Advanced Courses in Mathematics CRM Barcelona

Vlad Bally
Lucia Caramellino
Rama Cont

Stochastic Integration by Parts and Functional Itô Calculus

CENTRE DE RECERCA MATEMÀTICA

Birkhäuser

Vlad Bally • Lucia Caramellino • Rama Cont

Stochastic Integration by Parts and Functional Itô Calculus

Editors for this volume:
Frederic Utzet, Universitat Autònoma de Barcelona
Josep Vives, Universitat de Barcelona

Birkhäuser

Vlad Bally
Université de Marne-la-Vallée
Marne-la-Vallée, France

Lucia Caramellino
Dipartimento di Matematica
Università di Roma “Tor Vergata”
Roma, Italy

Rama Cont
Department of Mathematics
Imperial College
London, UK

and

Centre National de Recherche Scientifique (CNRS)
Paris, France

Rama Cont dedicates his contribution to Atossa, for her kindness and constant encouragement.

ISSN 2297-0304 ISSN 2297-0312 (electronic)
Advanced Courses in Mathematics - CRM Barcelona
ISBN 978-3-319-27127-9 ISBN 978-3-319-27128-6 (eBook)
DOI 10.1007/978-3-319-27128-6

Library of Congress Control Number: 2016933913

Printed on acid-free paper

This book is published under the trade name Birkhäuser.
The registered company is Springer International Publishing AG (www.birkhauser-science.com)

Advanced Courses in Mathematics
CRM Barcelona

Centre de Recerca Matemàtica

Managing Editor:
Enric Ventura

More information about this series at http://www.springer.com/series/5038

Foreword

During July 23th to 27th, 2012, the first session of the *Barcelona Summer School on Stochastic Analysis* was organized at the Centre de Recerca Matemàtica (CRM) in Bellaterra, Barcelona (Spain). This volume contains the lecture notes of the two courses given at the school by Vlad Bally and Rama Cont.

The notes of the course by Vlad Bally are co-authored with her collaborator Lucia Caramellino. They develop integration by parts formulas in an abstract setting, extending Malliavin's work on abstract Wiener spaces, and thereby being applicable to prove absolute continuity for a broad class of random vectors. Properties like regularity of the density, estimates of the tails, and approximation of densities in the total variation norm are considered. The last part of the notes is devoted to introducing a method to prove existence of density based on interpolation spaces. Examples either not covered by Malliavin's approach or requiring less regularity are in the scope of its applications.

Rama Cont's notes are on Functional Itô Calculus. This is a non-anticipative functional calculus extending the classical Itô calculus to path-dependent functionals of stochastic processes. In contrast to Malliavin Calculus, which leads to *anticipative* representation of functionals, with Functional Itô Calculus one obtains *non-anticipative* representations, which may be more natural in many applied problems. That calculus is first introduced using a pathwise approach (that is, without probabilities) based on a notion of directional derivative. Later, after the introduction of a probability on the space of paths, a weak functional calculus emerges that can be applied without regularity conditions on the functionals. Two applications are studied in depth; the representation of martingales formulas, and then a new class of path-dependent partial differential equations termed *functional Kolmogorov equations.*

We are deeply indebted to the authors for their valuable contributions. Warm thanks are due to the Centre de Recerca Matemàtica, for its invaluable support in the organization of the School, and to our colleagues, members of the Organizing Committee, Xavier Bardina and Marta Sanz-Solé. We extend our thanks to the following institutions: AGAUR (Generalitat de Catalunya) and Ministerio de Economía y Competitividad, for the financial support provided with the grants SGR 2009-01360, MTM 2009-08869 and MTM 2009-07203.

Frederic Utzet and Josep Vives

Contents

Part I

Integration by Parts Formulas, Malliavin Calculus, and Regularity of Probability Laws

Vlad Bally and Lucia Caramellino

Preface

The lectures we present here turn around the following integration by parts formula. On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we consider a random variable G, and a random vector $F = (F_1, \ldots, F_d)$ taking values in $\mathbb{R}^d$. Moreover, we consider a multiindex $\alpha = (\alpha_1, \ldots, \alpha_m) \in \{1, \ldots, d\}^m$ and we write $\partial_\alpha = \partial_{x^{\alpha_1}} \cdots \partial_{x^{\alpha_m}}$. We look for a random variable, which we denote by $H_\alpha(F; G) \in L^p(\Omega)$, such that the following integration by parts formula holds:

$$\mathrm{IBP}_{\alpha,p}(F,G) \qquad \mathbb{E}(\partial_\alpha f(F)G) = \mathbb{E}(f(F)H_\alpha(F;G)), \quad \forall f \in C_b^\infty(\mathbb{R}^d),$$

where $C_b^\infty(\mathbb{R}^d)$ denotes the set of infinitely differentiable functions $f\colon \mathbb{R}^d \to \mathbb{R}$ having bounded derivatives of any order. This is a set of test functions, that can be replaced by other test functions, such as the set $C_c^\infty(\mathbb{R}^d)$ of infinitely differentiable functions $f\colon \mathbb{R}^d \to \mathbb{R}$ with compact support.

The interest in this formula comes from the Malliavin calculus; this is an infinite-dimensional differential calculus associated to functionals on the Wiener space which permits to construct the weights $H_\alpha(F;G)$ in the above integration by parts formula. We will treat several problems related to $\mathrm{IBP}_{\alpha,p}(F,G)$ that we list now.

Problem 1

Suppose one is able to produce $\mathrm{IBP}_{\alpha,p}(F,G)$, does not matter how. What can we derive from this? In the classical Malliavin calculus such formulas with $G = 1$ have been used in order to:

(1) prove that the law of F is absolutely continuous and, moreover, to study the regularity of its density;

(2) give integral representation formulas for the density and its derivatives;

(3) obtain estimates for the tails of the density, as well;

(4) obtain similar results for the conditional expectation of G with respect to F, assuming If $\mathrm{IBP}_{\alpha,p}(F,G)$ holds (with a general G).

Our first aim is to derive such properties in the abstract framework that we describe now.

A first remark is that $\mathrm{IBP}_{\alpha,p}(F,G)$ does not involve random variables, but the law of the random variables: taking conditional expectations in the above formula we get $\mathbb{E}(\partial_\alpha f(F)\mathbb{E}(G \mid \sigma(F))) = \mathbb{E}(f(F)\mathbb{E}(H_\alpha(F;G) \mid \sigma(F))$. So, if we denote $g(x) = \mathbb{E}(G \mid F = x)$, $\theta_\alpha(g)(x) = \mathbb{E}(H_\alpha(F;G) \mid F = x)$, and if $\mu_F(dx)$ is the law of F, then the above formula reads

$$\int \partial_\alpha f(x)g(x)d\mu_F(x) = \int f(x)\theta_\alpha(g)(x)d\mu_F(x).$$

If $\mu_F(dx)$ is replaced by the Lebesgue measure, then this is a standard integration by parts formula and the theory of Sobolev spaces comes on. But here, we have the specific point that the reference measure is μ_F, the law of F. We denote $\partial_\alpha^{\mu_F} g = \theta_\alpha(g)$ and this represents somehow a weak derivative of g. But it does not verify the chain rule, so it is not a real derivative. Nevertheless, we may develop a formalism which is close to that of the Sobolev spaces, and express our results in this formalism. Shigekawa [45] has already introduced a very similar formalism in his book, and Malliavin introduced the concept of *covering vector fields*, which is somehow analogous. The idea of giving an abstract framework related to integration by parts formulas already appears in the book of Bichteler, Gravereaux and Jacod [16] concerning the Malliavin calculus for jump processes.

A second ingredient in our approach is to use the Riesz representation formula; this idea comes from the book of Malliavin and Thalmaier [36]. So, if Q_d is the Poisson kernel on $\mathbb{R}^d$, i.e., the solution of $\triangle Q_d = \delta_0$ (δ_0 denoting the Dirac mass), then a formal computation using $\mathrm{IBP}_{\alpha,p}(F,1)$ gives

$$p_F(x) = \mathbb{E}(\delta_0(F-x)) = \sum_{i=1}^d \mathbb{E}\big(\partial_i^2 Q_d(F-x)\big) = \sum_{i=1}^d \mathbb{E}\big(\partial_i Q_d(F-x)H_i(F,1)\big).$$

This is the so-called Malliavin–Thalmaier representation formula for the density. One may wonder why do we perform one integration by parts and not two? The answer is that Q_d is singular in zero and then $\partial_i^2 Q_d$ is "not integrable" while $\partial_i Q_d$ is "integrable"; so, as we expect, integration by parts permits to regularize things. As suggested before, the key issue when using Q_d is to handle the integrability problems, and this is an important point in our approach. The use of the above representation allows us to ask less regularity for the random variables at hand (this has already been remarked and proved by Shigekawa [45] and by Malliavin [33]).

All these problems —(1), (2), (3), and (4) mentioned above— are discussed in Chapter 1. Special attention is given to *localized integration by parts formulas* which, roughly speaking, means that we take G to be a smooth version of $1_{\{F\in A\}}$. This turns out to be extremely useful in a number of problems.

Problem 2

How to construct $\mathrm{IBP}_{\alpha,p}(F,G)$? The strategy in Malliavin calculus is roughly speaking as follows. One defines a differential operator D on a subspace of "regular objects" which is dense in the Wiener space, and then takes the standard extension of this unbounded operator. There are two main approaches: the first one is to consider the expansion in Wiener chaos of the functionals on the Wiener space. In this case the multiple stochastic integrals are the "regular objects" on which D is defined directly. We do not take this point of view here (see Nualart [41] for a complete theory in this sense). The second one is to consider the cylindrical functions of the Brownian motion W as "regular objects", and this is our point of view. We consider $F_n = f(\Delta_n^1, \dots, \Delta_n^{2^n})$ with $\Delta_n^k = W(k/2^n) - W((k-1)/2^n)$ and f being a smooth function, and we define $D_s F = \partial_k f(\Delta_n^1, \dots, \Delta_n^{2^n})$ if $(k-1)/2^n \leq s < k/2^n$. Then, given a general functional F, we look for a sequence F_n of simple functionals such that $F_n \to F$ in $L^2(\Omega)$, and if $DF_n \to U$ in $L^2([0,1] \times \Omega)$, then we define $D_s F = U_s$.

Looking to the "duality formula" for DF_n (which is the central point in Malliavin calculus), one can see that the important point is that we know explicitly the density p of the law of $(\Delta_n^1, \dots, \Delta_n^{2^n})$ (Gaussian), and this density comes on in the calculus by means of $\ln p$ and of its derivatives only. So we may mimic all the story for a general finite-dimensional random vector $V = (V_1, \dots, V_m)$ instead of $\Delta = (\Delta_n^1, \dots, \Delta_n^{2^n})$. This is true in the finite-dimensional case (for simple functionals), but does not go further to general functionals (the infinite-dimensional calculus). Nevertheless, we will see in a moment that, even without passing to the limit, integration by parts formulas are useful. This technology has already been used in the framework of jump type diffusions in [8, 10]. In Chapter 2, and specifically in Section 2.1, we present this "finite-dimensional abstract Malliavin calculus" and then we derive the standard infinite-dimensional Malliavin calculus. Moreover, we use these integration by parts formulas and the general results from Chapter 1 in order to study the regularity of the law in this concrete situation. But not only this: we also obtain quantitative estimates concerning the density of the law as mentioned in the points (1), (2), (3), and (4) from Problem 1.

Problem 3

There are many other applications of the integration by parts formulas in addition to the regularity of the law. We mention here few ones.

Malliavin calculus permits to obtain results concerning the rate of convergence of approximation schemes (for example the Euler approximation scheme for diffusion processes) for test functions which are not regular but only measurable and bounded. And moreover, under more restrictive hypotheses and with a little bit more effort, to prove the convergence of the density functions; see for example [12, 13, 27, 28, 30]. Another direction is the study of the density in short time (see, e.g., Arous and Leandre [14]), or the strict positivity of the density (see, e.g., Bally and Caramellino [3]). We do not treat here these problems, but we give a result which is central in such problems: in Section 2.4 we provide an estimate of the distance between the density functions of two random variables in terms of the weights $H_\alpha(F,G)$, which appear in the integration by parts formulas. Moreover we use this estimates in order to give sufficient conditions allowing one to obtain convergence in the total variation distance for a sequence of random variables which converge in distribution. The localization techniques developed in Chapter 1 play a key role here.

Problem 4

We present an alternative approach for the study of the regularity of the law of F. The starting point is the paper by Fournier and Printems [26]. Coming back to the approach presented above, we recall that we have defined $D_sF = \lim_n D_sF_n$ and then we used DF in order to construct $\mathrm{IBP}_{\alpha,p}(F,1)$. But one may proceed in an alternative way: since DF_n is easy to define (it is just a finite-dimensional gradient), one may use elementary integration by parts formulas (in finite-dimensional spaces) in order to obtain $\mathrm{IBP}_{\alpha,p}(F_n,1)$. Then passing to the limit $n \to \infty$ in $\mathrm{IBP}_{\alpha,p}(F_n,1)$, one obtains $\mathrm{IBP}_{\alpha,p}(F,1)$. If everything works well, then we are done —but we are not very far from the Malliavin calculus itself. The interesting fact is that sometimes this argument still works even if things are "going bad", that is, even when $H_\alpha(F_n,1) \uparrow \infty$. The idea is the following. We denote by $\widehat{p}_F$ the Fourier transform of F and by $\widehat{p}_{F_n}$ the Fourier transform of F_n. If we are able to prove that $\int |\widehat{p}_F(\xi)|^2 \, d\xi < \infty$, then the law of F is absolutely continuous. In order to do it we notice that $\partial_x^m e^{i\xi x} = (i\xi)^m e^{i\xi x}$ and we use $\mathrm{IBP}_{m,p}(F_n,1)$ in order to obtain

$$\widehat{p}_{F_n}(\xi) = \frac{1}{(i\xi)^m}\mathbb{E}\big(\partial_x^m e^{i\xi F_n}\big) = \frac{1}{(i\xi)^m}\mathbb{E}\big(e^{i\xi F_n} H_m(F_n,1)\big).$$

Then, for each $m \in \mathbb{N}$,

$$|\widehat{p}_F(\xi)| \le |\widehat{p}_F(\xi) - \widehat{p}_{F_n}(\xi)| + |\widehat{p}_{F_n}(\xi)| \le |\xi|\,\mathbb{E}\big(\,|F - F_n|\,\big) + \frac{1}{|\xi|^m}\mathbb{E}\big(\,|H_m(F_n,1)|\,\big).$$

So, if we obtain a good balance between $\mathbb{E}(|F-F_n|) \downarrow 0$ and $\mathbb{E}(|H_m(F_n,1))| \uparrow \infty$, we have a chance to prove that $\int_{\mathbb{R}^d} |\widehat{p}_F(\xi)|^2 \, d\xi < \infty$ and so to solve our problem. One unpleasant point in this approach is that it depends strongly on the dimension d of the space: the above integral is convergent if $|\widehat{p}_F(\xi)| \le \mathrm{Const}\,|\xi|^{-\alpha}$ with $\alpha > d/2$, and this may be rather restrictive for large d. In [20], Debussche and Romito presented an alternative approach which is based on certain relative compactness criterion in Besov spaces (instead of the Fourier transform criterion), and their method is much more performing. Here, in Chapter 3 we give a third approach based on an expansion in Hermite series. We also observe that our approach fits in the theory of interpolation spaces and this seems to be the natural framework in which the equilibrium between $\mathbb{E}(|F-F_n|) \downarrow 0$ and $\mathbb{E}(|H_m(F_n,1)|) \uparrow \infty$ has to be discussed.

The class of methods presented above goes in a completely different direction than the Malliavin calculus. One of the reasons is that Malliavin calculus is somehow a pathwise calculus —the approximations $F_n \to F$ and $DF_n \to DF$ are in some L^p spaces and so, passing to a subsequence, one can also achieve almost sure convergence. This allows one to use it as an efficient instrument for analysis on the Wiener space (the central example is the Clark–Ocone representation formula). In particular, one has to be sure that nothing blows up. In contrast, the argument presented above concerns the laws of the random variables at hand and, as mentioned, it allows to handle the blow-up. In this sense it is more flexible and permits to deal with problems which are out of reach for Malliavin calculus. On the other hand, it is clear that if Malliavin calculus may be used (and so real integration by parts formulas are obtained), then the results are more precise and deeper: because one does not only obtain the regularity of the law, but also integral representation formulas, tail estimates, lower bounds for densities and so on. All this seems out of reach with the asymptotic methods presented above.

Conclusion

The results presented in these notes may be seen as a complement to the classical theory of Sobolev spaces on $\mathbb{R}^d$ that are suited to treat problems of regularity of probability laws. We stress that this is different from Watanabe's distribution theory on the Wiener space. Let $F\colon \mathcal{W} \to \mathbb{R}^d$. The distribution theory of Watanabe deals with the infinite-dimensional space $\mathcal{W}$, while the results in these notes concern μ_F, the law of F, which lives in $\mathbb{R}^d$. The Malliavin calculus comes on in the construction of the weights $H_\alpha(F, G)$ in the integration by parts formula (1) and its estimates. We also stress that the point of view here is somehow different from the one in the classical theory of Sobolev spaces: there the underling measure is the Lebesgue measure on $\mathbb{R}^d$, whereas here, if F is given, then the basic measure in the integration by parts formula is μ_F. This point of view comes from Malliavin calculus (even if we are in a finite-dimensional framework). Along these lectures, almost all the time, we fix F. In contrast, the specific achievement of Malliavin calculus (and of Watanabe's distribution theory) is to provide a theory in which one can deal with all the "regular" functionals F in the same framework.

Vlad Bally, Lucia Caramellino

Chapter 1

Integration by parts formulas and the Riesz transform

The aim of this chapter is to develop a general theory allowing to study the existence and regularity of the density of a probability law starting from integration by parts type formulas (leading to general Sobolev spaces) and the Riesz transform, as done in [2]. The starting point is given by the results for densities and conditional expectations based on the Riesz transform given by Malliavin and Thalmaier [36]. Let us start by speaking in terms of random variables.

Let F and G denote random variables taking values on $\mathbb{R}^d$ and $\mathbb{R}$, respectively, and consider the following integration by parts formula: there exist some integrable random variables $H_i(F,G)$ such that for every test function $f \in C_c^\infty(\mathbb{R}^d)$

$$\mathrm{IBP}_i(F,G) \quad \mathbb{E}\big(\partial_i f(F)G\big) = -\mathbb{E}\big(f(F)H_i(F,G)\big), \quad i = 1,\dots,d.$$

Malliavin and Thalmaier proved that if $\mathrm{IBP}_i(F,1)$, $i = 1,\dots,d$, hold and the law of F has a continuous density p_F, then

$$p_F(x) = -\sum_{i=1}^d \mathbb{E}(\partial_i Q_d(F-x)H_i(F,1)),$$

where Q_d denotes the Poisson kernel on $\mathbb{R}^d$, that is, the fundamental solution of the Laplace operator. Moreover, they also proved that if $\mathrm{IBP}_i(F,G)$, $i = 1,\dots,d$, a similar representation formula holds also for the conditional expectation of G with respect to F. The interest of Malliavin and Thalmaier in these representations came from numerical reasons —they allow one to simplify the computation of densities and conditional expectations using a Monte Carlo method. This is crucial in order to implement numerical algorithms for solving nonlinear PDE's or optimal stopping problems. But there is a difficulty one runs into: the variance of the estimators produced by such a representation formula is infinite. Roughly speaking, this comes from the blowing up of the Poisson kernel around zero: $\partial_i Q_d \in L^p$ for $p < d/(d-1)$, so that $\partial_i Q_d \notin L^2$ for every $d \geq 2$. Hence, estimates of $\mathbb{E}(|\partial_i Q_d(F-x)|^p)$ are crucial in this framework and this is the central point of interest here. In [31, 32], Kohatsu-Higa and Yasuda proposed a solution to this problem using some cut-off arguments. And, in order to find the optimal cut-off level, they used the estimates of $\mathbb{E}(|\partial_i Q_d(F-x)|^p)$ which are proven in Theorem 1.4.1.