ESAIM: PS 16 (2012) 453–478
DOI: 10.1051/ps/2011107

ESAIM: Probability and Statistics
www.esaim-ps.org

Special Issue

STOCHASTIC TAYLOR EXPANSIONS AND HEAT KERNEL ASYMPTOTICS*

Fabrice Baudoin[1]

Abstract. These notes focus on the applications of the stochastic Taylor expansion of solutions of stochastic differential equations to the study of heat kernels in small times. As an illustration of these methods we provide a new heat kernel proof of the Chern–Gauss–Bonnet theorem.

Mathematics Subject Classification. 60H30, 58J20.

Received July 15, 2009. Revised April 27, 2010.

1. Introduction

The purpose of these notes is to provide to the reader an introduction to the theory of stochastic Taylor expansions with a view toward the study of heat kernels. They correspond to a five hours course given at a Spring school in June 2009.

In the first Section, we remind some basic facts about stochastic differential equations and introduce the language of vector fields. In the second Section, which is the heart of this course, we study stochastic Taylor expansions by means of the so-called formal Chen series. Parts of this section may be found in my book [4], but the proofs given in these notes are different and maybe more intuitive. In the third Section, we focus on the applications of study stochastic Taylor expansions to the study of the asymptotics in small times of heat kernels associated with elliptic diffusion operators. In the fourth Section, we extend the results of the third Section, to study heat kernels on vector bundles and provide a new proof of the Chern–Gauss–Bonnet theorem.

2. Stochastic differential equations: the language of vector fields

In this section we remind some preliminary results and definitions that will be used throughout the text. We focus on the connection between parabolic linear diffusion equations and stochastic differential equations and introduce the language of vector fields which is the most convenient when dealing with applications to geometry.

For further reading on the connection between diffusion equations and stochastic differential equations we refer to the book by Stroock and Varadhan [41], where the proofs of the below cited results may be found. For further reading on vector fields we refer to the Chapter 1 of [44] and for more explanations on the use of the language of vector fields for stochastic differential equations, we refer to the Chapter 1 of [26].

Keywords and phrases. Stochastic Taylor expansions, Index theorems.

* *Supported in part by NSF Grant DMS 0907326.*

[1] Department of Mathematics Purdue University, 504 Northwestern, Avenu West Lafayette, Indiana, USA.
fbaudoin@math.purdue.edu

Article published by EDP Sciences

© EDP Sciences, SMAI 2012

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

2.1. Heat kernels

A diffusion operator L on $\mathbb{R}^n$ is a second order differential operator that can be written

$$L = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x)\frac{\partial}{\partial x_i},$$

where b_i and a_{ij} are continuous functions on $\mathbb{R}^n$ such that for every $x \in \mathbb{R}^n$, the matrix $(a_{ij}(x))_{1\leq i,j\leq n}$ is symmetric and positive.

Associated to L, we may consider the following diffusion equation

$$\frac{\partial \Phi}{\partial t} = L\Phi, \quad \Phi(0,x) = f(x).$$

The function $\Phi : [0,+\infty) \times \mathbb{R}^n \to \mathbb{R}$ is the unknown of the equation and the function $f : \mathbb{R}^n \to \mathbb{R}$ is the initial datum.

Under mild conditions on the coefficients b_i and a_{ij} (for instance C^∞ bounded), it is well known that the above equation has one and only one solution Φ. It is usual to use the notation

$$\Phi(t,x) = \mathbf{P}_t f(x)$$

to stress how the solution Φ depends on the function f. The family of operators $(\mathbf{P}_t)_{t\geq 0}$ (acting on a convenient space of initial data[2]) is called the heat semigroup associated to the diffusion operator L. The terminology semigroups stems from the following easily checked property

$$\mathbf{P}_{t+s} = \mathbf{P}_t \mathbf{P}_s.$$

The diffusion operator L is said to be elliptic at a point $x_0 \in \mathbb{R}^n$ if the matrix $(a_{ij}(x_0))_{1\leq i,j\leq n}$ is invertible. We shall simply say that L is elliptic if it is elliptic at any point.

If L is an elliptic diffusion operator, the heat semigroup associated to it admits the following integral representation

$$\mathbf{P}_t f(x) = \int_{\mathbb{R}^n} p_t(x,y) f(y) \mathrm{d}y, \quad t > 0,$$

where $p : (0,+\infty) \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a smooth function that is called the heat kernel associated to L.

2.2. Stochastic differential equations and diffusion equations

Stochastic differential equations provide a powerful tool to study diffusion equations and associated heat kernels. Let us briefly recall below the main connection between these two types of equations.

Let

$$L = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x)\frac{\partial}{\partial x_i},$$

be a diffusion operator on $\mathbb{R}^n$. Since the matrix a is symmetric and positive, it admits a square root, that is, there exists a symmetric and positive matrix σ such that

$$\sigma^2 = a.$$

Let us consider a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathbb{P})$ which satisfies the usual conditions and on which is defined a n-dimensional Brownian motion $(B_t)_{t\geq 0}$.

[2] In probability theory, it is common to work with the space of bounded Borel functions.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

The following theorem is well known:

Theorem 2.1. Let us assume that b and σ are smooth, and that their derivatives of any order are bounded. Then, for every $x_0 \in \mathbb{R}^n$, there exists a unique and adapted process $(X_t^{x_0})_{t\geq 0}$ such that for $t \geq 0$

$$X_t^{x_0} = x_0 + \int_0^t b(X_s^{x_0})\mathrm{d}s + \int_0^t \sigma(X_s^{x_0})\mathrm{d}B_s. \tag{2.1}$$

Moreover, if $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth and compactly supported function, then the function

$$\phi(t,x) = \mathbb{E}\left(f(X_t^x)\right) = \mathbf{P}_t f(x)$$

is the unique bounded solution of the diffusion equation

$$\frac{\partial \phi}{\partial t}(t,x) = L\phi(t,x), \quad \phi(0,x) = f(x).$$

2.3. The language of vector fields

For geometric purposes, it is often more useful to use Stratonovitch integrals and the language of vector fields in the study of stochastic differential equations.

Let $\mathcal{O} \subset \mathbb{R}^n$ be a non empty open set. A smooth vector field V on $\mathcal{O}$ is a smooth map

$$V : \mathcal{O} \to \mathbb{R}^n$$

$$x \to (v_1(x), ..., v_n(x)).$$

A vector field V defines a differential operator acting on the smooth functions $f : \mathcal{O} \to \mathbb{R}$ as follows:

$$(Vf)(x) = \sum_{i=1}^n v_i(x)\frac{\partial f}{\partial x_i}.$$

We note that V is a derivation, that is a map on $\mathcal{C}^\infty(\mathcal{O},\mathbb{R})$, linear over $\mathbb{R}$, satisfying for $f, g \in \mathcal{C}^\infty(\mathcal{O},\mathbb{R})$,

$$V(fg) = (Vf)g + f(Vg).$$

Conversely, it may be shown that any derivation on $\mathcal{C}^\infty(\mathcal{O},\mathbb{R})$ is a vector field.

With these notations, it is readily checked that if $V_0, V_1, \ldots, V_d$ are smooth vector fields on $\mathbb{R}^n$, then the second order differential operator

$$L = V_0 + \frac{1}{2}\sum_{i=1}^d V_i^2$$

is a diffusion operator. Though it is always locally true, in general, a diffusion operator may not necessarily be globally written under the above form. If this is the case, the operator is said to be a Hörmander's type operator. We may observe as an easy exercise that the operator is elliptic if and only if for every $x \in \mathbb{R}^n$, the linear space generated by the vectors $V_1(x), \ldots, V_d(x)$ is equal to $\mathbb{R}^n$.

To associate with L a stochastic differential equation it is more convenient to use Stratonovitch integration than Itô's. Let us recall that if $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are two continuous semimartingales, the Stratonovitch integral of Y against X may be defined by

$$\int Y \circ \mathrm{d}X = \int Y\mathrm{d}X + \frac{1}{2}\langle X, Y\rangle,$$

where $\langle X, Y\rangle$ denotes the quadratic covariation between X and Y.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

With this language, we have the following translation of Theorem 2.1:

Theorem 2.2. Let $(B_t)_{t\geq 0}$ be a d-dimensional Brownian motion. Let us assume that $V_0, V_1, \dots, V_d$ are smooth vector fields on $\mathbb{R}^n$, and that their derivatives of any order are bounded.

Then, for every $x_0 \in \mathbb{R}^n$, there exists a unique and adapted process $(X_t^{x_0})_{t\geq 0}$ such that for $t \geq 0$

$$X_t^{x_0} = x_0 + \int_0^t V_0(X_s^{x_0})\mathrm{d}s + \sum_{i=1}^d \int_0^t V_i(X_s^{x_0}) \circ \mathrm{d}B_s^i. \tag{2.2}$$

Moreover, if $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth and compactly supported function, then the function

$$\phi(t,x) = \mathbb{E}\left(f(X_t^x)\right) = \mathbf{P}_t f(x)$$

is the unique bounded solution of the diffusion equation

$$\frac{\partial \phi}{\partial t}(t,x) = L\phi(t,x), \quad \phi(0,x) = f(x).$$

where

$$L = V_0 + \frac{1}{2}\sum_{i=1}^d V_i^2.$$

The main advantage of this language is the following simple change of variable formula which is nothing else but the celebrated Itô's formula:

Proposition 2.3. If $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth function and $(X_t^{x_0})_{t\geq 0}$ the solution of (2.2), then the following Itô's formula holds

$$f(X_t^{x_0}) = f(x_0) + \int_0^t V_0 f(X_s^{x_0})\mathrm{d}s + \sum_{i=1}^d \int_0^t V_i f(X_s^{x_0}) \circ \mathrm{d}B_s^i.$$

3. Stochastic Taylor expansions

Our goal is to study solutions of stochastic differential equations in small times. A powerful tool to do so is the stochastic Taylor expansion whose scheme is described in the first subsection below. After subtle algebraic manipulations of the stochastic Taylor expansion involving the so-called formal Chen series, it is possible to deduce an approximation in small times of the flow $(x \to X_t^x)_{t\geq 0}$ associated to the given stochastic differential equation. Let us mention here that the algebraic material presented lies at the heart of the rough paths theory of of Lyons [32]. More precisely considering the closure, with respect to a convenient topology, of the algebra of Chen series of absolutely continuous paths leads to the notion of geometric rough paths. So, the results of this section admits a counterpart in rough paths theory (See [22]).

3.1. Motivation

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^∞ bounded function and denote by $(X_t^{x_0})_{t\geq 0}$ the solution of (2.2) with initial condition $x \in \mathbb{R}^n$. First, by Itô's formula, we have

$$f(X_t^x) = f(x) + \sum_{i=0}^d \int_0^t (V_i f)(X_s^x) \circ \mathrm{d}B_s^i, \ t \geq 0,$$

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

where we use the notation $B_t^0 = t$. Now, a new application of Itô's formula to $V_i f(X_s^x)$ leads to

$$f(X_t^x) = f(x) + \sum_{i=0}^{d} (V_i f)(x) B_t^i + \sum_{i,j=0}^{d} \int_0^t \int_0^s (V_j V_i f)(X_u^x) \circ \mathrm{d}B_u^j \circ \mathrm{d}B_s^i.$$

We may iterate this process. For this, let us introduce the following notations:

1.
$$\Delta^k[0,t] = \{(t_1, ..., t_k) \in [0,t]^k, t_1 \leq ... \leq t_k\};$$

2. If $I = (i_1, ...i_k) \in \{0, ..., d\}^k$ is a word with length k,
$$\int_{\Delta^k[0,t]} \circ \mathrm{d}B^I = \int_{0 \leq t_1 \leq ... \leq t_k \leq t} \circ \mathrm{d}B_{t_1}^{i_1} \circ ... \circ \mathrm{d}B_{t_k}^{i_k},$$
and $n(I)$ is the number of 0's in I.

We can then continue the above procedure and get that for every $N \geq 1$

$$f(X_t^x) = f(x) + \sum_{k=1}^{N} \sum_{I \in \{0,...,d\}^k, k+n(I) \leq N} (V_{i_1} ... V_{i_k} f)(x) \int_{\Delta^k[0,t]} \circ \mathrm{d}B^I + \mathbf{R}_N(t, f, x),$$

for some remainder term $\mathbf{R}_N(t, f, x)$ which is easily computed, and shown to satisfy

$$\sup_{x \in \mathbb{R}^n} \sqrt{\mathbb{E}\left(\mathbf{R}_N(t, f, x)^2\right)} \leq C_N t^{\frac{N+1}{2}} \sup_{(i_1,...,i_k), k+n(I)=N+1 \text{ or } N+2} \|V_{i_1} \dots V_{i_k} f\|_\infty.$$

This shows that, in *small times*, the sum

$$f(x) + \sum_{k=1}^{N} \sum_{I \in \{0,...,d\}^k, k+n(I) \leq N} (V_{i_1} ... V_{i_k} f)(x) \int_{\Delta^k[0,t]} \circ \mathrm{d}B^I \tag{3.3}$$

is a more and more accurate approximation of $f(X_t^x)$ when $N \to +\infty$.

Remark 3.1. For further details on the above discussion, we refer to Ben Arous [10] and Kloeden–Platen [28]. Related discussions for the Taylor expansion of solutions of equations driven by fractional Brownian motions may be found in Baudoin–Coutin [6]. For the case of rough paths we refer to Inahama [27] and Friz–Victoir [21,22].

3.2. Chen series

Our goal is now transform the approximation given by the Taylor expansion (3.3) into an approximation of the stochastic flow associated to the equation (2.2). That is, at any order, we wish to construct an explicit random diffeomorphism Φ_t^N such that

$$(\Phi_t^N f)(x) := f\left(\Phi_t^N(x)\right) = f(x) + \sum_{k=1}^{N} \sum_{I \in \{0,...,d\}^k, k+n(I) \leq N} (V_{i_1} ... V_{i_k} f)(x) \int_{\Delta^k[0,t]} \circ \mathrm{d}B^I + \mathbf{R}_N^*(t, f, x),$$

for some remainder term $\mathbf{R}_N^*(t, f, x)$.

In order to do so, the main tool was introduced by Chen in [16] in his seminal paper of 1957. Chen considered formal Taylor series associated to paths (or currents) and proved that such series could be represented as the exponential of Lie series.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

We present in this subsection those results.

Let $\mathbb{R}[[X_0,...,X_d]]$ be the non commutative algebra over $\mathbb{R}$ of the formal series with $d+1$ indeterminates, that is the set of series

$$Y = y_0 + \sum_{k=1}^{+\infty} \sum_{I\in\{0,1,...,d\}^k} a_{i_1,...,i_k} X_{i_1}...X_{i_k}.$$

In what follows, we will denote by $\mathbb{R}_{\geq 0}$ the set $[0,+\infty)$.

Definition 3.2. If $x:\mathbb{R}_{\geq 0}\to\mathbb{R}^d$ is an absolutely continuous path, the Chen series of x is the formal series:

$$\mathfrak{S}(x)_t = 1 + \sum_{k=1}^{+\infty} \sum_{I\in\{0,1,...,d\}^k} \left(\int_{0\leq t_1\leq ...\leq t_k\leq t} \mathrm{d}x_{t_1}^{i_1}\dots\mathrm{d}x_{t_k}^{i_k}\right) X_{i_1}\dots X_{i_k}, \quad t\geq 0,$$

with the convention $x_t^0 = t$.

The exponential of $Y\in\mathbb{R}[[X_0,...,X_d]]$ is defined by

$$\exp(Y) = \sum_{k=0}^{+\infty} \frac{Y^k}{k!},$$

and the logarithm of Y by

$$\ln(Y) = \sum_{k=1}^{+\infty} \frac{(-1)^k}{k}(Y-1)^k.$$

The Chen–Strichartz formula that we will prove in this subsection, is an explicit formula for $\ln\mathfrak{S}(x)_t$.

Remark 3.3. As a preliminary, let us first try to understand a simple case: the commutative case.

We denote $\mathcal{S}_k$ the group of the permutations of the index set $\{1,...,k\}$ and if $\sigma\in\mathcal{S}_k$, we denote for a word $I=(i_1,...,i_k)$, $\sigma\cdot I$ the word $(i_{\sigma(1)},...,i_{\sigma(k)})$. If $X_0,X_1,...,X_d$ were commuting[3], we would have

$$\mathfrak{S}(x)_t = \mathbf{1} + \sum_{k=1}^{+\infty} \sum_{I=(i_1,...,i_k)} X_{i_1}...X_{i_k} \left(\frac{1}{k!}\sum_{\sigma\in\mathcal{S}_k}\int_{\Delta^k[0,t]} \mathrm{d}x^{\sigma\cdot I}\right).$$

Since

$$\sum_{\sigma\in\mathcal{S}_k}\int_{\Delta^k[0,t]} \mathrm{d}x^{\sigma\cdot I} = x_t^{i_1}...x_t^{i_k},$$

we get,

$$\mathfrak{S}(x)_t = \mathbf{1} + \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{I=(i_1,...,i_k)} X_{i_1}...X_{i_k} x_t^{i_1}...x_t^{i_k} = \exp\left(\sum_{i=0}^{d} X_i x_t^i\right).$$

We define the Lie bracket between two elements U and V of $\mathbb{R}[[X_0,...,X_d]]$ by

$$[U,V] = UV - VU.$$

Moreover, if $I=(i_1,...,i_k)\in\{0,...,d\}^k$ is a word, we denote by X_I the commutator defined by

$$X_I = [X_{i_1},[X_{i_2},...,[X_{i_{k-1}},X_{i_k}]...].$$

[3]Rigorously, this means that we work in $\mathbb{R}[[X_0,X_1,\dots,X_d]]/\mathcal{J}$ where $\mathcal{J}$ is the two-sided ideal generated by the relations $X_iX_j - X_jX_i = 0$.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

The universal Chen's theorem asserts that the Chen series of a path is the exponential of a Lie series.

Theorem 3.4 (Chen–Strichartz expansion theorem)**.** If $x : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$ is an absolutely continuous path, then

$$\mathfrak{S}(x)_t = \exp\left(\sum_{k\geq 1}\sum_{I\in\{0,1,\ldots,d\}^k} \Lambda_I(x)_t X_I\right),\ t \geq 0,$$

where for $k \geq 1,\ I \in \{0, 1, ..., d\}^k$:

- $\mathcal{S}_k$ is the set of the permutations of $\{0, ..., k\}$;
- If $\sigma \in \mathcal{S}_k$, $e(\sigma)$ is the cardinality of the set

$$\{j \in \{0, ..., k-1\}, \sigma(j) > \sigma(j+1)\};$$

- $$\Lambda_I(x)_t = \sum_{\sigma\in\mathcal{S}_k} \frac{(-1)^{e(\sigma)}}{k^2\binom{k-1}{e(\sigma)}} \int_{0\leq t_1\leq\ldots\leq t_k\leq t} \mathrm{d}x_{t_1}^{\sigma^{-1}(i_1)}\ldots\mathrm{d}x_{t_k}^{\sigma^{-1}(i_k)},\quad t \geq 0.$$

Remark 3.5. The first terms in the Chen–Strichartz formula are:

1. $$\sum_{I=(i_1)} \Lambda_I(x)_t X_I = \sum_{k=0}^{d} x_t^i X_i;$$
2. $$\sum_{I=(i_1,i_2)} \Lambda_I(x)_t X_I = \frac{1}{2}\sum_{0\leq i<j\leq d} [X_i, X_j]\int_0^t x_s^i \mathrm{d}x_s^j - x_s^j \mathrm{d}x_s^i.$$

We shall give the proof of this theorem in the case where the path x_t is piecewise affine that is

$$\mathrm{d}x_t = a_i \mathrm{d}t$$

on the interval $[t_i, t_{i+1})$ where $0 = t_0 \leq t_1 \leq \ldots \leq t_N = T$. Since any absolutely continuous path is limit of piecewise affine paths, we may then conclude by a limiting argument. The proof relies on several lemmas.

Lemma 3.6 (Chen's relations)**.** Let x_t be an absolutely continuous path. For any word $(i_1, ..., i_n) \in \{0, 1, ..., d\}^n$ and any $0 < s < t$,

$$\int_{\Delta^n[0,t]} \mathrm{d}x^{(i_1,\ldots,i_n)} = \sum_{k=0}^{n}\int_{\Delta^k[0,s]} \mathrm{d}x^{(i_1,\ldots,i_k)}\int_{\Delta^{n-k}[s,t]} \mathrm{d}x^{(i_{k+1},\ldots,i_n)},$$

where we used the following notations:

1. $$\int_{\Delta^k[s,t]} \mathrm{d}x^{(i_1,\ldots,i_k)} = \int_{s\leq t_1\leq\ldots\leq t_k\leq t} \mathrm{d}x_{t_1}^{i_1}\ldots\mathrm{d}x_{t_k}^{i_k};$$
2. if I is a word with length 0, then $\int_{\Delta^0[0,t]} \circ \mathrm{d}x^I = 1$.

Proof. It follows readily by induction on n by noticing that

$$\int_{\Delta^n[0,t]} \mathrm{d}x^{(i_1,\ldots,i_n)} = \int_0^t \left(\int_{\Delta^{n-1}[0,t_n]} \mathrm{d}x^{(i_1,\ldots,i_{n-1})}\right)\mathrm{d}x_{t_n}^{i_n}.$$

□

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

The previous lemma implies the following flow property for the signature:

Lemma 3.7. Let x_t be an absolutely continuous path. For $0 < s < t$,

$$\mathfrak{S}(x)_t = \mathfrak{S}(x)_s \left(\mathbf{1} + \sum_{k=1}^{+\infty} \sum_{I=(i_1,\dots i_k)} X_{i_1} \dots X_{i_k} \int_{\Delta^k[s,t]} \mathrm{d}x^I \right).$$

Proof. We have, thanks to the previous lemma,

$$\begin{aligned} &\mathfrak{S}(x)_s \left(\mathbf{1} + \sum_{k=1}^{+\infty} \sum_{I} X_{i_1} \dots X_{i_k} \int_{\Delta^k[s,t]} \mathrm{d}x^I \right) = \mathbf{1} \\ &+ \sum_{k,k'=1}^{+\infty} \sum_{I,I'} X_{i_1} \dots X_{i_k} X_{i'_1} \dots X_{i'_{k'}} \int_{\Delta^k[s,t]} \mathrm{d}x^I \int_{\Delta^{k'}[0,s]} \mathrm{d}x^{I'} = \mathbf{1} \\ &+ \sum_{k=1}^{+\infty} \sum_{I} X_{i_1} \dots X_{i_k} \int_{\Delta^k[0,t]} \mathrm{d}x^I = \mathfrak{S}(x)_t. \end{aligned}$$

□

□

With this in hands, we may now come back to the proof of the Chen–Strichartz expansion theorem in the case where x_t is piecewise affine. By using inductively the previous proposition, we obtain

$$\mathfrak{S}(x)_T = \prod_{n=0}^{N-1} \left(\mathbf{1} + \sum_{k=1}^{+\infty} \sum_{I=(i_1,\dots i_k)} X_{i_1} \dots X_{i_k} \int_{\Delta^k[t_n,t_{n+1}]} \mathrm{d}x^I \right).$$

Since, on $[t_n, t_{n+1})$,

$$\mathrm{d}x_t = a_n \mathrm{d}t,$$

we have

$$\int_{\Delta^k[t_n,t_{n+1}]} \mathrm{d}x^I = a_n^{i_1} \dots a_n^{i_k} \int_{\Delta^k[t_n,t_{n+1}]} \mathrm{d}t_{i_1} \dots \mathrm{d}t_{i_k} = a_n^{i_1} \dots a_n^{i_k} \frac{(t_{n+1}-t_n)^k}{k!}.$$

Therefore

$$\begin{aligned} \mathfrak{S}(x)_T &= \prod_{n=0}^{N-1} \left(\mathbf{1} + \sum_{k=1}^{+\infty} \sum_{I=(i_1,\dots i_k)} X_{i_1} \dots X_{i_k} a_n^{i_1} \dots a_n^{i_k} \frac{(t_{n+1}-t_n)^k}{k!} \right) \\ &= \prod_{n=0}^{N-1} \exp\left((t_{n+1}-t_n) \sum_{i=0}^{d} a_n^i X_i \right). \end{aligned}$$

We now use the Baker–Campbell–Hausdorff–Dynkin formula (see Dynkin [18] and Strichartz [42]):

Proposition 3.8. (Baker–Campbell–Hausdorff–Dynkin formula). If $y_1, \dots, y_N \in \mathbb{R}^{d+1}$ then,

$$\prod_{n=1}^{N} \exp\left(\sum_{i=0}^{d} y_n^i X_i \right) = \exp\left(\sum_{k\geq 1} \sum_{I\in\{0,1,\dots,d\}^k} \beta_I(y_1,\dots,y_N) X_I \right),$$

where for $k \geq 1$, $I \in \{0,1,...,d\}^k$:

$$\beta_I(y_1,\dots,y_N) = \sum_{\sigma\in\mathcal{S}_k} \sum_{0=j_0\leq j_1\leq\dots\leq j_{N-1}\leq k} \frac{(-1)^{e(\sigma)}}{j_1!\dots j_{N-1}! k^2 \binom{k-1}{e(\sigma)}} \prod_{\nu=1}^{N} y_\nu^{\sigma^{-1}(i_{j_{\nu-1}+1})} \dots y_\nu^{\sigma^{-1}(i_{j_\nu})}.$$

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

We get therefore:

$$\mathfrak{S}(x)_T = \exp\left(\sum_{k\geq 1}\sum_{I\in\{0,1,\dots,d\}^k} \beta_I(t_1 a_0,\dots,(t_N - t_{N-1})a_{N-1})X_I\right).$$

It is finally an easy exercise to check, by using the Chen relations, that:

$$\beta_I(t_1 a_0,\dots,(t_N - t_{N-1})a_{N-1}) = \sum_{\sigma\in\mathcal{S}_k} \frac{(-1)^{e(\sigma)}}{k^2\binom{k-1}{e(\sigma)}} \int_{0\leq t_1\leq\dots\leq t_k\leq t} \mathrm{d}x_{t_1}^{\sigma^{-1}(i_1)}\dots\mathrm{d}x_{t_k}^{\sigma^{-1}(i_k)}, \quad t\geq 0.$$

Remark 3.9. The seminal result of Chen [16] asserted that $\ln \mathfrak{S}(x)_T$ was a Lie series. The coefficients of this expansion were computed by Strichartz [42].

3.3. Brownian Chen series

Chen's theorem can actually be extended to Brownian paths (see Baudoin [4], Ben Arous [10], Castell [15], Fliess [19]) and even to rough paths (see Lyons [32], Friz–Victoir [22]).

Definition 3.10. If $(B_t)_{t\geq 0}$ is a d-dimensional Brownian motion, the Chen series of B is the formal series:

$$\mathfrak{S}(B)_t = 1 + \sum_{k=1}^{+\infty}\sum_{I\in\{0,1,\dots,d\}^k}\left(\int_{0\leq t_1\leq\dots\leq t_k\leq t} \circ\mathrm{d}B_{t_1}^{i_1}\dots\circ\,\mathrm{d}B_{t_k}^{i_k}\right)X_{i_1}\dots X_{i_k}, \quad t\geq 0,$$

with the convention $B_t^0 = t$, and $\circ$ denotes Stratonovitch integral.

Theorem 3.11. If $(B_t)_{t\geq 0}$ is a d-dimensional Brownian motion, then

$$\mathfrak{S}(B)_t = \exp\left(\sum_{k\geq 1}\sum_{I\in\{0,1,\dots,d\}^k} \Lambda_I(B)_t X_I\right),\ t\geq 0,$$

where for $k\geq 1,\ I\in\{0,1,\dots,d\}^k$,

$$\Lambda_I(B)_t = \sum_{\sigma\in\mathcal{S}_k} \frac{(-1)^{e(\sigma)}}{k^2\binom{k-1}{e(\sigma)}} \int_{0\leq t_1\leq\dots\leq t_k\leq t} \circ\mathrm{d}B_{t_1}^{\sigma^{-1}(i_1)}\dots\circ\,\mathrm{d}B_{t_k}^{\sigma^{-1}(i_k)}, \quad t\geq 0.$$

If

$$Y = y_0 + \sum_{k=1}^{+\infty}\sum_{I\in\{0,1,\dots,d\}^k} a_{i_1,\dots,i_k}X_{i_1}\dots X_{i_k}.$$

is a random series, that is if the coefficients are real random variables defined on a probability space, we will denote

$$\mathbb{E}(Y) = \mathbb{E}(y_0) + \sum_{k=1}^{+\infty}\sum_{I\in\{0,1,\dots,d\}^k} \mathbb{E}(a_{i_1,\dots,i_k})X_{i_1}\dots X_{i_k}.$$

as soon as the coefficients of Y are integrable, where $\mathbb{E}$ stands for the expectation.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

The following theorem gives the expectation (see Baudoin [4], Lyons–Victoir [33]) of the Brownian Chen series:

Theorem 3.12. For $t \geq 0$,

$$\mathbb{E}\left(\mathfrak{S}(B)_t\right) = \exp\left(t\left(X_0 + \frac{1}{2}\sum_{i=1}^{d} X_i^2\right)\right).$$

Proof. An easy computation shows that if $\mathcal{I}_n$ is the set of words with length n obtained by all the possible concatenations of the words

$$\{0\}, \{(i,i)\}, \quad i \in \{1, ..., d\},$$

1. If $I \notin \mathcal{I}_n$ then
$$\mathbb{E}\left(\int_{\Delta^n[0,t]} \circ \mathrm{d}B^I\right) = 0;$$
2. If $I \in \mathcal{I}_n$ then
$$\mathbb{E}\left(\int_{\Delta^n[0,t]} \circ \mathrm{d}B^I\right) = \frac{t^{\frac{n+n(I)}{2}}}{2^{\frac{n-n(I)}{2}}\left(\frac{n+n(I)}{2}\right)!},$$
where $n(I)$ is the number of 0 in I (observe that since $I \in \mathcal{I}_n$, n and $n(I)$ necessarily have the same parity).

Therefore,

$$\mathbb{E}\left(\mathfrak{S}(B)_t\right) = 1 + \sum_{k=1}^{+\infty}\sum_{I\in\mathcal{I}_k} \frac{t^{\frac{k+n(I)}{2}}}{2^{\frac{k-n(I)}{2}}\left(\frac{k+n(I)}{2}\right)!} X_{i_1}...X_{i_k}$$

□

□

3.4. Exponential of a vector field

With these new tools in hands we may now come back to our primary purpose, which was to transform the stochastic Taylor expansion into an approximation of the stochastic flow associated to a stochastic differential equation.

In order to use the previous formalism, we first need to understand what is the exponential of a vector field. Let $\mathcal{O} \subset \mathbb{R}^n$ be a non empty open set and V be a smooth vector field on $\mathcal{O}$. It is a basic result in the theory of ordinary differential equations that if $K \subset \mathcal{O}$ is compact, there exist $\varepsilon > 0$ and a smooth mapping

$$\Phi : (-\varepsilon, \varepsilon) \times K \to \mathcal{O},$$

such that for $x \in K$ and $-\varepsilon < t < \varepsilon$,

$$\frac{\partial \Phi}{\partial t}(t,x) = X(\Phi(t,x)), \ \Phi(0,x) = x.$$

Furthermore, if $y : (-\eta, \eta) \to \mathbb{R}^n$ is a C^1 path such that for $-\eta < t < \eta$, $y'(t) = X(y(t))$, then $y(t) = \Phi(t, y(0))$ for $-\min(\eta, \varepsilon) < t < \min(\eta, \varepsilon)$. From this characterization of Φ it is easily seen that for $x \in K$ and $t_1, t_2 \in \mathbb{R}$ such that $\mid t_1 \mid + \mid t_2 \mid < \varepsilon$,

$$\Phi(t_1, \Phi(t_2, x)) = \Phi(t_1 + t_2, x).$$

Because of this last property, the solution mapping $t \to \Phi(t,x)$ is called the exponential mapping, and we denote $\Phi(t,x) = \mathrm{e}^{tV}(x)$. It always exists if $\mid t \mid$ is sufficiently small. If e^{tV} can be defined for any $t \in \mathbb{R}$, then the vector field is said to be complete. For instance if $\mathcal{O} = \mathbb{R}^n$ and if V is C^∞-bounded then the vector field V is complete.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

3.5. Lie bracket of vector fields

We also need to introduce the notion of Lie bracket between two vector fields.

We have already stressed that a vector field V may be seen derivation, that is a map on $\mathcal{C}^\infty(\mathcal{O},\mathbb{R})$, linear over $\mathbb{R}$, satisfying for $f,g\in\mathcal{C}^\infty(\mathcal{O},\mathbb{R})$,

$$V(fg)=(Vf)g+f(Vg).$$

Also, conversely, any derivation on $\mathcal{C}^\infty(\mathcal{O},\mathbb{R})$ is a vector field. If V' is another smooth vector field on $\mathcal{O}$, then it is easily seen that the operator $VV'-V'V$ is a derivation. It therefore defines a smooth vector field on $\mathcal{O}$ which is called the Lie bracket of V and V' and denoted $[V,V']$. A straightforward computation shows that for $x\in\mathcal{O}$,

$$[V,V'](x)=\sum_{i=1}^{n}\left(\sum_{j=1}^{n}v_j(x)\frac{\partial v_i'}{\partial x_j}(x)-v_j'(x)\frac{\partial v_i}{\partial x_j}(x)\right)\frac{\partial}{\partial x_i}\cdot$$

Observe that the Lie bracket satisfies obviously $[V,V']=-[V',V]$ and the so-called Jacobi identity, that is:

$$[V,[V',V'']]+[V',[V'',V]]+[V'',[V,V']]=0.$$

3.6. Castell's approximation theorem

Combining the stochastic Taylor expansion with the Chen–Strichartz formula leads finally to the following result of approximation of stochastic flows which is due to Castell [15]:

Theorem 3.13 (Castell approximation theorem)**.** Let $(B_t)_{t\geq 0}$ be a d-dimensional Brownian motion. Let us assume that $V_0,V_1,\ldots,V_d$ are C^∞ bounded vector fields on $\mathbb{R}^n$, Then, for the solution $(X_t^{x_0})_{t\geq 0}$ of the following stochastic differential equation

$$X_t^{x_0}=x_0+\int_0^t V_0(X_s^{x_0})\mathrm{d}s+\sum_{i=1}^{d}\int_0^t V_i(X_s^{x_0})\circ\mathrm{d}B_s^i, \tag{3.4}$$

we have for every $N\geq 1$,

$$X_t^{x_0}=\exp\left(\sum_{k=1}^{N}\sum_{I\in\{0,1,\ldots,d\}^k,k+n(I)\leq N}A_I(B)_tX_I\right)(x_0)+t^{\frac{N+1}{2}}\mathbf{R}_N(t),$$

where $n(I)$ denotes the number of 0's in the word I and where the remainder term $\mathbf{R}_N(t)$ is bounded in probability when $t\to 0$. More precisely, $\exists\ \alpha,c>0$ such that $\forall A>c$,

$$\lim_{t\to 0}\mathbb{P}\left(\sup_{0\leq s\leq t}s^{\frac{N+1}{2}}\mid\mathbf{R}_N(s)\mid\geq At^{\frac{N+1}{2}}\right)\leq\exp\left(-\frac{A^\alpha}{c}\right)\cdot$$

Remark 3.14. Kunita, in [29], proved an extension of this theorem to stochastic flows generated by stochastic differential equations driven by Lévy processes. Also, the result may be stated in the context of rough paths theory (see [22]).

4. Approximation in small times of solutions of diffusion equations

In this section, we now turn to applications of the previous results on the stochastic Taylor expansion to the study in small times of parabolic diffusion equations.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

We consider the following linear partial differential equation

$$\frac{\partial \Phi}{\partial t} = \mathcal{L}\Phi, \quad \Phi(0,x) = f(x), \tag{4.5}$$

where $\mathcal{L}$ is a diffusion operator on $\mathbb{R}^n$ that can be written

$$\mathcal{L} = V_0 + \frac{1}{2}\sum_{i=1}^{d} V_i^2,$$

the V_i's being smooth and compactly supported[4] vector fields on $\mathbb{R}^n$. It is known that the solution of (4.5) can be written

$$\Phi(t,x) = (\mathrm{e}^{t\mathcal{L}} f)(x) = \mathbf{P}_t f(x).$$

If $I \in \{0,1,...,d\}^k$ is a word, we denote as before

$$V_I = [V_{i_1}, [V_{i_2}, ..., [V_{i_{k-1}}, V_{i_k}]...].$$

and

$$d(I) = k + n(I),$$

where $n(I)$ is the number of 0's in the word I.

For $N \geq 1$, let us consider

$$\mathbf{P}_t^N = \mathbb{E}\left(\exp\left(\sum_{I, d(I)\leq N} \Lambda_I(B)_t V_I\right)\right).$$

For instance

$$\mathbf{P}_t^1 = \mathbb{E}\left(\exp\left(\sum_{i=1}^{d} B_t^i V_i\right)\right),$$

and

$$\mathbf{P}_t^2 = \mathbb{E}\left(\exp\left(\sum_{i=0}^{d} B_t^i V_i + \frac{1}{2}\sum_{1\leq i<j\leq d}\int_0^t B_s^i \mathrm{d}B_s^j - B_s^j \mathrm{d}B_s^i [V_i, V_j]\right)\right).$$

The meaning of this last notation is the following. If f is a smooth and bounded function, then $(\mathbf{P}_t^N f)(x) = \mathbb{E}(\Psi(1,x))$, where $\Psi(\tau,x)$ is the solution of the first order partial differential equation with random coefficients:

$$\frac{\partial \Psi}{\partial \tau}(\tau,x) = \sum_{I, d(I)\leq N} \Lambda_I(B)_t (V_I \Psi)(\tau,x), \quad \Psi(0,x) = f(x).$$

Finally, let us consider the following family of norms: If f is a C^∞ bounded function, then for $k \geq 0$,

$$\| f \|_k = \sup_{0\leq l\leq k} \sup_{0\leq i_1,...,i_l\leq d} \sup_{x\in\mathbb{R}^n} \| V_{i_1} \dots V_{i_l} f(x) \| .$$

Theorem 4.1. Let $N \geq 1$ and $k \geq 0$. If f is a C^∞ bounded function, then

$$\| \mathbf{P}_t f - \mathbf{P}_t^N f \|_k = O\left(t^{\frac{N+1}{2}}\right), \quad t \to 0.$$

[4]This assumption will not be restrictive for us because we shall eventually be interested in local results.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Proof. First, by using the scaling property of Brownian motion and expanding out the exponential with Taylor formula we obtain

$$\exp\left(\sum_{I,d(I)\leq N} \Lambda_I(B)_t V_I\right) f = \left(\sum_{k=0}^{N} \frac{1}{k!}\left(\sum_{I,d(I)\leq N} \Lambda_I(B)_t V_I\right)^k\right) f + t^{\frac{N+1}{2}} \mathbf{R}_N^1(t),$$

where the remainder term $\mathbf{R}_N^1(t)$ is such that $\mathbb{E}\left(\| \mathbf{R}_N^1(t) \|_k\right)$ is bounded when $t \to 0$. We now observe that, due to Theorem 3.11, the rearrangement of terms in the previous formula gives

$$\left(\sum_{k=0}^{N} \frac{1}{k!}\left(\sum_{I,d(I)\leq N} \Lambda_I(B)_t V_I\right)^k\right) f = f + \sum_{I,d(I)\leq N} \int_{\Delta^{|I|}[0,t]} \circ \mathrm{d}B^I V_{i_1} ... V_{i_{|I|}} f + t^{\frac{N+1}{2}} \mathbf{R}_N^2(t),$$

where $\mathbb{E}\left(\| \mathbf{R}_N^2(t) \|_k\right)$ is bounded when $t \to 0$. Therefore

$$\exp\left(\sum_{I,d(I)\leq N} \Lambda_I(B)_t V_I\right) f = f + \sum_{I,d(I)\leq N} \int_{\Delta^{|I|}[0,t]} \circ \mathrm{d}B^I V_{i_1} ... V_{i_{|I|}} f + t^{\frac{N+1}{2}} \mathbf{R}_N^3(t),$$

and

$$\mathbf{P}_t^N f = f + \sum_{I,d(I)\leq N} \mathbb{E}\left(\int_{\Delta^{|I|}[0,t]} \circ \mathrm{d}B^I\right) V_{i_1} ... V_{i_{|I|}} f + t^{\frac{N+1}{2}} \mathbb{E}\left(\mathbf{R}_N^3(t)\right),$$

where $\mathbb{E}\left(\| \mathbf{R}_N^3(t) \|_k\right)$ is bounded when $t \to 0$. We now recall (see the Proof of Thm. 3.12) that if $\mathcal{I}_n$ is the set of words with length n obtained by all the possible concatenations of the words

$$\{0\}, \{(i,i)\}, \quad i \in \{1, ..., d\},$$

1. If $I \notin \mathcal{I}_n$ then
$$\mathbb{E}\left(\int_{\Delta^n[0,t]} \circ \mathrm{d}B^I\right) = 0.$$
2. If $I \in \mathcal{I}_n$ then
$$\mathbb{E}\left(\int_{\Delta^n[0,t]} \circ \mathrm{d}B^I\right) = \frac{t^{\frac{n+n(I)}{2}}}{2^{\frac{n-n(I)}{2}}\left(\frac{n+n(I)}{2}\right)!},$$
where $n(I)$ is the number of 0 in I (observe that since $I \in \mathcal{I}_n$, n and $n(I)$ necessarily have the same parity).

We conclude therefore

$$\| \mathbf{P}_t^N f - \sum_{k \leq \frac{N+1}{2}} \frac{t^k}{k!} \mathcal{L}^k f \|_k = O\left(t^{\frac{N+1}{2}}\right).$$

Since it is known that

$$\| \mathbf{P}_t f - \sum_{k \leq \frac{N+1}{2}} \frac{t^k}{k!} \mathcal{L}^k f \|_k = O\left(t^{\frac{N+1}{2}}\right),$$

the theorem is proved. □

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

The previous approximation theorem may be used to obtain small times heat kernel asymptotics.

From now on, we assume furthermore that the operator $\mathcal{L}$ is elliptic in the sense that for $x_0 \in \mathbb{R}^n$ the vector fields $V_1(x_0), ..., V_d(x_0)$ form a basis of $\mathbb{R}^n$ which of course implies $n = d$. In that case, it is known that the random variable $X_t^{x_0}$ admits a smooth density $p_t(x_0, \cdot)$ with respect to the Lebesgue measure of $\mathbb{R}^n$. In other words,

$$\mathbb{P}(X_t^{x_0} \in \mathrm{d}y) = p_t(x_0, y)\mathrm{d}y,$$

for some smooth function $p(x_0, \cdot) : (0, +\infty) \times \mathbb{R}^n \to \mathbb{R}_{\geq 0}$.

We are interested in $p_t(x_0, y)$ in small times. Before we give the Theorem on the asymptotics of $p_t(x_0, y)$, let us recall that given smooth vector fields $V_1, \ldots, V_n$ on $\mathbb{R}^n$ that form a basis at each point, it is well-known we may naturally associate the distance:

$$\mathrm{d}(x, y) = \inf_{\gamma \in \mathcal{C}(x,y)} \int_0^1 \|\gamma'(s)\| \mathrm{d}s$$

where $\mathcal{C}(x, y)$ denotes the set of absolutely continuous curves $\gamma : [0, 1] \to \mathbb{R}^n$, such that the solution of the ordinary differential equation

$$\mathrm{d}x_t = \sum_{i=1}^n V_i(x_t)\mathrm{d}\gamma_t^i,$$

satisfies $x_0 = x, x_1 = y$.

More analytically, this distance may also be defined as

$$\mathrm{d}(x, y) = \sup \left\{ f(x) - f(y), f \in C_b^\infty(\mathbb{R}^n), \sum_{i=1}^n (V_i f)^2 \leq 1 \right\},$$

where $C_b^\infty(\mathbb{R}^n)$ denotes the set of smooth and bounded functions on $\mathbb{R}^n$. If the vector fields $V_1, \ldots, V_n$ are Lipschitz then it is known that this metric is complete and that the Hopf-Rinow theorem holds (That is, the compact sets are the closed and bounded sets).

In this framework, the following heat kernel asymptotics holds:

Theorem 4.2. Let $x_0 \in \mathbb{R}^n$. There exists $\varepsilon > 0$ and smooth functions, $a_k(x_0, y)$, $y \in \mathbf{B}(x_0, \varepsilon)$ such that uniformly on $\mathbf{B}(x_0, \varepsilon)$, when $t \to 0$,

$$p_t(x_0, y) = \frac{\mathrm{e}^{-\frac{d(x_0,y)^2}{2t}}}{t^{n/2}} \left(\sum_{k=0}^N a_k(x_0, y)t^k + O(t^{N+1}) \right), \quad N \geq 0.$$

Proof. This theorem admits numerous proofs. A purely analytic proof based on the parametrix method may be found in Chapter 1 of the book by Friedman [20]. In the seminal works [2,3], Azencott introduced probabilistic methods to prove the result. These methods introduced by Azencott were then further developed by Ben Arous and Léandre in [7, 8, 11, 31], in order to cover the subelliptic case. Let us sketch the approach of [8] which is based on the Laplace method on the Wiener space.

The first idea is to consider the scaled stochastic differential equation

$$\mathrm{d}X_t^\varepsilon = \varepsilon \sum_{i=1}^n V_i(X_t^\varepsilon)\mathrm{d}B_t^i + \varepsilon^2 V_0(X_t^\varepsilon)\mathrm{d}t \qquad \text{with } X_0^\varepsilon = x_0.$$

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Then, we observe that there exist neighborhoods U, V of x_0 and a bounded smooth function $F(x,y,z)$ on $U \times V \times \mathbb{R}^n$ such that:

(1) For any $(x,y) \in U \times V$ the infimum

$$\inf\left\{F(x,y,z) + \frac{\mathrm{d}(x,z)^2}{2}, z \in \mathbb{R}^n\right\} = 0$$

is attained at the unique point y.

(2) For each $(x,y) \in U \times V$, there exists a ball centered at y with radius r independent of x, y such that $F(x,y,\cdot)$ is a constant outside of the ball.

So, by denoting $p_\varepsilon(x_0,y)$ the density of X_1^ε, by the Fourier inversion formula we have

$$\begin{aligned} p_\varepsilon(x_0,y)\mathrm{e}^{-\frac{F(x_0,y,y)}{\varepsilon^2}} &= \frac{1}{(2\pi)^d}\int \mathrm{e}^{-i\zeta\cdot y}\mathrm{d}\zeta \int \mathrm{e}^{i\zeta\cdot z}\mathrm{e}^{-\frac{F(x_0,y,z)}{\varepsilon^2}} p_\varepsilon(x_0,z)\mathrm{d}z \\ &= \frac{1}{(2\pi\varepsilon)^d}\int \mathrm{e}^{-i\frac{\zeta\cdot y}{\varepsilon}}\mathrm{d}\zeta \int \mathrm{e}^{i\frac{\zeta\cdot z}{\varepsilon}}\mathrm{e}^{-\frac{F(x_0,y,z)}{\varepsilon^2}} p_\varepsilon(x_0,z)\mathrm{d}z \\ &= \frac{1}{(2\pi\varepsilon)^d}\int \mathrm{d}\zeta \mathbb{E}\left(\mathrm{e}^{\frac{i\zeta\cdot(X_1^\varepsilon - y)}{\varepsilon}}\mathrm{e}^{\frac{F(x_0,y,X_1^\varepsilon)}{\varepsilon^2}}\right). \end{aligned}$$

Therefore, the asymptotics of $p_t(x_0,y)$ may be understood from the asymptotics when $\varepsilon \to 0$ of

$$J_\varepsilon(x_0,y) = \mathbb{E}\left(\mathrm{e}^{\frac{i\zeta\cdot(X_1^\varepsilon - y)}{\varepsilon}}\mathrm{e}^{\frac{F(x_0,y,X_1^\varepsilon)}{\varepsilon^2}}\right).$$

Then, by using the Laplace method on the Wiener space based on the large deviation principle (see [7]), we get an expansion in powers of ε of $J_\varepsilon(x_0,y)$, leading therefore to the expected result. □

As a consequence, we see that the following asymptotic expansion holds when $t \to 0$,

$$p_t(x_0,x_0) = \frac{1}{t^{\frac{d}{2}}}\left(\sum_{k=0}^N a_k(x_0)t^k\right) + O\left(t^{N+1-\frac{d}{2}}\right), \quad N \geq 0,$$

for some constants $a_0(x_0), \ldots, a_N(x_0)$. The following proposition provides an effective way to compute these constants.

Proposition 4.3. Assume that for every $x_0 \in \mathbb{R}^n$ the vector fields $V_1(x_0), ..., V_n(x_0)$ form a basis of $\mathbb{R}^n$, then for $N \geq 1$, when $t \to 0$,

$$p_t(x_0,x_0) = \mathrm{d}_t^N(x_0) + O\left(t^{\frac{N+1-d}{2}}\right),$$

where $\mathrm{d}_t^N(x_0)$ is the density at 0 of the random variable $\sum_{I,d(I)\leq N} \Lambda_I(B)_t V_I(x_0)$

Proof. This is a particular case of Theorem 5.2 which is proven below. However, since the geometric framework is simpler in this Section, let us outline the main ingredients of the proof.

The keypoint is that from Castell's approximation result, as we have seen in Theorem 4.1, we have for $N \geq 1$, $k \geq 0$ and f, C^∞ bounded

$$\| \mathbf{P}_t f - \mathbf{P}_t^N f \|_k = O\left(t^{\frac{N+1}{2}}\right), \quad t \to 0. \tag{4.6}$$

Then, we observe that due to the ellipticity assumption on the vector fields $V_1, ..., V_n$, it turns out that for $N \geq 1$ the approximation $\mathbf{P}_t^N$ admits a heat kernel, in the sense that there is a smooth function $p_t^N(x,y)$ such

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

that for every C^∞ bounded function

$$\mathbf{P}_t^N f(x_0) \int_{\mathbb{R}^n} p_t^N(x_0, y) f(y) \mathrm{d}y.$$

We therefore deduce from (4.6) that

$$p_t(x_0, x_0) = p_t^N(x_0, x_0) + O\left(t^{\frac{N+1-d}{2}}\right),$$

which gives the expected result from the fact that

$$\mathbf{P}_t^N f(x_0) = \mathbb{E}\left(f\left(\exp\left(\sum_{I, d(I)\leq N} \Lambda_I(B)_t V_I\right) x_0\right)\right).$$

□

Remark 4.4. This result and similar arguments may also be found in [29].

For instance, by applying the previous proposition with $N = 1$, we get

$$a_0(x_0) = \frac{1}{(2\pi)^{\frac{d}{2}}} \frac{1}{|\det(V_1(x_0), \ldots, V_d(x_0))|}.$$

The computation of $a_1(x)$ is technically more involved and relies on the so-called Lévy's area formula. In order to simplify the following computations we shall make some assumptions that avoid heavy computations, however the described methodology may be extended to the general case. First we assume $V_0 = 0$.

We wish to apply the previous proposition with $N = 2$. For that, we need to understand the law of the random variable

$$\Theta_t = \sum_{i=1}^d B_t^i V_i(x_0) + \frac{1}{2} \sum_{1\leq i<j\leq d} \int_0^t B_s^i \mathrm{d}B_s^j - B_s^j \mathrm{d}B_i^s [V_i, V_j](x_0).$$

Since L is assumed to be elliptic at x_0, we can find ω_{ij}^k such that $\omega_{ij}^k = -\omega_{ji}^k$ and

$$[V_i, V_j](x_0) = \sum_{k=1}^d \omega_{ij}^k V_k(x_0).$$

Our second assumption is the following skew-symmetry property[5] $\omega_{ij}^k = -\omega_{ik}^j$. With these notations, we therefore have

$$\Theta_t = \sum_{k=1}^d \left(B_t^k + \frac{1}{2} \sum_{1\leq i<j\leq d} \omega_{ij}^k \int_0^t B_s^i \mathrm{d}B_s^j - B_s^j \mathrm{d}B_s^i\right) V_k(x_0).$$

By a simple linear transformation, we are reduced to the problem of the computation of the law of the $\mathbb{R}^d$-valued random variable

$$\theta_t = \left(B_t^k + \frac{1}{2} \sum_{1\leq i<j\leq d} \omega_{ij}^k \int_0^t B_s^i \mathrm{d}B_s^j - B_s^j \mathrm{d}B_s^i\right)_{1\leq k\leq d}.$$

It is known from the Lévy's area formula that if A is a $d \times d$ skew-symmetric matrix, then, for $t > 0$,

$$\mathbb{E}\left(\mathrm{e}^{i\int_0^t (AB_s, \mathrm{d}B_s)} \mid B_t = z\right) = \det\left(\frac{tA}{\sin tA}\right)^{\frac{1}{2}} \exp\left(\frac{I - tA\cot tA}{2t} z, z\right).$$

[5] This assumption, together with $V_0 = 0$, for instance holds when L is the Laplace–Beltrami operator on a compact semi-simple Lie group.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Therefore, for $\lambda \in \mathbb{R}^d$,

$$\mathbb{E}\left(\mathrm{e}^{i(\lambda,\theta_t)}\right) = \int_{\mathbb{R}^d} \mathrm{e}^{i(\lambda,y)} \frac{\mathrm{e}^{-\frac{\|y\|^2}{2t}}}{(2\pi t)^{d/2}} \det\left(\frac{tA}{\sin tA}\right)^{\frac{1}{2}} \exp\left(\frac{I - tA\cot tA}{2t} y, y\right) \mathrm{d}y,$$

where

$$A_{ij} = \frac{1}{2}\sum_{k=1}^{d} \lambda_k \omega_{ij}^k.$$

Thus, from the inverse Fourier transform formula, the density of θ_t with respect to the Lebesgue measure is given by

$$q_t(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathrm{e}^{-i(\lambda,x)}\mathrm{e}^{i(\lambda,y)} \frac{\mathrm{e}^{-\frac{\|y\|^2}{2t}}}{(2\pi t)^{d/2}} \det\left(\frac{tA}{\sin tA}\right)^{\frac{1}{2}} \exp\left(\frac{I - tA\cot tA}{2t} y, y\right) \mathrm{d}y\mathrm{d}\lambda$$

According to Proposition 4.3, we are interested in

$$q_t(0) = \frac{1}{(2\pi)^{3d/2} t^{d/2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathrm{e}^{i(\lambda,y)} \det\left(\frac{tA}{\sin tA}\right)^{\frac{1}{2}} \exp\left(-\frac{A\cot tA}{2} y, y\right) \mathrm{d}y\mathrm{d}\lambda$$

and wish to perform an asymptotic development when $t \to 0$. By using the standard Laplace method, we are led to

$$q_t(0) = \frac{1}{(2\pi t)^{\frac{d}{2}}}\left(1 - \frac{1}{16}\sum_{i,j,k=1}^{d} (\omega_{ij}^k)^2 t + O(t^2)\right)$$

which implies therefore:

$$p_t(x_0,x_0) = \frac{1}{|\det(V_1(x_0),\ldots,V_d(x_0))|} \frac{1}{(2\pi t)^{\frac{d}{2}}}\left(1 - \frac{1}{16}\sum_{i,j,k=1}^{d} (\omega_{ij}^k)^2 t + O(t^2)\right).$$

Remark 4.5. At that time, up to the knowledge of the author, a similar method may not be achieved to compute for instance $a_2(x_0)$. This is due to the fact that the law of the random variable

$$\sum_{I, d(I)\le N} \Lambda_I(B)_t V_I(x_0)$$

is poorly understood when $N \ge 3$. However, Proposition 4.3 makes explicitly appear the geometric information contained in the coefficients $a_i(x_0)$.

5. Elliptic heat kernels asymptotics on vector bundles and the Chern–Gauss–Bonnet theorem

The goal of the present section is to extend the above results to the study of elliptic heat kernels on vector bundles. In particular, we will see that the methods that we developed are sharp enough to recover the celebrated heat kernel proof of the Chern–Gauss–Bonnet theorem. The relevant geometric quantity involved in this theorem is hidden very far in the asymptotic development of the heat kernel but will appear in a straightforward way by using stochastic Taylor expansions.

Our first goal is to adapt what we did before, to vector bundles over compact Riemannian manifolds. We assume here from the reader some basic knowledge on Riemannian manifolds, vector bundles and linear connections. An account on these materials which is designed for probabilists may be found in the book by Hsu [26].

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

A more detailed account may be found in the book by Berline–Getzler–Vergne [12]. In particular, in this last reference, the authors give in Chapter 1 the detailed construction of the heat kernel on vector bundles over compact Riemannian manifolds.

5.1. Elliptic heat kernels asymptotics on vector bundles

Let $\mathbb{M}$ be a d-dimensional compact smooth Riemannian manifold and let $\mathcal{E}$ be a finite-dimensional vector bundle over $\mathbb{M}$. We denote by $\Gamma(\mathbb{M},\mathcal{E})$ the space of smooth sections of this bundle. Let now ∇ denote a connection on $\mathcal{E}$.

We consider the following linear partial differential equation

$$\frac{\partial \Phi}{\partial t} = \mathcal{L}\Phi, \quad \Phi(0,x) = f(x), \tag{5.7}$$

where $\mathcal{L}$ is an operator on $\mathcal{E}$ that can be written

$$\mathcal{L} = \nabla_0 + \frac{1}{2}\sum_{i=1}^{d} \nabla_i^2,$$

with

$$\nabla_i = \mathcal{F}_i + \nabla_{V_i}, \quad 0 \leq i \leq d,$$

the V_i's being smooth vector fields on $\mathbb{M}$ and the $\mathcal{F}_i$'s being smooth potentials (that is smooth sections of the bundle $\mathbf{End}(\mathcal{E})$). As in the scalar case, it is known that the solution of (5.7) can be written

$$\Phi(t,x) = (\mathrm{e}^{t\mathcal{L}} f)(x) = \mathbf{P}_t f(x),$$

where $\mathbf{P}_t$ is a contraction semigroup of operators. If $I \in \{0,1,...,d\}^k$ is a word, we denote

$$\nabla_I = [\nabla_{i_1}, [\nabla_{i_2}, ..., [\nabla_{i_{k-1}}, \nabla_{i_k}]...].$$

and

$$d(I) = k + n(I),$$

where $n(I)$ is the number of 0 in the word I.

For $N \geq 1$, let us consider

$$\mathbf{P}_t^N = \mathbb{E}\left(\exp\left(\sum_{I, d(I)\leq N} \Lambda_I(B)_t \nabla_I\right)\right).$$

As before for the scalar case, the meaning of this last notation is the following. If $f \in \Gamma(\mathbb{M},\mathcal{E})$, then $(\mathbf{P}_t^N f)(x) = \mathbb{E}(\Psi(1,x))$, where $\Psi(\tau,x)$ is the solution of the first order partial differential equation with random coefficients:

$$\frac{\partial \Psi}{\partial \tau}(\tau,x) = \sum_{I, d(I)\leq N} \Lambda_I(B)_t (\nabla_I \Psi)(\tau,x), \quad \Psi(0,x) = f(x).$$

Let us consider the following family of norms: If $f \in \Gamma(\mathbb{M},\mathcal{E})$, for $k \geq 0$,

$$\| f \|_k = \sup_{0\leq l\leq k} \sup_{0\leq i_1,...,i_l\leq d} \sup_{x\in\mathbb{M}} \| \nabla_{i_1} \dots \nabla_{i_l} f(x) \| .$$

We have the following extension of Theorem 4.1 which may be proved in the very same way.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Theorem 5.1. Let $N \geq 1$ and $k \geq 0$. For $f \in \Gamma(\mathbb{M}, \mathcal{E})$,

$$\| \mathbf{P}_t f - \mathbf{P}_t^N f \|_k = O\left(t^{\frac{N+1}{2}}\right), \quad t \to 0.$$

Let us now assume that the operator $\mathcal{L}$ is elliptic at $x_0 \in \mathbb{M}$ in the sense that $(V_1(x_0), ..., V_d(x_0))$ is an orthonormal basis of the tangent space at x_0. In that case, $\mathbf{P}_t$ is known to admit a smooth Schwartz kernel at x_0. That is, there exists a smooth map

$$p(x_0, \cdot) : \mathbb{R}_{>0} \to \Gamma(\mathbb{M}, \mathbf{Hom}(\mathcal{E}))$$

such that for $f \in \Gamma(\mathbb{M}, \mathcal{E})$,

$$(\mathbf{P}_t f)(x_0) = \int_{\mathbb{M}} p_t(x_0, y) f(y) \mathrm{d}y.$$

Theorem 5.2. Let $N \geq 1$. There exists a map

$$p^N(x_0, \cdot) : \mathbb{R}_{>0} \to \Gamma(\mathbb{M}, \mathbf{Hom}(\mathcal{E}))$$

such that for $f \in \Gamma(\mathbb{M}, \mathcal{E})$,

$$(\mathbf{P}_t^N f)(x_0) = \int_{\mathbb{M}} p_t^N(x_0, y) f(y) \mathrm{d}y.$$

Moreover,

$$p_t(x_0, x_0) = p_t^N(x_0, x_0) + 0\left(t^{\frac{N+1-d}{2}}\right).$$

Proof. The proof is not simple. We shall proceed in several steps. In a first step, we shall show the existence of a kernel at x_0 for $\mathbf{P}_t^N$ acting on functions. In a second step we shall deduce by parallel transport, the existence of $p^N(x_0, \cdot)$. And finally, we shall prove the required estimate.

First step:

Let us define,

$$\mathbf{Q}_t^N = \mathbb{E}\left(\exp\left(\sum_{I, d(I) \leq N} \Lambda_I(B)_t V_I\right)\right).$$

In order to show that $\mathbf{Q}_t^N$ admits a kernel at x_0, we show that for $t > 0$, the stochastic process

$$Z_t^N = \exp\left(\sum_{I, d(I) \leq N} \Lambda_I(B)_t V_I\right)(x_0)$$

has a density with respect to the Riemannian measure of $\mathbb{M}$. To this end, from the well-known criterion of Malliavin (see [35,36]), we show that the Malliavin matrix of Z_t^N is invertible with probability one. A sufficient condition for that, is

$$\mathbb{D}_0^i Z_t^N, \quad i = 1, ..., d,$$

forms a basis of the tangent space at x_0 where $\mathbb{D}_0^i$ denotes the *ith* partial Malliavin's derivative taken at time 0. An easy computation shows that

$$\mathbb{D}_0^i Z_t^N = V_i(x_0), \quad t > 0.$$

Our ellipticity assumption gives therefore the existence of $q^N(x_0, \cdot) : \mathbb{R}_{>0} \times \mathbb{M} \to \mathbb{R}_{\geq 0}$, such that for every smooth $f : \mathbb{M} \to \mathbb{R}$,

$$(\mathbf{Q}_t^N f)(x_0) = \int_{\mathbb{M}} q_t^N(x_0, y) f(y) \mathrm{d}y.$$

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

Second step:

For $t>0$, let us consider the operator $\Theta_t^N(x_0)$ defined on $\Gamma(\mathbb{M},\mathcal{E})$ by the property that for $\eta\in\Gamma(\mathbb{M},\mathcal{E})$ and $y\in\mathcal{O}_{x_0}$,

$$(\Theta_t^N(x_0)\eta)(y)=\mathbb{E}\left(\left[\exp\left(\sum_{I,d(I)\leq N}\Lambda_I(B)_t\nabla_I\right)\eta\right](x_0)\,\middle|\,\exp\left(\sum_{I,d(I)\leq N}\Lambda_I(B)_tV_I\right)(x_0)=y\right).$$

We claim that $\Theta_t^N(x_0)$ is actually a potential, that is a section of the bundle $\mathbf{End}(\mathcal{E})$. For that, we have to show that for every smooth $f:\mathbb{M}\to\mathbb{R}$ and every $\eta\in\Gamma(\mathbb{M},\mathcal{E})$, $y\in\mathcal{O}_{x_0}$,

$$(\Theta_t^N(x_0)f\eta)(y)=f(y)(\Theta_t^N(x_0)\eta)(y).$$

If f is a smooth function on $\mathbb{M}$, we denote by $\mathcal{M}_f$ the operator on $\Gamma(\mathbb{M},\mathcal{E})$ that acts by multiplication by f. Due to the Leibniz rule for connections, we have for any word I:

$$[\nabla_I,\mathcal{M}_f]=\mathcal{M}_{V_If}.$$

Consequently,

$$\left[\sum_{I,d(I)\leq N}\Lambda_I(B)_t\nabla_I,\mathcal{M}_f\right]=\mathcal{M}_{\sum_{I,d(I)\leq N}\Lambda_I(B)_tV_If}.$$

The above commutation property implies the following one:

$$\exp\left(\sum_{I,d(I)\leq N}\Lambda_I(B)_t\nabla_I\right)\mathcal{M}_f=\mathcal{M}_{\exp\left(\sum_{I,d(I)\leq N}\Lambda_I(B)_tV_I\right)f}\exp\left(\sum_{I,d(I)\leq N}\Lambda_I(B)_t\nabla_I\right).$$

Therefore,

$$[\Theta_t^N(x_0),\mathcal{M}_f]=0,$$

so that $\Theta_t^N(x_0)$ is a section of the bundle $\mathbf{End}(\mathcal{E})$. We can now conclude with the disintegration formula that for every $\eta\in\Gamma(\mathbb{M},\mathcal{E})$,

$$(\mathbf{P}_t^N\eta)(x_0)=\int_{\mathbb{M}}p_t^N(x_0,y)\eta(y)\mathrm{d}y,$$

with

$$p_t^N(x_0,\cdot)=q_t^N(x_0,\cdot)\Theta_t^N(x_0).$$

Final step:

Let us now turn to the proof of the pointwise estimate

$$p_t(x_0,x_0)=p_t^N(x_0,x_0)+O\left(t^{\frac{N+1-d}{2}}\right),\quad t\to 0.$$

Let $y\in\mathbb{M}$ be sufficiently close to x_0. Since $\mathcal{L}$ is elliptic at x_0, it is known (see for instance Chapter 2 [12]) that $p_t(x_0,y)$ admits a development

$$p_t(x_0,y)=\frac{\mathrm{e}^{-\frac{\mathrm{d}^2(x_0,y)}{2t}}}{(2\pi t)^{d/2}}\left(\sum_{k=0}^N\Psi_k(x_0,y)t^{\frac{k}{2}}+t^{\frac{N+1}{2}}\mathbf{R}_N(t,x_0,y)\right),\tag{5.8}$$

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

where the remainder term $\mathbf{R}_N(t,x_0,y)$ is bounded when $t\to 0$, $\Psi_k(x_0,\cdot)$ is a section of $\mathbf{End}(\mathcal{E})$ defined around x_0 and $d(\cdot,\cdot)$ is the distance defined around x_0 by the vector fields $V_1,...,V_d$. By using the fact that for every smooth $f:\mathbb{M}\to\mathbb{R}$,

$$(\mathbf{Q}_t^N f)(x_0)=\mathbb{E}\left(f\left(\exp\left(\sum_{I,d(I)\le N}\Lambda_I(B)_t V_I\right)(x_0)\right)\right),\quad t\ge 0,$$

and classical results for asymptotic development in small times of subelliptic heat kernels (see for instance [9] and Chapter 3 of [4]), we get for $q_t^N(x_0,y)$ a development that is similar to (5.8). For $\Theta_t^N(x_0)$, the scaling property of Brownian motion implies that we have a short-time asymptotics in powers $t^{\frac{k}{2}}$, $k\in\mathbb{N}$. Since,

$$p_t^N(x_0,\cdot)=q_t^N(x_0,\cdot)\Theta_t^N(x_0),$$

we deduce that

$$p_t^N(x_0,y)=\frac{\mathrm{e}^{-\frac{\mathrm{d}^2(x_0,y)}{2t}}}{(2\pi t)^{d/2}}\left(\sum_{k=0}^{N}\tilde{\Psi}_k(x_0,y)t^{\frac{k}{2}}+t^{\frac{N+1}{2}}\tilde{\mathbf{R}}_N(t,x_0,y)\right),$$

where the remainder term $\tilde{\mathbf{R}}_N(t,x_0,y)$ is bounded when $t\to 0$. With Theorem 4.1, we obtain that $\Psi_k=\tilde{\Psi}_k$, $k=0,...,N$, and the required estimate easily follows. □

Remark 5.3. The question of the smoothness of p_t^N is not addressed here. It would require bounds on the inverse of the Malliavin matrix of Z_t^N.

From the previous theorem, we deduce an explicit asymptotic expansion of $p_t(x_0,x_0)$. If $I\in\{0,1,...,d\}^k$, $k\ge 2$, is a word, we denote

$$\mathcal{F}_I=\nabla_I-\nabla_{V_I}\in\Gamma(\mathbb{M},\mathbf{End}(\mathcal{E})).$$

Corollary 5.4. For $N\ge 1$, when $t\to 0$,

$$p_t(x_0,x_0)=\mathrm{d}_t^N(x_0)\mathbb{E}\left(\exp\left(\sum_{I,d(I)\le N}\Lambda_I(B)_t\mathcal{F}_I(x_0)\right)\middle|\sum_{I,d(I)\le N}\Lambda_I(B)_t V_I(x_0)=0\right)+O\left(t^{\frac{N+1-d}{2}}\right),$$

where $\mathrm{d}_t^N(x_0)$ is the density at 0 of the random variable $\sum_{I,d(I)\le N}\Lambda_I(B)_t V_I(x_0)$.

Proof. Let us first observe that for the same reason than in the proof of step 1 of the above theorem, the random process

$$\sum_{I,d(I)\le N}\Lambda_I(B)_t V_I(x_0)$$

has a density $\mathrm{d}_t^N(x_0,\cdot)$. Therefore, due to the disintegration formula, for every smooth $\eta\in\Gamma(\mathbb{M},\mathcal{E})$,

$$(\mathbf{P}_t^N\eta)(x_0)=\int_{\mathbf{T}_{x_0}\mathbb{M}}\mathbb{E}\left(\exp\left(\sum_{I,d(I)\le N}\Lambda_I(B)_t\nabla_I\right)\eta(x_0)\middle|\sum_{I,d(I)\le N}\Lambda_I(B)_t V_I(x_0)=0\right)\mathrm{d}_t^N(x_0,y)\mathrm{d}y,$$

and the proof follows by letting η converge to Dirac distribution at x_0. □

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

5.2. The Chern–Gauss–Bonnet theorem

In this Section we show an application of the materials presented in these notes to local index theorems. We focus here on the simplest one which is the Chern–Gauss–Bonnet theorem. A proof of the (more general) Atiyah–Singer local index theorem for the Dirac operator on spin manifolds that uses the above machinery can be found in [5].

Let us mention here that the heat equation approach to index theorems is not new and particularly fruitful: It was suggested by Atiyah–Bott [1] and McKean–Singer [34], and first carried out by Patodi [37] and Gilkey [25]. Bismut in [13] introduces stochastic methods based on Feynman–Kac formula. For probabilistic approaches, we also refer to [30] and Chapter 7 of [26]. For a complete survey on (non probabilistic) heat equation methods for index theorems, we refer to the book [12].

Let $\mathbb{M}$ be a d-dimensional Riemannian, compact, smooth manifold. The Chern–Gauss–Bonnet theorem proved by Chern [17] in 1944 is the following:

Theorem 5.5. Let $\chi(\mathbb{M})$ be the Euler characteristic of $\mathbb{M}$. If d is odd, then $\chi(\mathbb{M}) = 0$. If d is even then

$$\chi(\mathbb{M}) = \int_{\mathbb{M}} \omega(x)\mathrm{d}x,$$

where $\omega(x)\mathrm{d}x$ is the Euler form, that is the volume form given in a local orthonormal frame e_i by

$$\omega = \frac{(-1)^{d/2}}{(8\pi)^{d/2}(d/2)!} \sum_{\sigma,\tau\in\Sigma_d} \epsilon(\sigma)\epsilon(\tau) \prod_{i=1}^{d-1} R_{\sigma(i)\sigma(i+1)\tau(i)\tau(i+1)}\mathrm{d}x,$$

where Σ_d is the set of the permutations of the indices $\{1, ..., d\}$, ϵ the signature of a permutation, and

$$R_{ijkl} = \langle R(e_j, e_k)e_l, e_i\rangle,$$

with R Riemannian curvature of $\mathbb{M}$.

The striking feature of Chern–Gauss–Bonnet theorem that makes it so beautiful is that the Euler form depends on the Riemannian metric whereas $\chi(\mathbb{M})$ is only a topological invariant. We now turn to a short proof of it that uses the tools we developed in these notes. In the sequel, we always assume that the dimension d is even.

We first briefly recall some basic facts on Fermion calculus on the Clifford exterior algebra of a finite dimensional vector space, as can be found in Section 2.2.2 of [39] and that will be used in our proof. Let V be a d-dimensional Euclidean vector space. We denote V^* its dual and

$$\wedge V^* = \bigoplus_{k\geq 0} \wedge^k V^*,$$

the exterior algebra. If $u \in V^*$, we denote a_u^* the map $\wedge V^* \to \wedge V^*$, such that $a_u^*(\omega) = u \wedge \omega$. The dual map is denoted a_u. Let now θ_1, ..., θ_d be an orthonormal basis of V^*. We denote $a_i = a_{\theta_i}$. We have the basic rules of Fermion calculus

$$\{a_i, a_j\} = 0, \{a_i^*, a_j^*\} = 0, \{a_i, a_j^*\} = \delta_{ij},$$

where $\{\cdot,\cdot\}$ stands for the anti-commutator: $\{a_i, a_j\} = a_i a_j + a_j a_i$. If I and J are two words with $1 \leq i_1 < \ldots < i_k \leq d$ and $1 \leq j_1 < \ldots < j_l \leq d$, we denote

$$A_{IJ} = a_{i_1}^* \ldots a_{i_k}^* a_{j_1} \ldots a_{j_l}.$$

The family of all the possible A_{IJ} forms a basis of the 2^{2d}-dimensional vector space $\mathbf{End}\,(\wedge V^*)$.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

If $A = \sum_{I,J} c_{IJ} A_{IJ} \in \mathbf{End}\,(\wedge V^*)$, we shall say that

If $A \in \mathbf{End}\,(\wedge V^*)$, we define the supertrace $\mathbf{Str}(A)$ as the difference of the trace of A on even forms minus the trace of A on odd forms.

One of the the interests of Fermion calculus (which is equivalent to Clifford calculus) is that it makes easy to compute supertraces: If $A = \sum_{I,J} c_{IJ} A_{IJ}$, then

$$\mathbf{Str}(A) = (-1)^{\frac{d(d-1)}{2}} c_{\{1,\ldots,d\}\{1,\ldots,d\}}. \tag{5.9}$$

We now carry the Fermionic construction on the tangent spaces of our manifold $\mathbb{M}$. Let e_i be a local orthonormal frame and let θ_i be its dual frame. The curvature endomorphism is defined by

$$\mathcal{F} = -\sum_{ijkl} R_{ijkl} a_i^* a_j a_k^* a_l$$

where

$$R_{ijkl} = \langle R(e_j, e_k) e_l, e_i \rangle ,$$

with R Riemannian curvature of $\mathbb{M}$. This definition is actually intrinsic, *i.e.* does not depend on the choice of the local frame. In this setting, the celebrated Weitzenböck formula reads

$$\Box = \Delta + \mathcal{F},$$

where $\Box = d\mathrm{d}^* + \mathrm{d}^* d$ is the Hodge–DeRham Laplacian and Δ the Bochner Laplacian. Let us recall that if e_i is a local orthonormal frame, we have the following explicit formula for Δ:

$$\Delta = -\sum_{i=1}^{d} (\nabla_{e_i} \nabla_{e_i} - \nabla_{\nabla_{e_i} e_i}),$$

where ∇ is the Levi–Civita connection.

After these preliminaries, we can now turn to the proof of Chern–Gauss–Bonnet theorem. From now on, we suppose that the dimension d is even. The first crucial step is McKean–Singer formula [34] (A simple proof of it can be found in [39], p. 113). We have

$$\chi(\mathbb{M}) = \int_{\mathbb{M}} \mathbf{Str}\; p_t(x,x) \mathrm{d}x, \quad t > 0,$$

where $\mathbf{P}_t = \mathrm{e}^{-t\Box}$ and p_t is the corresponding Schwartz kernel (density). In other words, the supertrace is constant along the heat semigroup associated with the Hodge–DeRham Laplacian and this constant is equal to the Euler characteristic.

An easily proved and non surprising precise statement is the following: When $t \to 0$,

$$\sup_{x \in \mathbb{M}} \| \; p_t(x,x) - \frac{1}{(4\pi t)^{\frac{d}{2}}} \mathrm{e}^{-t\mathcal{F}}(x) \; \| = O\left(\frac{1}{t^{d/2-1}} \right).$$

But as seen in the next proposition, due to Corollary 5.4, *fantastic and subtle cancellations*[6] occur at the paths level when we take the supertrace:

Proposition 5.6. For every $x \in \mathbb{M}$,

$$\lim_{t \to 0} \mathbf{Str}\; p_t(x,x) = \omega(x).$$

[6]We quote here McKean–Singer [34] who conjectured these cancellations of terms that should involve many covariant derivatives of curvature terms.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

Proof. Let $x_0 \in \mathbb{M}$ be fixed once time or all in the following proof. We work in a synchronous local orthonormal frame e_i around x_0, that is $\nabla e_i = 0$ at x_0. At the point x_0, we have

$$\Delta = -\sum_{i=1}^{d} \nabla_{e_i} \nabla_{e_i}$$

and therefore

$$\Box = -\sum_{i=1}^{d} \nabla_{e_i} \nabla_{e_i} + \mathcal{F}(x_0).$$

We want to apply Corollary 5.4 in the present framework.

We denote $\mathcal{F}_0 = -\frac{1}{2}\mathcal{F}$, $\mathcal{F}_i = 0$, $1 \leq i \leq d$ and if $I \in \{0, 1, ..., d\}^k$ is a word,

$$\mathcal{F}_I = [\nabla_{i_1}, [\nabla_{i_2}, ..., [\nabla_{i_{k-1}}, \nabla_{i_k}]...] - \nabla_{[e_{i_1}, [e_{i_2}, ..., [e_{i_{k-1}}, e_{i_k}]...]} \in \Gamma(\mathbb{M}, \mathbf{End}\,(\wedge T^*\mathbb{M})),$$

with the convention $e_0 = 0$, $\nabla_0 = \mathcal{F}_0$, $\nabla_i = \nabla_{e_i}$, $1 \leq i \leq d$. According to Corollary 5.4, we thus have for $N \geq 1$, and $t \to 0$,

$$p_{t/2}(x_0, x_0) = q_t^N(x_0) \mathbb{E}\left(\exp\left(\sum_{I, d(I) \leq N} \Lambda_I(B)_t \mathcal{F}_I(x_0) \right) \middle| \sum_{I, d(I) \leq N} \Lambda_I(B)_t e_I(x_0) = 0 \right) + O\left(t^{\frac{N+1-d}{2}}\right),$$

where $q_t^N(x_0)$ is the density at 0 of the random variable $\sum_{I, d(I) \leq N} \Lambda_I(B)_t e_I(x_0)$. Applying this when $N = d$ gives

$$p_{t/2}(x_0, x_0) = q_t^d(x_0) \mathbb{E}\left(\exp\left(\sum_{I, d(I) \leq d} \Lambda_I(B)_t \mathcal{F}_I(x_0) \right) \middle| \sum_{I, d(I) \leq N} \Lambda_I(B)_t e_I(x_0) = 0 \right) + O\left(\sqrt{t}\right).$$

By using the scaling property of Brownian motion, it is easily seen that for $k \geq \frac{d}{2} + 1$,

$$q_t^d(x_0) \mathbb{E}\left(\left(\sum_{I, d(I) \leq d} \Lambda_I(B)_t \mathcal{F}_I(x_0) \right)^k \middle| \sum_{I, d(I) \leq N} \Lambda_I(B)_t e_I(x_0) = 0 \right) = O\left(\sqrt{t}\right).$$

Therefore

$$p_{t/2}(x_0, x_0) = q_t^d(x_0) \mathbb{E}\left(\sum_{k=0}^{d/2} \frac{1}{k!} \left(\sum_{I, d(I) \leq d} \Lambda_I(B)_t \mathcal{F}_I(x_0) \right)^k \middle| \sum_{I, d(I) \leq N} \Lambda_I(B)_t e_I(x_0) = 0 \right) + O\left(\sqrt{t}\right)$$

and

$$\mathbf{Str}\; p_{t/2}(x_0, x_0) = q_t^d(x_0) \mathbb{E}\left(\sum_{k=0}^{d/2} \frac{1}{k!} \mathbf{Str} \left(\sum_{I, d(I) \leq d} \Lambda_I(B)_t \mathcal{F}_I(x_0) \right)^k \middle| \sum_{I, d(I) \leq N} \Lambda_I(B)_t e_I(x_0) = 0 \right) + O\left(\sqrt{t}\right).$$

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Since a routine computation shows that $\sum_{I,d(I)\leq d} \Lambda_I(B)_t \mathcal{F}_I(x_0)$ may be written as a linear combination of terms $a_i^* a_j$, $a_i^* a_j a_k^* a_l$, due to the formula (5.9), we have for $k \leq \frac{d}{2} - 1$,

$$\mathbf{Str}\left(\sum_{I,d(I)\leq d} \Lambda_I(B)_t \mathcal{F}_I(x_0)\right)^k = 0.$$

Consequently,

$$\mathbf{Str}\ p_{t/2}(x_0,x_0) = q_t^d(x_0)\mathbb{E}\left(\frac{1}{(d/2)!}\mathbf{Str}\left(\sum_{I,d(I)\leq d} \Lambda_I(B)_t \mathcal{F}_I(x_0)\right)^{\frac{d}{2}} \middle| \sum_{I,d(I)\leq N} \Lambda_I(B)_t e_I(x_0) = 0\right) + O\left(\sqrt{t}\right).$$

By using again the scaling property of Brownian motion, we have

$$q_t^d(x_0)\mathbb{E}\left(\frac{1}{(d/2)!}\mathbf{Str}\left(\sum_{I,d(I)\leq d} \Lambda_I(B)_t \mathcal{F}_I(x_0)\right)^{\frac{d}{2}} \middle| \sum_{I,d(I)\leq N} \Lambda_I(B)_t e_I(x_0) = 0\right)$$
$$= q_t^d(x_0)\mathbb{E}\left(\frac{1}{(d/2)!}\mathbf{Str}\ \left(t\mathcal{F}_0(x_0)\right)^{\frac{d}{2}} \middle| \sum_{I,d(I)\leq N} \Lambda_I(B)_t e_I(x_0) = 0\right) + O\left(\sqrt{t}\right).$$

We finally end up with

$$\mathbf{Str}\ p_{t/2}(x_0,x_0) = \frac{1}{(d/2)!} t^{\frac{d}{2}} q_t^d(x_0)\mathbf{Str}\ \mathcal{F}_0(x_0)^{\frac{d}{2}} + O\left(\sqrt{t}\right),$$

which proves that

$$\lim_{t\to 0} \mathbf{Str}\ p_t(x_0,x_0) = \frac{(-1)^{d/2}}{(d/2)!(4\pi)^{d/2}}\mathbf{Str}\ \mathcal{F}(x_0)^{\frac{d}{2}}.$$

Heavy, but straightforward computations (see [39], Lem. 2.35) show that

$$\frac{(-1)^{d/2}}{(d/2)!(4\pi)^{d/2}}\mathbf{Str}\ \mathcal{F}(x_0)^{\frac{d}{2}} = \omega(x_0).$$

□

References

[1] M.F. Atiyah and R. Bott, A Lefschetz fixed point formula for elliptic complexes. *I. Ann. Math.* **86**. (1967) 374–407.

[2] R. Azencott, Formule de Taylor stochastique et développements asymptotiques d'intégrales de Feynman, in *Séminaire de probabilités XVI*, edited by J. Azema, M. Yor. *Lect. Notes. Math.* **921** (1982) 237–284.

[3] R. Azencott, Densité des diffusions en temps petit: développements asymptotiques (part I), *Sem. Prob.* **18** (1984) 402–498.

[4] F. Baudoin, *An Introduction to the Geometry of Stochastic Flows*. Imperial College Press (2004).

[5] F. Baudoin, Brownian Chen series and Atiyah–Singer theorem. *J. Funct. Anal.* **254** (2008) 301–317.

[6] F. Baudoin and L. Coutin, Operators associated with a stochastic differential equation driven by fractional Brownian motions. *Stoc. Proc. Appl.* **117** (2007) 550–574.

[7] G. Ben Arous, Méthodes de Laplace et de la phase stationnaire sur l'espace de Wiener (French) [The Laplace and stationary phase methods on Wiener space]. *Stochastics* **25** (1988) 125–153.

[8] G. Ben Arous, Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus (French) [Asymptotic expansion of the hypoelliptic heat kernel outside of the cut-locus]. *Ann. Sci. Cole Norm. Sup.* **21** (1988) 307–331.

[9] G. Ben Arous, Développement asymptotique du noyau de la chaleur hypoelliptique sur la diagonale. *Ann. Inst. Fourier* **39** (1989) 73–99.

[10] G. Ben Arous, Flots et séries de Taylor stochastiques. *J. Probab. Theory Relat. Fields* **81** (1989) 29–77.

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press

Special Issue

[11] G. Ben Arous and R. Léandre, Décroissance exponentielle du noyau de la chaleur sur la diagonale. II (French) [Exponential decay of the heat kernel on the diagonal II] Probab. *Theory Relat. Fields* **90** (1991) 377−402.
[12] N. Berline, E. Getzler and M. Vergne, *Heat kernels and Dirac operators*, 2nd edition. Grundlehren Text Editions, Springer (2003).
[13] J.M. Bismut, The Atiyah–Singer Theorems: A Probabilistic Approach. *J. Func. Anal., Part I, II* **57** (1984) 329−348.
[14] N. Bourbaki, *Groupes et Algèbres de Lie*, Chap. 1–3. Hermann (1972).
[15] F. Castell, Asymptotic expansion of stochastic flows. *Probab. Theory Relat. Fields* **96** (1993) 225−239.
[16] K.T. Chen, Integration of paths, Geometric invariants and a generalized Baker-Hausdorff formula. *Ann. Math.* **65** (1957).
[17] S.S. Chern, A simple intrinsic proof of the Gauss-Bonnet theorem for closed Riemannian manifolds. *Ann. Math.* **45** (1944) 747−752.
[18] E.B. Dynkin, Calculation of the coefficients in the Campbell-Hausdorff formula. *Dodakly Akad. Nauk SSSR* **57** (1947) 323−326, in Russian, English translation (1997).
[19] M. Fliess and D. Normand-Cyrot, Algèbres de Lie nilpotentes, formule de Baker-Campbell-Hausdorff et intégrales itérées de K.T. Chen, in *Séminaire de Probabilités. Lect. Notes Math.* **920** (1982).
[20] A. Friedman, *Partial differential equations of parabolic type*. Prentice-Hall, Inc., Englewood Cliffs, NJ (1964) xiv+347.
[21] P. Friz and N. Victoir, Euler estimates for rough differential equations. *J. Differ. Equ.* **244** (2008) 388–412.
[22] P. Friz and N. Victoir, Multidimensional stochastic processes as rough paths. Theory and Applications, Cambridge Studies in *Adv. Math.* (2009).
[23] B. Gaveau, Principe de moindre action, propagation de la chaleur et estimées sous-elliptiques sur certains groupes nilpotents. *Acta Math.* **139** (1977) 95−153.
[24] E. Getzler, A short proof of the Atiyah–Singer index theorem. *Topology* **25** (1986) 111−117.
[25] P.B. Gilkey, Curvature and the eigenvalues of the Laplacian for elliptic complexes. *Adv. Math.* **10** (1973) 344−382.
[26] E.P. Hsu, Stochastic Analysis on manifolds. AMS, Providence USA. *Grad. Texts Math.* **38** (2002).
[27] Y. Inahama, *A stochastic Taylor-like expansion in the rough path theory*. Preprint from Tokyo Institute of Technology (2007)
[28] P.E. Kloeden and E. Platen, Numerical solution of stochastic differential equations. *Appl. Math.* **23** (1992).
[29] H. Kunita, Asymptotic self-similarity and short time asymptotics of stochastic flows. *J. Math. Sci. Univ. Tokyo* **4** (1997) 595–619.
[30] R. Léandre, Sur le théorème d'Atiyah–Singer. *Probab. Theory Relat. Fields* **80** (1988) 119−137.
[31] R. Léandre, Développement asymptotique de la densité d'une diffusion dégénérée. *Forum Math.* **4** (1992) 45−75.
[32] T. Lyons, Differential equations driven by rough signals. *Revista Mathemàtica Iberio Americana* **14** (1998) 215−310.
[33] T. Lyons and N. Victoir, Cubature on Wiener space. *Proc. R. Soc. Lond. A* **460** (2004) 169−198.
[34] H. McKean and I.M. Singer, Curvature and the eigenvalues of the Laplacian. *J. Differ. Geom.* **1** (1967) 43−69.
[35] P. Malliavin, Stochastic calculus of variations and hypoelliptic operators, in *Proc. of Inter. Symp. Stoch. Differ. Equ.,* Kyoto 1976, edited by Wiley (1978) 195−263.
[36] P. Malliavin, Stochastic Analysis. *Grundlehren der Mathematischen Wissenschaften* **313** (1997).
[37] V.K. Patodi, An analytic proof of the Riemann-Roch-Hirzebruch theorem. *J. Differ. Geom.* **5** (1971) 251−283.
[38] C. Reutenauer, Free Lie algebras, London Mathematical Society Monographs. *New series* **7** (1993).
[39] S. Rosenberg, The Laplacian on a Riemannian manifold. *London Mathematical Society Student Texts* **31** (1997).
[40] L.P. Rotschild and E.M. Stein, Hypoelliptic differential operators and Nilpotent groups. *Acta Math.* **137** (1976) 247−320.
[41] D. Stroock and S.R.S. Varadhan, Multidimensional diffusion processes. Springer-Verlag, Berlin, New York. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **233** (1979) xii+338.
[42] R.S. Strichartz, The Campbell-Baker-Hausdorff-Dynkin formula and solutions of differential equations. *J. Func. Anal.* **72**. (1987) 320–345.
[43] S. Takanobu, Diagonal short time asymptotics of heat kernels for certain degenerate second order differential operators of Hörmander type. *Publ. Res. Inst. Math. Sci.* **24** (1988) 169−203.
[44] M.E. Taylor, Partial Differential Equations, Basic Theory, 2nd edition. *Appl. Math.* **23** (1999)
[45] M.E. Taylor, Partial Differential Equations, Qualitative Studies of Linear Equations. *Appl. Math. Sci.* **116** (1996).

https://doi.org/10.1051/ps/2011107 Published online by Cambridge University Press