Lecture Notes in Mathematics 1908

Subseries:
École d'Été de Probabilités de Saint-Flour

Saint-Flour Probability Summer School

The Saint-Flour volumes are reflections of the courses given at the Saint-Flour Probability Summer School. Founded in 1971, this school is organised every year by the Laboratoire de Mathématiques (CNRS and Université Blaise Pascal, Clermont-Ferrand, France). It is intended for PhD students, teachers and researchers who are interested in probability theory, statistics, and in their applications.

The duration of each school is 13 days (it was 17 days up to 2005), and up to 70 participants can attend it. The aim is to provide, in three high-level courses, a comprehensive study of some fields in probability theory or Statistics. The lecturers are chosen by an international scientific board. The participants themselves also have the opportunity to give short lectures about their research work.

Participants are lodged and work in the same building, a former seminary built in the 18th century in the city of Saint-Flour, at an altitude of 900 m. The pleasant surroundings facilitate scientific discussion and exchange.

The Saint-Flour Probability Summer School is supported by:

- Université Blaise Pascal
- Centre National de la Recherche Scientifique (C.N.R.S.)
- Ministère délégué à l'Enseignement supérieur et à la Recherche

For more information, see back pages of the book and
http://math.univ-bpclermont.fr/stflour/

Jean Picard
Summer School Chairman
Laboratoire de Mathématiques
Université Blaise Pascal
63177 Aubière Cedex
France

Terry J. Lyons
Michael Caruana · Thierry Lévy

Differential Equations Driven by Rough Paths

École d'Été de Probabilités
de Saint-Flour XXXIV - 2004

Springer

Authors

Terry J. Lyons
Michael Caruana
Mathematical Institute
University of Oxford
24-29 St. Giles'
Oxford OX1 3LB
United Kingdom
e-mail: tlyons@maths.ox.ac.uk
caruana@maths.ox.ac.uk

Thierry Lévy
Département de Mathématiques et
Applications
École Normale Supérieure
45, rue d'Ulm
75230 Paris Cedex 05
France
e-mail: thierry.levy@ens.fr

Cover: Blaise Pascal (1623-1662)

Library of Congress Control Number: 2007922938

Mathematics Subject Classification (2000): 60H10, 60H05, 34A99, 60H30

ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692
ISSN Ecole d'Eté de Probabilités de St. Flour, print edition: 0721-5363
ISBN-10 3-540-71284-4 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-71284-8 Springer Berlin Heidelberg New York
DOI 10.1007/978-3-540-71285-5

Springer is a part of Springer Science+Business Media
springer.com

Typesetting by the authors and SPi using a Springer LaTeX macro package
Cover design: *design & production* GmbH, Heidelberg

Printed on acid-free paper SPIN: 12031597 VA41/3100/SPi 5 4 3 2 1 0

A Word About the Summer School

The Saint-Flour Probability Summer School was founded in 1971. It is supported by CNRS, the "Ministère de la Recherche", and the "Université Blaise Pascal".

Three series of lectures were given at the 34th Probability Summer School in Saint-Flour (July 6–24, 2004), by the Professors Cerf, Lyons and Slade. We have decided to publish these courses separately. This volume contains the course of Professor Lyons; this final version has been written with two participants of the school, Michael Caruana and Thierry Lévy. We cordially thank them, as well as Professor Lyons for his performance at the summer school.

Sixty-nine participants have attended this school. Thirty-five of them have given a short lecture. The lists of participants and of short lectures are enclosed at the end of the volume.

Here are the references of Springer volumes which have been published prior to this one. All numbers refer to the *Lecture Notes in Mathematics* series, except S-50 which refers to volume 50 of the *Lecture Notes in Statistics* series.

1971: vol 307	1980: vol 929	1990: vol 1527	1998: vol 1738
1973: vol 390	1981: vol 976	1991: vol 1541	1999: vol 1781
1974: vol 480	1982: vol 1097	1992: vol 1581	2000: vol 1816
1975: vol 539	1983: vol 1117	1993: vol 1608	2001: vol 1837 & 1851
1976: vol 598	1984: vol 1180	1994: vol 1648	2002: vol 1840 & 1875
1977: vol 678	1985/86/87: vol 1362 & S-50	1995: vol 1690	2003: vol 1869 & 1896
1978: vol 774	1988: vol 1427	1996: vol 1665	2004: vol 1878, 1879 & 1908
1979: vol 876	1989: vol 1464	1997: vol 1717	2005: vol 1897

Further details can be found on the summer school web site
`http://math.univ-bpclermont.fr/stflour/`

Jean Picard
Clermont-Ferrand, September 2006

Foreword

I am ashamed to say that my first visit to the annual summer school at Saint Flour was to give the lectures recorded in the pages to follow. The attentive and supportive audience made the opportunity into a true privilege. I am grateful for the opportunity, to Jean Picard and his team for their considerable efforts and the audiences for their interest and patience.

I had been very busy in the weeks before the lectures and came equipped with a detailed outline and bibliography and a new tablet PC. Unlike my well-organised co-authors, who came with their printed notes, I wrote the detailed lecture notes as I went. Jean Picard, with his inevitable charm, tact and organisational skill, brought to the very rural setting of St Flour, a fast laser printer and every night there was a long period printing off and stapling the notes for the next days lecture (60 times over, as there was no Xerox machine!).

The notes were, in the main, hand-written on the tablet PC, supplemented by a substantial set of preprints and publications for the improvised library. This worked adequately and also provided some amusement for the audience: the computer placed a time stamp on each page permitting the audience to see which pages were written at 2.00 am and how long each page (that survived) took to write. Writing ten 90-min lectures in two weeks was a demanding but enjoyable task.

Two members of the meeting, Michael Caruana, and Thierry Lévy offered to convert the notes into the form you find here. They have taken my presentation to pieces, looked at it afresh, and produced a version that is cleaner and more coherent than I ever could have managed or imagined. I do not know how to express my gratitude. The original hand-written notes are, at the time of writing, to be found at:
`http://sag.maths.ox.ac.uk/tlyons/st_flour/index.htm`

The goal of these notes is to provide a straightforward and self-supporting but minimalist account of the key results forming the foundation of the theory of rough paths. The proofs are similar to those in the existing literature, but have been refined with the benefit of hindsight. We hope that the overall

presentation optimises transparency and provides an accessible entry point into the theory without side distractions. The key result (the convergence of Picard Iteration and the universal limit theorem) has a proof that is significantly more transparent than in the original papers.

We hope they provide a brief and reasonably motivated account of rough paths that would equip one to study the published work in the area or one of the books that have or are about to appear on the topic.

Mathematical Goal

The theory of rough paths aims to create the appropriate mathematical framework for expressing the relationships between evolving systems, by extending classical calculus to the natural models for noisy evolving systems, which are often far from differentiable.

A rather simple idea works; differential equations obviously have meaning when used to couple smoothly evolving systems. If one could find metrics on smooth paths for which these couplings were uniformly continuous, then one could complete the space. The completions of the space of smooth paths are not complicated or too abstract and considering these spaces of "generalised paths" as the key spaces where evolving systems can be defined, modelled and studied seems fruitful. This approach has a number of applications, a few of which are mentioned in the notes. But the minimalistic approach we have set ourselves means we limit such discussion severely – the applications seem to still be developing and quite distinctive so we would commit the reader into much extra work and defeat the overall goal of this text. In Saint Flour it was natural to give probabilistic applications. The hand-written notes give the first presentation of a proof for a quite precise extension of the support theorem not reproduced here.

In 1936 Young introduced an extension to Stieltjes integration which applies to paths of p-variation less than 2. In a separate line of development, Chen (1957, geometry) and more recently Fliess (control theory), E. Platen (stochastic differential equations) and many others were lead to consider the sequence of iterated integrals of a path x in order to obtain a pathwise Taylor series of arbitrary order for the solution y to the vector equation

$$dy_t = f(y_t)\, dx_t.$$

These notes develop the non-commutative analysis required to integrate these two developments into the theory of rough paths, a mathematical framework for modelling the interaction between evolving systems.

Oxford, March 2006 *Terry J. Lyons*

Introduction

These notes put on record a series of ten lectures delivered by Terry Lyons at the Saint Flour summer school in July 2004. Terry Lyons's declared purpose was to bring the audience to the central result of the theory of rough paths along the straightest possible path. These notes, which follow closely the content of the lectures, are thus primarily intended for a reader who has never been exposed to the theory of rough paths. This introduction gives an overview of the subject and presents the content of each chapter, especially the first three, in some detail.

The theory of rough paths can be briefly described as a non-linear extension of the classical theory of controlled differential equations which is robust enough to allow a deterministic treatment of stochastic differential equations, or even controlled differential equations driven by much rougher signals than semi-martingales.

Let us first explain what a controlled differential equation is. In a setting where everything is differentiable, it is a differential equation of the form

$$\dot{Y}_t = F(\dot{X}_t, Y_t), \quad Y_0 = \xi, \tag{1}$$

where X is a given function, ξ is an initial condition, and Y is the unknown. The mapping F is taken to be linear with respect to its first variable. If F did not depend on its first variable, this would be the most general first-order time-homogeneous differential equation. The function F would be a vector field and Y would be the integral curve of this vector field starting at ξ. If now we had $\dot{Y}_t = F(t, Y_t)$ instead of (1), this would be the most general first-order time-inhomogeneous differential equation. The solution Y would be an integral curve of the time-dependent vector field F. Equation (1) is really of this kind, except that the time-inhomogeneity has been made explicitly dependent on a path X, which is said to control the equation. The physical meaning of (1) is the following: at each time, Y describes the state of a complex system (for instance the brain, or a car) and it evolves as a function of its present state and the infinitesimal variation of an external parameter (like the air pressure near the ears or the angle of the steering wheel).

Stochastic differential equations are of the form (1), except that X is usually far from being differentiable. They are often written under the form

$$dY_t = f(Y_t)\,dX_t, \quad Y_0 = \xi, \tag{2}$$

which emphasises the linear dependence of the right-hand side with respect to dX_t. K. Itô, in introducing the concept of strong solution, emphasised the fact that the resolution of a stochastic differential equation amounts to the construction of a mapping between spaces of paths. For this reason, when the (deterministic) equation (2) admits a unique solution, we denote this solution by $Y = I_f(X, \xi)$, and we call I_f the Itô map associated with f.

The classical theory of differential equations tells us that if f is Lipschitz continuous, then (2) admits a unique solution as soon as X has bounded variation and this solution has bounded variation. Moreover, the Itô map I_f is continuous as a mapping between spaces of paths with bounded variation. The fundamental results of the theory of rough paths resolve the following two problems:

1. Identify a natural family of metrics on the space of paths with bounded variation such that the Itô map $X \mapsto I_f(X, \xi)$ is uniformly continuous with respect to these metrics, at least when f is regular enough.
2. Describe concretely the completion of the space of paths with bounded variation with respect to these metrics.

Let us give the solutions in a very condensed form. In the simplest setting, the appropriate metrics depend on a real parameter $p \in [1, +\infty)$ and two paths are close in the so-called *p-variation metric* if they are close in p-variation (a parameter-independent version of $\frac{1}{p}$-Hölder norm), as well as their first $\lfloor p \rfloor$ iterated integrals. An element of the completion of the space of paths with bounded variation on some interval $[0, T]$ with respect to the p-variation metric is called a *rough path* and consists in the data, for each sub-interval $[s, t]$ of $[0, T]$, of $\lfloor p \rfloor$ tensors – the first of which is the increment $x_t - x_s$ of some continuous path x – which summarise the behaviour of this rough path on $[s, t]$ in an efficient way, as far as controlled differential equations are concerned. This collection of tensors must satisfy some algebraic consistency relations and some analytic conditions similar to $\frac{1}{p}$-Hölder continuity.

The theorem which is proved in Chap. 5 of these notes states, under appropriate hypotheses, the existence and uniqueness of the solution of a differential equation controlled by a rough path.

It is worth noting that if the control X takes its values in a one-dimensional space, then the theory of rough paths becomes somehow trivial. Indeed, provided f is continuous, the Itô map is continuous with respect to the topology of uniform convergence and extends to the space of all continuous controls. The theory of rough paths is thus meaningful for *multi-dimensional* controls.

The coexistence of algebraic and analytic aspects in the definition of rough paths makes it somewhat difficult at first to get a general picture of the theory.

Chapters 1 and 2 of these notes present separately the main analytical and algebraic features of rough paths.

Chapter 1 is devoted to the concept of p-variation of a Banach-valued continuous function on an interval. Given a Banach space V and a real number $p \geq 1$, a continuous function $X : [0,T] \longrightarrow V$ is said to have finite p-variation if

$$\sup \sum_{\mathcal{D}} |X_{t_{i+1}} - X_{t_i}|^p < +\infty,$$

where the supremum is taken over all subdivisions $\mathcal{D}$ of $[0,T]$. This is equivalent to the fact that X can be reparametrised as a $\frac{1}{p}$-Hölder continuous function.

The central result of this chapter states that the classical Stieltjes integral $\int_0^t Y_u \, dX_u$, defined when Y is continuous and X has bounded variation, has an extension to the case where X and Y have finite p- and q-variation, respectively, provided $p^{-1} + q^{-1} > 1$. Moreover, in this case, the integral, as a function of t, has finite p-variation, like X. This was discovered by Young around 1930 and allows one to make sense of and even, if f is regular enough, to solve (2) when X has finite p-variation for $p < 2$.

On the other hand, if X is a real-valued path with finite p-variation for $p \geq 2$, then in general the Riemann sums $\sum X_{t_i}(X_{t_{i+1}} - X_{t_i})$ fail to converge as the mesh of the subdivision tends to 0 and one cannot anymore define $\int_0^t X_u \, dX_u$. One could think that this threshold at $p = 2$ is due to some weakness of the Young integral, but this is not the case. A simple and very concrete example shows that the mapping which to a path $X = (X_1, X_2) : [0,T] \longrightarrow \mathbb{R}^2$ with bounded variation associates the real number

$$\frac{1}{2} \int_0^T X_{1,u} \, dX_{2,u} - X_{2,u} \, dX_{1,u} \tag{3}$$

is *not* continuous in p-variation for $p > 2$.

The number defined by (3) is not just a funny counter-example. Firstly, it has a very natural geometric interpretation as the area enclosed by the curve X. Secondly, it is not very difficult to write it as the final value of the solution of a differential equation controlled by X, with a very regular, indeed a polynomial vector field f. So, even with a polynomial vector field, two paths which are close in p-variation for $p > 2$ do not determine close responses in the controlled equation (2). This is a first hint at the fact that it is natural to declare two paths with finite p-variation for $p > 2$ close to each other only if their difference has a small total p-variation and the areas that they determine are close.

It is interesting to note that stochastic differential equations stand just above this threshold $p = 2$. Indeed, almost surely Brownian paths have infinite two-variation and finite p-variation for every $p > 2$. The convergence in probability of the Riemann sums for stochastic integrals is, from the point of view of the Young integral, a miracle due to the very special stochastic structure of Brownian motion and the subtle cancellations it implies.

The quantity (3) has a sense as a stochastic integral when X is a Brownian motion: it is the Lévy area of X. It was conjectured by H. Föllmer that all SDEs driven by a Brownian motion can be solved at once, i.e. outside a single negligible set, once one has chosen a version of the Lévy area of this Brownian motion. The theory of rough paths gives a rigourous framework for this conjecture and proves it.

Chapter 2 explores the idea that (3) is the first of an infinite sequence of quantities which are canonically associated to a path with bounded variation in a Banach space. These quantities are the iterated integrals of the path. Let V be a Banach space. Let $X : [0,T] \longrightarrow V$ be a path with bounded variation. For every integer $n \geq 1$, and every (s,t) such that $0 \leq s \leq t \leq T$, the nth iterated integral of X over $[s,t]$ is the tensor of $V^{\otimes n}$ defined by

$$X^n_{s,t} = \int_{s<u_1<\ldots<u_n<t} dX_{u_1} \otimes \ldots \otimes dX_{u_n}. \tag{4}$$

When V is $\mathbb{R}^2$, (3) is just the antisymmetric part of $X^2_{0,T}$. The zeroth iterated integral is simply $X^0_{s,t} = 1 \in \mathbb{R} = V^{\otimes 0}$.

The importance of iterated integrals has been recognised by geometers, in particular K.T. Chen, a long time ago. In the context of controlled differential equations, their importance is most strikingly illustrated by the case of linear equations. Linear equations are those of the form (2) where the vector field f depends linearly on Y. When the control X has bounded variation, the resolution of the equation by Picard iteration leads to an expression of the solution Y as the sum of an infinite series of the form

$$Y_t = \left(\sum_{n=0}^{\infty} f^n X^n_{0,t} \right) Y_0, \tag{5}$$

where f^n is an operator depending on f and n and whose norm grows at most geometrically with n. It is not hard to check that the norm of the iterated integral $X^n_{0,t}$ of X decays like $\frac{1}{n!}$. Thus, the series (5) converges extremely fast. In typical numerical applications, a dozen of terms of the series suffice to provide an excellent approximation of the solution. What is even better is the following: once a dozen of iterated integrals of X have been stored on a computer, i.e. around d^{12} numbers if d is the dimension of V, it is possible to solve numerically very accurately any linear differential equation controlled by X with very little extra computation. The numerical error can be bounded by a simple function of the norm of the vector field.

The iterated integrals of X over an interval $[s,t]$ are thus extremely efficient statistics of X, in the sense that they determine very accurately the response of any linear system driven by X. It is in fact possible to understand exactly what it means geometrically for two paths with bounded variation and with the same origin to have the same iterated integrals: it means that they differ by a tree-like path. We state this result precisely in the notes, but do not include its proof.

However satisfying the results above are: one is missing something essential about iterated integrals until one considers them all as a single object. This object is called the *signature of the path*. More precisely, with the notation used in (4), the signature of X over the interval $[s,t]$ is the infinite sequence in $\mathbb{R} \oplus V \oplus V^{\otimes 2} \oplus \ldots$ defined as follows:

$$S(X)_{s,t} = (1, X^1_{s,t}, X^2_{s,t}, \ldots). \tag{6}$$

The infinite sequence space $\oplus_{n \geq 0} V^{\otimes n}$ is called the *extended tensor algebra* of V and it is denoted by $T((V))$. It is indeed an algebra for the multiplication induced by the tensor product. The reader not familiar with this kind of algebraic structures should keep the following dictionary in mind. Assume that V has finite dimension d and choose a basis $(v_1, \ldots, v_d)$ of V. Then it is a tautology that $V = V^{\otimes 1}$ is isomorphic to the space of homogeneous polynomials of degree 1 in the variables $X_1, \ldots, X_d$. It turns out that for every $n \geq 0$, $V^{\otimes n}$ is isomorphic to the space of homogeneous polynomials of degree n in the non-commuting variables $X_1, \ldots, X_d$. For instance, if $d = 2$, then a basis of $V^{\otimes 2}$ is $(X_1^2, X_1X_2, X_2X_1, X_2^2)$. Finally, $T((V))$ is isomorphic to the space of all formal power series in d non-commuting variables, not only as a vector space, but also as an algebra: the product of tensors corresponds exactly to the product of non-commuting polynomials.

The fundamental property of the signature is the following: if (s, u, t) are such that $0 \leq s \leq u \leq t \leq T$, then

$$S(X)_{s,t} = S(X)_{s,u} \otimes S(X)_{u,t}. \tag{7}$$

This multiplicativity property, although it encodes infinitely many relations between iterated integrals of X, can be proved in a very elementary way. However, the following abstract and informal point of view gives an interesting insight on the signature of a path. Among all differential equations that X can control, there is one which is more important than the others and in a certain sense universal. It is the following:

$$dS_t = S_t \otimes dX_t, \quad S_0 = 1, \quad S : [0,T] \longrightarrow T((V)). \tag{8}$$

The solution to (8) is nothing but the signature of X: $S_t = S(X)_{0,t}$. This suggests that the signature of a path should be thought of as a kind of universal non-commutative exponential of this path. Moreover, we deduce from (8) that the two sides of (7), which satisfy the same differential equation with the same initial value, are equal.

Chapter 3 focuses on collections $(S_{s,t})_{0 \leq s \leq t \leq T}$ of elements of $T((V))$ which satisfy (7). Such collections are called *multiplicative functionals* and the point of the theory of rough paths is to take them as the fundamental objects driving

differential equations. Rough paths are multiplicative functionals which satisfy some regularity property related to p-variation.

Like the multiplicativity property, the regularity property is inspired by the study of the signature of a path. If X has bounded variation, then $X^1_{s,t} = X_t - X_s \sim |t-s|$ and $X^n_{s,t}$ is of the order of $|t-s|^n$. If X has only finite p-variation for some $p \in (1,2)$, then it is still possible to define its iterated integrals as Young integrals, and we expect $X^n_{s,t}$ to be of the order of $|t-s|^{\frac{n}{p}}$.

Let $\triangle_T$ denote the set of pairs (s,t) such that $0 \leq s \leq t \leq T$. A multiplicative functional of degree n in V is a continuous mapping $X : \triangle_T \longrightarrow T^{(n)}(V) = \bigoplus_{i=0}^n V^{\otimes i}$. For each $(s,t) \in \triangle_T$, $X_{s,t}$ is thus a collection of $n+1$ tensors $(1, X^1_{s,t}, X^2_{s,t}, \ldots, X^n_{s,t})$. Let $p \geq 1$ be a number. A multiplicative functional X is said to have finite p-variation if

$$\sup_{0\leq i\leq n} \sup_{\mathcal{D}} \sum |X^i_{t_k,t_{k+1}}|^{\frac{p}{i}} < +\infty. \tag{9}$$

The first fundamental result of the theory expresses a deep connection between the multiplicativity property (7) and the finiteness of the p-variation (9). It states that a multiplicative functional with finite p-variation is determined by its truncature at level $\lfloor p \rfloor$. More precisely, if X and Y are multiplicative functionals of degree $n \geq \lfloor p \rfloor$ with finite p-variation and $X^i_{s,t} = Y^i_{s,t}$ for all $(s,t) \in \triangle_T$ and $i = 0, \ldots, \lfloor p \rfloor$, then $X = Y$. Conversely, any multiplicative functional of degree m with finite p-variation can be extended to a multiplicative functional of arbitrarily high degree with finite p-variation, provided $\lfloor p \rfloor \leq m$.

A p-rough path is then defined to be a multiplicative functional of finite p-variation and degree $\lfloor p \rfloor$.

Chapters 4 and 5 give a meaning to differential equations driven by rough paths and present a proof of the main theorem of the theory, named Universal Limit Theorem by P. Malliavin, which asserts that, provided f is smooth enough, (2) admits a unique solution when X is a p-rough path. The solution Y is then itself a p-rough path.

Let us conclude this introduction by explaining the part of the Universal Limit Theorem which can be stated without referring to rough paths, i.e. let us describe the metrics on the space of paths with bounded variation with respect to which the Itô map is uniformly continuous. Choose $p \geq 1$. Consider X and $\widetilde{X}$, both with bounded variation. For all (s,t) and all $n \geq 0$, let $X^n_{s,t}$ and $\widetilde{X}^n_{s,t}$ denote their iterated integrals of order n over $[s,t]$. Then the distance in the p-variation metric between X and $\widetilde{X}$ is defined by

$$d_p(X, \tilde{X}) = \sup_{0\leq i\leq \lfloor p \rfloor} \sup_{\mathcal{D}} \left[\sum |X^i_{t_k,t_{k+1}} - \widetilde{X}^i_{t_k,t_{k+1}}|^{\frac{p}{i}} \right]^{\frac{1}{p}}. \tag{10}$$

Now, provided the vector field f is of class $C^{\lfloor p \rfloor + \varepsilon}$, the Itô map I_f is uniformly continuous with respect to the distance d_p.

It has been a great pleasure for us to write these notes, not the least thanks to the countless hours of discussions we have had with Terry Lyons during their preparation. We would like to thank him warmly for his kind patience and communicative enthusiasm. Our greatest hope is that some of this enthusiasm, together with some of our own fascination for this theory, will permeate through these notes to the reader.

Contents