Rev. Mat. Iberoamericana **25** (2009), no. 3, 971–994

A signed measure on rough paths associated to a PDE of high order: results and conjectures

Daniel Levin and Terry Lyons

Abstract

Following old ideas of V. Yu. Krylov we consider the possibility that high order differential operators of dissipative type and constant coefficients might be associated, at least formally, with signed measures on path space in the same way that Wiener measure is associated with the Laplacian.

There are fundamental difficulties with this idea because the measure would always have locally infinite mass. However, this paper provides evidence that if one considers equivalence classes of paths corresponding to distinct parameterisations of the same path, the measures might really exist on this quotient space.

Precisely, we consider the measures on piecewise linear paths with given time partition defined using the semigroup associated to the differential operator and prove that these measures converge in distribution when the test functions on path space are the iterated integrals of the paths.

Given a "random" piecewise-linear path, we evaluate its "expected" signature in terms of an explicit tensor series in the tensor algebra. Our approach uses an integration by parts argument under very mild conditions on the polynomial corresponding to the PDE of high order.

1. Introduction

V. Yu. Krylov [14] introduced finitely additive signed measures on paths associated to some special PDEs of high order. We believe his idea has considerable unexploited potential and could ultimately produce the correct replacement for Brownian motion to discuss equations of higher order in the same way that one could discuss the second order case.

2000 Mathematics Subject Classification: 60H05, 60H15.
Keywords: High-order PDE, brownian motion, rough path, signature of a path.

Brownian motion is Markov process whose infinitesimal generator is the Laplace operator Δ. Harmonic functions are the continuous solutions to the constant coefficient PDE

$$\Delta u = 0.$$

An essential connection between Brownian motion and harmonic functions was provided by Doob [6]. Harmonic functions are the continuous functions $u[Z(t)]$ for which $Z(t)$ is a local martingale.

After Doob, the next big breakthrough connected more general PDEs and Probability and came from Itô [13] - whose Stochastic Differential Equation methods allowed one to construct new diffusions out of Brownian motion. These methods have been very powerful for the study of wide classes of second order PDEs with non-constant coefficients. For example, Stroock and Varadhan [29] used these methods to show that non-divergence form second order elliptic PDEs with continuous coefficients for the PDEs have existence and uniqueness of solutions. Malliavin [25] used the Itô perspective in his proof and extensions of Hörmander's Theorem [12]. These probabilistic methods (from Malliavin down) are now used in practical settings and, for example, significantly influence approaches to the numerical analysis of PDEs used in finance.

More recently, advances in our understanding of controlled differential equations have lead to the theory of rough paths. The essence of that theory is the study of controlled differential equations of the kind

$$\begin{aligned} dy_t &= \sum_i V^i\left(y_t\right) dx_t^i \\ y_0 &= a. \end{aligned}$$

These equations can be thought of as describing the response y of a system to a control x that varies in time. If the V^i are Lipschitz, and the control x is continuous and of bounded variation with values in a finite dimensional vector space, then it is an exercise in Picard iteration to prove that the solution y exists and is unique.

On the other hand many systems are subject to highly oscillatory external forcing. The theory of rough paths identifies novel metrics on the space of controls that makes the Itô functional $x \to y$ uniformly continuous. These metrics capture the possibility that paths may have effect if they are close in their global features, even though on the scale of microstructure they are quite different.

Key to this development was the use of the iterated integrals to provide an effective top down description of x [8] and a series of delicate analytic estimates on ordinary differential equations controlling the changes to y when the path x is varied within the class of paths with a common top down description to a certain level [24].

As a result a deterministic theory of stochastic differential equations has emerged. The dyadic piecewise linear approximations to semi-martingales, and hence to Brownian motion, are almost surely Cauchy sequences in these rough path metrics (at least for $p > 2$). Uniformly continuous functions extend to the completion and so provide a deterministic treatment of Itô-Stratonovich SDEs. Applications include the extension, in [2], of the results of Malliavin on the regularity of the density of a diffusion at a fixed time to be extended to processes driven by fractional Brownian motion with Hurst index greater than $1/4$.

We explain an application. Monte Carlo is an effective (if slow) method for solving linear second order PDEs in moderate dimensions. The very process of solution makes it clear that solving such a PDE is theoretically indistinguishable from integrating a functional of the solution to an SDE over the Wiener path space. When one integrates smooth functions one should always consider the possibility of cubature; that is to say replacing the idealised measure with a discrete measure and integrating the functional against this discrete measure. By choosing the discrete measure to integrate polynomials perfectly one can sometimes achieve excellent results with the substitute measure.

Extending this idea to path space requires that one can well approximate the Itô functional by a linear combination of special functionals analogous to polynomials. The first few terms in the signature are the appropriate analogue. It is now understood that the iterated integrals of a path are, through the signature, an effective structured tool for describing paths; moreover the expectation of the signature of a random path can fully characterise the law of the signature [7]. One chooses a finite collection of paths and weights so that the expected signature for this discrete measure matches the expected signature for the underlying Wiener process. In view of results by Kusuoka [18], Kusuoka and Stroock [15, 16, 17], Lyons and Victoir [22], these methods have produced effective higher order numerical methods for integrating second order PDEs.

This paper is about higher order operators, but it draws on this experience at second order even though the probabilities we study are not positive or defined on the full path space.

We study the "diffusions" associated with constant coefficient PDEs that are generalizations of hypo-elliptic equations in the sense of Laurent Schwartz and Lars Hörmander [12], [26]. Each equation we study has associated with it a semigroup whose kernel is in Schwartz class. If we introduce a finite partition of the time interval $[0, T]$ then it can be used with the semigroup to construct a measure on piecewise linear paths. One can then ask if there is a sense in which these measures converge (to give a "Brownian motion"). We will prove that, in a weak sense, there is a convergence result. It is compatible with our evidence that this "distributional" limit exists as a measure on rough paths without parameterization although we are not yet close to testing this. We think the positive intermediate results presented here already have some interest and might lead to effective numerical methods analogous to those described above.

Fix a partition, and consider the expected signature for a piecewise linear path chosen randomly according to the measure associated to the partition and semigroup. Our main result is that, if we take the limit as the mesh of the partition goes to zero, this expected signature converges to a non-trivial limit which is readily calculated. The limit of the expected signatures is an explicit tensor series in the tensor algebra.

The rigorous result is not so obvious and the proof relies on carefully structured logic with an integration by parts argument at its core. The total variation of the measure on piecewise linear paths associated to a partition explodes as one refines the partition. Our result can be interpreted as a statement about the convergence of these measures in a weak sense against the coordinate iterated integrals, which (following Chen [3]) are the natural test functions on the space of un-parameterized paths.

The closed form value for the limit of expectations should also prove valuable. The precise form we obtain, in the classical Brownian case, is a key step to developing the high order numerical methods of Kusuoka-Lyons-Victoir mentioned above. The explicit form suggests concrete algorithms for numerical analysis on non-constant coefficient hypo-elliptic PDEs.

2. Constant coefficient operators

If $f(x) = f(x_1, \ldots, x_d)$ is a smooth function on $\mathbb{R}^d$ then, for each $i \in (1, \ldots, d)$ we denote the differential operator $f \to \frac{\partial}{\partial x_i} f$ by D_i. Then D_i is translation invariant. As the D_i commute, any product $D_1^{i_1} \ldots D_d^{i_d}$ can be written in canonical form D^I where $I = (i_1, \ldots, i_d)$ and $i_1, \ldots, i_d$ are positive integers. We call such I multi-indices.

More generally we call any polynomial in the D_i a translation invariant differential operator on $\mathbb{R}^d$. Any such operator L can be expressed uniquely as a sum, over distinct multi-indices I

$$L = \sum_I a_I D^I,$$

where $a_I \in \mathbb{R}$ (or $\mathbb{C}$) and where D^I is interpreted to be 1 in the special case where $I = ()$. We will be concerned primarily with the case where $a_{()} = 0$ and L annihilates constant functions. The degree of L is the maximum cardinality of the multi-indices $\{I \,|a_I \neq 0\}$.

Remark 2.1. *Any translation invariant operator T can be conveniently expressed as a multiplier $Tf(x) = \left(\mathcal{M}_T \hat{f}(\xi)\right)^{\vee}(x)$ acting on the Fourier transform. We use the following convention:*

$$\begin{aligned} f &\to \hat{f} & \hat{f}(\xi) &= \int_{\mathbb{R}^d} e^{-2\pi i x.\xi} f(x)\, dx \\ g &\to \check{g} & \check{g}(x) &= \int_{\mathbb{R}^d} e^{2\pi i \xi.x} g(\xi)\, d\xi. \end{aligned}$$

In the case of L the multiplier is a polynomial.

Definition 2.2. *The Schwartz class $\mathcal{S}(\mathbb{R}^d)$ is the set of functions $f \in C^\infty\left(\mathbb{R}^d\right)$ such that*

$$\sup_{x\in\mathbb{R}^d} |x|^N \left|\left(D^I f\right)(x)\right| < \infty$$

for all integers $N \geq 0$ and all multi-indices I.

It is classical that the operators $^\wedge$ and $^\vee$ are defined on $\mathcal{S}\left(\mathbb{R}^d\right)$ and map $\mathcal{S}\left(\mathbb{R}^d\right)$ to itself.

Lemma 2.3. *Schwartz class $\mathcal{S}(\mathbb{R}^d)$ is an algebra.*

Proof. Let f and g be in $\mathcal{S}(\mathbb{R}^d)$ then

$$\begin{aligned} & |x|^N \left|\left(D^I fg\right)(x)\right| \leq \left(1+|x|^{2N}\right)\left|\left(D^I fg\right)(x)\right| \\ & \leq \left(1+|x|^{2N}\right)\left|\sum_{J\sqcup K=I} c_I \left(D^J f\right)(x)\left(D^K g\right)(x)\right| \\ & \leq 2^{|I|} \sum_{J\sqcup K=I} \left(\sup_{x\in\mathbb{R}^d}\left(1+|x|^N\right)\left|\left(D^J f\right)(x)\right|\right)\left(\sup_{x\in\mathbb{R}^d}\left(1+|x|^N\right)\left|\left(D^K g\right)(x)\right|\right) \\ & < \infty \end{aligned}$$

since the sum is finite. ■

Lemma 2.4. *The multiplier for $L = \sum_I a_I D^I$ is the polynomial*

$$\mathcal{P}(\xi) = \sum_I a_I (2\pi i \xi_1)^{i_1} \cdots (2\pi i \xi_d)^{i_d}. \tag{2.1}$$

That is to say

$$Lf(x) = \left(\mathcal{P}(\xi)\hat{f}(\xi)\right)^{\vee}(x).$$

For example if $L = -\frac{1}{2}\Delta$ then $\mathcal{P}(\xi) = 2\pi^2 \sum_{j=1}^d |\xi_j|^2$.

Remark 2.5. *This map taking constant coefficient differential operators to polynomials is an algebra isomorphism when both spaces are given the natural multiplication and addition operators.*

There is an important duality between differential operators and polynomials.

Proposition 2.6. *Consider the differential operator*

$$L := \sum_I a_I D_1^{i_1} \dots D_d^{i_d}$$

and the polynomial

$$\tilde{\mathcal{P}}(\xi) := \sum_I \tilde{a}_I \, (2\pi i \xi_1)^{i_1} \cdots (2\pi i \xi_d)^{i_d} \, .$$

Then

$$L\tilde{P}|_{\xi=0} = \tilde{L}P|_{\xi=0},$$

where

$$\mathcal{P}\,(\xi) = \sum_I a_I \, (2\pi i \xi_1)^{i_1} \cdots (2\pi i \xi_d)^{i_d}$$

and

$$\tilde{L} = \sum_I \tilde{a}_I D_1^{i_1} \dots D_d^{i_d} \, .$$

Proof. It is enough to prove this when L and $\tilde{\mathcal{P}}$ are monomials

$$\prod_k D_k^{i_k} \quad \text{and} \quad \prod_k \xi_k^{r_k}.$$

According to our notations, $\mathcal{P}$ and $\tilde{L}$ are respectively monomials

$$\prod_k (2\pi i \xi_k)^{i_k} \quad \text{and} \quad \prod_k \left(\frac{1}{2\pi i} D_k \right)^{r_k} .$$

Now one always has

$$\left(\prod_k D_k^{i_k} \right) \left(\prod_k \xi_k^{r_k} \right) = \prod_k D_k^{i_k} \xi_k^{r_k}$$

and if $i_k \neq r_k$ then

$$D_k^{i_k} \xi_k^{r_k} \big|_{\xi=0} = \left(\frac{1}{2\pi i} D_k \right)^{r_k} (2\pi i \xi_k)^{i_k} \bigg|_{\xi=0} = 0$$

and if $i_k = r_k$ then one has

$$D_k^{i_k} \xi_k^{i_k} \big|_{\xi=0} = \left(\frac{1}{2\pi i} D_k \right)^{i_k} (2\pi i \xi_k)^{i_k} \bigg|_{\xi=0} = i_k!$$

and taking the product one has the result. ∎

3. Semigroups

Definition 3.1. (Minimal growth condition). *A polynomial* $\mathcal{P}(\xi)$, $\xi \in \mathbb{R}^d$ *is said to satisfy the* minimal growth condition *if*

$$Re\ \mathcal{P}(\xi) \to +\infty, \quad |\xi| \to \infty.$$

Proposition 3.2. *If a real-valued polynomial* $\mathcal{P}(\xi)$, $\xi \in \mathbb{R}^d$ *satisfies the* minimal growth condition *then there exist* $A > 0$ *and* $\alpha > 0$ *such that*

$$\mathcal{P}(\xi) \geq A|\xi|^\alpha, \quad |\xi| \to \infty.$$

In particular,

$$\mathcal{P}(\xi)/\log|\xi| \to +\infty, \quad |\xi| \to \infty. \tag{3.1}$$

Proof. We adapt an example from [12, Example A.2.7]. The set

$$E = \left\{\left(1 + |\xi|^2, \mathcal{P}(\xi), \xi\right)\right\}$$

is a semi-algebraic set in $\mathbb{R}^{2+d}$, and so the function

$$f(x) = \inf_{1+|\xi|^2=x} \mathcal{P}(\xi) = \inf\{y; (x, y, \xi) \in E\}.$$

is also semi-algebraic [12, Corollary A.2.4]. This function is finite everywhere, positive and continuous, and converging to $+\infty$ as $x \to \infty$ since $\mathcal{P}(\xi) \to +\infty$, as $|\xi| \to \infty$. We now apply [12, Theorem A.2.5], to conclude that

$$f(x) = Ax^\alpha (1 + o(1)), \quad x \to +\infty.$$

We observe that α can not be negative or zero, otherwise the *minimal growth condition* is violated.[1] ■

Corollary 3.3. *If the polynomial* $\mathcal{P}(\xi)$, $\xi \in \mathbb{R}^d$ *satisfies the* minimal growth condition *3.1 then* $\exp(-\mathcal{P}(\xi))$ *is in* $\mathcal{S}(\mathbb{R}^d)$.

Remark 3.4. *In contrast to the one dimensional case, a polynomial* $\mathcal{P}(\xi)$, $\xi \in \mathbb{R}^d$ *can satisfy* $Re\,\mathcal{P}(\xi) \to +\infty, \quad |\xi| \to \infty$ *while* $m(r) = \inf_{|\xi|>r} Re\,\mathcal{P}(\xi)$ *can grow more slowly than any given power of* r *(consider* $\xi = (x, y) \in \mathbb{R}^2$ *and the polynomial* $(y - x^n)^2 + x^2$, $n \in \mathbb{N}$*).*

[1] We thank Lars Hormander and Boris Zilber for independently telling us how to prove this result.

Remark 3.5. *Radial behaviour is not a good indicator of the overall behaviour of a polynomial. For example, let* $\xi = (x,y) \in \mathbb{R}^2$ *and* $\mathcal{P}(\xi) = (y - x^n)^2$*, then* $Re\,\mathcal{P}(\xi)/\left|\xi\right|^{2-\varepsilon} \to \infty$ *along every ray but* $\mathcal{P}(\xi) = 0$ *on the curve* $y = x^n$*.*

Remark 3.6. *It might be worth remarking that there are other stronger conditions on polynomials (e.g. leading to Gevrey class PDEs) which were introduced by Hörmander* [26, Corollary 3.5.3] *which ensure that the principal term dominates the derivative pointwise for large* ξ*; these are very popular conditions used to capture ellipticity. Our original proofs of the main result assumed these conditions. However the integration by parts argument presented here, and suggested by Bruce Driver, avoids dependence on such pointwise estimates.*

Lemma 3.7. *If* $\mathcal{P}$ *satisfies the minimal growth condition, then the semigroup of operators* P_t *characterised by*

$$P_t : f \to (e^{-t\mathcal{P}(\xi)}\hat{f})^{\vee}$$

is well defined and continuous as a family of continuous operators on $\mathcal{S}(\mathbb{R}^d)$*.*

Proof. Now, if $f \in \mathcal{S}(\mathbb{R}^d)$ then $\hat{f} \in \mathcal{S}(\mathbb{R}^d)$. By the condition 3.1 and our remarks, $e^{-t\mathcal{P}(\xi)}$ is in $\mathcal{S}(\mathbb{R}^d)$. By Lemma 2.3 $\mathcal{S}(\mathbb{R}^d)$ is an algebra. So $e^{-t\mathcal{P}(\xi)}\hat{f}$ and hence $(e^{-t\mathcal{P}(\xi)}\hat{f})^{\vee}$ are also in $\mathcal{S}(\mathbb{R}^d)$. Each of these transformations is a continuous function from $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}(\mathbb{R}^d)$. ∎

Definition 3.8. *Let* L *be a constant coefficient differential operator on* $\mathcal{S}(\mathbb{R}^d)$ *whose multiplier is a polynomial* $\mathcal{P}$ *satisfying the minimal growth condition. If* $f \in \mathcal{S}(\mathbb{R}^d)$*, then we define* $e^{-tL}f$*, also written as* $P_t f$*, by*

$$e^{-tL}f(x) = (e^{-t\mathcal{P}(\xi)}\hat{f})^{\vee}.$$

In view of Lemma 3.7*, the operators* P_t*,* e^{-tL} *form a semigroup of continuous operators on* $\mathcal{S}(\mathbb{R}^d)$*.*

The following lemma is a consequence of our assumption that $\mathcal{P}$ satisfies the minimal growth condition 3.1.

Lemma 3.9. *The operator* P_t *can be represented as convolution with a kernel as follows:*

$$P_t f\,(x) = \int_{\mathbb{R}^d} f\,(y)\,\varphi_t\,(x-y)\,dy = \int_{\mathbb{R}^d} \varphi_t(y) f(x-y)dy,$$

where

$$\varphi_t\,(x) = \int_{\mathbb{R}^d} e^{-t\mathcal{P}(\xi)} e^{2\pi i x.\xi} d\xi$$

is in $\mathcal{S}(\mathbb{R}^d)$*.*

Proof.

$$P_t f\,(x) = (e^{-t\mathcal{P}(\xi)}\hat{f})^\vee = \int_{\mathbb{R}^d} e^{-t\mathcal{P}(\xi)} \int_{\mathbb{R}^d} f(y) e^{2\pi i (x-y).\xi} dy d\xi,$$

i.e. the kernel can be explicitly written as

$$\varphi_t(x) = \int_{\mathbb{R}^d} e^{-t\mathcal{P}(\xi)} e^{2\pi i x.\xi} d\xi.$$

■

We will use the notation φ_t frequently in the rest of the paper.

Corollary 3.10. *The operator P_t can be consistently defined for finite measures μ*

$$P_t\mu\,(x) = \int_{\mathbb{R}^d} \varphi_t\,(x-y)\,\mu\,(dy)\,,$$

and P_t acts from L^1 to $L^1 \cap L^\infty$ and it is also bounded as an operator $L^\infty \to L^\infty$.

We do not claim that this map is well behaved as $t \to 0$ or as $t \to \infty$ and in general it will not be [5].

4. The Integral Kernel acting on polynomials and its generator

This section contains the analytic result on which our main theorem depends. The semigroup P_t defined above naturally acts as a semigroup of integral operators on the Schwartz functions; it also acts on polynomials; it is easy to show that the semigroup has L as its infinitesimal generator on the space of Schwartz functions; we need to prove that the generator on polynomials is also L. Although it would be surprising if the result were not true it definitely needs a proof. It seems that the minimal growth condition is close to the natural condition on L for such a result.

One could hope to get the lemma by using the result for Schwartz functions, and identifying Schwartz functions that agree with the polynomials on large intervals. Such an approach would require tail estimates on φ_t as t goes to zero. The authors carried out this approximation method under the condition that the multiplier was in Gevrey class (see [26]). It might also be proved by using the infinitiesimal generator of the semigroup on Schwartz functions, and identifying the adjoint action. We give a direct proof. The crucial analytic component is an integration by parts.[2]

[2]We are very grateful to Bruce Driver for suggesting the outline of the integration by parts argument we give here –before his intervention we had no thought for this approach.

Lemma 4.1. (Integration by parts). *If u is a smooth function with the property that*

$$\sup_{0\leq|I|\leq N}\int\left|D_1^{i_1}\cdots D_d^{i_d}u\left(\xi\right)\right|d\xi<\infty \tag{4.1}$$

and if q is a polynomial of degree at most N then for each fixed y

$$\begin{aligned}\int_{\mathbb{R}^d} & e^{2\pi iy.\xi}q\left(-\frac{1}{2\pi i}D_1,\ldots,-\frac{1}{2\pi i}D_d\right)u\left(\xi\right)d\xi\\ &=\int_{\mathbb{R}^d}\left(q\left(\frac{1}{2\pi i}D_1,\ldots,\frac{1}{2\pi i}D_d\right)e^{2\pi iy.\xi}\right)u\left(\xi\right)d\xi\\ &=q\left(y_1,\ldots,y_d\right)\int_{\mathbb{R}^d}e^{2\pi iy.\xi}u\left(\xi\right)d\xi.\end{aligned}$$

Proof. This is a well known result. The heart of the proof is the observation that

$$-\int_{\mathbb{R}^d}D_ku\left(\xi\right)e^{2\pi iy.\xi}d\xi=\int_{\mathbb{R}^d}u\left(\xi\right)D_ke^{2\pi iy.\xi}d\xi=\int_{\mathbb{R}^d}u\left(\xi\right)2\pi iy_ke^{2\pi iy.\xi}d\xi.$$

To prove this one introduces a smooth function ϕ that is 1 on an open neighbourhood of 0 and of compact support. We put $\phi_r\left(x\right):=\phi\left(x/r\right)$ and observe that, by applying Green's theorem on a ball of large but finite radius, one knows that

$$\begin{aligned}-\int_{\mathbb{R}^d}(\phi_ru\left(\xi\right))D_ke^{2\pi iy.\xi}d\xi &=\int_{\mathbb{R}^d}D_k\left(\phi_ru\left(\xi\right)\right)e^{2\pi iy.\xi}d\xi\\ &=\int_{\mathbb{R}^d}\phi_rD_k\left(u\left(\xi\right)\right)e^{2\pi iy.\xi}d\xi+\int_{\mathbb{R}^d}u\left(\xi\right)D_k\left(\phi_r\right)e^{2\pi iy.\xi}d\xi.\end{aligned}$$

If $\int_{\mathbb{R}^d}\left|D^k\left(u\left(\xi\right)\right)\right|d\xi<\infty$ then the dominated convergence theorem implies that

$$\lim_{r\to\infty}\int_{\mathbb{R}^d}\phi_rD_k\left(u\left(\xi\right)\right)e^{2\pi iy.\xi}d\xi=\int_{\mathbb{R}^d}D_k\left(u\left(\xi\right)\right)e^{2\pi iy.\xi}d\xi$$

while $D_k\phi\left(x/r\right)=\frac{1}{r}\left(D_k\phi\right)\left(x/r\right)$ and $\left(D_k\phi\right)$ is bounded. So again, given $\int_{\mathbb{R}^d}\left|u\left(\xi\right)\right|d\xi<\infty$, using the dominated convergence theorem we see that

$$\lim_{r\to\infty}\int_{\mathbb{R}^d}u\left(\xi\right)D_k\left(\phi_r\right)e^{2\pi iy.\xi}d\xi=0.$$

■

Remark 4.2. *The hypotheses* (4.1) *of Lemma* 4.1 *are satisfied by $u\left(\xi\right)=e^{-t\mathcal{P}\left(\xi\right)}$ if the polynomial $\mathcal{P}$ satisfies the minimal growth condition.*

Remark 4.3. *It is elementary and well known that if L is a bounded linear operator on a Banach space E, the operator e^{-tL} defined by the series*

$$\sum_{n=0}^{\infty} \frac{(-tL)^n}{n!}$$

converges in operator norm in $Hom\,(E,E)$ and moreover for any $f \in E$ the limit of

$$\frac{e^{-tL} f - f}{t}$$

exists in E and equals $-Lf$.

Lemma 4.4. *Suppose L is a constant coefficient differential operator whose multiplier $\mathcal{P}$ (as defined in equation (2.1)) satisfies the minimal growth condition 3.1. Suppose further that $f(x)$ is a polynomial of degree at most N.*

1. *The differentiation operators D_k are bounded operators on polynomials of degree at most N and hence L is a bounded operator.*
2. *Therefore, $e^{-tL} f$ is well defined, through its series, as a polynomial of degree at most N.*
3. *The convergence of the series is in any norm that is finite for all polynomials of degree at most N; the resulting polynomial $e^{-tL} f$ is independent of N greater than the degree of f or any choice of norm.*
4. *The convolution $(\varphi_t * f)$ is also well defined as an integral.*
5. *For each point x*
$$(\varphi_t * f)(x) = e^{-tL} f(x).$$

Proof. By the condition (3.1) φ_t is in Schwartz class and so integrates polynomials. By the definition for convolution,

$$(\varphi_t * f)(x) = \int_{\mathbb{R}^d} f(x-y)\, \varphi_t(y) dy$$

when the right hand side makes sense. From the definition of φ_t

$$f(x-y)\, \varphi_t(y) = f(x-y) \int_{\mathbb{R}^d} e^{2\pi i y.\xi} e^{-t\mathcal{P}(\xi)} d\xi$$

and since f is a polynomial we can apply Lemma 4.1 to deduce that

$$\begin{aligned} f(x-y) &\int_{\mathbb{R}^d} e^{2\pi i y.\xi} e^{-t\mathcal{P}(\xi)} d\xi \\ &= \int_{\mathbb{R}^d} e^{2\pi i y.\xi} f\left(x_1 + \frac{1}{2\pi i} D_1, \ldots, x_d + \frac{1}{2\pi i} D_d\right) e^{-t\mathcal{P}(\xi)} d\xi. \end{aligned}$$

Let $G_x(\xi)$ be the smooth function

$$f\left(x_1+\frac{1}{2\pi i}D_1,\ldots,x_d+\frac{1}{2\pi i}D_d\right)e^{-t\mathcal{P}(\xi)}\in\mathcal{S}\left(\mathbb{R}^d\right)$$

and let $\hat{G}_x$ denote the Fourier transform of G_x. Then $\hat{G}_x$ is also in Schwartz space. From the above calculation one observes that

$$\begin{aligned}f(x-y)\,\varphi_t(y) &= \int_{\mathbb{R}^d} e^{2\pi i y.\xi}\left(f\left(x_1+\frac{1}{2\pi i}D_1,\ldots,x_d+\frac{1}{2\pi i}D_d\right)e^{-t\mathcal{P}(\xi)}\right)d\xi\\ &= \int_{\mathbb{R}^d} e^{2\pi i y.\xi}G_x(\xi)\,d\xi = \hat{G}_x(y).\end{aligned}$$

Since the inverse Fourier transform converges pointwise for functions in $\mathcal{S}\left(\mathbb{R}^d\right)$ one has $\int_{\mathbb{R}^d}\hat{G}_x(y)\,dy = G_x(0)$. Hence,

$$\begin{aligned}\int_{\mathbb{R}^d}\varphi_t(y)f(x-y)\,dy &= G_x(0)\\ &= f\left(x_1+\frac{1}{2\pi i}D_1,\ldots,x_d+\frac{1}{2\pi i}D_d\right)e^{-t\mathcal{P}(\xi)}\bigg|_{\xi=0}.\end{aligned}$$

So it is enough to prove that for any polynomial f and $x=(x_1,\ldots,x_d)$ one has

$$f\left(x_1+\frac{1}{2\pi i}D_1,\ldots,x_d+\frac{1}{2\pi i}D_d\right)e^{-t\mathcal{P}(\xi)}\bigg|_{\xi=0} = e^{-tL}f(x).$$

where all expressions are interpreted naively. Since L is translation invariant, one may change variable so that $x=0$ and f is another polynomial. Because all the expressions are linear we can treat only the monomial case. It is enough to prove that

$$\left(\frac{1}{2\pi i}D_1\right)^{k_1}\cdots\left(\frac{1}{2\pi i}D_d\right)^{k_d}e^{-t\mathcal{P}(\xi)}\bigg|_{\xi=0} = e^{-tL}(x_1)^{k_1}\cdots(x_d)^{k_d}\Big|_{x=0}.$$

It suffices to prove that for all r

$$\left(\frac{1}{2\pi i}D_1\right)^{k_1}\cdots\left(\frac{1}{2\pi i}D_d\right)^{k_d}\mathcal{P}(\xi)^r\bigg|_{\xi=0} = L^r(x_1)^{k_1}\cdots(x_d)^{k_d}\Big|_{x=0}.$$

The map $L\to\mathcal{P}$ is an algebra map (Remark 2.5), so one sees that it is enough to prove this identity for general L and $\mathcal{P}$ taking $r=1$. This case follows from Proposition 2.6. ■

Corollary 4.5. *If L is a constant coefficient differential operator whose multiplier $\mathcal{P}$ satisfies the minimal growth condition* 3.1*, P_t is the integral operator defined in Lemma* 3.9 *and f is a polynomial then*

$$\lim_{t\to 0}\frac{(P_t f - f)(x)}{t} = (-Lf)(x),$$

and indeed the limit

$$\lim_{t\to 0}\frac{(P_t f - f)}{t} = (-Lf)$$

exists in any norm that is bounded on polynomials of bounded degree.

Proof. Now, by definition, for each x, $P_t f\,(x) = (\varphi_t * f)\,(x)$. By the previous lemma the function $(\varphi_t * f)$ is a polynomial and equals $e^{-tL} f\,(x)$. Fix some integer N greater than the degree of f. By elementary functional analysis, the limit

$$\lim_{t\to 0}\frac{e^{-tL} f - f}{t}$$

exists in any norm on the space of polynomials of degree N that makes L a bounded operator. For example

$$\|g\| = \sup_{x\in\mathbb{R}^d}\frac{|g\,(x)|}{1 + |x|^{N+1}}$$

is such a norm and in particular $\frac{e^{-tL} f - f}{t}$ converges to $-Lf$ locally uniformly and pointwise. ■

Remark 4.6. *Linear operators T, such as P_t and L, are initially defined on scalar valued functions. However, they can be uniquely extended to (finite dimensional) vector valued functions so that $\langle e^*, Tx\rangle := T\,\langle e^*, x\rangle$ for every e^* in the dual space. We do this without further remark.*

Corollary 4.7. *For $x \in \mathbb{R}^d$, let $\mathbf{g}_n(x) = \underbrace{x \otimes \cdots \otimes x}_{n} \in \mathbb{R}^{dn}$ then*

$$\lim_{t\to 0}\frac{(P_t\mathbf{g}_n - \mathbf{g}_n)(x)}{t} = (-L\mathbf{g}_n)(x),$$

for every $x \in \mathbb{R}^d$. The convergence to the limit is locally uniform.

Proof. Fix a basis for the dual space to $\mathbb{R}^d$ and use the words ω of length of n in these basis elements to be the dual E^* of $\underbrace{\mathbb{R}^d \otimes \cdots \otimes \mathbb{R}^d}_{n}$. The limit

$$\lim_{t\to 0}\frac{(P_t\mathbf{g}_n - \mathbf{g}_n)}{t}$$

exists if and only if

$$\left\langle \omega, \lim_{t\to 0} \frac{(P_t \mathbf{g}_n - \mathbf{g}_n)}{t} \right\rangle = \lim_{t\to 0} \frac{P_t \langle \omega, \mathbf{g}_n \rangle - \langle \omega, \mathbf{g}_n \rangle}{t}$$

exists for all ω. But $\langle \omega, \mathbf{g}_n \rangle$ is a scalar polynomial of degree n and so the second limit exists by the previous result. ∎

5. Measures on paths

Definition 5.1. *Let $D = S \leq t_0 < t_1 < \cdots < t_n \leq T$ be a partition of the interval $[S,T]$. Then μ_D is the unique Borel measure on $\left(\mathbb{R}^d\right)^n$ which satisfies, for cylinder sets $C = A_1 \times \ldots \times A_n$, A_i a Borel subset of $\mathbb{R}^d$,*

$$\begin{aligned} \mu_D(C) \quad = \quad & \underset{x_1\in A_1,\ldots,x_n\in A_n}{\int\cdots\int} \varphi_{t_1-t_0}(x_1)\, \varphi_{t_2-t_1}(x_2 - x_1) \ldots \\ & \qquad\qquad \varphi_{t_n-t_{n-1}}(x_n - x_{n-1})\, dx_1 dx_2 \ldots dx_n. \end{aligned}$$

The semigroup property ensures that the measures μ_D are consistent as one enlarges D etc., and V. Yu. Krylov regards the family as defining a "finitely additive" measure $\mathbb{P}$ on paths:

$$\mathbb{P}\left(X_{t_i} \in A_i,\ i \in [1,n]\right) = \mu_D(C)$$

and one can consult [14] (see also Hochberg [9]) for details. However the total variation of μ_D typically increases to infinity exponentially in $\#D$ as one refines the partition. In this paper, we work directly with finite measures derived from μ_D and supported on the space $PL[0,T]$ of all continuous piecewise linear paths on $[0,T]$.

Definition 5.2. *Let $D = S \leq t_0 < t_1 < \cdots < t_n \leq T$ be a partition of $[S,T]$ and $(x_0,\ldots,x_n) \in (\mathbb{R}^d)^{n+1}$. Define $\pi_D\left((x_0,\ldots,x_n)\right)$ to be the path $\gamma \in PL[-\infty,\infty]$ defined on $[t_{j-1},t_j]$ by*

$$\gamma_t = \frac{(t_j - t)x_{j-1} + (t - t_{j-1})x_j}{t_j - t_{j-1}}, \qquad t \in [t_{j-1}, t_j]$$

and for t in $[-\infty, t_0]$ set $\gamma_t = x_0$ and in $[t_n, \infty]$, set $\gamma_t = x_n$.

Definition 5.3. *$\mathbb{P}_D$ is the measure $\pi_D\left(\mu_D\right)$.*

For every D, $\mathbb{P}_D$ is a signed measure supported on the space $PL[0,T]$ of continuous piecewise linear paths.

Definition 5.4. *If f is a Borel measurable function defined on $PL[0,T]$ then we define $\mathbb{E}_D(f)$ to be the integral $\int_{\omega\in PL[0,T]} f(\omega)\,\mathbb{P}_D(d\omega)$.*

Note that $\mathbb{E}_D(1) = 1$, and $\mathbb{P}_D$ is a finite measure. In general it is not positive, and the total variation of $\mathbb{P}_D$ is greater than 1.

6. The Signature of a path

We give several auxiliary algebraic definitions which will be used later.

Definition 6.1. *Let V be a vector space. Then we denote the space of formal power series over V (where $V^{\otimes 0}$ is $\mathbb{R}$, the field of scalars) by:*

$$T((V)) = \bigoplus_{k=0}^{\infty} V^{\otimes k}.$$

Then T is an associative algebra with unit.

If $\mathbf{a} = (a_0, a_1, a_2, \dots)$, $\mathbf{b} = (b_0, b_1, b_2, \dots)$, are two elements of T, where $a_k,\ b_k \in V^{\otimes k}$, $\lambda \in \mathbb{R}$ then the sum, tensor product and the action of scalars are given by

$$\begin{aligned}
\mathbf{a} + \mathbf{b} &= (a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots), \\
\mathbf{a} \otimes \mathbf{b} &= \left(\dots, \sum_{j=0}^{k} a_j \otimes b_{k-j}, \dots \right), \\
\lambda \mathbf{a} &= (\lambda a_0, \lambda a_1, \lambda a_2, \dots).
\end{aligned}$$

The exponential $\exp(\mathbf{a})$ of a tensor $\mathbf{a}$ is the formal power series

$$\exp(\mathbf{a}) = \sum_{n=0}^{\infty} \frac{\mathbf{a}^{\otimes n}}{n!}.$$

Definition 6.2. *The truncated tensor algebra $T^{(n)}$ is defined to be*

$$T^{(n)}(V) = \bigoplus_{k=0}^{n} V^{\otimes k},$$

equipped with the product for $\mathbf{a} = (a_0, a_1, \dots a_n)$, $\mathbf{b} = (b_0, b_1, \dots, b_n)$,

$$\mathbf{a} \otimes \mathbf{b} = \left(a_0 b_0, \dots, \sum_{j=0}^{n} a_j \otimes b_{n-j} \right).$$

We use π_n to denote the natural projection of $T((V))$ into $T^{(n)}(V)$. π_n is an algebra homomorphism.

Recall the definition for the signature of a path.

Definition 6.3. *Let* $0 \le s < t \le T$. *If* γ *is a continuous path of finite length in a Banach space then the series*

$$\mathbf{X}_{s,t}(\gamma) = \sum_{k=0}^{\infty} \int_{s<t_1<\cdots<t_k<t} d\gamma(t_1) \otimes \cdots \otimes d\gamma(t_k) \tag{6.1}$$

of iterated integrals is called the signature of the path γ *over the interval* $[s,t]$.

Remark 6.4. *In the sequel, we use* $\mathbf{X}$ *as a shorter notation for* $\mathbf{X}_{0,1}$.

One of the most important properties of the signature is that it is a homomorphism from the space of paths of finite length with concatenation to the tensor series with the tensor product as multiplication (for the proof see [3]). In particular,

Lemma 6.5. (Chen's Theorem). *The signature is multiplicative in the sense that for any* $u \in [s,t]$:

$$\mathbf{X}_{t,u}(\gamma) \otimes \mathbf{X}_{u,s}(\gamma) = \mathbf{X}_{t,s}(\gamma). \tag{6.2}$$

For the proof of the above lemma see [24] or [3].

7. The main result: Theorem 7.4

We now want to consider the $\mathbb{E}_D-$expectation of the signature of a path. This is a fundamental transform from measures on rectifiable paths into elements of the tensor algebra. Under certain restrictions it has been shown to determine the measure itself [7].

Remark 7.1. *As above, let* $\mathbf{g}_n(x) = \underbrace{x \otimes \cdots \otimes x}_{n}$, *and define*

$$L\Big(\underbrace{x \otimes \cdots \otimes x}_{n}\Big)\Big|_{x=0} = (L\mathbf{g}_n)(0).$$

Definition 7.2. *Henceforth* $D = \{0 \le t_0 \le \cdots \le t_r \le 1\}$ *is a partition of* $[0,1]$ *and*

$$\#D = \max\{t_{i+1} - t_i,\ i = 0, \ldots, r-1\}.$$

$\#D$ is an indicator of the refinement of the partition D.

Consider again the measures μ_D on paths introduced in Section 5.

Remark 7.3. *It is obviously an interesting question to ask if the measures* $\mathbb{P}_D$ *converge in some sense as* $\#D \to 0$. *We have the following positive theorem about weak existence:*

Theorem 7.4. *The limit, as we refine D, of the $\mathbb{P}_D$-expectation of the signature (as a function on path space) exists and can be computed explicitly in terms of the constant coefficient differential operator L:*

$$\mathbb{E}(\mathbf{X}) = \lim_{\#D\to 0} \mathbb{E}_D(\mathbf{X}) = \exp\left(-\sum_{n=0}^{\infty} \frac{1}{n!} L\Big(\underbrace{x\otimes\cdots\otimes x}_{n}\Big)\Big|_{x=0}\right).$$

Example 7.5. *Let $L = -\frac{1}{2}\sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}$. Then $L(x_{i_1}\dots x_{i_n})|_{(x_1,\dots,x_d)=0} = 0$ unless $n = 2$ and $i_1 = i_2$. Moreover, $L(x_i x_i)|_{(x_1,\dots,x_d)=0} = -1$. Thus, for the case where γ is distributed like a Brownian path defined on $[0,1]$ we recover the result of Fawcett [7] or Lyons-Victoir [22]:*

$$\begin{aligned}
\mathbb{E}(\mathbf{X}(\gamma)) &= \exp\left(-\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{i_1=1,\dots,i_n=1}^{d} L(x_{i_1}\dots x_{i_n})|_{(x_1,\dots,x_d)=0}\ e_{i_1}\otimes\dots\otimes e_{i_n}\right)\\
&= \exp\left(-\frac{1}{2}\sum_{i=1}^{d} L(x_i^2)\big|_{(x_1,\dots,x_d)=0}\, e_i\otimes e_i\right)\\
&= \exp\left(\frac{1}{2}\sum_{i=1}^{d}\ e_i\otimes e_i\right),
\end{aligned}$$

where e_i, $i = 1,\dots,d$ are an orthonormal basis for $\mathbb{R}^d$.

Now we proceed with the proof of our main result.

Definition 7.6. *Consider the ideal $T_0\left(\left(\mathbb{R}^d\right)\right)$ in $T\left(\left(\mathbb{R}^d\right)\right)$ of elements*

$$0\oplus\mathbb{R}^d\oplus\left(\mathbb{R}^d\otimes\mathbb{R}^d\right)\oplus\cdots\oplus\left(\mathbb{R}^d\otimes\cdots\otimes\mathbb{R}^d\right)\oplus\cdots$$

with zero constant term. For every x in $T_0\left(\left(\mathbb{R}^d\right)\right)$, the series

$$\exp(x) = 1 + x + \frac{x\otimes x}{2!} + \cdots$$

involves only finitely many terms with a given tensor degree, therefore it converges and so defines a map from $T_0\left(\left(\mathbb{R}^d\right)\right)$ to $T\left(\left(\mathbb{R}^d\right)\right)$. We identify $\mathbb{R}^d$ with $0\oplus\mathbb{R}^d\oplus 0\oplus 0\oplus\cdots\subset T_0\left(\left(\mathbb{R}^d\right)\right)$ and denote the restriction of exp *to this $\mathbb{R}^d$ by $\mathbf{g}$, so that if $e\in\mathbb{R}^d$ then*

$$\mathbf{g}(e) = \exp(e) = 1 + e + \frac{e\otimes e}{2!} + \cdots$$

Note that $\mathbf{g}$ is a tensor-valued function defined on $\mathbb{R}^d$.

Proof. Let $D = \{0 = t_0 < t_1 < \cdots < t_n = 1\}$ be a partition of $[0,1]$, and recall the definition of the measure $\mathbb{P}_D$ (definition 5.3) and its expectation operator $\mathbb{E}_D$ (definition 5.4) in terms of the semigroup P_t. We extend $\mathbb{E}_D$ to functions that take their values in the infinite tensor series (such as $(\mathbf{X}_{t,u})$ and $\mathbf{g}(\gamma_t)$) by integrating the terms in the series term by term to produce a new series.

Recall from the spatial translation invariance and the semigroup property of P_t that, for any D with $u \in D$, one has independence of $(\mathbf{X}_{t,u})_{t<u}$ and $(\mathbf{X}_{u,s})_{s>u}$ over the measure $\mathbb{P}_D$. By Lemma 3.9 the kernel φ_t for P_t is in Schwartz class. If the path γ is linear off D, then each term in its signature $\mathbf{X}_{t,u}$ is a polynomial in the values of $\gamma(t_i)$, $t_i \in D$, so it is obvious that $\mathbb{E}_D(\mathbf{X}_{t,u})$ is defined. Using Lemma 6.5, we have

$$\mathbf{X}_{t,u} \otimes \mathbf{X}_{u,s} = \mathbf{X}_{t,s}.$$

Suppose that $u \in D$. Then the independence mentioned above ensures that

$$\mathbb{E}_D(\mathbf{X}_{t,u}) \otimes \mathbb{E}_D(\mathbf{X}_{u,s}) = \mathbb{E}_D(\mathbf{X}_{t,s})$$

and

$$\mathbb{E}_D(\mathbf{X}) = \mathbb{E}_D(\mathbf{X}_{t_0,t_n}) = \mathbb{E}_D(\mathbf{X}_{t_0,t_1}) \otimes \cdots \otimes \mathbb{E}_D\left(\mathbf{X}_{t_{n-1},t_n}\right). \tag{7.1}$$

Since γ is linear on the interval (t_i, t_{i+1}), i.e. $\gamma_u = \gamma_{t_i} + (u - t_i)e$, $e \in \mathbb{R}^d$:

$$\begin{aligned}
\mathbf{X}_{t_i,t_{i+1}}(\gamma) &= 1 + \sum_{k=1}^{\infty} \int_{t_i<u_1<\cdots<u_k<t_{i+1}} d\gamma_{u_1} \otimes \cdots \otimes d\gamma_{u_k} \\
&= 1 + \sum_{k=1}^{\infty} \int_{t_i<u_1<\cdots<u_k<t_{i+1}} e\,du_1 \otimes \cdots \otimes e\,du_k \\
&= 1 + \sum_{k=1}^{\infty} e^{\otimes k} \int_{t_i<u_1<\cdots<u_k<t_{i+1}} du_1 \ldots du_k \\
&= 1 + \sum_{k=1}^{\infty} \frac{e^{\otimes k}}{k!}(t_{i+1} - t_i)^k \\
&= 1 + \sum_{k=1}^{\infty} \frac{1}{k!}\left(\gamma_{t_{i+1}} - \gamma_{t_i}\right)^{\otimes k} = \exp\left(\gamma_{t_{i+1}} - \gamma_{t_i}\right).
\end{aligned} \tag{7.2}$$

Where, here and in rest of this section

$$(\gamma_t - \gamma_s)^{\otimes k} := \underbrace{(\gamma_t - \gamma_s) \otimes \cdots \otimes (\gamma_t - \gamma_s)}_{k}.$$

Further, by equation (7.1) and by equation (7.2),

$$\mathbb{E}_D(\mathbf{X}) = \mathbb{E}_D(\exp(\gamma_{t_1} - \gamma_{t_0}))\,\mathbb{E}_D(\exp(\gamma_{t_2} - \gamma_{t_1})) \cdots \mathbb{E}_D\left(\exp\left(\gamma_{t_n} - \gamma_{t_{n-1}}\right)\right).$$

Let $D_i = \{t_i, t_{i+1}\}$ and $\tilde{D}_i = \{0, t_{i+1} - t_i\}$. Then $\exp\left(\gamma_{t_{i+1}} - \gamma_{t_i}\right)$ is $\sigma\{\gamma_t | t \in D_i\}$ measurable and $D_i \subset D$ so that

$$\mathbb{E}_D\left(\exp\left(\gamma_{t_{i+1}} - \gamma_{t_i}\right)\right) = \mathbb{E}_{D_i}\left(\exp\left(\gamma_{t_{i+1}} - \gamma_{t_i}\right)\right)$$

and translation-invariance leads one to

$$\begin{aligned}
\mathbb{E}_{D_i}\left(\exp\left(\gamma_{t_{i+1}} - \gamma_{t_i}\right)\right) &= \mathbb{E}_{\{0,\tau\}}\left(\exp\left(\gamma_\tau - \gamma_0\right)\right), \quad \tau = t_{i+1} - t_i, \\
&= \mathbb{E}_{\{0,\tau\}}\left(\exp\left(\gamma_\tau\right) | \gamma_0 = 0\right) \\
&= P_\tau\left(\mathbf{g}\left(x\right)\right)(0) \\
&= \sum_{k=1}^{\infty} \frac{1}{k!} P_\tau(x^{\otimes k})(0),
\end{aligned}$$

where $\{0, \tau\}$ is the trivial partition of $[0, \tau]$ and, by our previous remarks, the sum $\sum_{k=1}^{\infty} \frac{1}{k!} P_\tau(x^{\otimes k})(0)$ is well defined as a tensor series. Now, for $k = 0, \ldots, n-1$, define $\tau_k = t_{k+1} - t_k$; then $\sum_{k=0}^{n-1} \tau_k = 1$. Then, using the multiplicative property, we have

$$\mathbb{E}_D\left(\mathbf{X}\right) = \prod_{k=0}^{n-1} P_{\tau_k}\left(\mathbf{g}\left(x\right)\right)(0).$$

Let us consider the truncation $\mathbb{E}_D^{(N)}\left(\mathbf{X}\right)$ of $\mathbb{E}_D\left(\mathbf{X}\right)$ etc. Recall that P_τ preserves tensor degree (and also the space of polynomials). Then, working in $T^{(N)}\left(\mathbb{R}^d\right)$, obtained from $T\left(\left(\mathbb{R}^d\right)\right)$ by quotienting out the ideal of tensors of tensor degree strictly greater than N, one has

$$\mathbb{E}_D^{(N)}\left(\mathbf{X}\right) = \prod_{k=0}^{n-1} P_{\tau_k}\left(\mathbf{g}^{(N)}\right)(0).$$

We recall from Lemma 4.4 that for any x_0 and the polynomial $\mathbf{g}_n(x) = \underbrace{x \otimes \cdots \otimes x}_{n}$ one has

$$\lim_{t \to 0} \frac{(P_t \mathbf{g}_n - \mathbf{g}_n)(x_0)}{t} = (-L\mathbf{g}_n)(x_0).$$

Since $T^{(N)}\left(\mathbb{R}^d\right)$ is a finite dimensional algebra and L is a bounded linear operator on $T^{(N)}\left(\mathbb{R}^d\right)$ so the exponential function is well defined and the following identity

$$\lim_{t \to 0}\left(\frac{\left(\exp\left(-tL\right) - I\right)}{t} \mathbf{g}_n\right)(x_0) = (-L\mathbf{g}_n)(x_0)$$

is routine. Both of these identities extend to linear combinations of the g_n (such as $g^{(N)}$); x_0 can be set to 0. We may choose a norm $\|\cdot\|_{T^{(N)}\left(\mathbb{R}^d\right)}$ making it a Banach algebra. Taking the difference and setting

$$h(t) = \left\|P_t\left(\mathbf{g}^{(N)}\right)(0) - \left(\exp(-tL)\,\mathbf{g}^{(N)}\right)(0)\right\|_{T^{(N)}\left(\mathbb{R}^d\right)}$$

it follows that $h(t)/t \to 0$ as $t \to 0$. Because the space of polynomials of degree at most N is finite dimensional the linear map $\mathbf{v} \to L\mathbf{v}$, defined on $T^{(N)}\left(\mathbb{R}^d\right)$, must be bounded. Let its norm be K_N. Similarly, P_t is a linear semigroup on the finite dimensional space $T^{(N)}\left(\mathbb{R}^d\right)$ and so, by routine analysis $\left\|P_t\left(\mathbf{g}^{(N)}\right)(0)\right\| \leq e^{t\tilde{K}_N\|\mathbf{g}^{(N)}\|}$ for some $\tilde{K}$ and in particular we have for any m that

$$\left\|\prod_{k=0}^{m} P_{\tau_k}\left(\mathbf{g}^{(N)}\right)(0)\right\|_{T^{(N)}\left(\mathbb{R}^d\right)} \leq e^{\tilde{K}_N\left\|\mathbf{g}^{(N)}\right\|_{T^{(N)}\left(\mathbb{R}^d\right)}\sum_{k=0}^{m}\tau_k}.$$

Similarly we have for any m that

$$\left\|\prod_{k=m+1}^{n-1} \exp\left(-\tau_k L\mathbf{g}^{(N)}\right)(0)\right\|_{T^{(N)}\left(\mathbb{R}^d\right)} \leq e^{K_N\left\|\mathbf{g}^{(N)}\right\|_{T^{(N)}\left(\mathbb{R}^d\right)}\sum_{k=m+1}^{n-1}\tau_k}$$

and so, using the identity

$$\prod_{i=1}^{N} A_i - \prod_{i=1}^{N} B_i = \sum_{r=1}^{N}\left(\left(\prod_{i=1}^{r-1} A_i\right)(A_r - B_r)\left(\prod_{i=r+1}^{N} B_i\right)\right)$$

it follows that the difference

$$\begin{aligned}
&\left\|\prod_{k=0}^{n-1} \exp\left(-\tau_k L\mathbf{g}^{(N)}\right)(0) - \prod_{k=0}^{n-1} P_{\tau_k}\left(\mathbf{g}^{(N)}\right)(0)\right\|_{T^{(N)}\left(\mathbb{R}^d\right)} \\
&\leq e^{\max\left\{K_N,\tilde{K}_N\right\}\|\mathbf{g}^{(N)}\|}\sum_{k=0}^{n-1}\left\|P_{\tau_k}\left(\mathbf{g}^{(N)}\right)(0) - \exp\left(-\tau_k L\mathbf{g}^{(N)}\right)(0)\right\|_{T^{(N)}\left(\mathbb{R}^d\right)} \\
&\leq e^{\max\left\{K_N,\tilde{K}_N\right\}\|\mathbf{g}^{(N)}\|}\sum_{k=0}^{n-1} h(\tau_k) \\
&\leq e^{\max\left\{K_N,\tilde{K}_N\right\}\|\mathbf{g}^{(N)}\|}\sum_{k=0}^{n-1} \frac{h(\tau_k)}{\tau_k}\tau_k \\
&\leq \left(\max_{k=0\ldots n-1}\frac{h(\tau_k)}{\tau_k}\right) e^{\max\left\{K_N,\tilde{K}_N\right\}\|\mathbf{g}^{(N)}\|}
\end{aligned}$$

and $\frac{h(\tau_k)}{\tau_k} \to 0$ since the max $\tau_k \to 0$.

Finally, let $\tau_k \to 0$, $k = 0 \dots n-1$

$$\begin{aligned}
\lim_{\#D\to 0} \mathbb{E}_D^{(N)}(\mathbf{X}) &= \lim_{\max_{k=0\dots n-1}\tau_k \to 0} \prod_{k=0}^{n-1} P_{\tau_k}\left(\mathbf{g}^{(N)}\right)(0) \\
&= \lim_{\max_{k=0\dots n-1}\tau_k \to 0} \prod_{k=0}^{n-1} \exp\left(-\tau_k L\mathbf{g}^{(N)}\right)(0) \\
&= \exp\left(-\sum_{k=0}^{n-1} \tau_k L\mathbf{g}^{(N)}(0)\right) = \exp\left(-L\mathbf{g}^{(N)}(e)\Big|_{e=0}\right).
\end{aligned}$$

Since $L\mathbf{g}^{(N)}(e)\big|_{e=0}$ is eventually constant as $N \to \infty$ and at that point agrees with $L\mathbf{g}(e)|_{e=0}$ the proof of Theorem 7.4 is finished. ■

Remark 7.7. *The identity* (6.2) *holds for any $t < u < s$ and hence if we denote by $f(t,u) = \mathbb{E}(\mathbf{X}_{t,u})$, and exploit time-invariance, it yields*

$$f(t-u)f(u-s) = f(t-s).$$

Solving this functional equation, $\mathbb{E}(\mathbf{X}_{t,u}) = e^{(t-u)\alpha}$. So,

$$\alpha = \frac{d}{du}|_{u=0}\mathbb{E}(\mathbf{X}_{0,u}).$$

8. Open questions

1. We have explained in Section 5 that the semigroup e^{-tL} defines a measure π_D on piecewise linear paths partitioned at the times D. The signature is a map defined on all rough paths and we can ask about the existence of limits. We have observed that, in general, the total variation norm of π_D explodes with $\#D \to 0$. However, we have also established that there are many test functions (coordinate iterated integrals) ϕ for which

   $$\lim_{\#D\to 0} \int \phi(\gamma)\, \pi_D(d\gamma)$$

   exists. So it seems that a sort of distributional limit of the π_D does indeed exist. Now all these test functions are defined on the full space of rough paths, and also ignore the parameterisation of the path γ. So perhaps, it could happen that the measures π_D do converge to a finite signed measure on rough paths when one considers the filtration that ignores parameterisation. The paper of Hambly and Lyons [8] proves in a precise way that two paths of finite variation with the same coordinate iterated integrals are in fact tree-like re-parameterisations of one another. It seems reasonable, although it is still open, that this extends to p-rough paths for every $p > 1$. Even though we know all the moments of this "measure", at the time of writing, it does not seem straightforward to determine the existence or otherwise of a measure with these moments!

2. It is possible to ask for less than we do in 1. We can consider the push forwards $S(\pi_D)$ of the measures π_D into the tensor sequences $T\left(\left(\mathbb{R}^d\right)\right)$ and into the truncated tensor algebra $T^{(n)}\left(\left(\mathbb{R}^d\right)\right)$. For each D this measure $\pi_D^{(n)}$ is a well defined finite signed measure on $T^{(n)}\left(\left(\mathbb{R}^d\right)\right)$. We have proved for every n that every moment of $\pi_D^{(n)}$ has a limit and this limit is explicit. Does a finite measure μ_n on $T^{(n)}\left(\left(\mathbb{R}^d\right)\right)$ exist with these moments, and in this case does its total variation norm remain bounded as $n \to \infty$. Optimistically
$$\lim_{n\to\infty} \mu_n = \mu \in Meas\left(T\left(\left(\mathbb{R}^d\right)\right)\right)$$
and, for each L there a $p > 1$ s.t.$|\mu|\left(S\left(p\text{-rough paths}\right)^c\right) = 0$.

3. Suppose that V^i are vector fields, and I_V is the Itô Map taking paths in $\mathbb{R}^d$ to somewhere else via
$$dy = V^i(y)\, d\gamma,\ y_0 = a$$
then we can define an operator
$$v^D(a) := \mathbb{E}_D\left(f\left(y_T\right)\right)$$
and ask if it converges for smooth enough f and if the resulting function v solved the non-constant coefficient PDE (see [22]).

4. It would be interesting if one could use the knowledge of the expectation of the signature to extend the Kusuoka-Lyons-Victoir approach to give numerical and analytic approximations to the solutions of non-constant coefficient PDEs.

References

[1] BEGHIN, L., HOCHBERG, K. J. AND ORSINGHER, E.: Conditional maximal distributions of processes related to higher-order heat-type equations. *Stochastic Process. Appl.* **85** (2000), no. 2, 209–223.

[2] CASS, T. AND FRIZ, P.: Densities for rough differential equations under Hörmander's condition. To appear in *Ann. of Math.*

[3] Chen, K. T.: Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Ann. of Math. (2)* **65** (1957), 163–178.

[4] Davies, E. B.: *Spectral theory and differential operators.* Cambridge Studies in Advanced Mathematics **42**. Cambridge Univ. Press, Cambridge, 1995.

[5] Davies, E. B.: Long time asymptotics of fourth order parabolic equations. *J. Anal. Math.* **67** (1995), 323–345.

[6] Doob, J. L.: Semimartingales and subharmonic functions. *Trans. Amer. Math. Soc.* **77** (1954), 86–121.

[7] Fawcett, T.: *Problems in stochastic analysis. Connections between rough paths and non-commutative harmonic analysis.* Ph.D. Thesis, Oxford University, 2003.

[8] Hambly, B. M. and Lyons, T. J.: Uniqueness for the signature of a path of bounded variation and the reduced path group. To appear in *Ann. of Math.*

[9] Hochberg, K. J.: A signed measure on path space related to Wiener measure. *Ann. Probab.* **6** (1978), no. 3, 433–458.

[10] Hochberg, K. J.: Central limit theorem for signed distributions. *Proc. Amer. Math. Soc.* **79** (1980), no. 2, 298–302.

[11] Hochberg, K. J. and Orsingher, E.: Composition of stochastic processes governed by higher-order parabolic and hyperbolic equations. *J. Theoret. Probab.* **9** (1996), no. 2, 511–532.

[12] Hörmander, L.: *The analysis of linear partial differential operators. II. Differential operators with constant coefficients.* Grundlehren der Mathematischen Wissenschaften **257**. Springer-Verlag, Berlin, 1983.

[13] Itô, K.: Stochastic integral. *Proc. Imp. Acad. Tokyo* **20** (1944), 519–524.

[14] Krylov, V. J.: Some properties of the distribution corresponding to the equation $\partial u/\partial t = (-1)^{q+1}\partial^{2q}u/\partial x^{2q}$. *Dokl. Akad. Nauk SSSR* **132** 1254–1257; translated as *Soviet Math. Dokl.* **1** (1960), 760–763.

[15] Kusuoka, S. and Stroock, D.: Applications of the Malliavin calculus. III. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **34** (1987), no. 2, 391–442.

[16] Kusuoka, S. and Stroock, D.: Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **32** (1985), no. 1, 1–76.

[17] Kusuoka, S. and Stroock, D.: Applications of the Malliavin calculus. I. In *Stochastic analysis (Katata/Kyoto, 1982)*, 271–306. North-Holland Math. Library **32**. North-Holland, Amsterdam, 1984.

[18] Kusuoka, S.: Malliavin calculus revisited. *J. Math. Sci. Univ. Tokyo* **10** (2003), no. 2, 261–277.

[19] Levin, D. and Wildon, M.: A combinatorial method for calculating the moments of Lévy area. *Trans. Amer. Math. Soc.* **360** (2008), no. 12, 6695-6709.

[20] Lévy, P.: Wiener's random function, and other Laplacian random functions. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, 171–187. Univ. California Press, Berkeley and Los Angeles, 1951.

[21] Lojasiewicz, S.: *Ensembles semi-analytiques.* Cours Faculté des Sciences d'Orsay, Mimeographié I.H.E.S. Bures-sur-Yvette, July, 1965.

[22] Lyons, T. and Victoir, N.: Cubature on Wiener space. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **460** (2004), no. 2041, 169–198.

[23] Lyons, T. J.: Differential equations driven by rough signals. *Rev. Mat. Iberoamericana* **14** (1998), no. 2, 215–310.

[24] Lyons, T. J., Caruana, M. and Lévy, T.: *Differential equations driven by rough paths.* Lecture Notes in Mathematics **1908**. Springer, Berlin, 2007.

[25] Malliavin, P.: Infinite-dimensional analysis. *Bull. Sci. Math.* **117** (1993), no. 1, 63–90.

[26] Mascarello, M. and Rodino, L.: *Partial differential equations with multiple characteristics.* Math. Topics **13**. Akademie Verlag, Berlin, 1997.

[27] Ree, R.: Lie elements and an algebra associated with shuffles. *Ann. of Math. (2)* **68** (1958), 210–220.

[28] Reutenauer, C.: Free Lie algebras. London Mathematical Society Monographs, New Series **7**. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1993.

[29] Stroock, D. and Varadhan, S.: *Multidimensional diffusion processes.* Classics in Mathematics. Springer-Verlag, Berlin, 2006.

Recibido: 23 de enero de 2008
Revisado: 24 de octubre de 2008

Daniel Levin
Mathematical Institute
University of Oxford
24-29 St Giles', Oxford OX1 3LB, United Kingdom
levin@maths.ox.ac.uk

Terry Lyons
Mathematical Institute
University of Oxford
24-29 St Giles', Oxford OX1 3LB, United Kingdom
tlyons@maths.ox.ac.uk

Supported by EPSRC grant "PDE–A rough path approach" GR/S18526/01. The second author would also like to acknowledge the support of the Oxford-Man Institute and the Swedish Institute of Sciences for visits to Mittag-Leffler in 2007.