Università di Padova – Dipartimento di Matematica

Scuole di Dottorato in Matematica Pura, Matematica Computazionale e Informatica

Seminario Dottorato 2014/15

Preface

This document offers a large overview of the eight months' schedule of Seminario Dottorato 2014/15. 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 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, July 2nd, 2015

Corrado Marastoni, Tiziano Vargiolu

Abstracts (from Seminario Dottorato's web page)

Wednesday 5 November 2014

The stochastic mesh method to price swing contracts

Matteo BASEI (Padova, Dip. Mat.)

This talk is based on the results achieved during a six-month internship in the Risk Department of a leading energy company. Our goal is twofold: on the one hand we give a brief survey on the problem of pricing swing contracts by the stochastic mesh method, on the other hand we describe our experience in the use of advanced mathematics in a private company. Firstly, we consider the case of American options and study the original formulation of the stochastic mesh method, introduced by Broadie and Glasserman in 1997. Secondly, we try to improve the method by optimally calibrating the parameters, by a literature review and by the use of variance reduction techniques. Finally, we use the revised method to price swing options in energy markets.

Wednesday 19 November 2014

Singular perturbations of stochastic control problems with unbounded fast variables

Joao MEIRELES (Padova, Dip. Mat.)

In this talk, we first give a short introduction to singular perturbations problems and to the Hamilton-Jacobi approach to the singular limit $\epsilon \to 0$. And we will end by considering a specific singular perturbation problem of a class of optimal stochastic control problems with unbounded fast variables and discussing some recents results.

Wednesday 26 November 2014

An introduction to density estimates for diffusions

Paolo PIGATO (Padova, Dip. Mat.)

We recall some notions in Malliavin calculus and some general criteria for the absolute continuity and regularity of the density of a diffusion. We present some estimates for degenerate diffusions under a weak Hormander condition, obtained by starting from the Malliavin and Thalmaier representation formula for the density. As an example, we focus in particular on the stochastic differential equation used to price Asian Options.

Wednesday 17 December 2014

A gentle introduction to semistable degeneration of $K3$ surfaces

Genaro HERNANDEZ MADA (Padova, Dip. Mat.)

In this talk we give the elements to understand the definition and two results about semistable degenerations of $K3$ surfaces over the complex numbers. The first result is a description of the special fiber and the second one is a classification of it in terms of monodromy. If time allows, we shall also introduce the p-adic analogue of these results.

Wednesday 28 January 2015

Shape sensitivity analysis for vibrating plate models

Davide BUOSO (Padova, Dip. Mat.)

In this talk, we consider two different models for the vibration of a clamped plate: the Kirchhoff-Love model, which leads to the well known biharmonc operator, and the Reissner-Mindlin model, which instead gives a system of differential equations. We point out similarities and differences, showing the connections between these two problems. Then we show some results concerning the stability of the spectrum with respect to domain perturbations. After recalling the known results in shape optimization for the biharmomic operator, we state some analyticity results for the dependence of the eigenvalues upon domain perturbations and Hadamard-type formulas for shape derivatives. Using these formulas, we prove that balls are critical domains for the symmetric functions of the eigenvalues under volume constraint.

Wednesday 11 February 2015

Metastability of the Ising model on random graphs at zero temperature

Sander DOMMERS (Bologna, Dip. Mat.)

In this talk I will introduce a random graph model known as the configuration model. After this, I will discuss the Ising model, which is a model from statistical physics where a spin is assigned to each vertex in a graph and these spins tend to align, i.e., take the same value as their neighbors. It is especially interesting to study the Ising model on random graphs. I will discuss some properties of this model. In particular, I will talk about the dynamics and metastability in this model when the interaction strength goes to infinity. This corresponds to the zero temperature limit in physical terms.

Wednesday 25 February 2015

Automorphism-invariant modules

Khanh Tung NGUYEN (Padova, Dip. Mat.)

In this talk, after recalling some basic concepts, we mention the class of injective modules, the class of quasi-injective modules and their generalization, the class of automorphism-invariant modules. Next, we give some results related to the endomorphism rings of automophism-invariant modules and their injective envelopes. Finally, we show a connection between automorphism-invariant modules and bolean rings.

Wednesday 18 March 2015

Introduction to kernel-based methods

Gabriele SANTIN (Padova, Dip. Mat.)

In this talk we give an introduction to kernel-based methods and to their application in different fields of applied mathematics. We consider some examples that motivate the use of kernel-based techniques. Each example can be included in the same framework, but allows to show and discuss different features that arise naturally in the particular application. The examples deal with multivariate scattered data approximation, optimal recovery in Hilbert spaces, numerical solution of PDE, machine learning, and statistics. After building up the fundamental tools of kernel-based methods, we will introduce the problem of the determination of optimal subspaces for kernel-based multivariate approximation. We will give some insight into the problem and discuss possible applications.

Wednesday 1 April 2015

Zooming into p-adic curves

Velibor BOJKOVIC (Padova, Dip. Mat.)

The goal of the seminar is to introduce the audience to the basic notions of Berkovich geometry through a toy example of a p-adic projective curve. After recalling the basic properties of a p-adic field, we motivate Vladimir Berkovich's approach to studying geometry over such fields and go into describing the structure of compact p-adic curves.

Wednesday 15 April 2015

Sobolev spaces, differential operators and multipliers

Aigul MYRZAGALIYEVA (Padova, Dip. Mat. and Eurasian National Univ. Astana)

In this talk, after recalling some basic notions of Sobolev spaces we give some examples, then we introduce differential operators and multipliers in pair of Sobolev spaces. We give the statement and motivation of the problem. Morever, we also present some open problems.

Wednesday 29 April 2015

Preferences in AI

Cristina CORNELIO (Padova, Dip. Mat.)

Artificial Intelligence (AI) is a field that has a long history but still constantly and actively growing and changing. The applications of AI are several, for example web search, speech recognition, face recognition, machine translation, autonomous driving, automatic scheduling etc. These are all complex real-world problems, and the goal of artificial intelligence (AI) is to tackle these with rigorous mathematical tools: machine learning, search, game playing, Markov decision processes, constraint satisfaction, graphical models, and logic. Recently, a new concept became very important in AI: the use of preferences. Let's think about social networks, online shops, systems that suggest music or films. In this talk it is presented an overview on the main applications of preferences in AI, like recommender systems, multi-agent decision making, computational social choice, stable marriage problems, uncertainty in preferences and qualitative preferences.

Wednesday 6 May 2015

Variational methods in nonlinear elasticity: an introduction

Alice FIASCHI (Padova, Dip. Mat.)

After a brief introduction of the variational formulation for the standard model in nonlinear elasticity, we will consider the problem of finding the "right" space to describe the equilibrium configurations of an elastic body, from the point of view of the Calculus of Variations. In this framework, I will introduce the space of Young measures as a suitable space to describe materials exhibiting microstructures.

Wednesday 27 May 2015

Controllability and the numerical approximation of the minimum time function

Thien Thuy LE THI (Padova, Dip. Mat.)

In optimal control theory, minimum time problems are of interest since they appear in many applications such as robotics, automotive, car industry, etc.. The scope of this talk is to give a brief introduction of these problems. Controllability conditions under various settings are considered. Such conditions play a vital role in studying the regularity of the minimum time function $T(x)$. Moreover, we will also introduce the HJB equation associated with a minimum time problem and approaches to computing $T(x)$ approximately.

Wednesday 10 June 2015

An introduction to derived categories

Francesco MATTIELLO (Padova, Dip. Mat.)

Derived categories were introduced in the sixties by Grothendieck and Verdier and have proved to be of fundamental importance in Mathematics. Starting with a short review of the basic language of category theory, we will first introduce the notion of abelian category with the help of several examples. Then we will spend some time giving a thorough motivation for the construction of the derived category of an abelian category. Finally, we will look at a way to break a derived category into two pieces that permit (among other things) to recover the original abelian category.

Wednesday 24 June 2015

Why should people in approximation theory care about (pluri-)potential theory?

Federico PIAZZON (Padova, Dip. Mat.)

We give an introductory summary of results in (pluri-)potential theory that naturally come into play when considering classical approximation theory issues both in one and (very concisely) in several complex variables. We focus on Fekete points and the asymptotic of orthonormal polynomials for certain L^2 counterpart of Fekete measures. No specific knowledge on the topic is assumed.

The stochastic mesh method to price American options and swing contracts

Matteo Basei (*)

Abstract. This talk is based on the results achieved during a six-month internship in the Risk Department of a leading energy company. Our goal is twofold: on the one hand we give a brief survey on the problem of pricing swing contracts by the stochastic mesh method, on the other hand we describe our experience in the use of advanced mathematics in a private company.
Firstly, we consider the case of American options and study the original formulation of the stochastic mesh method, introduced by Broadie and Glasserman in 1997. Secondly, we try to improve the method by optimally calibrating the parameters, by a literature review and by the use of variance reduction techniques. Finally, we outline how to use the method to price swing options in energy markets.

Contents

1 Introduction

We here give the definitions of European and American options and outline some features of energy markets.

European and American options. The following definitions are fundamental in mathematical finance.

- A *European option* is a contract giving the holder the right (not the obligation!) to buy or sell an underlying asset at a prespecified price and on a prespecified date.

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: basei@math.unipd.it . Seminar held on November 5th, 2014.

Example. We have the right to buy one UniPd stock at 10€ on 10th November. On that date, we check the price of one UniPd stock; if the price is less than 10€, we do not exercise the option, if the price is greater than 10€ (say 15€), we do exercise the option (the gain being 15-10=5€).

- An *American option* is similar to a European option, but here the holder can choose, among a set of prespecified dates, when to exercise the right.

  Example. We have the right to buy one UniPd stock at 10€ on 10th, 11th, 12th or 13th November. On those dates, we check the price and decide if exercising or not. The problem of which date to choose is not easy at all and will be sketched below.

Given a function $h : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}$ and an option (either European or American), we say that the option has payoff h if $h(t, x)$ is what the owner would gain when exercising the option at time t and underlying price $S_t = x$. In the previous examples, $h(t, x) = (x-10)^+$, since we exercise if and only if $x \geq 10$ and, in that case, what we gain is $x - 10$.

A fundamental problem is how to compute the fair price of a European or American option with payoff h, where, without entering in technicalities, by fair price we mean that neither the buyer nor the seller can make sure profits. It can be proved that such a price is

- $Q = \mathbb{E}[h(T, S_T)]$ for European options, where T is the maturity;
- $Q = \sup_\tau \mathbb{E}[h(\tau, S_\tau)]$ for American options, where the τ's are stopping times taking value in the set of the possible exercise dates.

For the European price, closed formulas sometimes exist; otherwise, fast numerical schemes can be applied. On the contrary, the computation of the American price is a complicated problem, and several methods have been proposed. We here focus on the stochastic mesh method, introduced by Broadie and Glasserman in [2].

Finally, we remark that much more complicated options are present in the market: multiple exercise, constraints, and so on. This is the case, for example, of swing options.

Swing contracts. The price of energy is subject to remarkable fluctuations, mainly because the markets are influenced by many elements (peaks in consumes, breakdowns in power plants, etc.) and because energy storage is either costly or almost impossible. To hedge against the risk of sudden price rises, several options are traded in the market. In particular, swing contracts give the holder the right to buy energy at an agreed (floating) price, but with some local and global constraints: on the one hand the withdrawal intensity is bounded, on the other hand some final conditions must be satisfied (e.g. lower and upper bounds on the totally bought quantity).

The problem of pricing swing options has been a challenging research subject in the last few years. The difficulties are both theoretical (we deal with a constrained stochastic control problem) and numerical (we usually have dozens of underlyings, so that there is the need of designing algorithms which are efficient even with high-dimensional problems).

Basically, most of the methods proposed in the literature consist in a suitable adaptation of some existing techniques originally meant to price American options. In our case, we will adapt the stochastic mesh method.

Contents. We here mainly focus on the basic theory of the stochastic mesh method. However, we will sketch some research areas related to this subject. In Section 1 we describe the original formulation of the stochastic mesh procedure; in Section 2 we try to improve the method by optimally calibrating the parameters, by considering some changes in the algorithm and by implementing importance sampling techniques; in Section 3 we outline the problem of pricing swing contracts by the stochastic mesh method.

2 Pricing American options by the stochastic mesh method

We first give a precise formulation of the problem of pricing American options, sketched in the Introduction; then, we describe the pricing procedure proposed by Broadie and Glasserman.

Formulation of the problem. We consider a time interval $[t^i, t^f]$ and a filtered probability space, where a d-dimensional Markov process $S = \{S_t\}_{t\in[t^i,t^f]} \in \mathbb{R}^d$ is defined. We assume the initial state of the process to be deterministic and fixed: $S_{t^i} = S_0$, for a given $S_0 \in \mathbb{R}^d$. Let h be a function from $[t^i, t^f] \times \mathbb{R}^d$ to $\mathbb{R}$ representing the payoff, which means that $h(t, x)$ is what the holder gains if he exercises the option at time $t \in [t^i, t^f]$ and with $S_t = x$. Let $r \geq 0$ be the riskless interest rate, here assumed constant. It is known that the price at time $t \in [t^i, t^f]$ and state $x \in \mathbb{R}^d$ of the American option with payoff h and underlying S is

$$Q(t, x) = \sup_{\tau\in\mathcal{T}^{t,tf}} \mathbb{E}[e^{-r(\tau-t)}h(\tau, S_\tau)|S_t = x], \tag{2.1}$$

where $\mathcal{T}^{a,b}$ is the set of the $[a, b]$-valued stopping times. In particular, the initial price of the option is

$$Q(t^i, S_0) = \sup_{\tau\in\mathcal{T}^{t^i,tf}} \mathbb{E}[e^{-r(\tau-t^i)}h(\tau, S_\tau)]. \tag{2.2}$$

Instead of a continuous time interval $[t^i, t^f]$, a discrete time grid

$$t^i = t_0 < t_1 < \cdots < t_{T-1} < t_T = t^f$$

is usually considered, both for contractual issues and for numerical needs. In this case, the initial price of the option is

$$Q(t_0, S_0) = \sup_{\tau\in\mathcal{T}^{0,T}} \mathbb{E}[e^{-r(\tau-t_0)}h(\tau, S_\tau)]. \tag{2.3}$$

where $\mathcal{T}^{0,T}$ is the set of the $\{t_0, \ldots, t_T\}$-valued stopping times. It can be proved that, as $T \to +\infty$, the value in (2.3) converges to the value in (2.2). We remark that, even if we now consider a discrete set of exercise dates, the underlying is a continuous time process. Finally, we will still use the name American option to designate problems as in (2.3), although the name Bermudan option is sometimes used in such a framework.

It is well known that problem (2.3) admits the following dynamic programming representation:

$$Q(t_T, x) = h(t_T, x), \tag{2.4}$$

$$Q(t_i, x) = \max\{h(t_i, x), e^{-r(t_{i+1}-t_i)}\mathbb{E}[Q(t_{i+1}, S_{t_{i+1}})|S_{t_i} = x]\}, \tag{2.5}$$

for $x \in \mathbb{R}^d$ and $i \in \{0, \dots, T-1\}$. The discounted conditional expectation in (2.5), called the *continuation value*, has the following meaning: it is the discounted price of the option at time t_{i+1} if it is not exercised at time t_i. In other words, equation (2.5) is the mathematical formulation of the obvious principle that an American option should be exercised when the payoff is greater than the value one expects to gain if he decides not to exercise immediately.

As a consequence, we also have a formula for the optimal stopping time:

$$\tau^{\text{opt}} = \min\{i \in \{0, \dots, T\} : h(t_i, S_{t_i}) \geq Q(t_{i+1}, S_{t_{i+1}})\}, \tag{2.6}$$

so that

$$Q(t_0, S_0) = \mathbb{E}[e^{-r(\tau^{\text{opt}}-t_0)}h(\tau^{\text{opt}}, S_{\tau^{\text{opt}}})]. \tag{2.7}$$

At this level, formulas (2.6) and (2.7) are not useful from a practical point of view, since they require the knowledge of the price, which is exactly our aim.

To be consistent with [2], from now on we will write, with a small abuse of notation,

$$Q(i, x) = Q(t_i, x), \qquad h(i, x) = h(t_i, x).$$

Thus, the price function Q and the payoff h will take value in $\{0, \dots, T\} \times \mathbb{R}^d$, whereas the stopping times will be $\{0, \dots, T\}$-valued.

Unless the European case, even under simple assumptions (such as the Black-Scholes model) there are no closed formulas for the price of an American option and the problem of approximating (2.3) has been a challenging research topic in the last years. Providing a detailed list of the main methods would be beyond the scope of this notes: we refer the interested reader to the comprehensive book by Glasserman [4] and to the papers [3], [5].

Our aim is to study a particular method: the stochastic mesh (SM from now on) method, introduced by Broadie and Glasserman in [2]. Basically, the authors consider the dynamic programming formulation of the pricing problem and, as a core part of the method, approximate the conditional expectations by weighted sums depending on the density of the underlying. In so doing, they get a high-biased estimator (the mesh estimator), which is combined to a low-biased estimator (the path estimator) to produce a confidence interval. We now describe in detail this procedure.

First step: the mesh estimator. To simplify the notations, we henceforth assume $r = 0$. We first recall the dynamic programming formulation (2.4)-(2.5) of the pricing problem for an American option:

$$Q(T, x) = h(T, x), \tag{2.8}$$

$$Q(t, x) = \max\{h(t, x), \mathbb{E}[Q(t+1, S_{t+1})|S_t = x]\}, \tag{2.9}$$

where $(T,x) \in \{T\} \times \mathbb{R}^d$ and $(t,x) \in \{(0,S_0)\} \cup \{1,\dots,T-1\} \times \mathbb{R}^d$. Recall that we assume S to be a continuous-time Markov process with deterministic initial value $S_0 \in \mathbb{R}^d$.

Let $b \in \mathbb{N}$. The cornerstone of the SM approach is to consider the following net:

we simulate b paths of the underlying
and we forget the trajectory each point comes from.

In this way, we get a mesh (the stochastic mesh naming the method) made up of one deterministic point $X_0(1) = S_0$ at time $t = 0$ and of b stochastic i.i.d. points $X_t(1), \dots, X_t(b)$ at time $t \in \{1,\dots,T\}$. Notice that, by construction, the density of each point at time $t+1$, conditioned to the values of the point(s) at time t, is $g(1,\cdot) = f(0,S_0,\cdot)$ in the case $t = 0$ and $g(t+1,\cdot) = \frac{1}{b}\sum_{k=1}^{b} f\left(t, X_t(k), \cdot\right)$ in the case $t > 1$, where the function f is defined by

$$f(t,x,\cdot) = \text{Density}(S_{t+1} | S_t = x),$$

for each $t \in \{1,\dots,T\}$ and $x \in \mathbb{R}^d$ (for simplicity's sake, we omit to remark the conditioning values in the notations).

We link each point at time t to all the points at time $t+1$ and we tag the arcs with the following weights: $w\left(0, X_0(1), X_1(j)\right) = 1$, with $j \in \{1,\dots,b\}$, and

$$w\left(t, X_t(i), X_{t+1}(j)\right) = \frac{f\left(t, X_t(i), X_{t+1}(j)\right)}{\frac{1}{b}\sum_{k=1}^{b} f\left(t, X_t(k), X_{t+1}(j)\right)},$$

with $t \in \{1,\dots,T\}$ and $i,j \in \{1,\dots,b\}$. We then approximate the continuation value by

$$\mathbb{E}[Q(t+1, S_{t+1}) | S_t = x] \approx \frac{1}{b}\sum_{j=1}^{b} Q(t+1, X_{t+1}(j)) w(t,x,X_{t+1}(j));$$

this leads to estimate the option price $Q(0,S_0)$ by $\hat{Q}(0,S_0)$, where the function $\hat{Q}$ is recursively defined by

$$\begin{aligned}
&\hat{Q}(T, X_T(i)) = h(T, X_T(i)),\\
&\hat{Q}\big(t, X_t(i)\big) = \max\left(h(t, X_t(i)), \frac{1}{b}\sum_{j=1}^{b} \hat{Q}\big(t+1, X_{t+1}(j)\big) w\big(t, X_t(i), X_{t+1}(j)\big) \right),
\end{aligned}$$

with $t \in \{T-1,\dots,0\}$ and $i \in \{1,\dots,b\}$ (but if $t = 0$, then $i = 1$). We call $\hat{Q}(0,S_0)$ the *mesh estimator* of the price. We remark that $\hat{Q}(0,S_0)$ depends on b, even if, for the sake of simplicity, such dependence is not explicit in the notation we use. It can be proved that the mesh estimator is high-biased and convergent to the real price:

Proposition 2.1 *Under technical assumptions:*

- *the mesh estimator is high-biased, i.e.* $\mathbb{E}[\hat{Q}(0,S_0)] \geq Q(0,S_0)$*;*
- *for a suitable* $p > 1$*, the mesh estimator converges in* L^p *to the true price, i.e.* $\|\hat{Q}(0,S_0) - Q(0,S_0)\|_p \to 0$ *as* $b \to \infty$*.*

Since our goal is to provide a brief summary over the SM method, we do not report here all the assumptions and proofs (which are mainly based on induction): we refer the reader to [2].

For future use, we are interested in estimating the price function Q in points outside the mesh. The same idea as above leads to

$$\hat{Q}(T,x) = h(T,x),$$
$$\hat{Q}(t,x) = \max\left(h(t,x), \frac{1}{b}\sum_{j=1}^{b}\hat{Q}\big(t+1, X_{t+1}(j)\big)w\big(t,x,X_{t+1}(j)\big)\right),$$

where $(T,x) \in \{T\}\times\mathbb{R}^n$ and $(t,x) \in \{(0,S_0)\}\cup\{1,\dots,T-1\}\times\mathbb{R}^n$. Notice that the recursion involves all the mesh points from time $t+1$ on. The weights are defined by $w\left(0,S_0,X_1(j)\right)=1$, with $j\in\{1,\dots,b\}$, and

$$w\left(t,x,X_{t+1}(j)\right) = \frac{f\left(t,x,X_{t+1}(j)\right)}{\frac{1}{b}\sum_{k=1}^{b} f\left(t,X_t(k),X_{t+1}(j)\right)},$$

with $t\in\{1,\dots,T\}$ and $j\in\{1,\dots,b\}$. A result similar to Proposition 2.1 holds.

Second step: the path estimator. Since the mesh estimator $\hat{Q}(0,S_0)$ is high-biased, we cannot provide a confidence interval. The idea of Broadie and Glasserman is to look for a low-biased estimator (which is quite easy when a price approximation is already available) and to combine the two estimates.

Recall the characterization provided in (2.3):

$$Q(0,S_0) = \sup_{\tau}\mathbb{E}[h(\tau,S_\tau)].$$

Thus, we get low-biased estimators by simply considering $\hat{q}(0,S_0) = h(\hat{\tau},S_{\hat{\tau}})$, for any stopping time $\hat{\tau}$. How to choose $\hat{\tau}$ so as to assure convergence as $b\to\infty$? Recall that the optimal exercise strategy of an American option is

$$\tau^{\text{opt}} = \min\{t\in\{0,\dots,T\} : h(t,S_t)\geq Q(t,S_t)\}.$$

Of course, we do not know Q: however, we know a convergent estimator of such a function. So, we consider the estimator

$$\hat{q}(0,S_0) = h(\hat{\tau},S_{\hat{\tau}}),$$

where the stopping time $\hat{\tau}$ is defined by

$$\hat{\tau} = \min\{t\in\{0,\dots,T\} : h(t,S_t)\geq \hat{Q}(t,S_t)\},$$

$\hat{Q}$ being the approximating function previously defined. In order to have better estimates, we "fix" $\hat{Q}$ and combine the results of several path estimates: more in detail, we consider the average of the results from $n_p\in\mathbb{N}$ path estimates with respect to the same stochastic mesh (and then the same estimator $\hat{Q}$). The authors suggest to choose $n_p = 10b$; we are going to discuss this choice in the next section.

To sum up, we consider the estimator $\hat{q}(0, S_0)$, called the *path estimator* of the real price, defined by

$$\hat{q}(0, S_0) = \frac{1}{n_p} \sum_{i=1}^{n_p} h(\hat{\tau}^{(i)}, S^{(i)}_{\hat{\tau}^{(i)}}),$$

where the stopping times $\hat{\tau}^{(i)}$ are defined by

$$\hat{\tau}^{(i)} = \min\{t \in \{0, \dots, T\} : h(t, S^{(i)}_t) \geq \hat{Q}(t, S^{(i)}_t)\},$$

the stochastic mesh (and then the estimator $\hat{Q}$) is the same in the n_p simulations and $S^{(i)}$ $(i = 1, \dots, n_p)$ are independent sample paths of the underlying. Notice that this is not a recursive definition. The properties of the path estimator (which depends on b as well, even if not explicitly remarked by the notation we use) are summarized in the next proposition.

Proposition 2.2 *Under technical assumptions,*

- *the path estimator is low-biased, i.e.* $\mathbb{E}[\hat{q}(0, S_0)] \leq Q(0, S_0)$*;*
- *the path estimator is asymptotically unbiased , i.e.* $\mathbb{E}[\hat{q}(0, S_0)] \to Q(0, S_0)$ *as* $b \to \infty$*.*

Third step: the confidence interval. Since we have both a low-biased and a high-biased estimator, we can combine them to get a confidence interval.

First, we consider $N \in \mathbb{N}$ estimates $\hat{Q}^{(i)}(0, S_0)$ of the mesh estimator $(i = 1, \dots, N)$, and average them to obtain

$$\bar{Q}(N) = \frac{1}{N} \sum_{i=1}^{N} \hat{Q}^{(i)}(0, S_0).$$

Then, we consider N estimates $\hat{q}^{(i)}(0, S_0)$ of the path estimator $(i = 1, \dots, N)$, and average them to obtain

$$\bar{q}(N) = \frac{1}{N} \sum_{i=1}^{N} \hat{q}^{(i)}(0, S_0).$$

By jointly considering the averaged estimates, we have the following confidence interval

$$\left[\bar{q}(N) - z_{\alpha/2} \frac{std(\hat{q})}{\sqrt{N}}, \bar{Q}(N) + z_{\alpha/2} \frac{std(\hat{Q})}{\sqrt{N}} \right],$$

where $z_{\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution and $std(\hat{q})$, $std(\hat{Q})$ are the sample standard deviations of $\{\hat{q}^{(i)}(0, S_0)\}_i$, $\{\hat{Q}^{(i)}(0, S_0)\}_i$.

3 Improving the method

In the previous section, we outlined the original stochastic mesh method, as detailed in [2]. When testing this algorithm, it turns out to be quite slow and biased. As a consequence, we have tried to improve the performances of the method, in several ways. We here sketch some of the results.

First idea: parameter calibration. We had the idea that the values of the parameters proposed by Broadie and Glasserman are not the optimal ones. As a first attempt to improve the performance of the SM method, we then considered the problem of optimally calibrating the parameters of the method. Remarkable results can be achieved.

Second idea: literature review. We wondered if some changes can be made at a deeper level, and if the idea in [2] can be enhanced by improving the algorithm. Hence, we have looked in the literature for papers about the SM method.

Some papers drew our attention and we tested the enhancements therein proposed. We do not report the results here; however, we have not noticed any remarkable improvement, so that the original algorithm still remains the best one, in our opinion.

Third idea: variance reduction. Another way to improve the performance of the SM method is to use some variance reduction techniques. In their paper, Broadie and Glasserman focus on control variates; the results are remarkable, and other methods can be tested.

4 Pricing swing options by the stochastic mesh method

In this section we consider the problem of pricing swing contracts by (an adapted formulation of) the stochastic mesh method. We consider swings written on gas, but the framework can be, of course, adapted to other commodities.

Swing contracts: formal definition. Let us consider a finite set of dates $t_0 < t_1 < \cdots < t_T$. As usual, in order to simplify the notations, we will denote them just by the index. We assume the price of the gas to be a continuous-time Markov process P; moreover, let r be the risk-free interest rate.

In every date $t \in \{0, \dots, T\}$, the holder of the option has the right to buy gas at strike price K_t, instead of the market price P_t. Usually K_t consists in a weighted average, computed with respect to a basket of indexes (e.g. the prices of oil and gas in the preceding six months). Some conditions must be satisfied. First of all, the quantity of gas the owner buys at time t, denoted by u_t, must lie in a prespecified interval $[u_t^{\min}, u_t^{\max}]$. Moreover, some global constraints are present: the most common one consists in setting a lower and an upper bound on the total bought quantity $\sum_{s=0}^{T} u_s$, but other conditions can hold too (for example, constraints on the gas bought in every month, and so on). These constraints can be strict or not; in the latter case a penalty must be paid for every unattained condition.

Notice that the owner's exercise strategy is here modeled by a process $u = \{u_t\}_{t\in\{0,\dots,T\}}$. For every t, let $Z_t = \big(\sum_{s\in\{0,\dots,t\}\cap\mathcal{T}_1} u_s, \dots, \sum_{s\in\{0,\dots,t\}\cap\mathcal{T}_n} u_s\big) \in \mathbb{R}^n$ be a vector collecting the energy globally bought, up to time t, with respect to some sub-periods $\mathcal{T}_1, \dots, \mathcal{T}_n \subseteq \{0, \dots, T\}$, as set in the constraints present in the contract. For example, if the owner is asked to fulfill a condition on the whole period (say, one year) and on every semester (hence, two more constraints), Z_t will store the quantity which has been bought until time t in the first semester, in the second semester, in the whole period. We denote by $\mathcal{A}(t, p, z)$ the set of processes $\{u_t, \dots, u_T\}$ satisfying all the constraints (local and global) of the problem, under the conditions $P_t = p$ $Z_t = z$. Finally, we denote by ξ the penalty function: $\xi(t, p, z)$ denotes the penalty that the owner must pay at time t and state $P_t = p, Z_t = z$ because of possible unreached conditions. We can now write the pricing problem. At every time t, the owner's gain or loss is $(P_t - K_t)u_t$, so that the global expected earning with strategy u is

$$\mathbb{E}\left[\sum_{t=0}^{T} e^{-rt}[(P_t - K_t)u_t - \xi(t, P_t, Z_t)]\right].$$

As well known, the price is the supremum of the expected gain with respect to the set of admissible strategies, i.e.

$$Q = \sup_{u\in\mathcal{A}(0,P_0,Z_0)} \mathbb{E}\left[\sum_{t=0}^{T} e^{-rt}[(P_t - K_t)u_t - \xi(t, P_t, Z_t)]\right].$$

Swing contracts: pricing. The stochastic mesh method can be used here, since the price formula can be rewritten by the dynamic programming principle. For example, let $r = 0$ and consider a swing with constant local constraint $u_t \in [u^{\min}, u^{\max}]$, one global constraint over the whole period, $Z_T = \sum_{s=0}^{T} u_s \in [U^{\min}, U^{\max}]$, no penalties. Then, we have

$$Q(T, p, z) = (-\infty)\mathbb{1}_{\mathbb{R}\setminus[Q^{\min},Q^{\max}]}(z),$$
$$Q(t, p, z) = \max_{u\in\mathcal{U}(t,z)} \Big\{(p - K_t)u + \mathbb{E}\big[Q(t + 1, P_{t+1}, z + u)\big|P_t = p\big]\Big\},$$

$\mathcal{U}(t, z) = \big\{u \in [u^{\min}, u^{\max}] : z+u \in [(U^{\min}-(T+1-t)u^{\max})^+, (U^{\max}-(T+1-t)u^{\min})^+]\big\}$. We refer the interested reader to [1] for details.

References

[1] O. Bardou, S. Bouthemy, G. Pagès, *Optimal quantization for the pricing of swing options.* Applied Mathematical Finance 16/1-2 (2009), 183–217.

[2] M. Broadie, P. Glasserman, *A stochastic mesh method for pricing high-dimensional American options.* The Journal of Computational Finance 7/4 (2004), 35–72 (known since 1997 as a working paper: Graduate School of Business, Columbia University, New York, 1997).

[3] M. Broadie, P. Glasserman, *Pricing American-style securities by simulation.* Journal of Economic Dynamics and Control 21 (1997), 1323–1352.

[4] P. Glasserman, "Monte Carlo Methods in Financial Engineering". Springer-Verlag, New York, 2004.

[5] F. A. Longstaff, E. S. Schwartz, *Valuing American options by simulation: a simple least-squares approach.* Review of Financial Studies 14 (2001), 113–147.

Singular perturbations of stochastic control problems with unbounded fast variables

João Meireles (*)

Abstract. In this talk, we first give a short introduction to singular perturbations problems and to the Hamilton-Jacobi approach to the singular limit $\epsilon \to 0$. Then, we consider a specific singular perturbation problem of a class of optimal stochastic control problems with unbounded and controlled fast variables and we discuss (briefly) how to solve it.

1 Statement of the problem

In a classic singular perturbed system (SPS) the state variables evolve along two different time scales: a positive parameter ϵ appears in front of one of the time derivatives affecting its velocity - the equation of the *fast* variables. Our problem is to describe and understand the asymptotic behaviour of a (SPS) as the parameter ϵ vanishes.

A simple model is

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

where

- $x \in \mathbb{R}^n$ and y belongs to $\mathbb{T}^m \simeq \mathbb{R}^m/\mathbb{Z}^m$ (the *flat torus*);
- the functions $a(\cdot)$ are *controls* and they are measurable functions from $[0, \infty)$ to a compact metric space A (and we will denote by $\mathcal{A}$ the set of all these functions);
- f and g are continuous functions in all their variables and Lipchitz-continuous in (x, y) uniformly with respect to the control a.

(S_ϵ) is a good example of a singular perturbed control system. As mentioned before, the role of ϵ is obvious: it splits the state variables in two groups, one, a group of n slow

(*) Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: meireles@math.unipd.it. Seminar held on November 19th, 2014.

variables evolving in a macroscopic time scale, and, another, a group of m fast variables evolving along a microscopic time scale. This is the reason why the study of multiscale problems is an important issue in many applications of engineering, chemistry and physics. Many phenomena can be modelled by a (SPS).

Here we are interested in characterising and analysing the asymptotic behaviour of (S_ϵ) as $\epsilon \to 0$ (if this makes sense). It is expected that passing to the limit as $\epsilon \to 0^+$ in the the initial problem amounts to reducing the dimension of the state space and that the limit dynamics $(\bar{S})$ (if it exists) involves only the slow variables.

2 Some approaches to singular perturbations problems

There are some approaches to address our problem. Here we present the most relevant ones:

- (The Levinson-Tikhonov method) This approach consists in considering, as the natural candidate for the limit, the system obtained by setting $\epsilon = 0$ in (S_ϵ). The result is an ordinary differential equation combined with an algebraic equation. This approach gives the appropriate solution when the stationary points of the fast dynamics are attractive, a condition that may fail to be satisfied.
- (Limit of occupational measures) Other averaging approaches have been proposed by Artstein in the context of invariant measure theory (see [5]), and by Gaitsgory and Leizarowitz, using limit occupational measures (see [11]).
- A PDE method based in the theory of viscosity solutions and of the homogenisation of fully nonlinear PDEs was developed by Alvarez and Bardi in [1,2,3], also [4] for problems with an arbitrary number of scales. This the approach that we will consider in this note.

3 The PDE approach to singular perturbations problems

It is well known that under some regularity conditions the *value function* V^ϵ of (S_ϵ), i.e,

$$V^\epsilon(t,x,y) := \inf\{l(x(s),y(s),a(s))ds + h(x(s),y(s))\},$$

where l and h are given functions and the inf is being taken among all admissible controls, satisfies the Hamilton-Jacobi-Bellman equation

$$(HJB)_\epsilon \qquad \begin{cases} V_t^\epsilon + H\big(x,y,D_xV^\epsilon,\frac{D_yV^\epsilon}{\epsilon}\big) = 0 \\ V^\epsilon(0,x,y) = h(x,y) \end{cases} \tag{1}$$

where D_xV^ϵ, D_yV^ϵ stand for the gradient of V^ϵ with respect to x and y respectively, and

H is the Hamilton-Jacobi-Bellman operator

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

The PDE approach to (SPP) consists in passing to the limit in the PDE $(HJB)_\epsilon$. Under suitable conditions, it is possible to define an effective Hamiltonian $\bar{H}(x,p)$ such that the V^ϵ converges locally uniformly, as $\epsilon\to 0$, to a solution of

$$\begin{cases} V_t + \bar{H}(x, D_x V) = 0 \\ V(0,x,y) = \bar{h}(x). \end{cases} \tag{2}$$

The proof of the existence of such an operator $\bar{H}$, and the analysis of some of its properties, embody a wide line of research, going back to the first works on homogenisation of PDEs (see for example [10]) and, in particular, to the famous unpublished preprint [16] by Lions, Papanicolaou and Varadhan.

In the paper [2], two crucial properties about the convergence of the V^ϵ has been singled out by the authors. One is an *ergodicity* property of the operator H, and therefore $\bar{H}$, and another is a property called *stabilization to a constant* of the pair (H,h) that allows the possibility to define the effective initial datum $\bar{h}$ for the effective Cauchy problem (2). All these properties permit to establish the uniform convergence of V^ϵ to the solution of the effective equation and in some cases to prove that the effective Hamiltonian is the partial differential operator associated to the limit control problem $(\bar{S})$. However, this theory was developed mostly for fast variables restricted to a compact set (almost all in the case of the m-dimensional torus). Nonetheless, in many physical and financial models the a priori knowledge of the boundedness of the fast variables does not appear to be natural according to the empirical data. Very few has been done until now to treat the unboundedness case.

In the papers [7] and [8] the authors present an extension of the methods based on viscosity solutions showed in [1, 2, 3] to singular perturbation problems that have unbounded but uncontrolled fast variables.

4 Singular perturbations with unbounded fast variables

In my thesis, I study singular perturbations of a class of optimal stochastic control problems with finite time horizon and with unbounded and uncontrolled fast variables. The problem I treat is for $t\in[0,T]$ and given $\theta^*>1$ and $\epsilon>0$

minimize in u and ξ: $\quad \mathbb{E}^{x,y}[\int_t^T (l(X_s,Y_s,u_s) + \frac{1}{\theta^*}|\xi_s|^{\theta^*})ds + g(X_T)]$

subject to

$$\begin{cases} dX_s = F(X_s,Y_s,u_s)ds + \sqrt{2}\sigma(X_s,Y_s,u_s)dW_s, & X_{s_0} = x \\ dY_s = -\frac{1}{\epsilon}\xi_s ds + \sqrt{\frac{1}{\epsilon}}\tau(Y_s)dW_s, & Y_{s_0} = y \end{cases} \tag{3}$$

where l is a running cost function satisfying the following coercivity condition

$$-l_0 + l_0^{-1}|y|^\alpha \leq l(x,y,u) \leq l_0(1+|y|^\alpha) \text{ for some } l_0 > 0$$

where $\alpha > 1$, g represents a terminal cost and is continuous, bounded from below, and growing like

$$\exists C_g > 0 \text{ s.t. } g(x) \leq C_g(1+|x|^\alpha),$$

and $X_s \in \mathbb{R}^n$, $Y_s \in \mathbb{R}^m$, u_s is a control taking values in a given compact set U, $\xi = (\xi_s)_{0\leq s\leq T}$ denotes a control process taking its values in $\mathbb{R}^m$, and W_s is a multi-dimensional Brownian motion on some probability space.

Basic assumptions on the drift F and on the diffusion coefficient σ of the slow variables X_s are that they are Lipschitz continuos functions in (x,y) uniformly in u and satisfy the following growth condition at infinity

$$|F| + ||\sigma|| \leq C(1+|x|).$$

On the fast process Y_s we assume that $\tau\tau^T = \mathbb{I}$.

Calling $V^\epsilon(t,x,y)$ the value function of this optimal control problem, i.e.

$$V^\epsilon(t,x,y) = \inf_{u,\xi} \mathbb{E}^{x,y}[\int_t^T (l(X_s,Y_s,u_s) + \frac{1}{\theta^*}|\xi_s|^{\theta^*})ds + g(X_T)],$$

we are interested in the limit V as $\epsilon \to 0$ of V^ϵ and in particular in understanding the PDE satisfied by V. This is a singular perturbation problem for the system above and for the HJB equation associated to it. We treat it by PDE methods and a careful analysis of the associated ergodic stochastic control problem in the whole space $\mathbb{R}^m$.

In fact, in Theorem 10.1 (see [17]) I prove that if $V^\epsilon(t,x,y)$ is a viscosity solution of the HJB equation then the relaxed semilimits (no standard ones!)

$$\underline{V}(t,x) = \liminf_{(\epsilon,t',x')\to(0,t,x)} \inf_{y\in\mathbb{R}^m} V^\epsilon(t',x',y) \tag{4}$$

and

$$\bar{V} = (\sup_R \bar{V}_R)^* \tag{5}$$

(the upper semicontinuous envelope of $\sup_R \bar{V}_R$) where $\bar{V}_R$ is defined as

$$\bar{V}_R(t,x) = \limsup_{(\epsilon,t',x')\to(0,t,x)} \sup_{y\in B_R(0)} V^\epsilon(t',x',y) \tag{6}$$

are, respectively, a supersolution and a subsolution of the effective Cauchy Problem

$$\begin{cases} -V_t + \bar{H}(x, D_xV, D^2_{xx}V) = 0 & \text{in } (0,\infty)\times\mathbb{R}^n \\ V(T,x) = g(x) & \text{in } \mathbb{R}^n. \end{cases} \tag{7}$$

This procedure allow me to prove also that in some cases $V^\epsilon(t,x,y)$ converges locally uniformly, as $\epsilon \to 0$, to the only solution $V(t,x)$ of (7). Moreover, the effective Hamiltonian $\bar{H}(x,p,M)$ is the unique constant λ such that the following ergodic PDE

$$(EP) \qquad \lambda - \frac{1}{2}\Delta\phi(y) + \frac{1}{\theta}|D\phi(y)|^\theta = f(y) \quad \text{in } \mathbb{R}^m,$$

has a solution ϕ bounded from below, $\frac{1}{\theta} + \frac{1}{\theta^*} = 1$ and $f(y) = -H(x,y,p,M,0)$, where H is the Bellman Hamiltonian associated to the slow variables of (3) and its last entry is for the mixed derivatives D_{xy}. Such type of equations appear in utility maximisation problems in mathematical finance and were first studied by Naoyuki Ichihara in [Ichihara, 2012] using probabilistic and analytical arguments.

References

[1] O. Alvarez, M. Bardi, *Viscosity solutions methods for singular perturbations in deterministic and stochastic control.* SIAM J. Control Optim. 40 (2001/02), 1159–1188.

[2] O. Alvarez, M. Bardi, *Singular perturbations of nonlinear degenerate parabolic PDEs: a general convergence result.* Arch. Ration. Mech. Anal. 170 (2003), 17–61.

[3] O. Alvarez, M. Bardi, *Ergodicity, stabilisation, and singular perturbations for Bellman-Isaacs equations.* Mem. Amer. Math. Soc. (2010).

[4] O. Alvarez, M. Bardi, C. Marchi, *Multiscale problems and homogenisation for second-order Hamilton-Jacobi equations.* J. Differential Equations 243 (2007), 349–387.

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

[6] M. Bardi, I. Capuzzo Dolcetta, "Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations". Birkäuser, Boston, 1997.

[7] M. Bardi, A. Cesaroni, *Optimal control with random parameters: a multiscale approach.* J. Control 17/1 (2011), 30–45.

[8] M. Bardi, A. Cesaroni, L. Manca, *Convergence by viscosity methods in multiscale financial models with stochastic volatility.* SIAM J. Financial Math. 1 (2010), 230–265.

[9] V. S. Borkar, V. Gaitsgory, *Singular perturbations in ergodic control of diffusions.* SIAM J. Control Optim. 46 (2007), 1562–1577.

[10] L. C. Evans, *The perturbed test function method for viscosity solutions of nonlinear PDE.* Proc. Roy. Soc. Edinburgh Sect. A 111 (1989), 359–375.

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

[12] N. Ichihara, *Recurrence and transience of optimal feedback processes associated with Bellman equations of ergodic type.* SIAM J. Control Optim. 49/5 (2011), 1938–1960.

[13] N. Ichihara, *Large time asymptotic problems for optimal stochastic control with super linear cost*. Stochastic Processes and their Applications 122 (2012), 1248–1275.

[14] N. Ichihara and S.-J. Sheu, *Large time behaviour of solutions of Hamilton-Jacobi-Bellman equations with quadratic nonlinearity in gradients*. SIAM J. Math. Anal. 45/1 (2013), 279–306.

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

[16] P.-L. Lions, G. Papanicolaou, S. R. S. Varadhan, *Homogenisation of Hamilton-Jacobi equations*. Unpublished (1986).

[17] J. Meireles, "Singular Perturbations and Ergodic Problems for degenerate parabolic Bellman PDEs in $\mathbb{R}^m$ with Unbounded Data". Phd thesis. Università degli Studi di Padova, 2015.

An introduction to density estimates for diffusion processes

Paolo Pigato (*)

Abstract. We recall some notions in Malliavin calculus and some general criteria for the absolute continuity and regularity of the density of a diffusion. We present some estimates for degenerate diffusion processes under a weak Hörmander condition, obtained starting from the Malliavin and Thalmaier representation formula for the density. As an example, we focus in particular on the stochastic differential equation used to price Asian Options.

1 Elements of Malliavin Calculus

Malliavin calculus is an infinite-dimensional differential calculus on the Wiener space. It is useful to investigate regularity properties of solutions of stochastic differential equations, and its most important application is a probabilistic proof of Hörmander's theorem.

A crucial fact in this theory is the integration-by-parts formula, which relates the derivative operator on the Wiener space and the Skorohod extended stochastic integral. A consequence of this is a formula which links existence and regularity of the density of a random variable to the so-called "Malliavin covariance matrix".

We introduce some basic notions, referring to [8]. We consider a probability space $(\Omega, \mathcal{F}, P)$, a Brownian motion $W = (W_t^1, ..., W_t^d)_{t\geq 0}$ and the filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by W. For fixed $T > 0$, we denote with $\mathcal{H}$ the Hilbert space $L^2([0,T], \mathbb{R}^d)$. For $h \in \mathcal{H}$ we introduce this notation for the Itô integral of h: $W(h) = \sum_{j=1}^d \int_0^T h^j(s) dW_s^j$. We denote by $C_p^\infty(\mathbb{R}^n)$ the set of all infinitely continuously differentiable functions $f : \mathbb{R}^n \to \mathbb{R}$ such that f and all of its partial derivatives have polynomial growth. We also denote by $\mathcal{S}$ the class of random variables of the form

$$F = f(W(h_1), ..., W(h_n)),$$

for some $f \in C_p^\infty(\mathbb{R}^n)$, $h_1, ..., h_n$ in $\mathcal{H}$, $n \geq 1$. The Malliavin derivative of $F \in \mathcal{S}$ is defined

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: pigato.p@gmail.com. Seminar held on November 26th, 2014.

as the $\mathcal{H}$ valued random variable given by

$$DF = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(W(h_1), ..., W(h_n))h_i.$$

Remark that this implies $D\int_0^T h_i(s)dW_s = h_i$. Therefore, this expression can be seen as the germ of a chain-rule formula and is therefore a reasonable definition for a "derivative".

The extension of the definition of D on a wider domain requires the introduction of the Sobolev norm of F:

$$\|F\|_{1,p} = [\mathbb{E}|F|^p + \|DF\|_H^p]^{\frac{1}{p}} \quad \text{where} \quad \|DF\|_H = \left(\int_0^T |D_sF|^2 ds\right)^{\frac{1}{2}}.$$

It is possible to prove that D is a closable operator with respect to this norm an take the extension of D in the standard way. We can now define in the obvious way DF for any F in the closure of $\mathcal{S}$ with respect to this norm. Therefore, the domain of D will be the closure of $\mathcal{S}$.

The higher order derivative of F is obtained by iteration. For any $k \in \mathbb{N}$, for a multi-index $\alpha = (\alpha_1, ..., \alpha_k) \in \{1, ..., d\}^k$ and $(s_1, ..., s_k) \in [0, T]^k$, we can define

$$D^{\alpha}_{s_1,...,s_k}F := D^{\alpha_1}_{s_1}...D^{\alpha_k}_{s_k}F.$$

We denote with $|\alpha| = k$ the length of the multi-index. Remark that $D^{\alpha}_{s_1,...,s_k}F$, is a random variable with values in $\mathcal{H}^{\otimes k}$, and its Sobolev norm is defined as

$$\|F\|_{k,p} = [\mathbb{E}|F|^p + \sum_{j=1}^{k} |D^{(j)}F|^p]^{\frac{1}{p}}$$

where

$$|D^{(j)}F|^2 = \left(\sum_{|\alpha|=j} \int_{[0,T]^k} |D^{\alpha}_{s_1,...,s_k}F|^2 ds_1...ds_k\right)^{1/2}.$$

The extension to the closure of $\mathcal{S}$ with respect to this norm is analogous to the first order derivative. We denote by $\mathbb{D}^{k,p}$ the space of the random variables which are k times differentiable in the Malliavin sense in L^p, and $\mathbb{D}^{k,\infty} = \bigcap_{p=1}^{\infty} \mathbb{D}^{k,p}$.

We denote with δ the adjoint operator of D, the so-called Skorohod integral. It is possible to prove that δ coincides with the Ito integral for adapted integrands, and that the following formula holds, for any $F \in \mathbb{D}^{1,2}$ and $u \in Dom(\delta)$ such that $F \in L^2(\Omega, \mathcal{H})$:

$$\delta(Fu) = F\delta(u) - \mathbb{E}\langle DF, u\rangle_H. \tag{1}$$

We consider a random vector $F = (F_1, ..., F_n)$ in the domain of D. We define its *Malliavin covariance matrix* as follows:

$$\gamma_F^{i,j} = \langle DF_i, DF_j\rangle_{\mathcal{H}} = \sum_{k=1}^{d} \int_0^T D^k_s F_i \times D^k_s F_j ds.$$

We say that F is *non-degenerate* if its Malliavin covariance matrix is invertible and

$$\mathbb{E}(|\det \gamma_F|^{-p}) < \infty, \quad \forall p \in \mathbb{N}. \tag{2}$$

We denote with $\hat{\gamma}_F$ the inverse of γ_F. Using (1) it is possible to prove the following integration by parts formula. Let $F \in \mathbb{D}^{2,2}$ be a scalar r.v. such that (2) holds. Then for every $G \in \mathbb{D}^{2,2}$

$$\mathbb{E}[\phi'(F)G] = \mathbb{E}[\phi(F)H(F;G)], \quad \forall \phi \in C_c^\infty(\mathbb{R}) \tag{3}$$

where the Malliavin weights are given by

$$H(F;G) = -G\hat{\gamma}_F \times LF + \langle D(\hat{\gamma}_F G), DF \rangle.$$

Here $L = -\delta \circ D$ is the *Ornstein-Uhlenbeck* operator. This integration by parts formula is really important, because if it holds it tells us that F is absolutely continuous with respect to the Lebesgue measure and allows us to write this representation for the density:

$$p_F(x) = \mathbb{E}[1_{[x,\infty)}(F)H(F;1)].$$

The intuition behind this fact is the following: we first express the density as $p_F(x) = \mathbb{E}[\delta_0(F-x)]$. Then we formally write the Dirac delta in 0 as $\delta_0(y) = \partial_y 1_{[0,\infty)}(y)$, and apply (3)

$$\mathbb{E}[\delta_0(F-x)] = \mathbb{E}[\partial 1_{[0,\infty)}(F-x)] = \mathbb{E}[1_{[x,\infty)}(F)H(F,1)].$$

If higher order integration by parts formula are available, they can be employed to find analogous expressions for the derivatives of the density, iterating the procedure above.

The following multidimensional generalisation has been proved by Malliavin and Thalmaier in [7]:

$$p_F(x) = -\mathbb{E}\left[\nabla \mathcal{Q}_n(F-x)H(F;1)\right],$$

where $\mathcal{Q}_n$ denotes the Poisson kernel on $\mathbb{R}^n$, i.e. the fundamental solution of the Laplace operator $\Delta \mathcal{Q}_n = \delta_0$. This is given by

$$\mathcal{Q}_1(x) = \max(x,0); \quad \mathcal{Q}_2(x) = \mathcal{A}_2^{-1} \ln|x|; \quad \mathcal{Q}_n(x) = -\mathcal{A}_n^{-1}|x|^{2-n},\ n > 2,$$

where $\mathcal{A}_n$ is the area of the unit sphere in $\mathbb{R}^n$.

2 Application to diffusion processes

2.1 Hörmander theorem

The original motivation and the most important application of the integration by parts mentioned in the previous section is a probabilistic proof of the Hörmander theorem.

Let X_t, $t \in [0,T]$ be the solution of the following stochastic differential equation in $\mathbb{R}^n$:

$$X_0 = x, \quad dX_t = \sum_{i=1}^{d} \sigma^i(X_t)dW_t^i + b(X_t)dt,$$

with $W = (W^1, \dots, W^d) \in \mathbb{R}^d$, $X \in \mathbb{R}^n$. We assume infinitely differentiable coefficients with bounded partial derivatives of all orders. We denote $\sigma^0 = b$.

For $f, g : \mathbb{R}^n \to \mathbb{R}^n$ we recall the definition of the directional derivative of f in the direction g as

$$\partial_g f(x) = \sum_{i=1}^{n} g^i(x) \partial_{x_i} f(x).$$

The Lie bracket $[f, g]$ in x is defined as

$$[f, g](x) = \partial_g f(x) - \partial_f g(x).$$

We say that the *Hörmander condition* holds at point x if the vector space spanned by the vector fields

$$\sigma^1, \dots, \sigma^d, \quad [\sigma^i, \sigma^j],\ 0 \le i, j \le d, \quad [\sigma^i, [\sigma^j, \sigma^k]], 0 \le i, j, k \le d, \dots$$

at x is $\mathbb{R}^n$.

Theorem 1 *Assume that the Hörmander condition holds at the initial point x. Then for $t \in [0, T]$ the random vector X_t has a probability distribution that is absolutely continuous with respect to the Lebesgue measure, and the density is infinitely differentiable.*

This result can be viewed as a probabilistic version of Hörmander's theorem on the hypoellipticity of second-order differential operators. We refer to [8] for details. The proof relies on the fact that X_t is non-degenerate (meaning that (2) holds for $F = X_t$) for any $t > 0$, if the Hörmander condition at x is satisfied.

Remark 2 A similar theorem applies for the density $p_{X_t}(y)$ at y if the Hörmander condition holds at y.

2.2 Bounds for the density

If it exists, define $p_t(x, y)$ as the density of X_t in y, with initial condition in x. Malliavin calculus also allows to find quantitative estimates for $p_t(x, y)$. We expect these estimates to be Gaussian, since the SDE satisfied by X_t is a diffusion driven by a Brownian motion.

We now look closer at the Hörmander non-degeneracy assumption. We say that the *Strong Hörmander* condition holds at x if the vector space spanned by the vector fields

$$\sigma^1, \dots, \sigma^d, \quad [\sigma^i, \sigma^j],\ 1 \le i, j \le d, \quad [\sigma^i, [\sigma^j, \sigma^k]], 1 \le i, j, k \le d, \dots$$

at x is $\mathbb{R}^n$. On the other hand, we say that the *Weak Hörmander* condition at x holds if the vector space spanned by the vector fields

$$\sigma^1, \dots, \sigma^d, \quad [\sigma^i, \sigma^j],\ 0 \le i, j \le d, \quad [\sigma^i, [\sigma^j, \sigma^k]], 0 \le i, j, k \le d, \dots$$

at x is $\mathbb{R}^n$ (recall σ^0 is the drift term). Also recall that we say that our diffusion is *elliptic* if the vector space spanned by the vector fields $\sigma^1, \dots, \sigma^d$ is $\mathbb{R}^n$. Estimates of the density are well known under this assumption, but it is a quite demanding one. For instance,

ellipticity implies $n \leq d$, i.e. the dimension of the driving Brownian Motion must be at least the same as the dimension of the diffusion itself. This is not true if we just suppose the strong Hörmander condition. Indeed, we can use the brackets $[\sigma_i, \sigma_j]$, $1 \leq i, j \leq d$ and their iterations to span $\mathbb{R}^n$. Weak Hörmander condition allows us to use also the brackets between diffusion and drift coefficients, and therefore it is the weakest of the three. We have seen in the previous section that this condition is sufficient to prove existence and regularity of the density. But is it enough to find some Gaussian bounds?

The most celebrated work on this topic is a series of three papers by Kusuoka and Stroock in the the eighties. In [5] the authors prove, under the weak Hörmander condition, the following upper bounds:

$$p_t(x,y) \leq \frac{C_0(T)(1+|x|)^{m_0}}{t^{n_0}} \exp\left(-\frac{D_0(T)|y-x|^2}{t}\right)$$
$$D_y^\alpha p_t(x,y) \leq \frac{C_\alpha(T)(1+|x|)^{m_\alpha}}{t^{n_\alpha}} \exp\left(-\frac{D_\alpha(T)|y-x|^2}{t}\right)$$

where all the above constants depend on how many iterated Lie brackets we have to take to span $\mathbb{R}^n$, and on the final time T. Here D_y^α denotes the derivative of order α of $p_t(x,y)$ with respect to y. Lower bounds of this kind, in general, are not available.

Two-sided bounds in terms of a control metric are established in [6] under *strong* Hörmander conditions if the drift is generated by the vector fields of the diffusive part. The standard control metric is defined as follows. For $x, y \in \mathbb{R}^n$ we denote by $C(x,y)$ the set of controls $\psi \in L^2([0,1];\mathbb{R}^n)$ such that the corresponding skeleton $(u_t)_{t\in[0,1]}$ solution of

$$du_t(\psi) = \sum_{j=1}^{d} \sigma_j(u_t(\psi))\psi_t^j dt, \qquad u_0(\psi) = x$$

satisfies $u_1(\psi) = y$. Notice that the drift $b = \sigma_0$ does not appear in the equation of $u_t(\psi)$. We define the control (Caratheodory) distance as

$$d_c(x,y) = \inf\Big\{\Big(\int_0^1 |\psi_s|^2\, ds\Big)^{1/2} : \psi \in C(x,y)\Big\}.$$

These metrics are really important in various fields of mathematics, having fundamental role in particular in sub-Riemannian geometry. The following estimates are proved: there exist a constant $M \geq 1$ such that

$$\begin{aligned} &\frac{1}{M|B_d(x,t^{1/2})|}\exp\left(-\frac{Md(x,y)^2}{t}\right) \\ &\qquad \leq p_t(x,y) \leq \frac{M}{|B_d(x,t^{1/2})|}\exp\left(-\frac{d(x,y)^2}{Mt}\right) \end{aligned} \tag{4}$$

for $(t,x,y) \in (0,1] \times \mathbb{R}^n \times \mathbb{R}^n$, where $B_d(x,r) = \{y \in \mathbb{R}^n : d(x,y) < r\}$.

Remark 3 As we said before, this estimate holds if the drift is generated by the vector fields of the diffusive part, meaning $b = \sum_{k=1^d} \alpha_k \sigma_k$, for some $\alpha_1, \ldots, \alpha_d \in C_b^\infty(\mathbb{R}^n)$. This is a slight generalisation of the pure-diffusion case $b = 0$.

3 A diffusion process under a weak Hörmander condition

In this last section we present some original result, based on [9]. We consider a diffusion process

$$X_t = x + \int_0^t \sigma(X_s) \circ dW_s + \int_0^t b(X_s)ds, \tag{5}$$

where X is in dimension two, W is in dimension one, and $\circ dW_s$ denotes the Stratonovich integral. This clearly implies that our diffusion cannot be elliptic, but it also implies that the strong Hörmander condition cannot be satisfied, since σ is just a column vector. Our non-degeneracy assumption is indeed of weak Hörmander type: we suppose that σ and $[b, \sigma]$ span $\mathbb{R}^2$, and we suppose it just locally, in a sense that we will specify later. We are in a different framework respect to the classical result (4) (cf. Remark 3), since here we do not just allow a drift, but we actually need it to have the non-degeneracy. It is thanks to the existence of the drift that the randomness spreads in all directions and not just in the direction of σ. There has been some interest in recent years in density estimates for similar weak-Hörmander type models, see for instance [2] and [1].

The prototype of this kind of problems is a two dimensional system where the first component X^1 follows a stochastic dynamic, and the second component X^2 is a deterministic functional of X^1, so the randomness acts indirectly on X^2. A natural application is the equation used price the Asian option:

$$X_t^1 = x^1 + \int_0^t \sigma_1(X_s^1) \circ dW_s + \int_0^t b_1(X_s^1)ds, \qquad X_t^2 = \int_0^t b_2(X_s^1)ds.$$

Here X^1 represents the price of a stock following a local volatility model. X^2 is an average of the underlying price over some pre-set period of time, which determines the payoff of the so-called Asian options on X^1. In this case it is easy to see that σ and $[b, \sigma]$ span $\mathbb{R}^2$ if and only if $\sigma_1 \partial b_2 \neq 0$. This is what one would expect, since the randomness should act on X^1, and X^2 should see it through a dependence on X^1.

There are other possible application, such as in [3], [4]. These papers deal with a stochastic Hodgkin-Huxley model for the functioning of a neuron: X^2 is the concentration of some chemicals resulting from a reaction involving the first component X^1. Differently from our setting, though, there are several measurements corresponding to the input X^1, so X^2 is multi-dimensional. The pattern, however, is similar.

We take a control $\phi \in L^2[0, T]$, and the associated skeleton path solution of

$$x_t(\phi) = x + \int_0^t \sigma(x_s(\phi))\phi_s ds + \int_0^t b(x_s(\phi))ds. \tag{6}$$

We are interested in a tube estimate for (5), which is still an open problem under this weak non-degeneracy assumption. With tube estimate we mean that we are interested in $\mathbb{P}\left(\sup_{t \leq T} \|X_t - x_t(\phi)\| \leq R\right)$. Several works have considered this subject, starting from Stroock and Varadhan in [10]. For them $\|\cdot\|$ is the Euclidean norm, but later on different norms have been used to take into account the regularity of the trajectories. This is

somehow true also in our case. Indeed, because of the weak Hörmander framework, our diffusion is non-isotropic. This means that it moves with different speeds in different directions. More precisely, it moves with speed $t^{1/2}$ in the direction σ and $t^{3/2}$ in the direction $[b,\sigma]$. To account of this fact, we have to introduce a suitable norm. For any $R>0$, we denote with $A_R(x)$ the matrix $\left(R^{1/2}\sigma(x), R^{3/2}[b,\sigma](x)\right)$. For fixed R, since we suppose the weak Hörmander condition and $A_R(x)$ is invertible, we associate to $A_R(x)$ the norm

$$|\xi|_{A_R(x)} = \sqrt{\langle (A_R(x)A_R(x)^T)^{-1}\xi,\xi\rangle} = |A_R^{-1}(x)\xi|$$

on $\mathbb{R}^n$. This is what we need, because it allows us to weight the two time scales in the appropriate way. Let us have a closer look at this norm for the Asian option SDE, i.e. $\sigma^2(x)=0$:

$$A_R(x) = \left(R^{1/2}\sigma(x), R^{3/2}[b,\sigma](x)\right) = \begin{pmatrix} \sigma_1(x_1)R^{1/2} & (\dots)R^{3/2} \\ 0 & \sigma_1\partial b(x_1)R^{3/2} \end{pmatrix}.$$

In this case the associated norm is equivalent to $|\xi|_* = |(R^{-1/2}\xi^1, R^{-3/2}\xi^2)|$. Here it is easier to see what is going on. The first component, following a diffusive dynamic, must be weighted according to the time scale $t^{1/2}$. The second component, which is integrated in time once more, must be weighted according to the time scale $t^{1/2}\times t = t^{3/2}$. Analogous norms appear, in a multidimensional framework, in [2]. Just remark that the results that we are going to state here hold for a more general model in the sense that we allow $\sigma_2 \neq 0$.

We suppose σ, b differentiable three times, denote $l(x)$ the smallest eigenvalue of $A(x)=(\sigma,[b,\sigma])(x)$, and $n(x)=\sum_{k=0}^3\sum_{|\alpha|=k}(|\partial_x^\alpha b(x)|+|\partial_x^\alpha\sigma(x)|)$. We assume that:

H1 Locally uniform weak Hörmander condition: $l(y)\geq l_t$, along $(x_t(\phi))_{t\in[0,T]}$.

H2 Locally uniform bounds for derivatives: $n(y)\leq n_t$, along $(x_t)_{t\in[0,T]}$.

H3 Geometric condition on volatility: $\exists\kappa_\sigma:\mathbb{R}^2\to\mathbb{R}$ s. t.

$$\partial_\sigma\sigma(x) = \kappa_\sigma(x)\sigma(x).$$

We suppose w.l.o.g. that $|\kappa_\sigma(x)|\leq n(x)$, $|\kappa'_\sigma(x)|\leq n(x)$ (this is a consequence of **H2**). If $\sigma(x)=(\sigma_1(x),0)$, i.e. the Asian option stochastic differential equation, this property holds true with $\kappa_\sigma=\sigma_1'/\sigma_1$.

H4 Control on the growth of bounds: we suppose $|\phi.|^2$, $l.$, $n.\in L(\mu,h)$, for some $h\in\mathbb{R}_{>0}$, $\mu\geq 1$, where

$$L(\mu,h) = \{f(t)\leq \mu f(s) \quad \text{for } |t-s|\leq h\}$$

Notice that the above hypothesis do not involve global controls of our bounds on $\mathbb{R}^2$: they concern the behaviour of the coefficients only along the skeleton path.

Under assumptions **H1**, **H2**, **H3** we have the following Gaussian bounds for the density in short time. Define, for fixed δ, $\hat{x}=x+\delta b(x)$.

Theorem 4 *There exist constants $L, L_1, L_2, K_1, K_2, \delta^*$ such that: for any $r_* > 0$, for $\delta \leq \delta^* \exp\left(-Lr_*^2\right)$, for $|y - \hat{x}|_{A_\delta(x)} \leq r_*$,*

$$\frac{K_1}{\delta^2} \exp\left(-L_1 |y - \hat{x}|^2_{A_\delta(x)}\right) \leq p_{X_\delta}(z) \leq \frac{K_2}{\delta^2} \exp\left(-L_2 |y - \hat{x}|^2_{A_\delta(x)}\right).$$

Using this estimate we are able to prove the following result for the tube in the A_R-matrix norm:

Theorem 5 *We assume that* **H1**, **H2**, **H3**, **H4** *holds, with $x_t(\phi)$ given by (6). There exist K, q universal constants such that for $H_t = K\left(\frac{\mu n_t}{l_t}\right)^q$, for $R \leq R_*(\phi)$*

$$\exp\left(-\int_0^T H_t \left(\frac{1}{R} + |\phi_t|^2\right) dt\right) \leq \mathbb{P}\left(\sup_{t \leq T} |X_t - x_t(\phi)|_{A_R(x_t(\phi))} \leq 1\right) \leq \exp\left(-\int_0^T \frac{1}{H_t}\left(\frac{1}{R} + |\phi_t|^2\right) dt\right)$$

Both of these theorems can be stated in a control metric as well, which is a variant of the Caratheodory distance which looks appropriate to our framework. For $\phi \in L^2((0,1), \mathbb{R}^2)$, we define the norm

$$\|\phi\|_{(1,3)} = \left\|(\phi^1, \phi^2)\right\|_{(1,3)} = \left\|(|\phi^1|, |\phi^2|^{1/3})\right\|_{L^2(0,1)}.$$

and, given $A(x) = (\sigma(x), [b, \sigma](x))$, the set

$$C_A(x, y) = \{\phi \in L^2((0,1), \mathbb{R}^2) : dv_s = A(v_s)\phi_s ds,\ x = v_0,\ y = v_1\}.$$

We define the control norm as

$$d_c(x, y) = \inf\left\{\|\phi\|_{(1,3)} : \phi \in C_A(x, y)\right\}.$$

Remark that this distance accounts of the different speed in the $[b, \sigma]$ direction. We define also the following quasi-distance (which is naturally associated to the norm $|\cdot|_{A_R()}$):

$$d(x, y) \leq \sqrt{R} \Leftrightarrow |x - y|_{A_R(x)} \leq 1.$$

It is possible to prove that d and d_c are locally equivalent, and so we can re-state Theorem 5 as follows:

Corollary 6 *For $H_t = K\left(\frac{\mu n_t}{l_t}\right)^q$, with K, q universal constants, for small R it holds*

$$\exp\left(-\int_0^T H_t\left(\frac{1}{R}+|\phi_t|^2\right)dt\right) \leq \mathbb{P}\left(\sup_{0\leq t\leq T} d_c(X_t, x_t(\phi)) \leq \sqrt{R}\right) \leq \exp\left(-\int_0^T \frac{1}{H_t}\left(\frac{1}{R}+|\phi_t|^2\right)dt\right)$$

References

[1] V. Bally and A. Kohatsu-Higa, *Lower bounds for densities of Asian type stochastic differential equations.* J. Funct. Anal. 258/9 (2010), 3134– 3164.

[2] F. Delarue and S. Menozzi, *Density estimates for a random noise propagating through a chain of differential equations.* J. Funct. Anal. 259/6 (2010), 1577–1630.

[3] R. Höpfner, E. Löcherbach, and M. Thieullen, *Ergodicity for a stochastic Hodgkin-Huxley model driven by Ornstein-Uhlenbeck type input.* ArXiv e-prints.

[4] R. Höpfner, E. Löcherbach, and M. Thieullen, *Strongly degenerate time inhomogeneous sdes: densities and support properties. application to a Hodgkin-Huxley system with periodic input.* ArXiv e-prints.

[5] S. Kusuoka and D. Stroock, *Applications of the Malliavin calculus. II.* J. Fac. Sci. Univ. Tokyo Sect. IA Math. 32 (1985), 1–76.

[6] S. Kusuoka and D. Stroock, *Applications of the Malliavin calculus. III.* J. Fac. Sci. Univ. Tokyo Sect. IA Math. 34/2 (1987), 391–442.

[7] P. Malliavin and A. Thalmaier, "Stochastic Calculus of Variations in Mathematical Finance". Springer Verlag, Berlin, 2006.

[8] D. Nualart, "Malliavin Calculus and Related Topics". Springer, Berlin, 2006.

[9] P. Pigato, *Tube estimates for diffusion processes under a weak Hörmander condition.* Preprint (2014).

[10] D. Stroock and S. Varadhan, *On the support of diffusion processes with applications to the strong maximum principle.* In Lucien M. Le Cam, Jerzy Neyman, and Elizabeth L. Scott, editors, Proceedings of the sixth Berkeley symposium on mathematical statistics and probability, volume III: Probability theory, p. 333–359, Berkeley, CA, 1972. University of California Press. (Berkeley, CA, June 21–July 18, 1970).

Semistable degenerations of $K3$ surfaces

Genaro Hernández Mada (*)

The purpose of these notes is to present an introduction to the study of semistable degenerations of $K3$ surfaces over the complex numbers. More particularly, given a semistable degeneration of $K3$ surfaces

$$\pi : X \to \Delta,$$

we shall state a classification of the special fiber X_0 in terms of monodromy (see Theorem 6).

Since the purpose is not to give a detailed proof, we have omitted details about intermediate results. In particular, we don't give a proof for the exactness of the Clemens-Schmid sequence, which the most important tool to prove the main theorem. The main reference for this is [11].

We assume the reader familiar with basic notions of algebraic topology, and in particular with singular homology and cohomology with coefficients in a ring. For some Hodge-theoretical reasons, we mainly use rational coefficients. Good references for these notions are [5] and [8].

We have divided these notes in two sections. In the first one, we give the general geometric notions to understand the main result, such as complex manifolds, algebraic varieties and $K3$ surfaces. In the second one, we give the definition of semistable degeneration and we state the results concerning the particular case of $K3$ surfaces, getting at the end to the main theorem.

1 Geometric background

1.1 Complex manifolds

We begin with a basic definition in complex geometry:

Definition 1 A complex manifold of dimension n is a second countable Hausdorff topological space X such that there exists an open covering $\{U_i\}_{i\in I}$ together with homeomorphisms $\phi_i : U_i \to \mathbb{B}^n$, where

$$\mathbb{B}^n = \{(z_1, ..., z_n) \in \mathbb{C}^n | \ \ |z_1|^2 + \cdots + |z_n|^2 < 1\},$$

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: genarohm@math.unipd.it . Seminar held on December 17th, 2014.

with the property that for any pair $i \neq j$ such that $U_i \cap U_j \neq \varnothing$, the mapping $\phi_j \cdot \phi_i^{-1}$ is holomorphic.

Remark 1 One may define a C^∞-differentiable manifold (or more generally, C^k-differentiable manifold, for some $k \in \mathbb{N}$) by requiring $\mathbb{B}^n$ to be the real n-dimensional unit ball, and the transition maps $\phi_j \cdot \phi_i^{-1}$ to be C^∞ (resp. C^k). In particular, any complex manifold of dimension n is a C^∞-differentiable manifold of dimension $2n$.

Remark 2 The above definitions allow to speak of holomorphic maps between complex manifolds and of C^∞-differentiable functions between real manifolds.

Example 1 The simplest examples of complex manifolds are $\mathbb{B}^n$ itself, $\mathbb{C}^n$, and the polydisc $\mathbb{B}^1 \times \cdots \times \mathbb{B}^1$. It is important to note that these three examples are essentially different one from another as complex manifolds (i.e., we cannot find a holomorphic function with holomorphic inverse between two of them), while the real n-ball, $\mathbb{R}^n$ and the real polydisc are essentially the same (i.e, there exists a C^∞ bijection with C^∞ inverse between any two of them).

Example 2 If $k = \mathbb{R}$ or $\mathbb{C}$, we define the projective n-space $\mathbb{P}_k^n$ as the quotient $(k^{n+1} - \{0\})/\sim$, where $x \sim y$ iff $y = \lambda x$, for some $\lambda \in k - \{0\}$. Then, $\mathbb{P}_{\mathbb{C}}^n$ is a complex manifold of dimension n, $\mathbb{P}_{\mathbb{R}}^n$ is a C^∞ manifold of dimension n, and there is a natural embedding $\mathbb{P}_{\mathbb{R}}^n \hookrightarrow \mathbb{P}_{\mathbb{C}}^n$. In particular, we may identify $\mathbb{P}_{\mathbb{C}}^1$ with the Riemann sphere, and via this embedding, $\mathbb{P}_{\mathbb{R}}^1$ is an equator.

Example 3 Let $P \in \mathbb{C}[X_1, ..., X_n]$ be polynomials in n variables with complex coefficients. Suppose that for any $z = (z_1, ..., z_n) \in \mathbb{C}^n$, the vector of the partial derivatives $\frac{\partial P}{\partial X_i}$ evaluated at z is not 0. Then, the set

$$\{z \in \mathbb{C}^n | P(z) = 0\}$$

with its induced topology is a complex manifold.

This is an example of a non-singular (or smooth) affine algebraic variety. If we drop the assumption on the partial derivatives of P, we get a more general notion of affine algebraic variety, allowing singularities. This is not anymore an example of complex manifold, since these cannot have singularities, by definition. For example,

$$\{(x, y) \in \mathbb{C}^2 | xy = 0\}$$

is an affine algebraic variety, but it is not a complex manifold, since around the point $(0, 0)$, any open neighbourhood is not homeomorphic to an open ball.

1.2 Algebraic Varieties

Moreover, we can define algebraic varieties as the set of points in which a number of polynomials $P_1, ..., P_r \in \mathbb{C}[X_1, ..., X_n]$ vanish simultaneously.

If X is a complex manifold of dimension 1, we shall call it a curve, and if it is of dimension 2, we shall call it a surface. Note that, in this sense, a curve is of real dimension 2, so it would be a surface as a real manifold. In these notes, we shall work mainly with complex manifolds, and unless otherwise stated, when we refer to a curve or surface, it is in the complex sense. If X is also an algebraic variety, we shall say that it is an algebraic curve (resp. algebraic surface).

Example 4 An example of an algebraic curve is that of elliptic curve. One may define it as an algebraic curve defined by an equation of the type

$$y^2 = x^3 - px - q.$$

This is a very simple way of defining them, but they are object of study of current research. If the reader is interested, a good reference is [14].

In the context of algebraic geometry, an important notion is that of birational maps. To have the precise notion, see [4], but for our purposes, it is enough to think that two varieties X, Y are birationally equivalent (also called simply birational) if there are dense open subsets $U \subset X$, $V \subset Y$ isomorphic to each other.

Example 5 An example of an algebraic surface is that of ruled surfaces. If C is a smooth curve, then a ruled surface is a non-singular surface together with a map $X \to C$ such that all the fibers are birationally equivalent to $\mathbb{P}^1_k$.

Definition 2 An algebraic curve (resp. surface) X is rational if it is birationally equivalent to $\mathbb{P}^1_k$ (resp. $\mathbb{P}^2_k$).

1.3 Definition of $K3$ Surface

Now we shall give the definition of $K3$ surface. Here we shall denote by $\mathcal{O}_X$ the structural sheaf (i.e., the sheaf of holomorphic functions on X) and by Ω^n_X the sheaf of holomorphic n-forms on X. For these notions, see for example [16].

Definition 3 A $K3$ surface is a connected, compact complex surface X such that its first Betti number is $b_1(X) = 0$ and $\Omega^2_X \cong \mathcal{O}_X$.

The name of $K3$ surface was introduced by André Weil in [17], and it was in honor of three algebraic geometers: Kummer, Kähler and Kodaira; and also in honor of the mountain named $K2$ in Pakistan.

Example 6 Some examples of $K3$ surfaces are: any intersection of a quadric and a cubic in $\mathbb{P}^4$, or a quartic in $\mathbb{P}^3$.

2 Semistable degenerations

Let Δ be the complex open unit disc centered at 0, and X a complex manifold of dimension $n+1$.

Definition 4 A semistable degeneration is a proper, flat, holomorphic map $\pi : X \to \Delta$, such that $\pi^{-1}(t) = X_t$ is a smooth complex variety for all $t \neq 0$, and $X_0 = \pi^{-1}(0)$ has smooth irreducible components intersecting transversally in such a way that π is locally defined by the equation

$$t = x_1 \cdots x_k$$

Example 7 If X is defined by the equation $xy - t = 0$ in three variables, and π is defined by t, then it is clear that π is a semistable degeneration.

When the fibers are $K3$ surfaces we have the following result:

Theorem 1 *Let $\pi : X \to \Delta$ be a semistable degeneration with X_t a $K3$ surface for all $t \neq 0$. Then, X is birational to a semistable degeneration with special fiber X_0 being one of the followng:*

I) X_0 is smooth

II) $X_0 = Y_0 \cup Y_1 \cup \cdots \cup Y_k$, where Y_0, Y_k are rational surfaces, $Y_1, .., Y_{k-1}$ are elliptic ruled, with $Y_\alpha \cap Y_{\alpha-1}$ and $Y_\alpha \cap Y_{\alpha+1}$ sections of the ruling.

III) All components of X_0 are rational surfaces, $Y_i \cap (\cup_{j \neq i}) Y_j$ is a cycle of rational curves, and $|\Gamma| = \mathbb{S}^2$

A natural question now is how to get a criterion to decide which of the three types of special fiber we have, given a semistable degeneration of $K3$ surfaces. The one that we shall give here is in terms of cohomology, and more specifically, the monodromy operator on the second cohomology group of a generic fiber. The method that we use to get this is explained in [11].

Let X be a complex manifold of dimension $n+1$ and Δ the complex unit disc. Let $\pi : X \to \Delta$ be a semistable degeneration, i.e., a flat, proper, holomorphic map such for any $t \neq 0$, $X_t = \pi^{-1}(t)$ is a smooth complex variety, and $X_0 = \pi^{-1}(0)$ has smooth irreducible components intersecting transversally in such a way that π is locally defined by the equation

$$t = x_1 \cdots x_k.$$

In this case, the restriction of π to the punctured disc $\pi^* : X^* \to \Delta^*$ is a C^∞ fibration, so $\pi_1(\Delta^*)$ acts on $H^m(X_t)$ for any $t \neq 0$. The *Picard-Lefschetz transformation*, denoted by $T : H^m(X_t) \to H^m(X_t)$ is the map induced by the canonical generator of $\pi_1(\Delta^*)$. Then, one can prove (see for example [7]) that T is unipotent, i.e., $(T - I)^{m+1} = 0$, where I is

the identity operator. This allows to define a monodromy operator as

$$N := \log T = (T - I) - \frac{1}{2}(T - I)^2 + \frac{1}{3}(T - I)^3 - \cdots,$$

which is in fact a finite sum. It is also clear that N is nilpotent, hence we can endow $H^m(X_t)$ with an increasing filtration

$$0 \subset W_0 \subset W_1 \subset \cdots \subset W_{2m} = H^m(X_t)$$

which is the unique filtration such that:

1) $N(W_k) \subset W_{k-2}$.

2) N^k induces an isomorphism on the graded parts:

$$Gr_{m+k}(H^m(X_t)) \xrightarrow{\sim} Gr_{m-k}(H^m(X_t)).$$

One can make an explicit description of this filtration. See for example [11].

We define a filtration on $H^m(X_0)$ via a spectral sequence. Denote by $Y_1, ..., Y_r$ the irreducible components of X_0, which we assumed to be smooth and proper. We define the *codimension p stratum of X_0* as

$$X^{[p]} := \bigsqcup_{i_0 < \cdots < i_p} Y_{i_0} \cap \cdots \cap Y_{i_p}.$$

We define $E_0^{p,q} = A^q(X^{[p]})$, the C^∞ q-forms on $X^{[p]}$. Then, we have $d_0^{p,q} : E_0^{p,q} \to E_0^{p,q+1}$ the exterior derivative. We also have morphisms $\delta_0^{p,q} : E_0^{p,q} \to E_0^{p+1,q}$ induced by the combinatorial formula

$$(\delta_0^{p,q}\omega)|_{Y_{j_0} \cap \cdots \cap Y_{j_{p+1}}} = \sum_{k=0}^{p+1} (-1)^k \omega|_{Y_{j_0} \cap \cdots \cap \widehat{Y}_{j_k} \cap \cdots \cap Y_{j_{p+1}}}$$

where $\widehat{Y}_{j_k}$ means that we ignore this term. This defines a double complex $(E_0^{\bullet,\bullet}, d, \delta)$ and we have the following:

Theorem 2 *The spectral sequence with*

$$E_1^{p,q} = H^q(X^{[p]})$$

degenerates at level 2 and it converges to $H^*(X_0)$.

By letting

$$W_k = \bigoplus_{q \leq k} E_0^{*,q},$$

we get a filtration on the simple complex associated to the double complex $(E_0^{\bullet,\bullet}, d, \delta)$, and consequently a filtration on $H^m(X_0)$.

One can construct a retraction $r : X \to X_0$ which induces isomorphisms

$$r^* : H^m(X_0) \xrightarrow{\sim} H^m(X) \tag{1}$$

$$r_* : H_m(X) \xrightarrow{\sim} H_m(X_0). \tag{2}$$

The details of this construction are in [3].

The isomorphism (2) allows to define a filtration on $H_m(X_0) \cong H_m(X) =: H_m$. Indeed, we use Poincaré duality and define

$$W_{-k}(H_m) = \mathrm{Ann}(W_{k-1}(H^m)) = \{h \in H_m | (W_{k-1}(H^m), h) = 0\}.$$

Now that we have filtrations on H^m, H_m and $H^m_{\lim} := H^m(X_t)$. We shall define maps relating them and respecting the filtrations in the following sense:

Definition 5 Let H, H' vector spaces with filtrations that we denote by $W_\bullet$ for both. A morphism of filtered vector spaces of type r is a linear map $\phi : H \to H'$ such that for all k,

$$\phi(W_k(H)) = W_{k+2r}(H') \cap \mathrm{Im}(\phi).$$

We define a morphism $\alpha : H_{2n+2-m} \to H^m$ as the composite

$$H_{2n+2-m}(X_0) \xrightarrow{p} H^m(X, X - X_0) \longrightarrow H^m(X),$$

where p is the Poincaré duality map, and the second is the natural morphism.

We define $\beta : H^m_{\lim} \to H_{2n-m}$ as the composite

$$H^m(X_t) \xrightarrow{p_t} H_{2n-m}(X_t) \xrightarrow{i_*} H_{2n-m}(X),$$

where i_* is induced by the natural inclusion $X_t \hookrightarrow X$ and p_t is the Poincaré duality morphism. Then, we have the following:

Theorem 3 (Clemens-Schmid) *The maps α, i^*, N, β are morphisms of filtered vector spaces of type $n+1, 0, -1, -n$, respectively, and the sequence*

$$\cdots \to H_{2n+2-m} \xrightarrow{\alpha} H^m \xrightarrow{i^*} H^m_{\lim} \xrightarrow{N} H^m_{\lim} \xrightarrow{\beta} H_{2n-m} \xrightarrow{\alpha} H^{m+2} \to \cdots \tag{3}$$

is exact.

Remark 3 One can state the Clemens-Schmid exact sequence as an exact sequence of Mixed Hodge Structures. This explain the notation of $H^m_{\lim}$, since that term is considered with the limit Mixed Hodge structure, defined by Steenbrink in [15].

Since we are interested in studying surfaces, now we want to restrict ourselves to the case $n = 2$. By using the exact sequence (3) for H^2, restricted to the elements of the filtrations on each term, and the properties of the graded parts, one can prove the following *monodromy criteria*:

Theorem 4 *Let Γ be the dual graph of X_0 and denote*

$$\Phi = \dim \ker(H^1(X^{[0]}) \to H^1(X^{[1]})),$$

$$q = \frac{1}{2}h^1(X^{[0]}), \quad g = \frac{1}{2}h^1(X^{[1]}).$$

Then,

(a) $N = 0$ *on* $H^1_{\lim}$ *if and only if* $h^1(|\Gamma|) = 0$ *if and only if* $b_1(X_t) = \Phi$.

(b) $N^2 = 0$ *on* $H^2_{\lim}$ *if and only if* $h^2(|\Gamma|) = 0$.

(c) $N = 0$ *on* $H^2_{\lim}$ *if and only if* $h^2(|\Gamma|) = 0$ *and* $\Phi + 2g = 2q$.

Now we apply these monodromy criteria to the case of semistable degenerations of $K3$ surfaces. In order to use it, we first need the following classification:

Theorem 5 *A semistable degeneration of* $K3$ *surfaces is birational to one for which the central fiber* X_0 *is one of three types:*

- *Type I.* X_0 *is a smooth* $K3$ *surface.*
- *Type II.* $X_0 = Y_0 \cup \cdots \cup Y_{k+1}$, *where* Y_α *intersects only* $Y_{\alpha\pm1}$, *and each* $Y_\alpha \cap Y_{\alpha+1}$ *is an elliptic curve.* Y_0, Y_{k+1} *are rational surfaces, and for* $1 \leq \alpha \leq k$, Y_α *is ruled with* $Y_\alpha \cap Y_{\alpha+1}$ *and* $Y_\alpha \cap Y_{\alpha-1}$ *sections of the ruling.*
- *Type III. All components of* X_0 *are rational surfaces,* $Y_i \cap (\cup_{j\neq i} Y_j)$ *is a cycle of rational curves, and* $|\Gamma| = S^2$.

Finally by applying the monodromy criteria to these three cases, we get a classification of the special fiber in terms of the monodromy operator N:

Theorem 6

(a) X_0 *is of type I if and only if* $N = 0$ *on* $H^2_{\lim}$.

(b) X_0 *is of type II if and only if* $N \neq 0$, *but* $N^2 = 0$ *on* $H^2_{\lim}$.

(c) X_0 *is of type III if and only if* $N^2 \neq 0$.

Proof. We shall prove that if $N = 0$ on $H^2_{\lim}$, then X_0 is neccesarily of type I; if $N \neq 0$ and $N^2 = 0$, then X_0 is necessarily of type II; and if $N^2 \neq 0$, then X_0 is necessarily of type III. This shall prove the equivalence, since we know that we can be only in one of these three cases.

First assume that X_0 is of type I. Then, $X^{[0]} = X_s$, $X^{[1]} = \varnothing$ and the dual graph Γ is only one point. In this case, the spectral sequence has the form

$$E_\infty^{p,q} = E_1^{p,q} = H^q(X^{[p]}) = \begin{cases} 0 & \text{if } p \geq 1 \\ H^q(X_s) & \text{if } p = 0 \end{cases}$$

and this gives immediately that $\Phi = \dim Gr_1 H^1 = \dim E_2^{0,1} = 0$. Since $H^1(X_s) = H^1(X^{[1]}) = 0$, and $h^2(|\Gamma|) = 0$, we conclude that $N = 0$, by Theorem 4 (iii).

Now assume that X_0 is of type II. In this case, it is clear that the dual graph is homeomorphic to $[0,1]$. In particular, $h^2(|\Gamma|) = 0$ and $N^2 = 0$ by Theorem 4 (ii). By definition of the type II, $X^{[1]}$ is the disjoint union of $j+1$ elliptic curves, hence $h^1(X^{[1]}) = 2j+2$. Since X_0 and X_{j+1} are rational surfaces, we have

$$h^1(X^{[0]}) = \sum_{i=1}^{j} h^1(X_i),$$

but the X_i's are ruled, with the double curves rulings. Then, $h^1(X^{[0]}) = 2j$ and we get $h^1(X^{[0]}) - h^1(X^{[1]}) = -2$, but Φ cannot be negative, hence

$$\Phi \neq h^1(X^{[0]}) - h^1(X^{[1]})$$

and $N \neq 0$. Finally, assume that X_s is of type III. In this case, $h^2(|\Gamma|) = h^2(S^2) = 1 \neq 0$, hence $N^2 \neq 0$. This completes the proof. □

By the preceding theorem, we can conclude that for semistable degenerations of $K3$ surfaces, one can obtain a classification of the special fiber in terms of monodromy. Now we give an example of how this theorem can be used to study other kind of surfaces.

Example 8 Let $\pi : X \to \Delta$ be a semistable degeneration of Enriques surfaces. We assume moreover that π is weakly projective. One can check that in this case, the monodromy operator is always zero. However, one can get a classification as in the case of $K3$ surfaces, i.e., X is birational to one of the following cases (see [10]):

i) X_0 is a smooth Enriques surface.

ii) Flower pot.

iii) Elliptic chain with one rational component.

iv) Rational chain.

v) Polyhedral

vi) Polyhedral with boundary.

Moreover, one can construct a double cover $Y \to \Delta$, birational to a semistable degeneration of $K3$ surfaces. In particular, we have a correspondence: cases i) and ii) give a degeneration of type I); cases III) and iv) give a degeneration of type II); and cases v) and vi) give a degeneration of type III). This means that the type of degeneration is distinguished by the monodromy N on the associated family of $K3$ surfaces. Namely, in cases i) and ii), we have $N = 0$; in cases iii) and iv), we have $N \neq 0, N^2 = 0$; in cases v) and vi), we have $N^2 \neq 0, N^3 = 0$.
One can note, however, that if one wants a complete characterisation, one needs more information, unlike the case of $K3$ surfaces. A similar treatment to this can be done for hyperelliptic surfaces. For more details on both the case of Enriques and the case of hyperelliptic surfaces, see [10].

One may wonder if it is possible to do this entirely in an arithmetic setting, i.e., to get an arithmetic version of Theorem 6. This is beyond the purpose of these notes, but it is natural question that we are currently working on. Some references on how to address this problem are [1], [2], [6]. A result in this direction was obtained by Pérez Buendía in [13], where he uses a trascendental method to obtain the arithmetic result. In particular, he uses Theorem 6, so the classical result can be used to prove his main theorem.

References

[1] B. Chiarellotto, N. Tsuzuki, *Clemens-Schmid exact sequence in characteristic p.* ArXiv:1111.0779 (2012).

[2] B. Chiarellotto, *Rigid Cohomology and Invariant Cycles for a Semistable Log Scheme.* Duke Mathematical Journal 97 (1999), no.1, 155–169.

[3] C. H. Clemens, *Degeneration of Kähler manifolds.* Duke Math. J. 44 (1977), no. 2, 215–290.

[4] R. Hartshorne, "Algebraic Geometry". Springer Verlag, 1977.

[5] A. Hatcher, "Algebraic Topology". Available at `http://www.math.cornell.edu/~hatcher`.

[6] O. Hyodo, K. Kato, *Semi-stable Reduction and Crystalline Cohomology with Logarithmic Poles.* "Période p-adiques", Astérisque 223 (1994), 221–268.

[7] A. Landman, *On the Picard-Lefschetz transformation for algebraic manifolds acquiring general singularities.* Trans. Amer. Math. Soc. 181 (1973), 89–126.

[8] W. S. Massey, "A Basic Course in Algebraic Topology". Springer Verlag, 1991.

[9] A. Mokrane, *La suite spectrale des poids en cohomologie de Hyodo-Kato.* Duke Math. J. 72 (1993), 301–337.

[10] D. Morrison, *Semistable Degenerations of Enriques and Hyperelliptic Surfaces.* Duke Math. J. 48 (1981), no.1, 197–249.

[11] D. Morrison, *Clemens-Schmid Exact Sequence and Applications.* "Topics in Transcendental Algebraic Geometry", Annals of Mathematics Studies 106, Princeton University Press (1984), 101–119.

[12] Y. Nakajima, *Liftings of Simple Normal Crossing Log $K3$ and Log Enriques Surfaces in Mixed Characteristics.* Algebraic Geometry 9 (2000), 355–393.

[13] J. R. Pérez Buendía, *A Crystalline Criterion for Good Reduction on Semi-stable $K3$-Surfaces over a p-Adic Field.* Available at `http://spectrum.library.concordia.ca/978195/` (2014).

[14] J. Silverman, "The Arithmetic of Elliptic curves". Graduate Texts in Mathematics, Vol. 106, Springer New York, second edition, 2009.

[15] J. Steenbrink, *Limits of Hodge structures.* Invent. Math. 31 (1975/76), no. 3, 229–257.

[16] C. Voisin, "Hodge Theory and Complex Algebraic Geometry, I". Cambridge Studies in Advance Mathematis 76, 2002.

[17] A. Weil, "Final report on contract AF 18(603)-57". Scientific works. Collected papers II, Springer Verlag, Berlin, New York, pp. 390–395, 545–547.

Shape sensitivity analysis for vibrating plate models

Davide Buoso (*)

The classical formulation of the vibrating clamped plate problem is the following

$$\begin{cases} \Delta^2 u = \gamma u, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \\ \frac{\partial u}{\partial n} = 0, & \text{on } \partial\Omega, \end{cases} \tag{1}$$

in the unknowns u (the eigenfunction), γ (the eigenvalue). Here $\Omega \subset \mathbb{R}^2$ represents the shape of the plate, u the displacement, and γ the vibration frequency. The differential problem (1) comes from the study of a vibrating plate with clamped edges, within the so-called Kirchhoff-Love model. Recall that its weak formulation is

$$\int_\Omega \Delta u \Delta \varphi = \gamma \int_\Omega u\varphi, \ \forall \varphi \in H_0^2(\Omega).$$

We refer to [10, 11] for an introduction to the problem.

Using the Reissner-Mindlin model instead, the classical formulation of the vibrating clamped plate problem is

$$\begin{cases} -\frac{\mu}{12}\Delta\beta_t - \frac{\mu+\lambda}{12}\nabla \mathrm{div}\beta_t - \mu\frac{k}{t^2}(\nabla w_t - \beta_t) = \frac{t^2\gamma}{12}\beta_t, & \text{in } \Omega, \\ -\mu\frac{k}{t^2}(\Delta w_t - \mathrm{div}\beta_t) = \gamma w_t, & \text{in } \Omega, \\ \beta_t = 0, \ w_t = 0, & \text{on } \Omega, \end{cases} \tag{2}$$

in the unknowns (β_t, w_t) (the eigenfunction), γ (the eigenvalue). Here w_t represents the displacement, β_t the fiber rotation, μ, λ are the Lamé constants of the material, t is the thickness of the plate and k is a correction factor. We recall that the weak formulation of the Reissner-Mindlin vibrating clamped plate problem (2) is

$$\begin{aligned} \frac{\mu}{12}\int_\Omega \nabla\beta_t : \nabla\eta dx + \frac{\mu+\lambda}{12}\int_\Omega \mathrm{div}\beta_t \mathrm{div}\eta dx \\ + \frac{\mu k}{t^2}\int_\Omega (\nabla w_t - \beta_t)\cdot(\nabla v - \eta)dx = \gamma\int_\Omega \left(w_t v + \frac{t^2}{12}\beta_t\cdot\eta\right)dx, \end{aligned} \tag{3}$$

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: dbuoso@math.unipd.it . Seminar held on January 28h, 2015.

$\forall(\eta, v) \in (H_0^1(\Omega))^2 \times H_0^1(\Omega)$, where $A : B = \sum_{i,j=1}^2 a_{ij} b_{ij}$.

We remark that, even if the Reissner-Mindlin model gives a system of equation, from a numerical point of view it seems better than the Kirchhoff-Love model, because it is of the second order, and therefore easier to treat using finite elements methods. Note also that the Kirchhoff-Love model assumes $\beta_t = \nabla w_t$, while the Reissner-Mindlin one drops such hypotesis. This assumption is satisfied for extremely thin plates, and therefore problem (2) is better to study moderately thick plates. However, we recall the following result from [1] (see also [2]).

Theorem 1 *Suppose w_t, β_t and ζ_t satisfy*

$$\frac{\mu}{12}\int_\Omega \nabla\beta_t : \nabla\eta dx + \frac{\mu+\lambda}{12}\int_\Omega \mathrm{div}\beta_t \mathrm{div}\eta dx + \int_\Omega \zeta_t \cdot (\nabla v - \eta) dx \;=\; \int_\Omega g v dx,$$

where $\zeta_t = \frac{\mu k}{t^2}(\nabla w_t - \beta_t)$. Then

$$\beta_t \rightharpoonup \beta_0 \text{ in } (H_0^1(\Omega))^2,$$

$$w_t \rightharpoonup w_0 \text{ in } H_0^1(\Omega),$$

$$\zeta_t \rightharpoonup \zeta_0 \text{ in } H^{-1}(\mathrm{div};\Omega),$$

where w_0, β_0 and ζ_0 satisfy

$$\frac{\mu}{12}\int_\Omega \nabla\beta_0 : \nabla\eta dx + \frac{\mu+\lambda}{12}\int_\Omega \mathrm{div}\beta_0 \mathrm{div}\eta dx + \int_\Omega \zeta_0 \cdot (\nabla v - \eta) dx \;=\; \int_\Omega g v dx,$$

and

$$\beta_0 = \nabla w_0.$$

As a consequence we get

$$\frac{2\mu+\lambda}{12}\Delta^2 w_0 = g.$$

In particular, this theorem tells us that the solutions of the Reissner-Mindlin clamped plate problem converge to those of the Kirchhoff-Love one, as $t \to 0$. This holds also for eigenfunctions and eigenvalues (cf. [9, Section 2]).

Our main interest is to study the dependence of the eigenvalue of problem (3) with respect to shape perturbation. In particular, we aim at stability estimates in the spirit of [5, 6, 7]. Note that such estimates for problem (1) were already proved in [8]. We also remark that the results shown here appeared in [4].

Let us start by the study of the map

$$\phi \mapsto \gamma_n[\phi(\Omega)],$$

where $\phi : \Omega \mapsto \phi(\Omega)$ is a diffeomorphism. Which regularity should we impose on ϕ? It seems that, since only Sobolev spaces H^1 are involved (the problem is of the second order),

the use of bi-Lipschitz homeorphism should give the desired results. The corresponding pull-back operator from $\phi(\Omega)$ to Ω is

$$(\beta, w) \mapsto (\beta \circ \phi, w \circ \phi).$$

Note that such construction is the straight adaption to the case of systems of the arguments used in [8] (see also [3]). The best estimate that can be obtained using such a pull-back is

$$|\gamma_n[\phi(\Omega)] - \gamma_n[\Omega]| \leq \frac{c}{t^2}\gamma_n[\Omega]\|\nabla\phi - I\|_{L^\infty(\Omega)}.$$

We observe that, as $t \to 0$, the coefficients of problem (3) are diverging, and therefore in principle we could not obtain a better result. However, we already know a priori that the eigenvalues converge to those of problem (1) (up to a multiplicative constant), and therefore we should expect a better estimate to hold. To this aim, we note that the unknown β_t converge to the gradient of w_t, hence we should treat it as a gradient rather than a vector. So the good pull-back seems to be

$$\beta \mapsto (\beta \circ \phi) \cdot (\nabla\phi).$$

Note that, in order to use such transformation, we need ϕ to be at least of class $C^{1,1}$. We have the following

Theorem 2 *Let*

$$\delta(\phi) = \max_{1\leq|\alpha|\leq 2} \sup_{x\in\Omega} |D^\alpha(\phi(x) - x)|.$$

There exists a constant $c > 0$ independent of ϕ, n, t such that

$$|\gamma_n[\phi(\Omega)] - \gamma_n[\Omega]| \leq c\gamma_n[\Omega]\delta(\phi), \tag{4}$$

provided $\delta(\phi) < c^{-1}$.

Note that, thanks to the uniformity of estimate (4) with respect to t, as a bypass product we obtain again the stability estimates for problem (1) (cf. [8]).

In general, even if two open sets are known to be diffeomorphic, it is not easy to construct diffeomorphisms ϕ and to control the quantity $\delta(\phi)$ via explicit geometric quantities. However, such a construction is possible in the so-called atlas class. We refer to [8] for the definition of atlas class, and to [5] for the construction of diffeomorphisms among open sets in such a class.

For our purposes, we briefly recall that an atlas $\mathcal{A}$ is defined as

$$\mathcal{A} = (\rho, s, s', \{V_j\}_{j=1}^s, \{r_j\}_{j=1}^s),$$

where V_j are cuboids, r_j are rotations, and $\rho > 0$, $s, s' \in \mathbb{N}$ are other parameters. In particular, an open set Ω belongs to the class $C(\mathcal{A})$ if $r_j(\Omega \cap V_j)$ is the subgraph of a continuous function g_j for any $j = 1, \ldots, s$. We remark that, in a similar way, it is also possible to define the classes $C^1(\mathcal{A})$, $C^2(\mathcal{A})$, $C^{0,1}(\mathcal{A})$, $C^{0,\alpha}(\mathcal{A})$ and so forth. In $C(\mathcal{A})$ we define the atlas distance

$$d_{\mathcal{A}}(\Omega_1, \Omega_2) = \max_{j=1,\dots,s} \|g_{j1} - g_{j2}\|_\infty.$$

Note that $(C(\mathcal{A}), d_{\mathcal{A}})$ is a complete metric space.

Then we have the following

Theorem 3 *Let $\mathcal{A}$ be fixed. Then for each $n \in \mathbb{N}$ there exists $c > 0$ independent of n, t such that*

$$|\gamma_n[\Omega_1] - \gamma_n[\Omega_2]| \leq c \max\{\gamma_n[\Omega_1], \gamma_n[\Omega_2]\} d_{\mathcal{A}}(\Omega_1, \Omega_2),$$

for all $\Omega_1, \Omega_2 \in C(\mathcal{A})$ such that $d_{\mathcal{A}}(\Omega_1, \Omega_2) \leq c^{-1}$.

We remark that, even though the atlas distance is quite easily computable, it is not a clear geometric quantity, since it obviously depends on the chosen atlas. In this sense, we recall that, given $A, B \subset \mathbb{R}^2$, the Hausdorff distance between A and B is defined as

$$d^{\mathcal{H}}(A, B) = \max\left\{\sup_{x \in A} d(x, B), \sup_{x \in B} d(x, A)\right\}.$$

Another geometric quantity, which turns out to be interesting in the frame of stability estimates, is the so-called lower Hausdorff deviation, which is defined as

$$d_{\mathcal{H}}(A, B) = \min\left\{\sup_{x \in A} d(x, B), \sup_{x \in B} d(x, A)\right\}.$$

In order to state the following result, we recall that an open set Ω belongs to the class $C^{\omega}_{M}(\mathcal{A})$ if

$$|g_j(x) - g_j(y)| \leq M\omega(|x - y|), \ \forall x, y \in r_j(\Omega \cap V_j),$$

for any $j = 1, \dots, s$, where M is a positive constant, and ω is a modulus of continuity.

Lemma 4 (Burenkov, Lamberti) *There exists $K > 0$ such that*

$$d^{\mathcal{H}}(\Omega_1, \Omega_2) \leq d_{\mathcal{A}}(\Omega_1, \Omega_2) \leq K\omega(d_{\mathcal{H}}(\partial\Omega_1, \partial\Omega_2)),$$

for all $\Omega_1, \Omega_2 \in C^{\omega}_{M}(\mathcal{A})$.

By means of the previous lemma, it is easy to prove the following

Theorem 5 *Let $\mathcal{A}$, ω, M be fixed. Then there exists $c > 0$ independent of n, t such that*

$$|\gamma_n[\Omega_1] - \gamma_n[\Omega_2]| \leq c \max\{\gamma_n[\Omega_1], \gamma_n[\Omega_2]\} \omega(d_{\mathcal{H}}(\partial\Omega_1, \partial\Omega_2)),$$

for all $\Omega_1, \Omega_2 \in C^{\omega}_{M}(\mathcal{A})$ such that $d_{\mathcal{H}}(\partial\Omega_1, \partial\Omega_2) \leq c^{-1}$.

We conclude observing that, in the case $\Omega_1, \Omega_2 \in C^{\omega}_{M}(\mathcal{A})$ satisfy

$$(\Omega_1)_\epsilon \subset \Omega_2 \subset (\Omega_1)^\epsilon,$$

where

$$\begin{aligned}(\Omega_1)_\epsilon &= \{x \in \Omega_1 : \ d(x, \partial\Omega_1) > \epsilon\},\\ (\Omega_1)^\epsilon &= \{x \in \mathbb{R}^N : \ d(x, \Omega_1) < \epsilon\},\end{aligned}$$

then our estimates can be rewritten in a nicer form, namely

$$|\gamma_n[\Omega_1] - \gamma_n[\Omega_2]| \leq c_n \omega(\epsilon).$$

References

[1] F. Brezzi, M. Fortin, *Numerical approximation of Mindlin-Reissner plates.* Math. Comp. 47, no. 175 (1986), 151–158.

[2] F. Brezzi, M. Fortin, "Mixed and hybrid finite element methods". Springer Series in Computational Mathematics, 15. Springer-Verlag, New York, 1991.

[3] D. Buoso, "Shape sensitivity analysis of the eigenvalues of polyharmonic operators and elliptic systems". Ph.D. thesis, Università degli Studi di Padova, Padova, 2015.

[4] D. Buoso, P. D. Lamberti, *Shape sensitivity analysis of the eigenvalues of the Reissner-Mindlin system.* SIAM J. Math. Anal. 47 (2015), 407–426.

[5] V. I. Burenkov, E. B. Davies, *Spectral stability of the Neumann Laplacian.* J. Differential Equations 186 (2002), 485–508.

[6] V. I. Burenkov, P. D. Lamberti, *Spectral stability of general non-negative self-adjoint operators with applications to Neumann-type operators.* J. Differential Equations 233 (2007), 345–379.

[7] V. I. Burenkov, P. D. Lamberti, *Spectral stability of Dirichlet second order uniformly elliptic operators.* J. Differential Equations 244 (2008), 1712–1740.

[8] V. I. Burenkov, P. D. Lamberti, *Spectral stability of higher order uniformly elliptic operators.* In "Sobolev Spaces in Mathematics II. Applications in Analysis and Partial Differential Equations (to the centenary of Sergey Sobolev)", edited by V. Maz'ya, International Mathematical Series, Vol. 9, Springer, New York, 2009.

[9] R. G. Durán, L. Hervella-Nieto, E. Liberman, R. Rodríguez, J. Solomin, *Approximation of the vibration modes of a plate by Reissner-Mindlin equations.* Math. Comp. 68, no. 228 (1999), 1447–1463.

[10] A. Henrot, "Extremum problems for eigenvalues of elliptic operators". Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2006.

[11] J. W. S. Rayleigh, "The theory of sound". Dover Pub. New York, 1945 (republication of the 1894/96 edition).

Metastability of the Ising model on random graphs at zero temperature

Sander Dommers (*)

Abstract. In this note, which is mainly based on [6], a random graph model known as the configuration model is introduced. After this, we discuss the Ising model, which is a model from statistical physics where a spin is assigned to each vertex in a graph and these spins tend to align, i.e., take the same value as their neighbors. It is especially interesting to study the Ising model on random graphs. We discuss some properties of this model. In particular, we study the dynamics and metastability in this model when the interaction strength goes to infinity. This corresponds to the zero temperature limit in physical terms.

1 Introduction

In the past decades complex networks and their behavior have attracted much attention. In the real world many of such networks can be found, for instance as social, information, technological and biological networks. In [12], an overview of many networks and their properties is given. It turns out that many of these networks behave more or less similarly. For example, most of these networks are so-called *small worlds*, which means that it takes only small number of connections to go from any node in the network to any other node. They are also very inhomogeneous, many nodes only have a small number of connections, but there are also nodes, called *hubs*, with a huge number of connections. Such networks are also called *scale free*.

Since these networks are very complex, many random graph models have been proposed to study them. The most well known is the Erdős-Rényi random graph [9], where between every pair of vertices an edge is formed independently with a certain probability. One drawback of this model is that it does not produce scale-free graphs. In this note we focus on the *configuration model*, where this problem is easily solved, since the number of connections of all the vertices in the graph is specified in advance before constructing the graph, and hence the scale-free property can be given as an input to the model. The exact construction of the configuration model is described below.

(*)Università di Bologna, Dipartimento di Matematica, Piazza di Porta San Donato 5, 40126 Bologna, BO, Italy. E-mail: sander.dommers@unibo.it . Seminar held on February 11th, 2014.

Not only the structure of these networks is interesting, also the behavior of processes living on these networks is a fascinating subject. Processes one can think of are opinion formation, the spread of information and the spread of viruses.

In this note, we focus on a simple model for opinion formation known as the Ising model. This model was first used to describe magnetization in a simple way, but has later been used as a model for many other phenomena, see [13, 14, 15] for an extensive history. This model can, for example, be seen as a situation where people can have two different opinions. They might prefer to vote for left or right wing political parties, for example. Every individual's opinion is influenced by that of their friends. If, for example, most of someones friends vote for a right wing party, it is more likely this person will also vote for that party, and hence it is less likely that this person will vote for a left wing party. Also external sources such as the media can influence peoples opinions.

We focus on the dynamics of this model in the situation where it is very unlikely, but not impossible, that someone goes against the majority opinion among that persons friends and against the external influence. In this case the stable situation is when everyone has the same opinion that is also the same as the external influence dictates. If everyone has the same opinion, but this is different from the external influence, the system is in a *metastable* situation. Nobody likes to change their opinion first because that would cause a disagreement with their friends, but at the same time everyone wants to change their opinion because the external influence tells them to. It can therefore take a very long time for the system to reach the stable situation if the system starts from the metastable situation. How long this will take is a question we answer in this note for some specific choices of all the parameters involved.

This note is mainly based on [6] and this introduction is partly taken from [8].

2 Configuration model

In the configuration model the degrees of the vertices are prescribed and the graph is then chosen randomly among all graphs with these degrees. More precisely, the configuration model is constructed as follows [3].

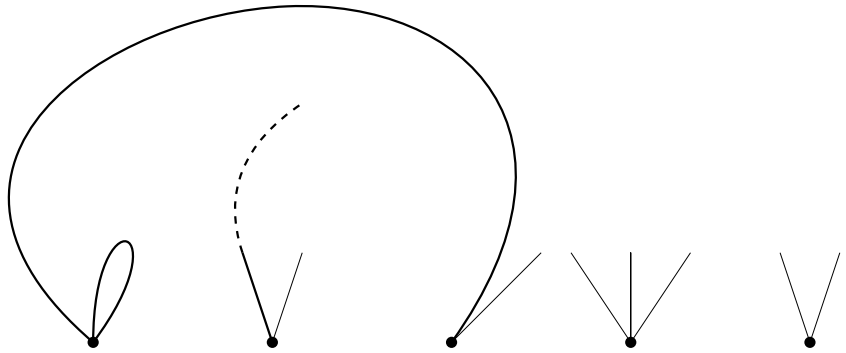
Start with n vertices labeled $1, \dots, n$ and write $\{1, \dots, n\} = [n]$. Let D be a random variable with

$$\mathbb{P}[D = k] = p_k, \qquad k = 1, 2, \dots. \tag{2.1}$$

Next, assign to assign to each vertex $i \in [n]$ D_i half-edges, where the D_i are independent and identically distributed with the same distribution as D, so that vertex i has degree D_i. Let $L_n = \sum_{i=1}^{n} D_i$ be the total degree and assume that L_n is even. If this is not the case, add 1 to D_n. Since we are interested in the limit $n \to \infty$ this will hardly change the graph.

With L_n even we can now construct a graph. We do this by connecting one of the half-edges to one of the other $L_n - 1$ half-edges uniformly at random. We repeat this procedure of pairing up unpaired half-edges uniformly at random until all half-edges have been connected and denote the resulting graph by G_n.

In the next figure, for example, the first two edges have been formed and the dashed line will form an edge with one of the remaining half-edges.



As becomes clear from this example, this construction does not necessarily result in a simple graph, because both self-loops and multiple edges between two vertices might occur. However, the number of times this happens is not growing with n [10] and hence is negligible.

An event A_n is said to hold *with high probability (whp)*, if

$$\lim_{n\to\infty} \mathbb{P}[A_n] = 1, \tag{2.2}$$

where $\mathbb{P}$ denotes the measure of selecting a random graph according to the configuration model described above.

For a graph G_n, the (edge) boundary $\partial_e A$ of a set $A \subseteq [n]$ consists of all edges between a vertex in A and a vertex outside of A, i.e.,

$$\partial_e A = \{(i,j) \in E_n \,|\, i \in A,\, j \notin A\}. \tag{2.3}$$

An important property of the graphs of interest is that the boundaries of any subset of the vertices is large compared to the number of vertices in this subset. Such graphs are called *expander graphs*. The formal definition of an expander graph is as follows.

Definition 2.1 (Expander graph) A graph G_n is a (δ, λ)-expander graph if for all $A \subset [n]$ with $\delta \leq \frac{|A|}{n} \leq \frac{1}{2}$,

$$\frac{|\partial_e A|}{|A|} \geq \lambda. \tag{2.4}$$

Graphs chosen according to the configuration model where the degrees are uniformly bounded and at least 3, are expander graphs, as was proved in [2]:

Lemma 2.2 (Expander graphs) *If $p_1 = p_2 = 0$ and the degrees are uniformly bounded, then for every $0 < \delta < \frac{1}{2}$ there exists a $\lambda_{\delta_0} > 0$ such that G_n is a $(\delta, \lambda_{\delta_0})$-expander graph w.h.p.*

A related quantity is the so-called *isoperimetric number*, which is defined as follows:

Definition 2.3 (Isoperimetric number) For a graph G_n, the (edge) *isoperimetric number* of G_n equals

$$i_e(G_n) = \min_{\substack{A\subset[n]\\ |A|\le n/2}} \frac{|\partial_e A|}{|A|}, \tag{2.5}$$

where $|A|$ denotes the cardinality of the set A.

This definition implies that any graph is a $(0, i_e(G_n))$-expander graph. When all degrees are equal to $r \geq 3$, such graphs are called r-regular graphs, then good bounds on the isoperimetric number are known as proved in [4, 1].

Lemma 2.4 (Bounds on isoperimetric number) *Let G_n be a random r-regular graph with $r \geq 3$. Then, whp, there exists a constant $C > 0$ such that*

$$\frac{r}{2} - \sqrt{\log 2}\sqrt{r} \leq i_e(G_n) \leq \frac{r}{2} - C\sqrt{r}. \tag{2.6}$$

The above shows that these random graphs behave in a significantly different way then, for example, the lattice $\mathbb{Z}^d$. If you look at a box $[\ell]^d \subset \mathbb{Z}^d$ then the boundary of this box grows like ℓ^{d-1} and the number of vertices in the box like ℓ^d. Hence, by letting $\ell \to \infty$, we get that the isoperimetric number of $\mathbb{Z}^d$ equals 0.

3 Ising model

The Ising model on a graph G_n is defined as follows. To each vertex $i \in [n]$ we assign a spin $\sigma_i \in \{-1, +1\}$ and we denote a configuration by $\sigma = (\sigma_i)_{i\in[n]}$. We define an energy function $H(\sigma) : \{-1, +1\}^n \mapsto \mathbb{R}$, as

$$H(\sigma) = -J \sum_{(i,j)\in E_n} \sigma_i\sigma_j - h \sum_{i\in[n]} \sigma_i, \tag{3.1}$$

where $J > 0$ is the interaction constant and $h \in \mathbb{R}$ is the external magnetic field. $H(\sigma)$ is called the Hamiltonian.

The probability that in equilibrium the system has configuration σ is then given by Boltzmann-Gibbs measure which is defined as

$$\mu_n(\sigma) = \frac{1}{Z_n} e^{-\beta H(\sigma)}, \tag{3.2}$$

where $\beta = 1/T \geq 0$ is the inverse temperature and Z_n is the normalization factor, called the partition function, i.e.,

$$Z_n = \sum_{\sigma\in\{-1,+1\}^n} e^{-\beta H(\sigma)}. \tag{3.3}$$

Without loss of generality, we assume that $J = 1$, since this is just a rescaling of β and h.

From this it becomes clear that spins in the graph that are neighbors of each other indeed tend to align, because if these spins have the same sign then the energy is lower, and hence the probability that this happens is higher.

For a set $A \subseteq [n]$, denote by σ^A the configuration where

$$\sigma_i^A = \begin{cases} +1, & \text{if } i \in A, \\ -1, & \text{if } i \notin A. \end{cases} \tag{3.4}$$

We also denote $\boxminus = \sigma^{\varnothing}$ and $\boxplus = \sigma^{[n]}$, the all minus and all plus configurations, respectively. We often identify the vertex and its spin, e.g., we say that vertex i has a $+$ neighbor if there is a vertex j such that $(i, j) \in E_n$ with $\sigma_j = +1$.

Many results about the equilibrium solution for the Ising model on random graphs have been obtained, see, for example, [5, 7, 8].

4 Dynamics of the Ising model and metastability

Besides looking at the Ising model at equilibrium, we can also look at this model when the spins evolve in time. We let the system evolve according to *Glauber dynamics* with Metropolis rates. That is, we consider a discrete time Markov chain where at every time step we select one of the n spins uniformly at random. If flipping this spin (changing it from $-$ to $+$ or from $+$ to $-$) results in a configuration with lower energy, this spin is always flipped. If the energy is going up by flipping this spin, the spin is flipped with probability $e^{-\beta \Delta H}$, where ΔH is the energy difference. If β is large this means that this is very unlikely to happen (but not impossible).

More formally, we can write the transition probabilities $c(\sigma^A, \sigma^B)$ from configuration σ^A to σ^B as

$$c(\sigma^A, \sigma^B) = \begin{cases} \frac{1}{n} e^{-\beta[H(\sigma^B) - H(\sigma^A)]^+}, & \text{if } |A \triangle B| = 1; \\ 1 - \sum_{B : |A \triangle B| = 1} \frac{1}{n} e^{-\beta[H(\sigma^B) - H(\sigma^A)]^+}, & \text{if } A = B, \\ 0, & \text{otherwise,} \end{cases} \tag{4.1}$$

where $A \triangle B$ is the symmetric difference between sets A and B, and $[a]^+ = \max\{a, 0\}$. We denote by $\mathbb{P}_\eta$ the law of the process starting from configuration η.

The time at which the process visits the configuration σ for the first time if the process starts from η is called the *hitting time* of σ and is denoted by τ_σ. When studying metastability, the problem is to find the hitting time of the stable configuration if the system starts in a metastable configuration. We now define what it means for a configuration to be (meta)stable.

The *stable state* is the state for which the Hamiltonian is minimal. Throughout the rest of this note we assume that $h > 0$, so that it is obvious from (3.1) that the stable state is $\boxplus$.

To define metastable states, we need to define the *communication height* between two configurations σ and σ' which is given by

$$\Phi(\sigma, \sigma') = \min_{\omega \text{ path from } \sigma \text{ to } \sigma'} \max_{\sigma'' \in \omega} H(\sigma''), \tag{4.2}$$

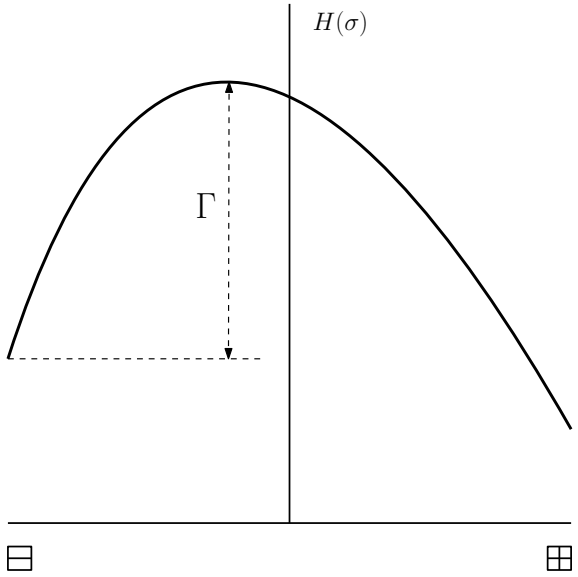
where we say that a sequence of configurations ω is a path from σ to σ' if $\omega = (\sigma = \sigma^{A_0}, \sigma^{A_1}, \dots, \sigma^{A_\ell} = \sigma')$ for some $\ell \geq 1$ and $|A_k \triangle A_{k+1}| = 1$ for all $0 \leq k < \ell$. We then define the *stability level* of a configuration σ as

$$V_\sigma = \min_{\sigma' : H(\sigma') < H(\sigma)} \Phi(\sigma, \sigma') - H(\sigma). \tag{4.3}$$

Note that $V_{\boxplus} = \infty$ since there are no configurations with smaller energy. The *maximal stability level* is defined as

$$\Gamma = \max_{\sigma \neq \boxplus} V_\sigma, \tag{4.4}$$

and the *metastable states* are those configurations η such that $V_\eta = \Gamma$. In general, the metastable state in the Ising model is $\boxminus$. (This has to be proved, but we assume this is true in this note.) This is depicted schematically in the following picture:



The maximal stability level Γ is an important quantity because of the following result from [11]:

Theorem 4.1 (Metastable time) *For all $\varepsilon > 0$,*

$$\lim_{\beta \to \infty} \mathbb{P}_{\boxminus}[e^{\beta(\Gamma - \varepsilon)} < \tau_{\boxplus} < e^{\beta(\Gamma + \varepsilon)}] = 1. \tag{4.5}$$

This theorem says that the time it takes for the system to go from the metastable configuration $\boxminus$ to the stable configuration $\boxplus$ is proportional to $e^{\beta \Gamma}$ when $\beta \to \infty$. Hence, if we are interested in the metastable time in the limit $\beta \to \infty$ (the zero temperature limit), it suffices to study the energy function $H(\sigma)$ to compute Γ.

4.1 Lower bound on the metastable time

We now present an example of the type of computations that need to be done to combine the above results to get bounds on the metastable time for the Ising model on random graphs.

To be specific, we derive a lower bound on the communication height between $\boxminus$ and $\boxplus$ for expander graphs.

Lemma 4.2 (Lower bound on communication height) *Let G_n be a (δ, λ)-expander graph for some $0 \leq \delta < \frac{1}{2}$ and $\lambda > 0$ and suppose that $0 < h < \lambda$. Then,*

$$\Phi(\boxminus, \boxplus) - H(\boxminus) \geq (\lambda - h)n. \tag{4.6}$$

Proof. For any subset $A \subseteq [n]$ with $|A| \leq n/2$, it holds that

$$H(\sigma^A) = -\sum_{(i,j)\in E_n} \sigma_i^A \sigma_j^A - h \sum_{i\in[n]} \sigma_i^A = -(|E_n| - |\partial_e A|) + |\partial_e A| - h|A| + h(n - |A|)$$

$$= 2|\partial_e A| - 2h|A| - |E_n| + hn. \tag{4.7}$$

Note that every path from $\boxminus$ to $\boxplus$ has to go through a configuration with $\lfloor n/2 \rfloor$ plus spins, because only one spin at a time can change. By the definition of expander graphs, for any such configuration A,

$$H(\sigma^A) = 2|\partial_e A| - 2h|A| - |E_n| + hn \geq 2(\lambda - h)\frac{n}{2} - |E_n| + hn. \tag{4.8}$$

The statement of the lemma now follows by observing that

$$H(\boxminus) = -|E_n| + hn. \tag{4.9}$$

□

Note that this lemma holds for general graphs, but that only gives useful information if λ stays strictly positive in the limit $n \to \infty$. As we mentioned this is not the case for $\mathbb{Z}^d$, for example, but it is true for the configuration model. Hence, combining the above lemma with Theorem 4.1 we get the following lower bound on the metastable time.

Corollary 4.3 (Lower bound on the metastable time) *Let G_n be a (δ, λ)-expander graph for some $0 \leq \delta < \frac{1}{2}$ and $\lambda > 0$ and suppose that $0 < h < \lambda$. Then,*

$$\lim_{\beta\to\infty} \mathbb{P}_{\boxminus}[\tau_{\boxplus} > e^{\beta((\lambda-h)n-\varepsilon)}] = 1. \tag{4.10}$$

Proof. If it is indeed true that $\boxminus$ is the metastable state, then

$$V_{\boxminus} = \Phi(\boxminus, \boxplus) - H(\boxminus). \tag{4.11}$$

From the definition of Γ and Lemma 4.2 it then follows that

$$\Gamma \geq \Phi(\boxminus, \boxplus) - H(\boxminus) \geq (\lambda - h)n. \tag{4.12}$$

The corollary now follows from Theorem 4.1. □

5 Concluding remarks

Above we gave a lower bound on the metastable time for the Ising model on random graphs chosen according to the configuration model. For r-regular graphs, Lemma 2.4 can be used to get a better lower bound. This lemma, or rather more precise results from [1], can also be used to get an upper bound on the metastable time. This gives the following theorem, which is the main result of [6].

Theorem 5.1 (Metastable time for random r-regular graphs) *Let G_n be a random r-regular graph with $r \geq 3$ and suppose that $0 < h < C_0\sqrt{r}$ for some uniform constant $C_0 > 0$ small enough. Then, there exist uniform constants $0 < C_1 < \sqrt{3}/2$ and $C_2 < \infty$ so that, whp,*

$$\lim_{\beta\to\infty} \mathbb{P}_{\boxminus}[e^{\beta(r/2-C_1\sqrt{r})n} < \tau_{\boxplus} < e^{\beta(r/2+C_2\sqrt{r})n}] = 1. \tag{5.1}$$

This gives bounds on the metastable time, but there is still a gap between the constants $-C_1$ and C_2. An interesting open problem is to identify the exact constant, or even to prove that such a constant exists.

It would also be interesting to generalize this results to general degree sequences. Above we showed how to obtain a lower bound also in this case, but how to obtain a matching upper bound is not known.

The above results only say something about the zero temperature limit $\beta \to \infty$. It would also be interesting to study the behavior of this model at low positive temperatures (large but finite β). This in general is a much more difficult problem, because not only the energy function $H(\sigma)$ is of importance, but also entropy effects have to be taken into account.

References

[1] N. Alon, *On the edge-expansion of graphs.* Combinatorics, Probability and Computing 6/2 (1997), 145–152.

[2] A. Basak and A. Dembo, *Ferromagnetic Ising measures on large locally tree-like graphs.* Preprint, arXiv:1205.4749, (2012).

[3] B. Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs.* European Journal of Combinatorics 1 (1980), 311–316.

[4] B. Bollobás, *The isoperimetric number of random regular graphs.* European Journal of Combinatorics 9/3 (1988), 241–244.

[5] A. Dembo and A. Montanari, *Ising models on locally tree-like graphs.* The Annals of Applied Probability 20/2 (2010), 565–592.

[6] S. Dommers, *Metastability of the Ising model on random regular graphs at zero temperature.* Preprint, arXiv:1411.6802 (2014).

[7] S. Dommers, C. Giardinà and R. van der Hofstad, *Ising models on power-law random graphs.* Journal of Statistical Physics 141/4 (2010),638–660.

[8] S. Dommers, C. Giardinà and R. van der Hofstad, *Ising critical exponents on random trees and graphs.* Communications in Mathematical Physics 328/1 (2014), 355–395.

[9] P. Erdős and A. Rényi, *On the evolution of random graphs.* Publications of the Mathematical Institute of the Hungarian Academy of Science 5 (1960), 17–61.

[10] R. van der Hofstad, "Random graphs and complex networks". Lecture notes available at the URL `http://www.win.tue.nl/~rhofstad/NotesRGCN.html` (2014).

[11] F. Manzo, F.R. Nardi, E. Olivieri and E. Scoppola, *On the essential features of metastability: tunnelling time and critical configurations.* Journal of Statistical Physics 115/1-2 (2004), 591–642.

[12] M. E. J. Newman, *The structure and function of complex networks.* SIAM Review 45/2 (2003),167–256.

[13] M. Niss, *History of the Lenz–Ising model 1920–1950: from ferromagnetic to cooperative phenomena.* Archive for History of Exact Sciences 59/3 (2005), 267–318.

[14] M. Niss, *History of the Lenz–Ising model 1950–1965: from irrelevance to relevance.* Archive for History of Exact Sciences 63/3 (2009), 243–287.

[15] M. Niss, *History of the Lenz–Ising model 1965–1971: the role of a simple model in understanding critical phenomena.* Archive for History of Exact Sciences 65/6 (2011), 625–658.

Automorphism-invariant modules

Nguyen Khanh Tung (*)

Abstract. In this note we mention the class of injective modules, the class of quasi-injective modules and their generalization, the class of automorphism-invariant modules. Then we give some results related to the endomorphism rings of automophism-invariant modules and their injective envelopes and show a connection between automorphism-invariant modules and bolean rings.

1 Introduction

The important of injective modules in Module Theory and more generally in Algebra became obvious in the 1960s and 1970s largely through the lecture note of Carl Faith [10]. Since that time there has been a continuing interest in such modules and their various generalization which arose directly from the study of injective modules. Many results obatined for injective modules could be transfered readily to quasi-injective modules. One of generalizations of the class of quasi-injective modules is the class of automorphism-invariant modules. A module M is called *automorphism-invariant* if it is invariant under automorphisms of its injective envelope, that is, if $\varphi(M) \subseteq M$ for every $\varphi \in \mathrm{Aut}(E(M))$ (equivalently, if $\varphi(M) = M$ for every $\varphi \in \mathrm{Aut}(E(M))$). It is showed that every direct summand of automorphism-invariant modules is automorphism-invariant and the class of automorphism-invariant modules satisfies Condition (C_2) and (C_3). However, an automorphism-invariant module satisfy Condition (C_1) if and only if it is quasi-injective. Moreover, we will see that if M is an automorphism-invariant module, then

$$\mathrm{End}(M)/J(\mathrm{End}(M))$$

turns out to be a rationally closed subring of

$$\mathrm{End}(E(M))/J(\mathrm{End}(E(M))).$$

Both the rings $\mathrm{End}(M)/J(\mathrm{End}(M))$ and $\mathrm{End}(E(M))/J(\mathrm{End}(E(M)))$ are von Neumann regular [12, Proposition 1]. We consider in particular the case of automorphism-invariant modules of finite Goldie dimension or indecomposable. Notice that automorphism-invariant modules have the exchange property [12], so that indecomposable automorphism-invariant

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: khanhtung06@yahoo.com . Seminar held on February 25th, 2015.

modules have a local endomorphism ring. Furthermore, idempotents can be lifted modulo every right ideal both in $\mathrm{End}(M)$ and in $\mathrm{End}(E(M))$ [17]. We then study the connection between automorphism-invariant modules and boolean rings. The existence of such a connection is from the results in Section 5 of [20], where Vámos considers modules whose endomorphism ring (or endomorphism ring modulo the Jacobson radical) is a boolean ring. He studies modules in which the identity endomorphism is the sum of two automorphisms. This condition is related to the existence of factors of the endomorphism ring isomorphic to the field $\mathbb{F}_2$ with two elements [13]. Notice that if M is an automorphism-invariant right R-module and $\mathrm{End}(M)$ has no factor isomorphic to $\mathbb{F}_2$, then M is quasi-injective [11, Theorem 3]. Every automorphism-invariant module is the direct sum of a quasi-injective module and a square-free module [7, Theorem 3]. This leads us to study, for automorphism-invariant square-free modules M, the relation between M being quasi-injective and the existence of factors isomorphic to $\mathbb{F}_2$ in $\mathrm{End}(M)$ and in $\mathrm{End}(E(M))$.

Throughout, all rings have identity element and modules are right unital. For a module M, $E(M)$ denotes the injective envelope of M.

2 Injective and Quasi-Injective Modules

In this section, we introduduce injective and quasi-injective modules as well as their basic properties that will be necessary to study automorphism-invariant modules. All results are well-known and can be found in most text-book of ring theory or module theory. We refer the reader to [1] and [14].

Definition 2.1 Let E_R, M_R be two modules. We say that E_R is *injective relative* to M_R (or E_R is M-injective) if, for each monomorphism $f : K_R \to M_R$ and each morphism $g : K_R \to E_R$, there exists a morphism $h : M_R \to E_R$ with $h \circ f = g$.

Definition 2.2 A module E is *injective* if E is injective relative to every module, that is, for every monomorphism $f : K \to M$ and every morphism $g : K \to E$, there exists a morphism $h : M \to E$ such that $g = h \circ f$. A ring R is *right selft-injective* if R_R is injective.

In the next Proposition, we give a further criterion to determine injective modules, that is, a further characterization of injective modules.

Proposition 2.3 (Baer's criterion) *The following about a module E are equivalent:*

(a) *E is injective.*

(b) *E is injective relative to R.*

(c) *For every right ideal $I \leq R_R$ and every morphism $h : I \to E$ there exists an $x \in E$ such that $h(a) = xa$ $(a \in I)$.*

Example 2.4 An abelian group G is *divisible* if $nG = G$ for every non zero integer n. Hence G is divisible if and only if, for every $g \in G$ and every $n > 0$, there exists $h \in G$

such that $g = nh$. A $\mathbb{Z}$-module G is injective if and only if it is a divisible abelian group. For instance, the abelian group $\mathbb{Z}$ is not divisible, and the abelian group $\mathbb{Q}$ is divisible. Hence, $\mathbb{Q}$ is injective.

Definition 2.5 An *injective envelope* of a module M_R is a pair (E_R, i) where E_R is an injective right R-module and $i : M_R \to E_R$ is an essential monomorphism.

Theorem 2.6 *Every right R-module has a unique injective envelope up to isomorphism.*

Definition 2.7 A module M_R is called *quasi-injective* if M is M-injective.

Theorem 2.8 (R. E. Johnson and E. T. Wong, 1961) *Let M be a module. Then M is quasi-injective if and only if it is invariant under every endomorphism of $E(M)$.*

Example 2.9 Let $R = \mathbb{Z}/4\mathbb{Z}$ and $M = 2\mathbb{Z}/4\mathbb{Z}$. Then M is quasi-injective but not injective.

3 Automorphism-Invariant Modules

The aim of this section is to give a presentation of some known and new results on automorphism-invariant modules and related notions.

Definition 3.1 A module M is called *automorphism-invariant* if it is invariant under automorphisms of its injective envelope, that is, if $\varphi(M) \subseteq M$ for every $\varphi \in \mathrm{Aut}(E(M))$ (equivalently, if $\varphi(M) = M$ for every $\varphi \in \mathrm{Aut}(E(M))$).

It is showed that a module M is quasi-injective if and only if it is invariant under endomorphisms of its injective envelopes (Theorem 2.8). So every quasi-injective is automorphism-invariant. But the converse is not true by the following example.

Example 3.2 Let R consists of all $(x_n)_{n\in\mathbb{N}} \in \prod_{n\in\mathbb{N}} \mathbb{Z}_2$ such that all except finitely many x_n are equal to some $a \in \mathbb{Z}_2$. Then R is a ring, and $E(R_R) = \prod_{n\in\mathbb{N}} \mathbb{Z}_2$. Because $End(S_R)$ has only one automorphism, namely the identity, R is automorphism-invariant but it is not quasi-injective.

Definition 3.3 A module M satisfies Condition (C_1) if every submodule of M is essential in a direct summand of M.

Definition 3.4 A module M satisfies Condition (C_2) if every submodule of M isomorphic to a direct summand of M is also a direct summand of M.

Definition 3.5 A module M satisfies Condition (C_3) if, for any two direct summands N_1, N_2 of M with $N_1 \cap N_2 = 0$, the direct sum $N_1 \oplus N_2$ is a direct summand of M.

It is well known that every injective (quasi-injective) module satisfies Condition (C_i) $(i = 1, 2, 3)$ (see [14] or [16]). The following results show that automorphism-invariant modules always satisfy Condition (C_2) and (C_3), but do not satisfy Condition (C_1) in general.

Proposition 3.6 (T.-K. Lee and Y. Zhou, 2013) *An automorphism-invariant module satisfies Condition* (C_1) *if and only if it is quasi-injective.*

Proposition 3.7 (T.-K. Lee and Y. Zhou, 2013) *Every automorphism-invariant module satisfies Condition* (C_3).

Definition 3.8 A module M is called pseudo-injective if, for any submodule A of M, every monomorphism $f : A \to M$ can be extended to an element of $End(M)$.

Theorem 3.9 (N. Er, S. Singh and A. K. Srivastava, 2013) *Automorphism-invariant modules are precisely pseudo-injective modules.*

Proposition 3.10 (H. Q. Dinh, 2005) *Every pseudo-injective satisfies Condition* (C_2).

Corollary 3.11 *Every automorphism-invariant satisfies Condition* (C_2).

As the cases of injective and quasi-injective modules, a direct summand of an automorphism-invariant module inherits the property by the next proposition.

Proposition 3.12 (T.-K. Lee and Y. Zhou, 2013) *Every direct summand of an automorphism-invariant is automorphism-invariant.*

Before stating some results about the endomorphism rings of automorphism-invariant modules, we review some necessary concepts.

Definition 3.13 A ring morphism $\varphi\colon R \to S$ is *local* if, for every $r \in R$, $\varphi(r)$ invertible in S implies r invertible in R.

Proposition 3.14 *A ring morphism* $\varphi\colon R \to S$ *is* local *if and only if* $\operatorname{Ker}\varphi \leq J(R)$ *and* $\varphi(U(R)) = \varphi(R) \cap U(S)$ *where* $U(R)$ *and* $U(S)$ *is the group of invertible elements of rings* R, S *respectively.*

Definition 3.15 A *rationally closed* subring of a ring S is a subring R of S such that the embedding $R \to S$ is a local morphism, that is, a subring R of S such that $U(R) = R \cap U(S)$.

Let M be a right R-module. From now we denote $E(M)$ the injective envelope of M. Let Δ be denote the set of all endomorphisms with essential kernel. The next proposition characterizes the Jacobson radical of the endomorphism rings of automorphism-invariant modules.

Proposition 3.16 (P. A. Guil Asensio and A. K. Srivastava, 2013) *Let M be an automorphism-invariant module. Then the Jacobson radical of $End(M)$ is Δ, $End(M)/J(End(M))$ is a von Neumann regular ring and idempotents can be lifted modulo $J(End(M))$.*

More properties about the endomorphism of automorphism-invariant modules are provided by the following results.

Theorem 3.17 (A. Alahmadi and A. Facchini and N. K. Tung, 2015) *Let M be an automorphism-invariant module and $E(M)$ be its injective envelope. Then*

(a) *There is a canonical local morphism*

$$\varphi\colon \mathrm{End}(M) \to \mathrm{End}(E(M))/J(\mathrm{End}(E(M)))$$

with kernel $J(\mathrm{End}(M))$, so that φ induces an embedding $\overline{\varphi}$, as a rationally closed subring, of the von Neumann regular ring $\mathrm{End}(M)/J(\mathrm{End}(M))$ into the von Neumann regular right self-injective ring

$$\mathrm{End}(E(M))/J(\mathrm{End}(E(M))).$$

(b) *For every invertible element v of the ring $\mathrm{End}(E(M))/J(\mathrm{End}(E(M)))$, there exists an invertible element u of $\mathrm{End}(M)/J(\mathrm{End}(M))$ such that $\overline{\varphi}(u) = v$.*

(c) *For every idempotent element f of the ring $\mathrm{End}(E(M))/J(\mathrm{End}(E(M)))$ there exists an idempotent element e of $\mathrm{End}(M)/J(\mathrm{End}(M))$ such that $\overline{\varphi}(e) = f$ if and only if the module M is quasi-injective.*

(d) *If M is quasi-injective, then $\overline{\varphi}$ is an isomorphism.*

Proposition 3.18 (A. Alahmadi and A. Facchini and N. K. Tung, 2015) *Let M be an automorphism-invariant module. Then*

(a) *If M is indecomposable, then $\mathrm{End}(M)$ is a local ring.*

(b) *If M has finite Goldie dimension, then every injective endomorphism of M is an automorphism of M and the endomorphism ring $\mathrm{End}(M)$ is a semiperfect ring.*

Corollary 3.19 (A. Alahmadi and A. Facchini and N. K. Tung, 2015) *If M, N are two automorphism-invariant R-modules of finite Goldie dimensions isomorphic to submodules of each other, then M is isomorphic to N.*

Proposition 3.20 (A. Alahmadi and A. Facchini and N. K. Tung, 2015) *The following conditions are equivalent for a ring R.*

(a) *Every automorphism-invariant R-module of finite Goldie dimension is quasi-injective.*

(b) *Every automorphism-invariant indecomposable R-module of finite Goldie dimension is uniform.*

(c) *Every automorphism-invariant indecomposable R-module of finite Goldie dimension is quasi-injective.*

A non-zero ring R is a boolean ring if every element of R is idempotent. A ring is boolean if and only if it is isomorphic to a subring of $\mathbb{F}_2^X$, where X is a non-empty set. The next theorem gives us a sufficient condition so that an automorphism-invariant module is quasi-injective.

Theorem 3.21 (P. A. Guil Asensio and A. K. Srivastava, 2014) *Let M be a right module such that $\mathrm{End}(M)$ has no factor isomorphic to $\mathbb{F}_2$. Then M is quasi-injective if and only if M is automorphism-invariant.*

Recall that two modules are said to be orthogonal if they do not contain nonzero isomorphic submodules.

Proposition 3.22 (A. Alahmadi and A. Facchini and N. K. Tung, 2015) *Let $M = M_1 \oplus M_2$ be an automorphism-invariant R-module where M_1 and M_2 are orthogonal. Then $\mathrm{End}(M)$ has no factor isomorphic to $\mathbb{F}_2$ if and only if each $\mathrm{End}(M_i)$ $(i = 1, 2)$ has no factor isormophic to $\mathbb{F}_2$.*

Definition 3.23 A module M is square-free if it does not contain a direct sum of two non-zero isomorphic submodules.

Theorem 3.24 (N. Er, S. Singh and A. K. Srivastava, 2013) *Every automorphism-invariant module M decomposes as a direct sum $M = X \oplus Y$, where X is quasi-injective, Y is a square-free module orthogonal to X, and X and Y are relatively injective modules.*

The previous theorem reduces studying automorphism-invariant modules to considering automophism-invariant square-free modules.

Lemma 3.24.1 (A. Alahmadi and A. Facchini and N. K. Tung, 2015) *If M_1, M_2 are two right modules over a ring R and M_1, M_2 have isomorphic injective envelopes, which are non-zero modules, then M_1 and M_2 have non-zero isomorphic submodules.*

Corollary 3.25 (A. Alahmadi and A. Facchini and N. K. Tung, 2015) *If M is an automorphism-invariant square-free module, then every injective endomorphism of M is an automorphism of M.*

Corollary 3.26 (A. Alahmadi and A. Facchini and N. K. Tung, 2015) *If M, N are two automorphism-invariant square-free R-modules isomorphic to submodules of each other, then M is isomorphic to N.*

Proposition 3.27 (A. Alahmadi and A. Facchini and N. K. Tung, 2015) *Let M be an automorphism-invariant module and $E(M)$ be its injective envelope. The following conditions are equivalent:*

(a) *M is square-free.*

(b) *$E(M)$ is square-free.*

(c) *The von Neumann regular ring $\mathrm{End}(M)/J(\mathrm{End}(M))$ is abelian.*

(d) *The von Neumann regular right self-injective ring $\mathrm{End}(E(M))/J(\mathrm{End}(E(M)))$ is abelian.*

Theorem 3.28 (A. Alahmadi and A. Facchini and N. K. Tung, 2015) *Let M be an automorphism-invariant module and let $E(M)$ be its injective envelope.*

(a) *If M is quasi-injective and $\mathrm{End}(M)$ has a factor isomorphic to $\mathbb{F}_2$, then $\mathrm{End}(E(M))$ has a factor isomorphic to $\mathbb{F}_2$.*

(b) *If M has finite Goldie dimension and $\mathrm{End}(M)$ has a factor isomorphic to $\mathbb{F}_2$, then the following conditions hold.*

  (i) *$\mathrm{End}(E(M))$ has a factor isomorphic to $\mathbb{F}_2$.*

  (ii) *$E(M)$ has a direct-sum decomposition $E(M) = E \oplus C$ with E orthogonal to C, E an indecomposable R-module and $\mathrm{End}(E)/J(\mathrm{End}(E)) \cong \mathbb{F}_2$.*

  (iii) *$\mathrm{Aut}(E) = 1 + J(\mathrm{End}(E))$, so that every automorphism of the R-module E is the identity on an essential R-submodule of E.*

  (iv) *E is the injective envelope of its non-zero R-submodule $\mathrm{ann}_E(2)$.*

References

[1] D. W. Anderson and K. R. Fuller, "Rings and Categories of Modules". Second Edition, GTM **13**, Springer-Verlag, New York, 1992.

[2] A. Alahmadi, A. Facchini and Nguyen Khanh Tung, *Automorphism-invariant modules*. To appear in Rend. Sem. Mat. Univ. Padova (2015).

[3] B. Brainerd and J. Lambek, *On the ring of quotients of a boolean ring*. Canad. Math. Bull. 2 (1959), 25–29.

[4] T. T. Bumby, *Modules which are isomorphic to submodules of each other*. Arch. Math. 16 (1965), 184–185.

[5] R. Camps and W. Dicks, *On semilocal rings*. Israel J. Math. 81 (1993), 203–211.

[6] H. Q. Dinh, *A note on pseudo-injective modules*. Comm. Algebra 33 (2005), 361–369.

[7] N. Er, S. Singh and A. K. Srivastava, *Rings and modules which are stable under automorphisms of their injective hulls*. J. Algebra 379 (2013), 223–229.

[8] A. Facchini, "Module Theory. Endomorphism rings and direct sum decompositions in some classes of modules". Progress in Math. **167**, Birkhäuser Verlag, Basel, 1998. Reprinted in Modern Birkhäuser Classics, Birkhäuser Verlag, Basel, 2010.

[9] A. Facchini and D. Herbera, *Local morphisms and modules with a semilocal endomorphism ring*. Algebr. Represent. Theory 9 (2006), 403–422.

[10] C. Faith, "Lectures on injective modules and quotient rings". Springer LNM 49 (1967).

[11] P. A. Guil Asensio and A. K. Srivastava, *Additive unit representations in endomorphism rings and an extension of a result of Dickson and Fuller*. In: "Ring Theory and its Applications", Contemp. Math. 609, Amer. Math. Soc., Providence 2014, D. V. Huynh, et al., eds., pp. 117–121.

[12] P. A. Guil Asensio and A. K. Srivastava, *Automorphism-invariant modules satisfy the exchange property*. J. Algebra 388 (2013), 101–106.

[13] D. Khurana and A. K. Srivastava, *Right self-injective rings in which every element is a sum of two units*. J. Algebra Appl. 6 (2007), 281–286.

[14] T. Y. Lam, "Lectures on Modules and Rings". Springer-Verlag, New York, 1998.

[15] T.-K. Lee and Y. Zhou, *Modules which are invariant under automorphisms of their injective hulls*. J. Algebra Appl. 12 (2013).

[16] S. H. Mohamed and B. J. Müller, "Continuous and Discrete Modules". London Mathematical Society, Lecture Notes Series 147, Cambridge University Press, 1990.

[17] W. K. Nicholson, *Lifting idempotents and exchange rings*. Trans. Amer. Math. Soc. 229 (1977), 269–278.

[18] S. Singh and A. K. Srivastava, *Rings of invariant module type and automorphism-invariant modules*. Contemp. Math. AMS, in press; available in `http://arxiv.org/pdf/1207.5370.pdf`.

[19] B. Stenström, "Rings of quotients". Springer-Verlag, Berlin-Heidelberg-New York, 1975.

[20] P. Vámos, *2-good rings*. Quart. J. Math. 56 (2005), 417–430.

Introduction to kernel-based methods

Gabriele Santin [*]

This note is a brief introduction to the theory of kernel based methods, with particular focus on approximation in kernel based spaces.

Kernel-based methods are emerging in various fields of applied mathematics as a technique to solve several kind of problems. Kernel-based algorithm are successfully applied in approximation theory, in machine learning, in geostatistics and in different type of numerical simulations in engineering. The use of such methods is motivated by reasons that vary from field to field, but it is mainly due to their flexibility and to the simple structure that is shared by all the particular algorithms. In particular, in the field of approximation theory, kernel-based methods can be proven to be error optimal both in solving interpolation problems and recovery of solution of certain PDEs.

In the following we will give an introduction to the use of kernels. The focus will be on approximation, but the underlying theory is the same used in all the applications. Further details on the topics presented in this note can be found in the monographs [1, 2, 3].

1 Statement of the problem and motivation

We are interested in the following recovery problem in $\mathbb{R}^d$. We are given a set of $n \in \mathbb{N}$ pairwise distinct points $X_n = \{x_1, \dots, x_n\} \subset \Omega \subset \mathbb{R}^d$, where Ω is a bounded set, and a vector of values $F_n = [f_1, \dots, f_n]^T \in R^n$. The vector F_n is understood as a collection of sampling of an unknown continuous function $f : \Omega \to \mathbb{R}$, but for now we consider only the n values f_i.

Given this set of data, we consider a linear scheme to find a continuous function $s_n : \Omega \to \mathbb{R}$ that *interpolates the data*, i.e., $s(x_i) = f_i$, $1 \leq i \leq n$. To do so, one considers a suitable n-dimensional *trial space* $V = \operatorname{span}\{v_i, \dots, v_n\}$, where $v_j : \Omega \to \mathbb{R}$ are linearly independent continuous functions, and an interpolant

$$s_n = \sum_{j=1}^{n} c_j v_j,$$

for an unknown coefficient vector $c = [c_1, \dots, c_n]^T \subset \mathbb{R}^n$ such that $s_n(x_i) = f_i$, $1 \leq i \leq n$.

[*]Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: gsantin@math.unipd.it . Seminar held on March 18th, 2015.

A minimal requirement for a scheme to be well posed is the existence of a unique function $s_n \in V$ that interpolates the data, for any given set of data. This is equivalent to require that V is an Haar space.

Definition 1 An n-dimensional linear space $V \subset \mathcal{C}(\Omega)$ is an *Haar space* of dimension n on Ω if for any set X_n of pairwise distinct point in Ω and any set of values $F_n \in \mathbb{R}^n$ there exist a unique function $s_n \in V$ such that $s_n(x_i) = f_i$, $1 \leq i \leq n$.

A typical example of an Haar space of dimension n in $\mathbb{R}$ is the space of univariate real polynomials of degree $n-1$, that is to say, the $(n-1)$-degree polynomial interpolant on a set of n distinct points is always uniquely defined.

The situation is completely different if $d > 1$.

Theorem 1 (Mairhuber - Curtis) *If $\Omega \subset \mathbb{R}^d$, $d > 1$, contains an interior point, there exists no Haar spaces on Ω of dimension $n > 1$.*

The theorem tells us that for $d > 1$, except for the trivial case $n = 1$, one can not choose a fixed trial space V, but it is necessary to adapt it to each particular set of data. A strategy to overcome this problem comes from the use of kernels and we will see that, in some circumstances, it is also the unique optimal strategy.

2 From kernel interpolation to Hilbert spaces

A *kernel* K on a set S is a positive definite and symmetric function $K : S \times S \to \mathbb{R}$, where positive definiteness is defined in the following sense.

Definition 2 A function $K : S \times S \to \mathbb{R}$ is *positive definite* if, for any $n \in \mathbb{N}$, for any set $\{x_1, \ldots, x_n\} \subset S$ of pairwise distinct elements and for any vector of coefficients $c = [c_1, \ldots, c_n]^T \in \mathbb{R}^n$,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j K(x_i, x_j) \geq 0.$$

If the above expression is zero only for $c = 0$, the function is *strictly positive definite*.

Although the theory can be developed for positive definite kernels, we will restrict the following discussion to *strictly* positive definite kernels, to which we will refer simply as *kernels*. Most of the results are true in both cases, but dealing with vanishing bilinear forms requires more technical details. Moreover, in this note we will only consider continuous kernels defined on bounded subspaces Ω of $\mathbb{R}^d$, even if kernels can be constructed on abstract sets S with almost no structure (e.g., kernels defined on sets of strings or on trees and graphs are commonly used in Machine Learning; also, complex valued kernels are studied in differential geometry and analysis).

To have a concrete example in mind, we recall that the Gaussian kernel $K(x, y) = e^{-\varepsilon^2 \|x-y\|}$, $\varepsilon > 0$, is in fact a kernel on $\mathbb{R}^d$, for all d.

2.1 Kernel-based interpolation

Given a kernel, the interpolation problem can be simply solved in the following way. We can build point-dependent trial spaces through translates of the kernel, i.e.,

$$V(X_n) = \text{span}\{K(\cdot, x_i) : 1 \leq i \leq n\},$$

hence look for interpolants in the form

$$s_n = \sum_{j=1}^{n} c_j K(\cdot, x_j),$$

for some unknown vector of coefficients $c \in \mathbb{R}^n$. The vector c is determined by the interpolation conditions $s_n(x_i) = f_i$, $1 \leq i \leq n$, i.e., by the solution of the linear system

$$Ac = F_n,$$

where the *kernel matrix* $A = [K(x_i, x_j)]_{i,j=1}^n \in \mathbb{R}^{n \times n}$ is symmetric and positive definite by the definition of kernel, and it is in particular invertible for any set of pairwise distinct points. In other words, the interpolation problem is well posed for any data.

We just mention here that the actual computation of the solution of the above linear system is in general an hard task since the kernel matrix is usually ill-conditioned, and it should not be solved directly. Nevertheless, several methods are available to compute c in a stable way, and the topic is currently an active field of research.

2.2 Kernel-based spaces

What we described so far is just a procedure to construct a continuous function with prescribed values at prescribed points. Now we would like to see if this is an effective approximation method. Namely, we assume that the data values are samples of a function $f : \Omega \to \mathbb{R}$, i.e., $f_i = f(x_i)$, $1 \leq i \leq n$, and we will investigate the relation between f and s_n. In particular, we will characterize the space of functions that can be recovered with arbitrary accuracy by the above procedure, provided n is sufficiently large.

First, observe that for any $x \in \Omega$ the functions $K(\cdot, x) : \Omega \to \mathbb{R}$ and all the finite linear combination of functions of this kind can be exactly recovered, simply by choosing as X_n all the points used to construct the kernel translates. We denote as $\mathcal{H}_0(\Omega)$ the set of all these functions, i.e.,

$$\mathcal{H}_0(\Omega) = \text{span}\{K(\cdot, x) : x \in \Omega\}.$$

On $\mathcal{H}_0(\Omega)$ is possible to define the bilinear form

$$\left(\sum_{j=1}^{n} c_j K(\cdot, x_j), \sum_{i=1}^{m} d_i K(\cdot, y_i) \right)_K := \sum_{j=1}^{n} \sum_{i=1}^{m} c_j d_i K(x_j, y_i).$$

and since K is a kernel, and it is in particular symmetric and strictly positive definite, this bilinear form defines an inner product on $\mathcal{H}_0(\Omega)$.

The closure of $\mathcal{H}_0(\Omega)$ with respect to the norm induced by $(\cdot,\cdot)_K$ is an Hilbert space $\mathcal{H}_K(\Omega)$ of functions on Ω to $\mathbb{R}$, and it is called the *native space* of the kernel K on Ω. We will still denote by $(\cdot,\cdot)_K$ its inner product.

The native space of K is in fact the natural space where kernel approximation takes place, and in enjoys some useful property that are listed in the following theorem.

Theorem 2 *The space $\mathcal{H}_K(\Omega)$ is a* reproducing kernel Hilbert space *(RKHS) with kernel K, i.e.,*

(a) $K(\cdot, y) \in \mathcal{H}_K(\Omega)$ *for all* $y \in \Omega$,

(b) $(f, K(\cdot, x))_K = f(x)$ *for all* $f \in \mathcal{H}_K(\Omega)$, $x \in \Omega$,

and it is the unique RKHS with kernel K on Ω. Moreover,

(c) *If* $f \in \mathcal{H}_K(\Omega)$ *then* $\|f - s_n\|_K \to 0$ *as* $n \to \infty$,

(d) *if there exists a positive constant C_f such that $\|s_n\|_K \leq C_f$ for any finite set $X_n \subset \Omega$, for any $n \in \mathbb{N}$, then $f \in \mathcal{H}_K(\Omega)$ and $\|f\|_K$ is the minimal possible value for the constant C_f.*

We now know how to solve the interpolation problem, how to build an Hilbert space of functions starting from a kernel, and that this space contains precisely (in the sense of properties (c), (d)) all the functions that can be approximated by the kernel method.

A question that naturally arises is if this is the only way to solve the interpolation problem. We will see in the next Section that one can in fact go the other way round, i.e., start with a quite general Hilbert space and prove that kernel interpolation is the optimal recovery strategy in that space.

Before moving to the next Section, observe that working in a RKHS is very useful to construct numerical approximation schemes. First, to compute norms of functions in the dense subspace $\mathcal{H}_0(\Omega)$ it is enough to evaluate the kernel at some points. Second, thanks to property (b) of Theorem 2, norm convergence in $\mathcal{H}_K(\Omega)$ implies pointwise convergence. Indeed, if $f, g \in \mathcal{H}_K(\Omega)$,

$$\begin{aligned} |(f-g)(x)| = |(f-g, K(\cdot,x)_K| &\leq \|f-g\|_K \|K(\cdot,x)\|_K \\ &\leq \|f-g\|_K \sqrt{K(x,x)} \quad \text{for all } x \in \Omega. \end{aligned}$$

3 From Hilbert spaces to kernel interpolation

We now assume to have an Hilbert space $\mathcal{H}$ of functions from Ω to $\mathbb{R}$, and we want to interpolate a given function in $\mathcal{H}$ at prescribed, but arbitrary, points $X_n \subset \Omega$. Given this task, since we want to recover a function from its pointwise samples, it make sense to assume that the pointwise evaluation functionals δ_x are continuous in $\mathcal{H}$. We are then in the following situation.

Theorem 3 *A Hilbert space $\mathcal{H}$ of functions from Ω to $\mathbb{R}$ is a reproducing kernel Hilbert space if and only if the pointwise evaluation functionals δ_x are continuous for any $x \in \Omega$.*

In this case, the kernel of the space is the Riesz representer of these functional, that is to say,

$$\delta_x(f) = (f, K(\cdot, x))_{\mathcal{H}} \text{ for all } f \in \mathcal{H},\ x \in \Omega.$$

Moreover, the kernel is strictly *positive definite if and only if all the functionals* $\{\delta_x : x \in \Omega\}$ *are linearly independent.*

As the Theorem proves, kernel-based spaces are quite common. Examples of such spaces are the following:

- any Hilbert space of finite dimension N is a RKHS. If $\{v_j\}_{j=1}^N$ is any orthonormal basis of this space, the kernel is $K(x,y) = \sum_{j=1}^N v_j(x)v_j(y)$, $x, y \in \Omega$;
- the Sobolev space $H_0^1((0,1))$ is a RKHS. Its kernel is the *Brownian bridge kernel* $K(x,y) = \min(x,y) - xy$, $x, y \in (0,1)$;
- the Sobolev spaces $H^\beta(\mathbb{R}^d)$, $\beta > d/2$ are RKHS, whose kernels are the *β-Matérn kernels*, which can be expressed in terms of Bessel's functions.

Under the assumption of continuity of the functionals δ_x, we are left to work in the native space of a given kernel K. We will hence still use the notation $\mathcal{H}_K(\Omega)$ for this space and $(\cdot,\cdot)_K$ for its inner product.

3.1 Optimal recovery

Going back to the interpolation problem, it is clear that the method described in the previous sections is well defined here. In particular, the trial space $V(X_n)$ is a finite dimensional subspace of $\mathcal{H}_K(\Omega)$, and we can define an interpolation operator

$$\begin{aligned} S_n : \mathcal{H}_K(\Omega) &\to V(X_n) \\ f &\mapsto s_n. \end{aligned}$$

where the interpolating function s_n is defined as above. The next Theorem shows that this is in fact the unique optimal interpolation strategy in $\mathcal{H}_K(\Omega)$.

Theorem 4 *The kernel interpolant s_n is the minimal norm interpolant in $\mathcal{H}_K(\Omega)$, i.e.,*

$$s_n = \min\{\|t_n\|_K : t_n \in \mathcal{H}_K(\Omega), t_n(x_i) = f(x_i), 1 \le i \le n\}.$$

Moreover, the interpolation operator S_n coincides with the $\mathcal{H}_K(\Omega)$-projection into the subspace $V(X_n)$, thus, in particular, the interpolant is also the $\|\cdot\|_K$-best approximant of f.

3.2 The power function and a simple error estimate

To conclude this note, we recall the basic error estimate on kernel approximation.

Namely, although we introduced this method in order to being able to deal with scattered data, there is obviously a dependence of the approximation error upon the distribution of the points in X_n. To measure this distribution it is customary to introduce the *fill distance* $h_{X_n,\Omega}$,

$$h_{X_n,\Omega} = \sup_{x\in\Omega} \min_{x_i\in X_n} \|x - x_i\|,$$

which corresponds to the radius of the biggest ball that can be put in Ω without points of X_n.

The error estimate is computed by bounding the *power function* $\mathcal{P}_n$, that is the norm of the pointwise approximation error. Namely, since the interpolation operator is a bounded operator in $\mathcal{H}_K(\Omega)$, we can compute the norm of the functional $\mathcal{E}_x := \delta_x \circ (id - S_n)$,

$$\begin{aligned} \mathcal{E}_x : \mathcal{H}_K(\Omega) &\to \mathcal{H}_K(\Omega) \\ f &\mapsto f(x) - s_n(x). \end{aligned}$$

The norm of this functional is the power function, and it can be computed as

$$\mathcal{P}_n^2(x) = K(x,x) - \sum_{j=1}^{n} v_j(x)^2,\ x \in \Omega,$$

where $\{v_j\}_{j=1}^n$ is any orthonormal basis of $V(X_n)$. Of course, for any $f \in \mathcal{H}_K(\Omega)$, it holds

$$|f(x) - s_n(x)| \le \mathcal{P}_n(x)\|f\|_K,\ \ x \in \Omega.$$

The next Theorem uses this relation and a bound on the power function to have a bound on the pointwise approximation error.

Theorem 5 *If Ω is bounded and satisfies an interior cone condition, and $K \in \mathcal{C}^{2k}(\Omega\times\Omega)$, there exists positive constants h_0 and C such that, for all X_n with $h_{X_n,\Omega} \le h_0$,*

$$|f(x) - s_n(x)| \le \mathcal{C} h_{X_n,\Omega}^k \|f\|_K,\ \ x \in \Omega.$$

References

[1] M. Buhmann, "Radial Basis Functions, Theory and Implementations". Cambridge University Press, 2003.

[2] G. F. Fasshauer, "Meshfree Approximation Methods with MATLAB". World Scientific Publishers, Singapore, 2007.

[3] H. Wendland, "Scattered Data Approximation". Cambridge University Press, Cambridge, 2005.

A short introduction to Berkovich affine line over the field $\mathbb{C}_p$

Velibor Bojković (*)

Abstract. In this note we present a construction of the Berkovich affine line over the field $\mathbb{C}_p$, which is a p-adic analogue of the complex plane $\mathbb{C}$. We emphasize on how one can visualize its tree-structure as well as explain what are its building blocks: both as a set and as a topological space. We end by describing a connection between general Berkovich compact, connected, smooth curves over $\mathbb{C}_p$ and (the reduction of) their semistable models over $\overline{\mathbb{F}}_p$.

1 Introduction: from $\mathbb{Q}$ to $\mathbb{C}_p$

1.1 p-adic norm on $\mathbb{Q}$

Let us start by recalling the definition of a norm on $\mathbb{Q}$.

Definition 1.1 A norm on $\mathbb{Q}$ is a mapping $|\cdot| : \mathbb{Q} \to \mathbb{R}_{\geq 0}$, $x \mapsto |x|$ which satisfies the following three conditions:

(a) $|x| = 0$ iff $x = 0$,

(b) $|xy| = |x||y|$,

(c) $|x + y| \leq |x| + |y|$ for all $x, y \in \mathbb{Q}$.

Example 1.2

(a) We all know the standard absolute value $|\cdot|_\infty$ on $\mathbb{Q}$ ($|-2|_\infty = 2$, $|\frac{9}{5}|_\infty = \frac{9}{5}$).

(b) Let $c \in (0, 1)$, then $|x|_{\infty,c} := |x|_\infty^c$ is also a norm on $\mathbb{Q}$.

(c) The trivial absolute value $|\cdot|_0$ ($|0| = 0$ $|x| = 1$, $x \neq 0$).

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: velibor.bojkovic@gmail.com. Seminar held on April 1st, 2015.

Exercise 1.3 Prove that $|\cdot|_{\infty,c}$ is really a norm on $\mathbb{Q}$.

The main example of a norm on $\mathbb{Q}$ in this note is the so-called p-adic norm, where p is a prime number.

Definition 1.4 Let $p \in \Pi$, where Π is the set of prime numbers, and let $x \in \mathbb{Q}$. Let us write $x = \frac{a}{b} = p^{\alpha}\frac{a'}{b'}$ where $(p, a') = (p, b') = 1$, and $\alpha \in \mathbb{Z}$. We define a function $|\cdot|_p : \mathbb{Q} \to \mathbb{R}_{\geq 0}$ as $x \mapsto |x|_p = p^{-\alpha}$.

Theorem 1.5 *The map $|\cdot|_p$ is a norm on $\mathbb{Q}$.*

Exercise 1.6 Prove the following:

(a) $|\cdot|_p$ is a norm on $\mathbb{Q}$;

(b) For $c \in (0, 1)$, $|\cdot|_{p,c} := |\cdot|_p^c$ is also a norm;

(c) p-adic norm is ultrametric, i.e. it satisfies *the strong triangle inequality* $|x + y|_p \leq \max(|x|_p, |y|_p)$.

One may notice that the norms $|\cdot|_p$ and $|\cdot|)_{p,c}$ are more "similar" than, for example, the norms $|\cdot|_p$ and $|\cdot|_\infty$, the similarity at this point being understood intuitively. For example, the sequence $(p^n)_n$ converges in each of the norms $|\cdot|_{p,c}$ but it doesn't converge in none of the norms $|\cdot|_{\infty,c}$, To make things formal, we recall the following.

Definition 1.7 We say that two norms $|\cdot|_1$ and $|\cdot|_2$ are equivalent if $|x|_1 < 1$ iff $|x|_2 < 1$.

The following celebrated theorem due to Ostrowski classifies all the norms on $\mathbb{Q}$ up to equivalence.

Theorem 1.8 (Ostrowski's Theorem) *Every norm on $\mathbb{Q}$ is either $|\cdot|_0$, or it is equivalent to $|\cdot|_\infty$ or to $|\cdot|_p$ for some $p \in \Pi$.*

1.2 The field $\mathbb{C}_p$

From the first course on mathematical analysis one knows that the fields $\mathbb{R}$ and $\mathbb{C}$ are obtained from $\mathbb{Q}$ by first completing with respect to the standard absolute value $|\cdot|_\infty$ (this gives $\mathbb{R}$) and then taking the algebraic closure of $\mathbb{R}$ (to obtain $\mathbb{C}$). Then, one proves as well that $\mathbb{C}$ is complete.

What happens if we do the same procedure for $\mathbb{Q}$ (*i.e.* completing and then taking the algebraic closure) with respect to the p-adic norm $|\cdot|_p$?

Definition 1.9 We define $\mathbb{Q}_p$ to be completion of $\mathbb{Q}_p$ with respect to the p-adic norm $|\cdot|_p$, and call it the p-adic field.

One may think about field $\mathbb{Q}_p$ as of an analogue of the field $\mathbb{R}$. However, when one takes the algebraic closure of $\mathbb{Q}_p$, the obtained field $\overline{\mathbb{Q}}_p$ doesn't really represent an analogue

of the field $\mathbb{C}$, as the following exercise suggests. Notice that $|\cdot|_p$ canonically extends to $\overline{\mathbb{Q}}_p$, and we call the extension as well the p-adic norm and use $|\cdot|_p$ to denote it.

Exercise 1.10 Prove that the algebraic closure $\overline{\mathbb{Q}}_p$ of $\mathbb{Q}_p$ is not complete.

Definition 1.11 The p-adic completion of the field $\overline{\mathbb{Q}}_p$ is denoted by $\mathbb{C}_p$.

It turns out that the field $\mathbb{C}_p$ is algebraically closed and complete, and as such one may think of it as the p-adic analogue of the field of complex numbers $\mathbb{C}$. But, it seems that all the analogy stops here, as the following short list of properties suggests, and which is left as an exercise.

Exercise 1.12

(a) Prove that the field $\mathbb{C}_p$ equiped with the topology induced by the norm $|\cdot|_p$ is totally disconnected. (Recall that the topological space is said to be totally disconnected if the maximal connected subsets are just points.)

(b) Let $B(a,r)$, where $a \in \mathbb{C}_p$ and $r \in \mathbb{R}_{>0}$ be the closed disc in $\mathbb{C}_p$ with center in a and of radius r, *i.e.* the set of points $\{b \in \mathbb{C}_p, |a-b|_p \leq r\}$. Prove that every point in $B(a,r)$ is a *center* of it.

(c) Let $a, b, c \in \mathbb{C}_p$. Prove that the "triangle" with vertices in a, b, c is isoscele.

(d) Prove that the rim $\{z \in \mathbb{C}_p, |z| = 1\}$ is both open and closed in p-adic topology.

(e) Let $B(0,r) \subset \mathbb{C}_p$ as above. Can you give an example of a *continuous, locally constant function* $f : B(0,1) \to \mathbb{R}$ which is not constant? Can this happen if we take the disc $B(0,r) \subset \mathbb{C}$? Why not?

And now, to the motivational question of this note: How one can do analytic geometry over a field such as $\mathbb{C}_p$? What would for example be an equivalent over $\mathbb{C}_p$ of the classical complex plane/affine line over $\mathbb{C}$?

In the next sections we present a short introduction to the Berkovich theory of analytic spaces over normed fields, which answers in a satisfactory way the previous questions.

2 Berkovich affine line $\mathbb{A}^1_{\mathbb{C}_p}$

2.1 Berkovich spectrum of a normed ring

Definition 2.1 A normed commutative ring A is a commutative ring A equipped with a function $|\cdot| : A \to \mathbb{R}_{\geq 0}$ which satisfies the following conditions:

(a) $|1| = 1$ and $|f| = 0$ iff $f = 0$,

(b) $|f \cdot g| \leq |f| \cdot |g|$,

(c) $|f+g| \leq |f| + |g|$,

Definition 2.2 Let A be a commutative ring (assumed to be with unit). A *multiplicative seminorm* on A is a function $|\cdot| : A \to \mathbb{R}_{\geq 0}$ such that

(a) $|0| = 0$ and $|1| = 1$,

(b) $|f \cdot g| = |f||g|$,

(c) $|f+g| \leq |f| + |g|$.

Definition 2.3 Let A be a normed ring, with norm $||\cdot||$. The Berkovich spectrum of the ring A is the set $\mathcal{M}(A)$ of all the multiplicative seminorms $|\cdot|$ on A which are bounded by the norm $||\cdot||$. We say that $|\cdot|$ is bounded by $||\cdot||$ if there exists a positive real constant C such that for all $f \in A$, $|f| < C||f||$. We equip the set $\mathcal{M}(A)$ with the weakest topology such that the functions, $f : \mathcal{M}(A) \to \mathbb{R}_{\geq 0}$, where $f \in A$ and given by $|\cdot| \in \mathcal{M} \mapsto |f| \in \mathbb{R}_{\geq 0}$ are continuous.

We have the following fundamental theorem by Berkovich (see [1]).

Theorem 2.4 *Let A be a commutative ring. Then,*

(a) *If A is nonzero, $\mathcal{M}(A)$ is nonempty.*

(b) *If A is complete, $\mathcal{M}(A)$ is compact.*

Exercise 2.5 Let A be the ring $\mathbb{Z}$ equipped with the standard norm $|\cdot|_\infty$. Then $\mathbb{Z}$ is a complete ring. Find $\mathcal{M}(\mathbb{Z})$. What is $\mathcal{M}(\mathbb{Z})$ if we equip $\mathbb{Z}$ with the trivial norm?

Remark 2.6 The general idea for doing analytic geometry over the field $\mathbb{C}_p$, which would be analogous to the complex analytic geometry over $\mathbb{C}$, is to start with $\mathbb{C}_p$-rings of functions (in the theory so-called $\mathbb{C}_p$-affinoid algebras) that would correspond to classical complex holomorphic functions, and then to assign Berkovich spectrum to such rings (of course, $\mathbb{C}_p$-affinoid algebras come already with a norm, so there is no problem here).

2.2 The set $\mathbb{A}^1_{\mathbb{C}_p}$

Definition 2.7 The Berkovich affine line over $\mathbb{C}_p$, $\mathbb{A}^1_{\mathbb{C}_p}$, is defined to be the set of multiplicative seminorms on the polynomial ring $\mathbb{C}_p[T]$ extending the p-adic norm of $\mathbb{C}_p$. The topology of $\mathbb{A}^1_{\mathbb{C}_p}$ is the weakest one such that all the functions $f(T) \in \mathbb{C}_p[T]$ which act on $\mathbb{A}^1_{\mathbb{C}_p}$ by $|\cdot| \mapsto |f(T)| \in \mathbb{R}_{\geq 0}$ are continuous.

Remark 2.8 Note a slight difference with respect to taking the Berkovich spectrum of the polynomial ring. First of all, we didn't assume any norm on $\mathbb{C}_p[T]$ so we couldn't just

take $\mathcal{M}(\mathbb{C}_p)$. However, it can be proved that $\mathbb{A}^1_k$ is a union of the Berkovich spectra of $\mathbb{C}_p$-affinoid algebras corresponding to the closed discs in $\mathbb{A}^1_{\mathbb{C}_p}$ (see the section on topology).

Remark 2.9 If we do the same procedure for the ring of polynomials $\mathbb{C}[T]$ we will obtain the complex plain/affine line over $\mathbb{C}$! (Recall Gelfand-Mazur theorem.) This gives a nice similarity between p-adic and complex worlds.

Let us describe some of the points in $\mathbb{A}^1_{\mathbb{C}_p}$ *i.e.* some of the multiplicative seminorms on $\mathbb{C}_p[T]$ which extend the p-adic norm on $\mathbb{C}_p$. As the reader may have already noted, we have a dual point of view on the elements of the set $\mathbb{A}^1_{\mathbb{C}_p}$, namely as on points and as on the multiplicative seminorms. To clarify in which mode we are, and to simplify the notions, we agree that for a point in $x \in \mathbb{A}^1_{\mathbb{C}_p}$ to denote the corresponding seminorm as $|\cdot|_x$.

Example 2.10 Let, as before, $B(a,r) := \{b \in \mathbb{C}_p, |b-a|_p \leq r\}$, where $a \in \mathbb{C}_p, r \geq 0$ (note that for $r = 0$, $B(a,r)$ is just the point a). Then, the map $|\cdot|_{B(a,r)}$ acting on $\mathbb{C}_p[T]$ as $|f|_{B(a,r)} := \sup_{b\in B(a,r)} |f(b)|$ is in $\mathbb{A}^1_{\mathbb{C}_p}$. The point corresponding to the disc $B(0,1)$ is called the Gauss point and is usually denoted by ζ_G.

Example 2.11 More generally, if $B_n := B(a_n, r_n)$, $n \in \mathbb{N}$ is a nested sequence of closed discs in $\mathbb{C}_p$,*i.e.* B_n are closed discs in $\mathbb{C}_p$ and for each $n \in \mathbb{N}$, $B_{n+1} \subseteq B_n$, then the map $|\cdot|_{(B_n)_n} : \mathbb{C}_p[T] \to \mathbb{R}_{\geq}$ given by $|f|_{(B_n)_n} := \lim_n |f|_{B_n}$ for $f \in \mathbb{C}_p[T]$ is a multiplicative seminorm on $\mathbb{C}_p[T]$ whose restriction to $\mathbb{C}_p$ is p-adic norm, and as such is in $\mathbb{A}^1_{\mathbb{C}_p}$.

Exercise 2.12 Prove the statements in the Examples 2.10 and 2.11.

It turns out that not many more examples can be made, as the following celebrated theorem of Berkovich tells us.

Theorem 2.13 (Berkovich classification theorem) *Every point/seminorm $x \sim |\cdot|_x \in \mathbb{A}^1_{\mathbb{C}_p}$ is of the form $|f|_x = \lim_n |f|_{B_n}$*

Proof. See [1]. □

2.3 Classification of points revisited

Even though we have classified all the points of $\mathbb{A}^1_{\mathbb{C}_p}$ in the Theorem 2.13 we can further refine the classification by saying something more about the nature of the nested sequences B_n that appear in the classification. The right thing to look at is the intersection $\cap_n B_n$.

Let $|\cdot|_x \in \mathbb{A}^1_{\mathbb{C}_p}$ be a point given by the nested sequence $(B_n)_n$ of closed discs in $\mathbb{C}_p$. Then, we have the following possibilities for the intersection $\cap_n B_n$:

(a) $\cap_n B_n = a \in \mathbb{C}_p$. In this case the seminorm is given by $|f|_{(B_n)} = |f(a)|$ and we agree to call $|\cdot|_x$ a *type I* or a *rational* point in $\mathbb{A}^1_k$.

(b) $\cap_n B_n = B(a,r)$ and $r \in p^{\mathbb{Q}} = |\mathbb{C}_p^*|$, where the later is the set $\{|\alpha|, 0 \neq \alpha \in C_p\}$. In this case the seminorm is given by $|f(T)|_x = \sup |f(T)|_{B(a,r)}$ or equivalently, by writing $f(T) = f_0 + f_1(T-a) + \cdots + f_n(T-a)^n$, we have $|f(T)| = \max_i |f_i| r^i$; These are called points of *type II*.

(c) $\cap_n B_n = B(a,r)$ and $r \notin p^{\mathbb{Q}} = |\mathbb{C}_p^*|$. These are called points of *type III*.

(d) $\cap_n B_n = \emptyset$; These are called points of *type IV*.

Remark 2.14 Contrary to the case of $\mathbb{R}$ or $\mathbb{C}$, it can happen that a sequence of nested closed discs in $\mathbb{C}_p$ has an empty intersection. This is due to the fact that $\mathbb{C}_p$ is not *spherically complete* and this is precisely the origin of the points of type IV.

2.4 Hierarchy

Before going to the topological description of $\mathbb{A}^1_k$ we use the fact that the values of polynomials in $\mathbb{C}_p[T]$ at different points in $\mathbb{A}^1_{\mathbb{C}_p}$ can be compared, and this can be used to put a partial ordering on the points in $\mathbb{A}^1_{\mathbb{C}_p}$.

Let us write $\zeta_{a,r}$ for the point corresponding to the disc $B(a,r)$ (here, a and r can also be sequences of numbers corresponding to a type IV point).

Definition 2.15 For two points $\zeta_{a_1,r_1}, \zeta_{a_2,r_3} \in \mathbb{A}^1_{\mathbb{C}_p}$,we write $\zeta_{a_1,r_1} \geq \zeta_{a_2,r_2}$ if for all $f(T) \in \mathbb{C}_p[T]$, $|f(T)|_{\zeta_{a_1,r_1}} \geq |f(T)|_{\zeta_{a_2,r_2}}$.

Exercise 2.16 Prove that $\zeta_{a_1,r_1} \geq \zeta_{a_2,r_2}$ iff $B(a_2,r_2) \subseteq B(a_1,r_1)$. Prove that in this case, $B(a_1,r_1) = B(a_2,r_1)$, that is $\zeta_{a_1,r_1} = \zeta_{a_2,r_1}$.

Exercise 2.17 Prove that the minimal points in $\mathbb{A}^1_{\mathbb{C}_p}$ are type I and type IV points.

One can reformulate the statements of the previous exercises, by saying that we have the "real semilines" inside $\mathbb{A}^1_{\mathbb{C}_p}$ given by parametrization $r \in [0,\infty) \leftrightarrow \zeta_{a,r}$, for an $a \in \mathbb{C}_p$. Furthermore, two such "lines" $\zeta_{a_1,r}$ and $\zeta_{a_2,r}$, r a parameter in $[0,\infty)$ intersect precisely in the point $\zeta_{a_1,|a_1-a_2|}$.

In the Figure 1. one may see the Berkovich affine line $\mathbb{A}^1_{\mathbb{C}_p}$. The "full" dots correspond to the rational points, the "empty" dots correspond to the type IV points (one can see that they are endpoints of an "infinite tree"), the points where the branching occurs are type II and finally the remaining points are type III. In the figure one may also notice the Gauss point ζ_G and the point "∞" which is an added point which by definition is bigger than any other point in $\mathbb{A}^1_{\mathbb{C}_p}$. (It can also be seen as a one-point compactification of $\mathbb{A}^1_{\mathbb{C}_p}$.)

2.5 Topology

The basis of topology on $\mathbb{A}^1_{\mathbb{C}_p}$ is given by the open and closed disc, which we define next.

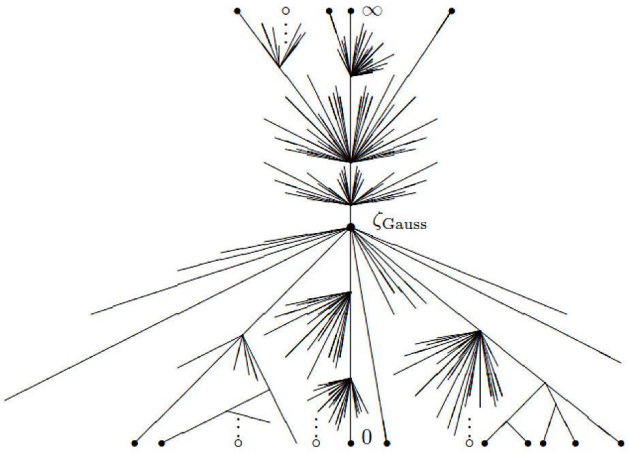


Figure 1. Compactification of the) Berkovich affine line $\mathbb{A}^1_{\mathbb{C}_p}$.

Definition 2.18 A Berkovich closed (resp. open) disc with center in $a \in \mathbb{C}_p$ and of radius $r \geq 0$ is a set

$$\mathbb{B}(a,r) := \{\zeta_{b,s}, \text{ s.t. } b \in B(a,r) \text{ and } s \leq r\}$$

(resp.)

$$\mathbb{B}(a,r^-) := \{\zeta_{b,s}, \text{ s.t. } b \in B(a,r) \text{ and } s < r\}.$$

In the Figures 2 and 3, adopting the representation of $\mathbb{A}^1_{\mathbb{C}_p}$ as in Figure 1, we show the Berkovich closed and open unit discs centered at $0 \in \mathbb{C}_p$.

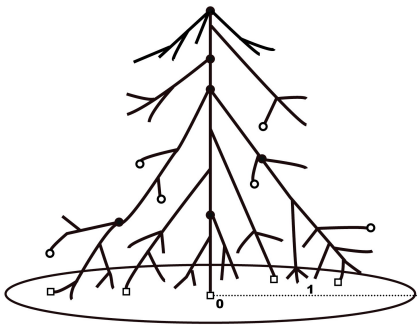


Figure 2. Closed Berkovich disc $\mathbb{B}(0,1)$.

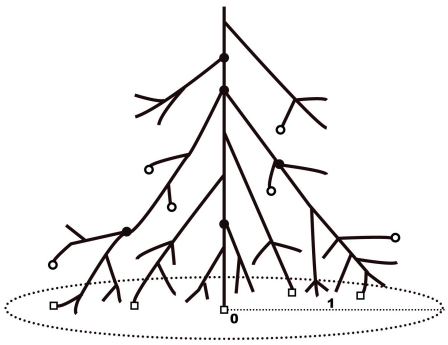


Figure 3. Open Berkovich disc $\mathbb{B}(0,1^-)$; The dashed points correspond to the branches that are taken out from the closed unit disc.

Notice that the difference between the closed and open Berkovich unit discs centered at

zero is that we removed from the closed disc all the branches which are attached to the Gauss point, except for branch which contains 0.

We may also present an open annulus which is a set difference between a closed Berkovich disc and an open Berkovich subdisc, as in the Figure 4.

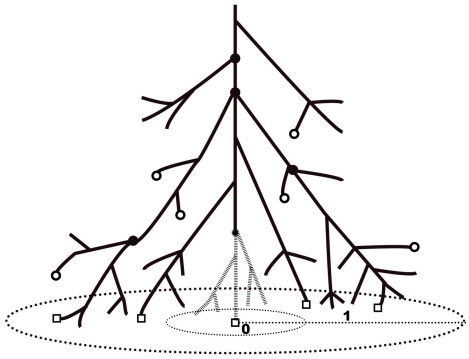


Figure 4. Open annulus: The dashed inner closed Berkovich disc represents the closed disc "taken out" from the bigger open disc.

3 Relation to curves in characteristic $p > 0$

3.1 Semistable curves

In this section we shortly present without going into details and definitions how one can "visualize" general smooth, connected, compact Berkovich curves using their semistable models. We begin with an exercise.

Exercise 3.1 Prove that the the set of points $\mathbb{C}_p^\circ := \{\alpha \in \mathbb{C}_p, |\alpha|_p\}$ is a ring in $\mathbb{C}_p$ having only one maximal ideal given by the set $\mathbb{C}_p^{\circ\circ} := \{\alpha \in \mathbb{C}_p, |\alpha|_p < 1\}$. We call the field $\mathbb{C}_p^\circ / C_p^{\circ\circ}$ the *residue field* of $\mathbb{C}_p$.

Remark 3.2 It can be shown that the residue field of $\mathbb{C}_p$ is isomorphic to the algebraic closure $\overline{\mathbb{F}}_p$ of the finite field $\mathbb{F}_p$.

Definition 3.3 We say that an $\overline{\mathbb{F}}_p$-curve C is semistable, if all the singularities on C are regular.

Theorem 3.4 (Bosch-Lütkebohmert-Berkovich) *Given a semistable curve C, we can assign to it a compact, connected, smooth Berkovich curve (over $\mathbb{C}_p$) C_B together with an*

(anticontinuous) map $C_B \to C$ and a set S consisting of finitely many type II points in C_B such that

(a) *The map induces a 1-1 correspondence between generic points and the set S in C_B.*

(b) *The preimage of a smooth point in C is an open disc attached to the corresponding point in S.*

(c) *The preimage of a singular point is an open annulus connecting the corresponding two points in S.*

Moreover, every compact, connected, smooth Berkovich curve over $\mathbb{C}_p$ comes in this way.

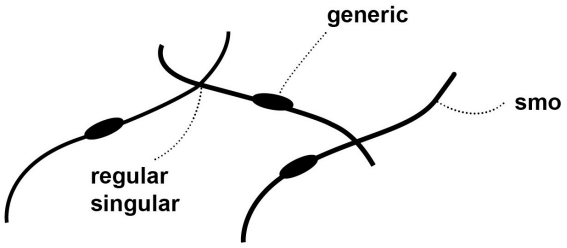


Figure 5. An example of a semistable $\overline{\mathbb{F}}_p$-curve.

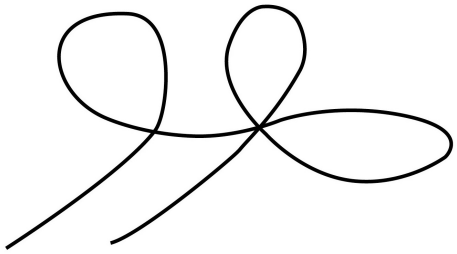


Figure 6. An example of a $\overline{\mathbb{F}}_p$-curve which is not semistable.

3.2 Example: Tate elliptic curve

In Figures 7 and 8 are given a semistable curve over $\overline{\mathbb{F}}_p$ and the corresponding Berkovich curve, which is called *Tate elliptic curve*.

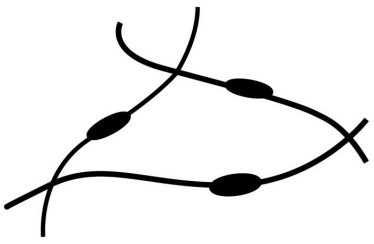


Figure 7. A semistable $\overline{\mathbb{F}}_p$-curve with emphasized generic points of the irreducible components.

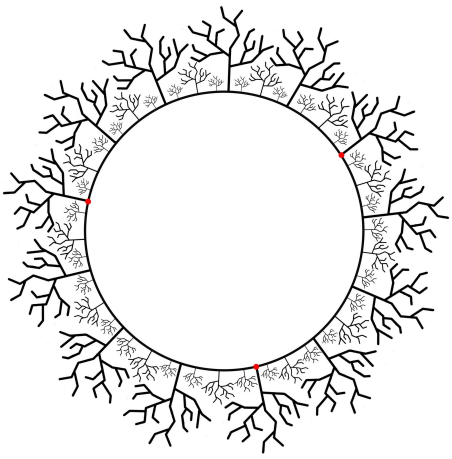


Figure 8. The corresponding Berkovich curve. The red points correspond to the set S in Theorem 3.4.

Perhaps, the reader will appreciate an artistic interpretation of the Tate elliptic curve presented in the next figure.

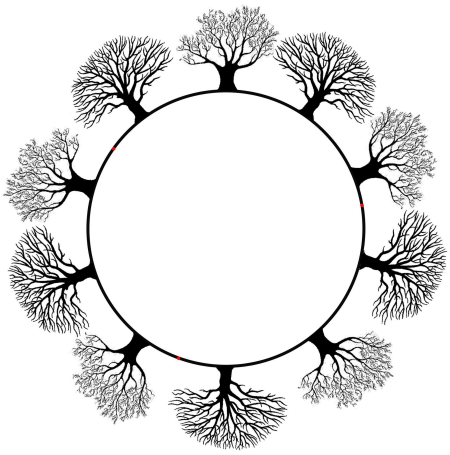


Figure 9. Tate elliptic curve: artistic version.

References

[1] Vladimir Berkovich, "Spectral theory and analytic geometry over non-archimedean fields". Mathematical surveys and monographs, vol. 33, American Mathematical Society, Providence RI, 1990.

On differential operators and multipliers in weighted Sobolev spaces

Aigul Myrzagaliyeva (*)

Abstract. In this talk, we introduce differential operators and pointwise multipliers in weighted Sobolev spaces. We give the statement and motivation of the problem. We obtain boundedness conditions and norm estimates for the differential operator L from the weighted Sobolev space $W^n_{p,v}$ to the weighted Lebesgue space $L_{q,\omega}$ and for the multiplication operator T from $W^n_{p,v}$ to $W^l_{q,\omega_0,\omega_1}$, where $W^l_{q,\omega_0,\omega_1}$ is a weighted Sobolev space on $\mathbb{R}$. Moreover, we also present some open problems.

1 Introduction

We consider the differential operator of the form

$$Ly = \sum_{k=0}^{l} \rho_k(x) y^{(k)} \quad (x \geq 0), \tag{1}$$

where $\rho_k(\cdot) \in L^{loc}(I),\ I = (0,\infty),\ l \geq 1$ is an integer. Here $L^{loc}(I)$ is the class of all locally summable functions in I. In the sequel, we assume that L is defined on a subspace $D(L)$ of $W^n_{p,v}$. We denote by $W^n_{p,v}(I)$ ($n \geq 1 -$ integer, $1 < p < \infty$) the weighted Sobolev space of all functions y, having absolutely continuous derivatives up to order $n-1$ in I and finite weighted norm in the following form

$$\left\| y; W^n_{p,v}(I) \right\| = \left(\int_I \left(\left| y^{(n)}(x) \right|^p + |y(x)|^p \, v(x) \right) dx \right)^{\frac{1}{p}},$$

$W^n_p = W^n_{p,v},\ v \equiv 1$. Here $v(\cdot)$ be a weight in I, i.e. it is an almost everywhere positive locally integrable function. We denote by $L_{q,\omega}(I)$ the weighted Lebesgue space of all measurable functions in I with the norm

$$\|f\|_{q,\omega} = \|f;\, L_{q,\omega}(I)\| = \left(\int_I |f(x)|^q \, \omega(x)\, dx \right)^{\frac{1}{q}} < \infty \quad (1 \leq q < \infty),$$

(*)The L.N. Gumilyov Eurasian National University, Astana, Kazakhstan, Munaitpasov St., 5, 010008, Astana, Kazakhstan; E-mail: mir_aigul@mail.ru . Seminar held on April 15th, 2015.

$L_q(I) = L_{q,\omega}(I),\ \omega \equiv 1$.

The purpose of this paper is to obtain conditions for the boundedness of the differential operator L as an operator acting from $W^n_{p,v}(I)$ to the space $L_{q,\omega}(I)$ and to apply these results to the study of multipliers in Sobolev spaces.

We define as length function in I any positive and right-continuous function $h(\cdot)$ ($h(\cdot)$ is a l.f.). We denote by $\Delta(x)$ the segment [x,x+h(x)] for the l.f. $h(\cdot)$.

This paper is organized as follows: In Sec. 2, we introduce Sobolev spaces used as function spaces and we give some examples to understand why such spaces are important. In Sec. 3, you find basic definitions and notation. In Sec. 4, we present some new results formulated as theorems.

2 Sobolev spaces

In mathematics, a Sobolev space is a vector space of functions equipped with a norm that is a combination of L_p-norms of the function itself as well as its derivatives up to a given order. The derivatives are understood in a suitable weak sense to make the space complete, thus a Banach space. Intuitively, the Sobolev space is a space of functions with sufficiently many derivatives for some application domain, such as partial differential equations. Their importance comes from the fact that solutions of partial differential equations are naturally found in Sobolev spaces, rather than in spaces of continuous functions. In the twentieth century, however, it was observed that the space C^1 (or C^2, etc.) was not exactly the right space to study solutions of differential equations. Sobolev spaces are the modern replacement for these spaces in which to look for solutions of partial differential equations.

Let us consider the simplest example the Dirichlet problem for the Laplace equation in a bounded domain $\Omega \subset \mathbb{R}^n$:

$$\begin{cases} \Delta u = 0, & x \in \Omega, \\ u(x) = \varphi(x), & x \in \partial\Omega, \end{cases}$$

where $\varphi(x)$ is a given function on the boundary $\partial\Omega$. It is known that the Laplace equation is the Euler equation for the functional

$$l(u) = \int\limits_{\Omega} \sum_{j=1}^{n} \left| \frac{\partial u}{\partial x_j} \right|^2 dx.$$

We can consider this problem as a variational problem: to find the minimum of $l(u)$ on the set of functions satisfying condition $u|_{\partial\Omega} = \varphi$. It is much easier to minimize this functional not in $C^1(\bar{\Omega})$, but in a larger class. Namely, in the Sobolev class $W_2^1(\Omega)$. $W_2^1(\Omega)$ consists of all functions $u \in L_2(\Omega)$, having the *weak derivatives* $\partial u \in L_2(\Omega),\ j = 1, \ldots, n$. If the boundary $\partial\Omega$ is smooth, then the trace of $u(x)$ on $\partial\Omega$ is well defined and relation $u|_{\partial\Omega} = \varphi$ makes sense. If we consider $l(u)$ on $W_2^1(\Omega)$, it is easy to prove the existence and uniqueness of solution of our variational problem. The function $u \in W_2^1(\Omega)$, that gives minimum to $l(u)$ under the condition $u|_{\partial\Omega} = \varphi$, is called the *weak solution* of the Dirichlet problem for the Laplacian.

Let us consider the following exterior Oseen problem to clarify the importance of the weighted Sobolev space:

$$\begin{cases} -\Delta u + k\frac{\partial u}{\partial x_1} + \nabla\pi = 0 \quad \text{in } \Omega^-, \\ \text{div} u = 0 \quad \text{in } \Omega^-, \\ u = u_* \quad \text{on } \partial\Omega^-, \end{cases}$$

where $\Omega^- \equiv \mathbb{R}^3 \backslash \bar{\Omega},\ \Omega = B(0,1)$. The data are the boundary value u^* and a real $k > 0$. The problem consists of looking for the velocity field u of the fluid and the pressure function π.

We introduce the weight function $\rho(x) = 1 + |x|$. Let us assume that $(\tilde{u}, \pi)$ is a solution of the problem and $(\tilde{u}, \pi)$ equal to the fundamental solution of the Oseen system, which is of class C^∞ in Ω^-. Then the boundary data u^* is just the restriction to $\partial\Omega^- = \partial\Omega$. As we can easily show, the component u of the fundamental solution behaves like $\frac{1}{\rho}$ at infinity and accordingly is not in $W_2^1(\Omega^-)$. However, it is in the weighted Sobolev space $W^1_{2,\rho}(\Omega^-)$.

3 Preliminaries and notation

Definition 1 (See [1]) A weighted function v in I is called admissible with respect to the length function $h(\cdot)$, if there exist $0 < \delta < 1$, $0 < \tau \leq 1$, such that the following inequality is true

$$h(x)^{n-\frac{1}{p}} \inf_{\{e\}} \left(\int_{\Delta(x)\backslash e} v(t)\, dt \right)^{\frac{1}{p}} \geq \tau \tag{2}$$

for all $x \in I$.

In (2) the infimum is taken over all measurable subset e of $\Delta(x)$ with Lebesgue measure $|e| \leq \delta\, |\Delta(x)|$. We denote by $\Pi_{n,p}(\delta, \tau)$ the set of admissible weights v with respect to the l.f. $h(\cdot)$.

We give some examples.

Example 1 Since

$$h(x)^{n-\frac{1}{p}} \inf_{\{e\}} \left(\int_{\Delta(x)\backslash e} v(t)\, dt \right)^{\frac{1}{p}} = \inf_{\{e\}} \left(\int_{[x,x+1]\backslash e} dt \right)^{\frac{1}{p}} \geq (1-\delta)^{\frac{1}{p}} = \tau,$$

the function $v = 1$ is admissible with respect to the l.f. $h(\cdot) = 1$.

Definition 2 We say that a function $\omega(\cdot)$ satisfies the slow variation condition with respect to the l.f. $h(\cdot)$, if there exist constants $0 < b_1 < 1 < b_2$ such that

$$b_1\omega(x) \leq \omega(t) \leq b_2\omega(x) \quad \text{for all} \quad t \in \Delta(x). \tag{3}$$

Example 2 Let v satisfy the slow variation condition (3) with respect to the l.f. $h(x) = v(x)^{-\frac{1}{np}}$. Then v is admissible with respect to the l.f. $h(x) = v(x)^{-\frac{1}{np}}$ when $\tau^p = b_1(1-\delta)$. The proof is trivial.

Every power function $v(x) = (1+|x|)^{\mu}$, $0 < \mu < +\infty$ satisfies the slow variation condition with respect to the l.f. $h(x) = (1+x)^{-\frac{\mu}{np}}$. Indeed,

$$\left(\frac{1+t}{1+x}\right)^{\mu} \le \left(\frac{1+x+h(x)}{1+x}\right)^{\mu} = \left(1+\left(\frac{1}{1+x}\right)^{1+\frac{\mu}{np}}\right)^{\mu} \le 2^{\mu} = b_2,$$
$$\left(\frac{1+t}{1+x}\right)^{\mu} = \left(\frac{1+x+t-x}{1+x}\right)^{\mu} = \left(1+\frac{t-x}{1+x}\right)^{\mu} \ge 1 > 2^{-\mu} = b_1$$

for all $t \in \Delta(x)$.

Definition 3 We say that a weight v satisfies the condition $A_{(\delta,\beta)}$ $(0 < \delta, \beta < 1)$ with respect to the length function $h(\cdot)$ in I if for any interval $\Delta = [a,b] \subset \Delta(x) = [x, x+h(x)]$ $(x \ge 0)$ and any measurable subset e of Δ with the Lebesgue measure $|e| \le \delta|\Delta|$ the following inequality holds

$$\int_e v(t)\,dt \le \beta \int_{\Delta} v(t)\,dt.$$

We denote by $A_{(\delta,\beta)}$ the set of all weights v which satisfy the condition $A_{(\delta,\beta)}$. For example, if $b_2 b_1^{-1}\delta < 1$ in (3), then $v \in A_{(\delta,\beta)}$ with $\beta = b_2 b_1^{-1}\delta$.

Let v^* be a Otelbayev function. Namely

$$v^*(x) = \sup_{h>0}\left\{h\colon h^{pn-1}\int_x^{x+h} v(t)\,dt \le 1\right\}$$

(see [2]). We first show that $0 < v^*(x) < \infty$ for all $x \ge 0$. To do this, we note that

$$M(x,h;v) \stackrel{\text{def}}{=} h^{pn-1}\int_x^{x+h} v(t)\,dt \xrightarrow[h\to 0+]{} 0$$

and that $M(x,h;v) \to \infty$ if $h \to \infty$. Hence, there exist $\delta_x > 0$ and $T_x > 0$, such that

$$M(x,h;v) \le 1, \quad \text{if} \quad 0 < h \le \delta_x,$$
$$M(x,h;v) > 1, \quad \text{if} \quad h \ge T_x.$$

Therefore, we obtain

$$(0,\delta_x) \subset H_{x,v} = \{h > 0\colon M(x,h;v) \le 1\} \subset (0,\, T_x),$$
$$\delta_x \le \sup H_{x,v} = v^*(x) \le T_x.$$

The function v^* is continuous in I. By using absolute continuity property of the integral we can imply that

$$v^*(x)^{pn-1} \int_x^{x+v^*(x)} v(t)\,dt = 1.$$

Example 3 Any weight $v \in A_{(\delta,\beta)}$ with respect to the l.f. $h(x) = v^*(x)$ in I is admissible with respect to the l.f. $h(x) = v^*(x)$. Thus, for all $e \subset \Delta^*(x) = [x, x + v^*(x)]$ with the measure $|e| \le \delta|\Delta^*(x)|$, we have

$$v^*(x)^{pn-1} \int_{\Delta^*(x)\setminus e} v(t)\,dt = v^*(x)^{pn-1} \left(\int_{\Delta^*(x)} v(t)\,dt - \int_e v(t)\,dt \right) \ge$$
$$\ge (1-\beta) v^*(x)^{pn-1} \int_{\Delta^*(x)} v(t)\,dt = 1 - \beta = \tau.$$

Let $C^n[a,b]$ $(-\infty < a < b < \infty)$ be the space of all functions y, having continuous derivative up to order n in $[a,b]$.

Proposition 1 (see [3]) *Let $1 \le p < \infty$. Then there exist constants $A(n,j)$, $A(n,j,p)$, $j = 0,\ 1,\ 2,\ \ldots,\ n-1$ such that*
(4)

$$\max_{a\le x\le b} \left| y^{(j)}(x) \right| \le A(n,j)(b-a)^{n-j-1/p} \left\{ \int_a^b \left| y^{(n)}(t) \right|^p \, dt + (b-a)^{-np} \int_a^b |y(t)|^p \, dt \right\}^{1/p},$$

$$(5) \qquad \left\| y^{(j)}; L_p[a,b] \right\| \le A(n,j,p)(b-a)^{n-j} \left\{ \left\| y^{(n)}; L_p[a,b] \right\| + (b-a)^{-n} \left\| y; L_p[a,b] \right\| \right\}$$

for all functions $y(\cdot) \in C^n[a,b]$. Here we understand that the symbols $A(n,j)$, $A(n,j,p)$ denote the best choice of the constants in (4), (5).

Lemma 1 *Let v belong to $\Pi_{n,p}(\delta,\tau)$ with respect to the l.f. $h(\cdot)$. Then there exists a constant $C^* = C^*(\delta,\tau) > 1$ independent of $v(\cdot)$ such that the following estimate is true*

$$(6) \qquad h(x)^{-np} \int_x^{x+h(x)} |y|^p \, dt \le C^* \int_x^{x+h(x)} \left(\left| y^{(n)} \right|^p + |y|^p v(t) \right) dt \quad (x \ge 0)$$

for all $y \in C^n(\Delta)$, where $\Delta = [x, x+h(x)]$.

4 Main results

4.1 One dimensional differential operators in weighted Sobolev spaces $W^n_{p,v}$

Theorem 1 *Let $n > 1$ be an integer. Let $1 < p \le q < \infty$. Let v belong to $\Pi_{n,p}(\delta, \tau)$ with respect to the l.f. $h(\cdot)$. Let*

$$R_k = \sup_{x \ge 0} h(x)^{n-k-\frac{1}{p}} \left\{ \int_x^{x+h(x)} |\rho_k(t)|^q \, d\omega(t) \right\}^{\frac{1}{q}} < \infty$$

for $k = 0, 1, \dots, l$ $(d\omega(t) = \omega(t)dt)$. Then the operator L in (1) is bounded from $W^n_{p,v}(I)$ to $L_{q,\omega}(I)$. Here the norm has the following form

$$\left\| L;\, W^n_{p,v}(I) \to L_{q,\omega}(I) \right\| \le C \sum_{k=0}^{l} R_k, \tag{7}$$

where $C = (1 + C^*)^{\frac{1}{p}} \sum_{k=0}^{l} A(n,k)$.

Let us assume that the operator L in (1) is bounded as an operator from $W^n_{p,v}$ to $L_{q,\omega}$, i.e. $D(L) \subset W^n_{p,v}$ and there exists a constant $K > 0$ such that

$$\left(\int_I |Ly|^q \, d\omega(t) \right)^{\frac{1}{q}} \le K \| y;\, W^n_{p,v} \| \quad (y \in D(L)). \tag{8}$$

Theorem 2 *Let $1 < p, q < \infty$. Let the operator L in (1) be bounded from $W^n_{p,v}$ to $L_{q,\omega}$. Then*

$$\widetilde{R}_k = \sup_{x \ge 0} v^*(x)^{n-k-\frac{1}{p}} \left\{ \int_{x+\frac{v^*(x)}{4}}^{x+\frac{3v^*(x)}{4}} |\rho_k(t)|^q \, d\omega(t) \right\}^{\frac{1}{q}} \le \widetilde{C}_k \left\| L;\, W^n_{p,v} \to L_{q,\omega} \right\|, \tag{9}$$

where the constant $\widetilde{C}_k > 0$ does not depend on ρ_k $(0 \le k \le l)$, ω, v.

We set

$$\overline{R^*} = \sum_{k=0}^{l} \sup_{x \ge 0} v^*(x)^{n-k-\frac{1}{p}} \left\{ \int_x^{x+v^*(x)} |\rho_k(t)|^q \, d\omega \right\}^{\frac{1}{q}},$$

$$\underline{R^*} = \sum_{k=0}^{l} \sup_{x \ge 1} v^*(x)^{n-k-\frac{1}{p}} \left\{ \int_{x+\frac{v^*(x)}{4}}^{x+\frac{3v^*(x)}{4}} |\rho_k(t)|^q \, d\omega \right\}^{\frac{1}{q}}.$$

Theorem 3 *Let v belong to $A_{(\delta,\beta)}$. Let $\overline{R^*} < \infty$. Then the operator L in (1) is bounded from $W^n_{p,v}$ to $L_{q,\omega}$. Moreover,*

$$C_0 \underline{R^*} \le \left\| L;\, W^n_{p,v} \to L_{q,\omega} \right\| \le C_1 \overline{R^*}.$$

The constants $0 < C_i < 1$ depend only on n, p, l, δ, β.

The statements of Theorem 3 are direct consequences of Theorem 1 and Theorem 2. The following result has been obtained in the monograph [4].

Corollary 1 *Let $1 < p \le q < \infty$. The following statements are true:*

a) *Let the operator L in (1) be bounded from W^n_p to $L_{q,\omega}$. Then*

$$\underline{A} = \sum_{k=0}^{l} \sup_{x \ge 1} \int_x^{x+1} |\rho_k|^q \, d\omega < \infty.$$

b) *Let*

$$\overline{A} = \sum_{k=0}^{l} \sup_{x \ge 0} \int_x^{x+1} |\rho_k|^q \, d\omega < \infty.$$

Then the operator L in (1) is bounded from W^n_p to $L_{q,\omega}$. Moreover,

$$C_0 \left(\underline{A}\right)^{\frac{1}{q}} \le \left\| L;\, W^n_p \to L_{q,\omega} \right\| \le C_1 \left(\overline{A}\right)^{\frac{1}{q}},$$

where the constants $0 < C_0 < C_1$ and C_i $(i = 0, 1)$ depend only on n, l, p and q.

The statement *b*) follows from Theorem 1, because $v(x) = 1 \in \Pi_{n,p}\left(2^{-1}, 2^{-\frac{1}{p}}\right) \cap A_{\left(\frac{1}{2},\frac{1}{2}\right)}$ with respect to the l.f. $h(x) = 1$. The statement *a*) follows from Theorem 2, because $v^*(x) = 1$ and

$$\int_x^{x+1} |\rho_k|^q \, d\omega = \int_x^{x+\frac{1}{2}} |\rho_k|^q \, d\omega + \int_{x+\frac{1}{2}}^{x+1} |\rho_k|^q \, d\omega = \int_{\widetilde{\Delta}\left(x-\frac{1}{4}\right)} |\rho_k|^q \, d\omega + \int_{\widetilde{\Delta}\left(x+\frac{1}{4}\right)} |\rho_k|^q \, d\omega.$$

Corollary 2 *Let $v(x) = \left(1 + x^2\right)^s$, $\omega(x) = \left(1 + x^2\right)^\beta$, $s, \beta > 0$. Then the operator*

$$Ly = \sum_{k=0}^{l} \left(1 + x^2\right)^{\alpha_k} y^{(k)}, \; -\infty < \alpha_k < \infty,$$

is bounded from $W^n_{p,v}$ to $L_{q,\omega}$ if and only if

$$\alpha_k + \frac{\beta}{q} \le \frac{s}{np}\left(n - k - \frac{1}{p} + \frac{1}{q}\right) \quad (k = 0, 1, \ldots, l). \tag{10}$$

Here if $\rho_k(x) = \left(1+x^2\right)^{\alpha_k}$ the statements of Theorem 3 with respect to the l.f.

$$h(x) \;=\; v^*(x) \;\sim\; \left(1+x^2\right)^{-\frac{s}{np}}$$

are equivalent to the following conditions:

$$\sup_{x>0}\left(1+x^2\right)^{-\frac{s}{np}\left(n-k-\frac{1}{p}+\frac{1}{q}\right)+\alpha_k+\frac{\beta}{q}} < \infty \quad (k=0,1,\ldots,l).$$

Such conditions are verified if and only if relation (10) holds.

4.2 Multipliers in weighted Sobolev spaces $W^n_{p,v}$

The results obtained in this paper can be regarded as a natural extension of certain results (in dimension one) of the monograph "Theory of multipliers in spaces of differentiable functions" by the authors V.G. Maz'ya and T.O. Shaposhnikova. Such a book is currently the only work in which the theory of pointwise multipliers in unweighted spaces of differentiable functions is treated systematically. A part of the chapters of this work are devoted to multipliers in classical Sobolev spaces $W^n_p,\ n \geq 0 -$ integer, $1 \leq p < \infty$.

Let X, Y be Banach spaces whose elements are functions $y\colon \Omega \to \mathbb{R}\ (\mathbb{C})$. We say that a function $z\colon \Omega \to \mathbb{R}\ (\mathbb{C})$ such that the multiplication operator

$$Ty = zy,\ y \in X,$$

is bounded from X to Y, is a multiplier for a pair (X,Y). We denote by $M(X \to Y)$ the space of all multipliers for the pair (X,Y). We introduce the norm

$$\|z;\, M(X \to Y)\| = \|T;\, X \to Y\|\,,$$

in $M(X \to Y)$.

Below as an application, we obtain the description of the space $M(W_1 \to W_2)$ for a pair of weighted Sobolev spaces (W_1, W_2) with weights of general type.

It should be noted that there is enough active study of multipliers in function spaces at present. Let us point out some specific directions through the works [5]-[13].

We denote by $A^l(I)$ the class of all functions y in I, having absolutely continuous derivatives up to order $l-1$ in I.

Let $\omega_0,\ \omega_1$ be weighted functions in I. Let $l \geq 1$ be an integer. We denote by $W^l_{q,\omega_0,\omega_1}(I)$ the weighted Sobolev space of all functions $y \in A^l(I)$ equipped with the following weighted norm

$$\left\|y;\, W^l_{q,\omega_0,\omega_1}(I)\right\| = \left\|y^{(l)};\, L_{q,\omega_1}(I)\right\| + \|y;\, L_{q,\omega_0}(I)\|\,.$$

Theorem 4 *Let $n > l \geq 1$ be integers, $1 \leq p \leq q < \infty$. Let $v \in \Pi_{n,p}(\delta,\tau)$ with respect to the l.f. $h(\cdot)$ in I. Let $\mu \in A^l(I)$. If*

$$M_{k,\mu,\omega_1} = \sup_{x \geq 0} h(x)^{n-k-\frac{1}{p}} \left\{ \int\limits_x^{x+h(x)} \left|\mu^{(l-k)}(t)\right|^q \, d\omega_1(t) \right\}^{\frac{1}{q}} < \infty \quad (k=0,1,\ldots,l),$$

$$M_{0,\mu,\omega_0} = \sup_{x\geq 0} h(x)^{n-\frac{1}{p}} \left\{ \int\limits_{x}^{x+h(x)} |\mu(t)|^q \, d\omega_0(t) \right\}^{\frac{1}{q}} < \infty,$$

then $\mu \in M(W^n_{p,v} \to W^l_{q,\omega_0,\omega_1})$. *Moreover,*

$$\left\| \mu;\, M(W^n_{p,v} \to W^l_{q,\omega_0,\omega_1}) \right\| \leq C \left[\sum_{k=0}^{l} M_{k,\mu,\omega_1} + M_{0,\mu,\omega_0} \right],$$

where $C = C(n,l,p,q) > 0$.

Corollary 3 *Let* $n > l \geq 1,\ 1 < p \leq q < \infty$. *Let* $\mu \in A^l$. *Then* $\mu \in M(W^n_p \to W^l_{q,\omega_0,\omega_1})$ *if and only if*

$$U_k = \sup_{x\geq 1} \int\limits_{x}^{x+1} |\mu^{(k)}|^q \, d\omega_1 < \infty \quad (k = 0, 1, \ldots, l),$$

$$V = \sup_{x\geq 1} \int\limits_{x}^{x+1} |\mu|^q \, d\omega_0 < \infty.$$

Previously, the description of the space $M(W^n_{p,v} \to W^l_{q,\omega_0,\omega_1})$ was obtained in [14]. There were used more complex terms, which included special maximal operators $M^*\omega_i,\ i = 1, 2$.

Acknowledgement. The author thanks Università Degli Studi di Padova for the opportunity to have an internship, as well as expresses her gratitude to Professor Leili Kussainova and Professor Massimo Lanza de Cristoforis for providing cooperative support and valuable assistance. The paper was done under partial financial support by the grant of RK MES, the grant No.2989/3.

References

[1] L. Kussainova, *On embedding the weighted space* $W^l_p(\Omega; v)$ *to the space* $L_p(\Omega; \omega)$. Matem. Sb. 2 (2000), 132–148 (Russian).

[2] M. Otelbaev,, "Spectrum estimates of the Sturm-Liouville operator". Gyilyim, Alma-ata, 1990 (Russian).

[3] S. M. Nikolskiy, "Approximation of functions of several variables and embedding theorems". Nauka, Moscow, 1977 (Russian).

[4] V. G. Maz'ya, T. O. Shaposhnikova, "Theory of multipliers in spaces of differentiable functions". Editore, anno.

[5] G. Nariman, *Smooth pointwise multipliers of modulation spaces.* An. Stiint. Univ. "Ovidius" Constanta Ser. Mat. 20/1 (2012), 317–327.

[6] E. Nakai, *Pointwise multipliers on weighted BMO spaces.* Studia Math. 125/1 (1997), 35–56.

[7] K. Yabuta, *Pointwise multipliers of weighted BMO spaces.* Proc. Amer. Math. Soc. 117/3 (1993), 737–744.

[8] S. Bloom, *Pointwise multipliers of weighted BMO spaces.* Proc. Amer. Math. Soc. 105/4 (1989), 950–960.

[9] L. Maligranda, E. Nakai, *Pointwise multipliers of Orlicz spaces.* Arch. Math. 95/3 (2010), 251–256.

[10] E. Nakai, *Pointwise multipliers for functions of weighted bounded mean oscillation.* Studia Math. 105/2 (1993), 105–119.

[11] F. Beatrous, J. Burbea, *On multipliers for Hardy-Sobolev spaces.* Proc. Amer. Math. Soc. 136/6 (2008), 2125–2133.

[12] E. Nakai, *Pointwise multipliers on the Lorentz spaces.* Mem. Osaka Kyoiku Univ. III Natur. Sci. Appl. Sci. 45/1 (1996), 17.

[13] M.-R. Pierre Gilles, *Multipliers and Morrey spaces.* Potential Anal. 38/3 (2013), 741–752.

[14] L. Kussainova, *On multipliers in weighted Sobolev spaces.* Matem. Sb. 196/8 (2000), 21–48 (Russian).

Preferences in AI

Cristina Cornelio (*)

Abstract. Artificial Intelligence (AI) is a field that has a long history but still constantly and actively growing and changing. The applications of AI are several, for example web search, speech recognition, face recognition, machine translation, autonomous driving, automatic scheduling etc. These are all complex real-world problems, and the goal of artificial intelligence is to tackle them with rigorous mathematical tools: machine learning, search, game playing, Markov decision processes, constraint satisfaction, graphical models, and logic. Recently, a new concept became very important in AI: the use of preferences. Let's think about social networks, online shops, systems that suggest music or films. In this document it is presented an overview on the main formalisms of qualitative preferences in AI, a particular subset of models to represent preferences, very useful in real word applications.

1 Artificial Intelligence: an overview

Artificial Intelligence (AI) is a big field and there are many important applications of AI technology in the real world. Most important are:

- **Robotics.** Robotics is widely used in industrial automation, but they are also used as medical surgery support to help doctors in complicated surgical procedures. Artificial body components (arms, legs) and also artificial organs (heart) was developed to interact directly with the human brain. Robotics is widely used in space exploration or to develop exploration robots for difficult conditions, for example with too hight/low temperature or pressure (e.g. deep sea).
- **Natural language processing.** The goal of this subfield is the direct interaction human-machines. A machine has to understand a text and to make a summary, to translate a text in a proper way, to recognize if two texts are written by the same person, or to search text following the meaning and not following only the words in the name.
- **Autonomous control systems and planning.** Autonomous schedule and planning is mainly useful for robotic application (e.g. industrial robot and management of the production line), but also in control system. Let's think about the electricity

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: corneliocristina@gmail.com. Seminar held on April 29th, 2015.

industries: Enel uses neural networks to manage the electricity switches between the centrals, the amount of produced energy and the temperature of the areas of production. This subfield is also used to solve conflicts with the instructions given by different machines/sensors.

- **Games.** IBM's Deep Blue became the first computer program that wins again the world chess champion Garry Kasparov. Let's think about the video-games market: video-games are becoming increasingly realistic using AI components.
- **Machine Learning.** Machine Learning is a system that improves its performance on a particular task using past experience. Examples of machine learning application are auto-parking cars , driver-less cars (e.g. Google car), handwriting recognition and speech recognition.
- **Preferences.** Preferences play an increasing role in the web application. Recommender systems are hidden behind many web sites (e.g. Amazon or Netflix) to suggest music, films, or products.

2 Qualitative Preferences on combinatorial domains

The ability to express preferences in a faithful way, which can be handled efficiently, is essential in many reasoning tasks. In settings such as e-commerce, on demand video, and other settings where supply outstrips an individuals ability to view all the available choices, we require an efficient formalism to model and reason with complex, interdependent preferences. We may also use these preferences to make decisions about joint plans, actions, or items in multi-agent environments. Agents express their preferences over a set of candidate decisions, these preferences are aggregated into one decision which satisfies as many agents as possible. Often multi-attribute preference modeling and reasoning causes a combinatorial explosion. The set of candidates is often described as a product of multiple features, for example, a users preferences over a set of cars, which can be described by their colors, technical specifications, cost, reliability, etc. A number of compact representation languages have been developed in the literature to tackle the computational challenges arising from these problems. The main tool for representing and reasoning with conditional preference statements of a single agent are the CP-nets. A more general and complete version of CP-nets are the GCP-nets, introduced in the following section (Section 3). Then we will introduce the other main formulations of qualitative preferences on combinatorial domains.

3 GCP-nets

GCP-nets ([7] and [6]) allow a general form of conditional and qualitative preferences to be modeled compactly.

Definition 3.1 A **Generalized CP-net** (**GCP-net**) C over Var is a set of conditional preference rules. A **conditional preference rule** is an expression $p : l > \bar{l}$, where l is

a literal of some atom $X \in Var$ and p is a propositional formula over Var that does not involve variable X. A GCP-net corresponds to a directed graph (dependency graph) where each node is associated with a feature and the edges are pairs (Y, X) where Y appears in p in some rule $p : x > \overline{x}$ or $p : \overline{x} > x$. Each node X is associated with a *CP-table* which expresses the user preference over the values of X. Each row of the CP-table corresponds to a conditional preference rule.

The CP-tables of a GCP-net can be *incomplete* (i.e. for some values of some variables' parents, the preferred value of X may not be specified) and/or *locally inconsistent* (i.e. for some values of some variables' parents, the table may both contain the information $x > \overline{x}$ and $\overline{x} > x$). CP-nets [2] are a special case of GCP-net in which the preferences are locally consistent and locally complete.

An *outcome* in a GCP-net is a complete assignment to all features. For example, given $Var = \{X_1, X_2\}$ and binary domains $D_1 = D_2 = \{T, F\}$, all the possible outcomes are TT, TF, FT and FF.

A *worsening flip* is a change in the value of a feature to a value which is less preferred according to the cp-statement for that feature. This concept defines an order over the set of outcomes such that one outcome o is *preferred* to another outcome o' ($o \succ o'$) if and only if there is a chain of worsening flips from o to o'. The notion of *worsening flip* induces a preorder over the set of outcomes. This preorder allows maximal elements that correspond to the so-called *optimal outcomes*, which are outcomes that have no other outcome better than them.

Given any GCP-net and CP-net the problems of consistency checking and finding optimal outcomes are PSPACE-complete [7]. Moreover, there could be several different maximal elements. When the dependency graph has no cycle the CP-net is called *acyclic*. The optimal outcomes for such nets are unique and can be found in polynomial time in N. The procedure used to this purpose is usually called a *sweep forward* and takes N steps [2].

The problem of dominance testing (i.e. determining if one outcome is preferred to another) is PSPACE-complete for both GCP-nets and CP-nets. It is polynomial if the CP-nets are tree structured or poly-tree structured [5, 7].

4 CP-nets

CP-nets are a graphical model for compactly representing conditional and qualitative preference relations [2]. They exploit conditional preferential independence by decomposing an agent's preferences via the *ceteris paribus* (cp) assumption (all other things being equal).

CP-nets correspond to complete and locally consistent GCP-nets.

CP-nets bear some similarity to Bayesian networks [4]: both use directed graphs where each node stands for a domain variable, and assume a set of features $F = \{X_1, \ldots, X_n\}$ with finite domains $D_1, \ldots, D_n$.

Definition 4.1 A CP-net (conditional preference network) over features $F = \{X_1, \ldots, X_n\}$ with finite domains $D_1, \ldots, D_n$ is a direct graph G over F in which:

- each node corresponds to a variable $X_i \in F$;
- a set of direct edges connects pairs of nodes (if there is an edge from node X_i to node X_j, X_i is said to be a parent of X_j , $X_i \in Pa(X_j)$); this defines a *dependency graph* in which each node X_i has $Pa(X_i)$ as its immediate predecessors.
- each node X_i has a conditional preference tables (CP-table) in which the user explicitly specifies her preference over the values of X_i for *each complete assignment* on $Pa(X_i)$. This preference is a total or partial order over the domain of X_i.

Note that the number of complete assignments over a set of variables is exponential in the size of the set. Throughout this paper, we assume there is an implicit constant that specifies the maximum number of parent features, $|Pa(X)|$, that any feature may have. With this restriction, and an implicit bound on $|D_X|$ (the domain of the variable X), we can and do treat the size of the conditional preference representation for any X as a constant.

An *acyclic* CP-net is one in which the dependency graph is acyclic. A CP-net need not be acyclic. For example, my preference for the entree may depend on the choice of the main course, and my preference for the main course may depend on the choice of the entree. However, in this paper we focus on acyclic CP-nets.

Example 4.1 Consider a CP-net (Figure 1) whose features are X_1, X_2, X_3, and X_4, with binary domains containing x and $\overline{x}$ if X is the name of the feature, and with the preference statements as follows: $x_1 > \overline{x}_1$, $x_2 > \overline{x}_2$, $(x_1 \wedge x_2) \vee (\overline{x}_1 \wedge \overline{x}_2) : x_3 > \overline{x}_3$, $(x_1 \wedge \overline{x}_2) \vee (\overline{x}_1 \wedge x_2) : \overline{x}_3 > x_3$, $x_3 : x_4 > \overline{x}_4$, $\overline{x}_3 : \overline{x}_4 > x_4$. Here, statement $x_1 > \overline{x}_1$ represents the unconditional preference for $X_1 = x_1$ over $X_1 = \overline{x}_1$, while statement $x_3 : x_4 > \overline{x}_4$ states that $X_4 = x_4$ is preferred to $X_4 = \overline{x}_4$, given that $X_3 = x_3$.

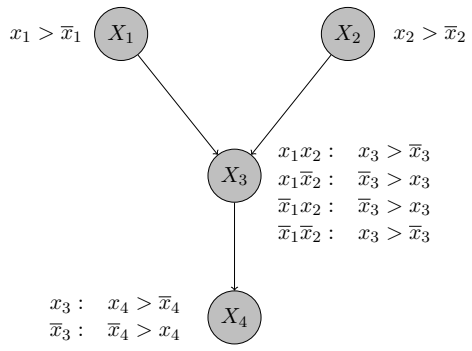


Figure 1. CP-net of Example 4.1.

The semantics of CP-nets depends on the notion of a *worsening flip*. A worsening flip is a change in the value of a variable to a value which is less preferred by the cp-statement for that variable. We say that one outcome α is *better* than another outcome β (written

$\alpha > \beta$) if and only if there is a chain of worsening flips from α to β. This definition induces a preorder over the outcomes.

In general, finding optimal outcomes and testing for optimality in this ordering is NP-hard. However, in acyclic CP-nets, there is only one optimal outcome and this can be found in as many steps as the number of features via a *sweep forward procedure* [2]. We sweep through the CP-net, following the arrows in the dependency graph and assigning at each step the most preferred value in the preference table. More formally:

Definition 4.2 The sweep forward procedure, in acyclic CP-nets, to achieve the optimal value is as follows:

- **Step 0 :** We choose for the CP-nets independent features the value that it is ranked first in the ordering of the domains values.
- **Step i :** For each feature that it is not yet assigned and that has the values assignments of all the parents, we choose the value that it is most preferred (so ranked first) in the ordering of the domains values (in the rows corresponding to that particular assignments for the parentss values).

Example 4.2 For instance, in the CP-net above (Example 4.1), we would choose $X_1 = x_1$ and $X_2 = x_2$, then, because the parents $X_1X_2 = x_1x_2$, we choose $X_3 = x_3$ and then, because $X_3 = x_3$ we obtain $X_4 = x_4$. The optimal outcome is therefore $x_1x_2x_3x_4$.

Each step in the sweep forward procedure is exponential in the number of parents of the current feature, and there are as many steps as features. In this paper we assume the number of parents is bounded, so this algorithm takes time polynomial in the size of the CP-net.

In the general case the optimal outcome coincides with the solutions of a constraint problem obtained replacing each cp-statement with a constraint. For example, the following cp-statement (concerning the Example 4.1) $(x_1 \wedge x_2) \vee (\overline{x}_1 \wedge \overline{x}_2) : x_3 > \overline{x}_3$ would be replaced by the constraint $(x_1 \wedge x_2) \vee (\overline{x}_1 \wedge \overline{x}_2) \Rightarrow x_3 > \overline{x}_3$.

Determining if one outcome is better than another according to this ordering (called a dominance query) is NP-hard even for acyclic CP-nets [5, 7]. Whilst tractable special cases exist, there are also acyclic CP-nets in which there are exponentially long chains of worsening flips between two outcomes.

Example 4.3 In the CP-net of the Example 4.1, $\overline{x}_1x_2\overline{x}_3\overline{x}_4$ is worse than $x_1x_2x_3x_4$.

4.1 Variants of CP-nets

CP-nets are deeply studied and widely used, thus many variants and generalization was developed. The main are the following:

- **TCP-nets** (Tradeoffs-enhanced CP-nets) [3] are a extension of CP-net that permits to represent also the notion of *relative importance* and *conditional relative importance*. A variable X is relative more important respect to a variable Y ($X \rhd Y$) if

the it is more important that the value of X be high than that the value of Y be high. For instance "The length of the journey is more important to me than the choice of airline". A variable X is relative more important respect to a variable Y given an assignment of a variable Z ($Z = z_0 : X \rhd Y, Z = z_1 : Y \rhd X$) if the it is more important that the value of X be high than that the value of Y be high, given the evidence that $Z = z_0$. For instance "The length of the journey is more important to me than the choice of airline provided that I am lecturing the following day. Otherwise the choice of airline is more important". This formulation introduce the *CI-tables (conditional importance tables)* to the dependency graph, to express the conditional importance statements.

- **UCP-nets** (Utility Conditional Preference networks) [1] are an extension of CP-nets in which the preferences are expressed in a quantitative formulation, and not only qualitative as in CP-nets. Each varible express the preferences of the user in a table that contains the user utility values of the values in the variable domain. A CP-statement of the form $x : y > \overline{y}$ is represented in UCP-net by $f_Y(y|x) = u_1$, $f_Y(\overline{y}|x) = u_2$ and $u_1 > u_2$ where f_Y is the utility function of the variable Y.

- **mCP-nets** (multi-agent Conditional Preference networks) [8] are an extension of CP-net formalism to model and handle the preferences of multiple agents. A mCP-net is a set of m partial CP-nets (CP-net with partial or missing information) which can share some features. A mCP-net dependency graph is obtained by combining the graphs of the partial of the parial CP-nets having one occurrence of each shared feature. A feature can be *shared* between the CP-nets, *visible* to other CP-nets that can use this feature as precondition in the CP-statements, or *private* if is ranked in only one CP-net and is not visible to the other CP-nets.

5 CP-theories and Comparative Preference languages

Other important generalizations of CP-nets are CP-theories and comparative preference languages, two different logic formulation of qualitative preferences. The corresponding theories and languages are defined below.

5.1 CP-theories

CP-theories are introduced in [9] as a logic of conditional preference which generalizes CP-nets.

Definition 5.1 Given a set of variables $Var = \{X_1, \cdots, X_N\}$ with domains D_i, $i = 1, \ldots, n$, the language L_{Var} is defined by all the statements of the form: $u : x_i \succ x_i'[W]$ where u is an assignment of a set of variables $U \subseteq Var \setminus \{X_i\}$, $x_i \neq x_i' \in D_i$ and W is a set of variables such that $W \subseteq (Var \setminus U \setminus \{X_i\})$.

Definition 5.2 Given a language L_{Var} as defined above, a *conditional preference theory (CP-theory)* Γ on Var is a subset of L_{Var}. Γ generates a set of preferences that cor-

responds to the set $\Gamma^* = \bigcup_{\varphi \in \Gamma} \varphi^*$ where given $\varphi = u : x_i \succ x_i'[W]$, φ^* is defined as $\varphi^* = \{(tuxw, tux'w') : t \in Var \setminus (\{X_i \cup U \cup W\}),\ w, w' \in W\}$.

A CP-net is a particular case of a CP-theory where $W = \emptyset$ for all $\varphi \in \Gamma$.

Two graphs are associated to a CP-theory: $H(\Gamma) = \{(X_j, X_i) | \exists \varphi \in \Gamma \text{ s.t. } \varphi = u : x_i \succ x_i'[W] \text{ and } X_j \in U\}$ and $G(\Gamma) = H \cup \{(X_i, X_j) | \exists \varphi \in \Gamma \text{ s.t. } \varphi = u : x_i \succ x_i'[W] \text{ and } X_j \in W\}$.

The semantics of CP-theories depends on the notion of a *worsening swap*, which is a change in the assignment of a set of variables to an assignment which is less preferred by a rule $\varphi \in \Gamma$. We say that one outcome o is better than another outcome o' ($o \succ o'$) if and only if there is a chain of worsening swaps (a *worsening swapping sequence*) from o to o'.

Definition 5.3 A CP-theory Γ is *locally consistent* if and only if for all $X_i \in Var$ and $u \in Pa(X_i)$ in the graph $H(\Gamma)$, $\succ_u^{X_i}$ is irreflexive.

Local consistency can be determined in time proportional to $|\Gamma|^2 N$. Given a CP-theory Γ, if the graph $G(\Gamma)$ is acyclic, Γ is consistent if and only if Γ is locally consistent, thus global consistency has the same complexity as local consistency given an acyclic graph $G(\Gamma)$.

5.2 Comparative Preference languages

Comparative Preference theories [10] are an extension of CP-theories.

Definition 5.4 The comparative preference language $\mathcal{CL}_{Var}$ is defined by all statements of the form: $p > q || T$ where P, Q and T are subsets of Var and p and q are assignments respectively of the variables in P and in Q.

Definition 5.5 Given a language $\mathcal{CL}_{Var}$ as defined above, a *comparative preference theory* Λ on Var is a subset of $\mathcal{CL}_{Var}$. Λ generates a set of preferences that corresponds to the set $\Lambda^* = \bigcup_{\varphi \in \Lambda} \varphi^*$ where if $\varphi = p > q || T$, φ^* is defined as a pair (α, β) of outcomes such that α extends p and β extends q and α and β agree on T: $\alpha \restriction_T = \beta \restriction_T$.

References

[1] C. Boutilier, F. Bacchus, and R. I. Brafman, *Ucp-networks: A directed graphical rep- resentation of conditional utilities.* In UAI ?01: Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence, University of Washington, Seattle, Washington, USA, August 2-5, 2001, pages 56?64, 2001.

[2] C. Boutilier, R. Brafman, C. Domshlak, H. Hoos, and D. Poole, *CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements.* Journal of Artificial Intelligence Research 21 (2004), 135–191.

[3] R. Brafman and C. Domshlak, *Introducing variable importance tradeoffs into cp-nets.* In UAI ?02, pages 69–76, 2002.

[4] B. D?Ambrosio, *Inference in bayesian networks.* AI Magazine, 20(2):21, 1999.

[5] C. Domshlak and R. Brafman, *CP-nets: Reasoning and consistency testing.* In Proc. 8th International Conference on Principles of Knowledge Representation and Reasoning (KR), 2002.

[6] C. Domshlak, F. Rossi, K. Venable, and T. Walsh, *Reasoning about soft constraints and conditional preferences: complexity results and approximation techniques.* In Proc. of the 18th International Joint Conference on Artificial Intelligence (IJCAI), 2003.

[7] J. Goldsmith, J. Lang, M. Truszczynski, and N. Wilson, *The Computational Com- plexity of Dominance and Consistency in CP-nets.* Journal of Artificial Intelligence Research 33/1 (2008), 403–432.

[8] F. Rossi, K. Venable, and T. Walsh, *mCP nets: representing and reasoning with preferences of multiple agents.* In Proc. of the 19th AAAI Conference on Artificial Intelligence (AAAI), 2004.

[9] N. Wilson, *Extending CP-Nets with Stronger Conditional Preference Statements.* In Proceedings of AAAI-04, pages 735–741, 2004.

[10] N. Wilson, *Efficient Inference for Expressive Comparative Preference Languages.* In Proceedings of IJCAI-09, 2009.

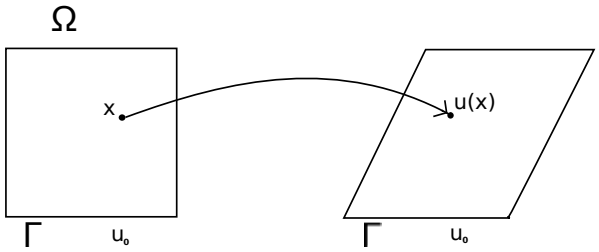
Variational methods in Nonlinear Elasticity: an introduction

Alice Fiaschi (*)

Abstract. After a brief introduction of the variational formulation for the standard model in nonlinear elasticity, we consider the problem of finding the "right" space to describe the equilibrium configurations of an elastic body, from the point of view of the Calculus of Variations. In this framework, we introduce the space of Young measures as a suitable space to describe materials exhibiting microstructures.

1 The standard model in nonlinear elasticity

We can describe the standard model in nonlinear elasticity as follows:



- *Reference configuration* (represents the space occupied by the undeformed body): a bounded connected domain $\Omega \subset \mathbb{R}^3$ with "smooth" boundary;
- *Deformation* (describes the transformation which occurs to our material): a function $u\colon \Omega \to \mathbb{R}^3$;
- *Prescribed boundary deformation* (imposes a constraint on the possible changes in the material configuration): a function $u_0\colon \Gamma \subseteq \partial\Omega \to \mathbb{R}$ such that $u = u_0$ on Γ;
- *Stored energy density* (represents the energy density associated to a deformation): a function $W\colon \Omega \times \mathbb{M}^{3\times 3} \to [0, +\infty]$;

(*)Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: afiaschi@math.unipd.it. Seminar held on May 6th, 2015.

- *Body forces density* (describes the other forces acting on the material): a function $f\colon \Omega \times \mathbb{R}^3 \to \mathbb{R}$.

To have a physically consistent model, we need to impose some conditions on the stored energy density W:

- *Frame-indifference*, to guarantee that no change of energy is associated with rotations:

$$W(x,F) = W(x,RF) \quad \text{for every } x \in \Omega,\ F \in \mathbb{M}^{3\times 3},\ R \in SO(3);$$

- *Orientation preservation*, to select only deformations which preserve the orientation:

$$W(x,F) = +\infty \quad \text{if } \det F \leq 0;$$

- *Non-interpenetration condition*, to exclude deformations which are not "invertible":

$$W(x,F) \to +\infty \quad \text{as } \det F \to 0^+;$$

- *Energy growth*, to guarantee that infinite energy is associated with infinite strain:

$$W(x,F) \to +\infty \quad \text{as } |F| \to +\infty.$$

Example $W(F) = a|F|^2 + g(\det F)$, with $\lim_{t\to 0} g(t) = +\infty$, is the energy of an Ogden material.

A possible way to guarantee *energy growth* is a p-growth condition from below:

$$W(x,F) \geq a|F|^p + b(x)\,.$$

We denote the *total energy* of the body by I, so that

$$I(u) = \int_\Omega W(x, \nabla u(x))\ dx - \int_\Omega f(x,u(x))\ dx\,,$$

where u is a deformation. The *equilibrium configurations* of our model are the minimizers of I satisfying the boundary condition $u = u_0$ on $\partial\Omega$. In particular, if w, f and u are smooth enough, the equilibrium configurations satisfy the Euler-Lagrange equation:

$$-\nabla \cdot \Big(\frac{\partial W}{\partial F}(x, \nabla u(x))\Big) = \frac{\partial f}{\partial u}(x,u(x))\,.$$

The questions are: *do minimizers exist*? And, if any, *in which space*?

A first attempt is *to look for minimizers among C^2 functions* in order to use Euler-Lagrange equation: but this approach is not always sucessful, as the following classical example shows.

Example We look for minimizers of

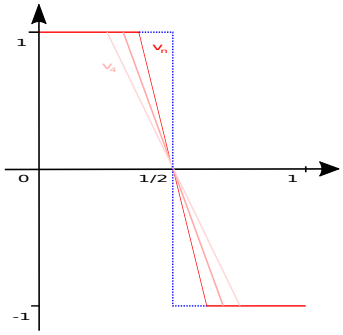
$$I(u) := \int_0^1 (u'^2(x) - 1)^2 \, dx$$

among $u \in C_0^1(0,1) := \{u \in C^1(0,1),\ u(0) = u(1) = 0\}$. We claim that

$$\inf_{C_0^1(0,1)} I(u) = 0\,.$$

Namely, let $u_n(x) := \int_0^x v_n(t)\, dt$, where

$$v_n(t) := \begin{cases} 1 & \text{for } t \in [0, \frac{1}{2} - \frac{1}{n}] \\ n\left(\frac{1}{2} - x\right) & \text{for } t \in [\frac{1}{2} - \frac{1}{n}, \frac{1}{2} + \frac{1}{n}] \\ -1 & \text{for } t \in [\frac{1}{2} + \frac{1}{n}, 1]\ : \end{cases}$$

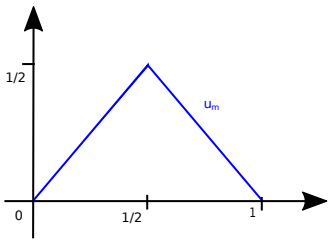


then it is easy to see that $I(u_n) = \int_0^1 (v_n^2 - 1)^2\, dx \to 0$ as $n \to +\infty$, and this proves the claim. But if $u \in C_0^1(0,1)$ is such that $I(u) = 0$, then $u'(x) \in \{1, -1\}$ for every $x \in [0,1]$, hence $u' \equiv 1$ or $u' \equiv -1$ (since u' is continuous), and therefore u could not satisfy the required boundary condition $u(0) = u(1) = 0$. Hence the minimum problem has NO solution in C^1, and *a fortiori* in C^2.

On the other hand, if we consider the pointwise limit $u_{\min}$ of the sequence u_n:

$$u_{\min}(x) := \begin{cases} x & \text{for } x \in [0, \frac{1}{2}] \\ 1 - x & \text{for } x \in [\frac{1}{2}, 1] \end{cases}$$

then $I(u_{\min}) = 0\,!$



Now, $u_{\min}$ clearly does not belong to $C_0^1(0,1)$ but rather to $W_0^{1,4}(0,1)$ (see below), so this function should be viewed as a minimizer of $I(u)$ in this larger space.

Recall that u is an element of the **Sobolev space** $W^{1,p}(0,1)$ if $u \in L^p(0,1)$ and there exists a function $v \in L^p(0,1)$ such that for every $\phi \in C_c^\infty(0,1)$ it holds $\int_0^1 u(x)\phi'(x)\, dx = -\int_0^1 v(x)\phi(x)\, dx$ (the function v is called the *weak derivative* of u and is denoted by u'). The norm on $W^{1,p}(0,1)$ is $\|u\|_{W^{1,p}} = \|u\|_{L^p} + \|u'\|_{L^p}$. The space $W_0^{1,p}(0,1)$ is defined as the closure of $C_0^\infty(0,1)$ in $W^{1,p}(0,1)$ with respect to the norm $\|u\|_{W^{1,p}}$.

The weak derivative of $u_{\min}$ is $v_{\min}(x) := \begin{cases} 1 & \text{for } x \in [0, \frac{1}{2}) \\ -1 & \text{for } x \in (\frac{1}{2}, 1] \end{cases}$: indeed

$$\begin{aligned} \int_0^1 u_{\min}(x)\phi'(x)\,dx &= \int_0^{\frac{1}{2}} x\phi'(x)\,dx + \int_{\frac{1}{2}}^1 (1-x)\phi'(x)\,dx \\ &= \Big[x\phi(x)\Big]_0^{1/2} - \int_0^{\frac{1}{2}} \phi(x)\,dx + \Big[(1-x)\phi(x)\Big]_{\frac{1}{2}}^1 - \int_{\frac{1}{2}}^1 (-1)\phi(x)\,dx \\ &= \frac{1}{2}\phi(\tfrac{1}{2}) - \Big(1-\tfrac{1}{2}\Big)\phi(\tfrac{1}{2}) - \int_0^1 v_{\min}(x)\phi(x)\,dx \;=\; -\int_0^1 v_{\min}(x)\phi(x)\,dx\,. \end{aligned}$$

Hence $u_{\min} \in W^{1,4}(0,1)$ and $I(u_{\min}) = \int_0^1 (v_{\min}(x)^2 - 1)\,dx = 0$, as claimed.

2 The Direct Method of the Calculus of Variations

The previous example is in fact an instance of a more general situation.

Theorem (Existence Theorem) *Let $p > 1$, and let $W\colon \Omega \times \mathbb{R}^3 \times \mathbb{M}^{3\times 3} \to [0, +\infty]$ be a continuous map such that*

- *$W(x,u,F) \geq a|F|^p + b|u|^q + c(x)$, for $a > 0$, $b \in \mathbb{R}$, $c \in L^1(\Omega)$, $1 \leq q < p$* (coercivenes*s*);
- *the map $F \mapsto W(x,u,F)$ is convex* (convexity).

Then there exists (at least) one minimizer of $I(u) = \int_\Omega W(x, u(x), \nabla u(x))\,dx$ in $u_0 + W_0^{1,p}(\Omega)$, provided $I(u_0) < +\infty$.

A *minimizing sequence* is a sequence $u_n \in u_0 + W_0^{1,p}(\Omega)$ with $I(u_n) \to \inf I$.

We just give an idea of the proof of the existence theorem:

- coerciveness implies the compactness of a minimizing sequence in the weak topology of $W^{1,p}$, hence $u_{n_k} \rightharpoonup u$ for some subsequence $(u_{n_k})_k$;
- convexity implies the lower semicontinuity of I with respect to the weak topology of $W^{1,p}$, hence $I(u) \leq \liminf I(u_{n_k}) = \inf I$.

The point is rather that *the hypothesis of convexity is incompatible with our physical assumptions on W*, as the following simple example shows.

Example Recall that among our physical assumptions we required that $W(RF) = W(F)$ for every $R \in SO(3)$, and that $W(F) \to +\infty$ if $\det F \to 0$. Now, for simplicity let us consider $F \in \mathbb{M}^{2\times 2}$ and suppose $W(\mathbf{1}) < +\infty$. Let $R = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$. Since $\frac{1}{2}R + \frac{1}{2}R^T = (\cos\alpha)\mathbf{1}$, convexity implies that $W((\cos\alpha)\mathbf{1}) \leq \frac{1}{2}W(R) + \frac{1}{2}W(R^T) = W(\mathbf{1}) < +\infty$; but for $\alpha \to \frac{\pi}{2}$, $\det((\cos\alpha)\mathbf{1}) \to 0$ although $W((\cos\alpha)\mathbf{1})$ is bounded from above by $W(\mathbf{1}) < +\infty$.

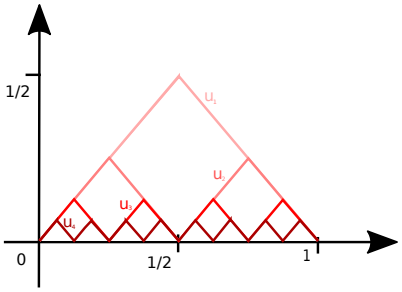
So, *could we do without convexity ?* If we are looking for a minimizer in the usual sense of function, the general answer is no, as the following examples show.

Example (Bolza example) We look for minimizers of

$$I(u) = \int_0^1 [(u'^2 - 1)^2 + u^2]\, dx \qquad \text{in } W_0^{1,4}(0,1).$$

Let

$$\begin{aligned} u_1(x) &:= \begin{cases} x & \text{for } x \in [0, \frac{1}{2}] \\ 1-x & \text{for } x \in [\frac{1}{2}, 1] \end{cases} \\ u_n(x) &:= \frac{1}{2^{n-1}}\, u_1(\mathrm{frac}(2^{n-1}x)) \quad \text{for } n \geq 1 \end{aligned}$$



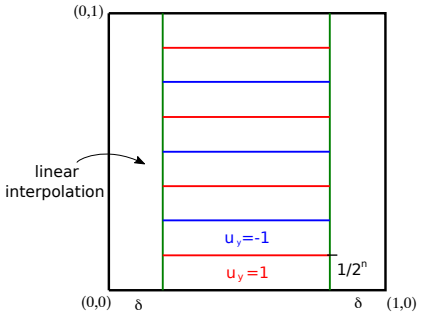
(here frac : $\mathbb{R} \to [0,1[$ denotes the fractional part). Since $I(u_n) \leq 1/2^{2n}$, it follows that $\inf I(u_n) = 0$; but $I(u) = 0$ implies $u \equiv 0$, hence $u' \equiv 0$. Therefore there are no minimizers for $I(u)$ in $W_0^{1,4}$!

Actually, in the previous example we deal with an energy density ($(\xi, F) \mapsto (|F|^2 - 1) + \xi^2$) depending not only on $(x, \nabla u)$ as our W does, but also on u. Unfortunately the lack of convexity of W gives troubles also in the case of energy density depending just on ∇u, as the following example shows.

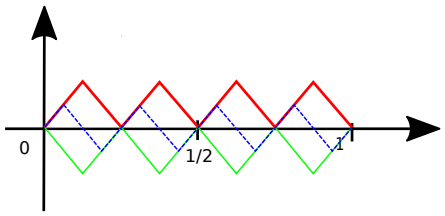
Example We look for minimizers of

$$I(u) = \int_0^1 \int_0^1 [(u_y^2 - 1)^2 + u_x^2]\, dx\, dy \qquad \text{in } W_0^{1,4}([0,1] \times [0,1]).$$

We consider the zig zag function u_1 of the previous example, and define a sequence $u_n^\delta(x,y) := \frac{1}{2^{n-1}} u_1(\mathrm{frac}(2^{n-1}y))$ for $\delta < x < 1 - \delta$ (depending only on y) with $0 < \delta \ll 1$, and we use linear interpolation to achieve the boundary values at $x = 0$ and $x = 1$. Considering first the limit as $n \to \infty$ and then as $\delta \to 0$ (with $\frac{1}{n\delta}$ remaining bounded) we obtain $\inf I(u) = 0$. But the requirements $u_x = 0$ and $u_y = \pm 1$ are incompatible with the boundary conditions. Therefore, once more, there are no minimizers for $I(u)$ in $W_0^{1,4}$.



If we cannot hope to find a function minimizing our energy functional, we are forced to describe our model through *minimizing sequences*. In this change of perspective, the first remark is that *for a given variational problem there are usually many minimizing sequences*: for example, the picture on the right shows different minimizing sequences for Bolza example.



The question is therefore the following: *is it possible to describe the "macroscopic" features" of these sequences ?*
An object describing the macroscopic behaviour of a sequence $u_n\colon \Omega \to \mathbb{R}^d$ should at least determine the limits of

$$\int_\Omega f(u_n)\ dx,$$

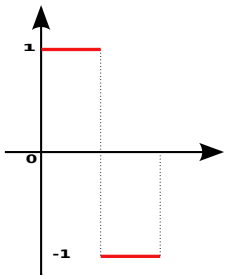
for continuous functions f. With this purpose in mind we introduce *Young measures*.

3 Young measures

Let us start with a particular case: what can we say of the limit of

$$\int_0^1 g(x) f(v(nx))\ dx \qquad \text{for } f \in C(\mathbb{R})\ ,\ g \in L^1(0,1)$$

as $n \to +\infty$, for v the function of $L^\infty(0,1)$ in the picture on the right?



A useful tool to treat such a question is the classical

Proposition (Riemann-Lebesgue Lemma) *Let $w \in L^\infty(a,b)$, extend it by periodicity to $\mathbb{R}$ and set $w_n(x) := w(nx)$. Then*

$$w_n \ \rightharpoonup^* \ \bar w := \frac{1}{b-a}\int_a^b w(y)\ dy \qquad \text{in } L^\infty(a,b).$$

In our case $w(x) = f(v(x))$: hence

$$\int_0^1 g(x) f(v(nx))\ dx \to \int_0^1 g(x) \int_0^1 f(v(y))\ dy\ dx$$
$$= \int_0^1 g(x)\Big[\frac{1}{2}f(1) + \frac{1}{2}f(-1)\Big]\ dx = \int_0^1 \int_{\mathbb{R}} g(x) f(\xi)\ d\Big[\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}\Big](\xi)\ dx,$$

where δ_{ξ_0} is the Delta of Dirac:

$$\delta_{\xi_0}(A) = \begin{cases} 1 & \text{if } \xi_0 \in A \\ 0 & \text{if } \xi_0 \notin A \end{cases}\ , \qquad \int_{\mathbb{R}} f(\xi)\ d\delta_{\xi_0}(\xi) = f(\xi_0).$$

In other words we have $\displaystyle\int_0^1 g(x) f(v(nx))\ dx \to \int_0^1 g(x) f(\xi)\ d\mu(x,\xi)$ where μ is the measure defined as

$$\int_{(0,1)\times\mathbb{R}} \varphi(x,\xi)\, d\mu(x,\xi) \ = \ \int_0^1 \int_{\mathbb{R}} \varphi(x,\xi)\ d\Big[\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}\Big](\xi)$$

for every bounded Borel function φ. We say that μ is the Young measure associated with the minimizing sequence v_n. In general, a *Young measure* μ on Ω with values in $\mathbb{R}^d$ can be identified with a measurable family $(\mu^x)_{x\in\Omega}$ of probability measures μ^x on $\mathbb{R}^d$ by

$$\int_{\Omega\times\mathbb{R}} \varphi(x,\xi)\, d\mu(x,\xi) \;=\; \int_\Omega \int_{\mathbb{R}} \varphi(x,\xi)\; d\mu^x(\xi)\, dx$$

for every bounded Borel function φ.

Coming back to our modeling problem, we need to translate it in terms of Young measures as summarized in the following table, where the question marks represent the still unclear translations:

Functions	**Young measures**
$v\colon \Omega \to R^d$	$(\delta_{v(x)})_{x\in\Omega}$ measure on $\Omega\times\mathbb{R}^d$
$\displaystyle\int_\Omega W(x,v(x))\, dx$	$\displaystyle\int_{\Omega\times\mathbb{R}^d} W(x,\xi)\; d\delta_{v(x)}(\xi)dx$
$\|v_n\|_p$ bounded $(p>1)$ $\Rightarrow v_n \rightharpoonup v$ weakly in L^p (up to a subsequence)	**?**
$\int_\Omega W(x,v_n(x))\, dx$ not lower semicontinuous with respect to weak convergence in L^p	$\int_{\Omega\times\mathbb{R}^N} W(x,\xi)\, d\delta_{v_n(x)}(\xi)dx$ **?**

To clear these question marks out, we must define a suitable topology on the space $Y(\Omega;\mathbb{R}^d)$ of Young measures and clarify some related notions: let us give a brief sketch of that.

Let $C_0(\Omega\times\mathbb{R}^d)$ be the space of continuous functions φ such that the sets $\{|\varphi|\geq\varepsilon\}$ are compact for any $\varepsilon>0$, and let $M_b(\Omega\times\mathbb{R}^d)$ be the space of bounded Radon measures, i.e. the dual space of $C_0(\Omega\times\mathbb{R}^d)$: therefore, for $\mu\in M_b(\Omega\times\mathbb{R}^d)$ and $\varphi\in C_0(\Omega\times\mathbb{R}^d)$ we have a duality represented by $\langle\varphi,\mu\rangle = \int_{\Omega\times\mathbb{R}^d}\varphi\, d\mu$. From this duality we inherit a notion of *weak* convergence*, as follows:

$$\mu_n \rightharpoonup^* \mu \quad\Longleftrightarrow\quad \int \varphi\; d\mu_n \to \int \varphi\; d\mu \qquad \text{for any } \varphi\in C_0(\Omega\times\mathbb{R}^d).$$

A Young measure $\mu=(\mu^x)_{x\in\Omega}$ is an element of $M_b^+(\Omega\times\mathbb{R}^d)$ satisfying $\mu(A\times\mathbb{R}^d)=\mathcal{L}^N(A)$ for every measurable set $A\subset\Omega$: hence in particular $Y(\Omega;\mathbb{R}^d)$ inherits the weak* topology

from $M_b(\Omega \times \mathbb{R}^d)$. Moreover, if φ is continuous and $\varphi \geq 0$, then the function $\mu_n \mapsto \int_{\Omega\times\mathbb{R}^d} \varphi(x,\xi)\, d\mu_n$ is *lower semicontinuous* with respect to weak* convergence.

The *p-moments* of a Young measure μ are defined as $\int_{\Omega\times\mathbb{R}^d} |\xi|^p \, d\mu$. If the p-moments of a sequence $(\mu_n)_n$ are equibounded we can deduce that $(\mu_n)_n$ converges weakly*, up to a subsequence.
Now, the p-moments of a sequence of Young measures μ_n associated with functions $(\nabla u_n)_n$ (i.e. $\mu_n = (\delta_{\nabla u_n(x)})_{x\in\Omega}$) are $\|\nabla u_n\|_{L^p}$. Since the growth condition $W(x,F) \geq a|F|^p + b(x)$ implies $\|\nabla u_n\|_{L^p} < C$ for any n, if u_n is a minimizing sequence we deduce that the p-moments of minimizing sequences are equibounded.
Hence, thanks to the compactness property and the lower semicontinuity property described here before, using the Direct Method we can find a minimizer in $Y(\Omega;\mathbb{R}^d)$.

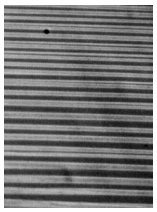
We can then complete the table above:

Functions	**Young measures**
$v\colon \Omega \to R^d$	$(\delta_{v(x)})_{x\in\Omega}$ measure on $\Omega \times \mathbb{R}^d$
$\displaystyle\int_\Omega W(x,v(x))\,dx$	$\displaystyle\int_{\Omega\times\mathbb{R}^d} W(x,\xi)\; d\delta_{v(x)}(\xi)dx$
$\|v_n\|_p$ bounded $(p>1)$ $\Rightarrow v_n \rightharpoonup v$ weakly in L^p (up to a subsequence)	$(\delta_{v_n(x)})_x$ with equibounded p-moments $\Rightarrow (\delta_{v_n(x)})_x \rightharpoonup \mu$ weakly* (up to a subsequence)
$\int_\Omega W(x,v_n(x))\,dx$ not lower semicontinuous with respect to weak convergence in L^p	$\int_{\Omega\times\mathbb{R}^N} W(x,\xi)\, d\delta_{v_n(x)}(\xi)dx$ lower semicontinuous with respect to weak* convergence

4 An example: Young measures and microstructures

A *microstructure* is any structure on a scale between macroscopic scale and the atomic scale. In a *variational model*, microstructures are formed to try to satisfy an optimality property.

Example Solid-solid phase transitions in certain elastic crystals (the picture shows a microstructure in Cu-Al-Ni).



- Above the critical temperature θ_0, $W_\theta(A) = 0$
- Below the critical temperature, change of stability of the crystal structure: $W_\theta(B) = 0$
- At the critical temperature both phases are stable: $W_{\theta_0}(A) = W_{\theta_0}(B) = 0$

Boundary conditions are typically caused by contacts with other part of the crystal where other deformation gradients prevail.

A nontrivial boundary condition $\lambda A + (1-\lambda)B$ causes no exact solutions: finer and finer mixtures of A and B are requested, and this is ilustrated by an equilibrium configuration given in terms of Young measures.

References

[1] B. Dacorogna, "Introduction to the Calculus of Variations". Third Edition. Imperial College Press, London, 2015.

[2] S. Müller, *Variational Models for microstructure and phase transitions.* Calculus of Variations and geometric evolution problems (Cetraro, 1996). Lecture Notes in Math., 1713, Springer, Berlin, 1999.

[3] P. Pedregal, "Parametrized Measures and Variational Principles". Birkhäuser-Verlag, Basel, Switzerland, 1997.

[4] P. Pedregal, "Variational Methods in Nonlinear Elasticity". Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2000.

Controllability and the numerical approximation of the minimum time function

Thuy T.T. Le (*)

Abstract. In optimal control theory, minimum time problems are of interest since they appear in many applications such as robotics, automotive, car industry, etc.. The scope of this talk is to give a brief introduction of these problems. Controllability conditions under various settings are considered. Such conditions play a vital role in studying the regularity of the minimum time function $T(x)$. Moreover, we will also introduce the HJB equation associated with a minimum time problem and approaches to computing $T(x)$ approximately.

1 Preliminary

1.1 Control theory

A standard reference is [1]. Consider the following controlled dynamics in $\mathbb{R}^n$

$$\begin{cases} \dot{y}(t) = f(y(t), u(t)), t \in [0, +\infty) \text{ a.e.} \\ y(0) = \xi \in \mathbb{R}^n, \end{cases} \tag{1.1}$$

where

- $U \in \mathbb{R}^m$ is a control set
- $u\colon [0, +\infty) \to U$ is a control function

Let $S \in \mathbb{R}^n$ be a target set and the set of admissible controls is defined as

$$\mathcal{U}_{ad} = \{u\colon [0, +\infty) \to U : u \text{ is measurable}\}.$$

The following assumptions are supposed to be satisfied in the sequel.

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: lttthuy@math.unipd.it . Seminar held on May 27th, 20145.

Assumptions 1.1

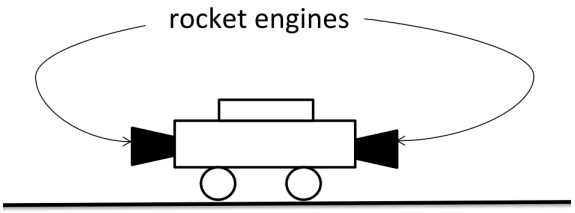
(1) $f(x,u)$ is C^∞ and all partial derivatives are Lipschitz with Lipschitz constant $L>0$ w.r.t x, uniformly continuous w.r.t u; moreover, $\|f(y,u)\| \leq K_0(1+\|y\|)$ for all $y \in \mathbb{R}^n$, where K_0 is a positive constant

(2) $S \in \mathbb{R}^n$ is compact

(3) $U \in \mathbb{R}^m$ is compact

The minimum time problem is the one that determines a control, subject to its constraints, which drives the initial state $\xi \in \mathbb{R}^n \setminus S$ to the given target S in the smallest amount of time.

Example 1.2 Imagine a railroad car powered by rocket engines on each side.



Let

+ $x(t)$ is the position at time t
+ $y(t) = \dot{x}(t)$ is the velocity at time t
+ $u(t) \in [-1,1]$ is the thrust from rockets, where its sign depends upon on which engine is firing.

Assume the car has unit mass, then the law of the motion is $\ddot{x}(t) = u(t)$. The goal is to figure out how to fire the rockets so that to arrive at the origin with zero velocity in a minimum amount of time.

Given any $\xi \in \mathbb{R}^n$ and $u \in \mathcal{U}_{ad}$, a solution of

$$\begin{cases} \dot{y}(t) = f(y(t), u(t)) \\ y(0) = \xi, \end{cases}$$

denoted by $y(\cdot, \xi, u)$ is called a trajectory starting from ξ associated with the control u. For fixed $\xi \in \mathbb{R}^n \setminus S$, the *minimum time starting from ξ to reach the target $\mathcal{S}$* for some $u \in \mathcal{U}_{ad}$ is as

$$t_S(\xi, u) = \min\{t \geq 0 : y(t, \xi, u) \in S\} \leq +\infty.$$

Then the minimum time function to reach S from ξ is defined as

$$T_S(\xi) = \inf_{u \in \mathcal{U}_{ad}} t_S(\xi, u).$$

We also define the reachable sets for time $t \geq 0$ as follows.

$$\mathcal{R}^S(t) = \{\xi \in \mathbb{R}^n : T_S(\xi) < t\}$$

is the set of starting points from which the system can reach the target in time less than t.

$$\mathcal{R}^S = \{\xi \in \mathbb{R}^n : T_S(\xi) < +\infty\}$$

is the set of starting points from which the system can reach the target in finite time.

1.2 Nonsmooth analysis

Now we will introduce some basic notions in nonsmooth analysis which are needed in this context, see [3].

Definition 1.3 Let $S \in \mathbb{R}^n$ be closed and $\rho > 0$ be given. S satisfies a ρ–internal sphere condition if for any $x \in S$ there exists y such that $x \in \overline{B(y,\rho)} \subset S$.

Definition 1.4 Let $\Omega \subset \mathbb{R}^n$ be an open set. A function $g\colon \Omega \to \mathbb{R}$ is locally semiconcave if for every $x \in \Omega$ there exists a ball $B(x,r)$ and a positive constant C such that

$$\lambda g(y) + (1-\lambda)g(y') \leq g(\lambda y + (1-\lambda)y') + C\left\|y - y'\right\|^2 \tag{1.2}$$

for all $y, y' \in B(x,r)$ and all $\lambda \in [0,1]$.

Global semiconcavity means that the above inequality is satisfied by every $y, y' \in \Omega$ such that the segment $[y,y'] \subset \Omega$ with the same constant C. The constant C appearing in (1.2) is labeled as semiconcavity constant.

Example 1.5 The distance function to $\overline{B(0,r)}$, $d_{\overline{B(0,r)}}(x) = |x| - r$, is semiconcave with semiconcavity $\frac{1}{r}$.

Definition 1.6 Let $\Omega \subset \mathbb{R}^n$ be an open set, $g\colon \Omega \to \mathbb{R}$. We say that a vector p belongs to the proximal superdifferential of g at x (notated by $p \in \partial^P g(x)$) if there exist $\sigma > 0$ such that

$$g(y) \leq g(x) + \langle p, y - x\rangle + \sigma\left\|y - x\right\|^2$$

for all y in a neighborhood of x.

2 Controllability

This section is devoted to small–time controllability of a given system and sufficient conditions to guarantee that property.

Definition 2.1 The system (f, U) is

(1) small–time controllable on S (briefly STCS) if $S \subset \mathrm{int}(\mathcal{R}^S(t))$ for all $t > 0$.

(2) fully controllable if $\mathcal{R}^S = \mathbb{R}^n$.

Let S_δ be an enlargement of S defined as follows $S_\delta = \{x \in \mathbb{R}^n : d_S(x) < \delta\}$. The following proposition shows the connection between small–time controllable property and the regularity of the minimum time function.

Proposition 2.2 (see [1]) *Under the standard assumptions and STCS,*

(1) $\mathcal{R}^S$ *is open;*

(2) T_S *is continuous in* $\mathcal{R}^S$*;*

(3) $T_S(x) \leq \omega(\mathrm{d}_S(x))$ *for all* $x \in S_\delta$*, where* $\omega\colon [0,\delta] \to [0,+\infty)$ *and* $\lim_{s\to 0^+} \omega(s) = 0$*.*

(4) $\lim_{x\to x_0} T_S(x) = +\infty$ *for any* $x_0 \in \partial\mathcal{R}^S$*.*

Remark 2.3 If $\omega(\mathrm{d}_S(x)) = C\ \mathrm{d}_S(x), C > 0$, $T_S(x)$ is locally Lipschitz continuous. If $\omega(\mathrm{d}_S(x)) = C\,\mathrm{d}_S(x)^\gamma, C > 0, 0 < \gamma < 1$, $T_S(x)$ is locally Hölder continuous with exponent γ.

Assume that ∂S is smooth enough, and the so called *Petrov condition* holds, i.e.

$$\inf_{u\in U} \langle \nabla d_S(x), f(x,u)\rangle \leq -\mu < 0,$$

for x in a neighborhood S_δ of S, then any point of S_δ can be steered to S in finite time $T_S(x)$ and $T_S(x)$ is Lipschitz continuous.
Let consider the following particular system

$$f(x,u) = g_1(x)u_1 + g_2(x)u_2, \tag{2.1}$$

by subsequently following the flows of $g_1(\cdot), g_2(\cdot), -g_1(\cdot), -g_2(\cdot)$, each one for time t, it is well known that

$$y(4t) = \xi + t^2[g_1,g_2](\xi) + O(t^3),$$

where $[g_1,g_2](x) = \nabla g_2(x)g_1(x) - \nabla g_1(x)g_2(x)$. Let ∂S be of class C^2 and assume

$$\langle \nabla d_S(\xi), [g_1,g_2](\xi)\rangle \leq -\mu,$$

for ξ in a neighborhood S_δ of S, then $d_S(\cdot)$ is of class $C^{1,1}$ in $S_\delta \setminus S$ and so

$$\begin{aligned} d_S(y(4t)) &= d_S(\xi) + \langle \nabla d_S(\xi), y(4t) - \xi\rangle + O(\|y(4t)-\xi\|^2) \\ &= d_S(\xi) + t^2\langle \nabla d_S(\xi), [g_1,g_2](\xi)\rangle + O(t^3) \leq d_S(\xi) - t^2\mu + O(t^3) \leq d_S(\xi) - t^2\frac{\mu}{2}. \end{aligned}$$

Moreover, $T_S(x)$ can be proved to be $\frac{1}{2}$- Hölder continuous on S_δ.
If ∂S is not smooth, but S satisfies a $\rho-$ internal sphere condition, then $d_S(\cdot)$ is semiconcave with the semiconcavity constant $\frac{1}{\rho}$ in $\overline{S^c}$, i.e for every $x, y \in \overline{S^c}$

$$d_S(y) \leq d_S(x) + \langle \partial^P d_S(x), y - x\rangle + \frac{1}{\rho}\|y-x\|^2.$$

The following theorem provides the conditions under which the system is small–time controllable

Theorem 2.4 (Controllability) *Let S be compact and let the following be valid*

IS.1 *let S be satisfying a ρ–internal sphere condition,*

IS.2 *there exist $\delta > 0$, $\mu > 0$, such that for every $\xi \in S_\delta \setminus S$, there exists $\zeta_\xi \in \partial^P d_S(\xi)$ with the property:*

$$\langle \zeta_\xi, [g_1, g_2](\xi) \rangle \leq -\mu < 0.$$

Then the minimum time function to reach S from ξ subject to the dynamics (2.1), $T_S(\xi)$ is (finite and) Hölder continuous with exponent $\frac{1}{2}$ on S_δ.

See e.g. [5], [6], [7].

3 A classical approach

We will formulate the Dynamic Programming Principle (DPP) for the minimum time problem.

Theorem 3.1 (see [1]) (DPP) *Under the standard assumptions, for $x \in \mathcal{R}^S$,*

$$\inf_{u \in \mathcal{U}_{ad}} \{t + T_S(y(t, x, u)\} = T_S(x), \text{ for } t \in [0, T_S(x)]. \tag{3.1}$$

A boundary value problem for the Hamilton–Jacobi–Bellman (HJB) equation associated with the minimum time function T_S is as follows.

Theorem 3.2 (see [1]) *Under the standard assumptions and STCS. Then T_S is the unique viscosity solution of*

$$\begin{cases} \sup_{u \in U}\{-f(x,u)DT_S\} - 1 = 0 \text{ in } \mathcal{R}^S \setminus S, \\ T_S(x) = 0 \text{ on } \partial S, \\ T_S(x) \to +\infty \text{ as } x \to x_0 \in \partial \mathcal{R}^S. \end{cases}$$

Then, defining the Kružkov transform

$$v_S(x) = \begin{cases} 1 - e^{-T_S(x)} & x \in \mathcal{R}^S, \\ 1 & x \notin \mathcal{R}^S, \end{cases}$$

$v_S(x)$ satisfies the dynamic programming principle

$$v_S(x) = \inf_{u \in \mathcal{U}_{ad}} \{\int_0^t e^{-s} ds + e^{-t} v_S(y(t, x, u))\} \tag{3.2}$$

for all $t \in [0, T_S(x)]$ and is the unique bounded viscosity solution of

$$\begin{cases} v_S(x) + \sup_{u\in U}\{-f(x,u)\nabla v_S(x)\} - 1 = 0 & \text{in } \mathbb{R}^n \setminus S \\ v_S(x) = 0 & \text{on } S. \end{cases} \tag{3.3}$$

Recover $T_S(x) = -\log(1 - v_S(x))$ and $\mathcal{R}^S = \{x \in \mathbb{R}^n \colon v_S(x) < 1\}$.
Given a fixed step $h > 0$ small enough, we approximate

$$\begin{cases} \dot{y}(t) = f(y(t), u(t)), t \in [0, +\infty) \text{ a.e.} \\ y(0) = \xi \in \mathbb{R}^n, \end{cases}$$

by a one step $(q+1)$-th order scheme which has the form

$$\begin{cases} y_{n+1} &= y_n + h\Phi(y_n, A_n, h) \\ y_0 &= \xi \end{cases} \tag{3.4}$$

where A_n is an $m \times l$ matrix, $A_n = (u_n^1, ..., u_n^l)$ with $u_n^i \in U$. Here $l > 0$ depends on the specific method. We make the following assumptions on the scheme to preserve the order of the method:

(A.1) For any $x \in \mathbb{R}^n$ and any measurable $u \colon [0, h) \to U$ there exists an $m \times l$ (where l depends on the chosen method) matrix $A \in U^l$ such that

$$\|y(h, x, u) - y_h(h, x, A)\| \le Ch^{q+2}, \tag{3.5}$$

where C is a constant, $q \ge k$, and $y(h, x, u)$ stands for the exact solution of (1.1) following the control u and $y_h(h, x, A) = x + h\Phi(x, A, h)$.

Conversely,

(A.2) for any matrix $A \in U^l$, there exists a measurable control $u \colon [0, h) \to U$ such that (3.5) holds.

For instance, if $q = 0$ or $q = 1$, we can simply take (3.4) as Euler or Heun 's method, respectively, i.e.

$$\Phi(x, u, h) = f(x, u),$$
$$\Phi(x, u_1, u_2, h) = \frac{f(x, u_1) + f(x + hf(x, u_1), u_2)}{2},$$

We define the function

$$n_h(\{A_i\}, \xi) = \min\{n \in \mathbb{N} : y_n \in S\} \le +\infty,$$

$$N_h(\xi) = \min_{\{A_i\}\in U^l}\{n_h(\{A_i\}, \xi)\}.$$

The discrete minimum time function is now defined by setting

$$T_h(\xi) = hN_h(\xi).$$

For a given stepsize $h > 0$, define

$$v_h(x) = 1 - e^{-T_h(x)},$$

Recall that v_h is the unique bounded solution of the following problem, see [2]:

$$\begin{cases} v_h(x) = \inf_{A\in U^l}\{e^{-h}v_h(x + h\Phi(x, A, h))\} + 1 - e^{-h} & \text{on } \mathbb{R}^n \setminus S \\ v_h(x) = 0 & \text{on } S. \end{cases} \tag{3.6}$$

For the convergence of v_h as well as T_h, see [2].
For the fully discrete scheme, the idea is to use the first order interpolation, which described briefly as follows. Let $\Gamma = \{x_{i,j} \colon i, j = 1, ..., I\}$ be a space grid for the domain $\Omega \subset S_\delta$. Now we construct a fully discrete version of (3.6) by substituting $v_h(x_i + h\Phi(x_i, A, h))$ with

$$I[v_h](x_i + h\Phi(x_i, A, h)) = \sum_j^I \lambda_j(A)v_h(x_j),$$

if $x_i + h\Phi(x_i, A, h) = \sum_j^I \lambda_j(A)x_j, \quad \lambda_j(A) \in [0, 1], \quad \sum_{j=1}^I \lambda_j(A) = 1$. More precisely,

$$\Gamma^\star := \{x \in \Gamma : \exists A \text{ such that } x + h\Phi(x, A, h) \in \Omega\}$$

and the fully discrete problem reads as

$$\begin{cases} v_h^{\Delta x}(x) = \min_{A\in U^l}\{e^{-h}I[v_h^{\Delta x}](x + h\Phi(x, A, h))\} + 1 - e^{-h} & \text{if } x \in \Gamma^\star \setminus S, \\ v_h^{\Delta x}(x) = 0 & \text{if } x \in \Gamma^\star \cap S, \\ v_h^{\Delta x}(x) = 1 & \text{if } x \in \Gamma \setminus \Gamma^\star \\ v_h^{\Delta x}(x) = I[v_h^{\Delta x}](x) & \text{if } x \in \Omega \setminus \Gamma. \end{cases} \tag{3.7}$$

Theorem 3.3 (see [4]) *Assume that*

$$T_S(x) \le C\sqrt[k]{d_S(x)},\ T_h(x) \le C\sqrt[k]{d_S(x)}$$

hold in a neighborhood S_δ of the target S, together with standard assumptions on the scheme. Then there exist $\bar{h}$ and C, C_1, $C_2 > 0$ such that

$$\|v_S - v_h\|_{\infty,\Omega} \le Ch^{\frac{q+1}{k}},$$
$$\left\|v_S - v_h^{\Delta x}\right\|_{\infty,\Omega} \le C_1 h^{\frac{q+1}{k}-1} + C_2\frac{(\Delta x)^{1/k}}{h}.$$

for $0 < h \le \bar{h}$.

3.1 Example

Consider the rocket car example $\ddot{x} = u$, $|u| \leq 1$. Here we used the third order Runge–Kutta method.

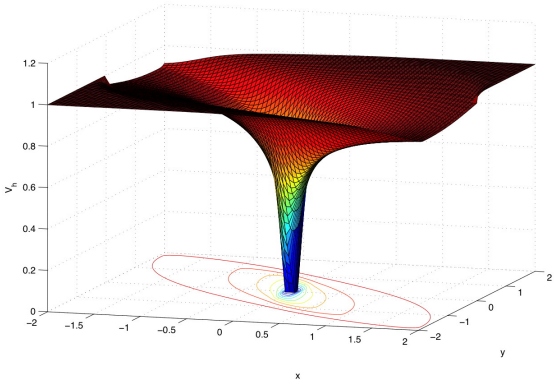


Figure 1. The value function.

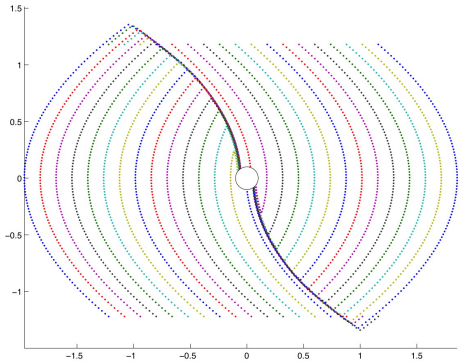


Figure 2. Computed discrete trajectories following discrete feedback controls

4 A new approach

Observe that the semi-discretization in time is a piecewise constant function with jumps of size $\approx h$ since

$$T_S(x) = \begin{cases} h & \text{for all } x \in \mathcal{R}^S(h) \setminus S \\ 0 & \text{on } \overset{o}{S} \end{cases}$$

Our goal is to reformulate the problem so that the jump size is reduced, at least, for some classes of the control systems.

To this aim, instead of letting $T_S(x)$ be zero in $\overset{o}{S}$, we do as in what follows. Consider the reverse dynamics of (1.1)

$$\begin{cases} \dot{y}^-(t) & = -f(y^-(t), u(t)) \\ y(0) & = \eta \in \overset{o}{S} \end{cases}. \tag{4.1}$$

We define the minimum time to $\bar{S^c}$ by following some $u \in \mathcal{U}_{ad}$ from $\eta \in \overset{o}{S}$

$$t_{S^c}(\eta, u) = \min\{t \geq 0 : y^-(t, \eta, u) \in \bar{S^c}\} \leq +\infty.$$

Then the minimum time function to $\bar{S^c}$ from $\eta \in \overset{o}{S}$ is defined as

$$T_{S^c}(\eta) = \inf_{u \in \mathcal{U}_{ad}} \{t_{S^c}(\eta, u)\}.$$

We also define

$$\mathcal{R}^{S^c}(t) = \{\eta \in \mathbb{R}^n : T_{S^c}(\eta) < t\},$$
$$\mathcal{R}^{S^c} = \{\eta \in \mathbb{R}^n : T_{S^c}(\eta) < +\infty\},$$

the reachable sets w.r.t. $\bar{S}^c$.

Besides the standard assumptions imposed before, the followings are assumed to be fulfilled from now on.

Assumptions 4.1

a) S is a compact set with C^2 boundary,

b) (f, U), $(-f, U)$ are *small time controllable* on S, $\bar{S}^c$ respectively.
Moreover, assume $T_S(\cdot)$, $T_{S^c}(\cdot)$ are locally Hölder continuous with exponent $\frac{1}{k}$ in $\mathcal{R}^S$, $\mathcal{R}^{S^c}$, $k \in \mathbb{N} \setminus \{0\}$

For $x \in \mathcal{R}^{S^c}$, consider the reverse dynamics (4.1) and the new target set as the closure of the complement of the original target set S, $\bar{S}^c$. Due to the same arguments as above, $T_{S^c}(x)$ is the unique viscosity solution of

$$\begin{cases} \sup_{u\in U}\{f(x,u)\nabla T_{S^c}(x)\} - 1 = 0 & \text{in } \mathcal{R}^{S^c} \setminus \bar{S}^c \\ T_{S^c}(x) = 0 & \text{on } \partial\bar{S}^c \\ T_{S^c}(x) = +\infty & \text{as } x \to x_0 \in \partial\mathcal{R}^{S^c}. \end{cases} \tag{4.2}$$

Now we are going to redefine the minimum time function as

$$T(x) = \begin{cases} T_S(x) & \text{if } x \in \mathcal{R}^S \\ 0 & \text{if } x \in \partial S \\ -T_{S^c}(x) & \text{if } x \in \mathcal{R}^{S^c} \\ +\infty & \text{if } x \to x_0 \in \partial\mathcal{R}^S \\ -\infty & \text{if } x \to x_0 \in \partial\mathcal{R}^{S^c} \end{cases} \tag{4.3}$$

and the value function as

$$v(x) = \begin{cases} 1 - e^{-T(x)} & \text{if } x \in \mathbb{R}^n \setminus S \\ 0 & \text{if } x \in \partial S \\ e^{T(x)} - 1 & \text{if } x \in \mathbb{R}^n \setminus \bar{S}^c. \end{cases} \tag{4.4}$$

Then the minimum time problem is reformulated as $T(x)$ is the unique viscosity solution of

$$\begin{cases} \sup_{u\in U}\{-f(x,u)\nabla T(x)\} - 1 = 0 & \text{in } \mathcal{R}^S \setminus S \text{ or } \mathcal{R}^{S^c} \setminus \bar{S}^c \\ T(x) = 0 & \text{on } \partial S \\ T(x) = +\infty & \text{as } x \to x_0 \in \partial\mathcal{R}^S \\ T(x) = -\infty & \text{as } x \to x_0 \in \partial\mathcal{R}^{S^c}. \end{cases} \tag{4.5}$$

thus $v(x)$ is the unique bounded viscosity solution of

$$\begin{cases} v(x) + \sup_{u\in U}\{-f(x,u)\nabla v(x)\} - 1 = 0 & \text{in } S^c \\ -v(x) + \sup_{u\in U}\{-f(x,u)\nabla v(x)\} - 1 = 0 & \text{in } \overset{o}{S} \\ v(x) = 0 & \text{on } \partial S, \end{cases} \tag{4.6}$$

It is easy to see that T and v satisfy the dynamic programming principles (3.1) and (3.2) whenever the optimal trajectories y on the right hand side of these principles stay in S^c. Likewise, it is straightforward to see that (3.1) and (3.2) with y^- in place of y hold for $-T$ and $-v$, respectively, whenever y^- stays in S. However, it remains to be clarified how these principles change for trajectories crossing ∂S.

Proposition 4.2 (Bridge DPP for T, see [8]) *Under Assumptions 1.1 and 4.1, there exits $\tau > 0$ such that*

$$T(x) = \inf_{\alpha \in \mathcal{U}_{ad}} \{t + T(y(t,x,\alpha))\} \quad \textit{for } x \in \mathcal{R}^S(\tau),\ 0 < T_S(x) < t < \tau,$$

$$T(x) = \sup_{\alpha \in \mathcal{U}_{ad}} \{-t + T(y^-(t,x,\alpha))\} \quad \textit{for } x \in \mathcal{R}^{S^c}(\tau),\ 0 < T_{S^c}(x) < t < \tau.$$

Proposition 4.3 (BDPP for v, see [8])
Under Assumptions 1.1 and 4.1, there exists $\tau > 0$ such that

$$v(x) = \inf_{\alpha \in \mathcal{U}_{ad}} \{\int_0^t e^{-s} ds + e^{-T(x)} v(y(t,x,\alpha))\} \qquad \textit{for } x \in \mathcal{R}^S(\tau),\ 0 < T_S(x) < t < \tau,$$

$$v(x) = \sup_{\alpha \in \mathcal{U}_{ad}} \{-\int_0^t e^{-s} ds + e^{-T(x)} v(y^-(t,x,\alpha))\} \qquad \textit{for } x \in \mathcal{R}^{S^c}(\tau),\ 0 < T_{S^c}(x) < t < \tau.$$

We discretize the problem defined in the interior of S in the same way as it is of the one defined exterior of S.

Consider the following problem as the discrete version of (4.6)

$$\text{(4.7)} \qquad \begin{cases} v_h(x) = \inf_{A \in U^l} \{e^{-h} v_h(x + h\Phi(x,A,h))\} + 1 - e^{-h} & \textit{for } x \in S^c \\ v_h(x) = \sup_{A \in U^l} \{e^{-h} v_h(x + h\Phi^-(x,A,h))\} + e^{-h} - 1 & \textit{for } x \in \overset{o}{S} \\ v_h(x) = 0 & \textit{for } x \in \partial S. \end{cases}$$

Theorem 4.4 (Uniqueness, see [8]) *The problem (4.7) admits a unique bounded solution $v \in L^\infty(\mathbb{R}^n)$. Moreover, $\|v\|_\infty \leq 1$.*

The fully discrete problem of (4.6) reads as

$$\text{(4.8)} \qquad \begin{cases} v_h^{\Delta x}(x) = \inf_{A \in U^l} \{e^{-h} I[v_h^{\Delta x}](x + h\Phi(x,A,h))\} + 1 - e^{-h} & \textit{for } x \in \Gamma \cap S^c \\ v_h^{\Delta x}(x) = \sup_{A \in U^l} \{e^{-h} I[v_h^{\Delta x}](x + h\Phi^-(x,A,h))\} + e^{-h} - 1 & \textit{for } x \in \Gamma \cap \overset{o}{S} \\ v_h^{\Delta x}(x) = 0 & \textit{for } x \in \Gamma \cap \partial S. \end{cases}$$

Theorem 4.5 (Error estimate, see [8]) *Under standard assumptions,*

$$\|v_h - v\|_{\infty,\Omega} \leq C_1 h^{\frac{q+1}{k} - 1} + C_2 h,$$

$$\left\|v_h^{\Delta x} - v\right\|_{\infty,\Omega} \leq C_1 \frac{\Delta x^{\frac{1}{k}}}{h} + C_2 h^{\frac{q+1}{k}-1} + C_3 h,$$

where v, v_h, $v_h^{\Delta x}$ *are solutions of (4.6), (4.7), (4.8) respectively and* C_1, C_2, C_3 *are positive constants.*

5 Comparison via examples

5.1 Example 1

The first test we perform uses the simple one dimensional dynamics

$$\dot{x} = u, \ u \in [-1, 1]. \tag{5.1}$$

on $\Omega = [-1, 1]$ with target $S = [-0.25, 0.25]$.

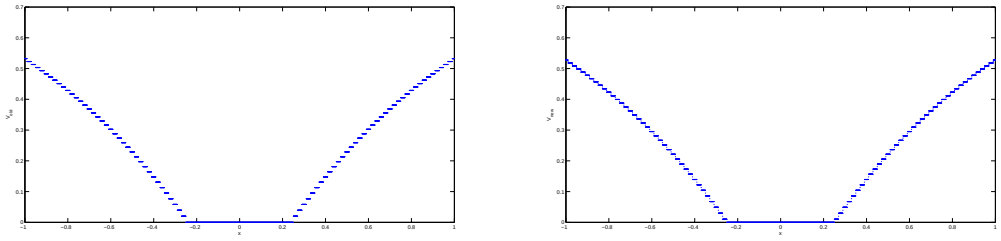


Figure 3. Positive part of the value function obtained by the classical approach (left) and by the proposed new one (right).

5.2 Example 2

The dynamics of the second example is

$$\dot{x}_1 = -x_2 + x_1\, u,\ \dot{x}_2 = x_1 + x_2\, u, \ u \in [-1, 1]. \tag{5.2}$$

It is easy to check that the Petrov condition holds for S, $\bar{S}^c$, thus T_S, T_{S^c} are Lipschitz continuous.

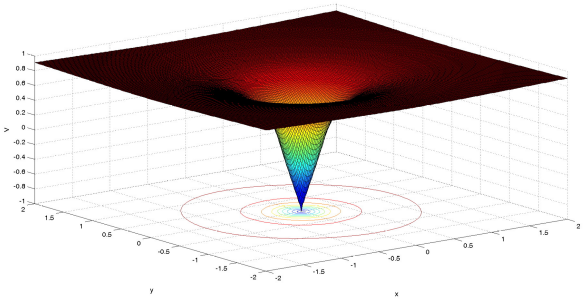


Figure 4. Value function on Ω obtained by the new approach (radius of the target $r = 0.25$, $h = 0.01$, $\Delta x = 0.01$, 3rd order Runge-Kutta scheme).

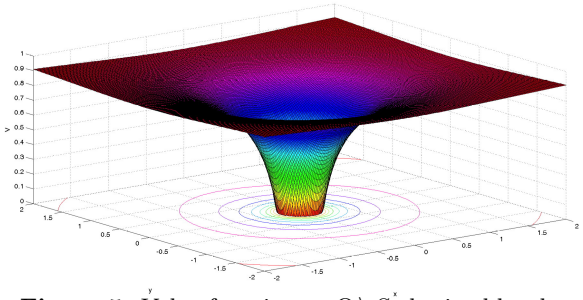


Figure 5. Value function on $\Omega \setminus S$ obtained by the new approach (radius of the target $r = 0.25$, $h = 0.01$, $\Delta x = 0.01$, 3rd order Runge-Kutta scheme).

Table 2: Comparison of error estimates for Example (5.2) $(r = 0.5)$

Δx	h	**New approach**	**Classical approach**
0.02	0.01	0.0089	0.0357
0.02	0.025	0.0059	0.0364
0.016	0.01	0.0077	0.0275
0.016	0.025	0.0068	0.0290

References

[1] M. Bardi and M. Capuzzo-Dolcetta, "Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations". Birkhäuser, Boston (1997).

[2] M. Bardi and M. Falcone, *An approximation scheme for the minimum time function.* SIAM Journal on Control and Optimization 28/4 (1990), 950–965.

[3] P. Cannarsa and C. Sinestrari, "Semiconcave Functions, Hamilton–Jacobi Equations, and Optimal Control". Birkhäuser, 2004.

[4] G. Colombo and Thuy T. T. Le, *Higher order discrete controllability and the approximation of the minimum time function.* Discrete and Continuous Dynamical Systems – Series A, 35/9 (2015), 4293–4322.

[5] A. Marigonda, *Second order conditions for the controllability of nonlinear systems with drift.* Commun. Pur. Appl. Anal. 5 (2006), no. 4, 861–885.

[6] A. Marigonda and S. Rigo, *Controllability of some nonlinear systems with drift via generalized curvature properties.* SIAM J. Control Optim., 53/1 (2015), 434–474.

[7] Thuy T. T. Le and A. Marigonda, *Small-time local attainability for a class of control systems with state constraints.* Preprint.

[8] L. Grüne and Thuy T. T. Le, *A new approach to the minimum time problem and its numerical approximation.* Preprint.

An introduction to derived categories

Francesco Mattiello (*)

Abstract. Derived categories were introduced in the sixties by Grothendieck and Verdier and have proved to be of fundamental importance in Mathematics. Starting with a short review of the basic language of category theory, we will first introduce the notion of abelian category with the help of several examples. Then we will spend some time giving a thorough motivation for the construction of the derived category of an abelian category.

Contents

Introduction

Derived categories were introduced in the sixties by Grothendieck and Verdier in the study of derived functors and spectral sequences.

Given an abelian category $\mathcal{C}$, its derived category $D(\mathcal{C})$ is obtained from the category $Ch(\mathcal{C})$ of (cochain) complexes in two stages. First one constructs a quotient $K(\mathcal{C})$ of $Ch(\mathcal{C})$ by identifying chain homotopy equivalent morphisms between complexes. Such a quotient

(*)Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: mattiell@math.unipd.it . Seminar held on June 10th, 2015.

category $K(\mathcal{C})$ is called the *homotopy category* of $\mathcal{C}$. It is an additive category in which the homotopy equivalences are invertible, but this advantage comes at a cost: $K(\mathcal{C})$ is not abelian anymore. Consequently, one has to look for an effective substitute for short exact sequences, that should still have the property that it induces long exact sequences on cohomology. The concept of a triangulated category with its "distinguished triangles" provides a solution. It turns out that the homotopy category $K(\mathcal{C})$ can be endowed with the structure of a triangulated category. The second step consists in "localizing" $K(\mathcal{C})$ by inverting quasi-isomorphisms using a calculus of fractions. The goal is the following: we want morphisms in $K(\mathcal{C})$ which induce isomorphisms on cohomology to be invertible in the category to be constructed. If we want a quasi-isomorphism to become an isomorphism, it has to have an inverse. Unfortunately, there is no candidate for an inverse around. So to define a derived category, we have to use a more elaborate process called localization of categories.
This short note is organized as follows. In Section 1 we recall the basic language of category theory; in Section 2, we introduce predictive, additive and abelian categories; finally, in Section 3 we give a sketch of the construction of the derived category of an abelian category.
The reader is referred to [1], [2], [3], [4], [5], [6], [7], [8].

1 The language of category theory

1.1 Categories

We recall that a *category* $\mathcal{C}$ consists of a class $Ob(\mathcal{C})$ of *objects*, a set of *morphisms* $\mathrm{Hom}_{\mathcal{C}}(A, B)$ (often denoted by $\mathrm{Hom}(A, B)$) for every ordered pair (A, B) of objects, and a *composition* $\mathrm{Hom}(A, B) \times \mathrm{Hom}(B, C) \to \mathrm{Hom}(A, C)$, denoted by $(f, g) \mapsto g\,f$, for every ordered triple A, B, C of objects. We often write $f\colon A \to B$ or $A \xrightarrow{f} B$ instead of $f \in \mathrm{Hom}(A, B)$. The following axioms must be satisfied:

(a) the Hom sets are pairwise disjoint, that is, each $f \in \mathrm{Hom}(A, B)$ has a unique *domain* A and a unique *target* B;

(b) for each object A, there is an *identity morphism* $1_A \in \mathrm{Hom}(A, A)$ such that $f\,1_A = f$ and $1_B\,f = f$ for all $f\colon A \to B$;

(c) composition is associative: given morphisms $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$, then $h\,(g\,f) = (h\,g)\,f$.

There is a bijection between objects A and their identity morphisms 1_A. Thus, we may regard a category as consisting only of morphisms. In almost all uses of categories, however, it is more natural to think of two sorts of entries: objects and morphisms.

If $\mathcal{C}$ is a category, its *opposite category* $\mathcal{C}^{op}$ is the category with $Ob(\mathcal{C}^{op}) = Ob(\mathcal{C})$, with morphisms $\mathrm{Hom}_{\mathcal{C}^{op}}(A, B) = \mathrm{Hom}_{\mathcal{C}}(B, A)$ (we may write morphisms in $\mathcal{C}^{op}$ as f^{op}, where f is a morphism in $\mathcal{C}$), and with composition the reverse of that in $\mathcal{C}$; that is, $g^{op}\,f^{op} = (f\,g)^{\mathrm{op}}$. We illustrate the composition in $\mathcal{C}^{op}$: a diagram $C \xrightarrow{f^{\mathrm{op}}} B \xrightarrow{g^{\mathrm{op}}} A$ in $\mathcal{C}^{op}$

corresponds to the diagram $A \xrightarrow{g} B \xrightarrow{f} C$ in $\mathcal{C}$. Any concept or statement about an arbitrary category admits a dual concept or statement (the process of reversing arrows).

A category $\mathcal{S}$ is a *subcategory* of a category $\mathcal{C}$ if: $Ob(\mathcal{S}) \subseteq Ob(\mathcal{C})$; $\mathrm{Hom}_{\mathcal{S}}(A, B) \subseteq \mathrm{Hom}_{\mathcal{C}}(A, B)$ for all $A, B \in Ob(\mathcal{S})$; if $f \in \mathrm{Hom}_{\mathcal{S}}(A, B)$ and $g \in \mathrm{Hom}_{\mathcal{S}}(B, C)$, then the composite $g\,f \in \mathrm{Hom}_{\mathcal{S}}(A, C)$ is equal to the composite $g\,f \in \mathrm{Hom}_{\mathcal{C}}(A, C)$; if $A \in Ob(\mathcal{S})$, then the identity $1_A \in \mathrm{Hom}_{\mathcal{S}}(A, A)$ is equal to the identity $1_A \in \mathrm{Hom}_{\mathcal{C}}(A, A)$.
A subcategory $\mathcal{S}$ of $\mathcal{C}$ is a *full subcategory* if, for all $A, B \in Ob(\mathcal{S})$, we have $\mathrm{Hom}_{\mathcal{S}}(A, B) = \mathrm{Hom}_{\mathcal{C}}(A, B)$.

A morphism $u\colon B \to C$ in a category $\mathcal{C}$ is a *monomorphism* (or is *monic*) if u can be canceled from the left; that is, for all objects A and all morphisms $f, g\colon A \to B$, we have that $u\,f = u\,g$ implies $f = g$.

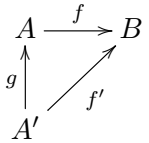
$$A \underset{g}{\overset{f}{\rightrightarrows}} B \xrightarrow{u} C.$$

It is clear that $u\colon B \to C$ is monic if and only if , for all A, the map $u_*\colon \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, C)$, $u_*(f) = u\,f$, is an injection.
A morphism $v\colon B \to C$ in a category $\mathcal{C}$ is an *epimorphism* (or is *epic*) if v can be canceled from the right; that is for all objects D and all morphisms $h, k\colon C \to D$, we have that $h\,v = k\,v$ implies $h = k$.

$$B \xrightarrow{v} C \underset{k}{\overset{h}{\rightrightarrows}} D.$$

It is clear that $v\colon B \to C$ is epic if and only if , for all D, the map $v^*\colon \mathrm{Hom}(C, D) \to \mathrm{Hom}(B, D)$, $v^*(f) = f\,v$, is an injection.
A morphism in a category $\mathcal{C}$ is a *bijection* if it is monic and epic.
We call a *left inverse* (respectively, a *right inverse*) of $u \in \mathrm{Hom}(A, B)$ a morphism $v \in \mathrm{Hom}(B, A)$ such that $v\,u = 1_A$ (respectively $u\,v = 1_B$); v is called the *inverse* of u if it is both a left inverse and a right inverse of u (in which case it is uniquely determined). The morphism u is called an *isomorphism* if it has an inverse. If u has a left inverse (respectively, a right inverse) it is monic (respectively epic). Thus an isomorphism is bijective (the converse being, in general, false).
The composite of two monomorphisms (respectively, epimorphisms) is a monomorphism (respectively, epimorphisms), hence the composite of two bijections is a bijection; similarly the composite of two isomorphisms is an isomorphism. If the composite $v\,u$ of two morphisms u, v is a monomorphism (respectively, an epimorphism), then u (respectively, v) is as well.

If B is an object in a category $\mathcal{C}$, consider all ordered pairs (A, f), where $f\colon A \to B$ is a monomorphism. Call two such pairs (A, f) and (A', f') *equivalent* if there exists an isomorphism $g\colon A' \to A$ with $f' = f\,g$.



A *subgadget of* B is an equivalence class $[(A, f)]$, and we call A a *subobject of* B. Note that if (A', f') and (A, f) are equivalent, then $A' \cong A$. There is also the dual notion of *quotient*. If B is an object in a category $\mathcal{C}$, consider all ordered pairs (f, C), where $f\colon B \to C$ is a epimorphism. Call two such pairs (f, C) and (f', C') *equivalent* if there exists an isomorphism $g\colon C \to C'$ with $f' = g\,f$. A *quotient of* B is an equivalence class $[(f, C)]$, and we call C a *quotient object of* B. Note that if (C', f') and (C, f) are equivalent, then $C' \cong C$.

Examples 1.1

(a) **Sets**. Objects are sets, morphisms are functions, and composition is the usual composition of functions.

(b) **Groups**. Objects are groups, morphisms are group homomorphisms, and composition is the usual composition (homomorphisms are functions).

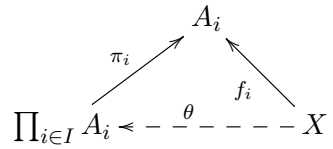
(c) R-Mod. If R is any ring (associative, with identity), the category of left R-modules has as its objects all left R-modules, as morphisms all R-linear homomorphisms, and as composition the usual composition of functions.

(d) **Top**. Objects are topological spaces, morphisms are continuous functions, and composition is the usual composition of functions.

(e) **Htp**. Objects are topological spaces, morphisms are homotopy classes of continuous functions (thus, a morphism here is not a function but a certain equivalence class of functions). Composition is defined by $[f][g] = [fg]$.

1.2 Products and coproducts

Let $\mathcal{C}$ be a category, and let $(A_i)_{i\in I}$ be a family of objects in $\mathcal{C}$ indexed by a set I. A *product* is an ordered pair $(\prod_{i\in I} A_i, (\pi_i)_{i\in I})$, consisting of an object $\prod_{i\in I} A_i$ and a family $\pi_i\colon \prod_{i\in I} A_i \to A_i$ of *projections*, satisfying the following universal property: for every object X equipped with morphisms $f_i\colon X \to A_i$, there exists a unique morphism $\theta\colon X \to \prod_{i\in I} A_i$ making the following diagram commute for each i.

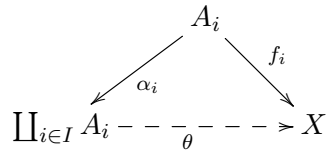


Note that the definition of $\prod_{i\in I} A_i$ can be summarized by the formula

$$\mathrm{Hom}(X, \prod_{i\in I} A_i) \cong \prod_{i\in I} \mathrm{Hom}(X, A_i), \tag{1}$$

where the second product is taken in the category of abelian groups. If it exists, a product is unique up to isomorphism, as a solution of a universal problem (cfr.).

Here is the dual notion.
Let $\mathcal{C}$ be a category, and let $(A_i)_{i\in I}$ be a family of objects in $\mathcal{C}$ indexed by a set I. A *coproduct* is an ordered pair $(\coprod_{i\in I} A_i, (\alpha_i)_{i\in I})$, consisting of an object $\coprod_{i\in I} A_i$ and a family $\alpha_i\colon A_i \to \coprod_{i\in I} A_i$ of *injections*, satisfying the following universal property: for every object X equipped with morphisms $f_i\colon A_i \to X$, there exists a unique morphism $\theta\colon \coprod_{i\in I} A_i \to X$ making the following diagram commute for each i.



Note that the definition of $\coprod_{i\in I} A_i$ can be summarized by the formula

$$\mathrm{Hom}(\coprod_{i\in I} A_i,\, X) \cong \prod_{i\in I} \mathrm{Hom}(A_i,\, X), \tag{2}$$

where the second product is taken in the category of abelian groups. If it exists, a coproduct is unique up to isomorphism, as a solution of a universal problem

1.3 Functors

If $\mathcal{C}$ and $\mathcal{D}$ are categories, then a (*covariant*) *functor* $T\colon \mathcal{C} \to \mathcal{D}$ is a function such that

(a) if $A \in Ob(\mathcal{C})$, then $T(A) \in Ob(\mathcal{D})$;

(b) if $f\colon A \to A'$ in $\mathcal{C}$, then $T(f)\colon T(A) \to T(A')$ in $\mathcal{D}$;

(c) if $A \xrightarrow{f} A' \xrightarrow{g} A''$ in $\mathcal{C}$, then $T(A) \xrightarrow{T(f)} T(A') \xrightarrow{T(g)} T(A'')$ in $\mathcal{D}$ and

$$T(g\,f) = T(g)\,T(f);$$

(d) $T(1_A) = 1_{T(A)}$ for every $A \in Ob(\mathcal{C})$.

We now recall an important example of functor. If $\mathcal{C}$ is a category and $A \in Ob(\mathcal{C})$, then the Hom *functor* $T_A\colon \mathcal{C} \to \mathbf{Sets}$, usually denoted by $\mathrm{Hom}(A, -)$, is defined by $T_A(B) = \mathrm{Hom}(A, B)$ for all $B \in Ob(\mathcal{C})$, and if $f\colon B \to B'$ in $\mathcal{C}$, then $T_A(f)\colon \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, B')$ is given by $T_A(f)\colon h \mapsto f\,h$. We call $T_A(f) = \mathrm{Hom}(A, f)$ the *induced map*, and we denote it by f_*; thus, $f_*\colon h \mapsto f\,h$.

A *controvariant functor* $T\colon \mathcal{C} \to \mathcal{D}$, where $\mathcal{C}$ and $\mathcal{D}$ are categories, is a function such that

(a) if $C \in Ob(\mathcal{C})$, then $T(C) \in Ob(\mathcal{D})$;

(b) if $f\colon C \to C'$ in $\mathcal{C}$, then $T(f)\colon T(C') \to T(C)$ in $\mathcal{D}$;

(c) if $C \xrightarrow{f} C' \xrightarrow{g} C''$ in $\mathcal{C}$, then $T(C'') \xrightarrow{T(g)} T(C') \xrightarrow{T(f)} T(C)$ in $\mathcal{D}$ and

$$T(g\,f) = T(f)\,T(g);$$

(d) $T(1_A) = 1_{T(A)}$ for every $A \in Ob(\mathcal{C})$.

It is clear that a controvariant functor $T\colon \mathcal{C} \to \mathcal{D}$ is the same thing as a (covariant) functor $T\colon \mathcal{C}^{op} \to \mathcal{D}$.
If $\mathcal{C}$ is a category and $B \in Ob(\mathcal{C})$, then the *controvariant* Hom *functor* $T^B\colon \mathcal{C} \to \mathbf{Sets}$, usually denoted by $\mathrm{Hom}(-, B)$, is defined by $T^B(C) = \mathrm{Hom}(C, B)$ for all $C \in Ob(\mathcal{C})$, and if $f\colon C \to C'$ in $\mathcal{C}$, then $T^B(f)\colon \mathrm{Hom}(C', B) \to \mathrm{Hom}(C, B)$ is given by $T^B(f)\colon h \mapsto h\,f$. We call $T^B(f) = \mathrm{Hom}(f, B)$ the *induced map*, and we denote it by f^*; thus, $f^*\colon h \mapsto h\,f$.

We similarly define functors of several variables (or *multifunctors*), covariant in some variables and contravariant in others. In order to simplify, we will generally limit ourselves to functors of one variable. Functors are composed in the same way as functions are, this composition is associative and "identify functors" play the role of units.

A functor $T\colon \mathcal{C} \to \mathcal{D}$, where $\mathcal{C}$, $\mathcal{D}$ are categories, is *faithful* if for each A, $B \in Ob(\mathcal{C})$, the function $\mathrm{Hom}_{\mathcal{C}}(A, B) \to \mathrm{Hom}_{\mathcal{D}}(T(A), T(B))$, given by $f \mapsto T(f)$, is an injection; T is *full* if the function $\mathrm{Hom}_{\mathcal{C}}(A, B) \to \mathrm{Hom}_{\mathcal{D}}(T(A), T(B))$ is surjective.

Examples 1.2

(a) If $\mathcal{C}$ is a category, then the *identity functor* $1_{\mathcal{C}}\colon \mathcal{C} \to \mathcal{C}$ is defined by $1_{\mathcal{C}}(A) = A$ for all objects A and $1_{\mathcal{C}}(f) = f$ for all morphisms f.

(b) If $\mathcal{S}$ is a subcategory of a category $\mathcal{C}$, then there is a canonical *inclusion functor* $\iota\colon \mathcal{S} \to \mathcal{C}$.

(c) The *forgetful functor* $F\colon \mathbf{Groups} \to \mathbf{Sets}$ is defined as follows: $F(G)$ is the underlying set of a group G and $F(f)$ is a homomorphism f regarded as a mere function.

(d) If $\mathcal{U}$ is the topology of a topological space X then a *presheaf of abelian groups over* X is a functor $\mathcal{P}\colon \mathcal{U}^{op} \to \mathbf{Ab}$. A *sheaf* is a presheaf satisfying certain axioms.
Presheaves and sheaves of abelian groups over X form categories: $\mathbf{pSh}(X, \mathbf{Ab})$ and $\mathbf{Sh}(X, \mathbf{Ab})$.

1.4 Natural transformations

Let $S, T\colon \mathcal{C} \to \mathcal{D}$ be covariant functors. A *natural transformation* $\tau\colon S \to T$ is a one-parameter family of morphisms in $\mathcal{D}$, $\tau = (\tau_A\colon S(A) \to T(A))_{A \in Ob(\mathcal{C})}$, making the following diagram commute for all $f\colon A \to A'$ in $\mathcal{C}$:

$$\begin{array}{ccc} S(A) & \xrightarrow{\tau_A} & T(A) \\ \Big\downarrow{\scriptstyle S(f)} & & \Big\downarrow{\scriptstyle T(f)} \\ S(A') & \xrightarrow[\tau_{A'}]{} & T(A'). \end{array}$$

Natural transformations between controvariant functors are defined similarly (replace $\mathcal{C}$ by $\mathcal{C}^{op}$). A *natural isomorphism* is a natural transformation τ for which each τ_A is an isomorphism.

Natural transformations can be composed. If $\tau\colon S \to T$ and $\sigma\colon T \to U$ are natural transformations, where S, T, $U\colon \mathcal{C} \to D$ are functors, then define $\sigma\,\tau\colon S \to U$ by $(\sigma\,\tau)_A = \sigma_A\,\tau_A$, for all $A \in Ob(C)$. It is easy to check that $\sigma\,\tau$ is a natural transformation (Rotman, Ex. 1.15).

For any functor $S\colon \mathcal{C} \to \mathcal{D}$, define the *identity natural transformation* $\omega_S\colon S \to S$ by setting $(\omega_S)_A\colon S(A) \to S(A)$ to be the identity morphism $1_{S(A)}$. The reader may check that a natural transformation $\tau\colon S \to T$ is a natural isomorphism if and only if there exists a natural transformation $\sigma\colon T \to S$ with $\sigma\,\tau = \omega_S$ and $\tau\,\sigma = \omega_T$.

A functor $S\colon \mathcal{C} \to \mathcal{D}$ is an *equivalence* if there exists a functor $T\colon \mathcal{D} \to \mathcal{C}$ and natural isomorphisms $T\,S \to 1_{\mathcal{C}}$, $S\,T \to 1_{\mathcal{D}}$.
The following description of equivalence is often more closed at hand:

Proposition 1.3 *A functor $S\colon \mathcal{C} \to \mathcal{D}$ is an equivalence if and only if it is full and faithful, and every object of $\mathcal{D}$ is isomorphic to an object of the form $S(C)$.*

2 Abelian categories

2.1 Additive categories

An object A in a category $\mathcal{C}$ is called an *initial object* (respectively, a *terminal object*) if, for every object $X \in \mathcal{C}$, there exists a unique morphism $A \to X$ (respectively, $X \to A$). The reader may check that any two initial (respectively, terminal) objects in a category $\mathcal{C}$, should they exist, are isomorphic and that this isomorphism is unique. A *zero object* is an object that is both an initial object and a terminal object. If A is a zero object in a category $\mathcal{C}$, then $\mathrm{Hom}(A, A)$ is reduced to 0, and for any $B \in Ob(\mathcal{C})$, $\mathrm{Hom}(A, B)$ (or $\mathrm{Hom}(B, A)$) is reduced to 0. If A and A' are zero objects, there exists a unique isomorphism of A to A' (that is, the unique zero element of $\mathrm{Hom}(A, A')$). We will identify all zero objects of $\mathcal{C}$ to a single one, denoted 0 by abuse of notation.

A category $\mathcal{C}$ is *preadditive* if

(a) $\mathrm{Hom}(A, B)$ is an (additive) abelian group for every A, $B \in Ob(\mathcal{C})$;

(b) the distributive laws hold: given morphisms

$$X \xrightarrow{a} A \underset{g}{\overset{f}{\rightrightarrows}} B \xrightarrow{b} Y,$$

where X and $Y \in Ob(\mathcal{C})$, then

$$b\,(f+g) = b\,f + b\,g \quad \text{and} \quad (f+g)\,a = f\,a + g\,a.$$

A preadditive category $\mathcal{C}$ is *additive* if the following two additional conditions holds:

(c) $\mathcal{C}$ has a zero object;

(d) $\mathcal{C}$ has finite products and finite coproducts.

The dual category of an additive category is still additive.

If $\mathcal{C}$ and $\mathcal{D}$ are additive categories, a functor $T\colon \mathcal{C} \to \mathcal{D}$ is *additive* if, for all A, B and all f, $g \in \mathrm{Hom}(A, B)$, we have $T(f+g) = T(f) + T(g)$; that is, the function $\mathrm{Hom}_{\mathcal{C}}(A, B) \to \mathrm{Hom}_{\mathcal{D}}(T(A), T(B))$, given by $f \mapsto T(f)$, is a homomorphism of abelian groups.
Clearly, if T is an additive functor, then $T(0) = 0$, where 0 is either a zero object or a zero morphism.

In all additive categories, finite products and finite coproducts coincide:

Proposition 2.1 *Let $\mathcal{C}$ be an additive category, and let M, A, $B \in Ob(\mathcal{C})$. Then $M \cong A \sqcap B$ if and only if there are morphisms $i\colon A \to M$, $j\colon B \to M$, $p\colon M \to A$, and $q\colon M \to B$ such that*

$$p\,i = 1_A, \qquad q\,j = 1_B, \qquad p\,j = 0, \qquad q\,i = 0, \qquad \text{and} \qquad i\,p + j\,q = 1_M.$$

Moreover, $A \sqcap B$ is also a coproduct with injections i and j, and so

$$A \sqcap B \cong A \sqcup B.$$

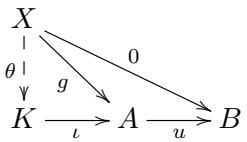
Proof. See Stenstrom, Prop. 3.2, Ch. IV.

If A and B are objects in an additive category, then $A \sqcap B \cong A \sqcup B$; their common value, denoted by $A \oplus B$, is called their *direct sum* (or *biproduct*).

The reader may check that, as a consequence of Proposition 2.1, if T is an additive functor, T transforms a finite direct sum of objects A_i into the direct sum of $T(A_i)$.

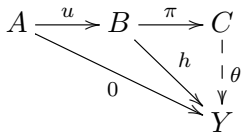
Let u, $v\colon B \to C$ be morphisms in an additive category $\mathcal{C}$. For all objects A, D in $\mathcal{C}$, $\mathrm{Hom}(A, B)$ and $\mathrm{Hom}(C, D)$ are abelian groups, and so u is monic if and only if $u\,g = 0$ implies $g = 0$, for any morphism $g\colon A \to B$ in $\mathcal{C}$; v is epic if and only if $h\,v = 0$ implies $h = 0$, for any morphism $h\colon C \to D$ in $\mathcal{C}$.

If $u\colon A \to B$ is a morphism in an additive category $\mathcal{C}$, then its *kernel* $\mathrm{Ker}\,u$ is a morphism $\iota\colon K \to A$ that satisfies the following universal mapping property: $u\,\iota = 0$ and, for every $g\colon X \to A$ with $u\,g = 0$, there exists a unique $\theta\colon X \to K$ with $i\,\theta = g$.



Any two kernels, should they exist, are isomorphic, as solutions of a universal problem.

There is a dual definition for *cokernel*, the morphism π in the following diagram:



Any two cokernels, should they exist, are isomorphic, as solutions of a universal problem.

Proposition 2.2 *Let $u\colon A \to B$ be a morphism in an additive category $\mathcal{C}$.*

(a) If $\operatorname{Ker} u$ *exists, then* u *is monic if and only if* $\operatorname{Ker} u = 0$.

(b) Dually, if $\operatorname{Coker} u$ *exists, then* u *is epic if and only if* $\operatorname{Coker} u = 0$.

Proof. We refer to the diagrams in the definitions of kernel and cokernel. Let $\operatorname{Ker} u$ be $\iota\colon K \to A$, and assume that $\iota = 0$. If $g\colon X \to A$ satisfies $u\,g = 0$, then the universal property of kernel provides a morphism $\theta\colon X \to K$ with $g = \iota\,\theta = 0$ (because $\iota = 0$). Hence, u is monic. Conversely, if u is monic, consider $K \underset{0}{\overset{\iota}{\rightrightarrows}} A \xrightarrow{u} B$. Since $u\,\iota = 0 = u\,0$, we have $\iota = 0$. The proof for epimorphisms and cokernels is dual. □

It is immediate to show that any kernel is monic and any cokernel is epic.

If ι is the kernel of some morphism and if $\operatorname{Coker}\iota$ exists, then $\iota = \operatorname{Ker}(\operatorname{Coker}\iota)$. Indeed, let $\iota\colon K \to A$ be the kernel of $u\colon A \to B$. Then $u\,\iota = 0$ implies that u factors over $\operatorname{Coker}\iota$. If now $\xi\colon X \to A$ is any morphism such that $\operatorname{Coker}\iota\,\xi = 0$, then it follows that $u\,\xi = 0$, and ξ factors uniquely over $\operatorname{Ker} u = \iota$. Therefore, ι is the kernel of $\operatorname{Coker}\iota$.

Let $f\colon A \to B$ be a morphism in an additive category $\mathcal{C}$, and assume that its cokernel $\operatorname{Coker} f\colon B \to C$ exists. Then its *image* is $\operatorname{Im} f = \operatorname{Ker}(\operatorname{Coker} f)$. In more suggestive notation, $\operatorname{Im}(A \xrightarrow{f} B) = \operatorname{Ker}(\operatorname{Coker} A \xrightarrow{f} B)$. We denote $\operatorname{Ker}(\operatorname{Coker} f)$ by $\operatorname{Im} f$. Dually, one defines the notion of *coimage*.

2.2 Abelian categories

An abelian category is an additive category $\mathcal{C}$ that satisfies the following two additional conditions (which are self-dual):

(e) Any morphism admits a kernel and a cokernel;

(f) Let u be a morphism in $\mathcal{C}$. Then the canonical morphism $\overline{u}\colon \operatorname{Coim} u \to \operatorname{Im} u$ is an isomorphism.

Let us show that given a morphism $u\colon A \to B$ in an additive category $\mathcal{C}$, there is a canonical morphism $\overline{u}\colon \operatorname{Coim} u \to \operatorname{Im} u$. Let $u\colon A \to B$, then there is a canonical factorization as indicated by the commutative diagram

$$\begin{array}{ccccccc} \operatorname{Ker} u & \longrightarrow & A & \xrightarrow{u} & B & \longrightarrow & \operatorname{Coker} u \\ & & \downarrow{\scriptstyle\lambda} & & \downarrow{\scriptstyle\mu} & & \\ & & \operatorname{Coim} u & \xrightarrow[\overline{u}]{} & \operatorname{Im} u & & \end{array}$$

where $\overline{u}$ is obtained as follows: $(\operatorname{Coker} u)\,u = 0$ implies that u factors as $u = \mu\,\alpha$ (by the universal property of the kernel), for some $\alpha : A \to \operatorname{Im} u$; then $\mu\,\alpha\,\operatorname{Ker} u = u\,\operatorname{Ker} u = 0$ implies that $\alpha\,\operatorname{Ker} u = 0$, since μ, being a kernel, is monic. Therefore, α factors as $\alpha = \overline{u}\,\lambda$ (by the universal property of the cokernel).

If $\mathcal{C}$ is an abelian category and $\mathcal{B}$ is a subcategory of $\mathcal{C}$ which is also abelian, then $\mathcal{B}$ is said to be an *abelian subcategory* if the inclusion functor $\mathcal{B} \to \mathcal{C}$ is exact.

If $\mathcal{C}$ is an abelian category, then the entire formalism of diagrams of homomorphisms between abelian groups can be carried over if we replace homomorphisms by morphisms in $\mathcal{C}$, insofar as we are looking at properties of finite type, i.e. not involving infinite products or coproducts. We content ourselves here with indicating a few particularly important facts. Note that the axiom (6) is equivalent to the following one:

(6') Every morphism u has a factorization as $u = \iota\pi$, where π is a cokernel and ι is a kernel.

Indeed, it is obvious that (6) implies (6′). Conversely, let $\mathcal{C}$ be a category satisfying (1)-(5) and (6'). If u factors as in (6'), then $\pi = \operatorname{Coim}\pi = \operatorname{Coker}(\operatorname{Ker}\pi)$ and $\iota = \operatorname{Im}\iota = \operatorname{Ker}(\operatorname{Coker}\iota)$. But $\operatorname{Ker}\pi = \operatorname{Ker}u$ since ι is a monomorphism, and similarly $\operatorname{Coker}\iota = \operatorname{Coker}u$. We conclude that $\operatorname{Coker}(\operatorname{Ker}u) = \operatorname{Ker}(\operatorname{Coker}u)$, that is $\operatorname{Coim}u = \operatorname{Im}u$.

For the rest of this section we assume that $\mathcal{C}$ is abelian. The previous discussion shows:

Proposition 2.3 *Let u be a morphism in $\mathcal{C}$.*

(a) *If u is both a monomorphism and an epimorphism, then u is an isomorphism;*

(b) *If u is a monomorphism, then $u = \operatorname{Ker}(\operatorname{Coker}u)$;*

(c) *If u is a epimorphism, then $u = \operatorname{Coker}(\operatorname{Ker}u)$.*

A sequence $\ldots \longrightarrow C_{i-1} \xrightarrow{\alpha_{i-1}} C_i \xrightarrow{\alpha_i} C_{i+1} \xrightarrow{\alpha_{i+1}} \ldots$ is *exact at* C_i if $\operatorname{Im}\alpha_{i-1} = \operatorname{Ker}\alpha_i$ (equals as subobjects of C_i). The whole sequence is *exact* if it is exact at each C_i.

Let $T\colon \mathcal{C} \to \mathcal{C}'$ be a (covariant) functor, where $\mathcal{C}'$ is an abelian category. We say that T is *left exact* (respectively, *right exact*) if T transforms any exact sequence $0 \longrightarrow A \longrightarrow B \longrightarrow C$ (respectively, $A \longrightarrow B \longrightarrow C \longrightarrow 0$) into an exact sequence $0 \longrightarrow T(A) \longrightarrow T(B) \longrightarrow T(C)$ (respectively, $T(C) \longrightarrow T(B) \longrightarrow T(A) \longrightarrow 0$). T is called an *exact functor* if it is both left exact and right exact; that is, T transforms any *short* exact sequence $0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$ into an exact sequence $0 \longrightarrow T(A) \longrightarrow T(B) \longrightarrow T(C) \longrightarrow 0$. If T is a controvariant functor, we say that T is *left exact* (respectively, *right exact*) if T has the corresponding property as a covariant functor $\mathcal{C}^{op} \to \mathcal{C}'$.
The composite of left exact, right exact, exact covariant functors is of the same type.
As a significant example, we note that $\operatorname{Hom}(-, -)$ is an additive bifunctor on $\mathcal{C}^{op} \times \mathcal{C}$, with values in **Ab**, controvariant in the first argument and covariant in the second argument, and left exact with respect to each argument.

Let $\mathcal{C} \in Ob(\mathcal{C})$, $(C_i)_{i\in I}$ a family of subobjects of $\mathcal{C}$, and assume that the coproduct $\coprod_{i\in I} C_i$ exists. The monomorphisms $C_i \to C$ induce a morphism $\alpha\colon \coprod_{i\in I} C_i \to C$. The image of α is called the *sum* of the subobjects C_i and it is denoted by $\sum_{i\in I} C_i$. If α is a monomorphism, then it induces an isomorphism $\coprod_{i\in I} C_i \cong \sum_{i\in I} C_i$, and in this case the sum is called *direct sum*. Dually, the epimorphisms $C \to C/C_i$ induce a morphism $\beta\colon C \to \prod_{i\in I} C/C_i$ (when the product exists). The kernel of β is the *intersection* of the subobjects C_i and it is denoted by $\bigcap_{i\in I} C_i$.

Examples 2.4

(a) R-Mod is abelian, for any ring R. In fact, it can be shown that any small abelian category is equivalent to a full subcategory of such a category of modules (*Mitchell's embedding theorem*). In particular:

(b) **Ab** is abelian.

(c) **Groups** is not abelian, as it is not additive.

(d) The full subcategory of **Ab** of all finitely generated abelian groups is an abelian category, as is the full subcategory of all torsion abelian groups.

(e) The full subcategory of **Ab** of all torsion-free abelian groups is not an abelian category, for there are morphisms having no cokernel; for example, the inclusion $2\mathbb{Z} \to \mathbb{Z}$ has cokernel $\mathbb{Z}/2\mathbb{Z}$, which is not torsion-free.

(f) $\mathbf{pSh}(X, \mathbf{Ab})$ and $\mathbf{Sh}(X, \mathbf{Ab})$ are abelian categories. More generally, if $\mathcal{C}$ is an abelian category, then $\mathbf{pSh}(X, \mathcal{C})$ and $\mathbf{Sh}(X, \mathcal{C})$ are abelian.

(g) If $\mathcal{C}$ is abelian, then its opposite category $\mathcal{C}^{op}$ is abelian. So, a theorem using only the axioms in its proof is true in every abelian category; moreover, its dual is also a theorem in every abelian category, and its proof is dual to the original proof.

3 Derived categories

3.1 The category of cochain complexes and the homotopy category

In this section we let $\mathcal{C}$ be an additive category. A *complex* $X^\bullet = (X^k, d_X^k)$ over $\mathcal{C}$ is a sequence of objects X^k and morphisms $d_X^k \colon X^k \to X^{k+1}$ $(k \in \mathbb{Z})$:

$$\ldots \longrightarrow X^{k-1} \xrightarrow{d_X^{k-1}} X^k \xrightarrow{d_X^k} X^{k+1} \longrightarrow \ldots$$

such that $d_X^{k+1} d_X^k = 0$ for all $k \in \mathbb{Z}$. A complex is *bounded below* if $X^k = 0$ for all but finitely many $k < 0$. It is *bounded above* if $X^k = 0$ for all but finitely many $k > 0$. It is *bounded* if it is bounded below and bounded above. We denote by $Ch(\mathcal{C})$ the additive category of complexes and by $Ch^*(\mathcal{C})$ $(* = b, +, -)$ the full additive subcategory of $Ch(\mathcal{C})$ consisting of bounded complexes (resp. bounded below, bounded above). We may consider $\mathcal{C}$ as a full subcategory of $Ch^b(\mathcal{C})$ by identifying each object X of $\mathcal{C}$ with the complex $\ldots \longrightarrow 0 \longrightarrow X \longrightarrow 0 \longrightarrow \ldots$ "concentrated in degree 0", where X stands in degree 0.

Let $X \in Ob(Ch(\mathcal{C}))$ and $p \in \mathbb{Z}$. The *shift functor* is defined by

$$(X[p])^n = X^{n+p}, \qquad d^n_{X[p]} = (-1)^p d_X^{n+p}$$

and if $f \colon X \to Y$ is a morphism in $Ch(\mathcal{C})$, then

$$(f[p])^n = f^{n+p}.$$

The shift functor $[1]\colon Ch(\mathcal{C}) \to Ch(\mathcal{C})$, $X \mapsto X[1]$ is an automorphism (that is, an invertible functor) of $Ch(\mathcal{C})$.

The *mapping cone* of a morphism $f\colon X \to Y$ in $Ch(\mathcal{C})$ is the complex $(Mc(f),\, d_{Mc(f)})$, where $(Mc(f))^k = (X[1])^k \oplus Y^k$ and $d^k_{Mc(f)} = \begin{pmatrix} d^k_{X[1]} & 0 \\ f^{k+1} & d^k_Y \end{pmatrix}$. There are natural morphisms of complexes $\alpha(f)\colon Y \to Mc(f)$, $\beta(f)\colon Mc(f) \to X[1]$ and $\beta(f)\alpha(f) = 0$.

A morphism $f\colon X \to Y$ in $Ch(\mathcal{C})$ is said to be *homotopic to zero* if for all $p \in \mathbb{Z}$ there exists a morphism $s^p\colon X^p \to Y^p$ such that $f^p = s^{p+1}d^p_x + d^{p-1}_Y s^p$. Two morphisms $f,\, g\colon X \to Y$ are *homotopic* if the morphism $f - g\colon X \to Y$ is homotopic to zero. A complex X is *homotopic to 0* if the identity morphism 1_X is homotopic to zero.

Let X, Y be two complexes and define

$$Ht(X, Y) = \{f\colon X \to Y \mid f \text{ is homotopic to zero}\}.$$

If $f\colon X \to Y$ and $g\colon Y \to Z$ are morphisms in $Ch(\mathcal{C})$ and if f or g is homotopic to zero, then $gf\colon X \to Z$ is homotopic to zero. This allow us to define a new category, the *homotopy category* $K(\mathcal{C})$, in which the objects are complexes over $\mathcal{C}$, and for all X, $Y \in Ob(K(\mathcal{C}))$, $\mathrm{Hom}_{K(\mathcal{C})}(X, Y) = \mathrm{Hom}_{Ch(\mathcal{C})}(X, Y)/Ht(x, Y)$. In other words, a morphism homotopic to zero in $Ch(\mathcal{C})$ becomes the zero morphism in $K(\mathcal{C})$ and a homotopy equivalence in $Ch(\mathcal{C})$ becomes an isomorphism in $K(\mathcal{C})$. Similarly, we define $K^*(\mathcal{C})$ for $* = b, +, -$. These categories are clearly additive, but in general not abelian.

3.1.1 Complexes in abelian categories

In this section $\mathcal{C}$ denotes an abelian category. One easily proves that that the categories $Ch^*(\mathcal{C})$ are abelian categories. Let $X \in Ob(Ch(\mathcal{C}))$. We define the following objects of $\mathcal{C}$:

$$Z^n(X) = \operatorname{Ker} d^n_X, \qquad B^n(X) = \operatorname{Im} d^{n-1}_X.$$

Note that there is a natural morphism $B^n(X) \to Z^n(X)$, so we can define the *n-th cohomology object* of X:

$$H^n(X) = Z^n(X)/B^n(X).$$

If $f\colon X \to Y$ is a morphism in $Ch(\mathcal{C})$, then it induces morphisms $Z^n(X) \to Z^n(Y)$ and $B^n(X) \to B^n(Y)$, hence a morphism $H^n(X) \to H^n(Y)$. Therefore, we have an additive functor $H^n\colon Ch(\mathcal{C}) \to \mathcal{C}$. This functor can be extended to a functor $H^n\colon K(\mathcal{C}) \to \mathcal{C}$. A morphism $f\colon X \to Y$ in $Ch(\mathcal{C})$ is said to be a *quasi-isomorphism* (for short, a *qis*) if $H^k(f)\colon H^k(X) \to H^k(Y)$ is an isomorphism for all $k \in \mathbb{Z}$. In this case we say that X and Y are *quasi-isomorphic*. Clearly, a complex X is exact if and only if it is quasi-isomorphic to zero.

Examples 3.1

(a) Let $R = \mathbb{Z}$. The morphism of complexes of abelian groups:

$$\begin{array}{ccccccccccc}
\cdots & \longrightarrow & 0 & \longrightarrow & \mathbb{Z} & \xrightarrow{4\cdot -} & \mathbb{Z} & \longrightarrow & 0 & \longrightarrow & \cdots \\
 & & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle \mathrm{proj}} & & \downarrow{\scriptstyle 0} & & \\
\cdots & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \mathbb{Z}/4\mathbb{Z} & \longrightarrow & 0 & \longrightarrow & \cdots
\end{array}$$

is a quasi-isomorphism in $C(\mathbb{Z})$. More generally:

(b) Any left R-module X can be considered as a complex $\cdots \to 0 \to 0 \to X \to 0 \to 0 \to \cdots$ (with X placed in degree zero). This complex has cohomology 0 outside zero, and its 0-th cohomology is isomorphic to X. Any projective resolution $P^\bullet \to X$ of the module X gives a quasi-isomorphism:

$$\begin{array}{ccccccccccc}
\cdots & \longrightarrow & P^2 & \longrightarrow & P^1 & \longrightarrow & P^0 & \longrightarrow & 0 & \longrightarrow & \cdots \\
 & & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle 0} & & \downarrow & & \downarrow{\scriptstyle 0} & & \\
\cdots & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & X & \longrightarrow & 0 & \longrightarrow & \cdots
\end{array}$$

3.2 The derived category

Theorem 3.2 *Let $\mathcal{C}$ be an abelian category. There exists a triangulated category $D^*(\mathcal{C})$ ($* = ub,\ b,\ +,\ -$), called the* derived category *of $\mathcal{C}$, and a functor $Q\colon K^*(\mathcal{C}) \to D^*(\mathcal{C})$ such that:*

(a) *$Q(s)$ is an isomorphism in whenever s is a quasi-isomorphism;*

(b) *for any functor $F\colon \mathcal{C} \to \mathcal{A}$ such that $F(s)$ is an isomorphism whenever s is a quasi-isomorphism, there exists a functor $\tilde{F}\colon D^*(\mathcal{C}) \to \mathcal{A}$ and a natural isomorphism $F \simeq \tilde{F} \circ Q$;*

(c) *if $G_1, G_2\colon \mathcal{C} \to \mathcal{A}$ are functors, then the natural map*

$$\mathrm{Hom}(G_1, G_2) \to \mathrm{Hom}(G_1 \circ Q, G_2 \circ Q)$$

is a bijection.

Proof. Sketch: we denote by Σ the class of all quasi-isomorphisms of the category $Ch(\mathcal{C})$. We construct the derived category $D(\mathcal{C})$ of $\mathcal{C}$ in two steps.

(a) First we take the homotopy category $K(\mathcal{C})$ of $\mathcal{C}$. The category $K(\mathcal{C})$ comes with a canonical functor $Q'\colon C(\mathcal{C}) \to K(\mathcal{C})$.

(b) In the second step, the morphisms from $Q'(\Sigma)$ are made invertible. That is, we construct $D(\mathcal{C})$ as:

$$D(\mathcal{C}) = K(\mathcal{C})[\Sigma^{-1}]$$

Similar constructions for $D^*(\mathcal{C})$ ($* =$b, $+$, $-$).

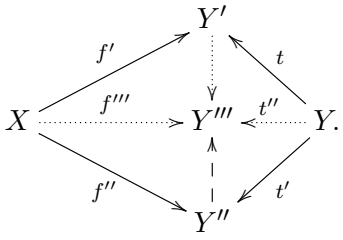
□

Condition (c) means that the functor $- \circ Q \colon Fun(D^*(\mathcal{C}), \mathcal{A}) \to Fun(\mathcal{C}, \mathcal{A})$ is fully faithful. This implies that the functor $\tilde{F}$ in (b) is unique up to unique isomorphism.
The objects in $D^*(\mathcal{C})$ are the objects in $K^*(\mathcal{C})$.

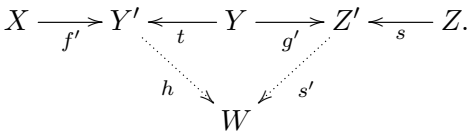
A morphism $f \colon X \to Y$ in $D^*(\mathcal{C})$ is an equivalence class of triplets (Y', t, f'), with $t \colon Y \to Y'$ a qis and $f' \colon X \to Y'$:

$$X \xrightarrow{f'} Y' \xleftarrow{t} Y \ ,$$

and the equivalence relation is defined as follows: $(Y', t, f') \sim (Y'', t', f'')$ if and only if there exists (Y''', t'', f''') (t, t', t'' quasi-isomorphisms) and a commutative diagram:



The composition of two morphisms $(Y', t, f') \colon X \to Y$ and $(Z', s, g') \colon Y \to Z$ is defined by the diagram below with t, s, s' quasi-isomorphisms :



The functor $Q \colon \mathcal{C} \to D^*(\mathcal{C})$ is given by $Q(X) = X$, for all $X \in Ob(\mathcal{C})$, $Q(f) = (Y, 1_Y, f) = X \xrightarrow{f} Y \xleftarrow{1_Y} Y$, for all $f \in \mathrm{Hom}_{\mathcal{C}}(X, Y)$. Note that for a morphism $f = (Y', t, f')$ in $D^*(\mathcal{C})$ we have

$$f = Q(t)^{-1} Q(f').$$

Moreover, for two parallel morphisms $f, g \colon X \to Y$ we have the equivalence

$$Q(f) = Q(g) \iff \text{there exists a qis } s \colon Y \to Y' \text{ such that } sf = sg.$$

References

[1] P. Gabriel, *Des catégories abéliennes.* Bulletin de la Société Mathématique de France 90 (1962), 323–448.

[2] A. Grothendieck, *Sur quelques points d'algèbre homologique.* Tohoku Mathematical Journal, Second Series 9.2 (1957), 119–183.

[3] M. Kashiwara and P. Schapira, "Categories and sheaves". Grundlehren 332, Springer, 2006.

[4] B. Keller, *Deriving DG categories.* Ann. Sci. École Norm. Sup. (4) 27 (1994), no. 1, 63–102.

[5] B. Keller, *Derived categories and tilting.* Handbook of tilting theory, London Math. Soc. Lecture Note Ser., vol. 332, Cambridge Univ. Press, Cambridge, 2007, pp. 49–104.

[6] A. Neeman, *The Grothendieck duality theorem via Bousfields techniques and Brown representability.* Journal of the American Mathematical Society 9 (1996), no. 1, 205–236.

[7] A. Neeman, "Triangulated categories". Princeton, NJ: Princeton University Press, 2001.

[8] J. L. Verdier, *Categories Derivees Quelques résultats (Etat 0).* In "Cohomologie étale" (P. Deligne ed.), Lecture Notes in Mathematics 569, Springer Berlin Heidelberg, 1977, pp. 262–311.

Why should people in approximation theory care about (pluri-)potential Theory?

Federico Piazzon (*)

Abstract. We give a summary of results in (pluri-)potential theory that naturally come into play when considering classical approximation theory issues both in one and (very concisely) in several complex variables. We focus on Fekete points and the asymptotic of orthonormal polynomials for certain L^2 counterpart of Fekete measures.

1 One variable case

1.1 Interpolation and Fekete Points

Let $K \subset \mathbb{C}$ be a compact set such that $\mathbb{C}_\infty \setminus K$ is connected and $f : K \to \mathbb{C}$ be a continuous function that is holomorphic in the interior of K, then Mergelyan Theorem ensures the existence of sequences of polynomials uniformly approximating f on K.

Two reasonable questions are whether we can compute such a sequence by interpolation or not, and how the interpolation nodes should be chosen. It turns out that these questions become much more difficult when suitably translated in the several variables context.

From here on we suppose K to be a polynomial determining set, that is if a polynomial vanish on K then it is the zero polynomial. If we consider the monomial basis $\{z^j\}_{j=1,2,\dots,k}$ for the space $\mathscr{P}^k$ of polynomials of degree at most k and we pick $k+1$ distinct points $\boldsymbol{z}_k := (z_0, z_1, \dots, z_k) \in K^{k+1}$ the interpolation problem can be written as $VDM_k(z_0, \dots, z_k)c = (f(z_0), \dots, f(z_k))^t$ where $c \in \mathbb{C}^{k+1}$ and

$$\underset{k}{\mathrm{VDM}}(z_0, \dots, z_k) := [z_i^j]_{i,j=0,\dots,k}$$

is the Vandermonde Matrix. Notice that for each $k+1$-tuple of distinct points $\det \mathrm{VDM}_k(\boldsymbol{z}_k) \neq 0$ so the problem is well posed.

(*) Ph.D. course, Università di Padova, Dip. Matematica, via Trieste 63, I-35121 Padova, Italy; E-mail: fpiazzon@math.unipd.it . Seminar held on June 24th, 2015.

It is a classical result that the norm of the operator $I_k : (\mathscr{C}(K), \|\cdot\|_K) \to \left(\mathscr{P}^k, \|\cdot\|_K\right)$ is the *Lebesgue Constant* $\Lambda_k := \max_{z\in K} \sum_{m=0}^{k} |l_{m,k}(z)|$, where $l_{m,k}(z) := \frac{\det \mathrm{VDM}_k(z_0,\ldots,z_{l-1},z_l,\ldots,z_k)}{\det \mathrm{VDM}_k(z_0,\ldots,z_k)}$.

Minimizing $\Lambda_k(\boldsymbol{z}_k)$ is an extremely hard task so one can consider the *simplified* problem of maximizing $|\det \mathrm{VDM}(\boldsymbol{z}_k)|$, though it is still a very hard one.

Definition 1.1 (Fekete Points) Let $K \subset \mathbb{C}$ be a compact set and $\boldsymbol{z}_k := (z_0, \ldots, z_k) \in K^{k+1}$. If we have $|\det \mathrm{VDM}_k(\boldsymbol{z}_k)| = \max_{\boldsymbol{\zeta}\in K^{k+1}} |\det \mathrm{VDM}_k(\boldsymbol{\zeta})|$, then $\boldsymbol{z}_k$ is said to be a Fekete array of order k, its elements are said Fekete points.

The relevance of Fekete points is easy to see: one has $\Lambda_k(\boldsymbol{F}_k) \leq k+1$ for any Fekete array $\boldsymbol{F}_k$; moreover notice that in general $\boldsymbol{F}_k$ is not unique.

An interesting property of Vandermonde determinants is that for any array of points $\boldsymbol{z}_k$ we have

$$|\det \underset{k}{\mathrm{VDM}}(\boldsymbol{z}_k)| = \prod_{i<j\leq k} |z_i - z_j|$$

that is the product of their distances; this is one of the main points in connecting interpolation with logarithmic potential theory. We introduce the number $d_k(K) := |\mathrm{VDM}_k(\boldsymbol{z}_k)|^{1/\binom{k}{2}}$ for one (thus any) Fekete array $\boldsymbol{z}_k$ of order k. In analogy with the case $k = 0$, $d_k(K)$ is called *k-th diameter* of K, it is a decreasing positive sequence, its limit $d(K)$ is called *transfinite diameter* of K.

Example 1.2 (Fekete points on the unit disc) Let $\mathbb{D}$ be the unit circle. Take any set of distinct points $\boldsymbol{z} = \{z_0, \ldots, z_k\}$ and consider the determinant of the matrix $V(\boldsymbol{z}) := \mathrm{VDM}_k(\boldsymbol{z})$. We have

$$\|V_{:,j}(\boldsymbol{z})\|_2 = \left\| \begin{pmatrix} z_0^j \\ z_1^j \\ \vdots \\ z_k^j \end{pmatrix} \right\|_2 = \sqrt{k+1}, \text{ for any } j = 0, 1, \ldots k.$$

Therefore Hadamard Inequality for determinants implies $|\det \mathrm{VDM}_k(\boldsymbol{z})| \leq \prod_{j=0}^{k} \|V_{:,j}(\boldsymbol{z})\|_2 = (k+1)^{\frac{k+1}{2}}$. This upper bound is achieved if and only if the columns of $V(\boldsymbol{z})$ are orthonormal and this condition is satisfied for $\boldsymbol{z}$ a set of $k+1$ roots of unity. Therefore $\{\frac{2i\pi j}{k+1}\}_{j=0,\ldots,k}$ is a Fekete set for $\mathbb{D}$.

1.2 Logarithmic Potential Theory

It is customary to introduce the differential operators $d := \partial + \bar{\partial}$ and $d^c := i(\bar{\partial} - \partial)$ such that (passing to real coordinates) one has $\Delta = 2i\partial\bar{\partial} = dd^c$.

We recall that using the Green Identity it is possible to show that $E(z) := \frac{1}{2\pi} \log|z|$ is a *fundamental solution* for the Laplacian, i.e., $\mathrm{dd^c}\, E(z) = \delta_0$ in the sense of distributions. We also recall that a real valued function $u \in \mathscr{C}^2$ on a domain Ω is said to be *harmonic* if $\Delta u = 0$ in Ω. A upper semi-continuous function v is said to be *subharmonic* in Ω if for any open relatively compact subset $\Omega_1 \subset \Omega$ and any harmonic function u on Ω such that

$v \leq u$ on $\partial\Omega_1$ we have $v \leq u$ on Ω_1. It follows that, given any compactly supported finite Borel measure μ, the function

$$U^\mu(z) := \mu * E(z) = \int \log \frac{1}{|z-\zeta|} d\mu(\zeta),$$

said the *logarithmic potential* of μ, is a superharmonic function (i.e. $-U^\mu$ is subharmonic) on $\mathbb{C}$ and in particular harmonic in $\mathbb{C} \setminus \operatorname{supp}\mu$, moreover we have $\mathrm{dd^c}\, U^\mu = \mu$.

To the Laplace operator is naturally attached an energy minimization problem; we refer the reader to [21], [23] and [22] for details.

Problem 1.1 (Logarithmic Energy Minimization) *Let K be a compact subset of $\mathbb{C}$, minimize*

$$\text{(LEM)} \qquad I[\mu] := \int\int \log \frac{1}{|z-\zeta|} d\mu(\zeta) d\mu(z) = \int U^\mu(z) d\mu(z),$$

among $\mu \in \mathcal{M}_1(K)$, the set of Borel probability measures on K endowed with the weak topology.*

We notice that if we consider the electrostatic interaction between $k+1$ unitary charges in the plan, the force acting on the i-th particle is $\sum_{j\neq i} \frac{(x_i,x_j,y_i-y_j)}{|z_i-z_j|^2} = -\nabla U^\mu$, where μ denotes the uniform probability measure associated to the charges distribution. The electrostatic energy associated to this charges distribution is precisely $\tilde{U}^\mu = \sum_{i\neq j} \sum \log \frac{1}{|z_i-z_j|}$, this minimization problem is very close to (LEM), we will see that in particular (LEM) is, in a suitable sense, its limit.

It turns out that $I[\cdot]$ is a lower semi-continuous functional on a locally compact space, one can use the Direct Methods of Calculus of Variation to prove that two situations may occur. Either $I[\mu] = +\infty$ for all $\mu \in \mathcal{M}_1(K)$, either there exists a unique minimizer that is the *equilibrium measure* of K and is usually denoted by μ_K.

In the latter case one has $U^{\mu_K}(z) = -\log c(K) - g_K(z)$, where $g_K := G_{\mathbb{C}_\infty \setminus K}$ is the generalized (e.g. possibly not continuous) Green function with logarithmic pole at ∞ for the complement of K. The number $c(K)$ called *logarithmic capacity* of K and is defined as $c(K) := \exp(-\inf_{\mu\in\mathcal{M}_1(K)} I[\mu])$, thus is non zero precisely when the minimization problem is well posed. It turns out that $U^{\mu_K}(z) = -\log c(K)$ quasi everywhere on K, that is for any $z \in K$ but for set of zero logarithmic capacity.

Sets of zero capacity are, roughly speaking, too small for logarithmic potential theory, they are termed *polar* and it can been shown that $c(K) = 0$ if and only if K is the $-\infty$ set of some subharmonic function defined in a neighbourhood of K.

The following result is due to Szego, Leja and Fekete.

Theorem 1.3 (Fundamental Theorem of Logarithmic Potential Theory) *Let $K \subset \mathbb{C}$ be a compact non polar set, we have*

$$c(K) = d(K).$$

Therefore, for any sequence of Fekete arrays $\{\boldsymbol{F}_k\}$, setting $\mu_k := \frac{1}{k+1}\sum_{j=0}^k \delta_{F_k^j}$, we have

$$\mu_k \rightharpoonup^* \mu_K.$$

Moreover locally uniformly on $\mathbb{C} \setminus K$ *we have*

$$\lim_k -U^{\mu_k}(z) := \lim_k \frac{1}{k+1} \log \prod_{j=0}^{k+1} |z - F_k^j| = g_K(z) - \log c(K) = -U^{\mu_K}.$$

The proof is based on lower semi-continuity and strict convexity of the energy functional and on the extremal property of Fekete points. Moreover the result holds true for any sequence of so called *asymptotically Fekete arrays* that is, arrays $\boldsymbol{L}_k$ such that $\lim_k |\operatorname{VDM}_k(\boldsymbol{L}_k)|^{1/\binom{n}{2}} = d(K)$.

1.3 Back to Approximation

Other deep connections between interpolation and logarithmic potential theory are given by the following two results. We recall that a compact set K is said to be *regular* if g_K is a continuous function.

Theorem 1.4 (Bernstein Walsh [25]) *Let* K *be a compact polynomially determining non polar set, then we have*

$$g_K(z) = \limsup_{\zeta \to z} \left(\left\{ \frac{1}{\deg p} \log^+ |p(\zeta)|, \|p\|_K \leq 1 \right\} \right).$$

Moreover

(Bernstein Wals Ineq.) $$|p(z)| \leq \|p\|_K \exp(\deg p \, g_K(z)) \ \ \forall p \in \mathscr{P}(\mathbb{C}).$$

The approximation theorem comes as an application of Hermite reminder formula and the previous theorem.

Theorem 1.5 (Bernstein-Walsh [25]) *Let* $K \subset \mathbb{C}$ *be a compact regular polynomially convex non polar set and* $f : K \to \mathbb{C}$ *be a bounded function. Let* $d_k(f, K) := \inf\{\|f - p\|_K : p \in \mathscr{P}^k\}$*, then for any real number* $R > 1$ *the following are equivalent*

(1) $\lim_k d_k(f, K)^{1/k} < 1/R$

(2) f *is the restriction to* K *of* $\tilde{f} \in hol(D_R)$*, where* $D_R := \{g_K < \log R\}$.

1.4 L^2 theory

We want to show that some analogues of results for Fekete points (that are L^∞ maximizers, in some sense) holds for particular measures in a L^2 fashion.

Definition 1.6 (Bernstein Markov Measures) Let $K \subset \mathbb{C}$ be a compact set and μ be a Borel measure such that $\operatorname{supp} \mu \subseteq K$, assume that

$$\limsup_k \left(\frac{\|p_k\|_K}{\|p_k\|_{L^2_\mu}} \right)^{1/\deg(p_k)} \leq 1,$$

for any sequence of non zero polynomials $\{p_k\}$. Then we say that (K,μ) has the Bernstein Markov property, BMP for short, or equivalently μ is a Bernstein Markov measure on K.

Example 1.7 We claim that $d\mu := \frac{1}{2\pi}d\theta$ is a Bernstein Markov measure for $\mathbb{S}^1$. To show that, one first notice that the monomials (up to degree k) are a orthonormal basis of $\left(\mathscr{P}^k, \langle\cdot;\cdot\rangle_{L^2_\mu}\right)$. Therefore, for any $p \in \mathscr{P}^k$ we have

$$|p(z)| = \left|\sum_{j=0}^{k}\langle p; z^j\rangle_{L^2_\mu} z^j\right| \leq \|p\|_{L^2_\mu}\left(\sum_{j=0}^{k}|z^j|^2\right)^{1/2} \leq \sqrt{k+1}\|p\|_{L^2_\mu},$$

hence $\limsup_k \left(\frac{\|p_k\|_K}{\|p_k\|_{L^2_\mu}}\right)^{1/\deg(p_k)} \leq \limsup_k (k+1)^{1/2k} = 1.$

Bernstein Markov measures on a given K are in general a very large set as the following sufficient condition shows. For a exhaustive treatment of *regular measures* (e.g. a weakened version of Bernstein Markov property) the reader is referred to [24]; generalization to the $\mathbb{C}^n$ version can be found in [5]. A survey with further development is [10].

Theorem 1.8 (Sufficient condition for BMP [24]) *Let K be a compact non polar regular set and μ a Borel measure such that* $\operatorname{supp}\mu = K$*, assume that there exists a positive number $t > 0$ such that*

$$\lim_{r\to 0^+} c\left(\{z \in K : \mu(B(z,r)) \geq r^t\}\right) = c(K). \tag{1}$$

Then (K,μ) satisfies the Bernstein Markov property.

Finding *necessary condition* is a relevant open question in the general theory of orthogonal polynomials, a necessary condition was stated as a conjecture by Erdös.

The first interest on BMP is from the approximation point of view. Let us take a orthonormal system $\{q_j\}_{j=0,1,\dots}$ for $\mathscr{P}$, then each $\mathscr{P}^k$ is a *Reproducing Kernel Hilbert Space*, being $K^\mu_k(z,\zeta) := \sum_{j=0}^k q_j(z)\bar{q}_j(\zeta)$ the kernel. We denote by $B^\mu_k(z)$ the diagonal of the kernel, say the *Bergman Function* $B^\mu_k(z) = K^\mu_k(z,z) = \sum_{j=0}^k |q_j(z)|^2$. It is not hard to see that the Bergman function represent the worst possible case for the l.h.s. of the definition of BMP, that is $\sup_{\deg p\leq k}\frac{\|p\|_K}{\|p\|_{L^2_\mu}} = \sqrt{\|B^\mu_k\|_K}$.

We consider natural the projection operator $L^\mu_k : \left(\mathscr{C}^0(K), \|\cdot\|_K\right) \to \left(\mathscr{P}^k, \|\cdot\|_K\right)$ defined by embedding the two spaces in L^2_μ, $\mathcal{L}^\mu_k[f](z) := \sum_{j=0}^k \langle f, q_j\rangle q_j(z)$. It follows that

$$\begin{aligned}\|\mathcal{L}^\mu_k[f]\|_K &\leq \left(\sum_{j=0}^k |\langle f, q_j\rangle|^2\right)^{1/2}\left\|\left(\sum_{j=0}^k |q_j(z)|^2\right)^{1/2}\right\|_K \\ &\leq \|f\|_{L^2_\mu}\sqrt{\|B^\mu_k(z)\|_K} \leq \|f\|_K\sqrt{\mu(K)\|B^\mu_k(z)\|_K}.\end{aligned}$$

Therefore we have $\|\mathcal{L}_k^\mu\| \le \sqrt{\mu(K)\|B_k^\mu(z)\|_K}$. This can be used to bound the error of polynomial approximation by least square projection, let p_k be the best uniform polynomial approximation of degree at most k to f

$$\begin{aligned}\|f-\mathcal{L}_k^\mu[f]\|_K &= \|f-p_k+p_k-\mathcal{L}_k^\mu[f]\|_K \\ &\le d_k(f,K)+\|\mathcal{L}_k^\mu[f-p_k]\|_K \ \le\ d_k(f,K)\left(1+\sqrt{\mu(K)\|B_k^\mu(z)\|_K}\right).\end{aligned}$$

This allows us to state a version of the Bernstein Walsh Lemma in a L^2 fashion for Bernstein Markov measures.

Theorem 1.9 (Bernstein Walsh L^2 version [17]) *Let K be a compact polynomially convex regular non polar set and μ a Borel probability measure such that* $\operatorname{supp}\mu = K$ *satisfying the Bernstein Markov property. Then the following are equivalent.*

(1) $\lim_k d_k(f,\mu)^{1/k} < 1/R$

(2) *f is the restriction to K of $\tilde f \in hol(D_R)$, where $D_R := \{g_K < \log R\}$.*

Here $\lim_k d_k(f,\mu)$ is the error of best L^2_μ polynomial approximation to f of degree not greater than k.

There are other interesting analogies between BM measures and measures associated to Fekete points. First we notice that if we pick a Fekete array F_k for K and we compute the squares of the modulus of the Vandermonde determinant on such points we can rewrite it as a L^2 norm w.r.t. the associated Fekete measure μ_k, that is

(2)
$$(k(k+1))!|\det \mathrm{VDM}(F_k)|^2 = \int\dots\int |\det \mathrm{VDM}(\zeta_0,\dots,\zeta_k)|^2 d\mu_k(\zeta_0)\dots d\mu_k(\zeta_k) := Z_k(\mu_k),$$

notice that the right hand side can be generalized to any measure on E. On the other hand if we perform the Gram-Schmidt ortogonalization of the Vandrmonde matrix one the right hand side of (2) we obtain (up to a normalizing constant $(k(k+1))!$) the product of L^2_μ norms of the monic orthonormal polynomials relative to μ and this is precisely the determinant of the Gram-matrix $G_k^{\mu_k}$ w.r.t. $\left(\mathscr{P}^k, \|\cdot\|_{L^2_{\mu_k}}\right)$ in the standard basis, that is $G_k^{\mu_k} = [\langle z^i \bar z^j\rangle_{L^2_{\mu_k}}]_{i,j}$.

This observation leads to a generalization of asymptotically Fekete points to any measure, namely $\mu \in \mathcal{M}_1(E)$ is *asymptotically Fekete* for E if $\lim_k Z_k(\mu,E)^{1/(k(k+1))} = d(E)$.

The following result, despite a not very difficult proof is fundamental, especially in more general contexts: Bernstein Markov measures are asymptotically Fekete; see [8, 6].

Proposition 1.1 *Let K be a compact non polar set and μ a Borel probability measure such that* $\operatorname{supp}\mu \subset K$ *satisfying the Bernstein Markov property. We have*

$$\lim_k \left(\int\dots\int |\det \mathrm{VDM}(\zeta_0,\dots,\zeta_k)|^2 d\mu(\zeta_0)\dots d\mu(\zeta_k)\right)^{\frac{1}{k(k+1)}} = \lim_k \left[(k(k+1))!\det G_k^\mu\right]^{\frac{1}{k(k+1)}} = d(K).$$

Morally speaking, Fekete points are L^∞ maximizers, while BM measures are L^2 maximizers.

Also we have that other interesting properties of Fekete points can be translated in this fashion.

Theorem 1.10 *Let K be a compact non polar set and μ a Borel probability measure such that* $\operatorname{supp}\mu \subset K$ *satisfying the Bernstein Markov property. We have*

i) $\lim_k \frac{1}{2k(k+1)} \log B_k^\mu(z) = g_K(z)$ *point-wise, locally uniformly if K is regular.*

ii) $\lim_k \frac{B_k^\mu}{k(k+1)}\mu = \mu_K$ *in the weak* sense.*

Notice that for Fekete measures $B_k^{\mu_k}$ is the sum of the squared modulus of Lagrange polynomials and $\frac{B_k^{\mu_k}}{k(k+1)} \equiv 1$ on the support of μ_k.

There are other applications of Bernstein Markov measures and potential theory tools concerning for instance random polynomials ensambles generalizing the classical result on Kac polynomials and asymptotic of zeroes of orthogonal polynomials; see for instance [11], [5], [9].

2 Several Variables Case

The extension of what we saw in the case of one complex variable is much more difficult and technical, but still the most of the relation in the previous section have their scv counterpart, provided a correct "translation".

The first difficulty is in defining the n-dimensional transfinite diameter for a given compact set, in particular showing the existence of the limit and its independence by the ordering of the monomial basis. The solution has been given by Zaharjuta [26] by a sophisticate procedure comparing the k-th diameter with certain integral mean of *directional Chebyshev constants*.

The second problem is that logarithmic energy is not related to maximization of Vandermonde determinants when $n > 1$. As a consequence, subharmonic functions are no more the "correct space" to look at; they are replaced in this context by plurisubharmonic ones.

Plurisubharmonic functions, PSH for short, are upper semi continuous functions being subharmonic along each complex line. This property is invariant under any holomorphic mapping, moreover there is a differential operator (the complex Monge Ampere) playing a role with PSH function similar to the one of Laplacian with respect to subharmonic functions in $\mathbb{C}$.

Let $u \in \mathrm{PSH}(\mathbb{C}^n) \cap \mathscr{C}^2$, then one can consider the continuous $(1,1)$ form $\mathrm{dd^c}\, u$,

$$\mathrm{dd^c}\, u := \sum_{i=1}^{n} 2i \frac{\partial^2}{\partial z_i \partial \bar{z}_j} u(z) dz_i \wedge d\bar{z}_j$$

and then take the wedge powers of it

$$(\mathrm{dd^c}\, u)^n := \mathrm{dd^c}\, u \wedge \cdots \wedge \mathrm{dd^c}\, u = \det[\frac{\partial^2}{\partial z_i \partial \bar{z}_j} u(z)]_{i,j} dV_n,$$

where dV_n is the standard volume form on $\mathbb{C}^n$.

It is a classical result that $\mathrm{dd^c}\, u$ can be defined as a *positive current* (i.e., an element of the dual of test forms) for any PSH function and, due to the seminal works of Bedford and Taylor [1] [2], for locally bounded PSH function the operator $(\mathrm{dd^c}\, u)^n$ is well defined as a positive Borel measure. This extension is termed the generalized complex *Monge Ampere operator*, *Pluripotential Theory* is the study of plurisubharmonic functions and Monge Ampere operator; we refer the reader to [15] for a detailed treatment of the subject.

In this context the role of Harmonic function is replaced by *maximal plurisubharmonic functions* that are defined requiring precisely the domination property that harmonic functions enjoy with respect to subharmonic functions in $\mathbb{C}$; they are characterized by $(\mathrm{dd^c}\, u)^n = 0$ in the sense of measures on the given domain. We denote by $L(\mathbb{C}^n)$ the class of plurisubharmonic functions with logarithmic pole at infinity, the Dirichlet problem for the Monge Ampere operator

$$\begin{cases} (\mathrm{dd^c}\, u)^n = 0 & \text{in } \mathbb{C}^n \setminus K \\ u =_{\text{q.e.}} 0 & \text{on } K, u \in L(\mathbb{C}^n) \cap L^\infty_{\text{loc}} \end{cases}$$

enjoys the role of the Dirichlet problem for the Laplacian in $\mathbb{C}$, its solution V_K^* is called *plurisubharmonic extremal function* or, by analogy, *pluricomplex Green function*. Being V_K^* representable, precisely as g_K, with the (upper-semicontinuous regularization of the) upper envelope of function v in $L(\mathbb{C}^n)$ such that $v \leq 0$ on K.

$$V_K^*(z) := \limsup_{\zeta \to z} V_K(\zeta)$$
$$V_K(\zeta) := \sup\{u(\zeta),\, u \in L(\mathbb{C}^n),\, u|_K \leq 1\}.$$

Again, as in the one dimension, one has $\Delta g_K = \mu_K$ here one has a *pluripotential equilibrium measure* $\mu_K := (\mathrm{dd^c}\, V_K^*)^n$, thus it is supported on K by definition.

Pluripotential theory has analogies with potential theory but also differences, first the Monge Ampere operator if fully non linear, there is no notion of potential, a suitable energy functional has been found only recently and there is no direct connection between polynomials and finitely supported measures. However there are plenty of good news as well.

First, the (Bernstein Wals Ineq.) goes precisely to $\mathbb{C}^n$ replacing g_K by V_K^*, due to that and a density result one has the scv counterpart (again replacing g_K by V_K^*) of the Bernstein Walsh Lemma 1.4, usually referred as the *Bernstein Walsh Siciak Theorem*.

For years the extension of the asymptotic property of Fekete points to $\mathbb{C}^n$ has been only conjectured. The work (see [4], [3]) of Berman Boucksom and Nymstrom finally proved that, despite the strong differences between potential and pluripotential theory, one has the same L^∞ and L^2 asymptotic results. More precisely, for a non pluri-polar (i.e., not contained in the $-\infty$ set of a plurisubharmonic function) compact set K the following holds.

(a) Fekete measures for K converge weakly* to μ_K.

(b) The same remains true for any sequence of asymptotically Fekete arrays.

(c) For any Bernstein Markov measure μ $\lim_k \frac{B_k^\mu}{\dim \mathscr{P}^k(\mathbb{C}^n)}\mu = \mu_K$.

(d) For any Bernstein Markov measure μ $\lim_k \frac{1}{2k}\log B_k^\mu = V_K^*$ locally uniformly if K is L-regular, e.g. V_K^* is continuous.

Moreover, the sufficient condition for the Bernstein Markov property can be translated to $\mathbb{C}^n$ by replacing the logarithmic capacity by a non linear "local" (e.g., relative to a open hold all set) capacity associated with the Monge Ampere operator, say the *relative capacity*; see [7].

3 A discrete approach

Admissible meshes, shortly AM, are sequences $\{A_k\}$ of finite subsets of a given compact set K such that

- there exists a positive real constant C such that for any $p \in \mathscr{P}^k$ we have

$$\max_K |p| \le C \max_{A_k} |p|.$$

- Card A_k increase at most polynomially.

They have been first introduced [14] as good sampling sets for uniform polynomial approximation by discrete least squares. The construction of such subsets has been studied for several cases, with emphasis in holding the cardinality growth rate; see for instances [12, 18, 20, 16].

Let associate the uniform probability measure μ_k to A_k, then we can see that, picking an orthonormal system $q_1, \dots, q_{N_k}$ of $\mathscr{P}^k$ we have

$$\sqrt{\| \sum_{j=1}^{N_k} |q_j| \|_K} = \sqrt{\|B_k^{\mu_k}\|_K} \le C \sup_{p\in\mathscr{P}^k} \frac{\|p\|_{A_k}}{\|p\|_{L^2_{\mu_k}}} \le C\sqrt{\operatorname{Card} A_k}.$$

As a consequence the error of uniform polynomial approximation by DLS on an AM has the (far to be sharp) upper bound $\|f - \mathcal{L}_k^{\mu_k}[f]\|_K \le (1 + C\sqrt{\operatorname{Card} A_k}) d_k(f, K)$.

Notice that in particular we shown that $\limsup_k \left(\frac{\|p_k\|_K}{\|p_k\|_{L^2_{\mu_k}}}\right)^{1/k} \le \left(C\sqrt{\operatorname{Card} A_k}\right)^{1/k} = 1$ for any sequence of polynomials p_k, $\deg p_k \le k$. In this sense *admissible meshes are a kind of discrete model of Bernstein Markov measures* suitable for applications since for each finite degree they are finitely supported, moreover in a variety of cases we can explicitly compute an admissible mesh for the given K.

Another analogy of these sequences of finitely supported measures is that it still hold true that *the sequence of uniform probability measures μ_k associated to an admissible mesh for K* s an asymptotically Fekete sequence of measures, namely

$$\lim_k \left(\int \dots \int |\det \mathrm{VDM}(z_1, \dots, z_{N_k})| d\mu_k(z_0) \dots d\mu_k(z_{N_k}) \right)^{\frac{n+1}{2nkN_k}} = d(K), \text{ where } N_k := \binom{k+d}{d}.$$

As a consequence it is possible to prove (following the case of a fixed Bernstein Markov measure; see [13] [19]) that one has

- $\lim_k \frac{B_k^{\mu_k}}{N_k} \mu_k = \mu_K$.
- $\lim_k \frac{1}{2N_k} \log B_k^{\mu_k} = V_K^*$ locally uniformly if K is L-regular.

Lastly we can extract (by numerical linear algebra) from an admissible mesh its Fekete points $F_k \subset A_k$, it turns out that they are asymptotically Fekete for K and thus

- $\lim_k \mu_{F_k} = \mu_K$ in the weak* sense; see [12].

References

[1] E. Bedford and B. A. Taylor, *The Dirichlet proiblem for a complex Monge Ampere equation.* Inventiones Mathematicae, 50 (1976), 129–134.

[2] E. Bedford and B. A. Taylor, *A new capacity for plurisubharmonic functions.* ActaMathematica,149/1 (1962), 1–40.

[3] R. Berman and S. Boucksom, *Growth of balls of holomorphic sections and energy at equilibrium.* Invent. Math. 181/2 (2010), 337–394.

[4] R. Berman, S. Boucksom, and D. W. Nymstrom, *Fekete points and convergence toward equilibrium on complex manifolds.* Acta Mat. 207 (2011), 1–27.

[5] T. Bloom, *Orthogonal polynomials in* $\mathbb{C}^n$. Indiana University Mathematical Journal 46/2 (1997), 427–451.

[6] T. Bloom, L. P. Bos, J. P. Calvi, and N. Levenberg, *Approximation in* $\mathbb{C}^n$. Annales Polonici Mathematici 106 (2012), 53–81.

[7] T. Bloom and N. Levenberg, *Strong asymptotics for Christoffel functions of planar measures.* J. Anal. Math. 106(2008), 353–371.

[8] T. Bloom and N. Levenberg, *Transfinite diameter notions in* $\mathbb{C}^n$ *and integrals of Vandermonde determinants.* Ark. Math. 48/1 (2010), 17–40.

[9] T. Bloom and N. Levenberg, *Pluripotential energy and large deviation.* Indiana Univ. Math. J., 62/2 (2013), 523–550.

[10] T. Bloom, N. Levenberg, F. Piazzon and F. Wielonsky, *Bernstein-Markov: a survey.* Dolomites Notes on Approximation 8 (2015), , to appear (special issue).

[11] T. Bloom, N. Levenberg and F. Wielonsky, *Vector energy and large deviations.* J. Anal. Math. 125/1 (2015), 139–174.

[12] L. Bos, J. P. Calvi, N. Levenberg, A. Sommariva, and M. Vianello, *Geometric weakly admissible meshes, discrete least squares approximation and approximate Fekete points.* Math. Comp. 80/275 (2011), 1623–1638.

[13] L. Bos, N. Levenberg, and S. Waldron, *On the convergence of optimal measures.* Constr. Approx. 32/1 (2010), 159–179.

[14] J. P. Calvi and N. Levenberg, *Uniform approximation by discrete least squares polynomials.* JAT 152 (2008), 82–100.

[15] M. Klimek, "Pluripotential Theory". Oxford Univ. Press, 1991.

[16] A. Kroó, *On optimal polinomial meshes.* JAT 163 (2011), 1107–1124.

[17] N. Levenberg, *Ten lectures on weighted pluripotential theory.* Dolomites Notes on Approximation 5 (2012), 1–59.

[18] F. Piazzon, *Optimal polynomial admissible meshes on compact subsets of* $\mathbb{R}^d$ *with mild boundary regularity.* Preprint submitted to JAT, arXiv:1302.4718 (2013).

[19] F. Piazzon, "Recent result in the theory of polynomial admissible meshes". Master?s thesis, Dep. of Mathematics, University of Padua, Italy. Supervisor: M. Vianello, 2013.

[20] F. Piazzon and M. Vianello, *Small perturbations of admissible meshes.* Appl. Anal. 92 (2013), 1063–1073.

[21] T. Ransford, "Potential Theory in the Complex Plane". Cambridge Univ. Press, 1995.

[22] E. B. Saff, *Logarithmic potential theory with applications to approximation theory.* Surv. Approx. Theory 5 (2010), 165–200.

[23] E. B. Saff and V. Totik, "Logarithmic potentials with external fields". Springer-Verlag Berlin, 1997.

[24] H. Stahl and V. Totik, "General Orthogonal Polynomials". Cambridge Univ. Press, 1992.

[25] J. L. Walsh, "Approximation by Rational function on complex domains". AMS, 1929.

[26] V. P. Zaharjuta, *Transfinite diameter, Chebyshev constant and capacity for compacta in* $\mathbb{C}^n$. USSR Sb. 25/350 (1975).