\underline{z}:=\mathsf{mAQ}-\min\left\{(D-j)\overline{\mathcal{M}}-\boldsymbol{M}_{j-1},\underline{\mathcal{M}}\right\} \tag{21}$$

  - If $\boldsymbol{M}_{j-1}$ is not admissible for year j, i.e. if $\boldsymbol{M}_{j-1}>(D-j)\overline{\mathcal{M}}$, then we must call back some make-up gas in order to obtain a final cumulated quantity $\boldsymbol{M}_j$ admissible for year j, i.e.

    $$\boldsymbol{M}_j=\boldsymbol{M}_{j-1}-U_j\leqslant(D-j)\overline{\mathcal{M}}\Rightarrow U_j\geqslant\boldsymbol{M}_{j-1}-(D-j)\overline{\mathcal{M}}$$

    So the lower bound for $z_{j,N}$ is now

    $$\underline{z}:=\mathsf{ACQ}+\boldsymbol{M}_{j-1}-(D-j)\overline{\mathcal{M}} \tag{22}$$

For the upper bound $\overline{z}$, we do not need to distinguish between the two previous cases and let

$$\overline{z}:=\mathsf{ACQ}+\min\left\{\boldsymbol{M}_{j-1},\overline{\mathcal{M}}\right\} \tag{23}$$

Finally, define

$$V_j(N,p,\iota,z,\boldsymbol{M}_{j-1}) := V_{j+1}(0,p,\iota,0,\boldsymbol{M}_{j-1}+(\mathsf{mAQ}-z)^+-(z-\mathsf{ACQ})^+)+ \\ +e^{-rt_{j,N}}\alpha\Gamma_j(\mathsf{mAQ}-z)^+, \tag{24}$$

and for $i=N-1,\dots,0$ define $V_j(i,\cdot,\cdot,\cdot,\cdot)$ exactly as in Equation (20).

Theorem 2.1

1. *The deterministic functions $V_j(\cdot,\cdot,\cdot,\cdot,\cdot)$, defined by the dynamic programming equations (19), (20) and (24) are such that $V_1(0, P_{1,0}, I_{1,0}, 0, 0)$ coincides with the value of the swing option with the make-up clause in Equation (16).*

2. *There exists an optimal Markovian consumption $u^*_{j,i} = u(t_{j,i}, P_{j,i}, I_{j,i}, z_{j,i}, \boldsymbol{M}_{j-1})$, where $u(\cdot,\cdot,\cdot,\cdot,\cdot)$ is given by the maximum argument in the dynamic programming equation (20).*

3. *If the quantities*

$$\underline{K} := \frac{\underline{\mathcal{M}}}{\mathsf{MDQ} - \mathsf{mDQ}} \qquad \text{and} \qquad \overline{K} := \frac{\overline{\mathcal{M}}}{\mathsf{MDQ} - \mathsf{mDQ}}, \tag{25}$$

are integer, then there exists an optimal bang-bang Markovian consumption $u^*_{j,i}$, i.e. $u^*_{j,i} = \mathsf{mDQ}$ or $u^*_{j,i} = \mathsf{MDQ}$ for all $j = 1, \dots, D$, $i = 0, \dots, N-1$. Moreover, $\boldsymbol{M}_j$ turns out to be an integer multiple of $\mathsf{MDQ} - \mathsf{mDQ}$ for all $j = 1, \dots, d$.

Proof. We proceed in analogy with [5] and [4].

(a) As $0 \leqslant t_{j,i} \leqslant T$, $0 \leqslant z \leqslant N \cdot \mathsf{MDQ}$, $\boldsymbol{M} \leqslant D \cdot \underline{\mathcal{M}}$ and $\mathbb{E}[P_{j,i}] = F^P_{j,i} < +\infty$, $\mathbb{E}[I_{j,i}] = F^I_{j,i} < +\infty$, then the assumptions ($\mathrm{F}^+$, F^-) in [7, Proposition 8.5] are satistied, so the argument follows.

(b) The right-hand side of Equation (20) is continuous in u and $\mathcal{A}(t_{j,i}, z, [\underline{z}, \overline{z}])$ is a compact set contained in $[\mathsf{mDQ}, \mathsf{MDQ}]$, thus the maximum is attained for $u \in \mathcal{A}(t_{j,i}, z, [\underline{z}, \overline{z}])$ again by applying [7, Proposition 8.5].

(c) As in [4], it can be proved that the functions $V_j(i,\cdot,\cdot,\cdot,\cdot)$, $j = 1, \dots, D$, $i = 0, \dots, N-1$ are continuous and concave on z and piecewise affine on the intervals

$$[k \cdot \mathsf{mDQ} + (i-k) \cdot \mathsf{MDQ}, (k+1) \cdot \mathsf{mDQ} + (i-k-1) \cdot \mathsf{MDQ}], \quad k = 1, \dots, i. \tag{26}$$

We now prove the claim by induction on $j = 1, \dots, D$. If $j = 1$, then $\boldsymbol{M}_0 = 0$ by definition. For a given year $j = 1, \dots, D$, assume for now that $\boldsymbol{M}_{j-1}$ is an integer multiple of $\mathsf{MDQ} - \mathsf{mDQ}$. Then this, together with the condition $\underline{K}, \overline{K} \in \mathbb{N}$ ensures that $[r_{\min}(t_{j,i}, \underline{z}), r_{\max}(t_{j,i}, \bar{z})]$ is exactly the union of suitable intervals of the kind of Equation (26). Thus, if $z = k \cdot \mathsf{mDQ} + (i-k) \cdot \mathsf{MDQ}$ for some $k = 0, \dots, i$, then the function to be maximised in Equation (20) is affine on u, thus its maximum point is $u^*_{j,i} = \mathsf{mDQ}$ or $u^*_{j,i} = \mathsf{MDQ}$. It can then be proved by induction that, since $z_{j,0} = 0$, the optimal $u^*_{j,i}$ is such that $z_{j,i} = k \cdot \mathsf{mDQ} + (i-k) \cdot \mathsf{MDQ}$ for some $k = 0, \dots, i$: this also implies that $\boldsymbol{M}_j$ will be also an integer multiple of $\mathsf{MDQ} - \mathsf{mDQ}$, and the conclusion follows.

□

Remark 2 Part 3. of the theorem above is essentially a consequence of the linear structure of the payoff function in Equations (15)-(16): the result is that in every year j, subperiod i, the optimal quantity $u_{j,i}$ can be safely chosen to be either the maximum (MDQ) or the minimum (mDQ) admissible for that substep. This kind of control is called of *bang-bang* type, and it was already found in [5] with smoother payoffs, and studied in deep detail in [4]. Qualitatively, this is due to the fact that, if the withdrawal is profitable in the subperiod, then the better choice is the maximum quantity we can take; conversely, if the withdrawal is not profitable, then the better choice is to take the minimum quantity we can.

2.5 Computational cost

As we have seen, the pricing problem for a swing option with make-up clause boils down to maximize the problem in Equation (17). Unfortunately, this maximization cannot be carried out by analytic means, as a closed form for V_j is not known even in the simplest case of a standard swing option without make-up clause. Thus, this maximization must be carried out via numerical methods.

The most efficient way to do this is to assume that the quantities $\underline{K}$ and $\overline{K}$ in Equation (25) are integer, so that the results of Theorem 2.1 hold. This induces a quantization in the candidate optimal make-up gas debt $(\boldsymbol{M}_j)_{j=1,\dots,D}$: in fact, since this process at optimality has values which are multiple integers of $\mathsf{MDQ} - \mathsf{mDQ}$, we obtain that the resulting candidate optimal quantities for $\boldsymbol{M}_j$, $j = 1, \dots, D$, are a finite number. More in detail, the sequence $(\boldsymbol{M}_j)_{j=1,\dots,D}$ is bound to have a finite number of nonnegative values in each year $j = 1, \dots, D-1$, with $-\overline{\mathcal{M}} \leqslant \boldsymbol{M}_j - \boldsymbol{M}_{j-1} \leqslant \underline{\mathcal{M}}$, i.e. the increments can have at most $\underline{K} + \overline{K} + 1$ distinct values, corresponding respectively to the cases when $M_j > 0$, $U_j > 0$ and $M_j = U_j = 0$.

With this in mind, we can calculate the computational cost needed to price a D-year swing option with make-up clause, and we do this by the same backward recursion used in the Dynamic Programming algorithm used in Section 2.4. In the D-th year, we can start with $\boldsymbol{M}_{D-1}$ taking at most $\overline{K} + 1$ different values, each one of this leading to a different optimization problem: since also the values of $(z_{j,i})_{j,i}$ are quantized via the bang-bang optimal process $(u_{j,i})_{j,i}$, for each one of this optimization problem we have a total of $O(N^2)$ states which can be assumed by $(z_{j,i})_{j,i}$ at optimality[(8)]. Having solved the $\overline{K}+1$ problems for the last year, we can attach the value functions thus obtained to the terminal nodes of year $D-1$: notice that also in this case, as now $\boldsymbol{M}_{D-2}$ can assume at most $2\overline{K}+1$ distinct values, we will have to model and solve at most $2\overline{K} + 1$ distinct optimization problem, each one having as terminal condition the values of the $\overline{K} + 1$ problems of year D. These numbers do not multiply, because once that we obtain the values for the $\overline{K} + 1$ problems of year D for each possible starting state $(P_{D-1,N}, I_{D-1,N}) = (P_{D,0}, I_{D,0})$, we can take these values as terminal values to use in the computation for year $D - 1$.

With this spirit, we are now ready for a result on the computational cost of the pricing of a swing option with make-up clause.

[(8)] precisely $\leqslant \frac{N(N+1)}{2}$, which is the number of nodes of a complete recombining binomial tree.

Theorem 2.2 *If the quantities $\underline{K}$ and $\overline{K}$, defined in Equation (25), are integer, then the order of distinct subproblems to be solved is $O(N^2D^2)$.*

Proof. First of all consider the 2-dimensional process $(j, \boldsymbol{M}_j)_{j=1,\dots,D}$: then the distinct states that this process can assume, at optimality, is in 1-1 correspondence with the integer solutions (x, y) of the system

$$\left\{ \begin{array}{rcl} x & \geqslant & 0, \\ x & \leqslant & \overline{K}(D-y). \\ x & \leq & \underline{K}y, \end{array} \right. \tag{27}$$

In fact, if (x, y) is such a solution, then $\boldsymbol{M}_x = (\mathsf{MDQ} - \mathsf{mDQ})y$ is a possible value, at time x, of an optimal path for $\boldsymbol{M}$ by Theorem 2.1. Conversely, by the same theorem, if $\underline{K}$ and $\overline{K}$ are integer then $\boldsymbol{M}_x$ is a integer multiple of $\mathsf{MDQ} - \mathsf{mDQ}$. Since for each of these possible states we must solve a separate optimization problem for the corresponding year, the number of optimization subproblems for all the values of $(z_{j,i})_{j,i}$ are of order $O(N^2)$, and their total number is the sum of these, the proof boils down to find the total number $\mathbf{N}$ of integer solutions of the system (27). By recalling the definition of $\bar{j}$ in Equation (13) and the discussion below, first of all we rewrite $\bar{j}$ as

$$\bar{j} = \frac{D\overline{K}}{\underline{K} + \overline{K}}$$

Now, the region of the solutions of the system (27) is the union of the two triangular regions $\{(x, y) \in \mathbb{N}^2 \mid x \geqslant 0, x \leqslant \underline{K}y, y \leqslant \bar{j}\}$ and $\{(x, y) \in \mathbb{N}^2 \mid x \geqslant 0, x \leqslant \overline{K}(D-y), y > \bar{j}\}$. It is then easy to see that

$$\begin{aligned} \mathbf{N} &= \sum_{\ell=0}^{[\bar{j}]}(1 + \underline{K}\ell) + \sum_{\ell=[\bar{j}]+1}^{D}(1 + \overline{K}(D-\ell)) = D + 1 + \underline{K}\sum_{\ell=1}^{[\bar{j}]}\ell + \overline{K}\sum_{\ell=[\bar{j}]+1}^{D-1}(D-\ell) = \\ &= D + 1 + \underline{K}\frac{[\bar{j}]\cdot[\bar{j}+1]}{2} + \overline{K}\frac{[D-\bar{j}-1]\cdot[D-\bar{j}]}{2} \end{aligned}$$

where $[x]$ denotes the integer part of x. By noticing that for all $x > 0$ we have $D[x] \leqslant [Dx]$ and that

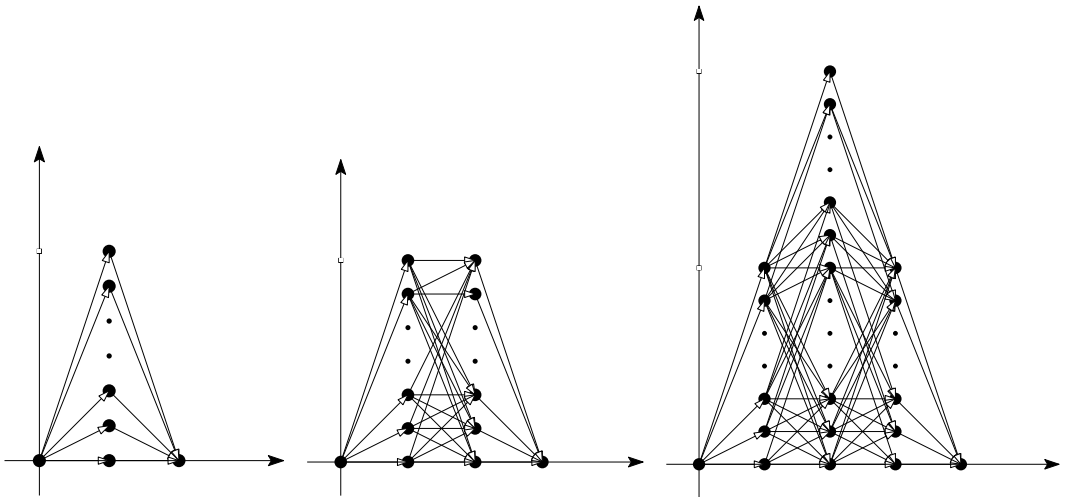
$$\underline{K}\cdot\bar{j} = \frac{D\underline{K}\overline{K}}{\underline{K} + \overline{K}} = \overline{K}(D - \bar{j})$$

then we have

$$\mathbf{N} \leqslant D + 1 + \frac{1}{2}\left[\frac{D\underline{K}\overline{K}}{\underline{K} + \overline{K}}\right]([\bar{j}+1] + [D-\bar{j}-1]) \leqslant \frac{1}{2}\frac{\underline{K}\overline{K}}{\underline{K} + \overline{K}}D^2 + D + 1$$

i.e. $\mathbf{N} = O(D^2)$. By multiplying this for $O(N^2)$ (the number of subproblems for given year $j = 1, \dots, D$ and state of make-up debt $\boldsymbol{M}_j$, we obtain that the computational cost is of order $O(N^2D^2)$, as desired. □

We show in Figure 3 an illustration of these numbers for $D = 2, 3, 4$, by making the simplifying assumption that $\underline{K} = \overline{K} =: K$.



(a) $D = 2$ (b) $D = 3$ (c) $D = 4$

Figure 3: In subfigure (a), we can only obtain $\boldsymbol{M}_1 (= M_1 = U_2)$ among $K + 1$ distinct values, and the corresponding value function is then used in the final values of the optimization problem of year 1, so the total number of optimization problems to be solved is $K + 1$. In subfigure (b) we can obtain $\boldsymbol{M}_2 (= U_3)$ among $K+1$ distinct values and then $\boldsymbol{M}_1 (= M_1)$ among $K+1$ distinct values, so the number of optimization problems to be solved in sequence is $2K + 2$. In subfigure (c) we can obtain $\boldsymbol{M}_3 (= U_4)$ among $K + 1$ different values, $\boldsymbol{M}_2$ among $2K + 1$ different values, and finally $\boldsymbol{M}_1 (= M_1)$ among $K + 1$ distinct values: so the number of optimization problems to be solved in sequence is now $4K + 3$.

We also present a numerical test which validates our result of a quadratic cost in the number of years. By taking the same parameters as in Section 5, we implemented the method on a Intel i7 workstation at 3.4GHz with 8GB RAM, with the following execution times.

Duration D of the make-up clause (years)	1	2	3	5	10	15
time (seconds)	0.06	0.12	0.31	1.25	7.60	23.65

Table 1: Execution times, in seconds, on a Intel i7 workstation at 3.4GHz with 8GB RAM. Notice that with 1 year there is no possibility *de facto* to exercise the make-up clause.

3 Sensitivity analysis of a three years contract

A swing contract is a derivative product whose value depends on two main classes of factor, namely market and volumetric. As previously explained in this paper, this kind of derivative shows an optionality value linked to the market price dynamics of the underlying commodity (exercise or not) and an optionality value linked to the volumetric structure of the product itself (how much to allocate with the make-up clause among the years and how much to withdraw in each subperiod). After having explained how to price a swing product on gas and how to determine the optimal exercise policy, it is now interesting to use the algorithm in order to explore and map the value of the contract with respect to some peculiar parameters of the contract and to market factors.

More in detail, we specify a trinomial dynamics for both the price P and the index I which approximates a geometric mean-reverting Ornstein-Uhlenbeck process as described in Appendix A of [11], and calibrate these models following [8], using historical data on TTF prices for the gas price P and the ENIGR07 formula[(9)] for the index price I. For ease of implementation, the average index price $\overline{\Gamma}_j$ of year j which appears in Equation (24) is substituted with the average of forward prices for that year. When not variable, the parameters used in this section are the ones in Table 2.

Parameter		*Value*	*Parameter*		*Value*
ACQ	=	$7.00 \cdot 10^6$	σ^P	=	0.6
mAQ	=	$6.00 \cdot 10^6$	a^P	=	2.95
MDQ	=	$8.75 \cdot 10^5$	σ^I	=	0.1
mDQ	=	$3.75 \cdot 10^5$	a^I	=	19.04
α	=	0.75	S	=	0
r	=	0.05	ρ	=	0

Table 2: Values of the parameters used for the analysis (when not variable).

We here present three analyses: the first one with respect to the volatility level σ^P of gas price, to the MDQ contract parameter, and to the level of market price decoupling. The second one is done with respect to the level of decoupling of the price term structure and to interest rates level. Finally, the third one is done with respect to correlation between P and I and level of decoupling.

The choice of these analyses have been done considering the aim of what we are pursuing, that is to analyse the flexibility given by the make-up clause in a decoupled market scenario. In view of this, we decided to change the parameters we believe to be more impactive on the value of the make-up clause. The volatility σ^P is representative of market uncertainty: in fact, σ^P is often much greater than σ^I, as the index I is calculated as a time average of a basket; as mentioned in the Introduction, this averaging is used to

[(9)] The ENIGR07 (ENI Gas Release 2007) index is a 9-months time average of a basket of three oil-related indexes, computed as in [1, Equation (1)] or in [14, Equation (1)]

reduce the volatility of the index and leads also to a pretty stable value for σ^I, so changes in σ^P are likely to influence the price more than ones in σ^I. The choice of MDQ is explained by the fact that this quantity is strictly linked with the maximum make-up $\overline{\mathcal{M}}$ the owner of the contract can call back in every year. In fact, the bigger MDQ is, the bigger $\overline{\mathcal{M}}$ becomes, and higher the possibility of the owner becomes to posticipate the calling back of the nominated make-up gas. This flexibility should increase the contract value, in particular when price decoupling is strong. We have decided not to move the minimum quantities. On one hand we set the minimum annual quantity and the minimum period quantity in such a way that the possible make-up one can nominate every year is very high ($1.5 \cdot 10^6$), so the stronger constraints are on $\overline{\mathcal{M}}$. On the other hand, we imposed the values of $\overline{K}, \underline{K}$ to be integer and we used values for MDQ in Table 3. The underlying idea is that any possible increase in the callable make-up quantity $\overline{\mathcal{M}} = \mathsf{mAQ} - N \cdot \mathsf{mDQ}$ is worthless if the upper bound of gas withdrawal per year $\overline{\mathcal{M}}$ is not enough to call back the nominated make-up quantity. Thus, we map the contract value for MDQ in the range between $\frac{\mathsf{ACQ}}{N} = \frac{7 \cdot 10^6}{12} \simeq 5.83 \cdot 10^5$, which reduces to the case of a standard contract without make-up clause[(10)], and a value big enough to ensure the withdrawal in the third year of the possible make-up gas nominated in the first and second year, i.e. bigger than $\frac{\mathsf{ACQ}+2(\mathsf{mAQ}-N\cdot\mathsf{mDQ})}{N} \simeq 8.3 \cdot 10^5$ and such that $\underline{K}, \overline{K}$ are integers.

MDQ	$\overline{\mathcal{M}}$	$\underline{\mathcal{M}}$	*Description*
$5.83 \cdot 10^5$	0	$1.5 \cdot 10^6$	No make up
$6.25 \cdot 10^5$	$5 \cdot 10^5$	$1.5 \cdot 10^6$	Low flexibility
$8.75 \cdot 10^5$	$3.5 \cdot 10^6$	$1.5 \cdot 10^6$	High flexibility

Table 3: Values of MDQ used in the analysis. All the other parameters, when not variable, are set as in Table 2.

The choice of changing MDQ and not other parameters is also a consequence of the practice: we think that the minimum annual quantity and the minimum period quantity are less negotiated than the maximum ones: the seller of the contract will never be willing to sell too much flexibility at the expense of its profits (he want to sell the physical gas), and the buyer will not pay too much for some flexibility he will probably not use in the future (he need the physical gas).

The second and third analyses mainly focus on market factors. As already stated, the make-up clause becomes profitable for the buyer of the contract only if the spread between market and index price $P_t - I_t$ is expected to be lower in the future than in the present. On the other hand, the make-up gas is paid in two different times and its price is affected by the interest rate, as seen in Eq. (14). Consequently, the benefits of the decoupling could be affected by high levels of interest rates, which potentially may vanish the power

[(10)] in fact, if if $\mathsf{MDQ} = \frac{\mathsf{ACQ}}{N}$ then, being not possible to call back any make-up gas before having reached ACQ, we are never able to call back any make-up gas, thus it is also impossible to nominate some.

of make-up clause. This is the focus of the second analysis. Also the correlation could potentially affect the benefits given by the decoupling: in principle, the decoupling should be enforced by negative correlation and weakened by positive one. This is the subject of the third analysis.

First Analysis. The first analysis studies how the contract value depends on the volatility level σ^P, on the MDQ contract parameter and on the level of decoupling. The latter is obtained by varying the initial forward prices used to calibrate the tree prices (see Appendix A of [11]), subtracting a level S from the forward prices F^P for the first year and adding the same quantity to the forward prices for the third year, as shown in Subfigure 6(b). Then we let S be a parameter and see how the swing price depends on it.

We expect the swing contract value to be increasing in σ^P, with a higher dependence when there is no flexibility given either by the absence of a make-up clause or by small values of MDQ. Figure 4 shows exactly these qualitative intuitions. The contract value is increasing with respect to σ^P also for high values of MDQ, but the range in y axes in the figure is so large that we may not appreciate the monotonicity of the curves. This also evidences the fact that the rights given by make-up reduce the risk given by market uncertainty.

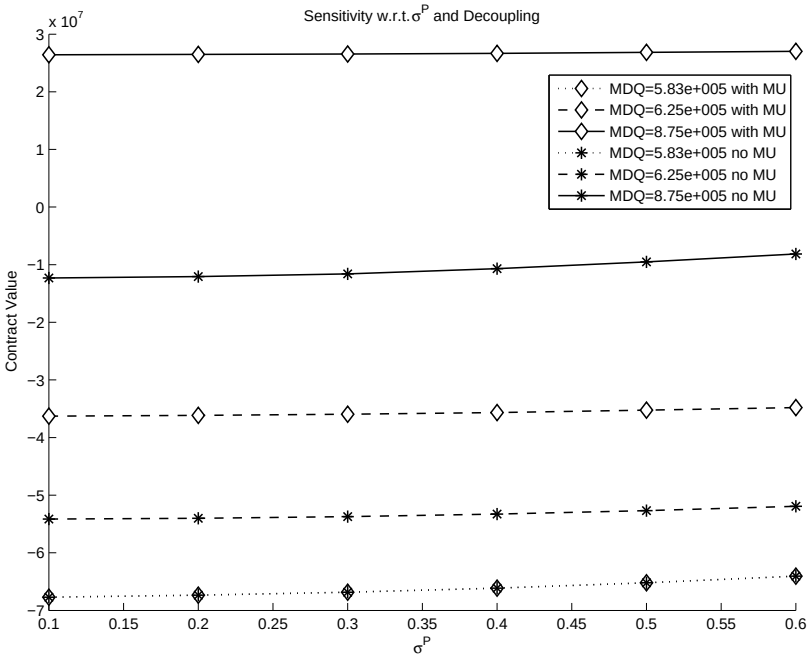


Figure 4: Sensitivity with respect to σ^P for three values of MDQ.

The dependence between contract value and decoupling parameter S is presented in Figure 5: make-up rights are useful when market decoupling is high. In these situations, we can nominate make-up gas at the beginning of the contract life and call it back in the future, when a positive market scenario shows up.

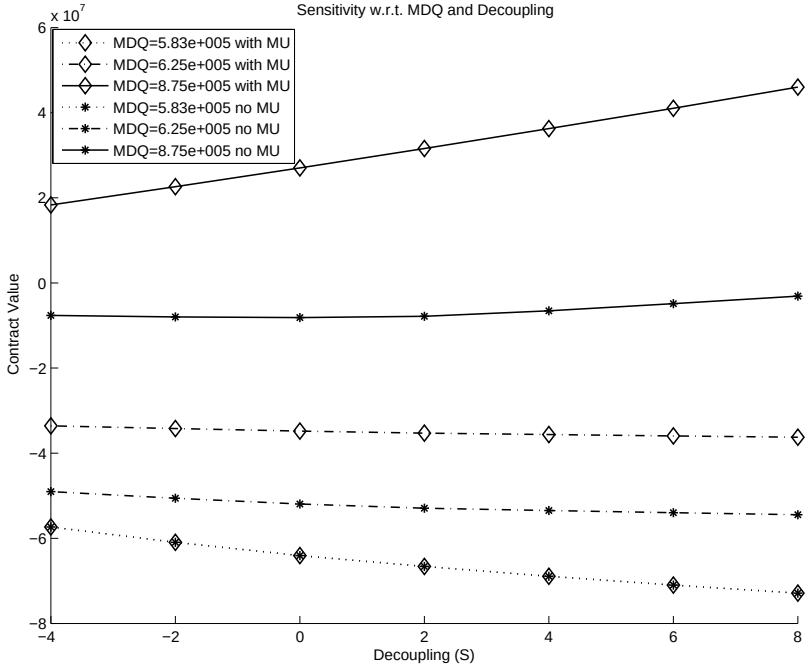
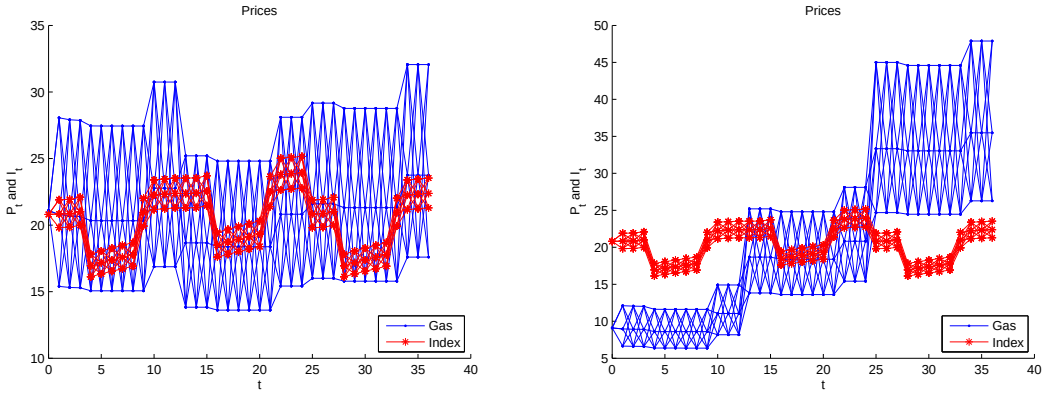


Figure 5: Sensitivity with respect to the decoupling S and three values of MDQ, from no make-up rights to very large flexibility. As expected, decoupling enforces make-up value.



(d) Low decoupling, swing option at the money in all the 3 years.

(e) High decoupling, swing option out of the money in the first year, at the money in the second year and in the money in the third year.

Figure 6: Scenarios for the term structure of gas and index prices for two levels of decoupling. In subfigure (a) make-up rights are typically not exercised, and prices are not decoupled, while in subfigure (b) typically make-up gas is declared in the first year and called back in the third year thanks to the decoupling.

Second Analysis. The second analysis is performed by mapping the swing value with respect to the decoupling parameter S and the interest rate r and reporting the corresponding prices in Figure 7.

The spirit of this analysis is that the make-up clause is exercised when a negative market scenario (typically, contractual price I higher than spot gas price P) is expected to change or disappear in the following years through a change in the slope of the index and the gas price forward term structure. On the other hand, as we saw in Subsection 2.3, the make-up gas nominated is paid partly immediately, and partly when the gas is withdrawn; this temporal mismatch implies a cash flow effects whose impact obviously depends also on the interest rate level: for higher interest rate levels, the benefit of the make-up clause is absorbed by the capitalization of the cost substained from the end of make-up nomination's year up to the withdrawal period. Conversely, in a standard contract without make-up clause, a higher interest rate in a market scenario with a low level of decoupling may lead to a higher contract value: in fact, if the decoupling is low, the present value of the contract in the long term, where the swing option is at or out of the money, is lower than the value in the short term, where the option is in the money. Figure 7 shows how any positive change in S is negatively compensated by an increase in the interest rates level.

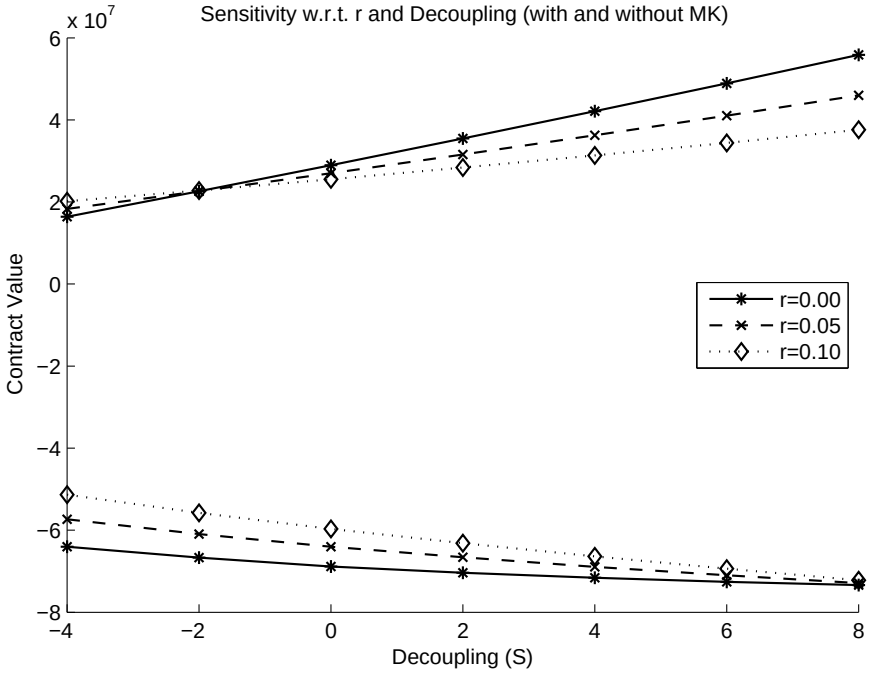
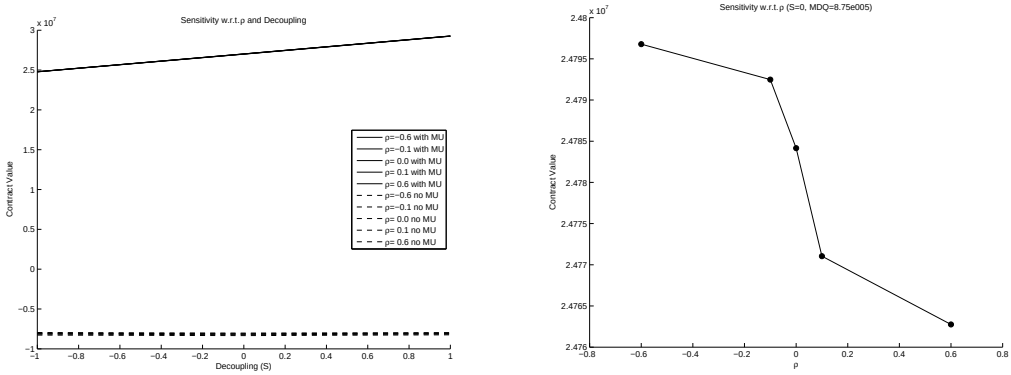


Figure 7: Sensitivities with respect to r and level of decoupling S in the forward prices of P. The first three cases on the top are with make-up, the last three in the bottom without.

Third Analysis. The third analysis maps the contract value with respect to the correlation ρ between the two prices P and I, and the level of decoupling S. In Figure 8(a) we see that decoupling knocks out correlation: in fact, the swing price's dependence on S is much greater than that on ρ, enforcing once again a strong dependence of the swing price on decoupling levels. Only a deeper analisys, performed for fixed values of decoupling, allows a better understanding of the impact of correlation: negative values of ρ leads to higher values of the contract. This is not a surprise: when ρ is negative the decoupling between prices is expected to be stronger (if P rises up then I falls down thanks to $\rho < 0$) and this increases the value of the contract. However, the changes due to correlation are still smaller than the changes due to decoupling, even for small values of S.



(f) Decoupling variable

(g) Fix value for decoupling

Figure 8: Contract value with respect to correlation ρ and level of decoupling. In Figure (a) the shift S vanquishes the effect of the correlation: in fact by varying ρ we obtain almost indistinguishable curves, both with or without make-up. In order to see the differences betwen curves, in Figure (b) the shift S is fixed and here we can see how correlation affecs contract value with make-up: negative values of ρ lead to higher contract values (negative ρ supports decoupling), but the stronger influence of the decoupling S is always evident.

References

[1] F. Aasche, P. Osmundsen, R. Tveterås, *European market integration for gas? Volume flexibility and political risk*. Energy Economics **24 (3)**, 249–265 (2002).

[2] Alba Soluzioni, `http://www.albasoluzioni.com` . Last accessed: 14/1/2011.

[3] O. Bardou, S. Bouthemy, G. Pagès, *Optimal quantization for the pricing of swing options*. Applied Mathematical Finance **16 (2)**, 183–217 (2009).

[4] O. Bardou, S. Bouthemy, G. Pagès, *When are swing options bangbang and how to use it?*. Pre-print LPMA-1141 (2007).

[5] C. Barrera-Esteve, F. Bergeret, C. Dossal, E. Gobet, A. Meziou, R. Munos, D. Reboul-Salze, *Numerical methods for the pricing of swing options: a stochastic control approach*. Methodology and Computing in Applied Probability **8(4)**, 517–540 (2006).

[6] F. E. Benth, J. Lempa, T. K. Nilssen, *On optimal exercise of swing options in electricity markets*. Working paper.

[7] D. P. Bertsekas, S. E. Shreve, "Stochastic Optimal Control: The Discrete Time Case". Mathematics in Science and Engineering, vol. 139. Academic: New York, [Harcourt Brace Jovanovich Publishers], 1978.

[8] D. Brigo, A. Dalessandro, M. Neugebauer, F. Triki, *A stochastic processes toolkit for risk management: Geometric Brownian motion, jumps, GARCH and variance gamma models*. Journal of Risk Management in Financial Institutions **2**, 365-393 (2009).

[9] D. Brigo, F. Mercurio, "Interest rate models - theory and practice". 2nd edition. Springer, 2006.

[10] L. Clewlow, C. Strickland, "Energy derivatives: pricing and risk management". Lacima Publications, 2000.

[11] E. Edoli, S. Fiorenzani, S. Ravelli, T. Vargiolu, *Modeling and valuing make-up clauses in gas swing contracts*. Energy Economics, In Press.

[12] European Commission's Eurostat, `http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database` . Last accessed: 6/9/2011.

[13] S. Fiorenzani, "Quantitative methods for electricity trading and risk management". Palgrave Macmillan, 2006.

[14] L. Holden, A. Løland, O. Lindqvist, *Valuation of Long-Term Flexible Gas Contracts*. The Journal of Derivatives, **18**, 75–85 (2011).

[15] J. C. Hull, "Options, futures and other derivatives". 6th edition. Prentice Hall, 2006.

[16] J. C. Hull, A. White, *Numerical procedures for implementing term structures I: single factor models*. Journal of Derivatives **2**, 7–16 (1994).

[17] J. C. Hull, A. White, *Numerical procedures for implementing term structures II: two factor models*. Journal of Derivatives **2**, 37–47 (1994).

[18] N. Cognome, N. Cognome, *Titoloarticolo*. Giornale (anno), pagine.

[19] IEA Natural Gas Information Statistics. OECD Publishing, 2010, `http://www.oecdbookshop.org/oecd/display.asp?sf1=identifiers&st1=SUB-64041S2` . Last accessed: 6/9/2011.

[20] P. Jaillet, E. I. Ronn, S. Tompaidis, *Valuation of commodity-based swing options*. Management Science **50**, 909–921 (2004).

[21] M. Kanai, *Decoupling the Oil and the Gas Prices*. IFRI papers (2011). `www.ifri.org/downloads/noteenergiemiharukanai.pdf` .

[22] A. Løland, O. Lindqvist, *Valuation of commodity-based swing options: a survey*. Norwegian Computing Center, Note SAMBA/38/80 (2008) `http://publications.nr.no/swingoptionsurvey.pdf` .

[23] M. Moreno, J. F. Navas, *On the Robustness of Least - Squares Monte Carlo (LSM) for Pricing American Derivatives*. Review of Derivatives Research, **6**, 2 (2003).

[24] A. Pokorná, "Pricing of Gas Swing Options". Diploma Thesis, Charles University, Prague, 2009. `http://ies.fsv.cuni.cz/default/file/download/id/11116` .

[25] E. Schwartz, J. E. Smith, *Short-Term Variations and Long-Term Dynamics in Commodity Prices*. Management Science **46 (7)**, 893–911 (2000).

[26] `http://www.theodora.com/pipelines`. Last accessed: 14/1/2011.

[27] *World Energy Outlook*. IEA 2009.

Singular limits of reaction-diffusion equations and propagation of interfaces

Cecilia De Zan (*)

Abstract. Interfacial phenomena are commonplace in physics, chemistry, biology. They occur, for example, whenever a continuum that can exists in at least two different chemical or physical "states" is present. The separation boundary is called an interface. In mathematics, interfaces appear in the study of the asymptotic limits of evolving systems, like reaction-diffusion equations. After a simple introduction about the connections between reaction-diffusion equations and the wavefronts they generates, we present some mathematical approaches to the study of evolving interfaces. We present the classical level set-approach and a geometrical approach introduced by Barles and Souganidis in 1998. Then we show how this second approach can be applied in the study of the asymptotic limits of reaction-diffusion equations. Finally we briefly show a generalization we obtained in the framework of Carnot Group for nonlinear and possibly degenerate diffusions.

Interfacial phenomena

Interfacial phenomena are commonplace in physics, chemistry, biology. They occur whenever a continuum is present that can exists in at least two different chemical or physical "states", and there is some mechanism that generates or enforces a spatial separation between these states.

The separation boundary is what we call an **interface**.

Some examples of physical processes where we can observe the generation of an interface are:

- the so-called **phase transition** that occurs whenever there is a double-well potential that drives a substance into one of two possible phases, such as solid or liquid;
- the **electrophoresis phenomenum** that is the motion of ions relatives to a fluid under the influence of an electric fields;
- in the **combustion phenomena** two different temperatures establish two different zones in the flame profile: the preheat zone, where the temperature is low enough so

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: cdezan@math.unipd.it. Seminar held on February 22nd, 2012.

that no chemical reaction has yet occurred and the burned zone where the gas has attained its final state.

From a mathematical point of view interfaces appear in the study of the asymptotic limits of evolving systems, like **reaction-diffusion equations**. To fix the ideas assume we have a smooth state variable u, a function of space and time depending also on a small parameter $\varepsilon > 0$,

$$u^\varepsilon(x,t) = u(x,t;\varepsilon), \quad x \in \mathbb{R}^n, t \in [0,+\infty).$$

Tipically u^ε is the solution of a semilinear reaction-diffusion equation

$$u^\varepsilon_t + \mathfrak{L}^\varepsilon(u^\varepsilon, x) = 0$$

satisfying an initial condition

$$u^\varepsilon(x,0) = g(x),$$

where g is a continuous bounded function. If such a solution u^ε exists for any $\varepsilon > 0$ one can try to look at the behavior of the family $(u^\varepsilon)_\varepsilon$ as $\varepsilon \to 0^+$.

The Allen Cahn equation

A famous example of semilinear reaction-diffusion equation is the so called **Allen Cahn equation**

$$u^\varepsilon_t(x,t) - \Delta u^\varepsilon(x,t) + \frac{f(u^\varepsilon(x,t))}{\varepsilon^2} = 0, \quad (x,t) \in \mathbb{R}^n \times (0,+\infty)\,,$$

where the reaction term f is the derivative of a double well potential W, $f = W'$.
In particular the function $f : \mathbb{R}^n \to \mathbb{R}$ has to satisfy the following assumptions,

$f \in C^2(\mathbb{R})$ has exactly three zeroes $m_- < m_0 < m_+$;
$f(q) < 0,\ \forall q \in (-\infty, m_-) \cup (m_0, m_+)$ and $f(q) > 0,\ \forall q \in (m_-, m_0) \cup (m_+, +\infty)$;
$f'(m_-) > 0, f'(m_+) > 0, f'(m_0) < 0$.

A tipical example for f is $f(q) = 2q(q^2 - 1)$.

Since the classical theory of elliptic equation assure us that, for any $\varepsilon > 0$ there exists a unique smooth solution of the Cauchy problem

$$(1) \qquad \begin{cases} u^\varepsilon_t(x,t) - \Delta u^\varepsilon(x,t) + \frac{f(u^\varepsilon(x,t))}{\varepsilon^2} = 0, \quad (x,t) \in \mathbb{R}^n \times (0,+\infty)\,, \\ u^\varepsilon(x,0) = g(x), \end{cases}$$

it is reasonable to study the asymptotic behavior of the u^ε's as $\varepsilon \to 0^+$. Numerical computations show that for any $t > 0$ there exists an interface Γ_t that separates two different regions $\mathcal{D}^-_t$ and $\mathcal{D}^+_t$ such that

$$\begin{aligned} \mathcal{D}^-_t &= \{x : \lim_{\varepsilon \to 0^+} u^\varepsilon(x,t) = m_-\} \\ \mathcal{D}^+_t &= \{x : \lim_{\varepsilon \to 0^+} u^\varepsilon(x,t) = m_+\}. \end{aligned}$$

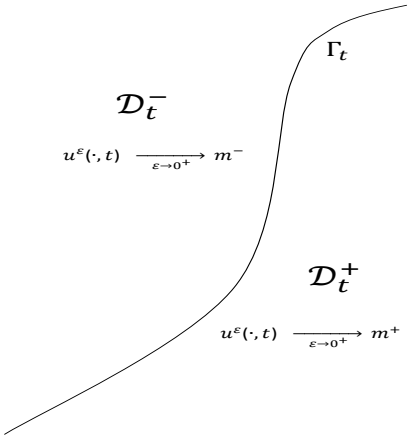


Figure 1: The limiting behavior of the u^ε's at time t.

Example For example, in the phase transition models u^ε is an order parameter such that a particular material is in the solid phase when u^ε is near m_- and it is in the liquid phase when it is near m_+. Therefore the interface Γ_t will be the boundary at time t of the solid region $\mathcal{D}_t^-$ from the liquid region $\mathcal{D}_t^+$.

In general, the asymptotic behavior of the solutions u^ε of the Cauchy problem for a reaction-diffusion equations

$$\begin{cases} u_t^\varepsilon + \mathfrak{L}(u^\varepsilon, x) = 0 \\ u^\varepsilon(x,0) = g(x) \end{cases} \tag{2}$$

is described by moving interfaces $t \longmapsto \Gamma_t$. Therefore to study the asymptotic behavior of the solutions of a reaction-diffusion equation one can study the evolution of the front they generates.

Propagation of fronts

Let Ω_t be an open subset of $\mathbb{R}^n$, we define the front Γ_t as

$$\Gamma_t = \partial\Omega_t.$$

We assume that for any $x \in \Gamma_t$ it is possible to consider the exterior normal vector at Ω_t in x, $\vec{n}(x)$, and that x evolves with normal velocity

$$v(D\vec{n}, \vec{n}, x, t).$$

An interesting **problem** to solve is the following. Let Ω_0 be an open subset of $\mathbb{R}^n$, study the evolution of the interfaces $t \longmapsto \Gamma_t$ moving with normal velocity

$$V = v(D\vec{n}(x), \vec{n}(x), x, t) \tag{3}$$

and starting at time $t = 0$ from $\Gamma_0 = \partial\Omega_0$. Two approaches are now possible,

- the so-called level set approach,
- the generalized flows approach.

The level set approach

The level set approach to front propagation can be summarize in the following three steps.

1. Let $\Gamma_0 = \partial\Omega_0$ be the front at time $t = 0$, we choose a real function $u_0 \in C(\mathbb{R}^n)$ such that

$$\Gamma_0 = \{x \in \mathbb{R}^n : u_0(x) = 0\}, \quad \Omega_0 = \{x \in \mathbb{R}^n : u_0(x) > 0\}$$
$$\text{(and therefore also } \overline{\Omega_0}^c = \{x \in \mathbb{R}^n : u_0(x) < 0\});$$

2. We call u the viscosity solution of the Cauchy problem

$$(\mathbf{Cp}) \qquad \begin{cases} u_t + F(D^2u, Du, x, t) = 0, & (x,t) \in \mathbb{R}^n \times (0,T), \\ u(x,0) = u_0(x), \end{cases}$$

where the function $F : \mathcal{S}^n \times (\mathbb{R}^n\backslash\{0\}) \times \mathbb{R}^n \times (0,+\infty) \to \mathbb{R}$ is related to the normal velocity v in the following way

$$F(X,p,x,t) = -|p|v(-\frac{1}{|p|}\Big(I - \frac{p\otimes p}{|p|^2}\Big)X, -\frac{p}{|p|}, x, t),$$

for any $X \in \mathcal{S}^n$, $p \in \mathbb{R}^n\backslash\{0\}$, $(x,t) \in \mathbb{R}^n \times (0,+\infty)$, $\mathcal{S}^n$ being the set of real symmetric $n \times n$ matrices.

3. Finally we define, for each $t > 0$, the triple $(\Gamma_t, \Omega_t^+, \Omega_t^-)$ as

$$\begin{aligned} \Omega_t^+ &= \{x \in \mathbb{R}^n : u(x,t) > 0\}, \\ \Omega_t^- &= \{x \in \mathbb{R}^n : u(x,t) < 0\}, \\ \Gamma_t &= \{x \in \mathbb{R}^n : u(x,t) = 0\}. \end{aligned}$$

The idea of the level set approach is to look at $(0,T) \ni t \longmapsto \big(\Gamma_t, \Omega_t^+, \Omega_t^-\big)$ to get the evolution of our interface Γ_0.

The main issues associated with this approach are the well-posedness of the Cauchy problem (**Cp**) and the well-posedness of the family $(\Gamma_t, \Omega_t^+, \Omega_t^-)_{t\in(0,T)}$. In other words one has to prove that there exists a unique viscosity solution of (**Cp**) an that the fronts Γ_t, $t > 0$, defined above depend only on Γ_0 and Ω_0 and not on the particular choice of u_0.

Among all the assumptions on F we want to point out the following basic assumptions:

(A1) F is a real-valued, locally bounded function on $\mathcal{S}^n \times \mathbb{R}^n \times \mathbb{R}^n \times (0,+\infty)$, continuous on $\mathcal{S}^n \times \mathbb{R}^n\backslash\{0\} \times \mathbb{R}^n \times (0,+\infty)$ and satisfies the *ellipticity condition*

$$F(X,p,x,t) \le F(Y,p,x,t) \qquad \text{whenever } X \ge Y,$$

for any $x \in \mathbb{R}^n$, $t \in (0,+\infty)$, $p \in \mathbb{R}^n\backslash\{0\}$ and $X, Y \in \mathcal{S}^n$,

(A2) for any $\lambda > 0$, $\mu \in \mathbb{R}$ and $(X,p,x,t) \in \mathcal{S}^n \times \mathbb{R}^n\backslash\{0\} \times \mathbb{R}^n \times (0,+\infty)$,

$$F(\lambda X + \mu p\otimes p, p, x, t) = \lambda F(X,p,x,t).$$

Generalized flows

In the study of the time evolution of the front Γ_0 by means of generalized flows one has to consider the evolution of the open subsets of $\mathbb{R}^n$, Ω_t and $\overline{\Omega}_t^c$, instead of the hypersurfaces Γ_t, $t > 0$. The main idea of this approach is to make a local monotonicity test against particular families of smooth open subsets of $\mathbb{R}^n$.

Let $(\Omega_t)_{t\in(0,T)}$ be a family of open subsets of $\mathbb{R}^n$. We call $(\Omega_t)_{t\in(0,T)}$ a *generalized flow*[(11)] with normal velocity $-F$ if and only if

- $(\Omega_t)_{t\in(0,T)}$ is a *generalized superflow* with normal velocity $-F$, i.e.
  - for any $B(x_0, r] \times [t, t + h]$
  - for any family of smooth open subsets $(A_t)_{t\in(0,T)}$ of $\mathbb{R}^n$, such that
    - $(A_t)_{t\in(0,T)}$ evolves with normal velocity smaller than $-F$,
    - $\overline{A}_t \cap B(x_0, r] \subset \Omega_t$,
    - $\overline{A}_s \cap \partial B(x_0, r] \subset \Omega_s$, for any $s \in [t, t + h]$,

  then
  $$A_s \cap B(x_0, r] \subset \Omega_s, \quad \text{for any } s \in [t, t + h).$$
- the family of closed subsets $(\overline{\Omega}_t)_{t\in(0,T)}$ is a *generalized superflow* with normal velocity $-F$, i.e.
  - for any $B(x_0, r] \times [t, t + h]$
  - for any family of smooth open subsets $(B_t)_{t\in(0,T)}$ of $\mathbb{R}^n$, such that
    - $(B_t)_{t\in(0,T)}$ evolves with normal velocity greater than $-F$,
    - $\overline{B}_t \cap B(x_0, r] \subset \overline{\Omega}_t^c$,
    - $\overline{B}_s \cap \partial B(x_0, r] \subset \overline{\Omega}_s^c$, for any $s \in [t, t + h]$,

  then
  $$B_s \cap B(x_0, r] \subset \overline{\Omega}_s^c, \quad \text{for any } s \in [t, t + h).$$

The level set approach and generalized flows

The level set approach and the approximation of an evolving interface Γ_t by means of generalized flows turn out to be equivalent, as the next theorem states.

Theorem 1 *Assume that* **(A1)** *and* **(A2)** *hold. We consider*

- *two families of open subsets of* $\mathbb{R}^n$, $(\Omega_t^1)_{t\in[0,T)}$ *and* $(\Omega_t^2)_{t\in[0,T)}$, *such that*
  - $(\Omega_t^1)_{t\in[0,T)}$ *is a* ***generalized superflow*** *with normal velocity* $-F$,
  - $((\Omega_t^2)^c)_{t\in[0,T)}$ *is a* ***generalized subflow*** *with normal velocity* $-F$,
  - $\cup_{t\in(0,T)}\Omega_t^1 \times \{t\}$, $\cup_{t\in(0,T)}\Omega_t^2 \times \{t\}$ *are open and disjoint,*

(11) For the precise definition we refer to [1].

- *an open set* Ω_0 *such that* $\Omega_0 \subseteq \Omega_0^1$ *and* $\overline{\Omega}_0^c \subseteq \Omega_0^2$.
- *the* ***level set evolution*** *of* $(\partial\Omega_0, \Omega_0, \overline{\Omega}_0^c)$, $(\Gamma_t, \Omega_t^+, \Omega_t^-)_{t\in(0,T)}$.

Then, for all $t \in [0,T)$,

$$\Omega_t^+ \subset \Omega_t^1 \subset \Omega_t^+ \cup \Gamma_t, \qquad \Omega_t^- \subset \Omega_t^2 \subset \Omega_t^- \cup \Gamma_t.$$

Moreover, if $\operatorname{Int}\Gamma_t = \emptyset$ *for any* $t \in [0,T)$, *then*

$$\Omega_t^+ = \Omega_t^1, \qquad \Omega_t^- = \Omega_t^2, \qquad \text{for any } t \in [0,T).$$

Application to reaction-diffusion equations

Generalized flows are very useful in the study of the asymptotics of solutions of reaction-diffusion equations. To this aim Theorem 1 is a crucial point. Barles and Souganidis in [3] and then Barles and Da Lio in [1] apply this theory to the study of the limiting behavior, as $\varepsilon \to 0^+$, of the solution of the Cauchy problem for the Allen Cahn equation (1). They prove the following theorem.

Theorem 2 *Let* u^ε *be the solution of the Cauchy problem*

$$\begin{cases} u_t^\varepsilon(x,t) - \Delta u^\varepsilon(x,t) + \frac{f(u^\varepsilon(x,t))}{\varepsilon^2} = 0, & (x,t) \in \mathbb{R}^n \times (0,+\infty), \\ u^\varepsilon(x,0) = g(x). \end{cases}$$

We define the interface at time $t = 0$ *in the following way*

$$\Gamma_0 = \{x : g(x) = m_0\}.$$

If d_0 *is the signed distance from* Γ_0 *so that*

$$d_0 = \begin{cases} > 0 & \text{in} \quad \{x : g(x) > m_0\}, \\ < 0 & \text{in} \quad \{x : g(x) < m_0\}, \end{cases}$$

and u *is the unique viscosity solution of*

$$\begin{cases} u_t - \operatorname{tr}\left[(I - \frac{Du}{|Du|} \otimes \frac{Du}{|Du|})D^2u\right] = 0, & (x,t) \in \mathbb{R}^n \times (0,T), \\ u(x,0) = d_0(x), & x \in \mathbb{R}^n, \end{cases}$$

then

$$u^\varepsilon(x,t) \underset{\varepsilon\to 0^+}{\longrightarrow} \begin{cases} m_+ & \\ & \text{locally uniformly in} \\ m_- & \end{cases} \begin{matrix} \{u > 0\} \\ \\ \{u < 0\}. \end{matrix}$$

Moreover, if

$$\operatorname{Int}\{x : u(x,t) = 0\} = \emptyset \quad \text{for any } t \in [0,T),$$

then

$$u^\varepsilon(x,t) \underset{\varepsilon\to 0^+}{\longrightarrow} \begin{cases} m_+ & \text{locally uniformly in } \{u>0\} \\ m_- & \text{locally uniformly in } \overline{\{u>0\}}^c. \end{cases}$$

In the proof of this Theorem becomes clear the importance of generalized sub- and superflow. Indeed to prove Theorem 2 one has to follow these three steps.

(a) Consider, for any $t \in (0,T)$

$$\begin{aligned} \Gamma_t &= \{u(\cdot,t)=0\}, \\ \Omega_t^+ &= \{u(\cdot,t)>0\}, \\ \Omega_t^- &= \{u(\cdot,t)<0\}, \end{aligned}$$

$(\Gamma_t, \Omega_t^+, \Omega_t^-)_{t\in(0,T)}$ is the **level set evolution** of $(\Gamma_0, \Omega_0^+, \Omega_0^-)$.

(b) Construct $(\Omega_t^1)_{t\in(0,T)}$, $(\Omega_t^2)_{t\in(0,T)}$ such that

- $(\Omega_t^1)_t$ is a **generalized superflow** with normal velocity $-F$,
- $((\Omega_t^2)^c)_t$ is a **generalized subflow** with normal velocity $-F$,
- $\begin{cases} u^\varepsilon(x,t) \to m_+ & \text{if } (x,t) \in \Omega^1 := \bigcup_{t\in(0,T)} \Omega_t^1 \times \{t\}, \\ u^\varepsilon(x,t) \to m_- & \text{if } (x,t) \in \Omega^2 := \bigcup_{t\in(0,T)} \Omega_t^2 \times \{t\}. \end{cases}$

(c) Determine the traces Ω_0^1 and Ω_0^2 of Ω^1 and Ω^2 for $t=0$ such that $\Omega_0^+ \subseteq \Omega_0^1$ and $\Omega_0^- \subseteq \Omega_0^2$.

(d) Use Theorem 1 to get the conclusion.

In [3] [1] the authors consider several types of reaction-diffusion equations. For example they study the limiting behavior, as $\varepsilon \to 0^+$, of the solution of the following nonlinear equation with $x-$dependent diffusion terms

$$\begin{cases} u_t^\varepsilon(x,t) - \operatorname{tr}\left(A(x)D^2u^\varepsilon(x,t)\right) + \frac{f(u^\varepsilon(x,t))}{\varepsilon^2} = 0, \quad t>0, \\ u^\varepsilon(x,0) = g(x), \quad x \in \mathbb{R}^n, \end{cases} \tag{4}$$

where f is a cubic function and $A = (a_{ij})_{ij} \in C^2(\mathbb{R}^n, \mathcal{S}^n) \cap W^{2,\infty}(\mathbb{R}^n, \mathcal{S}^n)$ is a matrix map such that the matrix $A(x)$ is positive definite for any $x \in \mathbb{R}^n$, i.e. there exists a constant $\nu > 0$ so that,

$$A(x)q \cdot q \geq \nu |q|^2, \qquad \text{for any } x \in \mathbb{R}^n.$$

From this last condition it follows the existence of a matrix map $\sigma = (\sigma_{ij})_{ij} \in C^2(\mathbb{R}^n, \mathcal{S}^n) \cap W^{2,\infty}(\mathbb{R}^n, \mathcal{S}^n)$ such that

$$\begin{gathered} A \equiv \sigma\sigma^t \\ \operatorname{rank}(\sigma(x)) = n, \text{ for any } x \in \mathbb{R}^n. \end{gathered}$$

We denote with $\sigma^{(1)}, \dots, \sigma^{(n)}$ the columns of the matrix σ and we consider, for any regular function h, the derivatives of h in the directions of the vector fields $\sigma^{(1)}, \dots, \sigma^{(n)}$, i.e. we define the *$\sigma$-gradient* of h,

$$D_\sigma h(x) := \sigma^t(x) Dh(x) = \begin{pmatrix} \sigma^{(1)} \cdot Dh(x) \\ \vdots \\ \sigma^{(n)} \cdot Dh(x) \end{pmatrix}$$

and the *σ-Hessian matrix* of h,

$$D^2_\sigma h(x) := \sigma^t(x) D^2 h(x) \sigma(x).$$

With these notations the reaction-diffusion equation (4) can be rewritten in the following way

$$\begin{cases} u^\varepsilon_t(x,t) - \operatorname{tr}\left(D^2_\sigma u^\varepsilon(x,t)\right) + \frac{f(u^\varepsilon(x,t))}{\varepsilon^2} = 0, \quad t > 0, \\ u^\varepsilon(x,0) = g(x), \quad x \in \mathbb{R}^n\,. \end{cases}$$

It is now clear the analogy with the Allen Cahn equation in (1). In fact in the Allen Cahn equation the diffusion term is the Laplacian of u^ε, that is the trace of the Hessian matrix of u^ε. Here as diffusion term we have the trace of the *σ-Hessian matrix* of u^ε, $D^2_\sigma u^\varepsilon$. The equation that describes the evolution of the interface Γ_t is therefore the usual mean curvature equation with the standard euclidean gradient and Hessian matrix Du and D^2u replaced by the *σ-gradient* and the *σ-Hessian matrix* $D_\sigma u$ and $D^2_\sigma u$, i.e.

$$\text{(5)} \qquad \begin{cases} u_t - \operatorname{tr}\left[(I - \frac{D_\sigma u}{|D_\sigma u|} \otimes \frac{D_\sigma u}{|D_\sigma u|}) D^2_\sigma u\right] = 0, & (x,t) \in \mathbb{R}^n \times (0,T) \\ u(x,0) = d_0(x), & x \in \mathbb{R}^n. \end{cases}$$

In our phd thesis we consider several reaction-diffusion equations. For example we look for the right reaction-diffusion equation that gives rise to a front evolving according to the Cauchy problem (5) when the matrix map $\sigma(\cdot)$ takes values in the space of the $m \times n$ matrices with $m < n$. Our first idea was to study again (4). The problem is that now A can be degenerate, i.e the constant ν can be zero

$$\nu \geq 0\,,$$

and therefore it is possible that a solution of (4) does not exist. This induces us to introduce a Riemannian approximation of A defined in the following way. For any $\varepsilon > 0$ we define

$$\sigma_\varepsilon(\cdot) = [\sigma(\cdot) | \varepsilon^k I_n] \in \mathbb{R}^{n \times (m+n)},$$

where $k > 0$ is a suitable fixed constant and I_n is the $n \times n$ identity matrix. Thus, if we put

$$A_\varepsilon(\cdot) = \sigma_\varepsilon(\cdot)\sigma^t_\varepsilon(\cdot) = A(\cdot) + \varepsilon^{2k} I_n,$$

we have that A_ε satisfies

$$A_\varepsilon(x)q \cdot q = A(x)q \cdot q + \varepsilon^{2k}|q|^2 \geq \underbrace{(\nu + \varepsilon^{2k})}_{>0}|q|^2.$$

Thus, by the classical theory of elliptic equation, for any $\varepsilon > 0$ there exists u^ε, smooth solution of

$$\begin{cases} u_t^\varepsilon(x,t) - \operatorname{tr}\left(A_\varepsilon(x) D^2 u^\varepsilon(x,t)\right) + \frac{f(u^\varepsilon(x,t))}{\varepsilon^2} = 0, \quad t > 0\,, \\ u^\varepsilon(x,0) = g(x), \quad x \in \mathbb{R}^n. \end{cases} \tag{6}$$

Following the notation above, if we define for any $h \in C^2(\mathbb{R}^n, \mathbb{R})$, the *$\varepsilon$-gradient* and the *$\varepsilon$-Hessian matrix*, $D_\varepsilon h$ and $D_\varepsilon^2 h$ as

$$\begin{aligned} D_\varepsilon h(x) &= \sigma_\varepsilon^t(x) Dh(x) = \begin{pmatrix} D_\sigma h(x) \\ \varepsilon^k Dh(x) \end{pmatrix}, \\ D_\varepsilon^2 h(x) &= \sigma_\varepsilon^t(x) D^2 h(x) \sigma_\varepsilon(x), \end{aligned}$$

we obtain that the equation (6) can be rewritten as

$$\begin{cases} u_t^\varepsilon(x,t) - \operatorname{tr}\left(D_\varepsilon^2 u^\varepsilon(x,t)\right) + \frac{f(u^\varepsilon(x,t))}{\varepsilon^2} = 0, \quad t > 0 \\ u^\varepsilon(x,0) = g(x), \quad x \in \mathbb{R}^n. \end{cases}$$

In our phd thesis we use the ideas of Barles, Souganidis and Da Lio to prove that, when the vector fields $\sigma^{(1)} \cdot \nabla, \ldots, \sigma^{(m)} \cdot \nabla$ generates a Carnot group of step two[(12)], then the front that describes the asymptotic behavior of the solution of (6) is governed by the geometric pde (5) where, we recall, σ is a matrix map that takes value into the space of the $m \times n$ matrices with $m < n$.

References

[1] Barles, G.; Da Lio, F., *A geometrical approach to front propagation problems in bounded domains with Neumann-type boundary conditions.* Interfaces Free Bound. **5** (2003), 239–274.

[2] Bonfiglioli, A., Lanconelli, U., Uguzzoni, F., "Stratified Lie groups and potential theory for their sub-laplacians". Springer, 2007.

[3] Barles, G.; Souganidis, P.E., *A new approach to front propagation problems: theory and applications.* Arch. Ration. Mech. Anal. **141** (1998), 237–296.

[4] Chen, Y.-G., Giga, Y., Goto, S., *Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations.* J. Differential Geom. **33** (1991), 749–786.

[(12)] for the definition of Carnot group see the book of Bonfiglioli, Lanconelli and Uguzzoni [2].

Smooth Asymptotics for a DIC option in a Binomial Tree Model

Jose Maria L. Escaner IV (*)

Abstract. The talk aims to compute for the smooth asymptotic expansion for a down-and-in call (DIC) barrier option that was modeled using the Cox-Ross-Rubenstein (CRR) binomial tree. For pricing option contracts, the most well-known model used both by practitioners and in the academe is the Black- Scholes continuous time model. Though less accurate, a simpler and easier understood way to model financial derivatives though would be to use discrete time models. Among the many different discrete time models, a simple and widely-used model is the CRR binomial tree model. It is well-known that the price of the Black-Scholes continuous time model is close to the price obtained with the CRR binomial tree model when the number of time steps is large, as the Black-Scholes price is the limit of the tree model price. As such, it is of interest to measure the convergence of the CRR model using asymptotic expansion. We follow the framework used by Diener and Diener in measuring the asymptotic expansion for the convergence of barrier options. For the purpose of finding the asymptotics, we make use of Andres symmetry principle in order to find the exact pricing formula of the DIC barrier option. By the guidelines set by Joshi, we specify the parameters and define our CRR binomial tree in such a way as to make the pricing formula symmetric. This would allow us to formulate a complete and smooth asymptotic expansion for our DIC barrier option.

1 Introduction

For the last three decades, the world of finance and banking has undergone radical changes. The shift started with the creation and introduction of derivatives. The most famous and probably most used derivative instrument is the *option*. An vanilla option is an agreement in which the buyer (holder) has the right, but not the obligation, to exercise by buying (call) or selling (put) an underlying asset at a set price (strike price) on (european style option) or before (american style option) a future date (the exercise date or expiration).

Barrier options, on the other hand, are exotic options where the payoff depends on whether the price of the underlying asset reaches (or does not reach) a certain level (the barrier) before exercise date. Barrier options are attractive to many investors because they are generally less expensive than vanilla options.

(*)University of the Philippines Diliman; E-mail: `jlescaner@up.edu.ph`. Seminar held on March 7th, 2012.

In pricing option contracts, the Black-Scholes formula for a vanilla option with no dividends has been the benchmark for all other types of options. This is given by

$$C(S,t) = N(d_1)S - N(d_2)Ke^{-r(T-t)}$$

where S is the price of the stock at time t, K is the strike price, T is the exercise time, r is the risk-free interest rate and

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}(T-t)\right)}{\sigma\sqrt{T-t}} \quad \text{and} \quad d_2 = d_1 - \sigma\sqrt{T-t}.$$

Though less accurate, a simpler formula uses discrete-time tree models such as the Cox-Ross-Rubenstein (CRR) binomial tree model given by

$$\begin{aligned} C_0 = \quad & e^{-rT}E((S_T-K)^+) = e^{-rT}\sum_{j=0}^{n}\binom{n}{j}p^j(1-p)^{n-j}\max(S_T-K,0) \\ (1) \quad = \quad & S_0\sum_{j=k}^{n}\binom{n}{j}q^j(1-q)^{n-j} - Ke^{-rT}\sum_{j=k}^{n}\binom{n}{j}p^j(1-p)^{n-j}, \end{aligned}$$

where $p = \dfrac{e^{r\frac{T}{n}} - d}{u-d}$ and $q = pue^{-r\frac{T}{n}} = \frac{u-e^{-r\frac{T}{n}}}{u-d}$.

It is known that the price obtained under the CRR model converges to that under the Black-Scholes continuous-time model as the number of time steps in the tree model becomes infinitely large. It is of interest then to measure the rate of convergence of tree models (in particular, the CRR model) as this helps in the numerics. Here, we use asymptotic expansion as our main tool in measuring the rate.

Researches pertaining to asymptotic expansions of tree models have been done previously. Diener and Diener [1, 2] have already started work on computing this asymptotic expansion. In particular, a general method was conceived for finding the asymptotic expansion of the vanilla european option and of barrier options in general. Joshi [3] have expanded on the works of Diener and Diener in finding the asymptotics of vanilla european options and have introduced a more refined and smooth asymptotic expansion.

2 Asymptotic Expansions for u, d, p and q

Definition 1 The formal series

$$\sum_{n=0}^{\infty} a_n g_n(x)$$

not necessarily convergent, is said to be an asymptotic expansion of $f(x)$ in the Poincare sense, with respect to the asymptotic sequence $\{g_n(x)\}$, if and only if, for every value of m,

$$f(x) - \sum_{n=0}^{m} a_n g_n(x) = o(g_m(x)) \text{ as } x \to x_0.$$

The partial sum $\sum_{n=0}^{m-1} a_n g_n(x)$ is an approximation of $f(x)$ with an error term of $O(g_m(x))$ as $x \to x_0$. If such an asymptotic expansion exists, it is unique, and we say that $f(x) \sim \sum_{n=0}^{\infty} a_n g_n(x)$.

In computing for the asymptotics of the movements u and d of the price of the stocks, we need to assume the following condition:

$$d < e^{r\frac{T}{n}} < u, \tag{2}$$

where d and u are the down and up factors, respectively, and $\frac{T}{n}$ represents one time step in the binomial tree. The assumption is needed in order to ensure *absence of arbitrage*, as pointed out by Cox, Ross, and Rubenstein in their original paper (see [4]).

Letting the votality of of stock prices be equal to σ, we can compute for the asymptotics of u and d using a simple Taylor formula expansion:

$$u = e^{\sigma\sqrt{\delta t}} = 1 + \sigma\sqrt{\frac{T}{n}} + \frac{\sigma^2}{2}\frac{T}{n} + \frac{\sigma^3}{6}\left(\frac{T}{n}\right)^{\frac{3}{2}} + \frac{\sigma^4}{24}\left(\frac{T}{n}\right)^2 + O\left(\left(\frac{1}{n}\right)^{\frac{5}{2}}\right), \tag{3}$$

$$d = e^{-\sigma\sqrt{\delta t}} = 1 - \sigma\sqrt{\frac{T}{n}} + \frac{\sigma^2}{2}\frac{T}{n} - \frac{\sigma^3}{6}\left(\frac{T}{n}\right)^{\frac{3}{2}} + \frac{\sigma^4}{24}\left(\frac{T}{n}\right)^2 + O\left(\left(\frac{1}{n}\right)^{\frac{5}{2}}\right). \tag{4}$$

Next, we compute for the asymptotics for risk-neutral probabilities p and q. Algebraic manupulations give us the following expansions:

$$p = \frac{1}{2} + \frac{\left(r - \frac{\sigma^2}{2}\right)}{2\sigma}\sqrt{\frac{T}{n}} + O\left(\frac{1}{n}\right)^{\frac{3}{2}}, \tag{5}$$

$$q = \frac{1}{2} + \frac{\left(r + \frac{\sigma^2}{2}\right)}{2\sigma}\sqrt{\frac{T}{n}} + O\left(\frac{1}{n}\right)^{\frac{3}{2}}. \tag{6}$$

Now that we have computed for the asymptotics of u, d, p, and q, we can proceed to the main task of computing for a smooth asymptotic expansion of the pricing formula for the down and in call (DIC)barrier option. A DIC barrier option is a contract which gives the right to buy the underlying asset at a certain price at a predefned date if and only if the asset price has dropped down at least to the pre-setbarrier level before the expiry date of the contract.

3 Incomplete Binomial Form of the DIC option

The first step towards finding a smooth asymptotic expansion for the price of barrier options involves the exact pricing formula of the DIC barrier option. Since the DIC

barrier option needs the whole trajectory of the stock S_t between $t = 0$ and $t = T$ in order to determine if the path crossed the barrier L, we introduce the stopping time $\tau_L(\omega)$ where

$$\tau_L(\omega) := \begin{cases} \min\{t \in T, S_t(\omega) \leq L\} & \text{if } \{t \in T, S_t(\omega) \leq L\} \neq \phi \\ T + 1(\text{or } +\infty) & \text{else .} \end{cases}$$

so that

$$DIC_T = (S_T - K)^+ \mathbf{I}_{\tau_L \leq T}.$$

To be able to get an incomplete binomial sum form of the DIC_T, a technique by Diener and Diener [1] is used. In their paper, they used the Andre's Symmetry/Reflection Principle. This allows us to determine the number of trajectories that have crossed the barrier level in our binomial tree after having having crossed that barrier level.

We define the barrier line L as the line located either on one of the horizontal lines or between two succeeding horizontal lines of the nodes in the binomial tree. The *line of L-points*, on the other hand is either the horizontal line in line with the barrier line or the one directly below the barrier line. Since the line of L-points represents the largest horizontal line of nodes that is equal or below L, any trajectory that would cross L would cross the line of L-points as well.

The y-intercept of the line of L-points is $S_0 u^l d^{n-l} = S_0 u^{2l-n}$ if the end point of the L-points has abscissa $n\delta t$; otherwise it is $S_0 u^l d^{(n-1)-l} = S_0 u^{2l-n+1}$ if end point has abscissa $(n-1)\delta t$. Here, l is the largest integer such that $S_0 u^{2l-n} \leq L$ or $S_0 u^{2l-n+1} \leq L$. That is,

$$l = \begin{cases} \left[\frac{n}{2} + \frac{\ln \frac{L}{S_0}}{2 \ln u}\right] = \frac{n}{2} + \frac{\ln \frac{L}{S_0}}{2 \ln u} - \{l\} & \text{if the abscissa is } n\delta t \\ \left[\frac{n-1}{2} + \frac{\ln \frac{L}{S_0}}{2 \ln u}\right] = \frac{n-1}{2} + \frac{\ln \frac{L}{S_0}}{2 \ln u} - \{l\} & \text{if the abscissa is } (n-1)\delta t. \end{cases} \tag{7}$$

where $\{l\}$ is the fractional component of l. Furthermore, if we define J such that

$$l = \begin{cases} J/2 & \text{if end point has abscissa } n\delta t \\ (J-1)/2 & \text{if end point has abscissa } (n-1)\delta t. \end{cases}$$

then we get

$$J := \left[n + \frac{\ln \frac{L}{S_0}}{\ln u}\right] = n + \frac{\ln \frac{L}{S_0}}{\ln u} - \{J\}.$$

We do a similar computation for the line of K points:

$$k = \begin{cases} \frac{n}{2} + \frac{\ln \frac{K}{S_0}}{2 \ln u} + 1 - \{k\} & \text{if the end point of the line of K-points has abscissa}(n)\delta t \\ \frac{n-1}{2} + \frac{\ln \frac{K}{S_0}}{2 \ln u} + 1 - \{k\} & \text{if the end point of the line of K-points has abscissa}(n-1)\delta t. \end{cases}$$

Lemma 1 [Diener and Diener] *Let n and j be positive integers, with $0 \leq J \leq n$. The number $A(j, J)$ of trajectories of the binary random walk $(S_t)_{t \in \mathbf{T}}$ such that $(S_T =) S_{n\delta t} =$*

S_0u^{2j-n} (i.e. such that at time $t = T$ they reach the level j) after having reached the node line with y-intercept S_0u^{J-n} at some time between $t = 0$ and $t = n\delta t$ is equal to:

$$A(j,J) = \begin{cases} 0 & \text{if } J < j \leq n \\ \binom{n}{J-j} & \text{if } J/2 \leq j \leq J \\ \binom{n}{j} & \text{if } 0 \leq j \leq J/2. \end{cases}$$

Theorem 1 [Diener and Diener] *The initial price of a down-and-in-call (DIC) option with exercise price K, expiry time T, and barrier $L < S_0$, written on a asset modeled by a random walk (S_t) is given by:*

For $L \leq K$:

$$DIC_0 = e^{-rT} \sum_{j=k}^{J} \binom{n}{J-j} p^j (1-p)^{n-j} (S_0 u^{2j-n} - K),$$

For $K \leq L$:

$$\begin{aligned} DIC_0 &= e^{-rT} \sum_{j=k}^{l} \binom{n}{j} p^j (1-p)^{n-j} (S_0 u^{2j-n} - K) \\ &\quad + e^{-rT} \sum_{j=l+1}^{J} \binom{n}{J-j} p^j (1-p)^{n-j} (S_0 u^{2j-n} - K). \end{aligned}$$

If we denote Φ to be the incomplete binomial sum

$$\Phi(n,k,p) := \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j}.$$

then for $L \leq K$,

$$DIC_0 = S_0 \left(\frac{q}{1-q}\right)^{J-n} \Phi(n, k+n-J, q) - Ke^{-rT} \left(\frac{p}{1-p}\right)^{J-n} \Phi(n, k+n-J, p),$$

and for $K \leq L$,

$$\begin{aligned} DIC_0 &= S_0 \left(\frac{q}{1-q}\right)^{J-n} \Phi(n, 1+n+l-J, q) \\ &\quad - Ke^{-rT} \left(\frac{p}{1-p}\right)^{J-n} \Phi(n, 1+n+l-J, p) + S_0 \Phi(n,k,q) \\ &\quad - Ke^{-rT} \Phi(n,k,p) - S_0 \Phi(n, l+1, q) + Ke^{-rT} \Phi(n, l+1, p). \end{aligned}$$

4 Transformation into Integrals

We now transform the incomplete binomial sums into integrals. The integral form allows us to write the formula into asymptotic expansions. We make use of the following general lemma, which can easily be shown by integration of parts of the right term by $(n-k)$ times.

Lemma 2 *For any $n \in \mathbf{N}$ and any $k, 0 < k \leq n$, one has:*

$$\sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} = k \binom{n}{k} \int_0^p y^{k-1} (1-y)^{n-k} \, dy.$$

Using this formula, the price DIC_0 can be written, for $L \leq K$, in the following integral form:

$$\begin{aligned} DIC_0 = \quad & (k+n-J) \binom{n}{k+n-J} \left(S_0 \left(\frac{q}{1-q} \right)^{J-n} \int_0^q y^{k+n-J-1} (1-y)^{J-k} \, dy \right. \\ & \left. -Ke^{-rT} \left(\frac{p}{1-p} \right)^{J-n} \int_0^p y^{k+n-J-1} (1-y)^{J-k} \, dy \right). \end{aligned} \tag{8}$$

In the case $K \leq L$, the expression of DIC_0 in terms of integrals can be obtained in a similar way, but this introduces four more integrals.

$$\begin{aligned} DIC_0 = {} & (1+n+l-J) \binom{n}{1+n+l-J} \left(S_0 \left(\frac{q}{1-q} \right)^{J-n} \int_0^q y^{n+l-J} (1-y)^{J-l-1} \, dy \right. \\ & \left. -Ke^{-rT} \left(\frac{p}{1-p} \right)^{J-n} \int_0^p y^{n+l-J} (1-y)^{J-l-1} \, dy \right) \\ & + (k) \binom{n}{k} \left(S_0 \left(\frac{q}{1-q} \right)^{J-n} \int_0^q y^{k-1} (1-y)^{n-k} \, dy \right. \\ & \left. -Ke^{-rT} \left(\frac{p}{1-p} \right)^{J-n} \int_0^p y^{k-1} (1-y)^{n-k} \, dy \right) \\ & - (l+1) \binom{n}{l+1} \left(S_0 \left(\frac{q}{1-q} \right)^{J-n} \int_0^q y^{l} (1-y)^{n-l-1} \, dy \right. \\ & \left. -Ke^{-rT} \left(\frac{p}{1-p} \right)^{J-n} \int_0^p y^{l} (1-y)^{n-l-1} \, dy \right). \end{aligned} \tag{9}$$

5 Choosing the Parameters

Diener and Diener (see [2]) have done extensive work in finding the asymptotic expansion for barrier options in a CRR tree. In their paper, Diener and Diener discovered that

the asymptotics for the pricing of barrier options in general does not allow for a normal smooth (or infinitely differentiable) expansion because of oscillations in its convergence.

The fractional components of J, l, and k in Section 3 cannot be expanded into asymptotics in the usual smooth sense. This is because these fractional components do not have limits as the number of time steps n goes to infinity. This leads to the inability of the variables containing these fractional components to be written as an analytic function. Therefore, the variables (and in effect, the pricing formula for the DIC barrier option) do not have an asymptotic expansion in the usual sense. Instead the convergence of the expansion of the variables (and the pricing formula) oscillates in value as n becomes larger.

It is for this reason that Diener and Diener have to introduce the concept of asymptotic expansions with bounded coefficients in order to accomodate the fractional components of these variables which would allow the variables to be transformed into a normal asymptotic expansion. We aim to improve on the results of Diener and Diener by finding a smooth asymptotic expansion for our DIC barrier options without resorting to their technique. Instead, we redefine the parameters J, l, and k and specify our binomial tree in such a way as to remove these fractional components of these variables. This would allow for a smooth asymptotic expansion.

The key step towards finding specific definitions for the paramters of our binomial tree involves transforming the integrals of the exact pricing formula for barrier options (found in equations (8) and (9) into one that will be symmetric around the half-way point. Doing this ensures that a smooth asymptotic expansion would occur using the parameters J, l, and k that were found. This is the general method used by Joshi that enabled him to find a smooth asymptotic expansion for vanilla european options (see [3]).

Specifically, in order to have a clean expansion, we want to equate the exponentials of y and $1-y$ inside the integral of equations (8) and (9). Equating all the exponentials, we would have:

For equation (8):

$$k+n-J-1=J-k \Longrightarrow n=2(J-k)+1. \tag{10}$$

For equation (9):

$$n+l-J=J-l-1 \Longrightarrow \qquad n=2(J-l)-1, \tag{11}$$

$$k-1=n-k \Longrightarrow \qquad n=2k-1, \tag{12}$$

$$l=n-l-1 \Longrightarrow \qquad n=2l+1. \tag{13}$$

Depending on the abscissa of the end point of the line of L-points, either $J=2l$ or $J=2l+1$. If $J=2l$, then this assumption contradicts equation (13). As it turns out, choosing $J=2l+1$ makes equation (11) equivalent to equation (13). With this as our only choice for J, letting l be equal to some integer N, we have the following redefinitions:

$$l= \qquad N \tag{14}$$

$$k= \qquad N+1 \tag{15}$$

$$n= \qquad J=2N+1. \tag{16}$$

By these redefinitions of n, J, k, and l, all the exponents of y and $1-y$ in equations (8) and (9) will become equivalent to N. Furthermore, solving for the position of the barrier level L,

$$J := n + \frac{\ln \frac{L}{S_0}}{\ln u} \Longrightarrow L = S_0.$$

We see here that the barrier level L is in line with the origin point and is centered in the middle of the tree.

Doing a similar process for finding the position of K, it is not difficult to show that $K = S_0$.

6 Results

We state at this point our results.

Proposition 1 *By letting the variables l, k, and J, and the number of time steps n be equivalent to the following:*

$$l = N, \qquad k = N+1, \qquad J = n = 2N+1,$$

where N is a natural number, we have the following results regarding our down and in call (DIC) barrier option as modeled in a Cox, Ross, and Rubenstein (CRR) binomial tree:

(a) *The barrier level L will be located in the same position as the strike price K in the binomial tree. Both will be centered in the middle of the tree, in line with the origin point S_0. In other words, $L = K = S_0$.*

(b) *The pricing formula of the DIC barrier option will be equivalent to the pricing formula of a vanilla european call option. The integral version of the pricing formula is given as follows:*

$$DIC_0 = (N+1)\binom{2N+1}{N+1}\left(S_0 \int_0^q y^N (1-y)^N \, dy - Ke^{-rT} \int_0^p y^N (1-y)^N \, dy\right).$$

(c) *A smooth asymptotic expansion for a DIC barrier option when using the method of symmetry of Joshi is given as follows:*

$$DIC_0 = \frac{S_0 - Ke^{-rT}}{2} + \frac{2(S_0 - Ke^{-rT})}{\sqrt{\pi}e}\left(\frac{2N+1}{2N}\right)^{2N+\frac{3}{2}} \int_0^{\tilde{g}\left(\frac{1}{N}\right)} e^{-w^2} \alpha(0) \, dw$$

$$+ \left(\frac{1}{N}\right)^{\frac{1}{2}} \frac{2(S_0 - Ke^{-rT})}{\sqrt{\pi}e}\left(\frac{2N+1}{2N}\right)^{2N+\frac{3}{2}} \int_0^{\tilde{f}\left(\frac{1}{N}\right)} e^{-w^2} \frac{d\alpha}{dw}(0) \, w dw + O\left(\frac{1}{N}\right),$$

where the functions $\tilde{g}$ and $\tilde{f}$ are smooth, the function α is analytic and smooth, and the variable w is a function of $\frac{1}{N}$.

The proofs of the first two have been discussed in Sections 2 and 3. The last item in the theorem stems from Theorem 2 of Joshi [3]:

Theorem 2 *Each of* $\phi_q(N) = \Phi(2N+1, N+1, q_{2N+1})$ *and* $\phi_p(N) = \Phi(2N+1, N+1, p_{2N+1})$ *has an asymptotic expansion in powers of* N^{-1}*, and in particular, there exist constants* C_0 *and* C_1 *such that*

$$\phi(N) = C_0 + \frac{1}{N}C_1 + \mathcal{O}(N^{-2}).$$

References

[1] F. Diener and M. Diener, *Asymptotics of the Price Oscillations of a European Call Option in a Tree Model*. Mathematical Finance 14/2 (2004), 271–293.

[2] F. Diener and M. Diener, *Asymptotics of the Price of a Barrier Option in a Tree Model*. Preprint du Laboratoire de Mathematiques Dieudonné, Nice (2006), 1–19.

[3] M. Joshi, *Achieving Smooth Asymptotics for the Prices of European Options in Binomial Trees*. Available at SSRN (2006): `http://ssrn.com/abstract=928186`.

[4] J. Cox, S. Ross, and M. Rubenstein, *Option Pricing: A Simplified Approach*. Journal of Financial Economics 7 (1979), 229–263.

[5] J. Hull, "Options, Futures, And Other Derivatives". Prentice Hall, 2002.

[6] S. Shreve, "Stochastic Calculus for Finance I: The Binomial Asset Pricing Model". Springer, 2000.

[7] T. Bjork, "Arbitrage Theory in Continuous Time". Oxford University Press, 2004.

[8] S. R. Pliska, "Introduction to Mathematical Finance, Discrete Time Models". Blackwell, 1997.

[9] E. T. Copson, "Asymptotic Expansions". Cambridge University Press, 1965.

[10] D. Andre, *Solution directe du probleme resolu par M. Bertrand*. Comptes Rendus Acad. Sci. Paris 105 (1887), 436–437.

[11] E. Catalan, *Note extraite dune lettre adressée*. J. Reine Angew. Math. 27:192.

[12] B. A. Odegaard, *Binomial Option Pricing Recipes*. `http://www.finance-old.bi.no/~bernt/gcc_prog/recipes/recipes/node6.html` .

[13] M. Abramowitz and I. Stegun, "Handbook of Mathematical Functions". `http://www.math.hkbu.edu.hk/support/aands/toc.htm` .

[14] R. B. Paris and D. Kaminsky, "Asymptotics and the Mellin-Barnes Integrals". Cambridge University Press, 2001.

An invitation to the study of Isoperimetric Inequalities

Francescopaolo Montefalcone (*)

Abstract. This is a very short introduction to Isoperimetric Inequalities. In particular, we shall discuss the so-called Michael-Simon Inequality (MSI) for compact hypersurfaces of class $\mathbf{C}^2$ immersed in the Euclidean space $\mathbb{R}^n$. In the last section, we give a generalization to (the non-Euclidean setting of) hypersurfaces immersed in sub-Riemannian Carnot groups.

1 Introduction: from Queen Dido to... J. Steiner

Problem 1 *Find the plane figure bounded by a line which has the maximum area for a given perimeter.*

It is well-known that the solution is a semicircle. Actually, this problem is based on a passage from Virgil's Aeneid. Closely related to this problem...

Theorem 1 (Isoperimetric Inequality in the Euclidean plane) *Let $\Omega \subsetneq \mathbb{R}^2$ be a bounded domain (i.e. open, connected) and assume that its boundary $\partial\Omega$ is* rectifiable *(i.e. covered, up to a negligible set, by a countable union of smooth curves; see Federer, [6]). Then*

$$4\pi \,\mathrm{Area}(\Omega) \leq \mathrm{Lenght}^2(\partial\Omega). \tag{1}$$

Under suitable assumptions, one has equality, if and only if, Ω is a disc.

Remark 2 (Curiosity: quick proof for $n = 2$; see Chavel, [3]) Let $\Omega \subset \mathbb{R}^2$ be a compact domain having $\mathbf{C}^1$-smooth boundary consisting of 1 component. Set $z := x + iy$ to denote a point in $\mathbb{R}^2 \cong \mathbb{C}$ and $\overline{z} := x - iy$. Then, the area element is given by

$$dx \wedge dy = \frac{i}{2} dz \wedge d\overline{z}.$$

(*)Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: montefal@math.unipd.it. Seminar held on March 21st, 2012.

One computes

$$\begin{aligned} 4\pi \,\mathrm{Area}(\Omega) &= \iint_\Omega 2\pi i dz \wedge d\overline{z} \\ &= \iint_\Omega dz \wedge d\overline{z} \int_{\partial\Omega} \frac{d\zeta}{\zeta - z} \\ &= \int_{\partial\Omega} d\zeta \iint_\Omega \frac{dz \wedge d\overline{z}}{\zeta - z} \\ &= \int_{\partial\Omega} d\zeta \int_{\partial\Omega} \frac{\overline{\zeta} - \overline{z}}{\zeta - z} dz \\ &\leq \mathrm{Lenght}^2(\partial\Omega), \end{aligned}$$

where we have used Cauchy formula together with the fact that the winding number of $\partial\Omega$ is 1, Fubini's Theorem and Green's formula.

The first attempt to give a rigorous proof of (1) was made by the Swiss geometer J. Steiner whose important contribution was the *Steiner symmetrization method*; see [2], [3], [5], [6]. It can be described as follows:

Remark 3 (The method in $\mathbb{R}^n$) Let $\Omega \subset \mathbb{R}^n$. Fix $V \in \mathbb{R}^n$ and let $V^\perp$ be the $(n-1)$-dimensional hyperplane passing from $0 \in \mathbb{R}^n$ and orthogonal to V; furthermore, let γ_p^V be the V-line starting from $p \in \mathcal{P}roj_V(\Omega)$. Denote by $l_p := \mathcal{H}^1(\gamma_p^V \cap \Omega)$ the 1-dimensional Hausdorff measure of $\gamma_p^V \cap \Omega$. Now let us define another set Ω^V by taking the segment with midpoint at p and length $l_p \; \forall \, p \in \mathcal{P}roj_V(\Omega)$. The set Ω^V is called the Steiner symmetrized of Ω w.r.t. the hyperplane $V^\perp$. *In particular, the following hold:*

- *Fubini's Theorem* $\Rightarrow \mathrm{Vol}(\Omega) = \mathrm{Vol}(\Omega^V)$*;*
- $\mathcal{H}^{n-1}(\partial\Omega^V) \leq \mathcal{H}^{n-1}(\partial\Omega)$*.*

The power of Steiner's idea was exploited by De Giorgi in 1958 (see [4]) in the (most general) setting of *finite perimeter sets in the Euclidean space* $\mathbb{R}^n$; see [2], [3], [5].

Let $A \subseteq \mathbb{R}^n$ be open and let $\Omega \subset \mathbb{R}^n$. The *perimeter* $|\partial\Omega|(A)$ of Ω in A is given by

$$|\partial\Omega|(A) := \sup\left\{ \int_{A\cap\Omega} div\,(\psi)\, dx : \psi \in C_0^1(\Omega, \mathbb{R}^n),\, |\psi| \leq 1 \right\}$$

The perimeter is a Radon measure that plays the role of *measure of the boundary.*

Theorem 4 (Isoperimetric Inequality; see [4]) Let $\Omega \subset \mathbb{R}^n$ be a set of finite perimeter. Then

$$\min\{\mathrm{Vol}(\Omega), \mathrm{Vol}(\mathbb{R}^n \setminus \Omega)\} \leq C_{Isop} |\partial\Omega|(\mathbb{R}^n), \tag{2}$$

where $C_{Isop} := \frac{1}{n^{\frac{n}{n-1}}} \times \frac{1}{\omega_n^{\frac{1}{n-1}}}$ and $\omega_n := \mathrm{Vol}(B_1(0))$ denotes the n-volume of the unit ball $B_1(0) \subset \mathbb{R}^n$. Equality holds, if and only if, Ω is $\mathcal{L}^1$-equivalent to a ball.

Apart from its "beauty", the Isoperimetric Inequality is a fundamental property of the Euclidean space and, for instance, it turns out to be equivalent to a Sobolev Inequality. This was shown in 1960 by Federer-Fleming, [7], and, independently, by Mazja, [10]. More precisely, (2) is equivalent to the inequality:

$$\|\varphi\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq C_{Isop} \|grad\, \varphi\|_{L^1(\mathbb{R}^n)} \quad \forall\ \varphi \in \mathbf{C}_0^\infty(\mathbb{R}^n). \tag{3}$$

Note that the proof of the equivalence of (2) and (3) is based on a careful use of the so-called *Fleming-Rishel Formula:*

$$\int_\Omega |grad\, \phi(x)|\, dx = \int_{\mathbb{R}} |\partial \Omega_t|(\mathbb{R}^n)\, dt, \tag{4}$$

where $\Omega_t = \{x \in \Omega : \phi > t\}$ and ϕ is a smooth function. This formula generalizes to the space BV of bounded variation functions; see [5], [16].

Some very general isoperimetric theorems for currents can be found in Federer's treatise [6].

Later on we shall discuss an important generalization of the Isoperimetric Inequality proven in the Seventies by Michael and Simon; see [11]. A similar result was independently proved by Allard for varifolds; see [1].

2 The Micheal-Simon Inequality (MSI)

Let $S \subset \mathbb{R}^n$ $(n > 2)$ be a compact $\mathbf{C}^2$-smooth hypersurface with $\mathbf{C}^1$-smooth boundary.

Theorem 5 (see [11]) *Under the previous assumptions, the following holds:*

$$\left(\sigma_R^{n-1}(S)\right)^{\frac{n-2}{n-1}} \leq C_{Isop} \left(\int_S |\mathcal{H}_R|\, \sigma_R^{n-1} + \sigma_R^{n-2}(\partial S) \right)$$

where $C_{Isop} > 0$ is a dimensional constant, $\mathcal{H}_R$ is the mean curvature of S and σ_R^k is the k-dimensional Riemannian measure.

The constant C_{Isop} in Theorem 5 is not the sharp one.

A more general presentation of these results can be found in the book [2].

2.1 Some preliminaries about hypersurfaces and measures

In this section we will collect some preliminaries concerning geometry of hypersurfaces, measures and differential forms for which we refer the reader to [3], [9] and [16].

Fix an o.n. frame $\underline{X} = \{X_1, ..., X_n\}$ in $\mathbb{R}^n$. By duality w.r.t. the (Euclidean) metric we define a co-frame $\underline{\omega} = \{\omega_1, ..., \omega_n\}$ by setting $\omega_i(X_j) = \langle X_i, X_j \rangle = \delta_i^j$, where δ_i^j is the Kronecker delta.

In the sequel, we shall denote by ∇ the *Levi-Civita connection on* $\mathbb{R}^n$, which is *flat,* i.e. $\nabla_{X_i} X_j = 0$ for all $i, j = 1, ..., n$, *compatible with the Euclidean metric* $\langle \cdot, \cdot \rangle$, i.e.

$$X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle \qquad \forall\, X, Y, Z \in \mathfrak{X}(\mathbb{R}^n) := \mathbf{C}^\infty(\mathbb{R}^n, \mathbb{R}^n),$$

and *torsion-free, i.e.* $\nabla_X Y - \nabla_Y X = [X,Y] \quad \forall\, X,Y,Z \in \mathfrak{X}(\mathbb{R}^n)$*;* see [3].

The *volume form* of $\mathbb{R}^n$ is given by $\sigma_R^n := \bigwedge_{i=1}^{n} \omega_i \in \bigwedge^n(T^*\mathbb{R}^n)$. By integration of σ_R^n, one obtains the Lebesgue measure $\mathcal{L}^n$ on $\mathbb{R}^n$.

Let $S \subset \mathbb{R}^n$ be a $\mathbf{C}^r$-smooth hypersurface, $r \geq 1$, and let ν be the unit normal vector to S. By definition, we have

$$\sigma_R^{n-1} := (\nu \lrcorner \sigma_R^n)|_S,$$

where $\lrcorner$ is the *contraction operator*; see [9]. We recall that $\lrcorner : \bigwedge^k(T^*\mathbb{R}^n) \to \bigwedge^{k-1}(T^*\mathbb{R}^n)$ is defined, for $X \in T\mathbb{R}^n$ and $\alpha \in \bigwedge^k(T^*\mathbb{R}^n)$, by

$$(X \lrcorner \alpha)(Y_1, ..., Y_{k-1}) := \alpha(X, Y_1, ..., Y_{k-1}).$$

Example 6 (see [9]) Let $X_1 = \mathrm{e}_1 = (1,0,0)$, $X_2 = \mathrm{e}_2 = (0,1,0)$, $X_3 = \mathrm{e}_3 = (0,0,1)$. In other words, let $\{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3\}$ be the standard basis of $\mathbb{R}^3$. Then $\omega_i = dx_i$ for any $i = 1,2,3$ and it follows that $\sigma_R^3 = dx_1 \wedge dx_2 \wedge dx_3$. Now let S be a $\mathbf{C}^1$-smooth surface and let $\nu = (\nu_1, \nu_2, \nu_3)$ be its unit normal. Then

$$\begin{aligned} \sigma_R^2 &:= \nu \lrcorner \sigma_R^3 \\ &= \nu_1 dx_2 \wedge dx_3 - \nu_2 dx_1 \wedge dx_3 + \nu_3 dx_1 \wedge dx_2. \end{aligned}$$

Let ∂S be $\mathbf{C}^1$-smooth and let $\eta \in \mathfrak{X}(TS)$ be its unit normal vector. As above, the Riemannian measure on ∂S can be defined by setting

$$\sigma_R^{n-2} := (\eta \lrcorner \sigma_R^{n-1})|_{\partial S}.$$

Furthermore, let ∇^{TS} be the induced connection from ∇ on TS. Later on, the TS-gradient and the TS-divergence operators will be denoted by $grad_{\,TS}$ and $div_{\,TS}$, respectively.

Proposition 7 (Coarea formula for the TS-gradient) *Let $S \subset \mathbb{R}^n$ be a $\mathbf{C}^2$-smooth hypersurface and let $\phi \in \mathbf{C}^1(S)$. Then*

$$\int_S |grad_{\,TS}\, \phi|\, \sigma_R^{n-1} = \int_{\mathbb{R}} \sigma_R^{n-2} \left\{\phi^{-1}[s] \cap S\right\} ds.$$

Remark 8 It is important to recall that the *2nd fundamental form* of a $\mathbf{C}^\infty$-smooth hypersurface S is the $\mathbf{C}^\infty$-smooth symmetric bilinear map

$$B_R(X,Y) := \left\langle \nabla_X^{TS} Y, \nu \right\rangle \qquad \forall\ X,\ Y \in \mathfrak{X}(TS) := \mathbf{C}^\infty(S, TS).$$

The *mean curvature* $\mathcal{H}_R$ is defined as $\mathcal{H}_R := \mathrm{Tr} B_R = -div_{\,TS}\, \nu$.

Definition 9 A variation $\vartheta : (-\epsilon, \epsilon) \times S \longrightarrow \mathbb{R}^n$ of S is a smooth map such that $\vartheta_t := \vartheta(t, \cdot) : S \longrightarrow \mathbb{R}^n$ is an immersion for every $t \in (-\epsilon, \epsilon)$ and $\vartheta_0 = \imath_S$. The variation

vector *of* ϑ_t *is defined as* $X := \frac{\partial \vartheta}{\partial t}\big|_{t=0}$. *Setting* $\Gamma(t) := \vartheta_t^* \sigma_R^{n-1}$, *the* 1st variation of σ_R^{n-1} *is given by*

$$I_S(X, \sigma_R^{n-1}) := \frac{d}{dt}\left(\int_S \Gamma(t)\right)\Bigg|_{t=0}.$$

Theorem 10 (1st variation for compact hypersurfaces) *Let* $S \subset \mathbb{R}^n$ *be a compact* $\mathbf{C}^2$*-smooth hypersurface with* $\mathbf{C}^1$*-smooth boundary. Under the previous notation, we have*

$$I_S(X, \sigma_R^{n-1}) = -\int_S \mathcal{H}_R \, \langle X, \nu \rangle \, \sigma_R^{n-1} + \int_{\partial S} \langle X, \eta \rangle \, \sigma_R^{n-2} \tag{5}$$

Remark 11 When div_{TS} acts on non-tangential vector fields, it is possible to reformulate (5) as an "integration by parts formula", i.e.

$$\int_S (div_{TS} X + \mathcal{H}_R \langle X, \nu \rangle) \, \sigma_R^{n-1} = \int_{\partial S} \langle X, \eta \rangle \, \sigma_R^{n-2}.$$

2.2 Sketch of Proof of MSI

First, we make a "suitable" choice of $X \in \mathfrak{X}(\mathbb{R}^n)$ into the last formula

$$\int_S (div_{TS} X + \mathcal{H}_R \langle X, \nu \rangle) \, \sigma_R^{n-1} = \int_{\partial S} \langle X, \eta \rangle \, \sigma_R^{n-2}.$$

So let us fix $x \in \mathbb{R}^n$ and choose $X(y) = y - x$. Using Cauchy-Schwarz yields

$$(n-1)\, \sigma_R^{n-1}(S) \leq r \left(\int_S |\mathcal{H}_R| \, \sigma_R^{n-1} + \sigma_R^{n-2}(\partial S) \right), \tag{6}$$

where r is the radius of a Euclidean ball $B_r(x)$ containing S. The inequality (6) is sometimes called *Linear Isoperimetric Inequality*. Using jointly (6) and Coarea formula, yields the next *key-property*:

Lemma 12 (Monotonicity Inequality) *The following inequality holds:*

$$-\frac{d}{dt}\left(\frac{\sigma_R^{n-1}(S_t)}{t^{n-1}}\right) \leq \frac{1}{t^{n-1}} \left(\mathcal{A}(t) + \mathcal{B}(t)\right)$$

for $\mathcal{L}^1$*-a.e.* $t > 0$, *where* $S_t := S \cap B_t(x)$ *and we have set*

$$\mathcal{A}(t) := \int_{S_t} |\mathcal{H}_R| \, \sigma_R^{n-1}, \qquad \mathcal{B}(t) := \sigma_R^{n-2}(\partial S \cap B_t(x)).$$

Remark 13 If S is minimal and $\partial S = \emptyset$, we have the (simpler) formula:

$$\frac{d}{dt}\left(\frac{\sigma_R^{n-1}(S_t)}{t^{n-1}}\right) \geq 0$$

for $\mathcal{L}^1$-a.e. $t > 0$. Moreover, we stress that $\sigma_R^{n-2}(\partial S_t) = \mathcal{B}(t) + \mathcal{B}_1(t)$, where $\mathcal{B}_1(t) := \sigma_R^{n-2}(\partial B_t(x) \cap S)$. The proof of Lemma 12 is based on the inequality:

$$\mathcal{B}_1(t) \leq \frac{d}{dt}\sigma_R^{n-1}(S_t) \tag{7}$$

for $\mathcal{L}^1$-a.e. $t > 0$. Note that, in turn, (7) follows from the well-known eikonal inequality:

$$|grad_{TS}\, d_{Eu}| \leq 1.$$

Using this estimate and the linear inequality for $S_t = S \cap B(x,t)$ yields:

$$(n-1)\,\sigma_R^{n-1}(S_t) \leq t\left(\mathcal{A}(t) + \mathcal{B}(t) + \frac{d}{dt}\sigma_R^{n-1}(S_t)\right),$$

which is equivalent to the monotonicity inequality.

Lemma 14 (Vitali-type Covering theorem) *Let (X, ϱ) be a compact metric space and $A \subset X$. Let C be a covering of A by closed d-balls with centers in A. We also assume that each point x of A is the center of (at least) one closed ϱ-ball belonging to C and that the radii of the balls of the covering are uniformly bounded by some positive constant. Then, for every $\lambda > 2$ there exists a no more than countable subset $\mathsf{C}' \subsetneq \mathsf{C}$ of pairwise non-intersecting closed balls $\overline{B}_{r_k}(x_k)$ such that*

$$A \subset \bigcup_k B_{\lambda\, r_k}(x_k).$$

Monotonicity implies an estimate modelled on the previous covering lemma.

Lemma 15 (Calculus Lemma) *Let $x \in \mathrm{Int}(S)$ and set*

$$r_0 = 2\left(\frac{\sigma_R^{n-1}(S)}{\omega_{n-1}}\right)^{1/n-1}.$$

Then, for every $\lambda \geq 2$ there exists $r \in]0, r_0[$ such that

$$\sigma_R^{n-1}(S_{\lambda r}) \leq \lambda^{Q-1} r_0 \left(\mathcal{A}(r) + \mathcal{B}(r)\right).$$

More precisely, the previous lemma follows from a contradiction argument based on the Monotonicity formula. Putting all together, the proof of MSI easily follows.

Remark 16 (Asymptotic) Monotonicity is equivalent to an exponential estimate:

$$\sigma_R^{n-1}(S_t) \geq \omega_{n-1}\, t^{n-1} e^{-\mathcal{H}_0 t}$$

as long as $t \to 0^+$, where $|\mathcal{H}_R| \leq \mathcal{H}_0$. For minimal hypersurfaces the estimate becomes

$$\sigma_R^{n-1}(S_t) \geq \omega_{n-1}\, t^{n-1} \qquad \text{as } t \to 0^+.$$

Finally, a classical argument due to Federer-Fleming and Maz'ja, implies the following:

Corollary 17 (Sobolev-type Inequality) *Let $S \subset \mathbb{R}^n$ be a $\mathbf{C}^2$-smooth closed hypersurface (without boundary). Then*

$$\left(\int_S |\psi|^{\frac{n-1}{n-2}} \sigma_R^{n-1}\right)^{\frac{n-2}{n-1}} \leq C_{Isop} \int_S (|\mathcal{H}_R||\psi| + |grad_{TS}\,\psi|)\, \sigma_R^{n-1}.$$

for every $\psi \in \mathbf{C}_0^\infty(S)$.

3 An extension of MSI to Carnot groups

Before stating our theorem, which generalizes MSI to hypersurfaces immersed in sub-Riemannian Carnot groups, we need some preliminaries; see [8], [12, 13], [14], [15].

A *k-step Carnot group* $(\mathbb{G}, \bullet)$ is a connected, simply connected, nilpotent and stratified Lie group. The Lie algebra $\mathfrak{g} \cong \mathbb{R}^n$ satisfies:

$$\mathfrak{g} = H \oplus H_2 ... \oplus H_k, \qquad [H_1, H_{i-1}] = H_i \ \ \forall\, i = 2, ..., k, \qquad H_{k+1} = \{0\}.$$

We also set $V := H^\perp$. Notice that H and V are smooth subbundles of $T\mathbb{G}$ called *horizontal* and *vertical*. Hereafter, we shall use the abbreviation L.I.= *left invariant*. Let $\underline{X_H} := \{X_1, ..., X_h\}$ be a L.I. frame for H. $\underline{X_H}$ can be completed to a global L.I. frame $\underline{X} := \{X_1, ..., X_n\}$ for $T\mathbb{G}$. We have $X_i(x) = L_* \mathrm{e}_i$, where $\{\mathrm{e}_i : i = 1, ..., n\}$ is a Euclidean basis of $\mathfrak{g} = \mathbb{R}^n$. We shall use exponential coordinates of the 1st kind, so that $\mathbb{G}$ identifies to $\mathfrak{g}$, via the exponential map exp.

Furthermore, let us fix a metric $g = \langle \cdot, \cdot \rangle$ on $\mathfrak{g} = T_0\mathbb{G}$ for which $\{X_i : i = 1, ..., n\}$ is an o.n. basis. Therefore, $(\mathbb{G}, g)$ is a Riemannian manifold, which has a (unique) L.I. Levi Civita connection ∇.

A *sub-Riemannian metric* g_H is a symmetric positive bilinear form on H. The *CC-distance* $d_{\mathbf{cc}}(x, y)$ between $x, y \in \mathbb{G}$ is

$$d_{\mathbf{cc}}(x, y) := \inf \int \sqrt{g_H(\dot\gamma, \dot\gamma)} dt,$$

where the infimum is taken over all smooth horizontal curves γ joining the points. We shall fix $g_H := g_{|H}$.

Carnot groups are *homogeneous groups*, i.e. they admit a 1-parameter group of dilations $\delta_t : \mathbb{G} \longrightarrow \mathbb{G}$ $(t \geq 0)$ given by

$$\delta_t x := \exp\left(\sum_{j,i_j} t^j \, x_{i_j} \mathrm{e}_{i_j}\right),$$

where $x = \exp\left(\sum_{j,i_j} x_{i_j} \mathrm{e}_{i_j}\right) \in \mathbb{G}$. The *homogeneous dimension* of $\mathbb{G}$ is $Q := \sum_{i=1}^{k} i\, h_i$, where $h_i = \dim H_i$. We stress that the integer Q turns out to be the *Hausdorff dimension* of $(\mathbb{G}, d_{\mathbf{cc}})$ as a metric space; see [8], [14], [15].

Example 18 (The 1st Heisenberg group $\mathbb{H}^1$) Let $(\mathbb{R}^3, \bullet)$ endowed with the L.I. frame $\{X, Y, T\}$,

$$X = \mathrm{e}_1 - \frac{y}{2}\mathrm{e}_3, \quad Y = \mathrm{e}_2 + \frac{x}{2}\mathrm{e}_3, \quad T = \mathrm{e}_3,$$

where $\bullet$ is a polynomial operation defined by

$$(x, y, t) \bullet (\xi, \eta, \tau) = \left(x + \xi, y + \eta, t + \tau + \frac{1}{2}\left(x\eta - y\xi\right)\right).$$

Obviously, we have $H = \mathrm{span}_{\mathbb{R}}\{X, Y\}$, $V = \mathrm{span}_{\mathbb{R}}\, T$ and $Q = 4$.

Remark 19 (A motivation) We recall that Pansu in his PhD Thesis (1982) proved that for every smooth compact domain $\Omega \subset \mathbb{H}^1$ one has the following non-Euclidean Isoperimetric Inequality:

$$(\mathrm{Vol}(\Omega))^{\frac{3}{4}} \leq C \mathrm{Area}(\partial\Omega)$$

for some dimensional constant $C > 0$, where Area is the intrinsic 3-dimensional Hausdorff measure w.r.t. the CC-distance $d_{\mathbf{cc}}$. This was the first example of a non-Euclidean isoperimetric inequality on (sub-Riemannian) Lie Groups; see also [15].

The L.I. volume of $\mathbb{G}$ is given by $\sigma_R^n := \bigwedge_{i=1}^{n} \omega_i \in \bigwedge^n(T^*\mathbb{G})$. Note that σ_R^n is a Haar measure and equals (up constants) the Lebesgue measure $\mathcal{L}^n$ on $\mathbb{R}^n$.
Let $S \subset \mathbb{G}$ be a $\mathbf{C}^1$-smooth hypersurface. We say that $x \in S$ is a *characteristic point* if

$$\dim\, H_x = \dim(H_x \cap T_x S).$$

We also set $C_S := \{x \in S : \dim\, H_x = \dim(H_x \cap T_x S)\}$ to denote the *characteristic set* of S. Equivalently, $x \in S$ is non-characteristic if, and only if, H intersects transversally S, i.e. $H \pitchfork S$ at x. We stress that $\mathcal{S}_{\varrho}^{Q-1}(C_S) = 0$, where $\mathcal{S}_{\varrho}^{k}$ denotes, more generally, the intrinsic k-dimensional spherical Hausdorff measure (w.r.t. the CC-distance $d_{\mathbf{cc}}$).

Let ν be the unit normal vector to S. Then, as in the Euclidean setting, we set

$$\sigma_R^{n-1} := (\nu \lrcorner \sigma_R^n)|_S.$$

At this point, let us define the *unit H-normal* as $\nu_H := \frac{\mathcal{P}roj_H \nu}{|\mathcal{P}roj_H \nu|}$. Accordingly, we define the *H-perimeter* $\sigma_H^{n-1} \in \bigwedge^{n-1}(T^*S)$ by setting

$$\sigma_H^{n-1} := (\nu_H \lrcorner \sigma_R^n)|_S.$$

The following hold:

- $\sigma_H^{n-1} = |\mathcal{P}roj_{\,H}\,\nu|\,\sigma_R^{n-1}$,
- $C_S = \{x \in S \,||\mathcal{P}roj_{\,H}\,\nu| = 0\}$,
- σ_H^{n-1} is equivalent to $\mathcal{S}_\varrho^{Q-1}$ up to a bounded density called *metric factor.*

Furthermore, the *unit HS-normal* along ∂S is given by $\eta_{HS} := \frac{\mathcal{P}roj_{HS}\,\eta}{|\mathcal{P}roj_{HS}\,\eta|}$ and we set

$$\sigma_H^{n-2} := \left(\eta_{HS} \lrcorner\, \sigma_H^{n-1}\right)\big|_{\partial S} \in \bigwedge^{n-2}(T^*\partial S).$$

Also in this case we have the representation formula:

$$\sigma_H^{n-2} = |\mathcal{P}roj_{\,H}\,\nu|\,|\mathcal{P}roj_{HS}\,\eta|\,\sigma_R^{n-2}.$$

It is important to remark that σ_H^{n-2} is equivalent to $\mathcal{S}_\varrho^{Q-2}$, up to a bounded density function, which is called *metric factor.*

The *horizontal tangent bundle* $HS \subset TS$ and the *horizontal normal bundle* $\nu_H S$ satisfy the splitting $H = \nu_H \oplus HS$.

From now on, let $S \subset \mathbb{G}$ be a $\mathbf{C}^2$-smooth hypersurface and let ∇^{TS} be the induced connection from ∇ on TS. The tangential connection ∇^{TS} induces a *partial connection* ∇^{HS} on HS as follows:

$$\nabla^{HS}{}_X Y := \mathcal{P}roj_{HS}\left(\nabla_X^{TS} Y\right) \qquad \forall\, X, Y \in \mathfrak{X}^1(HS) := \mathbf{C}^1(S, HS).$$

Finally, the *2nd horizontal fundamental form* is the map

$$B_H(X,Y) := \langle \nabla^H{}_X Y, \nu_H \rangle \qquad X,\, Y \in \mathfrak{X}^1(HS).$$

The trace of B_H is the *H-mean curvature* $\mathcal{H}_H$, i.e. $\mathcal{H}_H := \mathrm{Tr} B_H = -div_{\,H}\,\nu_H$. We remark that, in general, B_H is not symmetric: this is a big difference from the Riemannian case.

At this point, our main result can be stated as follows:

Theorem 20 (MS type inequality; see [13]) *Let $S \subset \mathbb{G}$ be a compact hypersurface of class $\mathbf{C}^2$ with piecewise $\mathbf{C}^2$-smooth boundary ∂S. Then*

$$\left(\sigma_H^{n-1}(S)\right)^{\frac{Q-2}{Q-1}} \leq C_{Isop}\left(\int_S |\mathcal{H}_H|\,\sigma_H^{n-1} + \sigma_H^{n-2}(\partial S)\right).$$

where C_{Isop} is a constant independent of S.

As in the Euclidean case, we can deduce asymptotic estimates for σ_H^{n-1} (but, in this case, of local type) and Sobolev-type inequalities involving *HS*-derivatives only.

References

[1] W. K. Allard, *On the first variation of a varifold.* Ann. Math. **95** (1972), 417-491.

[2] Yu. D. Burago, V.A. Zalgaller, "Geometric Inequalities". Springer Verlag, 1988.

[3] I. Chavel, "Isoperimetric inequalities: Differential geometric and analytic perspectives". University Press, Cambridge, UK, 2001.

[4] E. De Giorgi, *Sulla proprietà isoperimetrica dell'ipersfera, nella classe degli insiemi aventi frontiera orientata di misura finita.* Memorie Accad. Naz. Lincei, serie 8, 5 (1958).

[5] L. C. Evans and R.F. Gariepy, "Measure Theory and Fine Properties of Functions". CRC Press, 1992.

[6] H. Federer, "Geometric Measure Theory". Springer Verlag, 1969.

[7] H. Federer, W. H. Fleming, *Normal and Integral Currents.* Ann. Math. **72** (1960), 458-520.

[8] M. Gromov, *Carnot-Carathéodory spaces seen from within.* In "Subriemannian Geometry", Progress in Mathematics 144, ed. by A. Bellaiche and J. Risler, Birkhauser.

[9] J. M. Lee, "Introduction to smooth manifolds". Graduate Texts in Mathematics, Springer.

[10] V. G. Maz'ya, *Classes of domains and embedding theorem for functional spaces.* Dokl. Acad. Nauk SSSR, **133** (1960) no. 3, 527-530 (in Russian). Engl. transl. Soviet Math. Dokl., **1** (1961), 882–885.

[11] J. H. Michael, L. M. Simon, *Sobolev and Mean-value Inequalities on Generalized Submanifolds of* $\mathbb{R}^n$. Comm. on Pure and Applied Math., Vol. XXVI (1973), 361-379.

[12] F. Montefalcone, *Hypersurfaces and variational formulas in sub-Riemannian Carnot groups.* Jour. Math. Pures et Appliquées **87** (2007), 453-494.

[13] ————, *Isoperimetric and Sobolev inequalities on hypersurfaces in sub-Riemannian Carnot groups.* Preprint ArXiv (2010), 50 pages.

[14] R. Montgomery, "A Tour of Subriemannian Geometries, Their Geodesics and Applications". AMS, Math. Surveys and Monographs 91, 2002.

[15] P. Pansu, *Submanifolds and differential forms in Carnot manifolds, after M. Gromov et M. Rumin.* Preprint (2005), 41 pages.

[16] L. M. Simon, "Lectures on geometric measure theory". Proceedings of the Centre for Mathematical Analysis, Austr. Nat. Univ. 1983.

Moment Problems and Spin Correlation Matrices

Neeraja Sahasrabudhe (*)

Abstract. Moment problems (or more specifically, covariance realization problems) are about realizability of a given pair correlation function or covariances (or higher moments), namely whether a probability distribution is determined by its moments. That is, given $m_0, m_1, \ldots \in \mathbb{R}$, one wants to find a probability measure μ such that $\int_{-\infty}^{\infty} x^k d\mu(x) = m_k$ for $k = 0, 1, \ldots$
Generalized moment problems of this kind have been widely studied, mainly in the theoretical Engineering community, for continuous random variables.
We look at the specific case of covariance realization problem for spin systems and discuss the necessary and sufficient conditions for a correlation matrix of order $n \geq 2$ to be the correlation matrix of spin variables in the classical sense and finally try to give an algorithm to explicitly compute the probability measure that realizes the given correlations.

1 Introduction

1.1 Moment Problem

The moment problem is essentially the question of existence of a probability distribution that realizes the given set of moments. That is, given $m_0, m_1, \ldots \in \mathbb{R}$, one wants to find a probability measure μ such that

$$\int_{-\infty}^{\infty} x^k d\mu(x) = m_k \tag{1}$$

for $k = 0, 1, \ldots$.

If the probability measure that realizes the given moments is unique then the moment problem is called determinate, if not we call it indeterminate.

It was Thomas Jan Stieltjes who in 1894-1895 formulated the moment problem on positive real axis and used it as a means of studying the analytic behavior of continued fractions, in which connection he invented the important Stieltjes integral. Later, in

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: neeraja@math.unipd.it. Seminar held on April 5th, 2012.

1920, Hamburger extended the "Stieltjes moment problem," which was only defined on the positive real axis, to the "Hamburger moment problem," which is defined on the whole real axis. Hamburger also introduced the modern definition of determinateness and indeterminateness, which does not depend on the corresponding continued fractions.

This was the first profound and complete treatment of the moment problem. From being primarily a tool for the determination of convergence/divergence of continued fractions it now became established as an independent problem.

The moment problems are just a special case of what are known as the **Generalized moment Problems**. The Generalized moment Problems are of particular interest in linear optimization and can be stated as follows:

Let $(E, \mathcal{E})$ be a measurable space and $\mathcal{F}$ be a set of measurable functions

$$f : E \longrightarrow \mathbb{C} \tag{2}$$

Given, $c_f \in \mathbb{C}$, find a measure μ such that

$$\int_E f(t) d\mu(t) = c_f \tag{3}$$

Remark 1 It is obvious that the moment problems are a special case of the generalized moment problem defined above for $f(x) = x^k$.

1.2 Covariance Realization Problem

Covariance realizations are also a special case of Generalized Moment Problem and can be stated as follows:

Given a matrix $C = ((c_{ij}))$, find random variables $X_1, X_2, \ldots, X_n$ such that

$$c_{ij} = \mathrm{Cov}[X_i, X_j]. \tag{4}$$

The idea is to find a joint distribution of $(X_1, X_2, \ldots, X_n)$ which realizes the given covariance matrix. The main objective of this note is to discuss a covariance realization problem for spin systems.

2 Covariance Matrices

2.1 What are Covariance Matrices?

Consider a vector of random variables $X = \{X_1, X_2, \ldots, X_n\}$ with μ being a joint probabiliti distribution of X. We call $C = ((c_{ij}))_{n \times n}$ where $c_{ij} = Cov[X_i, X_j]$ as the covariance matrix of X. We define the Correlations matrix as

$$\mathrm{corr}_{ij} = \frac{\mathrm{Cov}[X_i, X_j]}{\sqrt{\mathrm{Var}(X_i)\,\mathrm{Var}(X_j)}} \tag{5}$$

Covariance matrices are necessarily symmetric and positive semi-definite.

Fact: Every symmetric positive semi-definite matrix is a covariance matrix.

Proof. Let M be a $n \times n$ symmetric, positive semi-definite matrix then, by the spectral theorem we know that there exists a matrix $M^{1/2}$ whose square is M. Now consider a vector of mutually independent random variables $X = (X_1, X_2, \ldots, X_n)$ such that the covariance matrix of X is an identity matrix. Then, the covariance matrix of $M^{1/2}X$ is given by M. □

So, being symmetric and positive semi-definite is necessary and sufficient for a matrix to be a covariance matrix.

Remark 2 Every symmetric positive semi-definite matrix with 1's on the diagonal is a correlation matrix and vice versa. So to understand the geometry of the set of correlation matrices is equivalent to understanding the geometry of the set of covariance matrices.

Denote by $\mathcal{C}_n$ the set of all such nth order correlation matrices. $\mathcal{C}_n$ is a convex set in the $n(n-1)$-dimensional Euclidean space of all symmetric matrices of order n with 1 on the diagonal.

We already have a description of this space in terms of hyperplanes. (the condition of positive semi-definiteness).

2.2 Extremal Correlations

Let $\mathcal{E}_n \subset \mathcal{C}_n$ be the subset of extremals of the set of correlation matrices of order n. What are the elements of $\mathcal{E}_n$?

A lot of work was done over the years to understand the geometry of this convex space and to find the extremals significantly by Christensen and Vesterstrom (1979), Grone (1990) and Loewy (1980), Grone (1990) and it was shown that the convex set of all nth order correlation matrices contains an extreme point of rank k if and only if $k(k+1) \leq n$.

In 1998, K. R. Parthasarthy gave a complete description of the set of all extreme points of this convex set and obtaining the above result as a corollary.

Theorem 1 (K. R. Parthasarathy) *Let R be an nth order correlation matrix of rank k. Then $R \in \mathcal{E}_n$ if and only if it admits the representation*

$$\left(\begin{array}{c|c} \Sigma & \Sigma A \\ \hline A'\Sigma & A'\Sigma A \end{array}\right) P^{-1} \tag{6}$$

where P is a permutation matrix of order n, Σ is a nonsingular correlation matrix of order k and A is a $k \times (n-k)$ matrix of the form $A = (\boldsymbol{a}^1, \boldsymbol{a}^2, \ldots, \boldsymbol{a}^{n-k})$, where $\boldsymbol{a}^i; (i = 1, 2, \ldots, (n-k))$ are column vectors of order k satisfying the following:

1. $\boldsymbol{a}^{i'}\Sigma\boldsymbol{a}^i = 1$ *for every i.*
2. *the rank of the set $\left\{\boldsymbol{a}^i \odot \boldsymbol{a}^i = 1, i = 1, 2, \ldots, (n-k)\right\}$ is $\binom{a}{b}$ where $\odot$ is defined as*

$$\boldsymbol{a} \odot \boldsymbol{a} = (a_1a_2, a_1a_3, \ldots, a_1a_k, a_2, a_3, \ldots, a_{k-1}a_k) \tag{7}$$

3 Spin Systems and Spin Correlation Matrices

3.1 Spin System

Let $\Omega_n = \{-1, 1\}^n$ be the space of n-length sequences denoted by $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$, where $\sigma_i = \pm 1$. Each of σ_i is called a spin (or a spin random variable) and we will refer to this set as a spin system.

Suppose we are (possibly partially) given the spin-spin correlations $c_{ij} := E(\sigma_i \sigma_j)$ of the above spin system (we assume $E(\sigma_i) \equiv 0$). Under what conditions a distribution with those correlations exists, and how to determine it?

Let's try to understand how the spin correlation matrices look like.

3.2 Spin Correlations

What are the necessary and sufficient conditions for a correlation matrix of order $n \geq 2$ to be the correlation matrix of spin variables in the classical sense?

Bell's inequalities: The well-known Bell's inequalities are:

$$1 + \varepsilon_i \varepsilon_j c_{ij} + \varepsilon_j \varepsilon_k c_{jk} + \varepsilon_k \varepsilon_i c_{ki} \geq 0 \qquad \forall\, i < j < k \leq n \tag{8}$$

where $\varepsilon_j \, \epsilon \, \{-1, 1\}$.
It is known that the Bells inequalities are necessary and sufficient for a correlation matrix of order ≤ 4 to be the correlation matrix of spin random variables in the classical sense. (Balasubramanian, Gupta and Parthasarathy, 1998)

Convex Polyhedra: Every convex polyhedron has two representations, one as the intersection of finite half-spaces and the other as Minkowski sum of the convex hull of finite points and the nonnegative hull of finite directions. These are called H-representation and V-representation, respectively.

We will now see that a complete V-representation description for the space of covariance matrices of spin systems is known.

Theorem 2 (J. C. Gupta, 1999) *The class of realizable correlation matrices is given by*

$$C_n = \text{ConvexHull}\,\{\Sigma^T : T \in \mathcal{T}\} \tag{9}$$

where Σ^T *and* $\mathcal{T}$ *are defined as follows:*

$$\mathcal{T} = \{T \subset \{1, 2, \ldots, n\} : 1 \in T\} \tag{10}$$

and

$$\Sigma^T = ((c_{ij}^T)) \tag{11}$$

where $c_{ii}^T = 1$ *for all* i *and* $c_{ij}^T = (-1)^{|T \cap \{i,j\}|}$ *for* $i \neq j$

Proof. On Ω_n, for $T \subset \{1, 2, \ldots, n\}$ denote by σ^T the sequence/configuration σ for which $\sigma_i = -1$ for all $i \in T$ and $\sigma_j = 1$ for all $j \in T'$. Define a probability as follows:

$$P^T\Big(\{\sigma^T\}\Big) = P^T\Big(\Big\{\sigma^{T'}\Big\}\Big) = \frac{1}{2} \tag{12}$$

Then, $\alpha_T \Sigma^T$ (where $\alpha_T \geq 0$ and $\sum \alpha_T = 1$) is the correlation matrix of the random variables σ_i with probability $P = \sum \alpha_T \Sigma^T$.

Conversely, if Δ is a correlation matrix of n spins ξ_i on $(\Omega, \mathcal{F}, Q)$, we can define a probability P on Ω_n in terms of Q such that $P = \sum_T \alpha_T P^T$ and under P, σ_i are spin variable with correlation matrix Δ such that $\Delta = \sum \alpha_T \Sigma^T \in C_n$ with $\alpha_T = P\Big(\{\sigma^T\} + \Big\{\sigma^{T'}\Big\}\Big)$. □

An example: Lets take n = 3. Then,

$$\mathcal{T} = \{\{1\}, \{1,2\}\{1,3\}\{1,2,3\}\} \tag{13}$$

the corresponding correlation matrices that are the extremals of the space of spin correlation matrices of order n are

$$\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \tag{14}$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \tag{15}$$

The corresponding Bell's inequalities are

$$1 + c_{12} + c_{23} + c_{13} \geq 0 \tag{16}$$

$$1 - c_{12} - c_{23} + c_{13} \geq 0 \tag{17}$$

$$1 - c_{12} + c_{23} - c_{13} \geq 0 \tag{18}$$

$$1 + c_{12} - c_{23} - c_{13} \geq 0 \tag{19}$$

Since we know that the Bell's inequalities are necessary and sufficient in this case, we have both H and V representation of the space of correlation matrices of order 3.

How do we find a complete set of inequalities that are necessary and sufficient for a correlation matrix to be a correlation matrix of a spin system? In other words, what is the H-representation of the space of correlation matrices of spin systems? We try to find this complete set of inequalities by solving the covariance realization problem using the maximum entropy method.

3.3 Maximum Entropy Solution for Spin Systems

We want to find a density function for the spin random variables such that it maximizes the entropy where the entropy is defined as:

$$\mathcal{S}(\mathcal{P}(\sigma)) = -\sum_{\sigma} \mathcal{P}(\sigma) \ln \mathcal{P}(\sigma) \tag{20}$$

We have the following constraints:

$$\sum_{\sigma} \sigma_h \sigma_k \mathcal{P}(\sigma) = c_{hk} \qquad \text{and} \qquad \sum_{\sigma} \mathcal{P}(\sigma) = 1 \tag{21}$$

Maximum Entropy Solution:

We use the method of Lagrange multipliers to solve the given optimization problem. As expected, we get a probability measure of the exponential family.

$$\mathcal{P}^*(\sigma) = \frac{1}{Z} \exp\left\{ \sum_{h,k} \lambda_{hk} \sigma_h \sigma_k \right\} \tag{22}$$

where Z is the normalization constant (it depends on λ where λ denotes the matrix $((\lambda_{ij}))$) and λ_{hk} are the Lagrange Multipliers.

Dual Problem:

The dual functional has the form:

$$\mathcal{J}(\lambda) = -\sum_{h,k} \lambda_{hk} c_{hk} + \ln \left[\sum_{\sigma} \exp\left\{ \sum_{h,k} \lambda_{hk} \sigma_h \sigma_k \right\} \right] \tag{23}$$

Note that, $\nabla \mathcal{J} = 0$ implies $\mathcal{P}^*(\sigma)$ satisfies the constraints. A critical point exists if $\mathcal{J}$ is proper, which means,

$$\lim_{|\lambda| \to \infty} \mathcal{J}(\lambda) = +\infty \tag{24}$$

It is clear that the following set of inequalities ensure the properness of $\mathcal{J}$:

$$\sum_{\{i,j\}} c_{ij} \lambda_{ij} \leq \max \left\{ \sum_{\{i,j\} \in E} \lambda_{ij} \sigma_i \sigma_j : \sigma \in \Omega \right\} \qquad \text{for every } \lambda \tag{25}$$

We denote this set of inequalities by Δ_n. We have the following theorem:

Theorem 3 *For E as above and $\lambda \in \mathbb{R}^E$, define*

$$M(\lambda) := \max \left\{ \sum_{\{i,j\} \in E} \lambda_{ij} \sigma_i \sigma_j : \sigma \in \Omega \right\} \tag{26}$$

and

$$\Delta_n := \left\{ C = ((c_{ij})) : \sum_{\{i,j\}} c_{ij}\lambda_{ij} \leq M(\lambda) \quad \text{for every } \lambda \right\} \tag{27}$$

Then,

$$\Delta_n = \mathrm{Cov}_n \tag{28}$$

We are not giving the proof of this theorem here but let us look a little closely at these inequalities. Given the closed form as above, it is not clear how these inequalities actually look. We are going to show here that for a suitable choice of λ_{ij}, these inequalities yield the inequalities that we are already familiar with.

(a) **Positivity:** Let $x \in \mathbb{R}^n$. Set $\lambda_{ij} = -x_i x_j$. Then,

$$\sum_{ij} \lambda_{ij}\sigma_i\sigma_j = -\frac{1}{2}\Big[\sum_i x_i\sigma_i\Big]^2 \leq 0 \tag{29}$$

for every $\sigma \in \Omega$. So, for $C \in \Delta_n$, $-\sum_{ij} x_i x_j c_{ij} \leq 0$ which implies positivity.

(b) **Bell's inequalities:** For $A \subset \{1, 2, \ldots, n\}$ such that $|A| = 3$ Let $\lambda_{ij} = -\varepsilon_i\varepsilon_j$ for $i, j \in A, i \neq j$ (0 otherwise). If $A = \{u, v, w\}$, then by re-writing $\sum_{ij} \lambda_{ij}\sigma_i\sigma_j$ we get that $\sum_{ij} \lambda_{ij}\sigma_i\sigma_j \leq 1$. This means, $M(\lambda) \leq 1$ and so for $C \in \Delta_n$ we get,

$$-\varepsilon_u\varepsilon_v c_{uv} - \varepsilon_v\varepsilon_w c_{vw} - \varepsilon_u\varepsilon_w c_{uw} \leq 1 \tag{30}$$

which are precisely the Bell's inequalities.

Δ_n is the complete set of inequalities that guarantee feasibility of existence of the maximum entropy measure. One would like to obtain a minimal set of inequalities that guarantee necessity and sufficiency. For that, we make special choice for λ_{ij} to get some interesting inequalities that we call the generalized Bell's inequalities.

$$\frac{|T| - 1}{2} + \sum_{i \neq j \in T} \varepsilon_i\varepsilon_j c_{ij} \geq 0 \tag{31}$$

where $T \subset \{1, 2, \ldots, n\}$ such that $|T|$ is odd.
These generalized Bell's inequalities are necessary and sufficient for $n = 5$. There are some generalizations of Bell's inequalities that turn out to be necessary and sufficient for the case $n = 6, 7$ as well.

4 Explicit Calculations

4.1 Setting

Given a vector of moments (spin covariances in this case), how does one find probability measure or equivalently the vector of parameters λ (Lagrange multipliers) that realizes the given covariances (if at all)?

As before, Ω_n is the set $\{-1,1\}^n$. Let $\sigma \in \Omega_n$, we call $\sigma = (\sigma_1, \ldots, \sigma_n)$ a configuration. Note that $|\Omega_n| = 2^n$.

Denote by $\mathcal{C}$ the space of real-valued (scalar) functions on Ω_n. We denote by $\mathcal{G}$ a $(N+1)$-dimensional subspace of $\mathcal{C}$ generated by a set of basis functions $(g_0, g_1, \ldots, g_N)$ (where $g_0 \equiv 1$). Define,

$$\mathfrak{G} = \left\{G(\sigma) = [g_0(\sigma), g_1(\sigma), \ldots, g_N(\sigma)]^T : \sigma \in \Omega_n\right\} \in \mathbb{R}^{N+1} \tag{32}$$

and

$$\mathfrak{R}(\mathfrak{G}) = \{R : R = \sum_{\Omega_n} G(\sigma)\mu(\sigma) \text{ where } \mu \text{ is a probability measure on } \Omega_n\} \tag{33}$$

4.2 Main Theorem

Theorem 4 *Consider $R_1 \in \mathbb{R}^N$. The differential equation -*

$$\frac{d\lambda(t)}{dt} = -M(\lambda)^{-1}\left(R_1 - \sum_{\Omega_n} G(\sigma)\frac{1}{Z(\lambda)}e^{-\langle\lambda, G(\sigma)\rangle}\right) \tag{34}$$

where,

$$M(\lambda) = \sum_{\Omega_n} G(\sigma)\frac{1}{Z(\lambda)}e^{-\langle\lambda, G(\sigma)\rangle}G(\sigma)^T \tag{35}$$

and $\lambda(0) = \lambda_0$. Note that $G(\sigma)^T$ denotes the transpose of the column vector $G(\sigma)$. If $R_1 \in$ int $(\mathfrak{R}(\mathfrak{G}))$, then as $t \longrightarrow \infty$ the solution $\lambda(t)$ of the differential equation tends to a limit λ_1 that satisfies

$$R_1 = \sum_{\Omega_n} G(\sigma)\frac{1}{Z(\lambda)}e^{-\langle\lambda_1, G(\sigma)\rangle} \tag{36}$$

Conversely, if $R_1 \notin \text{int}(\mathfrak{R}(\mathfrak{G}))$, then $\|\lambda(t)\| \longrightarrow \infty$.

Remark 3 In the case of covariance realization problem for spins $\mathcal{G}$ is a $(\frac{n(n-1)}{2} + 1)$ - dimensional subspace with basis functions $g_0, g_{1,2}, g_{1,3}, \ldots g_{2,3}, g_{2,4}, \ldots g_{n-1,n}$ defined as:

$$g_{i,j}(\sigma) = \sigma_i\sigma_j \tag{37}$$

The idea is to solve the differential equation

$$\frac{d\lambda(t)}{dt} = -M(\lambda)^{-1}\left(R_1 - \sum_{\Omega_n} G(\sigma)\frac{1}{Z(\lambda)}e^{-\langle\lambda, G(\sigma)\rangle}\right); \tag{38}$$

the solution λ_1 gives the required probability measure. We look at the corresponding Euler scheme to solve the above differential equation:

$$\lambda_{n+1} = \lambda_n - M(\lambda_n)^{-1}\left(R_1 - \sum_{\Omega_n} G(\sigma)\frac{1}{Z(\lambda_n)}e^{-\langle\lambda_n, G(\sigma)\rangle}\right). \tag{39}$$

Thus, one can explicitly construct the maximum entropy measure for the given covariance problem.

References

[1] K. R. Parthasarathy, *On Extremal Correlations.* (1998).

[2] Balasubramanian K., Gupta J. C. and K. R. Parthasarathy, *Remarks On Bell's Inequality For Spin Correlations.* (1998).

[3] Gupta J. C., *Characterization Of Correlation Matrices Of Spin Variables.* (1999).

[4] Tryphon T. Georgiou, *Solution Of The General Moment Problem Via a One-Parameter Imbedding.* (2005).

Harmonic functions in a domain with a small hole

Paolo Musolino (*)

Abstract. The asymptotic behaviour of the solutions of boundary value problems in domains with small holes has been largely investigated by many authors with different approaches. In this note, we consider a Dirichlet problem for the Laplace operator in a bounded domain of $\mathbb{R}^n$ where we make a small perforation, and we see what happens to the solutions when the hole collapses to a point. (Keywords: Singularly perturbed perforated domains; harmonic functions; real analytic continuation in Banach space. 2000 MSC: 31B05; 31B10; 35B25; 35C20; 35J25.)

1 Introduction

In this note, we investigate the asymptotic behaviour of the solutions of a Dirichlet problem for the Laplace operator in a bounded domain with a small hole. Some of the results presented here have been proved by Lanza de Cristoforis [14]. Others, instead, can be deduced by a general result by Dalla Riva and the author [7], concerning certain real analytic families of harmonic functions in a domain with a small hole. In this note, such a family will consist of the solutions of a Dirichlet problem in a perforated domain with a hole whose size depends on a parameter ϵ.

We fix once for all

$$n \in \mathbb{N} \setminus \{0,1,2\}\,, \qquad \alpha \in]0,1[\,.$$

Then we fix two sets Ω^i and Ω^o in the n-dimensional Euclidean space $\mathbb{R}^n$. We assume that Ω^i and Ω^o satisfy the following condition.

$$\begin{aligned} &\Omega^i \text{ and } \Omega^o \text{ are open bounded connected subsets of } \mathbb{R}^n \text{ of} \\ &\text{class } C^{1,\alpha} \text{ such that } \mathbb{R}^n \setminus \operatorname{cl}\Omega^i \text{ and } \mathbb{R}^n \setminus \operatorname{cl}\Omega^o \text{ are connected,} \\ &\text{and such that the origin 0 of } \mathbb{R}^n \text{ belongs both to } \Omega^i \text{ and } \Omega^o. \end{aligned} \tag{1.1}$$

Here $\operatorname{cl}\Omega$ denotes the closure of Ω for all $\Omega \subseteq \mathbb{R}^n$. For the definition of functions and sets of the usual Schauder classes $C^{0,\alpha}$ and $C^{1,\alpha}$, we refer for example to Gilbarg and

(*) Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy.
Current address: Institut de Recherche Mathématiques de Rennes (IRMAR), Université de Rennes 1, 263 Avenue du General Leclerc, 35042 Rennes, France. E-mail: paolo.musolino@univ-rennes1.fr.
Seminar held on May 16th, 2012.

Trudinger [10, §6.2]. The set Ω^o will represent the unperturbed domain where we shall make a hole. The set Ω^i, instead, will represent, in a sense, the shape of the perforation. Here, the letter 'i' stands for 'inner domain' and the letter 'o' stands for 'outer domain'.

We note that condition (1.1) implies that Ω^i and Ω^o have no holes and that there exists a real number ϵ_0 such that

$$\epsilon_0 > 0 \text{ and } \epsilon \mathrm{cl}\,\Omega^i \subseteq \Omega^o \text{ for all } \epsilon \in]-\epsilon_0, \epsilon_0[.$$

We can now introduce the perforated domain $\Omega(\epsilon)$ by setting

$$\Omega(\epsilon) \equiv \Omega^o \setminus (\epsilon \mathrm{cl}\,\Omega^i) \qquad \forall \epsilon \in]-\epsilon_0, \epsilon_0[\,.$$

In other words, the set $\Omega(\epsilon)$ is obtained by removing from Ω^o the closure of the set $\epsilon\Omega^i$, which can be seen as a hole.

A simple topological argument shows that $\Omega(\epsilon)$ is an open bounded connected subset of $\mathbb{R}^n$ of class $C^{1,\alpha}$ for all $\epsilon \in]-\epsilon_0, \epsilon_0[\setminus\{0\}$. Moreover, the boundary $\partial\Omega(\epsilon)$ of $\Omega(\epsilon)$ has exactly the two connected components $\partial\Omega^o$ and $\epsilon\partial\Omega^i$, for all $\epsilon \in]-\epsilon_0, \epsilon_0[$. In particular, $\partial\Omega^o$ is the 'outer boundary' of $\Omega(\epsilon)$, whereas $\epsilon\partial\Omega^i$ is the 'inner boundary'. We also note that $\Omega(0) = \Omega^o \setminus \{0\}$.

For each $\epsilon \in]-\epsilon_0, \epsilon_0[\setminus\{0\}$ we want to consider a Dirichlet problem for the Laplace operator $\Delta \equiv \sum_{j=1}^n \partial^2_{x_j}$ in the perforated domain $\Omega(\epsilon)$. In order to do so, we need to introduce a suitable Dirichlet datum defined on $\partial\Omega(\epsilon)$. So let $f^i \in C^{1,\alpha}(\partial\Omega^i)$ and $f^o \in C^{1,\alpha}(\partial\Omega^o)$. We note that if $\epsilon \in]-\epsilon_0, \epsilon_0[\setminus\{0\}$, then the function f_ϵ from $\partial\Omega(\epsilon)$ to $\mathbb{R}$ defined by

$$f_\epsilon(x) \equiv \begin{cases} f^i(x/\epsilon) & \text{if } x \in \epsilon\partial\Omega^i\,, \\ f^o(x) & \text{if } x \in \partial\Omega^o\,, \end{cases}$$

is in $C^{1,\alpha}(\partial\Omega(\epsilon))$.

Then for each $\epsilon \in]-\epsilon_0, \epsilon_0[\setminus\{0\}$, we consider the following boundary value problem

$$\begin{cases} \Delta u = 0 & \text{in } \Omega(\epsilon)\,, \\ u(x) = f^i(x/\epsilon) & \text{for } x \in \epsilon\partial\Omega^i\,, \\ u(x) = f^o(x) & \text{for } x \in \partial\Omega^o\,. \end{cases} \tag{1.2}$$

As is well known, the problem in (1.2) has a unique solution in $C^{1,\alpha}(\mathrm{cl}\,\Omega(\epsilon))$, and we denote such a solution by u_ϵ. Here, we recall that a function defined on $\mathrm{cl}\,\Omega(\epsilon)$ is in $C^{1,\alpha}(\mathrm{cl}\,\Omega(\epsilon))$ if its derivatives of order less or equal to 1 can be extended with continuity to $\mathrm{cl}\,\Omega(\epsilon)$ and its derivatives of order 1 are Hölder continuous with exponent α. Our aim is to investigate the behaviour of the solution u_ϵ of (1.2) as ϵ tends to 0. We observe that problem (1.2) is clearly singular when $\epsilon = 0$. Indeed, the domain $\Omega(\epsilon)$ is degenerate for $\epsilon = 0$ and also the second condition in (1.2) does not make sense for $\epsilon = 0$.

Therefore, in order to study the behaviour of u_ϵ, we can fix a point which belongs to $\Omega(\epsilon)$ for all ϵ that are close to 0, and see what happens to the value of the solution u_ϵ at this fixed point, as ϵ approaches 0. Also, we can choose to approach to the degenerate value $\epsilon = 0$, for example, from positive values of ϵ. So assume that

$$p \in \Omega^o \setminus \{0\}\,, \tag{1.3}$$

and that

$$\epsilon_p \in]0, \epsilon_0[\text{ is such that } p \in \Omega(\epsilon) \text{ for all } \epsilon \in]0, \epsilon_p[. \tag{1.4}$$

We note that (1.4) implies that the point p belongs to the domain of the function u_ϵ for all $\epsilon \in]0, \epsilon_p[$, and therefore it makes sense to consider $u_\epsilon(p)$ for $\epsilon \in]0, \epsilon_p[$. Thus we can ask the following question.

$$\begin{aligned} &\text{What can be said of the map from }]0, \epsilon_p[\text{ to } \mathbb{R} \\ &\text{which takes } \epsilon \text{ to } u_\epsilon(p) \text{ when } \epsilon \text{ is close to } 0? \end{aligned} \tag{1.5}$$

Questions of this type are not new and have long been investigated for example with the methods of Asymptotic Analysis.

Thus, one could resort to Asymptotic Analysis and may succeed to write out an asymptotic expansion of $u_\epsilon(p)$ in terms of the parameter ϵ. In this sense, we mention, *e.g.*, the work of Ammari and Kang [1, Ch. 5], Ammari, Kang, and Lee [2, Ch. 3], Dauge, Tordeux, and Vial [8], Kozlov, Maz'ya, and Movchan [11], Maz'ya, Nazarov, and Plamenewskij [17], Ozawa [20], Ward and Keller [21].

Here instead we wish to characterize the behaviour of u_ϵ at $\epsilon = 0$ by a different approach. Thus for example, if we consider a certain function relative to the solution (such as, for example, the value of u_ϵ at the point p, as in question (1.5)), we would try to represent it for $\epsilon > 0$ in terms of real analytic functions of the variable ϵ defined on a whole neighbourhood of 0, and of possibly singular at $\epsilon = 0$ but explicitly known functions of ϵ (such as $\log \epsilon$, ϵ^{-1}, *etc.*).

We observe that our approach does have certain advantages. Indeed, if we knew, for example, that $u_\epsilon(p)$ equals for positive values of ϵ a real analytic function of the variable ϵ defined on a whole neighbourhood of 0, then there would exist $\epsilon' \in]0, \epsilon_p[$ and a sequence $\{c_j\}_{j=0}^\infty$ of real numbers such that

$$u_\epsilon(p) = \sum_{j=0}^{\infty} c_j \epsilon^j \qquad \forall \epsilon \in]0, \epsilon'[\,,$$

where the series in the right hand side converges absolutely on $]-\epsilon', \epsilon'[$.

Such a project has been carried out by Lanza de Cristoforis in several papers for linear and nonlinear problems for the Laplace operator in a bounded domain with a small hole (cf., *e.g.*, Lanza [12, 13, 14, 15]). Later, it has been extended to singular perturbation problems for the Lamé equations (see Dalla Riva and Lanza [4, 5, 6]) and for the Stokes system (see Dalla Riva [3]). More recently, also singularly perturbed boundary value problems in an infinite periodically perforated domain have been considered (cf., *e.g.*, [18, 19] and Lanza and the author [16]).

2 What happens when ϵ is positive and close to 0?

In the following theorem, we answer question (1.5) on the behaviour of $u_\epsilon(p)$ as $\epsilon \to 0^+$, by exploiting the functional analytic approach proposed by Lanza de Cristoforis.

Theorem 2.1 [Lanza de Cristoforis [14]] *Let p be as in (1.3). Then there exist ϵ_p as in (1.4) and a real analytic function U_p from $]-\epsilon_p, \epsilon_p[$ to $\mathbb{R}$ such that*

$$u_\epsilon(p) = U_p(\epsilon)\,,$$

for all $\epsilon \in]0, \epsilon_p[$.

By Theorem 2.1, the function which takes $\epsilon \in]0, \epsilon_p[$ to $u_\epsilon(p)$ can be continued real analytically for negative values of ϵ, and accordingly $u_\epsilon(p)$ can be represented for ϵ small and positive in terms of a power series which converges absolutely on a whole neighbourhood of 0. We also note that for the validity of Theorem 2.1 the assumption $n \geq 3$ is crucial. Indeed, if we considered problem (1.2) in the two-dimensional case, then we would have

$$u_\epsilon(p) = U_p^{\#,1}(\epsilon) + \frac{U_p^{\#,2}(\epsilon)}{U_p^{\#,3}(\epsilon) + U_p^{\#,4}(\epsilon)\log\epsilon} \tag{2.2}$$

for ϵ small and positive, where $U_p^{\#,1}$, $U_p^{\#,2}$, $U_p^{\#,3}$, $U_p^{\#,4}$ are real analytic functions defined on a whole neighbourhood of 0 (cf. Lanza [14, Theorem 5.3]). As a consequence, due to the presence of the logarithmic term in the right hand side of (2.2), the left hand side of (2.2), in general, cannot be continued real analytically around $\epsilon = 0$.

Actually, we can go deeper into the analysis of the behaviour of u_ϵ. Namely, instead of considering the behaviour of the value of u_ϵ at a fixed point, as done in Theorem 2.1, we can consider the behaviour of the restriction of u_ϵ to the closure of a suitable open subset of $\Omega^o \setminus \{0\}$, and we can prove the following result.

Theorem 2.3 [Lanza de Cristoforis [14]] *Assume that*

$$\tilde{\Omega} \text{ is a bounded open subset of } \mathbb{R}^n \text{ such that } \operatorname{cl}\tilde{\Omega} \subseteq \Omega^o \setminus \{0\}\,. \tag{2.4}$$

Then there exist $\epsilon_{\tilde{\Omega}}$ such that

$$\epsilon_{\tilde{\Omega}} \in]0, \epsilon_0[\text{ and } \operatorname{cl}\tilde{\Omega} \subseteq \Omega(\epsilon) \text{ for all } \epsilon \in]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[\,, \tag{2.5}$$

and a real analytic operator $U_{\tilde{\Omega}}$ from $]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[$ to $C^0(\operatorname{cl}\tilde{\Omega})$ such that

$$u_\epsilon(x) = U_{\tilde{\Omega}}[\epsilon](x) \qquad \forall x \in \operatorname{cl}\tilde{\Omega}\,, \tag{2.6}$$

for all $\epsilon \in]0, \epsilon_{\tilde{\Omega}}[$.

We note that in Theorem 2.3 $U_{\tilde{\Omega}}$ is a real analytic operator from an open neighbourhood of 0 in $\mathbb{R}$ to the Banach space $C^0(\operatorname{cl}\tilde{\Omega})$ of continuous functions on $\operatorname{cl}\tilde{\Omega}$, endowed, as usual, with the 'supremum' norm. Here we just recall that if $\mathcal{X}$, $\mathcal{Y}$ are (real) Banach spaces, and if F is an operator from an open subset $\mathcal{W}$ of $\mathcal{X}$ to $\mathcal{Y}$, then F is real analytic in $\mathcal{W}$ if for every $x_0 \in \mathcal{W}$ there exist $r > 0$ and continuous symmetric j-linear operators A_j from $\mathcal{X}^j$ to $\mathcal{Y}$ such that $\sum_{j\geq 1} \|A_j\| r^j < \infty$ and $F(x_0 + h) = F(x_0) + \sum_{j\geq 1} A_j(h, \ldots, h)$ for $\|h\|_{\mathcal{X}} \leq r$ (cf., *e.g.*, Deimling [9, p. 150]).

We note that Theorem 2.3 can also be deduced by some more recent results of Dalla Riva and the author [7], concerning real analytic families of harmonic functions (cf. [7,

Proposition 4.1]). Moreover, in Lanza [14, Theorem 5.3] also real analyticity properties of the solution upon perturbations of Ω^o and Ω^i have been proved.

It is worth noting that Theorem 2.1 can be deduced by Theorem 2.3. Indeed, if $p \in \mathrm{cl}\,\tilde{\Omega}$, since the map which takes a function $u \in C^0(\mathrm{cl}\,\tilde{\Omega})$ to $u(p)$ is linear and continuous (and thus real analytic) and since the composition of real analytic operators is real analytic, we deduce that the function from $]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[$ to $\mathbb{R}$ which takes ϵ to $U_{\tilde{\Omega}}[\epsilon](p)$ is real analytic. As a consequence, Theorem 2.1 holds with $\epsilon_p \equiv \epsilon_{\tilde{\Omega}}$ and $U_p(\epsilon) \equiv U_{\tilde{\Omega}}[\epsilon](p)$ for all $\epsilon \in]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[$.

3 What happens for ϵ negative?

We now observe that both u_ϵ and $U_{\tilde{\Omega}}[\epsilon]$ in equality (2.6) are defined also for negative values of ϵ. However, by Theorem 2.3, we just know that the equality in (2.6) holds when ϵ is small and positive. As a consequence, it is natural to ask the following question.

(3.1) Does the equality in (2.6) hold also for ϵ negative?

In Dalla Riva and the author [7], it has been shown that the answer to the question in (3.1) depends on the parity of the dimension n.

The following theorem says that if the dimension n is even ($n = 4, 6, 8, \dots$), then the equality in (2.6) holds also for $\epsilon < 0$ (cf. Dalla Riva and the author [7, Theorem 3.1 and Proposition 4.1]). Moreover, we can introduce a function u_0 in such a way that equality (2.6) holds in a whole neighbourhood of 0, and in particolar also for $\epsilon = 0$.

Theorem 3.2 [Dalla Riva and M. [7]] *Let n be even (and $n \geq 3$). Let $\tilde{\Omega}$, $\epsilon_{\tilde{\Omega}}$ be as in (2.4), (2.5), respectively. Let u_0 be the unique solution in $C^{1,\alpha}(\mathrm{cl}\,\Omega^o)$ of the following Dirichlet problem*

$$\begin{cases} \Delta u = 0 & \text{in } \Omega^o\,, \\ u(x) = f^o(x) & \text{for } x \in \partial\Omega^o\,. \end{cases}$$

Then there exists a real analytic operator $U_{\tilde{\Omega}}$ from $]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[$ to $C^0(\mathrm{cl}\,\tilde{\Omega})$ such that

$$u_\epsilon(x) = U_{\tilde{\Omega}}[\epsilon](x) \qquad \forall x \in \mathrm{cl}\,\tilde{\Omega}\,,$$

for all $\epsilon \in]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[$.

We now turn to consider case n odd. As we shall see, if n is odd ($n = 3, 5, 7, \dots$), then the validity of the equality in (2.6) also for $\epsilon < 0$ has to be considered as a very exceptional situation. Indeed, we have the following theorem (cf. Dalla Riva and the author [7, Proposition 4.3]).

Theorem 3.3 [Dalla Riva and M. [7]] *Let n be odd (and $n \geq 3$). Then the following statements are equivalent.*

(i) *There exist $\tilde{\Omega}$, $\epsilon_{\tilde{\Omega}}$ as in (2.4), (2.5), respectively, and a real analytic operator $U_{\tilde{\Omega}}$ from $]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[$ to $C^0(\mathrm{cl}\,\tilde{\Omega})$ such that*

$$u_\epsilon(x) = U_{\tilde{\Omega}}[\epsilon](x) \qquad \forall x \in \mathrm{cl}\,\tilde{\Omega}\,,$$

for all $\epsilon \in]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[\setminus\{0\}$.

(ii) *There exists* $c \in \mathbb{R}$ *such that*

$$f^i(x) = c \quad \forall x \in \partial\Omega^i\,, \qquad f^o(x) = c \quad \forall x \in \partial\Omega^o$$

(so that $u_\epsilon(x) = c$ *for all* $x \in \mathrm{cl}\,\Omega(\epsilon)$ *and* $\epsilon \in]-\epsilon_0, \epsilon_0[\setminus\{0\}$*).*

Clearly, if statement (ii) of Theorem 3.3 holds and $\tilde{\Omega}$, $\epsilon_{\tilde{\Omega}}$ are as in (2.4), (2.5), respectively, then the operator $U_{\tilde{\Omega}}$ from $]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[$ to $C^0(\mathrm{cl}\,\tilde{\Omega})$ defined by

$$U_{\tilde{\Omega}}[\epsilon](x) = c \qquad \forall x \in \mathrm{cl}\,\tilde{\Omega}\,, \quad \forall \epsilon \in]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[\,,$$

is such that the equality in (2.6) holds for $\epsilon \in]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[\setminus\{0\}$, and therefore we deduce the validity of statement (i). On the other hand, Theorem 3.3 says in particular that if there exists at least one open subset $\tilde{\Omega}$ as in (2.4) for which we can find a small positive number $\epsilon_{\tilde{\Omega}}$ as in (2.5) and a real analytic operator $U_{\tilde{\Omega}}$ from $]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[$ to $C^0(\mathrm{cl}\,\Omega)$ such that the equality in (2.6) holds for $\epsilon \in]-\epsilon_{\tilde{\Omega}}, \epsilon_{\tilde{\Omega}}[\setminus\{0\}$, then we are in the very exceptional situation that f^i and f^o are equal to the same constant $c \in \mathbb{R}$ (and that accordingly $u_\epsilon = c$ on $\mathrm{cl}\,\Omega(\epsilon)$ for all $\epsilon \in]-\epsilon_0, \epsilon_0[\setminus\{0\}$). Hence, if n is odd the validity of equality (2.6) also for ϵ negative has to be considered as a very special situation which happens only in the very trivial case in which the functions u_ϵ for $\epsilon \in]-\epsilon_0, \epsilon_0[\setminus\{0\}$ are all equal to the same constant.

References

[1] H. Ammari and H. Kang, "Polarization and moment tensors". Springer, New York, 2007.

[2] H. Ammari, H. Kang, and H. Lee, "Layer potential techniques in spectral analysis". American Mathematical Society, Providence, RI, 2009.

[3] M. Dalla Riva, *Stokes flow in a singularly perturbed exterior domain.* Complex Var. Elliptic Equ., to appear.

[4] M. Dalla Riva and M. Lanza de Cristoforis, *A singularly perturbed nonlinear traction boundary value problem for linearized elastostatics. A functional analytic approach.* Analysis (Munich), **30** (2010), 67–92.

[5] M. Dalla Riva and M. Lanza de Cristoforis, *Microscopically weakly singularly perturbed loads for a nonlinear traction boundary value problem: a functional analytic approach.* Complex Var. Elliptic Equ., **55** (2010), 771–794.

[6] M. Dalla Riva and M. Lanza de Cristoforis, *Weakly singular and microscopically hypersingular load perturbation for a nonlinear traction boundary value problem: a functional analytic approach.* Complex Anal. Oper. Theory, **5** (2011), 811–833.

[7] M. Dalla Riva and P. Musolino, *Real analytic families of harmonic functions in a domain with a small hole*. J. Differential Equations, **252** (2012), 6337–6355.

[8] M. Dauge, S. Tordeux, and G. Vial, *Selfsimilar perturbation near a corner: matching versus multiscale expansions for a model problem*. In "Around the research of Vladimir Maz'ya. II", Int. Math. Ser. (N. Y.), 12, Springer, New York (2010), 95–134.

[9] K. Deimling, "Nonlinear functional analysis". Springer-Verlag, Berlin, 1985.

[10] D. Gilbarg and N.S. Trudinger, "Elliptic partial differential equations of second order". Springer Verlag, Berlin, 1983.

[11] V. Kozlov, V. Maz'ya, and A. Movchan, "Asymptotic analysis of fields in multi-structures". The Clarendon Press Oxford University Press, New York, 1999.

[12] M. 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** (2002), 1–27.

[13] M. Lanza de Cristoforis, *Asymptotic behavior of the solutions of a nonlinear Robin problem for the Laplace operator in a domain with a small hole: a functional analytic approach*. Complex Var. Elliptic Equ., **52** (2007), 945–977.

[14] M. Lanza de Cristoforise, *Asymptotic behavior of the solutions of the Dirichlet problem for the Laplace operator in a domain with a small hole. A functional analytic approach*. Analysis (Munich), **28** (2008), 63–93.

[15] M. Lanza de Cristoforis, *Simple Neumann eigenvalues for the Laplace operator in a domain with a small hole. A functional analytic approach*. Rev. Mat. Complut. **25** (2012), 369–412.

[16] M. Lanza de Cristoforis and P. Musolino, *A singularly perturbed nonlinear Robin problem in a periodically perforated domain: a functional analytic approach*. Complex Var. Elliptic Equ., to appear.

[17] V. Maz'ya, S. Nazarov, and B. Plamenevskij, "Asymptotic theory of elliptic boundary value problems in singularly perturbed domains. Vols. I, II". Birkhäuser Verlag, Basel, 2000.

[18] P. Musolino, *A singularly perturbed Dirichlet problem for the Laplace operator in a periodically perforated domain. A functional analytic approach*. Math. Methods Appl. Sci., **35** (2012), 334–349.

[19] P. Musolino, *A singularly perturbed Dirichlet problem for the Poisson equation in a periodically perforated domain. A functional analytic approach*. In A. Almeida, L. Castro, F.-O. Speck, editors, "Advances in Harmonic Analysis and Operator Theory, The Stefan Samko Anniversary Volume", Operator Theory: Advances and Applications, Birkhäuser Verlag, Basel, to appear.

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

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

Coherent states approximation and semiclassical limit

Simone Vazzoler (*)

Abstract. The problem of understanding the relation between Classical and Quantum Mechanics has a long history which dates back to Schrödinger in 1926. After an introductory part on Classical and Quantum Mechanics, we start giving the definition of Coherent State and consider the problem of approximating the exact solution to the Schrodinger equation using this particular class of wave functions. In the last part we discuss about the error term of this approximation and introduce the so called Ehrenfest time.

1 Classical and Quantum Mechanics

Classical Mechanics originated in 1687 with Isaac Newton's laws of motion in Principia Mathematica. It models real world objects as *point particles*, that is objects whose size is negligible. The state of the point particle P is then described as a point (q, p) moving in the phase space T^*M (the cotangent space of M). Here and in the following $q \in M$ will represent the position of P while $p \in T^*_q M$ will be the momentum of the particle. To keep things simple we will consider $M = \mathbb{R}^n$ and, consequently, $T^*M = \mathbb{R}^{2n}$.The time evolution of the particle is then described by the Hamiltonian function $h(q, p)$ and by Hamilton's equations

$$\begin{cases} \dot{q}(t) = \partial_p h(q,p) \\ q(0) = q \\ \dot{p}(t) = -\partial_q h(q,p) \\ p(0) = p \end{cases} \tag{1.1}$$

a first order system of ODE in $\mathbb{R}^{2n}$. We define the flow of the Hamiltonian vector field as $\varphi_h^t(q,p) \stackrel{\text{def}}{=} (q(t), p(t))$ (i.e. the solution of (1.1)) and for sake of simplicity we will always consider $h(q,p) = \frac{|p|^2}{2} + V(q)$. In Physics a system observable is a property of the state that can be determined by some sequence of physical operations (the *observations*

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: vazzoler@math.unipd.it. Seminar held on June 20th, 2012.

or measurements). In Classical Mechanics an observable is simply a function

$$\begin{aligned} a : T^*M &\to \mathbb{R} \\ (q,p) &\mapsto a(q,p) \end{aligned}$$

and if we define $a_t(q,p) \stackrel{\text{def}}{=} a(\varphi_h^t(q,p))$ then the time evolution of a is given by

$$\dot{a}_t = \frac{d}{dt} a_t(q,p) = -\{a,h\} \tag{1.2}$$

where $\{\cdot,\cdot\}$ are the Poisson brackets

$$\{f,g\} \stackrel{\text{def}}{=} \sum_{i=1}^{n} \partial_p f \partial_q g - \partial_q f \partial_p g \tag{1.3}$$

We want to point out that it is equivalent describing Hamiltonian Mechanics using equation (1.2) or (1.1).

The set of states in Quantum Mechanics is described by an Hilbert space $\mathcal{H}$. In the following we will always consider $\mathcal{H} = L^2(\mathbb{R}^n)$ but generally it is possible to take different spaces. Note that the wave functions (that are the elements of $\mathcal{H}$) are complex-valued functions, so that we identify $L^2(\mathbb{R}^n) = L^2(\mathbb{R}^n, \mathbb{C})$. We will denote the elements of $L^2(\mathbb{R}^n)$ with $\psi(x)$. To be more precise

$$L^2(\mathbb{R}^n) = \left\{ \psi : \mathbb{R}^n \to \mathbb{C} : \int_{\mathbb{R}^n} \psi^*(x)\psi(x)dx < +\infty \right\}$$

where ψ^* is the complex conjugation. We will always consider only the pure states, that means the states ψ such that $\|\psi\|_{L^2} = 1$. In this case $|\psi(x)|^2$ is the probability distribution of the particle.

The time evolution of the quantum states is described by Schrödinger equation

$$\begin{cases} \mathrm{i}\varepsilon \partial_t \psi(t,x) = H\psi(t,x) \\ \psi(0,x) = \psi_0(x) \end{cases} \tag{1.4}$$

where $H = -\frac{\varepsilon^2}{2}\Delta + V(x)$ is the operator acting on the state $\psi(t,x)$ representing the energy of the system. When H is not time dependent, a (formal) solution can be obtained as

$$\psi(t,x) \stackrel{\text{def}}{=} U_H(t)\psi_0(x) = e^{-\frac{\mathrm{i}}{\varepsilon}tH}\psi_0(x) \tag{1.5}$$

An observable in QM is represented by a self-adjoint operator A acting on the Hilbert space $L^2(\mathbb{R}^n)$. If we define $A(t) \stackrel{\text{def}}{=} U_H(-t)AU_H(t)$ then its time evolution satifies

$$\dot{A} = \frac{d}{dt}A(t) = -\frac{\mathrm{i}}{\varepsilon}[A(t), H] \tag{1.6}$$

where $[A,B] = AB - BA$ is the commutator. The two (most important) observables in Quantum Mechanics are the position X and the momentum P, where

$$\begin{aligned} X_i\psi(x) &= x_i\psi(x) \\ P_i\psi(x) &= -\mathrm{i}\varepsilon\partial_{x_i}\psi(x) \end{aligned}$$

In 1927 Heisenberg postulated the so called *uncertainty principle*:

$$\langle \Delta X_i \rangle_\psi \langle \Delta P_i \rangle_\psi \geq \frac{\varepsilon}{2}$$

where

$$\langle \Delta A \rangle_\psi = \sqrt{\langle \psi | A^2 \psi \rangle - \langle \psi | A \psi \rangle^2}$$

is the error of the measurement of A (its standard deviation). In Quantum Mechanics, the uncertainty principle asserts a fundamental lower bound on the precision with which the position x and the momentum p, can be simultaneously known. The uncertainty principle is a particular case of Robertson-Schrödinger inequality that states that if A and B are two operators then the following inequality holds true:

$$\langle \Delta A \rangle_\psi^2 \langle \Delta B \rangle_\psi^2 \geq \frac{1}{2} \left(\langle \{A, B\} \rangle_\psi - \langle A \rangle_\psi \langle B \rangle_\psi \right)^2 + \left(\frac{1}{2\mathrm{i}} \langle [A, B] \rangle \right)^2$$

where $\{A, B\} = AB + BA$ is the anti-commutator. To sum up we present in Table 1 the correspondence between Classical and Quantum Mechanics. As one can see there is no "classical version" of uncertainty principle. Only very recently in [4] the author proved that Robertson-Schrödinger inequality appears in Classical Mechanics starting from Gromov's symplectic non-squeezing theorem (also called "the Principle of the Symplectic Camel") and introducing "quantum blobs". The aim of *semiclassical analysis* and in particular of the *semiclassical limit* is to understand how the classical dynamics determine the behaviour of the solutions of Schrödinger equation as $\varepsilon \to 0$

$$\mathrm{i}\varepsilon \partial_t \psi(t, x) = -\frac{\varepsilon^2}{2} \Delta \psi(t, x) + V(x) \psi(t, x)$$

	Classical Mechanics	Quantum Mechanics
States	$(q, p) \in \mathbb{R}^{2n}$	$\psi(t, x) \in L^2(\mathbb{R}^n)$
Time evolution	$(q(t), p(t)) = \varphi_h^t(q, p)$	$\psi(t, x) = U_H(t)\psi(x)$
Observables	$a : \mathbb{R}^{2n} \to \mathbb{R}$	$A : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$
Time evolution	$\dot{a} = -\{a, h\}$	$\dot{A} = -\frac{\mathrm{i}}{\varepsilon}[A, H]$
Uncertainty principle	?	$\langle \Delta X \rangle_\psi \langle \Delta P \rangle_\psi \geq \frac{\varepsilon}{2}$

Table 1: The Classical-Quantum correspondence

2 Coherent States

We start searching for the "most classical" states in QM. These are the states ψ that minimizes Heisenberg inequality

$$\langle \Delta X_i \rangle_\psi \langle \Delta P_i \rangle_\psi = \frac{\varepsilon}{2} \tag{2.1}$$

We call such states coherent states. In literature these states are also called MUST (Minimum Uncertainty STates, e.g. [1]).

Proposition 2.1 *Let $(q,p) \in \mathbb{R}^{2n}$. A coherent state $\psi_{(q,p)}(x)$ has the following form*

$$\psi_{(q,p)}(x) = \left(\frac{1}{\pi\varepsilon}\right)^{n/4} e^{\frac{\mathrm{i}}{\varepsilon}S_{(q,p)}(x)} \prod_{i=1}^{n} e^{-\frac{(x_i-q_i)^2}{2\varepsilon}} \tag{2.2}$$

where $S_{(q,p)}(x)$ satisfies

$$[(\partial_{x_i} S_{(q,p)})^2 * \mathcal{G}](q) = [(\partial_{x_i} S_{(q,p)}) * \mathcal{G}]^2(q) \tag{2.3}$$

with $\mathcal{G}(x) = \left(\frac{1}{\pi\varepsilon}\right)^{n/2} \prod_{i=1}^{n} e^{-\frac{x_i^2}{\varepsilon}}$

It is easy to see that the simplest function $S_{(q,p)}$ that satisfies equation (2.3) is $S_{(q,p)}(x) = \langle p|x-q\rangle$. For this reason in the following a coherent state will be

$$\psi_{(q,p)}(x) = \left(\frac{1}{\pi\varepsilon}\right)^{n/4} e^{\frac{\mathrm{i}}{\varepsilon}\langle p|x-q\rangle} \prod_{i=1}^{n} e^{-\frac{(x_i-q_i)^2}{2\varepsilon}} \tag{2.4}$$

and we will rewrite it as

$$\psi_{(q,p)}(x) = \left(\frac{1}{\pi\varepsilon}\right)^{n/4} e^{\frac{\mathrm{i}}{\varepsilon}\left(\frac{1}{2}\langle x-q|Z(x-q)\rangle + \langle p|x-q\rangle\right)} \tag{2.5}$$

with $Z = \mathrm{i}\mathbb{I}$.
Since in the following our coherent state will evolve in time and will change its shape, we generalize the concept of coherent state with the following definition.

Definition 2.2 Let $a \in \mathbb{R}_{>0}$, $(q,p) \in \mathbb{R}^{2n}$, $\Theta \in \mathbb{R}$ and Z be a complex $n \times n$ matrix. We define a generalized coherent state centered in (q,p) as the following wave function

$$\psi^{a,Z,\Theta}_{(q,p)}(x) = a e^{\frac{\mathrm{i}}{\varepsilon}\left(\frac{1}{2}\langle x-q|Z(x-q)\rangle + \langle p|x-q\rangle + \Theta\right)} \tag{2.6}$$

where $\|a e^{-\frac{1}{2\varepsilon}\langle x-q|\Im Z(x-q)\rangle}\|_{L^2} = 1$.

Some remarks on the last definition:

(i) a generalized coherent state does not satisfy equation (2.1);

(ii) we will have $(q,p) = (q(t),p(t)), Z = Z(t), a = a(t), \Theta = \Theta(t)$;

(iii) if $Z = \mathrm{i}\mathbb{I}$, $\Theta = 0$ and $a = \left(\frac{1}{\pi\varepsilon}\right)^{n/4}$ then

$$\psi^{a,Z,\Theta}_{(q,p)}(x) = \left(\frac{1}{\pi\varepsilon}\right)^{n/4} e^{\frac{\mathrm{i}}{\varepsilon}\left(\frac{1}{2}\langle x-q|Z(x-q)\rangle + \langle p|x-q\rangle\right)} = \psi_{(q,p)}(x)$$

as in equation (2.5).

3 The main theorem

We consider a coherent state

$$\phi_{(q,p)}(x) = \left(\frac{1}{\pi\varepsilon}\right)^{n/4} e^{\frac{\mathrm{i}}{\varepsilon}\left(-\frac{1}{2}\langle x-q|Z(x-q)\rangle + \langle p|x-q\rangle\right)}$$

where $Z = \mathrm{i}\mathbb{I}$. We want to construct an approximate solution of the following Cauchy problem

$$\begin{cases} i\varepsilon\partial_t\phi(t,x) = -\frac{\varepsilon^2}{2}\Delta\phi(t,x) + V(x)\phi(t,x) \\ \phi(0,x) = \phi_{(q,p)}(x) \end{cases} \tag{3.1}$$

The exact solution will be given by $\phi(t,x) = e^{-\frac{\mathrm{i}}{\varepsilon}tH}\phi_{(q,p)}(x)$. Now we present the main theorem of these notes (the statement of this theorem can be found in [5] or in [6]).

Theorem 3.1 *There exist a generalized coherent state $\psi^{a(t),Z(t),\Theta(t)}_{(q(t),p(t))}$ such that*

$$\left\|\phi(t,x) - \psi^{a(t),Z(t),\Theta(t)}_{(q(t),p(t))}(x)\right\|_{L^2} \leq C\varepsilon^{1/2}t|\Im Z(t)|^{-3/2} \tag{3.2}$$

uniformly in x, where $(q(t),p(t))$ is the classical flow $\varphi^t_h(q,p)$, $Z(t)$ is a complex $n\times n$ matrix, $a(t)$ is a real function and the phase factor $\Theta(t)$ is given by

$$\Theta(t) = \int_0^t \Big[p(s)\dot{q}(s) - h(q(s),p(s))\Big]ds \tag{3.3}$$

Proof. We will give only the sketch of the proof.

(1) put $\psi^{a(t),Z(t),\Theta(t)}_{(q(t),p(t))}$ in the equation and use Taylor expansion of the potential $V(x)$ in powers of $(x-q)$: from imposing that order $0,1$ and 2 are all equal to 0 one gets three equation for the parameters

$$\mathrm{i}\varepsilon\frac{\dot{a}}{a} - \dot{\Theta} - \langle p,\dot{q}\rangle = -\frac{\mathrm{i}\varepsilon}{2}\mathrm{tr}\, Z + \frac{|p|^2}{2} + V(q) \tag{3.4}$$

$$-\dot{p} + Z\dot{q} = Zp + V'_x(q) \tag{3.5}$$

$$-\dot{Z} = Z^2 + V''_{x,x}(q) \tag{3.6}$$

(2) eq. (3.5) is easily solved by $(q(t),p(t)) = \varphi^t_h(q,p)$. The real part of eq. (3.4) is solved by $\Theta(t) = \int_0^t \Big[p(s)\dot{q}(s) - h(q(s),p(s))\Big]ds$

(3) Solve eq. (3.6): the solution is given by $Z(t) = \mathcal{S}(t)_*Z := (S_{21}+S_{22}Z)(S_{11}+S_{12}Z)^{-1}$ where $\mathcal{S}(t)$ solves

$$\begin{cases} \dot{\mathcal{S}} = (\mathcal{J}H''_h)\mathcal{S} = \begin{pmatrix} \mathbb{O} & \mathbb{I} \\ -V''_{x,x}(q(t)) & \mathbb{O} \end{pmatrix}\mathcal{S} \\ \mathcal{S}(0) = \mathbb{I} \end{cases} \tag{3.7}$$

and now one can also find a solution for the imaginary part of eq. (3.4), that is

$$a(t) = (\det \Im Z)^{1/4} e^{-\frac{1}{2}\int_0^t \operatorname{tr}[Z(s)]ds}$$

□

The theorem asserts that we can approximate the Quantum evolution using purely Classical objects. In a certain sense Classical Mechanics "emerges" from Quantum Mechanics.

4 Ehrenfest times

The error term is

$$C\varepsilon^{\frac{1}{2}} t |\Im Z(t)|_\infty^{-\frac{3}{2}}$$

and as one can see it depends on the sup-norm of the imaginary part of the matrix $Z(t)$ that is the solution of

$$-\dot{Z} = Z^2 + V''_{x,x}(q)$$

and it is given by $Z(t) = \mathcal{S}(t)_* Z(0)$ where $\mathcal{S}(t)$ is the linearized flow. In the next sections we will show that it depends on the properties of the classical flow.

4.1 Stable/unstable case

Here we will consider Coherent States centered on equilibrium points and we will try to estimate the error.

Example (part 1): the stable equilibrium

We consider

$$H = -\frac{\varepsilon^2}{2}\Delta + \cos x - 1 \tag{4.1}$$

and the "correspondent" classical Hamiltonian $h(q,p) = \frac{p^2}{2} + \cos q - 1$. The flow of h has a stable equilibrium point in $P_s = (\pi, 0)$ (i.e. $\varphi_h^t(\pi, 0) = (\pi, 0)$ for all $t \in \mathbb{R}$). We have to solve the following linearized system around P_s

$$\begin{cases} \dot{\mathcal{S}} = \begin{pmatrix} \dot{S}_{11} & \dot{S}_{12} \\ \dot{S}_{21} & \dot{S}_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \cos\pi & 0 \end{pmatrix} \mathcal{S} = \begin{pmatrix} S_{21} & S_{22} \\ -S_{11} & -S_{12} \end{pmatrix} \\ \mathcal{S}(0) = \mathbb{I} \end{cases} \tag{4.2}$$

It is easy to see that the solution is

$$\mathcal{S}(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \tag{4.3}$$

and consequently the function $Z(t)$ will be

$$\begin{aligned} Z(t) &= (S_{21} + \mathrm{i}S_{22})(S_{11} + \mathrm{i}S_{12})^{-1} \\ &= (\sin t + \mathrm{i}\cos t)(\cos t - \mathrm{i}\sin t)^{-1} \\ &= (\sin t + \mathrm{i}\cos t)(\cos t + \mathrm{i}\sin t) = \mathrm{i} \end{aligned}$$

since $Z_0 = \mathrm{i}$. We have $|\Im Z(t)|_\infty = 1$ that means that in this case the sup-norm of the matrix is less than a constant. This is true in general, as the theorem below shows.

Theorem 4.1 *If (q_s, p_s) is a stable equilibrium point and we consider the Cauchy problem with initial datum $\psi_{(q_s,p_s)}(x)$, then $|\Im Z(t)|_\infty^{-3/2} \leq C$ where $C \in \mathbb{R}$ is a constant. In this case the approximation of the main theorem is valid up to times of order $t \sim 1/\sqrt{\varepsilon}$.*

It is important to notice that the approximation (even in the case of a stable equilibrium) is good only up to times of order $t \sim \frac{1}{\sqrt{\varepsilon}}$.

Example (part 2): the unstable equilibrium

We consider the same example as before but we choose $P_u = (0, 0)$ (the unstable equilibrium point). Then the matrix $\mathcal{S}(t)$ is

$$\mathcal{S}(t) = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix} \tag{4.4}$$

and

$$Z(t) = \tanh 2t + \mathrm{i}\frac{1}{\sinh^2 t + \cosh^2 t} \tag{4.5}$$

We get the following upper bound for the sup-norm of $Z(t)$:

$$|\Im Z(t)|^{-3/2} = (\sinh^2 t + \cosh^2 t)^{3/2} = (\cosh 2t)^{3/2} \leq e^{3t} \tag{4.6}$$

Again this is true in the general case as stated in the following theorem.

Theorem 4.2 *If (q_u, p_u) is an unstable equilibrium point and we consider the Cauchy problem with initial datum $\psi_{(q_u,p_u)}(x)$, then $|\Im Z(t)|^{-3/2} \leq e^{3\lambda t}$ where $\lambda > 0$ is a constant. In this case the approximation of the theorem is valid up to times of order $t \sim \frac{1}{6\lambda} \ln \frac{1}{\varepsilon}$.*

This theorem gives us a different time scale for the validity of our approximation: in the unstable case the limit is $\frac{1}{6\lambda} \ln \frac{1}{\varepsilon}$. These time scales are called Ehrenfest times.

5 Conclusions

We want to know if there are classical properties that "come" from Quantum world. As we have alredy see semiclassical approximation is good only for times smaller than Ehrenfest times but some of these properties (for example ergodicity) appear only when $t \to \infty$. The (open) problem is: is it possible to go beyond Ehrenfest times?

References

[1] S. T. Ali, J. P. Antoine and J. P. Gazeau, "Coherent states, wavelets and their generalizations". Springer, 2000.

[2] Schrödinger, E., *Der stetige Übergang von der Mikro-zur Makromechanik.* Naturwissenschaften 14 (1926), 664–666.

[3] Comberscue, M. and Robert, D., "Coherent States and Applications in Mathematical Physics". Springer, 2012.

[4] De Gosson, M., *The symplectic camel and the uncertainty principle: the tip of an iceberg?* Found. Phys. 39/2 (2009), 194–214.

[5] Paul, T., *Echelles de temps pour l'évolution quantique à petite constante de Planck.* Preprint (2007).

[6] Paul, T., *Semiclassical analysis and sensitivity to initial conditions.* Information and Computation 207/5 (2009), 660–669.