

Sharp Gradient Bounds for the Diffusion Semigroup

Colm Nee

November 2011

I would like to thank my supervisor and mentor, Dan Crisan, whose support, both academic and otherwise, made this work possible. I gratefully acknowledge the doctoral training grant awarded to me by EPSRC: long may generous investment in the physical sciences continue. I would like to thank Claudia Klüppelberg for allowing and organising my working at the Technical University of Munich as a guest. I would also like to praise the numerous (sadly too numerous to give individual mention to) doctoral students, post-doctoral students and lecturers for the many interesting and stimulating discussions I enjoyed during my doctoral studies. And finally, but most importantly, I dedicate this work to my mother and father for their immeasurable support in my education.

Abstract

Precise regularity estimates on diffusion semigroups are more than a mere theoretical curiosity. They play a fundamental role in deducing sharp error bounds for higher-order particle methods. In this thesis error bounds which are of consequence in iterated applications of Wiener space cubature (Lyons and Victoir [29]) and a related higher-order method by Kusuoka [21] are considered. Regularity properties for a wide range of diffusion semigroups are deduced. In particular, semigroups corresponding to solutions of stochastic differential equations (SDEs) with non-smooth and degenerate coefficients. Precise derivative bounds for these semigroups are derived as functions of time, and are obtained under a condition, known as the UFG condition, which is much weaker than Hörmander's criterion for hypoellipticity. Moreover, very relaxed differentiability assumptions on the coefficients are imposed. Proofs of exact error bounds for the associated higher-order particle methods are deduced, where no such source already exists. In later chapters, a local version of the UFG condition - 'the LFG condition' - is introduced and is used to obtain local gradient bounds and local smoothness properties of the semigroup. The condition's generality is demonstrated. In later chapters, it is shown that the V_0 condition, proposed by Crisan and Ghazali [8], may be completely relaxed. Sobolev-type gradient bounds are established for the semigroup under very general differentiability assumptions of the vector fields. The problem of considering regularity properties for a semigroup which has been perturbed by a potential, and a Langrangian term are also considered. These prove important in the final chapter, in which we discuss existence and uniqueness of solutions to the Cauchy problem.

Contents

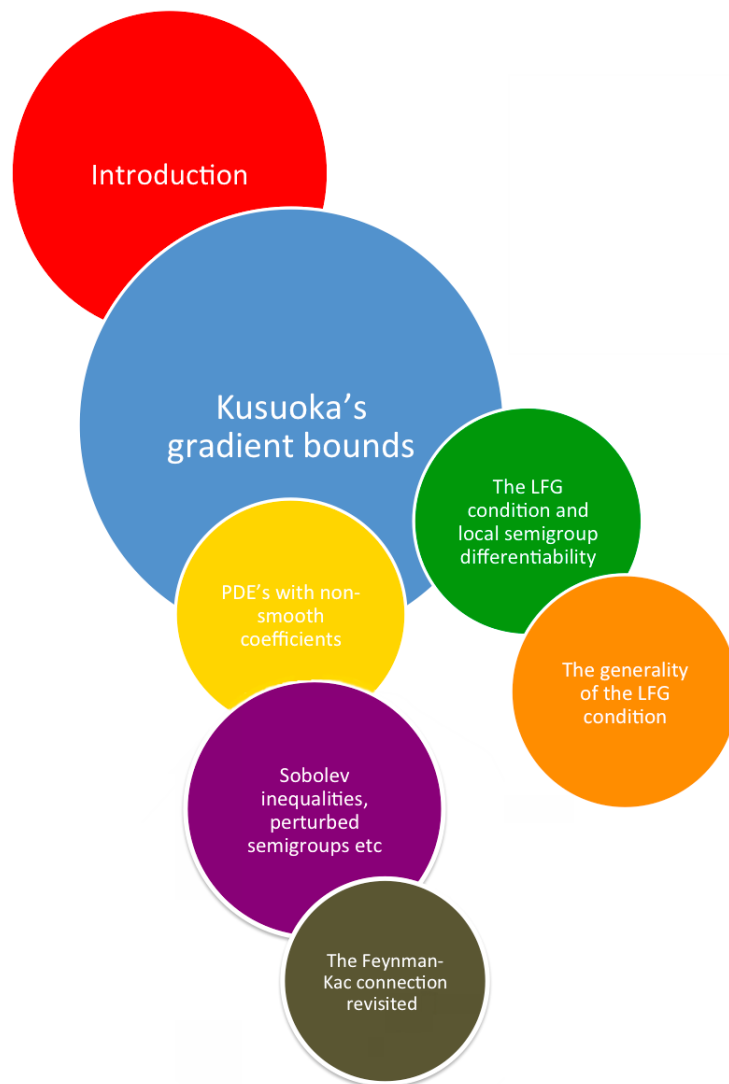
1. Introduction and background material	8
1.1. Motivation and Introduction	8
1.2. Gradient bounds	12
1.3. The role of Malliavin calculus in probability laws	13
1.4. The Malliavin calculus	16
1.5. The stochastic flow of diffeomorphisms	23
1.6. Multi-indices, Lie brackets and the UFG condition	27
1.7. The Kusuoka-Lyons-Victoir method	31
2. The gradient bounds of Kusuoka	37
2.1. Non-linearity breeds contempt	38
2.2. The integration by parts formula	44
2.3. Invertibility of the Malliavin covariance matrix	46
2.4. Diffuseness of iterated Stratonovich integrals	51
2.5. Precise gradient bounds	62
2.6. Hörmander's theorem and hypoellipticity	69
3. Regularity of PDEs with C_b^k coefficients	77
3.1. The UFG condition for non-smooth, non-bounded vector fields	77
3.2. An integration by parts formula of limited applicability	80
3.3. A finite number of sharp gradient bounds	82
3.4. Applications to gradient bounds	84
3.5. Application of the KLV method to non-smooth semigroups	88
4. The LFG condition and local differentiability of the semigroup	92
4.1. Adapting the integration by parts formula for local properties	94
4.2. Robustness of the invertibility proof	97
4.3. Several local integration by parts formulae	99
4.4. The deduction of local gradient bounds	102
5. The generality of the LFG condition	106
5.1. Some pathologies disproving <i>a priori</i> -ness of the LFG condition	107
5.2. Finite generation of the Lie algebra	110

5.3.	Constant rank of the analytic Lie algebra	115
5.4.	Satisfaction of the LFG condition under constant rank	116
6.	Sobolev inequalities, perturbed semigroups, the Lagrangian and relaxing the V_0 condition	119
6.1.	Weighted Sobolev spaces	119
6.2.	The perturbed semigroup	123
6.3.	The Lagrangian term	131
6.4.	Relaxing the V_0 condition	135
7.	The Feynman-Kac connection revisited	138
7.1.	Existence of solutions	140
7.2.	Uniqueness of solutions	144
A.	Appendix	149
A.1.	'Introduction and background material'	150
A.2.	Appendix for 'Kusuoka's gradient bounds'	150
A.3.	'Regularity of PDEs with C_b^k coefficients'	160
A.4.	'The LFG condition and local differentiability of the semigroup'	166

List of Figures

0.1. Interdependence and proportion of the chapters: intersection denotes dependence . .	7
1.1. Illustration of the smoothing properties of the diffusion semigroup	13
1.2. Mollifiers approximating the Dirac measure	14
1.3. Mollification of $(x - K)^+$ for $K = 1$	15
1.4. Cubature formula of degree 5 for one Brownian motion	35
5.1. A non-analytic smooth function	107
5.2. Extension of non-analytic smooth function to general dimensions	108
5.3. A lower-semicontinuous function	114

Figure 0.1.: Interdependence and proportion of the chapters: intersection denotes dependence



1. Introduction and background material

1.1. Motivation and Introduction

In finance, economics and the natural sciences one often seeks to model the key dynamics of random processes through the use of semimartingales. They impart a broad class of stochastic process, whose form is intuitive and tractable. As a substrate of this class, stochastic differential equations (SDEs) are widely studied by theoreticians and practitioners. Their key benefit is the ability to specify the behaviour through the choice of deterministic functions: the 'diffusion' and 'drift' coefficients. The ability to effect changes in dynamics through the choice of these factors - in such a transparent and instinctive way - elevates SDEs within the mathematical study of continuously observable phenomena. Although traditionally the domain of deterministic methodology, the Feynman-Kac formula, which equates the study of PDEs and their stochastic ordinary counterparts, has moved the latter out of the former's shadow and has further sharpened our understanding of nature's intimate secrets. This particle-based approach offers a fresh perspective on both old and new problems. If motivation were needed, one can be certain that even the minutest observation of their properties would have proved a worthwhile pursuit.

In physics, SDEs are employed to examine the proliferation of heat through matter over time. In quantum mechanics, they are used to describe Schrödinger's equation, which models the evolution of a physical system's quantum state. In neuroscience, the study of cable theory uses them to calculate the flow of electrical current along neuronal fibres. In finance, they are widely used as models in asset and derivative pricing. The list goes on, and is as impressive as it is long.

Unfortunately, the class of SDEs which admit explicit solutions is limited. Moreover, the SDEs which are explicitly solvable often prove to be inadequate reflections of reality. Consequently, to benefit from more realistic models, discrete time approximations of the solutions are frequently required.

There are two distinct types of approximations that one can choose to study following this unfortunate revelation. The first are classed as '*strong*'. This class is characterised by the pathwise approximation which is sought - that is: numerical approximation of the solution is carried out path-by-path. This is advantageous in areas like scenario analysis, filtering and hedge simulation. On the other hand, '*weak*' approximations seek to provide the distribution of a solution at (one or several) particular times. This is especially relevant in the aforementioned scientific branches, as these approximations respect the link with PDEs via Feynman-Kac.

This work only considers *path-continuous semimartingales* and is concerned with questions which are of use in weak approximations.

Much energy is spent on implementing efficient numerical algorithms which weakly approximate SDEs. The driving forces behind this effort are individual and subject specific, and the demands on the performance and accuracy of these algorithms will vary from person to person and industry to industry. What perhaps sets the financial world apart from other areas in which numerical approximations of natural phenomena play a significant role, is the need for very rapid convergence. Whilst the physicist at CERN might be content to allow his numerical scheme to run overnight, a bank's trading desk - owing to the liquidity, quantity, and parametric diversity of financial derivatives, as well as the desire to complete trades in constantly changing markets - need to produce results at a rate as small as thousandths of a second. It is partly for this reason that, in many cases, the numerical schemes of Euler and other related higher order approximations are judged impractical. They sometimes prove one, or a combination of: too slow, unstable, inaccurate¹. The driving force behind the theoretical slant of this work, is to place a class of modern numerical schemes, which avoid many of these pitfalls, on a firm mathematical footing.

OVERVIEW

In general, and unless specified otherwise, we consider SDEs written in the Stratonovich form:

$$X_{s,t}^x = x + \int_s^t V_0(u, X_{s,u}^x) du + \sum_{j=1}^d \int_s^t V_j(u, X_{s,u}^x) \circ dB_u^j, \quad (1.1)$$

where V_0 is called the *drift term*, and $V_1, \dots, V_d$ make up the diffusion term. One of the primary reasons to study SDE solutions is due to their connection with parabolic PDEs. In particular, we assume the drift and diffusion term are time-homogeneous and satisfy linear growth, and that $V_1, \dots, V_d \in C_b^2$. Moreover, if $f : \mathbb{R}^N \rightarrow \mathbb{R}$ is continuous and has polynomial growth, then the solution $u \in C^{1,2}([0, T] \times \mathbb{R}^N)$ of the 'Cauchy problem':

$$\begin{aligned} \frac{\partial u}{\partial t} &= \mathcal{L}u ; & \text{in } (0, T] \times \mathbb{R}^N, \\ u(0, \cdot) &= f ; & \text{on } \mathbb{R}^N, \end{aligned} \quad (1.2)$$

where $\mathcal{L} := \sum_{i=1}^d V_i^2 + V_0$, has a stochastic representation in terms of the one-parameter 'diffusion semigroup', $P_t f(x)$, defined as:

$$u(t, x) := (P_t f)(x) := \mathbb{E} [f(X_{0,t}^x)]. \quad (1.3)$$

¹For a survey of such schemes, there are few sources more comprehensive than the book of Kloeden and Platen [18].

Note that the drift and diffusion terms in the expression for the second order differential operator $\mathcal{L}$, are viewed as being first order differential operators. For more details on the Feynman-Kac formula consult, for example, Karatzas and Shreve [17], or Rogers and Williams [35]. Chapter 6 of this thesis is also devoted to the presentation and discussion of this problem.

The Feynman-Kac formula creates a duality, which permits the study of regularity properties of parabolic PDEs through a deterministic or stochastic approach. The classical theory on PDE-related regularity properties was centred around SDE solutions for which the *uniform ellipticity* or more generally, the *uniform Hörmander* condition hold. These are well-known criteria on the drift and diffusion coefficients. For example, it is well-documented that under the assumption of time-homogeneity², smoothness and ‘uniform Hörmanderness’ of the coefficients, for any uniformly continuous function f , the diffusion semigroup applied to f , $P_t f(\cdot)$, is a smooth function for any $t > 0$.

Many of the classical schemes rely on such properties and so Hörmander’s wonderful 1967 paper [14] is seen to be a considerable contribution to this field. However, probabilists prefer not to rely on analytical proofs wherever possible, and it is this desire which led Malliavin [30] to develop a calculus through which he was able to prove, probabilistically, the sufficiency of Hörmander’s condition. Although a remarkable achievement in itself, this calculus is proving to have far reaching consequences beyond its initial purpose. The work of Kusuoka and Stroock [23], [24] & [25], and Kusuoka [22], in the 80’s proved an extension of Malliavin’s work and provided precise, time-explicit regularity estimates for semigroups and densities of laws (where they exist) of SDE solutions. As shall be demonstrated, these estimates have very important applications in deducing convergence of particular modern numerical schemes and algorithms, which approximate expectations of SDE functionals. The classical schemes based on Monte-Carlo simulation simply do not converge to a satisfactory level of accuracy fast enough for many cases of interest. The work of Kusuoka [21] and Lyons & Victoir [29] provide new numerical schemes, whose proof of rapid convergence is derived from the precise regularity estimates of Kusuoka. In [22], Kusuoka extended the class of SDEs one may apply these schemes to by introducing a new, more general, condition: the so-called UFG condition is shown to imply Hörmander’s uniform hypoellipticity condition (cf. Hörmander [14]), and similar explicit gradient bounds may be proved under this assumption. There are also examples for which Hörmander’s theorem fails to hold, but the UFG condition does, making the latter demonstrably more general. This is discussed in more detail later.

In what remains of this introductory chapter, we present the notion of a gradient bound and discuss how integration by parts formulae are used to derive them. We give a comprehensive introduction to the Malliavin calculus, having motivated this exhibition with a dialogue of its worth within the subject of integration by parts formula and densities of probability laws. Regularity of SDE solutions, an important component of the techniques employed throughout, is demonstrated

²For an comprehensive extension of regularity properties of PDEs with time-inhomogeneous coefficients, using stochastic methods, see Cattiaux and Mesnager [6]

by way of a summary of the important results from the literature. Important notation, which shall be used throughout the work, is familiarised in the subsequent section, along with a statement and extensive discussion of the UFG condition. Finally, the impetus for the theoretical inclinations of the work is established via a treatise on the Kusuoka-Lyons-Victoir (KLV) method.

In the second chapter we give a full and self-contained account of Kusuoka's visionary results (cf [22]) on regularity of the one-parameter (i.e. time-homogeneous) diffusion semigroup as an explicit function of time. Many of the techniques demonstrated here will be employed in subsequent extensions.

In the third chapter the restrictive assumptions of smoothness of the coefficients is relaxed. Regularity of non-smooth equations and their corresponding gradient bounds are deduced. This regularity depends intrinsically on the level of coefficient differentiability, and the 'order' of the UFG condition.

In the fourth chapter we present of local version of the UFG condition, termed the '*LFG condition*'. An integration by parts formula is derived and is used to prove local regularity properties of the diffusion semigroup. The work is also applied to the more familiar Hörmander setting and we show local smoothness of the semigroup, under a local version of Hörmander's criterion.

In the fifth chapter we discuss how the very general LFG condition can be shown to hold *a priori* in certain situations. We treat two cases separately: analytic coefficients and smooth coefficients. In particular, we show that for analytic coefficients the Lie algebra is finitely generated on compact sets, and the LFG condition holds across points of '*constant rank*'. If the coefficients are 'merely' smooth, then finite generation holds on compact sets for which the rank of the Lie algebra is constant, and the LFG condition again holds across points of constant rank. Counter-examples for non-analytic coefficients not satisfying the LFG condition are provided, dispelling the notion that the LFG condition holds on compact sets.

In the sixth chapter we extend the class of gradient bounds deduced from the integration by parts formulae. This includes bounds for weighted Sobolev-type norms. We also consider the case where the semigroup has been perturbed by a potential term, and these calculations are used to consider an additional Lagrangian term. Several gradient bounds are deduced for the perturbed semigroup and Lagrangian term. These are proved in preparation for the final chapter on the connection of semigroup theory with parabolic PDEs. Finally, the V_0 condition is relaxed, which was proposed by Crisan and Ghazali in [8] to overcome a problem in proving theoretical efficacy of certain cubature methods.

In the seventh and final chapter we discuss the connection between semigroup theory and solutions of parabolic PDEs. We redefine the notion of a strong solution to the Cauchy problem, based on the regularity results proved throughout the thesis. A corresponding notion of a weak solution is

given, and existence of a solution is proved under various criteria. Finally uniqueness is demonstrated under the same conditions.

1.2. Gradient bounds

In this thesis, when we discuss gradient bounds for the diffusion semigroup, we will be exclusively considering the family of operators: $\{f \mapsto P_t f(x) : t \geq 0\}$, defined as

$$(P_t f)(x) := \mathbb{E} [f(X_{0,t}^x)],$$

where f belongs to a suitable class of test functions, and $\{X_{s,t}^x\}_{s \leq t}$ is the solution to (1.1) in the case where the vector fields $V_0, \dots, V_d$ are time-homogeneous. In general, the study of the solution to (1.1) would lead to consideration of a two-parameter semigroup, but when the vector fields are time-homogeneous the two-parameter semigroup depends on s and t only through their difference $t - s$, and hence it is equivalent to view it as a one-parameter semigroup. Although Diffusion semigroups are known - via the Feynman-Kac formula - to be solutions of parabolic PDEs. The estimation of the global error of certain numerical schemes depends intrinsically on the smoothness of $P_t f(\cdot)$. It is for this reason that regularity properties of diffusion semigroups play an important role in proving convergence. For much of this work, we shall be concentrating on gradient bounds of the following type:

$$\|V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t f\|_{L^p} \leq C_t \|f\|_{L^p}, \quad p \in [1, \infty] \quad (1.4)$$

where $V_{[\alpha_i]}$ are vector fields, along which derivatives are taken. These shall be defined shortly and at this stage it suffices to say that they are intrinsically linked to the drift and diffusion coefficients. It is well-known that the diffusion semigroup has very strong smoothing properties. This can be compared to the tendency of an irregularly heated metal bar to unify in temperature as time passes. Figure 1.1 illustrates this smoothing effect through the passage of time. For this reason we seek to quantify this propensity by identifying, as sharply as possible, C_t in the above. Apart from being fascinated by the diffusion semigroup's smoothing characteristics, one may ask: *why is it worthwhile to study gradient bounds?* As has already been alluded to, one of the primary reasons is that they facilitate the mathematical proof of the extremely fast convergence of certain numerical schemes. The main point to note is that we can approximate the expected value of an SDE functional by considering its Stratonovich Taylor expansion. By doing this, the error of the numerical scheme can be expressed in terms of the derivatives appearing on the RHS of (1.4). As we shall see, Malliavin calculus is an ideal piece of machinery for studying gradient bounds³ as it provides the mechanics to deduce an integration by parts formula, from which these gradient bounds are deduced.

³as well as its better known uses in existence and regularity of densities for laws of random variables.

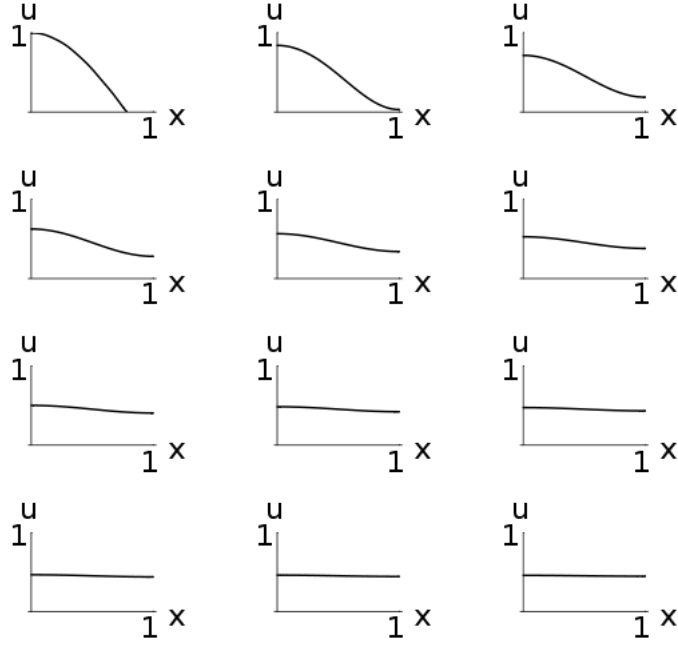


Figure 1.1.: Illustration of the smoothing properties of the diffusion semigroup

1.3. The role of Malliavin calculus in probability laws

Malliavin Calculus is tailor-made to provide the tools to answer the questions which an experimental probabilist might stumble upon. Here a brief rationale is given behind its original purpose.

PROBABILITY DENSITIES

Suppose we have a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we have a random vector, $X : \Omega \rightarrow \mathbb{R}^d$. In seeking an expression for the density of the law of such a vector one would undoubtedly start to examine expressions such as: $\mathbb{E} \varphi(X)$. Assume without loss of generality that $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ is a smooth, compactly-supported function, i.e. $\varphi \in C_0^\infty(\mathbb{R}^d)$. We can approximate general, bounded measurable functions with those in the said space. We could sequentially approximate $\varphi(X)$ by convoluting it with compactly-supported smooth functions, which approximate the Dirac measure, then Fubini's theorem may be used to interchange expectation and Lebesgue integration. This may lead to an expression for the density. These approximating functions are known as *kernel functions* or *mollifiers*. Let $\rho \in C_0^\infty(\mathbb{R}^d)$ be such that:

$$\int_{\mathbb{R}^d} \rho(x) dx = 1, \quad \text{Supp}(\rho) \subset B_1(0),$$

Then define a sequence $\{\rho_n\} \in C_0^\infty(\mathbb{R}^d)$ by:

$$\rho_n(x) := n^d \rho(nx),$$

It is easy to see that $\rho_n \in C_0^\infty(\mathbb{R}^d)$ and that $\int_{\mathbb{R}^d} \rho_n(x) dx = 1$. Moreover, define

$$\varphi_n(x) := (\varphi * \rho_n)(x) := \int_{\mathbb{R}^d} \varphi(y) \rho_n(x - y) dy.$$

It may be shown that $\varphi_n \in C_0^\infty(\mathbb{R}^d)$, and that $\varphi_n \xrightarrow{L^p(dx)} \varphi$, for $p \in [1, \infty)$. Moreover, if $\varphi \in C(\mathbb{R}^d)$ then $\varphi_n \rightarrow \varphi$ uniformly on compact sets⁴. Now one may attempt to derive an expression for the density based on these mollifiers. The dominated convergence theorem and Fubini's theorem can be applied to deduce the following

$$\begin{aligned} \int_{\mathbb{R}^N} \varphi(y) p_X(y) dy &= \mathbb{E}[\varphi(X)] = \lim_{n \rightarrow \infty} \mathbb{E}[\varphi_n X] \\ &= \lim_{n \rightarrow \infty} \mathbb{E} \int_{\mathbb{R}^d} \varphi(y) \rho_n(X - y) dy \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} \varphi(y) \mathbb{E} \rho_n(X - y) dy \\ &= \int_{\mathbb{R}^d} \varphi(y) \lim_{n \rightarrow \infty} \mathbb{E}[\rho_n(X - y)] dy \\ &\stackrel{(\dagger)}{=} \int_{\mathbb{R}^d} \varphi(y) \mathbb{E} \left[\lim_{n \rightarrow \infty} \rho_n(X - y) \right] dy \\ &= \int_{\mathbb{R}^d} \varphi(y) \mathbb{E}[\delta_0(X - y)] dy \\ &\Rightarrow p_X(y) = \mathbb{E} \delta_0(X - y), \quad \text{for almost all } y \in \mathbb{R}^N. \end{aligned}$$

Figure 1.3 illustrates how the mollifiers smooth the function for $\varphi(x) := (x - K)^+$: Unfortunately,

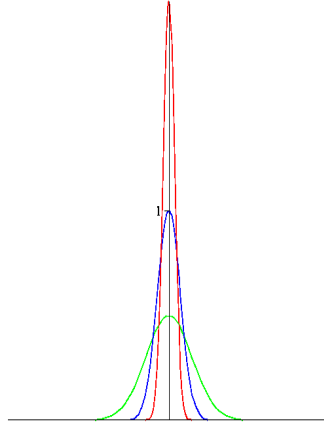


Figure 1.2.: Mollifiers approximating the Dirac measure

the expression for the density given above is untractable at best. Moreover, the step $(\dagger)$ may only be

⁴for more information about mollification see Jost [16, pp250-253]

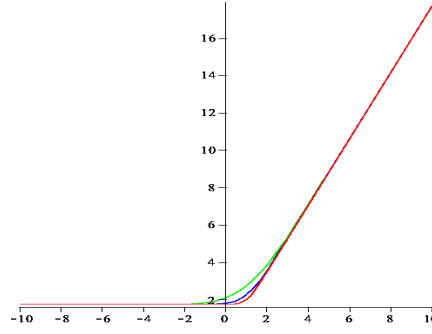


Figure 1.3.: Mollification of $(x - K)^+$ for $K = 1$

understood on a heuristic level. This is why an integration by parts formula is sought - by recognising the Dirac delta function in the current expression as the derivative of the Heaviside function - such a formula would provide *not only* a more tractable expression for the density, but also a rigorous way to prove the step (†). The following abstract definition clears up what is meant by 'an integration by parts formula'. For a comprehensive summary of the relationships between integration by parts formulae one should consult, for example: Nualart [33] or Sanz-Solé [37].

Definition 1.1 (Integration by Parts Formula) Let X be a random vector with values in $\mathbb{R}^N$. Moreover, let Y be an integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. Let α be a multi-index (i.e. $\alpha = (\alpha_1, \dots, \alpha_N)$ for $\alpha_i \in \mathbb{N}$). The pair (X, Y) is said to satisfy an integration by parts formula (IBPF) of order $|\alpha|$ if there is an integrable random variable $H_\alpha(X, Y)$ such that

$$\mathbb{E}[\partial_\alpha \varphi(X) Y] = \mathbb{E}[\varphi(X) H_\alpha(X, Y)], \quad (1.5)$$

for all $\varphi \in C_b^\infty(\mathbb{R}^N)$.

Notice that for the immediate purposes one only seeks an integration by parts formula for $\alpha = (1, \dots, 1)$ and $Y = 1$. In this case, if it is assumed that (1.5) holds for $Y = 1$, then by considering $\psi_n(x) := \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_N} \rho_n(y) dy$, i.e. so that $\partial_\alpha \psi_n = \rho_n$ it holds that:

$$\mathbb{E}[\rho_n(X - y)] = \mathbb{E}[\psi_n(X - y) H_{(1,1,\dots,1)}(X, 1)].$$

Now the situation is different, as the dominated convergence theorem can be applied to the RHS (note that the 'dominator' i.e. integrable random variable which dominates $H_{(1,1,\dots,1)}(X, 1) \forall n \in \mathbb{N}$, is $|H_{(1,1,\dots,1)}(X, 1)|$). Hence,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}[\rho_n(X - y)] &= \lim_{n \rightarrow \infty} \mathbb{E}[\psi_n(X - y) H_{(1,1,\dots,1)}(X, 1)] \\ &= \mathbb{E}\left[\lim_{n \rightarrow \infty} \psi_n(X - y) H_{(1,1,\dots,1)}(X, 1)\right] \end{aligned}$$

Note that we have used that $\{\psi_n\}_n$ are bounded uniformly in n . We now note that $\psi_n(x) \rightarrow 1_{[0,\infty)}(x)$ where

$$1_{[0,\infty)}(x) = \begin{cases} 1 & \text{if } x_i \geq 0, \quad \forall i = 1, \dots, d, \\ 0 & \text{otherwise.} \end{cases}$$

Hence one has the following expression for the density of the law of X .

$$p_X(y) = \mathbb{E}(1_{[y,\infty)}(X)H_{(1,1,\dots,1)}(X, 1)).$$

Of course, the integrable function H which allows this calculation to be performed is unknown at this stage. It will not come as a surprise to the reader that Malliavin calculus is turned to for help, as it is precisely Malliavin calculus which provides the sought after integration by parts formula.

The definition was stated in generality, which allows scope for proving regularity results about the density. Indeed, if the formula holds for any multi-index α then it can be shown that the density is smooth.

Although Malliavin calculus can be easily motivated from the perspective of densities for law of random variables, another important use for it is for deducing gradient bounds.

1.4. The Malliavin calculus

In this section the notion of a differential calculus on a probability space, formally known as the Malliavin calculus, is introduced. Although the Malliavin calculus is exhibited in this section with Brownian motion on the space of continuous functions, more general presentations of the theory are possible. Indeed, Malliavin calculus can be constructed using a more general *Gaussian process* on a more general construction of the Wiener space. For details of such topics one may consult, for example, Nualart [33].

The probability space will be the standard Wiener space with paths in $\mathbb{R}^d$. That is, consider the space $(\Omega, \mathcal{B}, \mathbb{P})$ where $\Omega = \{\omega \in C_0([0, \infty) ; \mathbb{R}^d)\}$, $\mathcal{B} = \mathcal{B}(C_0([0, \infty) ; \mathbb{R}^d))$ and $\mathbb{P}$ is the Wiener measure, i.e. the measure such that the coordinate mapping process: $B = \{B_t, t \in [0, \infty)\}$, $B_t(\omega) := \omega(t)$ is a d -dimensional Brownian motion.

The Wiener space has an important subspace, which is fundamental to the development of the Malliavin calculus. Namely, $H = \{h \in \Omega ; h' \in L^2([0, \infty) ; \mathbb{R}^d)\}$. i.e. we consider those continuous functions with values in $\mathbb{R}^d$ which are absolutely continuous with respect to the Lebesgue measure on $[0, \infty)$. The fact that absolutely continuous paths are differentiable almost everywhere, permits the use of the notation h' for its derivative. The Hilbert space property of H , under the inner product $\langle h, g \rangle_H := \langle h', g' \rangle_{L^2([0, \infty); \mathbb{R}^d)} := \int_0^\infty h'(u) \cdot g'(u) du$, is inherited from the Hilbert space $L^2([0, \infty) ; \mathbb{R}^d)$ and the almost everywhere uniqueness of their densities. i.e. if two absolutely continuous functions differ on a set of positive measure, then their $L^2([0, \infty) ; \mathbb{R}^d)$ representations must differ on a set of positive measure. H is known as the *Cameron-Martin Space* and plays an

important part in Malliavin calculus.

In developing a rigorous theory of stochastic integration (which, in turn leads to stochastic differential equations) some care is taken to define the integral. In particular, the theory is developed with respect to a probability measure, and functionals of stochastic integrals are always defined $\mathbb{P}$ -a.s. It may have been tempting for those initially seeking a theory of differentiation on the Wiener space to consider Fréchet differentiation on $(\Omega, \|\cdot\|_\infty)$. The following result shows that this is not wise.

Proposition 1.2 *Consider the functional $F : \Omega \rightarrow \mathbb{R}$ given by $F(\omega) := \int_0^1 h_t dW_t(\omega)$, for some $h : [0, \infty) \rightarrow \mathbb{R}$. Then F has a continuous modification if and only if there exists a signed measure μ on $[0, 1]$ such that $h(t) = \mu((t, 1])$ for a.a. $t \in [0, 1]$.*

Proof : See Nualart [33, p34] ■

This result shows us that there is no hope of developing a tractable theory of differentiation on $(\Omega, \|\cdot\|_\infty)$, where functionals of stochastic integrals are considered, as the price for continuity is a high one, i.e. one can only consider integrands which have bounded variation.

One is then confronted with a problem; whether meaning can be assigned to objects like $F(\omega + \epsilon \tilde{\omega})$. These would form an important part of any theory of differentiation as they are the basis of directional differentiation. One must tread carefully here, as stochastic integration is constructed $\mathbb{P}$ -a.s. One would ideally like to know that if F is defined $\mathbb{P}$ -a.s. that $F(\cdot + \epsilon \tilde{\omega})$ is too, for a.e. $\tilde{\omega} \in \Omega$ and $\epsilon > 0$. This seemingly innocuous problem prevents some difficulty. Indeed, one may show that $\mathbb{P}$ and $\mathbb{P}_{\tilde{\omega}}$, the Wiener measure and the shifted Wiener measure are singular for $\tilde{\omega} \in \Omega \setminus H$. This is bad news, as it means that one cannot deduce anything about the well-definedness of $F(\cdot + \epsilon \tilde{\omega})$ from $F(\cdot)$. The positive news (in fact, the reason for restricting attention to H) is that the exact opposite is true when $\tilde{\omega} \in H$. This is a corollary of the *Cameron-Martin Theorem*, which, incidentally, also gives us a formula for the directional derivative of F (in directions of the Cameron-Martin space) and an integration by parts formula. A derivative shall be defined on a restrictive class of random variables, and will then be generalised and extended to a much broader class. Solutions of stochastic differential equations will form a part of this extended class. From now on: $\langle \cdot, \cdot \rangle := \langle \cdot, \cdot \rangle_{L^2([0, \infty); \mathbb{R}^d)}$

THE MALLIAVIN DERIVATIVE

Definition 1.3 (Malliavin derivative) *We consider for $f \in C_0^\infty(\mathbb{R}^n)$ the functional $F : \Omega \rightarrow \mathbb{R}$ and $h'_1, \dots, h'_n \in L^2([0, \infty) ; \mathbb{R}^d)$ given by:*

$$F(\omega) = f \left(\int_0^\infty h'_1(t) dB_t(\omega), \dots, \int_0^\infty h'_n(t) dB_t(\omega) \right),$$

where $\int_0^\infty h'(t) dB_t := \sum_{i=0}^d \int_0^\infty h'_i(t) dB_t^i$. Such random variables will be called 'smooth', and denoted by $F \in \mathcal{S}$. Then the Malliavin derivative of F , denoted $DF \in L^2(\Omega ; H) \simeq L^2(\Omega \times [0, \infty) ;$

$\mathbb{R}^d$) is given by:

$$DF = \sum_{i=1}^n \partial_i f \left(\int_0^\infty h'_1(u) dB_u, \dots, \int_0^\infty h'_n(u) dB_u \right) \int_0^\cdot h'_i(s) ds. \quad (1.6)$$

The Malliavin derivative may be equivalently viewed as an element of the Cameron-Martin space, or its L^2 kernel. Denote

$$D_h F := \langle DF, h \rangle_H,$$

This will be called the directional derivative of F in the direction h . In order to see this interpretation of DF , consider for $h = \int_0^\cdot h'(s) ds$ where $h' \in L^2$:

$$\begin{aligned} D_h F &= \sum_{i=1}^d \partial_i f \left(\int_0^\infty h'_1(u) dB_u, \dots, \int_0^\infty h'_n(u) dB_u \right) \left\langle \int_0^\cdot h'_i, h \right\rangle_H \\ &= \frac{d}{d\epsilon} f \left(\int_0^\infty h'_1(u) dB_u + \epsilon \langle h'_1, h' \rangle, \dots, \int_0^\infty h'_n(u) dB_u + \epsilon \langle h'_n, h' \rangle \right) \Big|_{\epsilon=0} \\ &= \frac{d}{d\epsilon} f \left(\int_0^\infty h'_i(u) d(\omega + \epsilon \int_0^\cdot h'(s) ds)(u) \right)_{i=1, \dots, n} \Big|_{\epsilon=0} \\ &= \frac{d}{d\epsilon} F(\omega + \epsilon h) \Big|_{\epsilon=0}. \end{aligned}$$

This equality encourages the use of the Cameron-Martin theorem, which gives an explicit formula for the RHS of the above equation. This provides a (very basic) integration by parts formula, from which the rest of the theory flourishes.

Theorem 1.4 (Cameron-Martin Theorem) *Let F, G be smooth random variables. Let $h(\cdot) := \int_0^\cdot h'(u) du \in H$. There holds the following formula (Note that ω -dependence is not, as convention dictates, suppressed; for emphasis):*

$$\begin{aligned} \mathbb{E}F(\omega + \epsilon h)G(\omega) &= \mathbb{E}F(\omega)G(\omega - \epsilon h) \\ &\quad \cdot \exp \left(\epsilon \int_0^\infty h'(u) dB_u - \frac{1}{2} \epsilon^2 \int_0^\infty h'(u) du \right). \end{aligned}$$

Proof : See, for example, Øksendal [34]. ■

This theorem can be applied to give a basic integration by parts formula.

Theorem 1.5 (Basic Integration by Parts Formula) *Assume F, G are smooth random variables, and let $h' \in H$. Then the following equality holds.*

$$\mathbb{E}(D_h F \cdot G) = \mathbb{E}(FG \int_0^\infty h'(u) dB_u - F \cdot D_h G).$$

This fact is fundamental to the program. Without this integration by parts formula, Malliavin calculus would be destined to fail in its most basic task. It should be noted at this point that the integration by parts formula has a very pleasing corollary for the Malliavin derivative as an operator.

Questions may be understandably raised at the somewhat restrictive choice of ‘test’ random variable to which the Malliavin derivative may be applied. However, the above integration by parts formula goes a long way towards extending the class of random variables and justifies the choice of *smooth* random variables. One reason for choosing to initially consider the class of *smooth* random variables is the following:

Proposition 1.6 (Density of Smooth random variables in $L^2(\Omega)$)

$\mathcal{S}$ is dense in $L^2(\Omega)$. That is, for any $F \in L^2(\Omega)$ there exists $\{F_n\} \subset \mathcal{S}$ such that

$$\|F_n - F\|_{L^2(\Omega)} \rightarrow 0.$$

Proof : The details of this are available in, for example, Nualart [33]. It is shown that a set, which elements of $\mathcal{S}$ are dense in (namely the Wiener polynomials), are themselves dense in $L^2(\Omega)$. This is done by using *Hermite Polynomials* and the *Wiener-Ito Chaos Expansion*. ■

If the Malliavin derivatives of two convergent sequences of smooth random variables, which converge to the same $L^2(\Omega)$ -limit, have a different $L^2([0, \infty) \times \Omega)$ -limit, then how could one possibly assign a definition to the Malliavin derivative of this limit? This is vital for extending the Malliavin derivative to a broader class of random variables. This amounts to showing the following:

Corollary 1.7 (Closability of the Malliavin Derivative operator)

The Malliavin derivative, a linear unbounded operator $D : \mathcal{S} \rightarrow L^2([0, \infty) \times \Omega ; \mathbb{R}^d)$ is closable as an operator from $L^2(\Omega ; \mathbb{R}^d)$ into $L^2([0, \infty) \times \Omega ; \mathbb{R}^d)$.

Proof : Since the operator is linear (in generality one should apply the following to the sequence $\{F_n - G_n\}$), the problem reduces to showing that for $\{F_n\} \subset \mathcal{S}$ such that:

$$\|F_n\|_{L^2(\Omega)} \rightarrow 0,$$

and

$$\|DF_n\|_{L^2([0, \infty) \times \Omega)} \text{ is convergent.}$$

then it follows that

$$\|DF_n\|_{L^2([0, \infty) \times \Omega)} \rightarrow 0.$$

This can be shown using the integration by parts formula. Note that by virtue of the density of smooth random variables in $L^2(\Omega)$, it is sufficient to prove that

$$\mathbb{E}(DF_n(h)\varphi) \rightarrow 0,$$

for all $\varphi \in \mathcal{S}$ and h' . One may also reduce the consideration to smooth random variables φ_ϵ , where $\epsilon > 0$, such that $\varphi_\epsilon \int_0^\infty h'(u)dB_u$ is bounded. For example, $\varphi_\epsilon = \varphi \exp(-\epsilon \{\int_0^\infty h'(u)dB_u\}^2)$. Indeed, by taking ϵ arbitrarily close to zero, φ_ϵ may be made arbitrarily close to φ . This makes the

analysis slightly easier. Observe, from the integration by parts formula:

$$\mathbb{E}[D_h F_n \varphi_\epsilon] = \mathbb{E}\left[F_n \varphi_\epsilon \int_0^\infty h'(u) dB_u - F_n D_h \varphi_\epsilon\right].$$

It is now remarked that both $\varphi_\epsilon \int_0^\infty h'(u) dB_u$ and $D_h \varphi_\epsilon$ are bounded. Moreover, since it has been assumed that F_n converges to zero in $L^2(\Omega)$ it follows that:

$$\mathbb{E}[D_h F_n \varphi_\epsilon] \rightarrow 0,$$

as required. ■

One may also extend the domain of D to L^p . Indeed, smooth random variables are also dense in L^p for $p \geq 1$. It should be noted that for $p \neq 2$ the norm would be defined as:

$$\|DF\|_{L^p(\Omega; H)}^p := \mathbb{E}[\|DF\|_H^p]. \quad (1.7)$$

The closability property of D holds from $L^p(\Omega)$ to $L^p(\Omega; H)$. Denote the domain of D by $\mathbb{D}^{1,p}$, meaning that $\mathbb{D}^{1,p}$ is the closure of smooth random variables $\mathcal{S}$ with respect to the norm:

$$\|F\|_{1,p} = (\mathbb{E}|F|^p + \mathbb{E}\|DF\|_H^p)^{\frac{1}{p}}$$

The iteration of the Malliavin derivative D may also be defined in such a way that for smooth random variables, the iterated derivative $D^k F$ is a random variable with values in $H^{\otimes k}$. Define

$$D^k F := \sum_{i_1, \dots, i_k=1}^n \partial_{i_1, \dots, i_k} f \left(\int_0^\infty h'_1(u) dB_u, \dots, \int_0^\infty h'_n(u) dB_u \right) h_{i_1} \otimes \dots \otimes h_{i_k},$$

where $h_i(\cdot) := \int_0^\cdot h'_i(s) ds$. *Why is this an intuitive definition?* It corresponds to iterative applications of the directional Malliavin differentiation. Indeed, for $h \in H$, $F \in \mathcal{S}$, it is easily seen that $D_h F \in \mathcal{S}$ and it can be shown that,

$$D(D(\dots DF(h_1)) \dots, h_{k-1})(h_k) = \langle D_k F, h_1 \otimes \dots \otimes h_k \rangle_{H^{\otimes k}}.$$

In an analogous way, one can close the operator D^k from $L^p(\Omega)$ to $L^p(\Omega; H^{\otimes k})$. So, for any $p \geq 1$ and integer $k \geq 1$, define $\mathbb{D}^{k,p}$ to be the closure of $\mathcal{S}$ with respect to the norm:

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

For $p = 2$ there holds an isometry $L^2(\Omega \times [0, \infty)^k; \mathbb{R}^d) \simeq L^2(\Omega; H^{\otimes k})$. Hence one may identify $D^k F$ as a process: $D_{t_1, \dots, t_k}^k F$.

A random variable is said to be *smooth in the Malliavin sense* or *smooth with respect to the Malliavin derivative* if $F \in \mathbb{D}^{k,p}$ for all $p \geq 1$ and $k \in \mathbb{N}$. The set of all such random variables is denoted by $\mathbb{D}^\infty := \bigcap_{p \geq 1, k \in \mathbb{N}} \mathbb{D}^{k,p}$. Moreover, there is nothing which pins consideration to $\mathbb{R}^d$ -valued random variables. Indeed, one could consider more general Hilbert space-valued random variables, and the theory would extend in an appropriate way. To this end, denote $\mathbb{D}^{k,p}(E)$ to be the appropriate space of E -valued random variables where E is some separable Hilbert space. For more details, please consult Nualart [33].

THE CHAIN RULE

The chain rule provides a simple way to extend the class of differentiable random variables.

Proposition 1.8 (Chain Rule for the Malliavin Derivative) *If $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}$ is a continuously differentiable function with bounded partial derivatives, and $F = (F_1, \dots, F_m)$ is a random vector with components belonging to $\mathbb{D}^{1,p}$ for some $p \geq 1$. Then $\varphi(F) \in \mathbb{D}^{1,p}$, with*

$$D\varphi(F) = \nabla\varphi(F)DF = \sum_{i=1}^m \partial_i\varphi(F)DF_i,$$

where $\nabla\varphi := (\partial_1\varphi, \dots, \partial_m\varphi)$ and $DF = (DF_1, \dots, DF_m)^T \in \mathbb{R}^m \otimes \mathbb{R}^d$.

Proof : For details, consult Nualart [33]. ■

DIFFERENTIATING INTEGRALS

When considering SDEs there is a prominent need to investigate how differential operators interact with integrals. Indeed, efforts to show Malliavin differentiability of SDE solutions without knowing how to differentiate stochastic integrals would prove futile.

Lemma 1.9 (The Malliavin derivative and integration) *Assume that E is a separable real Hilbert space. Consider $f : [0, \infty) \times \Omega \rightarrow E$, and suppose for each $t \in [0, T]$ that $f(t) \in \mathbb{D}^{1,2}(E)$. Define $B_t^0 := t$. Moreover, suppose that:*

$$\mathbb{E} \int_0^T \|f(t)\|_E^2 dt < \infty \quad \mathbb{E} \int_0^T \|Df(t)\|_{E \otimes H}^2 dt < \infty. \quad (1.8)$$

Then $F_i(T) := \int_0^T f(t)dB_t^i \in \mathbb{D}^{1,2}(E)$ for all $i = 0, 1, \dots, d$, with

$$\begin{aligned} DF_0(T) &= \int_0^T Df(t)dB_t^0 \\ DF_i(T) &= \int_0^T Df(t)dB_t^i + \int_0^{T \wedge \cdot} f(s) \otimes e_i ds, \quad i = 1, \dots, d, \end{aligned}$$

Or, analogously:

$$\begin{aligned} D_h F_0(T) &= \int_0^T D_h f(t) dB_t^0 \\ D_h F_i(T) &= \int_0^T D_h f(t) dB_t^i + \int_0^T f(t) h_i(t) dt, \quad i = 1, \dots, d. \end{aligned}$$

Moreover, assuming that

$$\mathbb{E} \int_0^T \left\| D^{k-1} f(t) \right\|_{E \otimes H^{\otimes(k-1)}}^2 dt < \infty \quad \mathbb{E} \int_0^T \left\| D^k f(t) \right\|_{E \otimes H^{\otimes k}}^2 dt < \infty.$$

Then for the iterated Malliavin derivative operator D^k :

$$\begin{aligned} D^k F_0 &= \int_0^T D^k f(t) dB_t^0, \\ D^k F_i(T) &= \int_0^T D^k f(t) dB_t^i + \int_0^{T \wedge \cdot} D^{k-1} f(s) \otimes e_i ds, \quad i = 1, \dots, d. \end{aligned}$$

Proof : Proof is done using an induction argument. See Kusuoka and Stroock [23]. ■

THE DIVERGENCE OPERATOR

A brief discussion follows about the adjoint of the Malliavin derivative, which shall complete the differential calculus. The divergence operator shall play a vital role in the construction of an integration by parts formula. This operator is sometimes called the Skorohod integral because it coincides with a generalisation of the Itô integral to anticipating integrands introduced by Skorohod. A detailed discussion of the divergence operator can be found in Nualart [33].

Definition 1.10 (Divergence operator) Denote by δ the adjoint of the operator D . That is, δ is an unbounded operator on $L^2(\Omega \times [0, \infty) ; \mathbb{R}^d)$ with values in $L^2(\Omega)$ such that:

1. $\text{Dom } \delta = \{u \in L^2(\Omega \times [0, \infty) ; \mathbb{R}^d) ; |\mathbb{E}(\langle DF, u \rangle)| \leq c \|F\|_{L^2(\Omega)} \forall F \in \mathbb{D}^{1,2}\}.$
2. If $u \in \text{Dom } \delta$, then $\delta(u) \in L^2(\Omega)$ and the following basic integration by parts formula is satisfied:

$$\mathbb{E}(F \delta(u)) = \mathbb{E}(\langle DF, u \rangle).$$

CONTINUITY OF OPERATORS AND EQUIVALENCE OF NORMS

The following important results are shown in for example Section 1.5 of Nualart [33], and are important results for repeated applications of the integration by parts formula:

1. D is continuous from $\mathbb{D}^{k,p}(V)$ into $\mathbb{D}^{k-1,p}(H \otimes V).$

2. $\langle DF, DG \rangle_H \in \mathbb{D}^\infty$ if $F, G \in \mathbb{D}^\infty$.
3. δ is continuous from $\mathbb{D}^\infty(H)$ into $\mathbb{D}^\infty$.

1.5. The stochastic flow of diffeomorphisms

The rather cryptically named title of this section belies a straightforward and elegant principle. Solutions of stochastic differential equations possess numerous regularity properties. This triumph owes much to the prescience of Gihman [12],[13] and Blagovescenskii and Freidlin [5], who studied the regularity of solutions with respect to the initial data. In the late 1970s attention was given to the diffeomorphic property of the solution. Contributors to this field included Elworthy [9], Bismut [4], Ikeda and Watanabe [15], Kunita [20], and others. That solutions of stochastic differential equations - with suitably regular coefficients - can be chosen to be smooth with respect to changes in initial condition, and are Hölder continuous with respect to time, or, that *the stochastic flow is diffeomorphic*, is of fundamental influence in this work. This section is devoted to reviewing and presenting the literature on the topic. Malliavin differentiability of the stochastic flow is also discussed and exhibited.

Before we get ahead of ourselves, a probability space on which to define a stochastic flow is required. Let $(\Omega, \mathcal{B}, \mathbb{P})$ be the standard Wiener space with paths in $\mathbb{R}^d$. For purposes of notational simplicity, we define $B_t^0 := t$. Although the work will be considering time-homogeneous SDEs, this section will adopt as much generality as is provided in the literature. To this end we permit the coefficients of the SDE, $V_0, \dots, V_d : [0, \infty) \times \mathbb{R}^N \rightarrow \mathbb{R}^N$, to be time-inhomogeneous.

Definition 1.11 ($C^{k,\alpha}$ -function, C^k -diffeomorphism)

Let k be a non-negative integer and let $\alpha \in (0, 1]$. A function $V : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is called a $C^{k,\alpha}$ -function if it is k -times continuously differentiable with k th-order derivatives, which are locally Hölder continuous with exponent α . If the k th order derivatives are globally Hölder continuous, the function is called a $C_g^{k,\alpha}$ -function. A bijective map $f : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is called a C^k -diffeomorphism if both, f and f^{-1} are k -times continuously differentiable.

For this section the following *minimal* assumptions shall be made (although attention should be paid to where stronger conditions are required). For some $k \in \mathbb{N}$

1. $V_i(t, \cdot) \in C_g^{k+1,\alpha}(\mathbb{R}^N, \mathbb{R}^N)$, $\forall t \geq 0$, and $V_i(\cdot, x) \in C^1((0, \infty) ; \mathbb{R}^N)$, and is locally bounded on $[0, \infty)$, for all $x \in \mathbb{R}^N$, $i = 1, \dots, d$,
2. $V_0(t, \cdot) \in C_g^{k,\alpha}(\mathbb{R}^N ; \mathbb{R}^N)$, $\forall t \geq 0$, and $t \rightarrow V_0(t, x)$ is locally bounded on $[0, \infty)$,

for some $k \in \mathbb{N}$. We regard these vector-valued functions - and often refer to them - as vector fields. Consider the family of Stratonovich stochastic differential equations, which we shall often refer to as *the stochastic flow*, given by:

$$\begin{cases} dX_{s,t}^x &= \sum_{i=0}^d V_i(t, X_{s,t}^x) \circ dB_t^i, & s \leq t. \\ X_{s,s}^x &= x. \end{cases} \quad (1.9)$$

The existence and uniqueness of a solution to this equation is thoroughly addressed in the literature. See, for example, Karatzas and Shreve [17]. Most techniques for proving results about SDEs involve expressing the equation in Itô form. Indeed, the difference in the assumptions of these two sets of vector fields stems indirectly from the desired regularity of the corresponding Itô equation. The solution of (1.9) is given by the corresponding Itô equation:

$$\begin{cases} dX_{s,t}^x = \sum_{i=1}^d V_i(t, X_{s,t}^x) dB_t^i + \tilde{V}_0(t, X_{s,t}^x) dt, & s \leq t, \\ X_{s,s}^x = x, \end{cases} \quad (1.10)$$

where $\tilde{V}_0 = V_0 + \frac{1}{2} \sum_{i=1}^d (\partial_t V_i + \sum_{j=1}^d V_i^j \partial_j V_i)$. Moreover, based on our assumptions, $\tilde{V}_0$ satisfies property 2. Indeed, the reason different levels of differentiability were chosen for V_0 and $V_1, \dots, V_d$, is so that the coefficients of the corresponding Itô equation have the same level of differentiability. From Kunita [19], we have the following important theorem:

Theorem 1.12 *The solution $X_{s,t}^{(\cdot)}$ of (1.9) is a $C^{k,\beta}$ -function for any $\beta < \alpha$ and $s < t$. Furthermore, if the derivatives of $V_0(u, \cdot), \dots, V_d(u, \cdot)$ up to order k are also bounded, for all $u \in [s, t]$, then $X_{s,t}$ is a C^k -diffeomorphism for any $s < t$. Moreover, the Jacobian matrix, $J_{s,t}^x$, of $X_{s,t}^x$ satisfies the following Stratonovich matrix SDE*

$$\begin{cases} dJ_{s,t}^x = \sum_{i=0}^d \partial V_i(t, X_{s,t}^x) J_{s,t}^x \circ dB_t^i, & s \leq t, \\ J_{s,s}^x = I. \end{cases} \quad (1.11)$$

The Jacobian is almost surely invertible as a matrix and its inverse, $Z_{s,t}^x$, satisfies the SDE

$$\begin{cases} dZ_{s,t}^x = - \sum_{i=0}^d Z_{s,t}^x \partial V_i(t, X_{s,t}^x) \circ dB_t^i, & s \leq t, \\ Z_{s,s}^x = I. \end{cases} \quad (1.12)$$

If the vector fields of (1.9) are globally Lipschitz continuous functions, then for each $p \geq 1$, and $T > s$ there exists a constant $A_{s,T,p}$ such that

$$\sup_{t \in [s, T]} \mathbb{E} |X_{s,t}^x|^p < A_{s,T,p} (1 + |x|)^p, \quad (1.13)$$

If the vector fields of (1.9) are C_b^k -functions, then for each $p \geq 1$, and $T > s$ there exists a constant $B_{s,T,p}$, such that for each $|\gamma| \leq k$

$$\sup_{t \in [s, T]} \mathbb{E} \left| \frac{\partial^{|\gamma|} X_{s,t}^x}{\partial x^\gamma} \right|^p < B_{s,T,p}, \quad (1.14)$$

Moreover, if for an SDE with a general initial condition $X_0^x = f(x)$, it holds that f is uniformly bounded, then for each $p \geq 1$, and $T > s$ there exists a constant $C_{s,T,p}$, such that for each $0 \leq |\gamma| \leq k$

$$\sup_{t \in [0, T], x \in \mathbb{R}^N} \mathbb{E} |X_{s,t}^x|^p < C_{s,T,p}. \quad (1.15)$$

For completeness we note that the solution to (1.9) is jointly Hölder continuous in (s, t, x) with respective Hölder exponents (β, β, α) , where $\beta \in (0, 1/2)$.

Proof : Everything apart from equation (1.14) is directly addressed by Kunita in [19]. For a proof of (1.14), one should refer to Nualart [33, Corollary 2.1.1, p119]. ■

N.B. From now on, we write $(J_{s,t}^x)^{-1} := Z_{s,t}^x$.

Much of the analysis will rely heavily on the fact that the solution of an SDE is differentiable with respect to the Malliavin derivative, as well as with respect to its initial condition. Analogous to the previous theorem, this regularity depends intrinsically on the smoothness of the coefficients. Moreover, the expressions for the Malliavin derivative and the Jacobian matrix satisfy similar SDEs; they differ by an inhomogeneous term. The exploitation of this similarity is fundamental.

Theorem 1.13 *Assume X is the stochastic flow which solves (1.9), where the coefficients are assumed to satisfy the same conditions (1. and 2. on previous page), with derivatives of all lower orders also being globally Lipschitz continuous. Then $X_{s,t}^{i,x} \in \mathbb{D}^{k+1,p}$ for all $t \in [0, \infty)$, $i = 1, \dots, N$ and $p \geq 1$. Furthermore, the matrix $DX_t^x := (DX_{s,t}^{1,x}, \dots, DX_{s,t}^{N,x})$ satisfies the linear stochastic differential equation:*

$$DX_{s,t}^x = \sum_{i=0}^d \int_s^t \partial V_i(u, X_{s,u}^x) DX_{s,u}^x \circ dB_u^i + \left(\int_{s \wedge \cdot}^{t \wedge \cdot} V_j(u, X_u^x) du \right)_{j=1, \dots, d}. \quad (1.16)$$

where, by writing $t \wedge \cdot$, etc, we are interpreting the Malliavin derivative as an element of the Cameron-Martin space; a process in L^2 . We may also write the Malliavin derivative as the H -inner-product with $h \in H$:

$$D_h X_{s,t}^x = \sum_{i=0}^d \int_s^t \partial V_i(u, X_{s,u}^x) D_h X_{s,u}^x \circ dB_u^i + \sum_{k=1}^d \int_s^t V_k(u, X_{s,u}^x) h_i(u) du. \quad (1.17)$$

Note, if the vector fields are merely globally Lipschitz continuous, then (1.13) holds with ∂V_i replaced by some bounded random variables G_i . Moreover, if the vector fields $V_0, \dots, V_d$ are uniformly bounded then the following bound on the norms of the derivatives can be shown to hold:

$$\sup_{\substack{t \in [0, T] \\ x \in \mathbb{R}^N}} \mathbb{E} \left\| D^k X_t^x \right\|_{H^{\otimes k}}^p < C_{k,p}, \quad \forall p \in [1, \infty), T > 0, \quad (1.18)$$

If, however, the vector fields $V_0, \dots, V_d$ are globally Lipschitz continuous, then it may only be deduced that the following holds:

$$\sup_{t \in [0, T]} \mathbb{E} \left\| D^k X_t^x \right\|_{H^{\otimes k}}^p < C_{k,p} (1 + |x|)^p, \quad \forall p \in [1, \infty), T > 0, \quad (1.19)$$

Proof : The extra level of differentiability over the standard result comes from the fact that the k th order derivative of the coefficients is also assumed to be globally Lipschitz continuous. This is

enough to guarantee one additional layer of Malliavin differentiability. For details, consult Nualart [33, p119-124]. ■

As was briefly mentioned, much of the analysis relies on the observation that the expressions for the Malliavin derivative and the Jacobian of the stochastic flow are similar.

Corollary 1.14 *There holds for any $(s, t, x) \in [0, \infty)^2 \times \mathbb{R}^N$, where $s \leq t$:*

$$(J_{s,t}^x)^{-1}DX_{s,t}^x = \left(\int_{s \wedge \cdot}^{t \wedge \cdot} (J_{s,u}^x)^{-1}V_j(u, X_{s,u}^x)du \right)_{j=1, \dots, d}. \quad (1.20)$$

Proof : This is shown in, for example, Nualart [33, Section 2.3.1]. ■

This alternative representation for $(J_{s,t}^x)^{-1}DX_{s,t}^x$ appears naturally in any attempt to construct an integration by parts formula. Indeed, the development of this expression is the basis for forming an integration by parts formula (IBPF). At this point we cease the review of the basic theory of SDE regularity, as any further development would compromise the varying goals of the later chapters. In the next section some requisite notation is introduced, along with some concepts which are basic to the geometry of the problem; including a very general condition, known as the UFG condition.

THE BEGINNINGS OF AN INTEGRATION BY PARTS FORMULA

The connections developed at the end of the last section are now used to compute an integration by parts formula. To begin with, let $f \in C_b^\infty(\mathbb{R}^N)$, then by using the chain rule for the Malliavin derivative, the chain rule for the gradient operator, differentiability properties of the stochastic flow (note it is assumed that $k \geq 1$) and the representation (1.20), deduce:

$$\begin{aligned} Df(X_{s,t}^x) &= (\nabla f)(X_{s,t}^x)DX_{s,t}^x \\ &= \nabla(f \circ X_{s,t})(x)(J_{s,t}^x)^{-1}DX_{s,t}^x \\ &= \nabla(f \circ X_{s,t})(x) \left(\int_{s \wedge \cdot}^{t \wedge \cdot} (J_{s,u}^x)^{-1}V_i(u, X_{s,u}^x)du \right)_{i=1, \dots, d}, \end{aligned} \quad (1.21)$$

where, as is the case throughout, the gradient $\nabla f = (\partial_1 f, \dots, \partial_N f)$ is taken to be a row vector. The idea is as follows: to develop the preceding equality to isolate terms involving $\nabla(f \circ X_{s,t})(x)$ ⁵. Once isolated the operators of the Malliavin calculus will be used to derive an integration by parts formula. Before this is done, some new concepts and notation need to be introduced.

⁵as we shall see, whether the gradient itself may be isolated depends on the vector fields $V_0, \dots, V_d$.

1.6. Multi-indices, Lie brackets and the UFG condition

This ideas and notations introduced in this section play a central rôle in the thesis as a whole. The UFG condition in particular is a condition which will appear throughout in various variants and forms.

MULTI-INDICES

Multi-indices are a notational necessity in what follows. Define

$$\mathcal{A} := \{1, \dots, d\} \cup \bigcup_{k \in \mathbb{N}} \{0, 1, \dots, d\}^k,$$

$$\mathcal{A}_{0,\emptyset} := \mathcal{A} \cup \{0\} \cup \{\emptyset\}.$$

It is prudent to introduce a product, $*$, defining it as follows:

$$\alpha * \beta := (\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_l),$$

where $\alpha = (\alpha_1, \dots, \alpha_k), \beta = (\beta_1, \dots, \beta_l)$. Combined with this operation, $(\mathcal{A}_{0,\emptyset}, *)$ forms a semi-group. It shall be very useful to assign several notions of size to the multi-indices. Consider the following:

$$|\alpha| := \begin{cases} k & \text{if } \alpha = (\alpha_1, \dots, \alpha_k), \\ 0 & \text{if } \alpha = \emptyset. \end{cases}$$

and also the related:

$$\|\alpha\| := |\alpha| + \text{card } \{i : \alpha_i = 0, i = 1, \dots, d\}.$$

The reason for these choices should become more obvious when gradient bounds are discussed. As shall be demonstrated, the asymptotic rates derived in the gradient bounds are expressible in terms of these ‘sizes’. For a preview, note that it is necessary to compute expectations of iterated Stratonovich integrals, and the behaviour of stochastic integrals and normal Lebesgue-Stieltjes integrals is quantitatively different. The latter is designed to allow for this fact. Denote several more families of multi-indices related to these, namely:

$$\mathcal{A}(m) := \{\alpha \in \mathcal{A} : \|\alpha\| \leq m\},$$

$$\mathcal{A}_{0,\emptyset}(m) := \{\alpha \in \mathcal{A}_{0,\emptyset} : \|\alpha\| \leq m\}.$$

It will also be useful to have a set notation for the cardinality of $\mathcal{A}(m)$ and $\mathcal{A}_{0,\emptyset}(m)$. Define:

$$N_m := |\mathcal{A}(m)|,$$

$$N_m^{0,\emptyset} := |\mathcal{A}_{0,\emptyset}(m)|.$$

It may be shown that:

$$N_m := \begin{cases} M_m - M_{m-2} & \text{if } m \geq 2 \\ M_m & \text{if } m = 1 \end{cases}$$

where

$$M_m = \sum_{i=1}^m \sum_{j=0}^{\lfloor i/2 \rfloor} \binom{i-j}{j} d^{i-2j}.$$

To complete this section, observe that, trivially: $N_m^{0,\emptyset} = N_m + 2$.

LIE BRACKETS AND THEIR ALGEBRA

Lie brackets, and the family they form, play an important role in smoothness results of the semigroup. Taking the Lie bracket involves forming a new vector field from two existing vector fields. Moreover, it involves considering the derivatives of two existing vector fields. For this section, it is assumed that the operations involved in taking Lie brackets are well-defined. That is, that the vector fields in question are 'smooth enough'.

Definition 1.15 (Lie bracket of vector fields) *Given two vector fields: V, W . The Lie bracket of V and W is a third vector field, $[V, W]$, and is defined by:*

$$[V, W] := \partial W V - \partial V W,$$

where $\partial V := (\partial_j V^i)_{1 \leq i, j \leq N}$ and the multiplication is of a vector by a matrix.

It is equivalent to think of the vector field as a first order differential operator, given by: $Vf(x) := V(x) \cdot \nabla f(x)$ for $f \in C^1$. Where no confusion is possible, V shall be used to denote both the vector-valued function, and the differential operator. In the latter case the Lie bracket - a first order differential operator - has an even simpler representation:

$$[V, W] = VW - WV.$$

The Lie bracket possess the following basic properties:

$$[V, W] = -[W, V] \quad \text{and} \quad [U, [V, W]] + [W, [U, V]] + [V, [W, U]] = 0.$$

The latter is referred to as *the Jacobi Identity*.

Although it is not yet clear why one would care to do this, the Lie bracket operation may be iterated to form a family of vector fields, indexed by multi-index notation. That is, define $V_{[\alpha]}$,

$\alpha \in \mathcal{A}$ inductively, as follows:

$$\begin{aligned} V_{[i]} &:= V_i \quad i = 1, \dots, d. \\ V_{[\alpha * i]} &:= [V_{[\alpha]}, V_i], \quad i = 0, 1, \dots, d \end{aligned}$$

In a similar vein to the above, one can inductively define an iterated Stratonovich integral, $\hat{B}_{s,t}^{\circ\alpha}$, for $s, t \in [0, \infty)$, $s < t$, and $\alpha \in \mathcal{A}$:

$$\begin{aligned} \hat{B}_{s,t}^{\circ i} &:= \int_s^t \circ dB_u^i = B_t^i - B_s^i \\ \hat{B}_{s,t}^{\circ\alpha * i} &:= \int_s^t \hat{B}_u^{\circ\alpha} \circ dB_u^i, \end{aligned}$$

where $\hat{B}_t^{\circ\alpha} := \hat{B}_{0,t}^{\circ\alpha}$. For the corresponding iterated Itô integral, $\hat{B}_{s,t}^\alpha$ or $\hat{B}_t^\alpha$ shall be written.

THE UFG CONDITION

The purpose of the UFG condition, in its purest form, is to truncate the expansion obtained when considering the expression $(J_{s,t}^x)^{-1}V_i(t, X_{s,t}^x)$, for $i = 1, \dots, d$. Recalling the work of the previous section, this appeared when considering the ‘difference’, $(J_{s,t}^x)^{-1}DX_{s,t}^x$, between the Malliavin derivative and the Jacobian of the stochastic flow. The UFG condition is a ‘finite generation’ assumption, which - through careful application and study - leads to an equation which is used to form an integration by parts formula. We state the condition and briefly discuss its form:

Definition 1.16 (UFG Condition) *A system of vector fields $\{V_i : i = 0, \dots, d\}$ satisfy the UFG condition if, for any $\alpha \in \mathcal{A}$, there exists $m \in \mathbb{N}$ and $\varphi_{\alpha,\beta} \in C_b^\infty(\mathbb{R}^N)$, uniformly bounded, with $\beta \in \mathcal{A}(m)$ such that*

$$V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha,\beta}(x) V_{[\beta]}(x). \quad (1.22)$$

We write $(\mathbf{UFG}, m)$ to denote that the UFG condition holds for $m \in \mathbb{N}$.

The following example is taken from Kusuoka [22]:

Example 1.17 *Assume $d = 1$ and $N = 2$. Let $V_0, V_1 \in C_b^\infty(\mathbb{R}^2; \mathbb{R}^2)$ be given by*

$$V_0(x_1, x_2) = \sin x_1 \frac{\partial}{\partial x_1}, \quad V_1(x_1, x_2) = \sin x_1 \frac{\partial}{\partial x_2}.$$

*One should note that the Hörmander condition (see **(H')** on p74) is not satisfied. But the UFG condition is (for $m = 4$, see definition).*

The above form of UFG condition is, although concise, somewhat impractical. The statement requires one to verify that all elements of the Lie algebra (a countably infinite number) may be expressed in terms of a finite number. We seek to show that this isn't necessary, and that one needs

only check all Lie brackets on some 'level'. Note: Lemma 1.20 and the auxillary UFG conditions which it is based on is new.

Definition 1.18 (UFG1 Condition) We say that a system of vector fields $\{V_i : i = 0, \dots, d\}$ satisfy the UFG1 condition if there exists $m \in \mathbb{N}$ and $\varphi_{\alpha,\beta} \in C_b^\infty(\mathbb{R}^N)$, uniformly bounded, with $\beta \in \mathcal{A}(m)$ such that for any $\alpha \in \mathcal{A}$ satisfying $\|\alpha\| \in \{m+1, m+2\}$,

$$V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha,\beta}(x) V_{[\beta]}(x).$$

We write $(\mathbf{UFG1}, m)$ to denote that the UFG1 condition holds for $m \in \mathbb{N}$.

Definition 1.19 (UFG2 Condition) We say that a system of vector fields $\{V_i : i = 0, \dots, d\}$ satisfy the UFG2 condition if there exists $m \in \mathbb{N}$ and $\varphi_{\alpha,\beta} \in C_b^\infty(\mathbb{R}^N)$, uniformly bounded, with $\beta \in \mathcal{A}$ satisfying $|\beta| \leq m$, such that for any $\alpha \in \mathcal{A}$ satisfying $|\alpha| = m+1$,

$$V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha,\beta}(x) V_{[\beta]}(x).$$

We write $(\mathbf{UFG2}, m)$ to denote that the UFG2 condition holds for $m \in \mathbb{N}$.

Lemma 1.20 (Equivalence of various forms of UFG condition) There holds the following relationships:

1. $(\mathbf{UFG}, m) \Rightarrow (\mathbf{UFG}, n)$, for $m \leq n$.
2. $(\mathbf{UFG}, m) \Leftrightarrow (\mathbf{UFG1}, m)$.
3. $(\mathbf{UFG}, 2m-1) \Leftarrow (\mathbf{UFG2}, m)$.
4. $(\mathbf{UFG}, m) \Rightarrow (\mathbf{UFG2}, m)$.

Proof : 1. This is obvious once one notes that for all $\alpha \in \mathcal{A}$ such that $|\alpha| \notin \mathcal{A}(n)$, $V_{[\alpha]}$ may be expressed as the same linear combination for either case, i.e. if

$$V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha,\beta}(x) V_{[\beta]}(x).$$

Then

$$V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}(n)} \varphi_{\alpha,\beta}(x) V_{[\beta]}(x),$$

where $\varphi_{\alpha,\beta} \equiv 0$ for all $\beta \in \mathcal{A}(n) \setminus \mathcal{A}(m)$.

2. The implication $(\mathbf{UFG}, m) \Rightarrow (\mathbf{UFG1}, m)$ is trivial, as $\{\alpha ; \|\alpha\| \in \{m+1, m+2\}\} \subset \mathcal{A}$.

For the reverse case, we observe the decomposition $\mathcal{A} = \mathcal{A}(m) \cup \{\alpha ; \|\alpha\| \in \{m+1, m+2\}\} \cup \{\alpha ; \|\alpha\| > m+2\}$. If $\alpha \in \mathcal{A}(m)$, then it is clear that

$$V_{[\alpha]} = \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha, \beta} V_{[\beta]}$$

where $\varphi_{\alpha, \beta} \equiv 1$ if $\alpha = \beta$ and $\varphi_{\alpha, \beta} \equiv 0$ if $\alpha \neq \beta$. If α satisfies $\alpha \in \{m+1, m+2\}$, then (1.22) holds directly from (UFG1). If $\alpha \in \{\alpha ; \|\alpha\| > m+2\}$, then the result follows from induction. We do the first step for illustration purposes. Assume $\|\alpha\| = m+3$, then α has the form $\alpha = \alpha' * \alpha^*$ where $\|\alpha'\| = m+1$ if $\alpha^* = 0$, and $\|\alpha'\| = m+2$ if $\alpha^* \in \{1, \dots, d\}$. Therefore,

$$\begin{aligned} V_{[\alpha]} &= [V_{[\alpha']}, V_{\alpha^*}] \\ &= \left[\sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha', \beta} V_{[\beta]}, V_{\alpha^*} \right] \\ &= \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha', \beta} [V_{[\beta]}, V_{\alpha^*}] - (V_{\alpha^*} \varphi_{\alpha, \beta}) V_{[\beta]} \\ &= \left[\sum_{\substack{\beta \in \mathcal{A}(m) \\ \beta * \alpha' \in \mathcal{A}(m+2) \setminus \mathcal{A}(m)}} \varphi_{\alpha', \beta} V_{[\beta * \alpha']} \right] + \left[\sum_{\substack{\beta \in \mathcal{A}(m) \\ \beta * \alpha' \in \mathcal{A}(m)}} \varphi_{\alpha', \beta} V_{[\beta * \alpha']} - (V_{\alpha^*} \varphi_{\alpha, \beta}) V_{[\beta]} \right]. \end{aligned}$$

Again, we may apply the UFG1 condition, this time to the first term in the above (the undermost expression) to complete the argument. The proof for general α follows by induction.

3. This proof is analogous to the proof in **2**. The most important difference here is that $\{\alpha \in \mathcal{A} : |\alpha| \leq m\} \subset \{\alpha \in \mathcal{A} : \|\alpha\| \leq 2m-1\}$. This is because $\|\cdot\|$ assigns extra weight to multi-indices which contain zeroes (and α may contain as many as $m-1$ zeroes).

4. This is clear once one notes that $\{\alpha \in \mathcal{A} : \|\alpha\| \leq m\} \subset \{\alpha \in \mathcal{A} : |\alpha| \leq m\}$

■

This lemma permits various ways to proof to satisfaction of the UFG condition. In the next section we discuss the primary motivation of the gradient bounds.

1.7. The Kusuoka-Lyons-Victoir method

CUBATURE ON THE WIENER SPACE

This numerical method, outlined in Lyons and Victoir [29], is an extension of Stroud's (cf.[38]) cubature formula for finite-dimensional positive measures to measures on the *infinite-dimensional* Wiener space. It seeks to significantly reduce the numerical effort required to compute expectations over an infinite dimensional space, by proving that computing expectations of functionals on this

space reduces to a consideration of certain key points. These points are then identified by way of applying the algorithm described in the paper [29].

Cubature on the Wiener space has many notable advantages over conventional numerical methods. In particular, when computing the expectation of functionals of diffusions whose coefficients do not satisfy the Hörmander condition, the conventional numerical methods prove naïve as they are blind to the irregularity of such solutions.

It is certainly easy to see why cubature would be useful in mathematical finance, where expectation computation of functionals of diffusions is fundamental to the pricing of derivatives. But many problems in physics and other fields also require such methods. If one focusses on the equivalence between solving parabolic partial differential equations and the integration of functionals on the Wiener space - in particular, recall 1.2 and 1.3 - then it is easy to see why this method transcends disciplinary boundaries. We shall denote by $C_0^{BV}([0, T] ; \mathbb{R}^d)$ the subset of Ω consisting of all the paths of bounded variation. The definition of a cubature formula is, as follows:

Definition 1.21 Fix $m \in \mathbb{N}$. It is said that the paths $\omega_1, \dots, \omega_n \in C_0^{BV}([0, T] ; \mathbb{R}^{d+1})$ and the positive weights $\lambda_1, \dots, \lambda_n$ define a 'cubature formula on the Wiener space of degree m at time T ' if and only if, for all $\alpha = (\alpha_1, \dots, \alpha_k) \in \mathcal{A}(m)$,

$$\mathbb{E} \left[\hat{B}_T^{\circ\alpha} \right] = \sum_{j=1}^n \lambda_j \int_{0 < t_1 < \dots < t_k < T} d\omega_j^{\alpha_1}(t_1) \dots d\omega_j^{\alpha_k}(t_k) =: \mathbb{E}_{\mathbb{Q}} \left[\hat{B}_T^{\circ\alpha} \right]$$

i.e. expectations under the Wiener measure of iterated Stratonovich integrals of degree $\leq m$ are the same as under the finitely supported measure $\mathbb{Q} := \sum_{j=1}^n \lambda_j \delta_{\omega_j}$. Note: $\omega_0(t) := t$ in the above.

Theorem 1.22 (Lyons-Victoir Theorem) There exists a cubature formula of degree m at time T , $\forall m \in \mathbb{N}$, $T \in (0, \infty)$, such that $n \leq N_m + 1$, where n is the number of paths of the cubature formula.

Proof : See Lyons and Victoir [29]. ■

Remark 1.23 Extending a cubature formula of a given degree and time T to another cubature formula of same degree, but different time, is obtained through the scaling invariance of Brownian motion. Indeed, suppose we have a cubature formula of degree m at time $T = 1$, then the measure supported on the paths $w_{T,i}$ given by $\omega_{T,i}^0(t) := t$ and $\omega_{T,i}^j(t) := \sqrt{T} \omega_i^j(t/T)$ for $j = 1, \dots, d$, and unchanged weights, is a cubature formula for general T . Thus, one only needs to find a cubature formula for $T = 1$.

The authors then proceed to provide an algorithm for computing such cubature formulae. The work, and proofs of its global convergence are based on the Stratonovich Taylor expansion of $f(X_t^x)$. Define:

$$V_{\alpha} f := \begin{cases} f & \text{if } \alpha = \emptyset \\ V_{\alpha_1} \dots V_{\alpha_k} f & \text{if } \alpha = (\alpha_1, \dots, \alpha_k) \end{cases}.$$

Lemma 1.24 Assume $f \in C_b^\infty$, and $V_0, \dots, V_d \in C_b^\infty$, uniformly bounded, then

$$f(X_t^x) = \sum_{\alpha \in \mathcal{A}_m} V_\alpha f(x) \hat{B}_t^{\circ \alpha} + R_m(t, x, f), \quad (1.23)$$

where

$$R_m(t, x, f) = \sum_{\substack{i=0, \dots, d \\ \alpha \in \mathcal{A}_m \\ \text{s.t. } i * \alpha \notin \mathcal{A}_m}} \int_{0 < t_0 < \dots < t_k < t} V_{i * \alpha} f(X_{t_0}^x) \circ dB_{t_0}^i \circ dB_{t_1}^{\alpha_1} \dots \circ dB_{t_k}^{\alpha_k},$$

satisfies:

$$\sup_{x \in \mathbb{R}^N} \left(\mathbb{E} R_m(t, x, f)^2 \right)^{1/2} \leq C \sum_{j=m+1}^{m+2} t^{j/2} \sup_{\alpha \in \mathcal{A}(j) \setminus \mathcal{A}(j-1)} \|V_\alpha f\|_\infty.$$

Proof : See Lyons and Victoir [29]. ■

When one has a cubature formula, estimation of $\mathbb{E}[f(X_t^x)]$ will depend on the values of X_t^x at each of the bounded variation paths forming the support cubature. i.e. $X_t^x(\omega_i)$ will be the solution at time t of the ODE:

$$dy_t(x) = \sum_{i=0}^d V_i(y_t(x)) d\omega_i(t), \quad y_0(x) = x.$$

This surprising transformation may be emphasised by writing $\Phi_{t,x}(\omega_i) := X_t^x(\omega_i)$. It may be easily shown that the global error of this numerical approximation in its current form is:

Proposition 1.25

$$\left\| \mathbb{E} [f(X_t^{(\cdot)})] - \mathbb{E}_{\mathbb{Q}} [f(X_t^{(\cdot)})] \right\|_\infty \leq C \sum_{j=m+1}^{m+2} t^{j/2} \sup_{\alpha \in \mathcal{A}_j \setminus \mathcal{A}_{j-1}} \|V_\alpha f\|_\infty. \quad (1.24)$$

Proof : Consult Lyons and Victoir [29] ■

Unfortunately, the term on the right hand side of the above inequality is small only if t and/or the driving noise of the SDE is small. For this reason the approach is refined and the time interval is partitioned, and the iterated application of cubature on each subinterval is considered. One will have noticed that the gradient bounds have yet to be utilised. It is for this iteration procedure that the gradient bounds are vital.

ITERATED APPLICATION OF CUBATURE: THE KLV METHOD

It is not always the case that one wishes to compute expectations of SDE functionals for smooth test functions. Indeed, in finance, the price of a European call option can be expressed by terms of:

$$\mathbb{E}(S_t - K)^+ \quad \text{where } S_t \text{ is the underlying stock price.}$$

Notice that the functional here is not smooth (but it is Lipschitz). In this case, the upper bound given in (1.24) is meaningless. One does know, however, that even if f is not smooth, that $P_t f$ is (at least in the direction of the vector fields of the SDE, under the UFG condition). An extension of the cubature method which uses this property is needed.

The algebraic description of this algorithm, as presented here, is taken from Litterer [27]. The original algorithm is described in Lyons and Victoir [29]

Definition 1.26 (Kusuoka-Lyons-Victoir operation) *Given a positive measure $\mu = \sum_{i=1}^l \mu_i \delta_{x_i}$ on $\mathbb{R}^N$, and a cubature measure $\mathbb{Q} = \sum_{j=1}^n \lambda_j \delta_{\omega_j}$, we define the KLV operation wrt μ over a time-step s by:*

$$KLV(\mu, s) := \sum_{i=1}^l \sum_{j=1}^n \mu_i \lambda_j \delta_{X_s^{x_i}(\omega_{s,j})}. \quad (1.25)$$

Note that KLV takes discrete measures on $\mathbb{R}^N$ to discrete measures on $\mathbb{R}^N$. It follows directly from this definition, that:

$$\mathbb{E}_{\mathbb{Q}} f(X_s^x) = \mathbb{E}_{KLV(\delta_x, s)} f.$$

Now consider a partition, $\mathcal{D}$, of the interval $[0, T]$, into $k - 1$ sub-intervals:

$$0 = t_0 < t_1 < \dots < t_k = T.$$

Define $s_j := t_j - t_{j-1}$, and denote by $\mathcal{D}^j \subset \mathcal{D}$, the subpartitions:

$$0 = t_0 < t_1 < \dots < t_j,$$

for $j = 1, \dots, k$.

Definition 1.27 *Define the KLV operation with respect to a partition $\mathcal{D}$, starting from the pointmass x , recursively by:*

$$\begin{aligned} KLV(\mathcal{D}^1, x) &:= KLV(\delta_x, s_1) \\ KLV(\mathcal{D}^{j+1}, x) &:= KLV(\mathcal{D}^j, s_{j+1}). \end{aligned} \quad (1.26)$$

Remark 1.28 *Given the previous discussion, this iterated application of cubature can be interpreted as a particle system on $\mathbb{R}^N$, which branches in the form of an n -ary tree. To compute the k th step of the KLV approximation, one needs to have solved $(n^{k+1} - 1)/(n - 1)$ ODEs, each determined by the cubature path ω_{i,s_j} for $1 \leq i \leq n$, $1 \leq j \leq k$. Indeed, this is one drawback of cubature-based methods. The computational effort required to solve an exploding number of ODEs can quickly become a problem. Amendments to cubature-based methods to address this problem have been made by Litterer [27] using a method called ‘recombination’. Figure 1.4 illustrates an example of a cubature path.*

The KLV operation satisfies a Markovian property (due to the Markovian property of the underlying SDE). This allows one to control the error over the global interval $[0, T]$, in terms of a constant

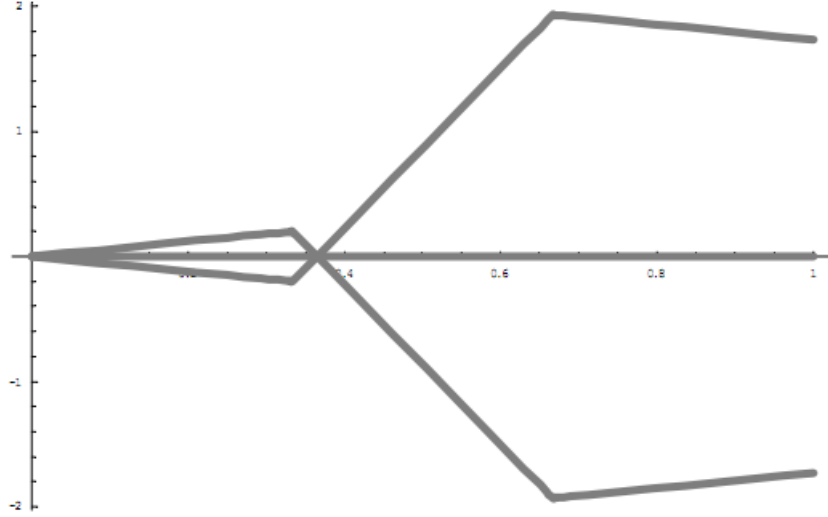


Figure 1.4.: Cubature formula of degree 5 for one Brownian motion

multiplied by the sum of the errors over the subintervals of the partition. This fact can be expressed rigorously as:

Proposition 1.29 *The KLV approximation of degree m satisfies*

$$\|(P_T f)(\cdot) - \mathbb{E}_{KLV(\mathcal{D}, \cdot)} f\|_\infty \leq C \sum_{l=1}^k \sum_{j=m+1}^{m+2} s_l^{j/2} \sup_{\alpha \in \mathcal{A}_j \setminus \mathcal{A}_{j-1}} \|V_\alpha P_{T-t_l} f\|_\infty. \quad (1.27)$$

Proof : See Lyons and Victoir [29] or Litterer and Lyons [28]) for full details. ■

A PROBLEM WITH THE STRATONOVICH TAYLOR EXPANSION

One now seeks to use the gradient bounds to provide an upper bound for the terms $\|V_\alpha P_{T-t_l} f\|_\infty$, $\alpha \in \mathcal{A}(m+2) \setminus \mathcal{A}(m)$. It is noted, however, that such $V_\alpha = V_{\alpha_1} \dots V_{\alpha_k}$ include the vector field V_0 for some V_α . The UFG condition is indeed very specific about which $V_{[\alpha]}$ one obtains a gradient bound for. This, as we shall see, certainly does not include V_0 . This issue was raised in Crisan and Ghazali [8], and the authors rectified this situation by assuming an extra condition; the so-called **V0** condition, which is stated below. In the Chapter 6 we adapt the following proof so that the **V0** condition can be completely relaxed.

Definition 1.30 (V0 condition) *A family of vector fields $V_i, 0 \leq i \leq d$ is said to satisfy the **V0** condition if for some $u_\beta \in C_b^\infty(\mathbb{R}^N)$, uniformly bounded, $\beta \in \mathcal{A}(2)$, there holds:*

$$V_0 = \sum_{\beta \in \mathcal{A}(2)} u_\beta V_{[\beta]}.$$

Therefore, if the **V0** condition is assumed it is easy to obtain a relevant gradient bound and to apply this with the aim of obtaining an upper bound for the global error. Indeed, from a gradient bound of the form (1.4) one can show the following:

Corollary 1.31 *There exists a constant C_m such that for all $\alpha \in \tilde{\mathcal{A}}_0(m)$*

$$\|V_\alpha P_t f\|_\infty \leq C_m \frac{t^{1/2}}{t^{\|\alpha\|/2}} \|\nabla f\|_\infty.$$

In even more recent developments (see Literer [27]), the author has shown that the **V0** condition on the drift can be relaxed and that one still obtains the same rate of convergence as shown above.

The following global bound for the error of the KLV method may be deduced:

Theorem 1.32 *Suppose the vector fields $V_i, 0 \leq i \leq d$ satisfy (UFG, m), then*

$$\|P_T f(\cdot) - \mathbb{E}_{KLV(\mathcal{D}, \cdot)} f\|_\infty \leq C_T \|\nabla f\|_\infty \left(s_k^{1/2} + \sum_{j=m}^{m+1} \sum_{i=1}^{k-1} \frac{s_i^{(j+1)/2}}{(T - t_i)^{j/2}} \right),$$

where the constant C_T is independent of the number of subintervals in the partition of $[0, T]$

Proof : For details consult for example, Ghazali [11]. ■

Finally, by taking uneven partitions of the interval $[0, T]$ one can derive high order convergence of the KLV method. Define, for $\gamma > 0$ and $0 \leq j \leq M$

$$t_j := T \left(1 - \left(1 - \frac{j}{M} \right)^\gamma \right).$$

This was pointed out in Kusuoka [21].

Corollary 1.33 *There holds the following global convergence rates:*

$$\begin{aligned} \|P_T f(\cdot) - \mathbb{E}_{KLV(\mathcal{D}, \cdot)} f\|_\infty &\leq K n^{-\gamma/2} \|\nabla f\|_\infty, & \text{if } 0 < \gamma < m - 1, \\ &\leq K n^{-(m-1)/2} \log(n) \|\nabla f\|_\infty, & \text{if } \gamma = m - 1, \\ &\leq K n^{-(m-1)/2} \|\nabla f\|_\infty, & \text{if } \gamma > m - 1. \end{aligned}$$

Proof : See Ghazali [11]. ■

Remark 1.34 *In Chapter 5, we state a local version of the UFG condition, called the LFG condition, which is used to derive local regularity results. Unfortunately, the author found no way to apply the local regularity results to prove fast convergence of the KLV method, apart from in very simple situations. For this reason, Chapter 5 does not contain a section on numerical schemes.*

2. The gradient bounds of Kusuoka

In this chapter the equations developed in the introductory section are expanded to give a full and self-contained proof of Kusuoka's gradient bounds (cf [22]). In later chapters, we shall apply these methods to new, more general problems, but this one remains the core on which the later calculations are based: both technically and ideologically.

Diffusion semigroups and bounds on their derivatives can, and have been, studied from many different approaches. They can broadly be separated into two different categories: those techniques which employ Malliavin calculus, and those which do not. In recent times, the main thrust from the latter group has come from Bakry and his collaborators. This group approach the geometric investigations of Markovian diffusion operators through functional analytic techniques, whose results and conditions are phrased in terms of a symmetric bilinear operator defined on $C^\infty \times C^\infty$, which is known as the "carré du champ" and denoted Γ . This operator was introduced by Roth [36] and Meyer [31]. If $\mathcal{L}$ is the generator of the semigroup P_t , i.e. $\mathcal{L}$ is a second order differential operator and is given by:

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

then the carré du champ is defined as follows:

$$\Gamma(f, g) := \mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f.$$

Roughly speaking, Γ can be thought of as measuring how far $\mathcal{L}$ is from being a derivation. The gradients bounds they deduce are phrased in terms of this operator, and its so-called "iteration", which replaces multiplication by taking the Γ 'product'. That is, define Γ_2 by:

$$2\Gamma_2(f, g) := \mathcal{L}\Gamma(f, g) - \Gamma(f, \mathcal{L}g) - \Gamma(g, \mathcal{L}f),$$

then it follows that $\Gamma_2 \geq R\Gamma$ for some $R \in \mathbb{R}$ if and only if for each $f \in C^\infty$ and $t \geq 0$ there holds: $\Gamma(P_t f) \leq e^{-2Rt} P_t(\Gamma f)$. Examples of operators which satisfy this condition are the classic heat semigroup, driven by the Laplace operator (for $R = 0$) and the Ornstein-Uhlenbeck semigroup (for $R = 1$). This result is proved in Bakry and Ledoux [3], and excellent summaries of these and related results can be found in, for example, Bakry [2] (in French) or Ledoux [26] (in English). Whilst these result are very interesting, they are not gradient bounds of the type we seek, and are not of the generality we wish for. The Malliavin calculus approach to gradient bounds has been pioneered

by Kusuoka and Stroock in [23], [24] and [25], and Kusuoka in [22]. Their utilisation of Malliavin calculus focusses on procuring an integration by parts formula. Moreover, the approach Kusuoka adopts in [22] is not only sensitive to the rates at which a general diffusion semigroup smooths, but is also dependent on a demonstrably general condition.

As most of the work in this section is an exhibition of the work of Kusuoka [22] and Kusuoka and Stroock [24], one should assume unless stated otherwise, that the proofs and results are owed to them. To aid clarity, we note the results from this section which are original: **Example 2.1**, **Theorem 2.40**.

2.1. Non-linearity breeds contempt

The idea behind theory is simple: to construct an integration by parts formula for the diffusion semigroup by using the operators of the Malliavin calculus. In this section the groundwork is laid for the deduction of an integration by parts formula. Key to the development of the formula is the creation of a linear system of equations, as permitted by the UFG condition.

USING THE UFG CONDITION

Care was taken in the first section to introduce the problem within a very general context. In this section one may assume the relative luxury of the following simplifications:

1. $V_0, \dots, V_d$ are time-homogeneous. That is, $V_i = V_i(x)$ for $i = 0, \dots, d$.
2. $V_i \in C_b^\infty$, and are uniformly bounded, for $i = 0, \dots, d$.

Based on these simplifying assumptions it is intuitive that the solution of (1.9) is a time-homogeneous Markov process and that $X_{s,t}^x = X_{0,t-s}^x$, $\mathbb{P}$ -a.s. This means that the resulting semigroup is a one-parameter semigroup, and to simplify the notation we write X_t^x instead of $X_{0,t}^x$, J_t^x instead of $J_{0,t}^x$, etc. Throughout this section it shall be assumed that **(UFG, m)** holds. Before developing the integration by parts formula, we provide an example of a stochastic differential equation for which the UFG condition holds, but the uniform Hörmander condition (see **(H')** on p74) does not.

Example 2.1 Consider the three-dimensional diffusion driven by the SDE

$$d \begin{bmatrix} X_t^{1,x} \\ X_t^{2,x} \\ X_t^{3,x} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} dt + \begin{bmatrix} X_t^{1,x} \\ 0 \\ 0 \end{bmatrix} dB_t^1 + \begin{bmatrix} 0 \\ X_t^{3,x} \\ 0 \end{bmatrix} dB_t^2 + \begin{bmatrix} 0 \\ 0 \\ X_t^{2,x} \end{bmatrix} dB_t^3,$$

$$\begin{bmatrix} X_0^{1,x} \\ X_0^{2,x} \\ X_0^{3,x} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

We now verify that the UFG condition is satisfied. Indeed, after identifying the vector fields, $V_0, \dots, V_3$ as:

$$V_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad V_1 = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 0 \\ x_3 \\ 0 \end{bmatrix}, \quad V_3 = \begin{bmatrix} 0 \\ 0 \\ x_2 \end{bmatrix}.$$

Then it is easy to verify that the Lie brackets of order $\|\alpha\| = 2, 3$ are given by:

$$V_{[1,0]}, V_{[1,2]}, V_{[1,3]}, V_{[2,0]} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad V_{[3,0]} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \quad V_{[2,3]} = \begin{bmatrix} 0 \\ -x_2 \\ x_3 \end{bmatrix}.$$

We claim that the UFG condition holds with $m = 3$. This is done by recalling the work from the previous chapter (cf Lemma 1.20), and instead showing the the UFG2 condition holds for $m = 2$. i.e. we show that that all Lie brackets of order $|\alpha| = 3$ can be expressed in the form (1.22):

$$\begin{aligned} V_{[\alpha]} &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \\ V_{[3,2,3]}, V_{[2,3,3]} &= \begin{bmatrix} 0 \\ 0 \\ \pm 2x_2 \end{bmatrix} = \pm 2V_3, \\ V_{[2,3,0]}, V_{[3,0,2]} &= \begin{bmatrix} 0 \\ 0 \\ \pm 1 \end{bmatrix} = \pm V_{[3,0]}. \end{aligned}$$

for $\alpha \in \{(2, 0, 0), (2, 0, 2), (2, 0, 3), (2, 2, 0), (2, 2, 2), (2, 2, 3), (2, 3, 2), (3, 0, 0), (3, 0, 3), (3, 2, 2), (3, 3, 2)\}$, or for α of the form $\alpha_i = 1$, for some i . Vector fields which are identically zero trivially satisfy equation (1.22) of the UFG condition. Hence we have demonstrated that the UFG condition holds.

Recalling the results developed in the preceding section, it holds that X_t is a flow of C^∞ -diffeomorphisms on $\mathbb{R}^N$ satisfying $X_t \in C^\infty(\mathbb{R}^N)$, and $X_t^x \in \mathbb{D}^\infty$. The goal is to create an integration by parts formula, and in particular recall (1.21):

$$Df(X_t^x) = (\nabla f)(X_t^x)DX_t^x = \nabla(f \circ X_t)(x) \left(\int_0^{t \wedge \cdot} (J_u^x)^{-1} V_i(X_u^x) du \right)_{i=1, \dots, d}.$$

It was mentioned that the basic idea was to isolate terms containing $\nabla(f \circ X_t)(x)$ and use the operators of the Malliavin calculus to deduce the IBPF. It is clear that any such attempt will have to explore the terms $(J_u^x)^{-1} V_i(X_u^x)$, $i = 1, \dots, d$. Indeed, by applying Itô's lemma for Stratonovich

integrals (cf. Nualart [33, p130,(2.63)]) one obtains:

$$\begin{aligned} d[(J_t^x)^{-1}V_i(X_t^x)] &= - \sum_{j=0}^d (J_t^x)^{-1}[V_i, V_j](X_t^x) \circ dB_t^j \\ &= - \sum_{j=0}^d (J_t^x)^{-1}V_{[i*j]}(X_t^x) \circ dB_t^j. \end{aligned}$$

This equation goes some way to exhibiting why the Lie bracket plays such an important role in this problem. Moreover, it is clear that Itô's lemma may be iteratively applied. Fix $\alpha \in \mathcal{A}$, then one has:

$$\begin{aligned} d[(J_t^x)^{-1}V_{[\alpha]}(X_t^x)] &= - \sum_{i=0}^d (J_t^x)^{-1}[V_{[\alpha]}, V_i](X_t^x) \circ dB_t^i \\ &= - \sum_{i=0}^d (J_t^x)^{-1}V_{[\alpha*i]}(X_t^x) \circ dB_t^i. \end{aligned}$$

Note that the application of Itô's lemma assumes certain differentiability of the vector fields $V_{[\alpha]}$, and hence of $V_0, \dots, V_d$. In later sections this is something we are careful of, but based on the simplifying assumptions that have been made, this presents no problem. A fair question at this stage is: *what do we do with this (potentially) infinite recursion?* Itô's lemma may be applied as many times as is wished, but this will gain no extra wisdom. This is where the UFG condition fulfills its task. The UFG condition curtails the infinite recursion, and creates a very tractable closed system of linear equations. It is through this linear system that the IBPF is born. For expressions which involve terms such as $V_{[\alpha]}$, where $\|\alpha\| > m$, as the UFG condition permits, the relationship:

$$V_{[\alpha]} = \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha, \beta} V_{[\beta]}$$

shall be used. Expressions which involve only Lie brackets such as $V_{[\alpha]}$, where $\|\alpha\| \leq m$, shall be left untouched. That is:

$$\begin{aligned} d[(J_t^x)^{-1}V_{[\alpha]}(X_t^x)] &= - \sum_{i=0}^d (J_t^x)^{-1}[V_{[\alpha]}, V_i](X_t^x) \circ dB_t^i \\ &= - \sum_{i=0}^d (J_t^x)^{-1}V_{[\alpha*i]}(X_t^x) \circ dB_t^i \\ &= \sum_{i=0}^d \sum_{\beta \in \mathcal{A}(m)} c_{\alpha, \beta}^i(X_t^x) (J_t^x)^{-1}V_{[\beta]}(X_t^x) \circ dB_t^i, \end{aligned}$$

where $c_{\alpha,\beta}^i$ have been chosen in line with the aforementioned rationale. Precisely,

$$c_{\alpha,\beta}^i(x) = \begin{cases} -1 & \text{if } \alpha * i \in \mathcal{A}(m) \text{ and } \beta = \alpha * i, \\ 0 & \text{if } \alpha * i \in \mathcal{A}(m) \text{ and } \beta \neq \alpha * i, \\ -\varphi_{\alpha * i, \beta} & \text{if } \alpha * i \notin \mathcal{A}(m). \end{cases} \quad (2.1)$$

In particular, it is noted that $c_{\alpha,\beta}^i \in C_b^\infty(\mathbb{R}^N)$, uniformly bounded. The key aspect to this representation is that a closed linear system of equations has been formed. That is, one may express any $(J_t^x)^{-1}V_{[\alpha]}(X_t^x)$ in terms of a finite linear combination of similar terms. Indeed, this construction can be abstractified by considering the above as the solution to the vector SDE:

$$\begin{cases} dY_t^x &= \sum_{i=0}^d C^i(X_t^x) Y_t^x \circ dB_t^i, \\ Y_0^x &= V(x), \end{cases} \quad (2.2)$$

where $Y(0, x) = V(x) := (V_{[\alpha]}(x))_{\alpha \in \mathcal{A}(m)} \in \mathbb{R}^{N_m} \otimes \mathbb{R}^N$ and $C^i : \mathbb{R}^N \rightarrow \mathbb{R}^{N_m} \otimes \mathbb{R}^{N_m}$, and is given by:

$$C^i(x) := (c_{\alpha,\beta}^i(x))_{\alpha,\beta \in \mathcal{A}(m)}.$$

That is, the terms $(J_t^x)^{-1}V_{[\alpha]}(X_t^x)$ are entries in the random vector Y_t^x . One is able to take advantage of the linear nature of this system of SDEs even further, by considering the dynamics of the matrix which drives this random vector.

Lemma 2.2 *Assume that $A(t, x)$ is, for each $(t, x) \in [0, \infty) \times \mathbb{R}^N$, the unique solution to the matrix stochastic differential equation.*

$$\begin{cases} dA(t, x) &= \sum_{i=0}^d C^i(X_t^x) A(t, x) \circ dB_t^i, \\ A(0, x) &= I. \end{cases} \quad (2.3)$$

Then $Y(t, x) = A(t, x)Y(0, x)$. In particular,

$$(J_t^x)^{-1}V_{[\alpha]}(X_t^x) = (A(t, x)Y(0, x))_\alpha = \sum_{\beta \in \mathcal{A}(m)} a_{\alpha,\beta}(t, x) V_{[\beta]}(x),$$

where $a_{\alpha,\beta}(t, x) := (A(t, x))_{\alpha,\beta}$.

Proof : It is easy to show $A(t, x)Y(0, x)$ solves equation (2.2), then by uniqueness of SDE solutions (see, for example Karatzas and Shreve [17]), the result follows. ■

The above results demonstrate that all the relevant information about the solution (2.2) is captured by (2.3). It is clear that a careful consideration of the random matrix $A(t, x)$ is required. The same classical results from Kunita [19] about SDE solutions, used in the first section for the stochastic flow (cf Theorem 1.12), may be applied to obtain the following proposition.

Proposition 2.3 *The matrix stochastic differential equation (2.3) has a unique solution, and its components $a_{\alpha,\beta} : [0, \infty) \times \mathbb{R}^N \rightarrow \mathbb{R}$, $\alpha, \beta \in \mathcal{A}(m)$ satisfy the mutually dependent SDEs:*

$$a_{\alpha,\beta}(t, x) = \delta_{\alpha,\beta} + \sum_{i=0}^d \sum_{\gamma \in \mathcal{A}(m)} \int_0^t c_{\alpha,\gamma}^i(X_s^x) a_{\gamma,\beta}(s, x) \circ dB_s^i.$$

Alternatively in Itô form:

$$\begin{aligned} a_{\alpha,\beta}(t, x) &= \delta_{\alpha,\beta} + \sum_{i=0}^d \sum_{\xi \in \mathcal{A}(m)} \int_0^t c_{\alpha,\xi}^i(X_s^x) a_{\xi,\beta}(s, x) dB_s^i \\ &\quad + \frac{1}{2} \sum_{i=1}^d \sum_{\xi \in \mathcal{A}(m)} \int_0^t (V_i c_{\alpha,\xi}^i)(X_s^x) a_{\xi,\beta}(s, x) ds \\ &\quad + \frac{1}{2} \int_0^t \sum_{i=1}^d \sum_{\xi, \tilde{\xi} \in \mathcal{A}(m)} (c_{\alpha,\xi}^i c_{\tilde{\xi},\beta}^i)(X_s^x) a_{\xi,\beta}(s, x) ds. \end{aligned}$$

Moreover, there is a version of $a_{\alpha,\beta}(t, \cdot) : \mathbb{R}^N \rightarrow \mathbb{R}$ which is smooth: both with respect to standard differentiation and Malliavin differentiation, that $a_{\alpha,\beta}(\cdot, \cdot)$ is jointly continuous in $[0, \infty) \times \mathbb{R}^N$ with probability one, for each $\alpha, \beta \in \mathcal{A}(m)$, and

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in [0, T]}} \mathbb{E} \left| \frac{\partial^{|\gamma|}}{\partial x^\gamma} a_{\alpha,\beta}(t, x) \right|^p < \infty, \quad \forall p \in [1, \infty), T > 0, \quad (2.4)$$

for any multi-index γ . Finally, for any $k \in \mathbb{N}$

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in [0, T]}} \mathbb{E} \left\| D^k a_{\alpha,\beta}(t, x) \right\|_{H^{\otimes k}}^p < \infty, \quad \forall p \in [1, \infty), T > 0. \quad (2.5)$$

It is also clear that such a matrix A must be invertible, as the inverse can be easily identified.

Corollary 2.4 *The matrix $A = (a_{\alpha,\beta})_{\alpha,\beta \in \mathcal{A}(m)}$ in Lemma 2.2 is invertible, and its inverse satisfies the matrix SDE:*

$$B(t, x) = I - \sum_{i=0}^d \int_0^t B(s, x) C^i(X_s^x) \circ dB_s^i.$$

Moreover, the components $b_{\alpha,\beta}$ of B , $\alpha, \beta \in \mathcal{A}(m)$, are $\mathbb{P}$ -a.s. smooth in x for fixed $t \in [0, \infty)$ and jointly continuous on $[0, \infty) \times \mathbb{R}^N$, with

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in [0, T]}} \mathbb{E} \left| \frac{\partial^{|\gamma|}}{\partial x^\gamma} b_{\alpha,\beta}(t, x) \right|^p < \infty, \quad p \in [1, \infty), T > 0. \quad (2.6)$$

Finally, for any $k \in \mathbb{N}$

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in [0, T]}} \mathbb{E} \left\| D^k b_{\alpha, \beta}(t, x) \right\|_{H^{\otimes k}}^p < \infty, \quad \forall p \in [1, \infty), T > 0, \quad (2.7)$$

The solution (2.3) must now be studied, as its elements are fundamental to the analysis. It is noted initially, that although this matrix is potentially very large, with potentially significant mutual dependence, many of the terms which make up this mutual dependence are zero. Note that for fixed $\alpha, \beta \in \mathcal{A}(m)$ there holds

$$a_{\alpha, \beta}(t, x) = \delta_{\alpha\beta} + \sum_{i=0}^d \sum_{\gamma \in \mathcal{A}(m)} \int_0^t c_{\alpha, \gamma}^i(X_s^x) a_{\gamma, \beta}(s, x) \circ dB_s^i. \quad (2.8)$$

But based on the coefficients, described in (2.1), the following may be identified:

For $\|\alpha\| \leq m - 2$ there holds: $\|\alpha * i\| \leq m$ for all $i = 0, \dots, d$, so $c_{\alpha, \gamma}^i \neq 0$ only when $\gamma = \alpha * i$, in which case $c_{\alpha, \gamma}^i = -1$. i.e.

$$a_{\alpha, \beta}(t, x) = \delta_{\alpha\beta} + \sum_{i=0}^d \int_0^t (-1) a_{\alpha * i, \beta}(s, x) \circ dB_s^i.$$

For $\|\alpha\| = m - 1$ there holds: $\|\alpha * i\| = m$ for $i = 1, \dots, d$, with $\|\alpha * 0\| = m + 1$. Hence $\alpha * i \in \mathcal{A}(m)$ for $i = 1, \dots, d$, and $\alpha * 0 \notin \mathcal{A}(m)$. i.e.

$$\begin{aligned} a_{\alpha, \beta}(t, x) &= \delta_{\alpha\beta} + \sum_{i=1}^d \int_0^t (-1) a_{\alpha * i, \beta}(s, x) \circ dB_s^i \\ &\quad - \sum_{\gamma \in \mathcal{A}(m)} \int_0^t \varphi_{\alpha * 0, \gamma}(X_s^x) a_{\gamma, \beta}(s, x) ds. \end{aligned}$$

For $\|\alpha\| = m$ there holds: $\|\alpha * i\| > m$ for $i = 0, \dots, d$. Hence $\alpha * i \notin \mathcal{A}(m)$ for $i = 0, \dots, d$. i.e.

$$a_{\alpha, \beta}(t, x) = \delta_{\alpha\beta} + \sum_{i=0}^d \sum_{\gamma \in \mathcal{A}(m)} \int_0^t (-1) \varphi_{\alpha * i, \gamma}(X_s^x) a_{\gamma, \beta}(s, x) \circ dB_s^i.$$

An explicit form for $a_{\alpha, \beta}$ is sought and is easy to identify from (2.8). In fact, each element of the matrix may be split up into a sum of two terms: the term which arises from $\delta_{\alpha\beta}$ - a iterated Stratonovich integral - and a remainder term. That is, for any $\alpha, \beta \in \mathcal{A}(m)$,

$$a_{\alpha, \beta}(t, x) = a_{\alpha, \beta}^0(t, x) + r_{\alpha, \beta}(t, x), \quad (2.9)$$

where

$$a_{\alpha,\beta}^0(t,x) = \begin{cases} (-1)^{|\gamma|} \hat{B}_t^{\circ\gamma}, & \text{if } \beta = \alpha * \gamma \text{ for some } \gamma \in \mathcal{A}(m), \\ 0, & \text{otherwise.} \end{cases}$$

And

$$r_{\alpha,\beta}(t,x) = \sum_{\substack{\gamma \in \mathcal{A}, j=0,\dots,d \\ \text{s.t. } \|\alpha * \gamma\| \leq m \\ \|\gamma * j\| \geq m+1 - \|\alpha\| \\ \delta \in \mathcal{A}(m)}} \int \dots \int_{S_{k+1}} (-1)^{|\gamma|} c_{\alpha * \gamma, \delta}^j(X_s^x) a_{\delta,\beta}(s,x) \circ dB_s^j \circ dB_{s_1}^{\gamma_1} \dots dB_{s_k}^{\gamma_k},$$

where $S_{k+1} := \{(s_1, \dots, s_{k+1}) \in [0, t]^{k+1} ; 0 \leq s_1 \leq \dots \leq s_{k+1} \leq t\}$. It will be beneficial to know the following about remainder terms:

Proposition 2.5 *For any $T > 0$, $p \in [1, \infty)$, $\alpha, \beta \in \mathcal{A}(m)$ and $\gamma \in \mathcal{A}$, the following hold*

$$\sup_{t \in (0, T]} \mathbb{E} \left[t^{-\|\gamma\|/2} \left| \hat{B}_t^{\circ\gamma} \right| \right]^p < \infty, \quad (2.10)$$

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in (0, T]}} \mathbb{E} \left[t^{-(m+1-\|\alpha\|)/2} |r_{\alpha,\beta}(t,x)| \right]^p < \infty. \quad (2.11)$$

Proof : The proof of the first part is straightforward once we observe, by the scaling invariance of Brownian motion, that the following relationship holds:

$$\hat{B}_{st}^{\circ\gamma} \stackrel{\mathcal{D}}{=} s^{\frac{\|\gamma\|}{2}} \hat{B}_t^{\circ\gamma},$$

and so $\hat{B}_t^{\circ\gamma} \stackrel{\mathcal{D}}{=} t^{\frac{\|\gamma\|}{2}} \hat{B}_1^{\circ\gamma}$. (2.10) follows easily. The proof of the second part is left to the appendix. ■

An equation which shall directly lead to the integration by parts formula may now be formed.

2.2. The integration by parts formula

The techniques in this section are guided by Kusuoka's original paper [22]. Assume initially $f \in C_b^\infty(\mathbb{R}^N)$, then from (1.21) we have

$$Df(X_t^x) = \nabla(f \circ X_t)(x) \left(\int_0^{t \wedge \cdot} (J_u^x)^{-1} V_i(X_u^x) du \right)_{i=1,\dots,d}.$$

The linear representation of Lemma 2.2 is used:

$$(J_u^x)^{-1} V_i(X_u^x) = (A(u, x) V(x))_i = \sum_{\beta \in \mathcal{A}(m)} a_{i,\beta}(u, x) V_{[\beta]}(x).$$

Hence,

$$\begin{aligned}
Df(X_t^x) &= \nabla(f \circ X_t)(x) \left(\int_0^{t \wedge \cdot} \sum_{\beta \in \mathcal{A}(m)} a_{i,\beta}(u, x) V_{[\beta]}(x) du \right)_{i=1, \dots, d} \\
&= \nabla(f \circ X_t)(x) \sum_{\beta \in \mathcal{A}(m)} V_{[\beta]}(x) \left(\int_0^{t \wedge \cdot} a_{i,\beta}(u, x) du \right)_{i=1, \dots, d} \\
&= \sum_{\beta \in \mathcal{A}(m)} V_{[\beta]}(f \circ X_t)(x) k_\beta(t, x),
\end{aligned}$$

where

$$\begin{aligned}
k_\beta(t, x) &:= \left(\int_0^{t \wedge \cdot} a_{i,\beta}(u, x) du \right)_{i=1, \dots, d} \\
V_{[\beta]}(f \circ X_t)(x) &:= \sum_{j=1}^d V_{[\beta]}^j(x) \partial_j(f \circ X_t)(x).
\end{aligned}$$

This single equation can be made into a linear system of equations by taking the H inner product with $k_\alpha(t, x)$ for all $\alpha \in \mathcal{A}(m)$. i.e. by taking $\alpha_1, \dots, \alpha_{N_m}$ to be an enumeration of the multi-indices of $\mathcal{A}(m)$:

$$\begin{aligned}
\langle Df(X_t^x), k_{\alpha_1}(t, x) \rangle_H &= \sum_{\beta \in \mathcal{A}(m)} V_{[\beta]}(f \circ X_t)(x) \langle k_\beta(t, x), k_{\alpha_1}(t, x) \rangle_H \\
&\vdots \\
\langle Df(X_t^x), k_{\alpha_n}(t, x) \rangle_H &= \sum_{\beta \in \mathcal{A}(m)} V_{[\beta]}(f \circ X_t)(x) \langle k_\beta(t, x), k_{\alpha_n}(t, x) \rangle_H \\
&\vdots \\
\langle Df(X_t^x), k_{\alpha_{N_m}}(t, x) \rangle_H &= \sum_{\beta \in \mathcal{A}(m)} V_{[\beta]}(f \circ X_t)(x) \langle k_\beta(t, x), k_{\alpha_{N_m}}(t, x) \rangle_H.
\end{aligned}$$

Define, for $\alpha \in \mathcal{A}(m)$, for notational simplicity:

$$\begin{aligned}
D^{(\alpha)} f(X_t^x) &:= \langle Df(X_t^x), k_\alpha(t, x) \rangle_H \\
M_{\alpha, \beta}(t, x) &:= t^{-\frac{\|\alpha\| + \|\beta\|}{2}} \langle k_\alpha(t, x), k_\beta(t, x) \rangle_H \\
&= t^{-\frac{\|\alpha\| + \|\beta\|}{2}} \sum_{i=1}^d \int_0^t a_{i,\alpha}(s, x) a_{i,\beta}(s, x) ds,
\end{aligned}$$

so that

$$D^{(\alpha)} f(X_t^x) = \sum_{\beta \in \mathcal{A}(m)} t^{\frac{\|\alpha\| + \|\beta\|}{2}} M_{\alpha, \beta}(t, x) V_{[\beta]}(f \circ X_t)(x).$$

A brief remark is made about the decision to multiply and divide by $t^{-\frac{\|\alpha\|+\|\beta\|}{2}}$. It will be seen that the presence of the iterated Stratonovich integrals in the expression for $a_{i,\alpha}(t, x)$ [cf (2.9)] will allow offsetting the behaviour of $M_{\alpha,\beta}(t, x)$ against $t^{-\frac{\|\alpha\|+\|\beta\|}{2}}$. What remains will form the aforementioned *rate* for the integration by parts formula. Although this step is technically unnecessary, it makes the proof easier and more transparent.

The above can be seen as a linear system of equations driven by a random matrix $M(t, x) = (M_{\alpha,\beta}(t, x))_{\alpha,\beta}$. It is hopefully clear that if this matrix can be inverted, then the IBPF will be very close. For then there would hold, $\mathbb{P}$ -a.s:

$$V_{[\alpha]}(f \circ X_t)(x) = t^{-\frac{\|\alpha\|}{2}} \sum_{\beta \in \mathcal{A}(m)} t^{-\frac{\|\beta\|}{2}} M_{\alpha,\beta}^{-1}(t, x) D^{(\beta)} f(X_t^x),$$

and the operators of the Malliavin calculus could be employed to obtain an IBPF for the derivatives of the diffusion semigroup **along the directions of the Lie algebra**. The next section is devoted to proving this result.

2.3. Invertibility of the Malliavin covariance matrix

The aim of this section is to prove the following proposition.

Proposition 2.6 *$M(t, x)$ is $\mathbb{P}$ -a.s. invertible. Moreover, for $p \in [1, \infty)$, $\alpha, \beta \in \mathcal{A}(m)$,*

$$\sup_{t \in (0,1], x \in \mathbb{R}^N} \mathbb{E} \left[M_{\alpha,\beta}^{-1}(t, x) \right]^p < \infty. \quad (2.12)$$

For real-symmetric matrices such as $M(t, x)$ there is an elegant representation of the minimal eigenvalue. The following lemma utilises this to simplify the requirements for invertibility.

Lemma 2.7 *The statement of the previous proposition holds, providing the following can be shown for each $p \in [1, \infty)$: there exists $C > 0$ s.t.*

$$\mathbb{P} \left(\inf_{|\xi|=1} (\xi, M(t, x)\xi) < \frac{1}{n} \right) < Cn^{-p},$$

for all $n \geq 1$, $t \in (0, 1]$, and $x \in \mathbb{R}^N$.

Proof : See appendix. ■

In view of these results, consider $(\xi, M(t, x)\xi)$. The determinant of $M(t, x)$ is non-negative and increasing with t . This means that if $M(t, x)$ is a.s. invertible for some $t > 0$, then it must be

invertible thereafter. Let $y \geq 1$.

$$\begin{aligned}
(\xi, M(t, x)\xi) &= \sum_{\alpha, \beta \in \mathcal{A}(m)} \xi_\alpha \xi_\beta M_{\alpha, \beta}(t, x) \\
&= \sum_{\alpha, \beta \in \mathcal{A}(m)} \xi_\alpha \xi_\beta t^{-(\frac{\|\alpha\|}{2} + \frac{\|\beta\|}{2})} \langle k_\alpha(t, x), k_\beta(t, x) \rangle_H \\
&= \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha t^{-\frac{\|\alpha\|}{2}} k_\alpha(t, x) \right\|_H^2 \\
&= \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha t^{-\frac{\|\alpha\|}{2}} \int_0^{t \wedge \cdot} (a_{\cdot, \alpha}^0 + r_{\cdot, \alpha})(u, x) du \right\|_H^2 \\
&\geq \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha t^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} (a_{\cdot, \alpha}^0 + r_{\cdot, \alpha})(u, x) du \right\|_H^2. \tag{2.13}
\end{aligned}$$

Observe that, since $y \geq 1$, using the notation: $S^n := \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$,

$$\begin{aligned}
&\inf_{\xi \in S^{N_m-1}} \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha t^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} a_{\cdot, \alpha}(u, x) du \right\|_H^2 \\
&\geq \inf_{\xi \in S^{N_m-1}} \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha \left[\frac{t}{y} \right]^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} a_{\cdot, \alpha}(u, x) du \right\|_H^2.
\end{aligned}$$

Now focus on the term appearing on the RHS:

$$\begin{aligned}
&\left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha \left[\frac{t}{y} \right]^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} a_{\cdot, \alpha}(u, x) du \right\|_H^2 \\
&\geq \frac{1}{2} \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha \left[\frac{t}{y} \right]^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} a_{\cdot, \alpha}^0(u, x) du \right\|_H^2 \\
&\quad - \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha \left[\frac{t}{y} \right]^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} r_{\cdot, \alpha}(u, x) du \right\|_H^2 \\
&\geq \frac{1}{2} \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha \left[\frac{t}{y} \right]^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} a_{\cdot, \alpha}^0(u, x) du \right\|_H^2 \\
&\quad - \left(\sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha^2 \right) \left(\int_0^{t/y} \sum_{\alpha \in \mathcal{A}(m)} \sum_{i=1}^d \left[\frac{t}{y} \right]^{-\|\alpha\|} r_{i, \alpha}(u, x)^2 du \right).
\end{aligned}$$

Recall that $a_{i,\alpha}^0(u, x) = 0$ whenever $\alpha \neq i * \gamma$ for all multiindices γ . Moreover, $a_{i,\alpha}^0(u, x) = \hat{B}_u^{\circ\gamma}$ when $i * \gamma = \alpha$. That is, as each multindex $\alpha \in \mathcal{A}(m)$ satisfies $\alpha_1 \in \{1, \dots, d\}$:

$$a_{.,\alpha}^0(u, x) = (0, \dots, \hat{B}_u^{\circ\gamma}, \dots, 0).$$

It is now necessary to briefly discuss the first term on the RHS. The following result is taken from Kusuoka and Stroock [24], but a comprehensive proof is provided in the next section.

Proposition 2.8 *Given $m \in \mathbb{N}$, there exist constants $C_m, \mu_m \in (0, \infty)$ such that for all $T > 0$*

$$\mathbb{P} \left(\inf_{a \in S^{N_{m-1}^{0,\emptyset}-1}} \int_0^T \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} T^{-\frac{\|\gamma\|}{2} - \frac{1}{2}} a_\gamma \hat{B}_t^{\circ\gamma} \right]^2 dt \leq \frac{1}{n} \right) \leq C_m \exp\{-n^{\mu_m}\}. \quad (2.14)$$

Proof : The proof of this result is very demanding. For a detailed proof, consult the next section. ■

As a result of this strong bound, which is incidentally much stronger than that which is required for invertibility, it is very easy to deduce the following two equivalent properties:

Corollary 2.9 *For any $m \in \mathbb{N}$, and $p \in [1, \infty)$, there holds:*

$$\mathbb{E} \left(\inf_{a \in S^{N_{m-1}^{0,\emptyset}-1}} \int_0^T \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} T^{-\frac{\|\gamma\|}{2} - \frac{1}{2}} a_\gamma \hat{B}_t^{\circ\gamma} \right]^2 dt \right)^{-p} < \infty. \quad (2.15)$$

And, equivalently, for all $q \in [1, \infty)$

$$\mathbb{P} \left(\inf_{a \in S^{N_{m-1}^{0,\emptyset}-1}} \int_0^t \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} T^{-\frac{\|\gamma\|}{2} - \frac{1}{2}} a_\gamma \hat{B}_t^{\circ\gamma} \right]^2 dt \leq \frac{1}{n} \right) < C_{m,q} n^{-q}. \quad (2.16)$$

The usefulness of the above might not be immediately clear, so turn attention back to the lower bound obtained for $(\xi, M(t, x)\xi)$. The fact that any $\alpha \in \mathcal{A}(m)$ can be expressed as $\alpha = j * \gamma$ for some $1 \leq j \leq d$ and $\gamma \in \mathcal{A}_{0,\emptyset}(m-1)$ is used. This allows the effective utilisation of the structure

of $a_{\cdot,\alpha}^0(t, x)$.

$$\begin{aligned}
& \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha \left[\frac{t}{y} \right]^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} a_{\cdot,\alpha}^0(u, x) du \right\|_H^2 \\
&= \left\| \sum_{j=1}^d \sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} \xi_{j*\gamma} \left[\frac{t}{y} \right]^{-\frac{\|\gamma\|+1}{2}} \int_0^{t/y \wedge \cdot} a_{\cdot,j*\gamma}^0(u, x) du \right\|_H^2 \\
&= \left\| \sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} \left[\frac{t}{y} \right]^{-\frac{\|\gamma\|+1}{2}} \int_0^{t/y \wedge \cdot} (\xi_{j*\gamma} \hat{B}_u^{\circ\gamma})_{j=1,\dots,d} du \right\|_H^2 \\
&= \left\| \int_0^{t/y \wedge \cdot} \left(\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} \left[\frac{t}{y} \right]^{-\frac{\|\gamma\|+1}{2}} \xi_{j*\gamma} \hat{B}_u^{\circ\gamma} \right)_{j=1,\dots,d} du \right\|_H^2 \\
&= \sum_{j=1}^d \int_0^{t/y \wedge \cdot} \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} \left[\frac{t}{y} \right]^{-\frac{\|\gamma\|+1}{2}} \xi_{j*\gamma} \hat{B}_u^{\circ\gamma} \right]^2 du.
\end{aligned}$$

It can also easily be shown that by taking $\inf_{\xi \in S^{N_{m-1}}}$ of both sides:

$$\begin{aligned}
& \inf_{\xi \in S^{N_{m-1}}} \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha \left[\frac{t}{y} \right]^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} a_{\cdot,\alpha}^0(u, x) du \right\|_H^2 \\
&= \inf_{\xi \in S^{N_{m-1}}} \sum_{j=1}^d \int_0^{t/y \wedge \cdot} \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} \left[\frac{t}{y} \right]^{-\frac{\|\gamma\|+1}{2}} \xi_{j*\gamma} \hat{B}_{s,u}^{\circ\gamma} \right]^2 du \\
&= \inf_{a \in S^{N_{m-1}^{0,\emptyset}-1}} \int_0^{t/y \wedge \cdot} \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} \left[\frac{t}{y} \right]^{-\frac{\|\gamma\|+1}{2}} a_\gamma \hat{B}_u^{\circ\gamma} \right]^2 du,
\end{aligned}$$

recalling that $N_{m-1}^{0,\emptyset} = N_{m-1} + 2$. This is precisely why the upper bound derived in Proposition 2.8 was introduced. It enables a precise control over the tail behaviour of $(\xi, M(t, x)\xi)$. The various pieces of analysis are now synthesised. In what follows, note that

$$\begin{aligned}
\mathbb{P} \left(\frac{1}{2}X - Y \leq \frac{1}{n} \right) &= \mathbb{P} \left(\frac{1}{2}X - Y \leq \frac{1}{n}, Y < \frac{1}{n} \right) + \mathbb{P} \left(\frac{1}{2}X - Y \leq \frac{1}{n}, Y \geq \frac{1}{n} \right) \\
&\leq \mathbb{P} \left(Y \geq \frac{1}{n} \right) + \mathbb{P} \left(X \leq \frac{4}{n} \right).
\end{aligned}$$

This gives:

$$\begin{aligned}
& \mathbb{P} \left(\inf_{\xi \in S^{N_{m-1}}} (\xi, M(t, x) \xi) < \frac{1}{n} \right) \\
& \leq \mathbb{P} \left(\inf_{a \in S^{N_{m-1}^{0, \emptyset} + 1 - 1}} \int_0^{t/y} \left[\sum_{\gamma \in \mathcal{A}_{0, \emptyset}(m-1)} \left[\frac{t}{y} \right]^{-\frac{\|\gamma\|+1}{2}} a_\gamma \hat{B}_u^{\circ \gamma} \right]^2 du < \frac{4}{n} \right) \\
& \quad + \mathbb{P} \left(\inf_{\xi \in S^{N_{m-1}}} \left[\sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha^2 \right] \left[\int_0^{t/y} \sum_{\alpha \in \mathcal{A}(m)} \sum_{i=1}^d \left[\frac{t}{y} \right]^{-\|\alpha\|} r_{i, \alpha}(u, x)^2 du \right] \geq \frac{1}{n} \right) \\
& = \mathbb{P} \left(\inf_{a \in S^{N_{m-1}^{0, \emptyset} + 1 - 1}} \int_0^{t/y} \left[\sum_{\gamma \in \mathcal{A}_{0, \emptyset}(m-1)} \left[\frac{t}{y} \right]^{-\frac{\|\gamma\|+1}{2}} a_\gamma \hat{B}_u^{\circ \gamma} \right]^2 du < \frac{4}{n} \right) \\
& \quad + \mathbb{P} \left(\int_0^{t/y} \sum_{\alpha \in \mathcal{A}(m)} \sum_{i=1}^d \left[\frac{t}{y} \right]^{-\|\alpha\|} r_{i, \alpha}(u, x)^2 du \geq \frac{1}{n} \right).
\end{aligned}$$

The program is almost complete. The following is deduced from Proposition 2.5,

Lemma 2.10 *There holds, for all $p \in [1, \infty)$,*

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in (0, 1]}} \mathbb{E} \left(\int_0^t \sum_{\alpha \in \mathcal{A}(m)} \sum_{i=1}^d t^{-\|\alpha\|-1} r_{i, \alpha}(u, x)^2 du \right)^p < \infty.$$

Proof : See appendix. ■

The proof can now be completed.

$$\begin{aligned}
& \mathbb{P} \left(\inf_{a \in S^{N_{m-1}^{0, \emptyset} + 1 - 1}} \left\{ \int_0^{t/y} \left[\sum_{\gamma \in \mathcal{A}_{0, \emptyset}(m-1)} \left[\frac{t}{y} \right]^{-\frac{\|\gamma\|+1}{2}} a_\gamma \hat{B}_u^{\circ \gamma} \right]^2 du \right\} < \frac{4}{n} \right) \\
& \quad + \mathbb{P} \left(\int_0^{t/y} \sum_{\alpha \in \mathcal{A}(m)} \sum_{i=1}^d \left[\frac{t}{y} \right]^{-\|\alpha\|} r_{i, \alpha}(u, x)^2 du \geq \frac{1}{n} \right) \\
& = \mathbb{P} \left(\inf_{a \in S^{N_{m-1}^{0, \emptyset} + 1 - 1}} \left\{ \int_0^{t/y} \left[\sum_{\gamma \in \mathcal{A}_{0, \emptyset}(m-1)} \left[\frac{t}{y} \right]^{-\frac{\|\gamma\|+1}{2}} a_\gamma \hat{B}_u^{\circ \gamma} \right]^2 du \right\} < \frac{4}{n} \right) \\
& \quad + \mathbb{P} \left(\int_0^{t/y} \sum_{\alpha \in \mathcal{A}(m)} \sum_{i=1}^d \left[\frac{t}{y} \right]^{-\|\alpha\|-1} r_{i, \alpha}(u, x)^2 du \geq \frac{y}{nt} \right) \\
& < C_{m, q} \left(\frac{4}{n} \right)^q + \tilde{C}_{m, q} \left(\frac{nt}{y} \right)^q \leq C_{m, q} \left(\frac{4}{n} \right)^q + \tilde{C}_{m, q} \left(\frac{n}{y} \right)^q.
\end{aligned}$$

It is important to note that the above bounds hold $\forall t \in (0, 1]$ and $\forall x \in \mathbb{R}^N$. The decision to introduce $y \geq 1$ should now become clear. Without it, the analysis would fail. Indeed, there is a

clever choice of y such that Lemma 2.7 holds. Set

$$y = \frac{n^2}{4},$$

so that

$$\frac{n}{y} = \frac{4}{n}.$$

And finally, combining this with the above we obtain:

$$\mathbb{P} \left(\inf_{\xi \in S^{N_{m-1}}} (\xi, M(t, x)\xi) < \frac{1}{n} \right) < \tilde{C}_{m,p} \frac{1}{n^q},$$

as required.

In the next section regularity results about the inverse of the matrix are proved. These results shall be fundamental to the integration by parts formula.

2.4. Diffuseness of iterated Stratonovich integrals

It was seen in the last section that invertibility of the Malliavin covariance matrix can be achieved if Proposition 2.8 holds. Its statement is recalled and it is sought to prove this result using the work of Kusuoka/Stroock in [24] as a guide.

Proposition 2.11 *For any $m \in \mathbb{N}$, and $p \in [1, \infty)$, there holds:*

$$\mathbb{E} \left(\inf_{a \in S^{N_{m-1}^{0, \emptyset}-1}} \int_0^T \left[\sum_{\gamma \in \mathcal{A}_{0, \emptyset}(m-1)} T^{-\frac{\|\gamma\|}{2} - \frac{1}{2}} a_\gamma \hat{B}_t^{\circ \gamma} \right]^2 dt \right)^{-p} = C_{m,p} < \infty. \quad (2.17)$$

And, equivalently, for all $q \in [1, \infty)$

$$\mathbb{P} \left(\inf_{a \in S^{N_{m-1}^{0, \emptyset}-1}} \int_0^T \left[\sum_{\gamma \in \mathcal{A}_{0, \emptyset}(m-1)} T^{-\frac{\|\gamma\|}{2} - \frac{1}{2}} a_\gamma \hat{B}_t^{\circ \gamma} \right]^2 dt \leq \frac{1}{n} \right) < C_{m,q} n^{-q}. \quad (2.18)$$

Proof : The proof of this important result is begun through simplification of the problem. By considering the distribution of the iterated Stratonovich integrals one is able to make a change of variable to the integral. Indeed, note that:

$$\hat{B}_{st}^{\circ \gamma} \stackrel{\mathcal{D}}{=} s^{\frac{\|\gamma\|}{2}} \hat{B}_t^{\circ \gamma},$$

Hence it may be deduced:

$$\begin{aligned}
& \int_0^T \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} T^{-\frac{\|\gamma\|}{2} - \frac{1}{2}} a_\gamma \hat{B}_t^{\circ\gamma} \right]^2 dt \\
& \stackrel{\mathcal{D}}{=} \int_0^T \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} T^{-\frac{1}{2}} a_\gamma \hat{B}_{\frac{t}{T}}^{\circ\gamma} \right]^2 dt \\
& \stackrel{u=t/T}{=} \int_0^1 \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} a_\gamma \hat{B}_u^{\circ\gamma} \right]^2 du.
\end{aligned}$$

Hence, proof of Proposition 2.11 is reduced to showing that for each $p \geq 1$, there exists $C > 0$ s.t.

$$\mathbb{P} \left(\inf_{a \in S_{N_{m-1}^{0,\emptyset}}^{-1}} \int_0^1 \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} a_\gamma \hat{B}_u^{\circ\gamma} \right]^2 du < \frac{1}{n} \right) < Cn^{-p}, \quad (2.19)$$

for all $n \geq 1$.

Iterated Stratonovich integrals arise in a very natural way from the geometry of this problem. That said, one must often turn to the more established results in stochastic integration to do an accurate analysis of them. These results are almost always phrased in terms of Itô integration and the semimartingales resulting therefrom. Hence, attention is switched to iterated Itô integrals via the following proposition. The moral of the story is that, although undoubtedly different objects, iterated Itô and Stratonovich integrals are equally as diffuse.

Proposition 2.12 Define $\hat{B}_t^{\circ L} := (\hat{B}_t^{\circ\alpha})_{\|\alpha\| \leq L}$ and $\hat{B}_t^L := (\hat{B}_t^\alpha)_{\|\alpha\| \leq L}$. Then, for all $L \in \mathbb{N}$ there exist constant matrices $A_L, \tilde{A}_L \in \mathbb{R}^{N_L \times N_L}$ such that

$$(i): \quad \hat{B}_t^{\circ L} = A_L \hat{B}_t^L \quad \text{and} \quad (ii): \quad \hat{B}_t^L = \tilde{A}_L \hat{B}_t^{\circ L}.$$

i.e. A_L is invertible with $A_L^{-1} = \tilde{A}_L$.

Moreover, it follows that the existence of constants $C_m, \mu_m \in (0, \infty)$

$$\mathbb{P} \left(\inf_{\sum a_\gamma^2 = 1} \int_0^1 \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} a_\gamma \hat{B}_t^{\circ\gamma} \right]^2 dt \leq \frac{1}{n} \right) \leq C_m \exp\{-n^{\mu_m}\},$$

is equivalent to the existence of constants $\tilde{C}_m, \tilde{\mu}_m \in (0, \infty)$ such that

$$\mathbb{P} \left(\inf_{\sum a_\gamma^2 = 1} \int_0^1 \left[\sum_{\gamma \in \mathcal{A}_{0,\emptyset}(m-1)} a_\gamma \hat{B}_t^\gamma \right]^2 dt \leq \frac{1}{n} \right) \leq \tilde{C}_m \exp\{-n^{\tilde{\mu}_m}\}.$$

Proof of (i) (adapted from the proof of Lemma A.12 in Kusuoka and Stroock [24]) :

(i) is approached by using an induction argument on L . Clearly if $L = 1$ then there is little to prove as $\hat{B}_t^{\circ L} = \hat{B}_t^L$. Hence, as $A_L = I_{d \times d} = \tilde{A}_L$. Now assume that the result holds for $L \leq k$. i.e. for all α such that $\|\alpha\| \leq k$ there holds, for some deterministic constants: $a_{\alpha,\beta}^k, \|\beta\| \leq k$.

$$\hat{B}_t^{\circ \alpha} = \sum_{\|\beta\| \leq k} a_{\alpha,\beta}^k \hat{B}_t^\beta.$$

It is clear one need only prove, for suitable constants $a_{\alpha,\beta}^{k+1}, \|\beta\| \leq k+1$ for $\|\alpha\| = k+1$

$$\hat{B}_t^{\circ \alpha} = \sum_{\|\beta\| \leq k+1} a_{\alpha,\beta}^{k+1} \hat{B}_t^\beta.$$

Let $\alpha = (\alpha', \alpha^*)$ where $\|\alpha'\| = k-1$ if $\alpha^* = 0$, and $\|\alpha'\| = k$ if $\alpha^* \in \{1, \dots, d\}$. The cases $\alpha^* = 0$ and $\alpha^* \in \{1, \dots, d\}$ are treated separately. Assume first that $\alpha^* = 0$. Then

$$\begin{aligned} \hat{B}_t^{\circ \alpha} &= \int_0^t \hat{B}_s^{\circ \alpha'} ds = \int_0^t \sum_{\|\beta\| \leq k} a_{\alpha',\beta}^k \hat{B}_s^{\circ \beta} ds \\ &= \sum_{\|\beta\| \leq k} a_{\alpha',\beta}^k \hat{B}_t^{\beta*0} \\ &= \sum_{\substack{\|\beta\| \leq k+1 \\ \beta^*=0}} a_{\alpha',\beta}^k \hat{B}_t^{\beta*0} \\ &= \sum_{\|\beta\| \leq k+1} a_{\alpha,\beta}^{k+1} \hat{B}_t^\beta, \end{aligned}$$

$$\text{where } a_{\alpha,\beta}^{k+1} = \begin{cases} a_{\alpha',\beta'}^k & \text{if } \beta^* = 0 \\ 0 & \text{if } \beta^* \neq 0. \end{cases}$$

Now assume $\alpha^* \in \{1, \dots, d\}$:

$$\begin{aligned} \hat{B}_t^{\circ \alpha} &= \int_0^t \hat{B}_s^{\circ \alpha'} \circ dB_s^{\alpha^*} = \int_0^t \hat{B}_s^{\circ \alpha'} dB_s^{\alpha^*} + \frac{1}{2} \int_0^t \hat{B}_s^{\circ \alpha''} ds 1_{\{\alpha^*=(\alpha')^*\}} \\ &= \int_0^t \sum_{\|\beta\| \leq k} a_{\alpha',\beta}^k \hat{B}_s^{\circ \beta} dB_s^{\alpha^*} + \frac{1}{2} \int_0^t \sum_{\|\beta\| \leq k} a_{\alpha'',\beta}^k \hat{B}_s^{\circ \beta} ds 1_{\{\alpha^*=(\alpha')^*\}} \\ &= \sum_{\substack{\|\beta\| \leq k+1 \\ \beta^*=\alpha^*}} a_{\alpha',\beta'}^k \hat{B}_t^{\beta*0} + \frac{1_{\{\alpha^*=(\alpha')^*\}}}{2} \sum_{\substack{\|\beta\| \leq k+1 \\ \beta^*=0}} a_{\alpha'',\beta}^k \hat{B}_t^{\beta*0} \\ &= \sum_{\|\beta\| \leq k+1} a_{\alpha,\beta}^{k+1} \hat{B}_t^{\beta*0}, \end{aligned}$$

$$\text{where } a_{\alpha,\beta}^{k+1} = \begin{cases} a_{\alpha',\beta'}^k & \text{if } \alpha^* = \beta^*, \\ \frac{1}{2}a_{\alpha'',\beta'}^{k-1} & \text{if } \beta^* = 0, \alpha^* = (\alpha')^*, \\ 0 & \text{otherwise.} \end{cases}$$

This completes the argument. As (ii) can be proved in an analogous manner, its proof is omitted. It is now shown how (i), (ii) imply the remaining equivalence result. Note that if A_L is invertible, then A_L^T is also invertible with $(A_L^T)^{-1} = (A_L^{-1})^T$. Moreover, from invertibility

$$0 < c_{\min} := \min_{|\xi|=1} |A_L^T \xi|.$$

Adopting the shorthand notation $\hat{B}_t^{\circ L}, \hat{B}_t^L$ employed above, there holds:

$$\begin{aligned} \inf_{|\xi|=1} \int_0^1 (\xi, \hat{B}_t^{\circ L})^2 dt &= \inf_{|\xi|=1} \int_0^1 (\xi, A_L \hat{B}_t^L)^2 dt \\ &= \inf_{|\xi|=1} \int_0^1 (A_L^T \xi, \hat{B}_t^L)^2 dt \\ &\leq \inf_{|\nu|=1} \int_0^1 (\nu, \hat{B}_t^L)^2 dt \, c_{\min}^2. \end{aligned}$$

A similar estimate can be made from (ii). These estimates prove the remaining claim of the proposition. ■

Before tackling Proposition 2.8 in earnest, some supplementary results about iterated Itô integrals are required.

Lemma 2.13 *Fix $l \in \mathbb{N}$. There exists $C_l < \infty$ and $\nu_l > 0$ such that for all $\alpha \in \mathcal{A}$ with $\|\alpha\| = l$, there holds:*

$$\mathbb{P} \left(\sup_{t \in (0,1]} |\hat{B}_t^\alpha| \geq n \right) \leq C_l \exp \left(-\frac{1}{2} n^{\nu_l} \right), \quad (2.20)$$

for all $n \geq 1$.

Proof (adapted from the proof of Lemma A.7 in Kusuoka and Stroock [24]) : Fundamental use of the following martingale inequality is made. For $K_1, K_2 \geq 0$

$$\mathbb{P} \left(\sup_{t \in (0,T]} |M_t| \geq K_1, \langle M \rangle_T \leq K_2 \right) \leq 2 \exp \left\{ -\frac{K_1^2}{2K_2} \right\}.$$

This result is proved by expressing the above martingale as time-changed Brownian motion (run at the ‘speed’ of its quadratic variation, see Karatzas and Shreve [17, Thm 3.4.6]), and then using the

following two relationships:

$$\mathbb{P}\left(\sup_{t \in (0, T]} |B_t| \geq K\right) = 2\mathbb{P}(B_T \geq K),$$

$$\int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du \leq e^{-x^2/2}, \quad x \geq 0.$$

The latter is seen by splitting consideration into two cases: $x \in [0, 1)$ and $x \geq 1$.

The relation in question can be obtained by iterative applications of this martingale inequality. Define $\nu_N \equiv 2$, and in what follows allow ν_i to be chosen optimally afterwards. First assume that $\alpha \in \{1, \dots, d\}^N$.

$$\begin{aligned} \mathbb{P}\left[\sup_{t \in (0, 1]} |\hat{B}_t^\alpha| \geq K\right] &\leq \mathbb{P}\left[\sup_{t \in (0, 1]} |\hat{B}_t^\alpha| \geq K, \langle \hat{B}^{\alpha'} \rangle_1 < K^{\nu_N}\right] + \mathbb{P}\left[\langle \hat{B}^\alpha \rangle_1 \geq K^{\nu_N}\right] \\ &= \mathbb{P}\left[\sup_{t \in (0, 1]} |\hat{B}_t^\alpha| \geq K, \langle \hat{B}^{\alpha'} \rangle_1 < K^{\nu_N}\right] + \mathbb{P}\left[\int_0^1 |\hat{B}_t^{\alpha'}|^2 dt \geq K^{\nu_N}\right] \\ &\leq \mathbb{P}\left[\sup_{t \in (0, 1]} |\hat{B}_t^\alpha| \geq K, \langle \hat{B}^{\alpha'} \rangle_1 < K^{\nu_N}\right] + \mathbb{P}\left[\sup_{t \in (0, 1]} |\hat{B}_t^{\alpha'}| \geq K^{\nu_N/2}\right] \\ &\stackrel{(*)}{\leq} \sum_{i=1}^N \mathbb{P}\left[\sup_{t \in (0, 1]} |\hat{B}_t^{\alpha^{(N+1-i)}}| \geq K^{\nu_i/2}, \langle \hat{B}^{\alpha^{(i-1)}} \rangle_1 < K^{\nu_{i-1}}\right] \\ &\leq \sum_{i=1}^N 2 \exp\left(-\frac{1}{2} K^{\nu_i - \nu_{i-1}}\right), \end{aligned}$$

where $\alpha^{(i)}$ denotes that shortening of the multi-index $\alpha = (\alpha_1, \dots, \alpha_N)$ by i . i.e. $\alpha^{(i)} = (\alpha_1, \dots, \alpha_{N-i})$ (additionally: $\alpha^{(0)} = \alpha$). Note that the $(*)$ follows from iteratively applying the preceding inequality.

Now choose ν_i for $i = 1, \dots, N$ given that $\nu_N = 2$ and $\nu_0 \geq 0$. In fact, ν_0 can be chosen arbitrarily for $K \geq 1$. If it is assumed that $\nu_i - \nu_{i-1} \equiv \delta > 0$ for $i = 1, \dots, N$, then

$$\sum_{i=1}^N \nu_i - \nu_{i-1} = N\delta \quad \Rightarrow \quad \delta = \frac{2}{N},$$

and $\nu_i = \frac{2i}{N}$. Assembling these facts gives:

$$\mathbb{P}\left(\sup_{t \in (0, 1]} |\hat{B}_t^\alpha| \geq K\right) \leq 2N \exp\left(-\frac{1}{2} K^{\frac{2}{N}}\right), \quad (2.21)$$

for arbitrary $|\alpha| = N$. Assuming instead that $\|\alpha\| = l$ and noting that $|\alpha| \leq \|\alpha\|$ so that $\frac{l}{2} \leq |\alpha| \leq l$ gives the same upper bound with N replaced by l . i.e. $C_l = 2l$ and $\nu_l = 2/l$.

Now observe that if $\alpha_i = 0$ for some $i = 1, \dots, N$ the situation is even simpler:

$$\mathbb{P}\left(\sup_{t \in (0,1]} \left| \hat{B}_t^{(\alpha_1, \dots, \alpha_i)} \right| \geq K\right) \leq \mathbb{P}\left(\sup_{t \in (0,1]} \left| \hat{B}_t^{(\alpha_1, \dots, \alpha_{i-1})} \right| \geq K\right),$$

as $\sup_{t \in (0,T]} \left| \int_0^t \hat{B}_s^\alpha dt \right| \leq T \sup_{t \in (0,T]} \left| \hat{B}_t^\alpha \right|$. Therefore, one needs only apply the martingale inequality $\text{Card} \{i : \alpha_i \neq 0\}$ times. i.e. $(2|\alpha| - \|\alpha\|)$ times. Hence, for a general α ,

$$\mathbb{P}\left(\sup_{t \in (0,1]} \left| \hat{B}_t^\alpha \right| \geq K\right) \leq 2(2|\alpha| - \|\alpha\|) \exp\left(-\frac{1}{2} K^{\frac{2}{2|\alpha| - \|\alpha\|}}\right).$$

However, for any α such that $\|\alpha\| = l$ the identified constants in (2.21) are still appropriate, as $\sup_{\|\alpha\|=l} (2|\alpha| - \|\alpha\|) = l$. ■

The main consequence of the above lemma is the following:

Proposition 2.14 *To prove Proposition 2.11, it suffices to show the existence of C_m, μ_m such that for all $n \geq 1$, there holds*

$$\sup_{a \in S^{N_{m-1}^{0,\emptyset}-1}} \mathbb{P}\left(\int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha \hat{B}_t^\alpha \right]^2 dt \leq \frac{1}{n}\right) \leq C_m \exp\{-n^{\mu_m}\}. \quad (2.22)$$

Adapted from the proof of Lemma 2.3.1 in Nualart [33] : There is some constant M_m such that for all $n \geq 1$, $S^{N_{m-1}^{0,\emptyset}-1}$ contains some finite set $\Sigma(n)$ with

$$|\Sigma(n)| \leq M_m n^{2N_m} \quad \text{and} \quad S^{N_{m-1}^{0,\emptyset}-1} \subset \bigcup_{c \in \Sigma(n)} B_{1/\sqrt{2n}}.$$

Observe, for fixed $a^c \in S^{N_{m-1}^{0,\emptyset}-1} \cap B_{1/\sqrt{2}n}(c)$, there holds

$$\begin{aligned}
& \min_{c \in \Sigma(n)} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} c_\alpha \hat{B}_t^\alpha \right]^2 dt \\
&= \min_{c \in \Sigma(n)} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} (c_\alpha - a_\alpha^c + a_\alpha^c) \hat{B}_t^\alpha \right]^2 dt \\
&\leq 2 \min_{c \in \Sigma(n)} \left(\int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} (c_\alpha - a_\alpha^c) \hat{B}_t^\alpha \right]^2 dt + \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha^c \hat{B}_t^\alpha \right]^2 dt \right) \\
&\leq 2 \min_{c \in \Sigma(n)} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} (c_\alpha - a_\alpha^c) \hat{B}_t^\alpha \right]^2 dt + 2 \min_{c \in \Sigma(n)} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha^c \hat{B}_t^\alpha \right]^2 dt \\
&\leq 2 \min_{c \in \Sigma(n)} |c - a^c|^2 \int_0^1 \sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} |\hat{B}_t^\alpha|^2 dt + 2 \min_{c \in \Sigma(n)} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha^c \hat{B}_t^\alpha \right]^2 dt \\
&\leq 2 \frac{1}{2n^2} \int_0^1 \sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} |\hat{B}_t^\alpha|^2 dt + 2 \min_{c \in \Sigma(n)} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha^c \hat{B}_t^\alpha \right]^2 dt.
\end{aligned}$$

Now, the above upper bound holds for any $a^c \in S^{N_{m-1}^{0,\emptyset}-1} \cap B_{1/\sqrt{2}n}(c)$, in particular, it must hold upon taking the infimum over all $a \in S^{N_{m-1}^{0,\emptyset}-1}$, as $S^{N_{m-1}^{0,\emptyset}-1} = \bigcup_{c \in \Sigma(n)} S^{N_{m-1}^{0,\emptyset}-1} \cap B_{1/2n^2}(c)$. This gives:

$$\begin{aligned}
\min_{c \in \Sigma(n)} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} c_\alpha \hat{B}_t^\alpha \right]^2 dt &\leq 2 \frac{1}{2n^2} \int_0^1 \sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} |\hat{B}_t^\alpha|^2 dt \\
&\quad + 2 \inf_{a \in S^{N_{m-1}^{0,\emptyset}-1}} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha \hat{B}_t^\alpha \right]^2 dt.
\end{aligned} \tag{2.23}$$

Furthermore, it is evident that:

$$\begin{aligned}
& \mathbb{P} \left[\inf_{a \in S^{N_{m-1}^{0,\emptyset}-1}} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha \hat{B}_t^\alpha \right]^2 dt \leq \frac{1}{n} \right] \\
&\leq \mathbb{P} \left(\min_{c \in \Sigma(n)} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha^c \hat{B}_t^\alpha \right]^2 dt \leq \frac{3}{n} \right) \\
&\quad + \mathbb{P} \left(\int_0^1 \sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} |\hat{B}_t^\alpha|^2 dt \geq n \right).
\end{aligned}$$

Using (2.23) to proceed, it is seen that:

$$\begin{aligned}
& \mathbb{P} \left[\inf_{a \in S^{N_{m-1}^{0,\emptyset}-1}} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha \hat{B}_t^\alpha \right]^2 dt \leq \frac{1}{n} \right] \\
& \leq \mathbb{P} \left[\min_{c \in \Sigma(n)} \int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} c_\alpha \hat{B}_t^\alpha \right]^2 dt \leq \frac{3}{n} \right] + \mathbb{P} \left[\int_0^1 \sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} |\hat{B}_t^\alpha|^2 dt \geq n \right] \\
& \leq \sum_{c \in \Sigma(n)} \mathbb{P} \left[\int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} c_\alpha \hat{B}_t^\alpha \right]^2 dt \leq \frac{3}{n} \right] + \mathbb{P} \left[\sup_{t \in (0,1]} \sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} |\hat{B}_t^\alpha|^2 \geq n \right] \\
& \leq M_m K^{2N_{m-1}^{0,\emptyset}} \sup_{\xi \in S^{N_{m-1}^{0,\emptyset}-1}} \mathbb{P} \left[\int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha \hat{B}_t^\alpha \right]^2 dt \leq \frac{3}{n} \right] \\
& \quad + N_{m-1}^{0,\emptyset} \max_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} \mathbb{P} \left[\sup_{t \in (0,1]} |\hat{B}_t^\alpha|^2 \geq \frac{n}{N_{m-1}^{0,\emptyset}} \right] \\
& \leq M_m n^{2N_{m-1}^{0,\emptyset}} B_m \exp \left(- \left[\frac{n}{3} \right]^{\mu_m} \right) + N_{m-1}^{0,\emptyset} \max_{\substack{k=0,\dots,m-1 \\ \|\alpha\|=k}} \mathbb{P} \left[\sup_{t \in (0,1]} |\hat{B}_t^\alpha| \geq \sqrt{\frac{n}{N_{m-1}^{0,\emptyset}}} \right] \\
& \leq M_m n^{2N_{m-1}^{0,\emptyset}} B_m \exp \left(- \left[\frac{n}{3} \right]^{\mu_m} \right) + N_{m-1}^{0,\emptyset} \max_{k=0,\dots,m-1} C_m \exp \left(- \frac{1}{2} \left[\frac{n}{N_{m-1}^{0,\emptyset}} \right]^{\frac{\nu_m}{2}} \right) \\
& \leq A_m \exp \left(-n^{\lambda_m} \right),
\end{aligned}$$

for some (large) constant A_m and (small) $\lambda_m > 0$, for all $n \geq 1$. Both (2.20) and (2.22) have been used. ■

The goal is now reasonably clear. If inequality (2.22) can be proved, then (2.19) will have been justified. Before turning to this proof in earnest, another supporting result is proved. Note that the rest of the proof is, unless otherwise stated, taken from the appendix (p73 and onwards) of Kusuoka and Stroock [24].

Lemma 2.15 *Assume $a \in S^{N_{m-1}^{0,\emptyset}-1}$ such that $|a_\emptyset| < 1$.¹ Then there are constants $Q_m < \infty$ and $\nu_m > 0$ such that:*

$$\mathbb{P} \left(\int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha \hat{B}_t^\alpha \right]^2 dt \leq \frac{1}{n} \right) \leq Q_m \exp \left\{ - \frac{1}{2} \left(\frac{\left[|a_\emptyset| - \frac{1}{\sqrt{n}} \right] \vee 0}{\sqrt{1 - a_\emptyset^2}} \right)^{2\nu_m} \right\}. \quad (2.24)$$

¹Indeed, the consideration is trivial if this condition is violated.

Proof : The starting point is noting that:

$$\begin{aligned}
\left(\int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha \hat{B}_t^\alpha \right]^2 dt \right)^{\frac{1}{2}} &\geq |a_\emptyset| - \left(\int_0^1 \left[\sum_{1 \leq \|\alpha\| \leq m-1} a_\alpha \hat{B}_t^\alpha \right]^2 dt \right)^{\frac{1}{2}} \\
&\geq |a_\emptyset| - \sqrt{1 - a_\emptyset^2} \int_0^1 \left[\sum_{1 \leq \|\alpha\| \leq m-1} |\hat{B}_t^\alpha|^2 dt \right]^{\frac{1}{2}} \\
&\geq |a_\emptyset| - \sqrt{1 - a_\emptyset^2} \sup_{t \in (0,1]} \left[\sum_{1 \leq \|\alpha\| \leq m-1} |\hat{B}_t^\alpha|^2 \right]^{\frac{1}{2}}.
\end{aligned}$$

Consequently,

$$\sup_{t \in (0,1]} \sum_{1 \leq \|\alpha\| \leq m-1} |\hat{B}_t^\alpha|^2 \geq \left(\frac{\left[|a_\emptyset| - \left(\int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha \hat{B}_t^\alpha \right]^2 dt \right)^{\frac{1}{2}} \right] \vee 0}{\sqrt{1 - a_\emptyset^2}} \right)^2.$$

In particular,

$$\begin{aligned}
&\mathbb{P} \left(\int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(m-1)} a_\alpha \hat{B}_t^\alpha \right]^2 dt \leq \frac{1}{n} \right) \\
&\leq \mathbb{P} \left(\sup_{t \in (0,1]} \sum_{1 \leq \|\alpha\| \leq m-1} |\hat{B}_t^\alpha|^2 \geq \left(\frac{\left(|a_\emptyset| - \frac{1}{\sqrt{n}} \right) \vee 0}{\sqrt{1 - a_\emptyset^2}} \right)^2 \right) \\
&\leq Q_m \exp \left\{ -\frac{1}{2} \left(\frac{\left(|a_\emptyset| - \frac{1}{\sqrt{n}} \right) \vee 0}{\sqrt{1 - a_\emptyset^2}} \right)^{2\nu_m} \right\},
\end{aligned}$$

for some Q_m, ν_m , where (2.20) has been used. ■

A semimartingale inequality from Norris [32] is now recalled, which plays an identical role to a similar martingale inequality in Kusuoka and Stroock [24].

Lemma 2.16 Assume $a, y \in \mathbb{R}$. Let $\beta = (\beta)_{t \geq 0}$ be a one-dimensional previsible process, and let $\gamma = (\gamma_t := (\gamma_t^1, \dots, \gamma_t^d))_{t \geq 0}$, $u = (u_t := (u_t^1, \dots, u_t^d))_{t \geq 0}$ be d -dimensional previsible processes. Moreover, assume $B = (B_t)_{t \geq 0}$ is a d -dimensional Brownian motion. Define,

$$\begin{aligned}
b_t &= b + \int_0^t \beta_s ds + \int_0^t \gamma_s^i dB_s^i, \\
Y_t &= y + \int_0^t b_s ds + \int_0^t u_s^i dB_s^i.
\end{aligned}$$

Then for any $q > 8$ and some $\nu < (q - 8)/9$, there is a constant $C = C(q, \nu)$ (independent of K)

such that

$$\mathbb{P} \left[\int_0^1 Y_t^2 dt < \frac{1}{n}, \int_0^1 |b_t|^2 + |u_t|^2 dt \geq \frac{1}{n^{1/q}}, \sup_{t \in (0,1]} |\beta_t| \vee |\gamma_t| \vee |b_t| \vee |u_t| \leq n \right] \leq C \exp\{-n^\nu\}. \quad (2.25)$$

Remark 2.17 Upon checking the above result in Norris [32], the keen reader would observe that the result is stated in a different fashion. Namely, the bound

$$\sup_{t \in (0,T]} |\beta_t| \vee |\gamma_t| \vee |b_t| \vee |u_t| \leq M,$$

is assumed up to some bounded stopping time T , as an extra condition. The resulting statement is then phrased in terms of some constant, which depends on M . i.e. $C = C(q, \nu, M)$ in (2.25). This constraint has been circumvented by letting the constant M depend also on n (indeed: $M = n$). The observation that C is then of the form $C = \hat{C}(q, \nu)n^l$ for some $l \in \mathbb{N}$, is then made. This observation is a result of tracking the constant in the proof of the lemma. This does not affect (2.25) as there is some larger constant $\tilde{C}$ and smaller $\tilde{\nu}$, which can be chosen such that $\hat{C}(q, \nu)n^l \exp\{-n^\nu\} \leq \tilde{C}(q, \nu) \exp\{-n^{\tilde{\nu}}\}$, for all $n \geq 1$.

The proof of the bound in Proposition 2.14 is done via an induction argument. The base case $m - 1 = 0$ is trivial. Assume therefore, that (2.22) holds for $0 \leq m - 1 \leq k - 1$. Let $a \in S^{N_k^{0,\emptyset}-1}$. Define, using the notation of Lemma 2.16, the following:

$$\begin{aligned} Y_t &:= \sum_{\|\alpha\| \leq k} a_\alpha \hat{B}_t^\alpha, \\ b_t &:= \sum_{\substack{1 \leq \|\alpha\| \leq k \\ \alpha^* = 0}} a_\alpha \hat{B}_t^{\alpha'}, \\ u_t^i &:= \sum_{\substack{1 \leq \|\alpha\| \leq k \\ \alpha^* = i}} a_\alpha \hat{B}_t^{\alpha'}, \\ \beta_t &:= \sum_{\substack{1 \leq \|\alpha\| \leq k \\ \alpha^* = 0, (\alpha')^* = 0}} a_\alpha \hat{B}_t^{\alpha''}, \quad \text{for } |\alpha| \geq 2, \\ \gamma_t^i &:= \sum_{\substack{1 \leq \|\alpha\| \leq k \\ \alpha^* = 0, (\alpha')^* = i}} a_\alpha \hat{B}_t^{\alpha''}, \quad \text{for } |\alpha| \geq 2, \\ y &:= a_\emptyset, \\ b &:= 0. \end{aligned}$$

With these definitions it is easy to see

$$\begin{aligned} b_t &= b + \int_0^t \beta_s ds + \int_0^t \gamma_s^i dB_s^i, \\ Y_t &= y + \int_0^t a_s ds + \int_0^t u_s^i dB_s^i. \end{aligned}$$

Using Lemma (2.15) consideration may be split into two separate cases. Assume first that $1 - a_\emptyset^2 \leq n^{-1/2q}$, where $q \geq 1$. So that

$$\sqrt{1 - a_\emptyset^2} \leq n^{-1/4q},$$

and

$$\begin{aligned} |a_\emptyset| &\geq \{(1 - n^{-1/2q}) \vee 0\}^{1/2} \\ \Rightarrow \left[|a_\emptyset| - \frac{1}{\sqrt{n}} \right] \vee 0 &\geq (1 - 2n^{-1/2q}) \vee 0. \end{aligned}$$

Then, by (2.24):

$$\begin{aligned} \mathbb{P}\left(\int_0^1 \left[\sum_{\alpha \in \mathcal{A}_{0,\emptyset}(k)} a_\alpha \hat{B}_t^\alpha \right]^2 dt \leq \frac{1}{n}\right) &\leq Q_k \exp \left\{ -\frac{1}{2} n^{\nu_k/2q} \left(\left[1 - \frac{2}{n^{1/2q}} \right] \vee 0 \right)^{2\nu_k} \right\} \\ &\leq P_k \exp \left\{ -n^{\lambda_k} \right\}, \end{aligned}$$

for some (large) constant P_k and (small) λ_k , as required. Suppose now that $1 - a_\emptyset^2 \geq 1/n^{1/2q}$.

Then it is clear that

$$\left\{ \int_0^1 \left[\sum_{\|\alpha\| \leq k} a_\alpha \hat{B}_t^\alpha \right]^2 dt \leq \frac{1}{n} \right\} \subset E_1 \cup E_2 \cup E_3,$$

where

$$\begin{aligned} E_1 &= \left\{ \int_0^1 Y_t^2 \leq \frac{1}{n}, \int_0^1 |b_t|^2 + |u_t|^2 dt \geq \frac{1}{n^{1/q}}, \sup_{t \in (0,1]} |\beta_t| \vee |\gamma_t| \vee |b_t| \vee |u_t| \leq n \right\}, \\ E_2 &= \left\{ \sup_{t \in (0,1]} |\beta_t| \vee |\gamma_t| \vee |b_t| \vee |u_t| > n \right\}, \\ E_3 &= \left\{ \int_0^1 |b_t|^2 + |u_t|^2 dt < \frac{1}{n^{1/q}} \right\}. \end{aligned}$$

It is now shown that $\mathbb{P}(E_i) \leq C_i \exp\{-n^{\nu_i}\}$ for $i = 1, 2, 3$. For $i = 1, 2$, Lemma 2.16 and Lemma 2.13 imply respectively, the required bounds (i.e. independent of $a \in S^{N_{m-1}^{0,\emptyset}-1}$). The case $i = 3$ is handled using the inductive hypothesis.

Define

$$N_j := \sum_{\substack{1 \leq \|\alpha\| \leq k - \|(j)\| \\ \alpha^* = j}} a_\alpha^2$$

As $\sum_{j=0}^d N_j = 1 - a_\emptyset^2 \geq 1/n^{1/2q}$, there exists $j_0 \in \{0, \dots, d\}$ such that $N_{j_0} \geq 1/(d+1)n^{1/2q}$. Moreover, $|b_t|^2 + |u_t|^2 \geq \left| \sum_{\substack{1 \leq \|\alpha\| \leq k - \|(j_0)\| \\ \alpha^* = j_0}} a_\alpha \hat{B}_t^{\alpha'} \right|^2$. Thus, using the inductive hypothesis,

$$\begin{aligned} \mathbb{P}(E_3) &\leq \mathbb{P}\left(\int_0^1 \left| \sum_{\substack{1 \leq \|\alpha\| \leq k - \|(k_0)\| \\ \alpha^* = j_0}} a_\alpha \hat{B}_t^{\alpha'} \right|^2 dt \leq \frac{1}{n^{1/q}}\right) \\ &= \mathbb{P}\left(\frac{1}{N_{j_0}} \int_0^1 \left| \sum_{\substack{1 \leq \|\alpha\| \leq k - \|(k_0)\| \\ \alpha^* = j_0}} a_\alpha \hat{B}_t^{\alpha'} \right|^2 dt \leq \frac{1}{N_{j_0} n^{1/q}}\right) \\ &\leq C_{k-1} \exp\{-(N_{j_0} n^{1/q})^{\nu_{k-1}}\} \\ &\leq C_{k-1} \exp\{-(n^{1/2q}/(d+1))^{\nu_{k-1}}\} \\ &\leq C_k \exp\{-n^{\nu_k}\}, \end{aligned}$$

for some C_k, ν_k . In applying the inductive hypothesis, care has been taken to check that $\left(\sum_{\substack{1 \leq \|\alpha\| \leq k - \|(k_0)\| \\ \alpha^* = k_0}} a_\alpha^2\right)/N_{k_0} = 1$. This finishes the proof. ■

2.5. Precise gradient bounds

KUSUOKA-STROOCK PROCESSES

In this section an important space of functions are introduced. These functions are tailor-made for the generation of an integration by parts formula with asymptotic rates. It is claimed that one is able to identify Φ_α , for $\alpha \in \mathcal{A}(m)$, such that:

$$\mathbb{E}[\Phi(t, x) V_{[\alpha]}(f \circ X_t)(x)] = t^{-\|\alpha\|/2} \mathbb{E}[\Phi_\alpha(t, x) f(X_t^x)].$$

The reason the following space of functions is important, is because it describes the common properties of Φ and Φ_α , and is closed under the operations which are taken during the formation of the IBPF. As a result this space supports iterative applications of the formula.

Definition 2.18 (Kusuoka-Stroock Processes) Let E be a separable Hilbert space and let $r \in \mathbb{R}$. Denote by $\mathcal{K}_r(E)$ the set of functions: $f : (0, 1] \times \mathbb{R}^N \rightarrow \mathbb{D}^\infty(E)$ satisfying the following:

1. $f(t, \cdot)$ is smooth and $\frac{\partial^\alpha f}{\partial x^\alpha}(\cdot, \cdot)$ is continuous in $(t, x) \in (0, 1] \times \mathbb{R}^N$ a.s. for any multi-index α .
2. $\sup_{t \in (0, 1], x \in \mathbb{R}^N} t^{-r/2} \left\| \frac{\partial^\alpha f}{\partial x^\alpha} \right\|_{k, p, E} < \infty$, for all $k \in \mathbb{N}$ and $p \in [1, \infty)$.

Define $\mathcal{K}_r := \mathcal{K}_r(\mathbb{R})$

The process are named after the authors of [24] in which they first appeared. The following properties shall prove extremely useful for computing the integration by parts formula.

Lemma 2.19 (Properties of Kusuoka-Stroock Smooth Processes)

The following hold,

1. Suppose $f \in \mathcal{K}_r(E)$, where $r \geq 0$. Then, for $i = 1, \dots, d$,

$$\int_0^\cdot f(s, x) dB_s^i \in \mathcal{K}_{r+1}(E) \quad \text{and} \quad \int_0^\cdot f(s, x) ds \in \mathcal{K}_{r+2}(E).$$

2. $a_{\alpha, \beta}, b_{\alpha, \beta} \in \mathcal{K}_{(\|\beta\| - \|\alpha\|) \vee 0}$ where $\alpha, \beta \in \mathcal{A}(m)$, where $a_{\alpha, \beta}, b_{\alpha, \beta}$ are defined as in Proposition 2.3 and Corollary 2.4, resp.
3. $k_\alpha \in \mathcal{K}_{\|\alpha\|}(H)$, where $\alpha \in \mathcal{A}(m)$, where k_α is defined on p44.
4. $D^{(\alpha)}u := \langle Du(t, x), k_\alpha \rangle_H \in \mathcal{K}_{r+\|\alpha\|}$ where $u \in \mathcal{K}_r$ and $\alpha \in \mathcal{A}(m)$.
5. If M^{-1} is the inverse matrix of M , defined on p44, then $M_{\alpha, \beta}^{-1} \in \mathcal{K}_0$, where $\alpha, \beta \in \mathcal{A}(m)$.
6. If $f_i \in \mathcal{K}_{r_i}$ for $i = 1, \dots, N$, then

$$\prod_{i=1}^N f_i \in \mathcal{K}_{r_1 + \dots + r_N} \quad \text{and} \quad \sum_{i=1}^N f_i \in \mathcal{K}_{\min(r_1, \dots, r_N)}.$$

Proof : See appendix. ■

The integration by parts formulae will be phrased in terms of Kusuoka-Stroock processes.

SEVERAL INTEGRATION BY PARTS FORMULAE

The following theorem can finally be proved

Theorem 2.20 (Integration by Parts formula I) Assuming $V_0, \dots, V_d \in C_b^\infty$ are uniformly bounded, and the the UFG condition holds for some $m \in \mathbb{N}$. Then, for all $\Phi \in \mathcal{K}_r$, $r \in \mathbb{R}$ and $\alpha \in \mathcal{A}$ there exists $\Phi_\alpha \in \mathcal{K}_r$ such that

$$\mathbb{E}[\Phi(t, x) V_{[\alpha]}(f \circ X_t^{(\cdot)})(x)] = t^{-\|\alpha\|/2} \mathbb{E}[\Phi_\alpha(t, x) f(X_t^x)], \quad (2.26)$$

for any $f \in C_b^\infty(\mathbb{R}^N; \mathbb{R})$, $t > 0$, $x \in \mathbb{R}^N$.

From this it is possible to prove other related integration by parts formulae. In particular (cf. Corollary 2.22), one should take note of the non-trivial role the UFG condition plays in these other derivations - testament to its strength.

Corollary 2.21 (Integration by Parts formula II) Under the assumptions of Theorem 2.20 the following holds:

$$\mathbb{E}[\Phi(t, x) (V_{[\alpha]} f)(X_t^x)] = t^{-\|\alpha\|/2} \mathbb{E}[\Phi'_\alpha(t, x) f(X_t^x)],$$

where $\Phi'_\alpha \in \mathcal{K}_r$, for any $f \in C_b^\infty(\mathbb{R}^N; \mathbb{R})$, $t > 0$, $x \in \mathbb{R}^N$.

Corollary 2.22 (Integration by Parts formula III) *Under the assumptions of Theorem 2.20 the following holds:*

$$V_{[\alpha]} \mathbb{E}[\Phi(t, x) f(X_t^x)] = t^{-\|\alpha\|/2} \mathbb{E}[\Phi_\alpha''(t, x) f(X_t^x)],$$

where $\Phi_\alpha \in \mathcal{K}_r$, for any $f \in C_b^\infty(\mathbb{R}^N; \mathbb{R})$, $t > 0$, $x \in \mathbb{R}^N$.

For $\Phi \in \mathcal{K}_r$, $r \in \mathbb{R}$, define the family of operators $P^\Phi = \{P_t^\Phi\}_{t \geq 0} \subset \{L : L : C_b^\infty \rightarrow C_b^\infty, L \text{ linear}\}$ defined by:

$$(P_t^\Phi f)(x) := \mathbb{E}[\Phi(t, x) f(X_t^x)].$$

Corollary 2.23 (Integration by Parts formula IV) *Under the assumptions of Theorem 2.20 the following holds:*

$$V_{[\alpha_1]} \cdots V_{[\alpha_n]} P_t^\Phi V_{[\alpha_{n+1}]} \cdots V_{[\alpha_{n+m}]} f(x) = t^{-\frac{\|\alpha_1\| + \cdots + \|\alpha_{n+m}\|}{2}} \mathbb{E}[\Phi_\alpha(t, x) f(X_t^x)],$$

where $\Phi_\alpha \in \mathcal{K}_r$, for any $f \in C_b^\infty(\mathbb{R}^N; \mathbb{R})$, $t > 0$, $x \in \mathbb{R}^N$.

Proof of Theorem 2.20 : It was demonstrated in the previous section that:

$$V_{[\alpha]}(f \circ X_t)(x) = t^{-\|\alpha\|/2} \sum_{\beta \in \mathcal{A}(m)} t^{-\|\beta\|/2} M_{\alpha, \beta}^{-1}(t, x) D^{(\beta)}(f(X_t^x))$$

holds $\mathbb{P}$ -a.s. The product rule for the Malliavin derivative can be used to deduce the integration by parts formula:

$$\begin{aligned} D^{(\beta)}(\Phi(t, x) M_{\alpha, \beta}^{-1}(t, x) f(X_t^x)) &= D^{(\beta)}\Phi(t, x) M_{\alpha, \beta}^{-1}(t, x) f(X_t^x) \\ &\quad + \Phi(t, x) D^{(\beta)}M_{\alpha, \beta}^{-1}(t, x) f(X_t^x) \\ &\quad + \Phi(t, x) M_{\alpha, \beta}^{-1}(t, x) D^{(\beta)}f(X_t^x). \end{aligned}$$

Then, it is clear that

$$\begin{aligned} \mathbb{E}[\Phi(t, x) V_{[\alpha]}(f \circ X_t)(x)] &= t^{-\frac{\|\alpha\|}{2}} \sum_{\beta \in \mathcal{A}(m)} t^{-\frac{\|\beta\|}{2}} \mathbb{E}[\Phi(t, x) M_{\alpha, \beta}^{-1}(t, x) D^{(\beta)}(f(X_t^x))] \\ &= t^{-\frac{\|\beta\|}{2}} \sum_{\beta \in \mathcal{A}(m)} t^{-\frac{\|\beta\|}{2}} \mathbb{E}\left[f(X_t^x) \{\Phi(t, x) M_{\alpha, \beta}^{-1}(t, x) \delta(k_\beta(t, x)) \right. \\ &\quad \left. - \Phi(t, x) D^{(\beta)}M_{\alpha, \beta}^{-1}(t, x) - D^{(\beta)}\Phi(t, x) M_{\alpha, \beta}^{-1}(t, x)\}\right]. \end{aligned}$$

And so

$$\begin{aligned}\Phi_\alpha(t, x) = \sum_{\beta \in \mathcal{A}(m)} t^{-\|\beta\|/2} \Big\{ & \Phi(t, x) M_{\alpha, \beta}^{-1}(t, x) \delta(k_\beta(t, x)) \\ & - \Phi(t, x) D^{(\beta)} M_{\alpha, \beta}^{-1}(t, x) - D^{(\beta)} \Phi(t, x) M_{\alpha, \beta}^{-1}(t, x) \Big\}.\end{aligned}$$

The claim $\Phi_\alpha \in \mathcal{K}_r$ follows from a diligent application of Lemma 2.19. Namely, parts 3, 4, 5, 6.

Remark 2.24 *As the process $k_\gamma(t, x)(\cdot)$ is $\mathcal{F}_s$ -adapted for a.a. $(t, x) \in (0, 1] \times \mathbb{R}^N$, the adjoint, $\delta(k_\gamma(t, x))$, is nothing more than the Itô integral of $k_\gamma(t, x)$ with respect to the d -dimensional Brownian motion $B_t = (B_t^1, \dots, B_t^d)$. i.e.*

$$\delta(k_\gamma(t, x)) = \sum_{i=1}^d \int_0^1 k_\gamma(t, x)(s) dB_s^i.$$

Thus, it follows that for processes $f \in K^r(H)$ which are a.e. adapted as stochastic processes in H , that $\delta(f) := \delta(f(\cdot, \cdot)) \in \mathcal{K}_{r+1}$.

Proof of Corollary 2.21: The first observation is the following relationship:

$$\begin{aligned}(V_{[\alpha]}f)(X_t^x) &= \nabla f(X_t^x) V_{[\alpha]}(X_t^x) \\ &= (J_t^x)^{-T} \nabla(f \circ X_t)(x) V_{[\alpha]}(X_t^x) \\ &= \nabla(f \circ X_t)(x) (J_t^x)^{-1} V_{[\alpha]}(X_t^x),\end{aligned}$$

where $(J_t^x)^{-T} := ((J_t^x)^{-1})^T$. At this point refer back to the closed linear system of equations, which induced the expression:

$$(J_t^x)^{-1} V_{[\alpha]}(X_t^x) = \sum_{\beta \in \mathcal{A}(m)} a_{\alpha, \beta}(t, x) V_{[\beta]}(x).$$

Again, the central position of the UFG condition is emphasised.

$$\begin{aligned}\nabla(f \circ X_t)(x) (J_t^x)^{-1} V_{[\alpha]}(X_t^x) &= \sum_{\beta \in \mathcal{A}(m)} a_{\alpha, \beta}(t, x) \nabla(f \circ X_t)(x) V_{[\beta]}(x) \\ &= \sum_{\beta \in \mathcal{A}(m)} a_{\alpha, \beta}(t, x) V_{[\beta]}(f \circ X_t)(x).\end{aligned}$$

From Lemma 2.19, $a_{\alpha, \beta} \in \mathcal{K}_{(\|\beta\| - \|\alpha\|) \vee 0}$. Hence, it has been shown that:

$$\mathbb{E} [\Phi(t, x) V_{[\alpha]} f(X_t^x)] = \sum_{\beta \in \mathcal{A}(m)} \mathbb{E} [\Phi(t, x) a_{\alpha, \beta}(t, x) V_{[\beta]}(f \circ X_t)(x)].$$

The integration by parts formula (2.26) can then be applied N_m times, after noting that the product $\Phi a_{\alpha,\beta} \in \mathcal{K}_{r+[(\|\beta\|-\|\alpha\|)\vee 0]}$. And so,

$$\begin{aligned}\mathbb{E} [\Phi(t, x) V_{[\alpha]} f(X_t^x)] &= \sum_{\beta \in \mathcal{A}(m)} t^{-\frac{\|\beta\|}{2}} \mathbb{E} [\Psi_\beta(t, x) f(X_t^x)] \\ &= \sum_{\beta \in \mathcal{A}(m)} t^{-\frac{\|\beta\|}{2}} t^{-\frac{\|\alpha\|-\|\beta\|}{2}} \mathbb{E} \left[t^{\frac{\|\alpha\|-\|\beta\|}{2}} \Psi_\beta(t, x) f(X_t^x) \right] \\ &= t^{-\frac{\|\alpha\|}{2}} \mathbb{E} [\Phi_\alpha(t, x) f(X_t^x)],\end{aligned}$$

where $\Phi_\alpha = \sum_{\beta \in \mathcal{A}(m)} t^{\frac{\|\alpha\|-\|\beta\|}{2}} \Psi_\beta \in \mathcal{K}_r$

Proof of Corollary 2.22:

Observe that

$$\begin{aligned}V_{[\alpha]} \mathbb{E} [\Phi(t, x) f(X_t^x)] &= \mathbb{E} [V_{[\alpha]}(\Phi(t, x)) f(X_t^x) + \Phi(t, x) V_{[\alpha]}(f \circ X_t)(x)] \\ &= \mathbb{E} [V_{[\alpha]}(\Phi(t, x)) f(X_t^x) + t^{-\|\alpha\|/2} \Phi_\alpha(t, x) f(X_t^x)] \\ &= t^{-\|\alpha\|/2} \mathbb{E} [\Phi_\alpha''(t, x) f(X_t^x)],\end{aligned}$$

where $\Phi_\alpha(t, x) = t^{\|\alpha\|/2} V_{[\alpha]}(\Phi(t, x)) + \Phi_\alpha(t, x) \in \mathcal{K}_r$.

Proof of Corollary 2.23:

Once it is noted that $\text{Id} \in \mathcal{K}_0$, the proof follows from n applications of Theorem 2.20 followed by m applications of Corollary 2.21. ■

APPLICATIONS TO GRADIENT BOUNDS

It is now discussed how the integration by parts formulae allow the acquisition of explicit gradient bounds.

Lemma 2.25 *Assume that $\Phi \in \mathcal{K}_r, r \in \mathbb{R}$. Then there is a constant $C < \infty$ such that:*

$$\|\mathbb{E} [\Phi(t, \cdot) f(X_t^x)]\|_\infty \leq C t^{r/2} \|f\|_\infty,$$

for all $t \in (0, 1]$ and $f \in C_b^\infty(\mathbb{R}^N)$.

Proof : By Hölder's inequality, for each $p, q \in [1, \infty]$, such that $p^{-1} + q^{-1} = 1$, there holds

$$|\mathbb{E} [\Phi(t, x) f(X_t^x)]| \leq \|\Phi(t, x)\|_{L^p(\mathbb{P})} \|f(X_t^x)\|_{L^q(\mathbb{P})} \leq \|\Phi(t, x)\|_{L^p(\mathbb{P})} \|f\|_\infty.$$

Moreover, from Definition 2.18,

$$\|\Phi(t, x)\|_{L^p(\mathbb{P})} < C t^{r/2},$$

uniformly, for $(t, x) \in (0, 1] \times \mathbb{R}^N$. This proves the result. ■

Observe the following important corollary of the integration by parts theorems:

Corollary 2.26 Assume that $\Phi \in \mathcal{K}_r, r \in \mathbb{R}$. Let $\alpha_1, \dots, \alpha_{n+m} \in \mathcal{A}(m)$. Then there is a constant $C < \infty$ such that:

$$\|V_{[\alpha_1]} \dots V_{[\alpha_n]} P_t^\Phi V_{[\alpha_{n+1}]} \dots V_{[\alpha_{n+m}]} f\| \leq t^{-\frac{\|\alpha_1 * \dots * \alpha_{n+m}\|}{2}} \|f\|_\infty. \quad (2.27)$$

One may wish to consider functions which are not uniformly bounded. In particular, Lipschitz functions may also be considered.

Corollary 2.27 There is a constant $C < \infty$ such that for $\alpha_1, \dots, \alpha_n \in \mathcal{A}(m)$:

$$\|V_{[\alpha_1]} \dots V_{[\alpha_n]} P_t f\|_\infty \leq \frac{C t^{1/2}}{t^{(\|\alpha_1\| + \dots + \|\alpha_n\|)/2}} L_f. \quad (2.28)$$

where L_f is the Lipschitz constant for the function f .

Proof : The idea behind this gradient bound is that one can ‘sacrifice’ the derivative along $V_{[\alpha_n]}$ to obtain a new integration by parts formula involving the gradient of f . Let $\{f_n\}_n \subset C_b^\infty$ be a sequence of smooth functions such that $f_n \rightarrow f$, uniformly. Observe,

$$\begin{aligned} V_{[\alpha_n]} P_t f_n &= \sum_{i=1}^N V_{[\alpha_n]}^i(x) \partial_i \mathbb{E}[(f_n \circ X_t)(x)] \\ &= \sum_{j=1}^N \mathbb{E} \left[\partial_j f_n(X_t^x) \sum_{i=1}^N V_{[\alpha_n]}^i(x) (J_t^x)_{j,i} \right] \\ &= \sum_{j=1}^N \mathbb{E} [\partial_j f_n(X_t^x) \Phi^j(t, x)], \end{aligned}$$

where $\Phi^j(t, x) = \sum_{i=1}^N V_{[\alpha_n]}^i(x) (J_t^x)_{j,i} \in \mathcal{K}_0$. Hence, following $n - 1$ applications of Theorem 2.20 to the above expression, we see that:

$$V_{[\alpha_1]} \dots V_{[\alpha_n]} P_t f_n(x) = t^{-(\|\alpha_1\| + \dots + \|\alpha_{n-1}\|)/2} \sum_{j=1}^N \mathbb{E} [\partial_j f_n(X_t^x) \Phi_{\alpha_1, \dots, \alpha_{n-1}}^j(t, x)].$$

And therefore

$$\begin{aligned} \|V_{[\alpha_1]} \dots V_{[\alpha_n]} P_t f_n\|_\infty &\leq C t^{-(\|\alpha_1\| + \dots + \|\alpha_{n-1}\|)/2} \|\nabla f_n\|_\infty \\ &\leq \frac{C t^{1/2}}{t^{(\|\alpha_1\| + \dots + \|\alpha_n\|)/2}} \|\nabla f_n\|_\infty. \end{aligned}$$

The last inequality follows because $t^{(1 - \|\alpha_n\|)/2} \geq 1$ for all $\alpha_n \in \mathcal{A}$. The argument is completed for general Lipschitz f by letting $n \rightarrow \infty$. ■

The gradient bounds which have been proved enable simple proofs of differentiability along vector fields which can be expressed as linear combinations of those in the Lie algebra. Indeed, this includes - through the UFG condition - any elements of the Lie algebra.

Corollary 2.28 *Let $W \in C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$, uniformly bounded. Assume that $W = \sum_{i=1}^M \varphi_{\alpha_i} V_{[\alpha_i]}$ for $\alpha_1, \dots, \alpha_M \in \mathcal{A}(m)$, $\varphi_{\alpha_i} \in C_b^\infty$ are also uniformly bounded. Then,*

$$\|WP_t^\Phi f\|_\infty \leq Ct^{-r/2} \|f\|_\infty,$$

where $r = \min\{\|\alpha_1\|, \dots, \|\alpha_M\|\}$.

Proof : We see that

$$\begin{aligned} \|WP_t^\Phi f\|_\infty &= \left\| \sum_{i=1}^M \varphi_{\alpha_i}(\cdot) V_{[\alpha_i]} P_t^\Phi f \right\|_\infty \\ &\leq \sum_{i=1}^M \tilde{C}_{\alpha_i} \|V_{[\alpha_i]} P_t^\Phi f\|_\infty \leq Ct^{-r/2} \|f\|_\infty. \end{aligned}$$

■

Consideration is now turned to a semigroup with a potential term. That, is define $(P_t^c)_{t \geq 0}$ by

$$(P_t^c f)(x) = \mathbb{E} \left(\exp \left\{ \int_0^t c(X_s^x) ds \right\} f(X_t^x) \right).$$

This is the solution of a parabolic PDE which has been perturbed by a smooth function, c . Then one can deduce the following:

Theorem 2.29 *Assume $\alpha_1, \dots, \alpha_{n+m} \in \mathcal{A}(m)$. Then there is a constant C_p such that:*

$$\|V_{[\alpha_1]} \dots V_{[\alpha_n]} P_t^c V_{[\alpha_{n+1}]} \dots V_{[\alpha_{n+m}]} f\|_{L^p(dx)} \leq C_p t^{-(\|\alpha_1\| + \dots + \|\alpha_{n+m}\|)/2} \|f\|_{L^p(dx)},$$

for any $f \in C_b^\infty \cap L^p$, $t \in (0, 1]$ and $p \in [1, \infty]$.

Proof : It is easy to see that P_t^c satisfies an integration by parts formula of the form (2.23), having noted that $\Phi(t, x) := \exp \left\{ \int_0^t c(X_s^x) ds \right\}$, satisfies $\Phi \in \mathcal{K}_0$. The rest of the proof is based on an interpolation argument. The cases $p = \infty$ and $p = 1$ are considered separately and then the *Riesz-Thorin interpolation theorem*, see for example Triebel [40], is used applied to obtain $p \in (1, \infty)$. The case $p = \infty$ follows trivially from applying Corollary 2.23 and Lemma 2.25. For the proof of the case $p = 1$, and for the rest of the proof, please see Kusuoka [22, p268-270]. ■

REGULARITY OF THE SEMIGROUP

In this section we apply the gradient bounds of the previous section to deduce regularity results about the semigroup.

We recall a basic result about uniform approximation with smooth functions:

Lemma 2.30 (Approximating uniformly cts with smooth functions) *Assume that $f \in C(\mathbb{R}^N)$ is uniformly continuous and vanishing at infinity. Then there exists $\{f_n\} \subset C_0^\infty(\mathbb{R}^N)$ such that f_n converges uniformly to f on $\mathbb{R}^N$.*

Proof : See, for example, Jost [16]

■

The next Proposition summaries some regularity results which may be deduced from the gradient bounds.

Proposition 2.31 (Regularity of the semigroup)

1. Assume that $f \in L^p$, where $p \in [1, \infty)$. Then for any $\alpha_1, \dots, \alpha_{n+m} \in \mathcal{A}(m)$, it holds that $V_{[\alpha_1]} \dots V_{[\alpha_n]} P_t^c V_{[\alpha_{k+1}]} \dots V_{[\alpha_{n+m}]} f$ is an L^p function.
2. Now assume that f is uniformly continuous. Then for any $\alpha_1, \dots, \alpha_{n+m} \in \mathcal{A}(m)$, it holds that $V_{[\alpha_1]} \dots V_{[\alpha_n]} P_t^c V_{[\alpha_{k+1}]} \dots V_{[\alpha_{n+m}]} f$ exists and is a smooth function.

Proof :

1. Theorem 2.29 holds for $f \in C_b^\infty$. From Jost [16, Corollary 19.24], we know that $C_b^\infty \cap L^p$ is dense in L^p for any $p \in [1, \infty)$. We choose a sequence $\{f_n\} \subset C_b^\infty$ such that $\|f_n - f\|_{L^p} \rightarrow 0$. It follows from the first gradient bound of Theorem 2.29 that the sequence $V_{[\alpha_1]} \dots V_{[\alpha_n]} P_t^c V_{[\alpha_{k+1}]} \dots V_{[\alpha_{n+m}]} f_n$ is uniformly Cauchy in L^p . As L^p is complete, the sequence converges to an element of L^p .

2. We apply the same argument as in 1 but instead choose a sequence of smooth functions with compact support with converge to ∇f in L^p . We then use the second gradient bound of Theorem 2.29 to complete the argument.

3. We observe from Lemma 2.30 that C_0^∞ is dense in the space of uniformly continuous functions. We then follow the same arguments as in part 1 to deduce smoothness. We have noted that the uniform limit of smooth functions is continuous.

■

2.6. Hörmander's theorem and hypoellipticity

In this section we introduce the notion of a hypoelliptic operator and discuss their almost complete treatment through Hörmander's theorem. As the diffusion semigroup is the solution of a parabolic

PDE, this theorem has important applications in semigroup regularity. By assuming Hörmander's condition for hypoellipticity, it can be easily proved that the semigroup is a smooth function, under certain criterion. In the first part of this section we prove this result without resorting to Hörmander's theorem. Of course, parabolic hypoellipticity of an operator is a stronger property than smoothness of the solution to a parabolic PDE with the corresponding operator. Due to a result in Kusuoka and Stroock [24], which gives probabilistic criterion for hypoellipticity of a parabolic sum of squares operator, we are able to deduce hypoellipticity using solely probabilistic techniques; again under Hörmander's condition. Hence, we recover Hörmander's theorem using solely probabilistic techniques.

RECOVERING SMOOTHNESS AND EXISTENCE OF A DENSITY

A closely related result about existence and smoothness of densities for the law of a diffusion can also be addressed using similar conditions. To show the latter, one needs only show the following proposition holds true:

Proposition 2.32 *Assume that for every multi-index α , and every $f \in C_0^\infty$ there is a constant K_α , such that:*

$$|\mathbb{E}[(\partial_\alpha f)(X_t^x)]| \leq K_\alpha \|f\|_\infty. \quad (2.29)$$

Then the law of X_t^x has a density, which is smooth on $\mathbb{R}^N$

Proof : This criteria was provided by Malliavin in [30]. ■

Gradient bounds such as (2.29) may be deduced from the techniques of Kusuoka, provided some extra assumptions are made.

Theorem 2.33 *Assume that the following holds for all $x \in \mathbb{R}^N$:*

$$\text{Span}\{V_{[\alpha]}(x) : \alpha \in \mathcal{A}(m)\} = \mathbb{R}^N. \quad (2.30)$$

Then (2.29) holds. Moreover, the diffusion semigroup maps uniformly continuous functions to smooth functions.

Note that we may restate (2.30), as $\exists \epsilon > 0$ s.t.:

$$\sum_{\alpha \in \mathcal{A}(m)} (V_{[\alpha]}(x), \xi)^2 \geq \epsilon |\xi|^2, \quad (2.31)$$

$\forall \xi \in \mathbb{R}^N$, or equivalently: the matrix $(VV^T)(x)$ is invertible $\forall x \in \mathbb{R}^N$, where $V(x) := (V_\alpha^j)_{j=1, \dots, N, \alpha \in \mathcal{A}(m)}$.

Note: upon taking the infimum over all $|\xi| = 1$, the LHS of (2.31) is the minimum eigenvalue of this matrix. The inverse must have C_b^∞ , bounded entries by the inverse function theorem.

Proof : Showing (2.29), amounts to deriving an integration by parts formula for the partial

derivatives ∂_i . This can easily be iterated to obtain any combination of partial derivatives. We claim that there exist uniformly bounded functions $C_\alpha^i \in C_b^\infty$ such that:

$$e_i = \sum_{\alpha \in \mathcal{A}(m)} C_\alpha^i(x) V_{[\alpha]}(x),$$

$\forall x \in \mathbb{R}$. This can be re-written in matrix form as

$$e_i = VC^i$$

where $V(x) := (V_{[\alpha]}^j(x))_{\substack{j=1,\dots,N \\ \alpha \in \mathcal{A}(m)}}$, and $C^i(x) = (C_\alpha^i(x))_{\alpha \in \mathcal{A}(m)}$. But it holds that $(VV^T)(x)$ is invertible $\forall x \in \mathbb{R}^N$. Therefore, we may choose

$$C^i = V^T(VV^T)^{-1}e_i,$$

that is, $C_\alpha^i(x) = (V^T(VV^T)^{-1}e_i)_\alpha(x)$. Clearly, $C_\alpha^i \in C_b^\infty$, bounded, by the inverse function theorem. Now observe that

$$\begin{aligned} |\mathbb{E}((\partial_i \varphi)(X_t^x))| &= \left| \sum_{\alpha \in \mathcal{A}(m)} \mathbb{E}[C_\alpha^i(X_t^x)(V_{[\alpha]} \varphi)(X_t^x)] \right| \\ &= \left| \sum_{\alpha \in \mathcal{A}(m)} \mathbb{E}[\Phi_\alpha^i(t, x) \varphi(X_t^x)] \right| \\ &\leq K_i \|\varphi\|_\infty. \end{aligned}$$

The second equality follows from Corollary 2.21. This may be done for any partial derivative and the procedure can be iterated for any multi-index α . To show the second part, we again note that there exist $C_\alpha^i \in C_b^\infty$ such that:

$$e_i = \sum_{\alpha \in \mathcal{A}(m)} C_\alpha^i V_{[\alpha]},$$

on $\mathbb{R}^N$. That is,

$$\partial_i(f \circ X_t)(x) = \sum_{\alpha \in \mathcal{A}(m)} c_\alpha^i(x) V_{[\alpha]}(f \circ X_t)(x).$$

One may then apply Corollaries 2.26 and 2.27 to deduce the gradient bounds:

$$\begin{aligned} \|\partial_{i_1} \dots \partial_{i_M} P_t f\|_\infty &\leq C t^{-Mr/2} \|f\|_\infty \\ \|\partial_{i_1} \dots \partial_{i_M} P_t f\|_\infty &\leq C \frac{t^{1/2}}{t^{Mr/2}} \|\nabla f\|_\infty, \end{aligned}$$

for any $i_1, \dots, i_N \in \{1, \dots, N\}$, $M \in \mathbb{N}$. This is sufficient to imply the required smoothness result, by Proposition 2.31. ■

HYPOELLIPTICITY OF PARABOLIC ‘SUM OF SQUARES’ OPERATORS

In this section we recall what it is for an operator to be hypoelliptic, and review a criterion from Kusuoka and Stroock [24] for parabolic hypoellipticity of an operator. We finally prove using the results of this chapter that this criterion is satisfied. Note that the review of distributions, hypoellipticity and the corresponding examples are taken from Williams’ excellent review of the area in [42].

Define the second order operator, $\mathcal{L}$, in sum of squares form as follows:

$$\mathcal{L} := \sum_{i=1}^d V_i^2 + V_0 \quad (2.32)$$

where $V_0, \dots, V_d$ are vector fields, which are viewed in this instance as differential operators. The main purpose of crystallising a notion of ‘hypoellipticity’ is to study certain regularity properties of solutions to (parabolic) partial differential equations. That is, to consider the solution, u , to the initial value problem:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \mathcal{L}u + cu + g, & \text{on } (0, \infty) \times U \\ u &= f, & \text{on } \{0\} \times U \end{aligned}$$

where $U \subset \mathbb{R}^N$ is open, $g, c \in C_b^\infty(\mathbb{R}^N)$ and $f \in C^0(\mathbb{R}^N)$ is uniformly continuous. The notion of hypoellipticity is also related to that of a ‘distribution’. These notion of a distribution is related to that of a ‘test functions’. Denote by $\mathcal{D}(U)$, the space of test functions on $U \subset \mathbb{R}^N$, open, with the following two properties:

1. Elements of $\mathcal{D}(U)$ are smooth with a compact support contained in U .
2. $\mathcal{D}(U)$ is equipped with a notion of sequential convergence (and corresponding topology τ). Namely, $\{\phi_n\}_n \subset \mathcal{D}(U)$ converges to 0 if there is a common compact subset of U , within which all supports of ϕ_n are contained. Moreover, for any multi-index α

$$\sup_{x \in U} |\partial_\alpha \phi_n(x)| \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

A distribution, Λ , on $U \subset \mathbb{R}^N$, is then a linear functional $\Lambda : \mathcal{D}(U) \rightarrow \mathbb{R}$, which is continuous with respect to the topology on $\mathcal{D}(U)$. i.e.

$$\text{If } \phi_n \rightarrow 0, \text{ w.r.t. } \tau \quad \text{then} \quad \Lambda \phi_n \rightarrow 0.$$

In particular, if the distribution Λ is given by $\Lambda(\phi) = \int_U \phi(x) f(x) dx$ is some smooth, integrable function $f \in C^\infty(U)$, then we may identify Λ with f and say Λ is a **smooth function**.

Definition 2.34 (Hypoelliptic operator) *An operator $\mathcal{A}$ is called hypoelliptic if, whenever u is a distribution on U , then u is actually a smooth function on any set for which $\mathcal{A}u$ is a smooth function.*

In his seminal paper, [14], Hörmander gave what is for most practical purposes a complete solution to the problem of identifying operator hypoellipticity. His criterion (see below) is phrased in terms of Lie algebras, and we use the notation already adopted in previous sections.

Theorem 2.35 (Hörmander (1967)) *Consider the operator $\mathcal{L}$ given by:*

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

Then $\mathcal{L}$ is hypoelliptic, if for some $m \in \mathbb{N}$ the following holds for all $x \in \mathbb{R}^N$

$$\text{Span} \{V_{[\alpha]}(x) : \alpha \in \mathcal{A}(m) \cup \{0\}\} = \mathbb{R}^N \quad (\mathbf{H})$$

The observant reader will have noted that this condition is different from the one adopted in the previous section, and indeed from the form of the UFG condition. In particular, in the above theorem the vector V_0 is considered in the span of the Lie algebra: in previous sections this vector field has not played a role. The reason for this is simple. In previous sections consideration has focussed on the diffusion semigroup, which is the solution of a parabolic PDE with the operator $\frac{\partial}{\partial t} - \mathcal{L}$, rather than $\mathcal{L}$ itself. The following corollary clarifies matters.

Corollary 2.36 *The operator $\frac{\partial}{\partial t} - \mathcal{L}$ is hypoelliptic providing for some $m \in \mathbb{N}$ the following holds for all $x \in \mathbb{R}^N$*

$$\text{Span} \{V_{[\alpha]}(x) : \alpha \in \mathcal{A}(m)\} = \mathbb{R}^N \quad (\mathbf{H}')$$

Proof : We can apply Hörmander's theorem by increasing the dimension by one, and considering the operator $\frac{\partial}{\partial t} - \mathcal{L}$ on $\mathbb{R}^{N+1}$. In this situation the condition **(H)** may be simplified, as the component $\frac{\partial}{\partial t}$ of the first order term, $\frac{\partial}{\partial t} - V_0$, of the sum of squares operator is in some sense 'on its own' in that it commutes with the vector fields $V_0, \dots, V_d$. This means that in order for **(H)** to hold in this case, $\frac{\partial}{\partial t} - V_0$ will need to span the time coordinate of $\mathbb{R}^{N+1}$. The effect of this is that V_0 is removed from consideration in the remaining $\mathbb{R}^N$ space components and **(H)** simplifies to **(H')**. ■

We consider now a concrete example, which has obvious connections with Brownian motion.

Example 2.37 *Let $\mathcal{A} := \lambda - \frac{d}{dx^2}$, where $\lambda > 0$, and consider the equation:*

$$\mathcal{A}u = \delta_0, \quad \text{on } \mathbb{R}.$$

*Then it is straightforward to verify that $\mathcal{A}$ satisfies Hörmander's condition **(H)**, and it is clear that $\mathcal{A}u$ is smooth on $\mathbb{R} \setminus \{0\}$. Hence u must be smooth on $\mathbb{R} \setminus \{0\}$. Indeed, by direct computation:*

$$u(x) = \gamma^{-1}e^{-\gamma|x|} + d_1e^{\gamma x} + d_2e^{-\gamma x}$$

where $\gamma = \sqrt{2\lambda}$ and d_1 and d_2 are given constants..

We now demonstrate through an example how the behaviour of sum of squares operator can be transformed by consideration of a time derivative $\frac{\partial}{\partial t}$ (i.e. the operator changes from elliptic to parabolic).

Example 2.38 Consider the example where $V_1, \dots, V_d \equiv 0$ and $V_0 = \frac{d}{dx}$. It is clear from (H) that $\mathcal{L} := V_0$ is hypoelliptic (this is easy to verify manually, of course), but it is also clear that (H') does not hold, and hence we cannot say that $\frac{\partial}{\partial t} - \mathcal{L}$ is hypoelliptic. Indeed, in this case we may identify the functional representation of u as $u(t, x) = \delta_0(t + x - \cdot)$, which is **not** a smooth function on $\mathbb{R}^N$. In the distributional sense this is:

$$u(\phi) := \int_{\mathbb{R}^N} \phi(y) \delta_0(x + t - y) dy$$

It is also clear by studying the corresponding ODE (or degenerate SDE), that this situation corresponds to $X_t^x = x + t$, whose law does not have a density. Indeed, the law is determined by the degenerate function u .

We now move away from the review of hypoellipticity as studied through the work of Hörmander and attempt to deduce hypoellipticity, through entirely probabilistic methods. Whilst probabilists can be rightly proud of the work of Malliavin, Bismut, Kusuoka, Stroock and co. in this area, it is also proper to retain a respect of the contribution of deterministic methods in this area. Indeed, the original contributor (Hörmander himself) was one such person and the area has since been significantly extended by others.

The following matrix, inkeeping with the notations set out in Kusuoka and Stroock [24], is closely related to the 'Malliavin covariance matrix' after its discovery and use by Malliavin in [30].

$$A(t, x) := \sum_{i=1}^d \int_0^t [(J_s^x)^{-1} V_i(X_s^x)]^{\otimes 2} ds$$

The invertibility of this matrix, which can be connected explicitly to the satisfaction of **(H')**, is of fundamental importance. Indeed, we shall write its determinant as

$$\Delta(t, x) := \det A(t, x)$$

The next theorem (Theorem 8.13 in [24]) provides a purely probabilistic criteria for hypoellipticity of parabolic PDEs in sum of squares form:

Theorem 2.39 (Kusuoka and Stroock (1985))

Assume $V_0, \dots, V_d \in C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$. Assume there is a non-decreasing $\rho : (0, 1] \rightarrow (0, \infty)$ such that for some constant C_p ,

$$\left\| \frac{1}{\Delta(t, x)} \right\|_{L^p(\Omega)} \leq \frac{C_p}{\rho_t},$$

for all $p \geq 1$ and $(t, x) \in (0, 1] \times \mathbb{R}^N$. Then if $t \log \rho_t \rightarrow 0$, as $t \rightarrow 0$, it holds that for each $c \in C_b^\infty$ the operator $\frac{\partial}{\partial t} - \mathcal{L} - c$ is hypoelliptic.

In the aforementioned paper Kusuoka and Stroock show, under the assumption of Hörmander's criterion for parabolic hypoellipticity, **(H')**, that this is indeed the case. In what follows, we seek to show that the criterion of the above theorem holds, by using the results which have been deduced solely via the UFG approach. This has not been done before, as far as the author is aware.

Theorem 2.40 (Probabilistic proof of Hörmander's theorem using UFG approach)

Assume **(H')** holds. Then the operator $\frac{\partial}{\partial t} - \mathcal{L} - c$ is hypoelliptic.

Proof : We have already shown in the previous section how the UFG condition is implied by **(H')**. Hence, we may assume that the UFG condition holds, and from its application we recall the following:

$$(J_s^x)^{-1} V_i(X_s^x) = \sum_{\alpha \in \mathcal{A}(m)} a_{i,\alpha}(s, x) V_{[\alpha]}(x)$$

By substituting this into the expression for $A(t, x)$ we get

$$\begin{aligned} A(t, x) &= \sum_{i=1}^d \int_0^t [(J_s^x)^{-1} V_i(X_s^x)]^{\otimes 2} ds \\ &= \sum_{\alpha, \beta \in \mathcal{A}(m)} \int_0^t \sum_{i=1}^d a_{i,\alpha}(s, x) a_{i,\beta}(s, x) ds \ V_{[\alpha]}(x) \otimes V_{[\beta]}(x) \\ &= \sum_{\alpha, \beta \in \mathcal{A}(m)} t^{(\|\alpha\| + \|\beta\|)/2} M_{\alpha, \beta}(t, x) \ V_{[\alpha]}(x) \otimes V_{[\beta]}(x) \end{aligned}$$

we denote for notational convenience $\tilde{M}(t, x) := (t^{(\|\alpha\| + \|\beta\|)/2} M_{\alpha, \beta}(t, x))_{\alpha, \beta \in \mathcal{A}(m)}$. Hence,

$$\begin{aligned} \Delta(t, x) &\geq \left(\inf_{\substack{\xi \in \mathbb{R}^N \\ |\xi|=1}} \xi A(t, x) \xi \right)^N \\ &= \left(\inf_{\substack{\xi \in \mathbb{R}^N \\ |\xi|=1}} y_{x, \xi}^T \tilde{M}(t, x) y_{x, \xi} \right)^N \end{aligned}$$

where $y_{x, \xi} := (\xi^T V_{[\alpha]}(x))_{\alpha \in \mathcal{A}(m)} \in \mathbb{R}^{N_m}$. Now, since **(H')** holds, it follows that $\inf_{x \in \mathbb{R}^N, |\xi|=1} |y_{x, \xi}|^2 =: \epsilon > 0$. Indeed, this point was made in the previous section in (2.31). Intuitively, if the span of the vectors in condition **(H')** is the whole of $\mathbb{R}^N$ then there cannot be a ξ for which $y_{x, \xi} = 0$. Hence

$$\begin{aligned} \inf_{\substack{\xi \in \mathbb{R}^N \\ |\xi|=1}} y_{x, \xi}^T \tilde{M}(t, x) y_{x, \xi} &\geq \epsilon \inf_{y \in \mathbb{R}^{N_m}} y^T \tilde{M}(t, x) y \\ &\geq \epsilon t^m \inf_{y \in \mathbb{R}^{N_m}} y^T M(t, x) y \end{aligned}$$

where we have used the fact that $t^{\|\alpha\|/2} \geq t^{m/2}$ for all $\alpha \in \mathcal{A}(m)$. We worked hard earlier in the

chapter to show that the RHS of the above satisfies some nice invertibility properties. Namely,

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in (0,1]}} \left\| \left(\inf_{y \in \mathbb{R}^{N_m}} y^T M(t, x) y \right)^{-N} \right\|_{L^p(\Omega)} \leq C_p$$

Hence,

$$\left\| \frac{1}{\Delta(t, x)} \right\|_{L^p(\Omega)} \leq t^{-mN} \left\| \left(\inf_{y \in \mathbb{R}^{N_m}} y^T M(t, x) y \right)^{-N} \right\|_{L^p(\Omega)} \leq C_p t^{-mN}$$

for all $(t, x) \in (0, 1] \times \mathbb{R}^N$, and we may identify

$$\rho_t := t^{mN}$$

and it follows that $t \log \rho_t = mN t \log t \rightarrow 0$ as $t \rightarrow 0$, as required. ■

3. Regularity of PDEs with C_b^k coefficients

In this chapter we consider the problem of inferring gradient bounds for stochastic flows which do **not** necessarily possess smooth, uniformly bounded coefficients. A UFG-type condition is imposed, but the finite differentiability restricts this condition. Due to non-smoothness, only a finite number of elements of the Lie ‘algebra’ are well-defined. In particular, we show that if the UFG condition is satisfied for this finite number of Lie brackets, then some regularity properties of the diffusion semigroup may be deduced. The derived differentiability property depends on the difference between the ‘order’ of the UFG condition and the order of coefficient smoothness. More than this, we are able to deduce the explicit rates of the previous sections. The generality of the UFG condition allows us to deduce a similar result for the case where the uniform Hörmander condition applies. In a later section we considerably extend the gradient bounds of Kusuoka to a very wide class of test functions and semigroups. The work in this section is used in a later chapter on existence and uniqueness to solutions of the Cauchy problem. As an application we consider the theoretical efficacy of the KLV method in approximating the diffusion semigroup; deriving precise convergence rates.

3.1. The UFG condition for non-smooth, non-bounded vector fields

In this section we again consider a stochastic flow of diffeomorphisms, on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, given by the family of SDE’s:

$$\begin{cases} dX_t^x = \sum_{i=0}^d V_i(X_t^x) \circ dB_t^i, & t > 0, \\ X_0^x = x, \end{cases} \quad (3.1)$$

where $B_t = (B_t^i)_{i=1,\dots,d}$ is a d -dimensional Brownian motion, and $V_0, \dots, V_d : \mathbb{R}^N \rightarrow \mathbb{R}^N$ are possibly **non-smooth** and **non-uniformly bounded** vector fields. The resultant diffusion semigroup, a one-parameter semigroup due to the time-homogeneity of the vector fields, is formed by taking

$$P_t f(x) := \mathbb{E} [f(X_t^x)].$$

It is known, through Itô’s lemma, that $(t, x) \mapsto P_t f(x)$ satisfies the parabolic PDE:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \mathcal{L}u, & \text{in } (0, T] \times \mathbb{R}^N, \\ u(\cdot, x) &= f, & \text{on } \mathbb{R}^N, \end{aligned}$$

in at least a weak sense, where $\mathcal{L} = \sum_{i=1}^d (V_i)^2 + V_0$ is a second-order differential operator, the *generator*, which has been perturbed by adding a *potential*, c . We shall be making the following assumptions on the potential term and the vector fields driving the diffusion:

$$V_i \in C_b^{k+1}, \quad i = 1, \dots, d, \quad \text{and } V_0 \in C_b^k, \quad \text{for some } k \geq 2. \quad (\mathbf{C})$$

We recall from Section 1.6 the definition of $V_{[\alpha]}$. We make an important observation - owing to the differentiability restrictions on the vector fields - that $V_{[\alpha]}$ is well-defined only for $|\alpha| \leq k + 1$. The key condition in this chapter is again a UFG-type condition. As was the case in the second chapter, the techniques employed are based on the analysis from the first chapter, which comes from the visionary paper of Kusuoka [22].

Definition 3.1 (UFG condition) *There exists $m \in \mathbb{N}$, with $m \leq k - 1$, such that for all $\alpha \in \mathcal{A}$ s.t. $\|\alpha\| = \{m, m + 1\}$ there exists $\varphi_{\alpha, \beta} \in C_b^{k+1-|\alpha|}(\mathbb{R}^N; \mathbb{R})$, uniformly bounded, where $\beta \in \mathcal{A}(m)$, such that:*

$$V_{[\alpha]} = \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha, \beta} V_{[\beta]}.$$

Or, in words: higher order Lie brackets can be expressed as a finite linear combination of lower order Lie brackets, for some fixed order m .

Remark 3.2

1. *The condition can be used inductively to obtain similar equalities for those higher-order Lie vector fields which are well-defined. In particular, we can use **(UFG)** to prove an analogous result for $V_{[\alpha]}$, for any $\alpha \in \mathcal{A}(k + 1)$.*
2. *The regularity of the coefficients $\varphi_{\alpha, \beta}$ is made in accordance with what one would expect, given the regularity of $V_{[\alpha]}$. In particular, notice that they are still required to be bounded. This condition should not be viewed as being restrictive. Indeed, the elements of the Lie algebra themselves have at most linear growth, and so it should normally be possible to make a choice of uniformly bounded UFG coefficients.*
3. *The demand in **(C)**, for more differentiability of the diffusion terms than of the drift term, is to ensure that the drift and diffusion terms for the corresponding Itô equation have the same level of differentiability.*

The goal of the first sections of the chapter is to prove the following representation and bounds for the derivatives of the diffusion semigroup (although this is significantly extended in later sections):

Theorem 3.3

$$V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t f(\cdot) = t^{-(\|\alpha_1\| + \dots + \|\alpha_N\|)/2} \mathbb{E}[f(X_t^x) \Phi_{\alpha_1, \dots, \alpha_N}(t, x)],$$

where $N \leq k - m$, $\alpha_1, \dots, \alpha_N \in \mathcal{A}(m)$, and $\Phi_{\alpha_1, \dots, \alpha_N}(\cdot, \cdot, \cdot)$ is an ‘integration by parts factor’. This provides the following gradient bound as a corollary:

$$\sup_{x \in K} |V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t f(x)| \leq C_K t^{-(\|\alpha_1\| + \dots + \|\alpha_N\|)/2} \|f\|_\infty.$$

for each $K \subset \mathbb{R}^N$, compact. In particular, for any uniformly continuous function, $f \in C_b^{BV}(\mathbb{R}^N)$, there holds $P_t f$ is $(k - m)$ -times differentiable along the vector fields of the Lie algebra.

There also holds, as a corollary of the preceding integration by parts formula, the following gradient bound:

$$\sup_{x \in K} |V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t f(x)| \leq C_K \frac{t^{\|\alpha_N\|/2}}{t^{(\|\alpha_1\| + \dots + \|\alpha_N\|)/2}} \|\nabla f\|_\infty.$$

In particular, for any globally Lipschitz continuous function, f , there holds $P_t^c f$ is $(k - m)$ -times differentiable along the vector fields of the Lie algebra.

The emphasis is placed once more on the explicit asymptotic rates of differentiability decay as powers of t .

Example 3.4 Consider the two-dimensional diffusion given by the SDE

$$\begin{aligned} d \begin{bmatrix} X_t^{1,x} \\ X_t^{2,x} \end{bmatrix} &= \begin{bmatrix} 0 \\ \theta(\mu - X_t^{2,x}) \end{bmatrix} dt + \begin{bmatrix} X_t^{1,x} \\ 0 \end{bmatrix} \circ dB_t^1 + \begin{bmatrix} 0 \\ \sqrt[k+2]{|X_t^{2,x}|^{k+2} + 1} \end{bmatrix} \circ dB_t^2 \\ \begin{bmatrix} X_0^{1,x} \\ X_0^{2,x} \end{bmatrix} &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \end{aligned}$$

where $\theta, \mu \in \mathbb{R}$. We now verify that the UFG condition is satisfied. Indeed, after identifying the vector fields, V_0, V_1, V_2 as:

$$V_0 = \begin{bmatrix} 0 \\ \theta(\mu - x_2) \end{bmatrix}, \quad V_1 = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 0 \\ \sqrt[k+1]{|x_2|^{k+1} + 1} \end{bmatrix}.$$

It is clear that $V_0, V_1 \in C_b^\infty$ and that $V_2 \in C_b^{k+1}$. Moreover, the only non-zero Lie bracket of order $|\alpha| \geq 2$ is $V_{[2,0]}$, and is given by:

$$V_{[2,0]} = \begin{bmatrix} 0 \\ \theta \left\{ 1 + \frac{(\mu - x_2)|x_2|^{k+1}}{|x_2|^{k+2} + 1} \right\} \sqrt[k+2]{|x_2|^{k+2} + 1} \end{bmatrix} = \varphi(\cdot) V_2,$$

It is clear that $\varphi = \varphi_{[2,0],[2]} \in C_b^k$ and hence the UFG condition holds for $m = 1$.

3.2. An integration by parts formula of limited applicability

As has been demonstrated and discussed in the previous sections, the gradient bounds are obtained by deriving an integration by parts formula. Let $f \in C_b^\infty(\mathbb{R}^N)$, then by following the same steps as in Section 2.1, we arrive at:

$$D^{(\alpha)} f(X_t^x) = \sum_{\beta \in \mathcal{A}(m)} t^{(\|\alpha\| + \|\beta\|)/2} M_{\alpha, \beta}(t, x) V_{[\beta]}(f \circ X_t)(x),$$

where

$$\begin{aligned} D^{(\alpha)} f(X_t^x) &:= \langle Df(X_t^x), k_\alpha(t, x) \rangle_H, \\ M_{\alpha, \beta}(t, x) &:= t^{-(\|\alpha\| + \|\beta\|)/2} \langle k_\alpha(t, x), k_\beta(t, x) \rangle_H \\ &= t^{-(\|\alpha\| + \|\beta\|)/2} \sum_{i=1}^d \int_0^t a_{i, \alpha}(u, x) a_{i, \beta}(u, x) du. \end{aligned}$$

Indeed, the finite differentiability assumptions do not interfere with this derivation. We only require $X_t^{(\cdot)}$ to be continuously differentiable, with a $\mathbb{P}$ -a.s. invertible Jacobian, X_t^x to be at least one-time Malliavin differentiable, and we require the vector fields to satisfy a UFG-type condition. We have insisted that the order of the UFG condition is smaller than the order of coefficient regularity, so this also poses no problem. Invertibility of this matrix was deduced in Chapter 1, Section 2.3. The lack of smoothness of this new matrix - in x and as a random variable - does not effect its invertibility. The proof is analogous and is hence merely referenced from Chapter 1.

Proposition 3.5 $M(t, x)$ is $\mathbb{P}$ -a.s. invertible. Moreover, for $p \in [1, \infty)$, $\alpha, \beta \in \mathcal{A}(m)$

$$\sup_{t \in (0, 1]} \mathbb{E} \left[M_{\alpha, \beta}^{-1}(t, x) \right]^p < C_p. \quad (3.2)$$

for some constant C_p .

Proof : Proof is analogous to that of Proposition 2.6 in Chapter 1. ■

From this result we may deduce, $\mathbb{P}$ -a.s:

$$V_{[\alpha]}(f \circ X_t)(x) = t^{-\|\alpha\|/2} \sum_{\beta \in \mathcal{A}(m)} t^{-\|\beta\|/2} M_{\alpha, \beta}^{-1}(t, x) D^{(\beta)} f(X_t^x),$$

and we shall use the operators of the Malliavin calculus to derive the integration by parts formula. The number of times we are able to use this relation to iterate the integration by parts formula, will depend on the regularity of $M(t, x)$ and $k_\alpha(t, x)$ for $\alpha \in \mathcal{A}(m)$. The classical results about SDE solutions can be applied and will prove important in quantifying this regularity (cf Proposition 2.3). The matrix $A(t, x) = (a_{\alpha, \beta}(t, x))_{\alpha, \beta}$, which is embedded in the definition of M , is the same matrix as (2.3), the only difference is in the regularity.

Proposition 3.6 *The matrix stochastic differential equation (2.3) has a unique solution, and its components $a_{\alpha,\beta} : [0, \infty) \times \mathbb{R}^N \rightarrow \mathbb{R}$, $\alpha, \beta \in \mathcal{A}(m)$ satisfy the mutually dependent SDEs:*

$$a_{\alpha,\beta}(t, x) = \delta_{\alpha,\beta} + \sum_{i=0}^d \sum_{\gamma \in \mathcal{A}(m)} \int_0^t c_{\alpha,\gamma}^i(X_u^x) a_{\gamma,\beta}(u, x) \circ dB_u^i.$$

Moreover, it may be assumed that $a_{\alpha,\beta}(t, \cdot) : \mathbb{R}^N \rightarrow \mathbb{R}$ is smooth: both with respect to standard differentiation and Malliavin differentiation, that $a_{\alpha,\beta}(\cdot, \cdot)$ is jointly continuous in $[0, \infty) \times \mathbb{R}^N$ with probability one, for each $\alpha, \beta \in \mathcal{A}(m)$, and

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in [0, T]}} \mathbb{E} \left| \frac{\partial^{|\gamma|}}{\partial x^\gamma} a_{\alpha,\beta}(t, x) \right|^p < \infty, \quad \forall p \in [1, \infty), T > 0, \quad (3.3)$$

for any multi-index γ satisfying $|\gamma| < k - m$. Finally, for any $k \in \mathbb{N}$

$$\sup_{t \in [0, T]} \mathbb{E} \left\| D^k a_{\alpha,\beta}(t, x) \right\|_{H^{\otimes k}}^p < C_{k,p} (1 + |x|)^p \quad \forall p \in [1, \infty), T > 0, \quad (3.4)$$

Furthermore, the matrix $A = (a_{\alpha,\beta})_{\alpha,\beta \in \mathcal{A}(m)}$ is invertible, and its inverse satisfies the matrix SDE:

$$B(t, x) = I - \sum_{i=0}^d \int_s^t B(u, x) C^i(X_u^x) \circ dB_u^i.$$

Moreover, the components $b_{\alpha,\beta}$ of B , $\alpha, \beta \in \mathcal{A}(m)$, are a.s. $k - m$ times differentiable in x for fixed $t \in [0, \infty)$, and jointly continuous in (t, x) , similarly with

$$\sup_{\substack{t \in [0, T] \\ x \in \mathbb{R}^N}} \mathbb{E} \left| \frac{\partial^{|\gamma|}}{\partial x^\gamma} b_{\alpha,\beta}(t, x) \right|^p < C_{T,p}, \quad (3.5)$$

for each $p \in [1, \infty)$, $T > 0$ and some constant $C_{T,p}$, where $|\gamma| \leq k - m$. Finally, for any $k \in \mathbb{N}$

$$\sup_{t \in [0, T]} \mathbb{E} \left\| D^k b_{\alpha,\beta}(t, x) \right\|_{H^{\otimes k}}^p < C_{k,p} (1 + |x|)^p \quad \forall p \in [1, \infty), T > 0, \quad (3.6)$$

Remark 3.7 *The above proposition highlights an idiosyncratic difference between the Malliavin derivative and the normal derivative for the solutions of such SDEs. It stems from the fact that the Malliavin derivative for the SDE of X_t^x has an unbounded norm¹ over $x \in \mathbb{R}^N$, as it has Lipschitz continuous coefficients. However, the same result for the norm of the standard derivatives is bounded*

¹ Recall Theorem 1.13

over $x \in \mathbb{R}^N$. This difference did not appear in the first chapter as we assumed the vector fields were uniformly bounded.

Proof : This is very similar to Proposition 2.3. The only difference here is that the bounds on the norms of the iterated Malliavin derivatives are now bounded only linearly in $|x|$. This is obvious once one considers Theorem 1.13 and in particular equation (1.16). It is clear from this equation that the norm of the Malliavin derivatives inherits the linear growth of the vector fields. All higher order Malliavin derivatives inherit this linearity from the first order Malliavin derivative, but given the boundedness of the derivatives of the vector fields, have no worse than linear growth. ■

3.3. A finite number of sharp gradient bounds

The importance of deducing regularity properties of the inverse of the covariance matrix and related objects, has already been exhibited in the previous sections. The elements of the integration by parts formula are again called Kusuoka-Stroock processes, despite being slightly different to those in the first two chapters.

Definition 3.8 (Local Kusuoka-Stroock Processes) Let E be a separable Hilbert space. We denote by $\mathcal{K}_r^{loc}(E, k)$ the set of functions: $f : (0, 1] \times \mathbb{R}^N \rightarrow \mathbb{D}^{k, \infty}(E)$ satisfying

1. $f(t, \cdot)$ is k -times continuously differentiable and $\frac{\partial^\alpha f}{\partial x^\alpha}(\cdot, \cdot)$ is continuous in (t, x) $\mathbb{P}$ -a.s. for any multi-index $|\alpha| \leq k$, $t \in (0, 1]$.
2. $\sup_{t \in (0, 1], x \in K} t^{-r/2} \left\| \frac{\partial^\alpha f}{\partial x^\alpha} \right\|_{n, p, E} < \infty$, for all $n \leq k - |\alpha|$ and $p \in [1, \infty)$ for each $K \subset \mathbb{R}^N$, compact.

Lemma 3.9 (Properties of Kusuoka-Stroock Processes) The following is a list of some properties of K-S Processes:

1. Suppose $f \in \mathcal{K}_r^{loc}(E, n)$, where $r \geq 0$. Then, for $i = 1, \dots, d$,

$$\int_0^\cdot f(s, x) dB_s^i \in \mathcal{K}_{r+1}^{loc}(E, n) \quad \text{and} \quad \int_0^\cdot f(s, x) ds \in \mathcal{K}_{r+2}^{loc}(E, n).$$

2. $a_{\alpha, \beta}, b_{\alpha, \beta} \in \mathcal{K}_{(\|\beta\| - \|\alpha\|) \vee 0}^{loc}(\mathbb{R}, k - m)$ where $\alpha, \beta \in \mathcal{A}(m)$.
3. $k_\alpha \in \mathcal{K}_{\|\alpha\|}^{loc}(H, k - m)$, where $\alpha \in \mathcal{A}(m)$.
4. $D^{(\alpha)}u := \langle Du(t, x), k_\alpha \rangle_H \in \mathcal{K}_{r+\|\alpha\|}^{loc}(\mathbb{R}, n \wedge (k - m))$ where $u \in \mathcal{K}_r^{loc}(\mathbb{R}, n)$ and $\alpha \in \mathcal{A}(m)$.
5. If $M^{-1}(t, x)$ is the inverse matrix of $M(t, x)$, then $M_{\alpha, \beta}^{-1} \in \mathcal{K}_0^{loc}(\mathbb{R}, k - m)$, $\alpha, \beta \in \mathcal{A}(m)$.
6. If $f_i \in \mathcal{K}_{r_i}^{loc}(\mathbb{R}, n_i)$ for $i = 1, \dots, N$, then

$$\prod_{i=1}^N f_i \in \mathcal{K}_{r_1 + \dots + r_N}^{loc}(\mathbb{R}, \min_i n_i) \quad \text{and} \quad \sum_{i=1}^N f_i \in \mathcal{K}_{\min_i r_i}^{loc}(\mathbb{R}, \min_i n_i).$$

Proof : The proof of this is similar to the analogous proof in the first chapter. See appendix for further details. ■

SEVERAL INTEGRATION BY PARTS FORMULAE

In this section we synthesise the developed results to obtain various integration by parts formulae, in a way which should now be familiar. The proofs are omitted as they are analogous to the proofs in the first and second chapters. In what follows, and unless otherwise stated, we assume that $f \in C_b^\infty(\mathbb{R}^N)$:

Theorem 3.10 (Integration by Parts formula I) *Under the conditions (C, UFG) the following integration by parts formula holds for $\Phi \in \mathcal{K}_r^{loc}(\mathbb{R}, n)$, and for any $\alpha \in \mathcal{A}(m)$:*

$$\mathbb{E} [\Phi(t, x) V_{[\alpha]}(f \circ X_t)(x)] = t^{-\|\alpha\|/2} \mathbb{E} [\Phi_\alpha(t, x) f(X_t^x)], \quad (3.7)$$

where $\Phi_\alpha \in \mathcal{K}_r^{loc}(\mathbb{R}, (n-1) \wedge (k-m-1))$. Moreover, for any $q > p$

$$\sup_{t \in (0,1]} \mathbb{E} |\Phi_\alpha(t, x)|^p \leq C_{p,q} (1 + |x|)^p \sup_{t \in (0,1]} \mathbb{E} \|\Phi(t, x)\|_{2,q}^p$$

Corollary 3.11 (Integration by Parts formula II) *Under the assumptions of Theorem 3.10 the following holds:*

$$\mathbb{E} [\Phi(t, x) (V_{[\alpha]} f)(X_t^x)] = t^{-\|\alpha\|/2} \mathbb{E} [\Phi'_\alpha(t, x) f(X_t^x)],$$

where $\Phi'_\alpha \in \mathcal{K}_r$, for any $f \in C_b^\infty(\mathbb{R}^N; \mathbb{R})$, $t > 0$, $x \in \mathbb{R}^N$. Moreover, for any $q > p$

$$\sup_{t \in (0,1]} \mathbb{E} |\Phi'_\alpha(t, x)|^p \leq C_{p,q} (1 + |x|)^p \sup_{t \in (0,1]} \mathbb{E} \|\Phi(t, x)\|_{2,q}^p$$

Corollary 3.12 (Integration by Parts formula III) *Under the same conditions as Theorem 3.10, the following integration by parts formula holds for $\Phi \in \mathcal{K}_r^{loc}(\mathbb{R}, n)$, and for any $\alpha \in \mathcal{A}(m)$:*

$$V_{[\alpha]} \mathbb{E} [\Phi(t, x) f(X_t^x)] = t^{-\|\alpha\|/2} \mathbb{E} [\Phi''_\alpha(t, x) f(X_t^x)], \quad (3.8)$$

where $\Phi''_\alpha \in \mathcal{K}_r^{loc}(\mathbb{R}, (n-1) \wedge (k-m-1))$. Moreover, for any $q > p$:

$$\sup_{t \in (0,1]} \mathbb{E} |\Phi''_\alpha(t, x)|^p \leq C_{p,q} (1 + |x|)^p \sup_{t \in (0,1]} \|\Phi(t, x)\|_{2,q}^p \quad (3.9)$$

Corollary 3.13 (Integration by Parts formula IV) *Under the same conditions as Theorem 3.10, the following integration by parts formula holds for $N + M \leq k - m$ and $\alpha_1, \dots, \alpha_N \in \mathcal{A}(m)$:*

$$V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t (V_{[\alpha_{N+1}]} \dots V_{[\alpha_{N+M}]} f)(x) = t^{-(\|\alpha_1\| + \dots + \|\alpha_{N+M}\|)/2} \mathbb{E} [\Phi_\alpha(t, x) f(X_t^x)], \quad (3.10)$$

where $\Phi_{\alpha_1, \dots, \alpha_N} \in \mathcal{K}_0^{loc}(\mathbb{R}, (k - m - N))$. Moreover,

$$\sup_{t \in (0,1]} \mathbb{E} |\Phi_\alpha(t, x)|^p \leq C_p (1 + |x|)^{Np}$$

Corollary 3.14 (IBPF for Semigroup with a Potential) *Under the conditions (C, UFG), and assuming $c \in C^{k+m-1}(\mathbb{R}^N)$ the following integration by parts formula holds for $N \leq k - m$ and $\alpha_1, \dots, \alpha_N \in \mathcal{A}(m)$:*

$$V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t^c f(x) = t^{-(\|\alpha_1\| + \dots + \|\alpha_N\|)/2} \mathbb{E} [\Phi_{\alpha_1, \dots, \alpha_N}^c(t, x) f(X_t^x)], \quad (3.11)$$

where $\Phi_{\alpha_1, \dots, \alpha_N}^c \in \mathcal{K}_0^{loc}(\mathbb{R}, (k - m - N))$. Moreover,

$$\sup_{t \in (0,1]} \mathbb{E} |\Phi_{\alpha_1, \dots, \alpha_N}^c(t, x)|^p \leq C_p (1 + |x|)^{Np}$$

Proof : The proofs of these are similar to the proofs in Chapter 1, and are thus left to the appendix.

■

Remark 3.15 *Observe that we are able to quantify exactly how the derivatives explode - as functions of x - based on an analysis of the integration by parts factors. In Chapter 5 we shall use the bounds on $\mathbb{E} |\Phi_\alpha(t, x)|$, etc, to prove gradient bounds on weighted L^p spaces, which compensate for this explosion. For now we derive local gradient bounds.*

3.4. Applications to gradient bounds

THE BASIC GRADIENT BOUNDS

In the previous chapter it was relatively straightforward to deduce explicit gradient bounds from the integration by parts formulae. This still holds true in the more general setting, but the types of gradient bounds we are able to deduce changes.

Lemma 3.16 *Assume that $\Phi \in \mathcal{K}_r^{loc}(n)$, where $r \in \mathbb{R}$, $n \in \mathbb{N}$. Then for each compact $K \subset \mathbb{R}^N$, there is a constant $C_K < \infty$ such that:*

$$\sup_{x \in K} |P_t^\Phi f(x)| \leq C_K t^{r/2} \|f\|_\infty,$$

for all $t \in (0, 1]$ and bounded functions $f \in C_0^\infty(\mathbb{R}^N)$.

Proof : The proof is analogous to Lemma 2.25, the only difference comes from the fact that Φ is a local Kusuoka-Stroock process, and hence,

$$\sup_{x \in K} \|\Phi(t, x)\|_{L^p(\mathbb{P})} < C_K t^{r/2},$$

for $t \in (0, 1]$. ■

The lack of global boundedness of Kusuoka-Stroock processes has a knock-on effect in the type of gradient bounds the IBPF lead to.

Corollary 3.17 *Assume that $\Phi \in \mathcal{K}_r^{loc}(k - m)$, $r \in \mathbb{R}$, and let $f \in C_0^\infty(\mathbb{R}^N)$. Let $\alpha_1, \dots, \alpha_{N+M} \in \mathcal{A}(m)$, and assume $N + M \leq k - m$. Then for each compact subset K of $\mathbb{R}^N$, there is a constant $C_K < \infty$ such that:*

$$\sup_{x \in K} |V_{[\alpha_1]} \dots V_{[\alpha_N]}(P_t^\Phi V_{[\alpha_{N+1}]} \dots V_{[\alpha_{N+M}]} f)(x)| \leq C_K t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{N+M}\|}{2}} \|f\|_\infty. \quad (3.12)$$

As we have already seen through the discussion on the KLV method, one may also wish to consider test functions, f , which are not uniformly bounded. In particular, Lipschitz functions may also be considered.

Corollary 3.18 *Assume $N \leq k - m$, and let $f \in C_b^\infty(\mathbb{R}^N)$ such that $\text{Supp } \nabla f \subset K$. Then for each compact $K \subset \mathbb{R}^N$, there is a constant $C_K < \infty$ such that for $\alpha_1, \dots, \alpha_N \in \mathcal{A}(m)$:*

$$\sup_{x \in K} |V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t f(x)| \leq \frac{C_K t^{1/2}}{t^{(\|\alpha_1\| + \dots + \|\alpha_N\|)/2}} \|\nabla f\|_\infty. \quad (3.13)$$

Proof : Proof is the same as that of Corollary 2.27 ■

The gradient bounds need not be restricted to the L^∞ norm. For general gradient bounds, we must use test functions of compact support, and the gradient bounds extend to L^p norms.

Theorem 3.19 *Assume $\alpha_1, \dots, \alpha_{N+M} \in \mathcal{A}(m)$, where $N + M \leq m - k$. Let $K \subset \mathbb{R}^N$ be compact. Then there is a constant $C_{p,K}$ such that:*

$$\|V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t V_{[\alpha_{N+1}]} \dots V_{[\alpha_{N+M}]} f\|_{L^p(K, dx)} \leq C_{p,K} t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{N+M}\|}{2}} \|f\|_{L^p(K, dx)}$$

for any $t \in (0, 1]$, $p \in [1, \infty]$ and $f \in C_0^\infty$ such that $\text{Supp } f \subset K$.

Proof : The proof is very similar to that of Theorem 2.29 and based on an interpolation argument. Unlike the previous Theorem, here we provide the argument from Kusuoka and Stroock [24, pp16-17]. The cases $p = \infty$ and $p = 1$ are considered separately and then the *Riesz-Thorin interpolation theorem* is used applied to obtain $p \in (1, \infty)$. The case $p = \infty$ follows trivially from applying Corollary 3.13 and Lemma 3.16. We prove the case $p = 1$. Define

$$\begin{aligned} \tilde{V}_0 &= -V_0 + \frac{1}{2} \sum_{j=1}^d \text{div}(V_j) V_j, \\ \tilde{V}_i &= V_i, \quad i = 1, \dots, d, \\ \tilde{c} &= c - \text{div}(V_0) + \frac{1}{2} \sum_{j=1}^d \text{div}(V_j) + \frac{1}{2} \sum_{j=1}^d [\text{div}(V_j)]^2 \end{aligned}$$

Moreover, $\{\tilde{V}_i, i = 0, \dots, d\}$ satisfies the UFG condition. Consider the SDE

$$\tilde{X}_t^x = x + \sum_{i=0}^d \int_0^t \tilde{V}_i(X_s^x) \circ dB_s^i.$$

Then $\tilde{P}_t^{\tilde{c}}$ given by

$$\tilde{P}_t^{\tilde{c}} f(x) := \mathbb{E} \left[\exp \left\{ \int_0^t \tilde{c}(\tilde{X}_s^x) ds \right\} f(\tilde{X}_t^x) \right]$$

is the formal adjoint of P_t^c and so

$$\int_{\mathbb{R}^N} (P_t^c f)(x) g(x) dx = \int_{\mathbb{R}^N} f(x) (\tilde{P}_t^{\tilde{c}} g)(x) dx, \quad f, g \in C_0^\infty, \text{ Supp } f, g \subset K$$

Hence,

$$\begin{aligned} & \left\| V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t^c V_{[\alpha_{N+1}]} \dots V_{[\alpha_{N+M}]} f \right\|_{L^1(K)} \\ &= \sup_{\substack{g \in C_0^\infty(\mathbb{R}^N) \\ \|g\|_\infty \leq 1 \\ \text{Supp } g \subset K}} \left\| V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t^c V_{[\alpha_{N+1}]} \dots V_{[\alpha_{N+M}]} f g \right\|_{L^1} \\ &= \sup_{\substack{g \in C_0^\infty(\mathbb{R}^N) \\ \|g\|_\infty \leq 1 \\ \text{Supp } g \subset K}} \left\| f V_{[\alpha_{N+M}]}^* \dots V_{[\alpha_{N+1}]}^* \tilde{P}_t^{\tilde{c}} V_{[\alpha_N]}^* \dots V_{[\alpha_1]}^* g \right\|_{L^1} \\ &\leq \sup_{\substack{x \in K \\ g \in C_0^\infty(\mathbb{R}^N) \\ \|g\|_\infty \leq 1 \\ \text{Supp } g \subset K}} \left| V_{[\alpha_{N+M}]}^* \dots V_{[\alpha_{N+1}]}^* \tilde{P}_t^{\tilde{c}} V_{[\alpha_N]}^* \dots V_{[\alpha_1]}^* g(x) \right| \|f\|_{L^1} \\ &\leq \tilde{C}_K t^{-(\|\alpha_1\| + \dots + \|\alpha_{N+M}\|)/2} \|f\|_{L^1} \end{aligned}$$

as required. The last inequality follows from the application of the result to $\tilde{P}_t^{\tilde{c}}$ for the case $p = \infty$.

■

RECOVERING HÖRMANDER'S THEOREM

In the first chapter, we demonstrated how the probabilistic interpretation of Hörmander's theorem on hypoellipticity of parabolic second order operators - smoothness of the diffusion semigroup - may be recovered. We even showed how the hypoellipticity property itself could be shown using the techniques of Kusuoka and Stroock in [24]. Whilst we do not consider the problem of hypoellipticity in our more general setting, we are still able to deduce strong semigroup regularity results under the assumption of a Hörmander-type condition. In the case where the vector fields of the underlying SDE are non-smooth, it is not reasonable to expect smoothness of the semigroup. It is unsurprising given the results of the previous section, however, that it is possible to show differentiability up to some level. We now note a variant on Malliavin's criterion for the existence and regularity of densities for the law of the diffusion:

Theorem 3.20 Assume that the following holds for all $x \in \mathbb{R}^N$, and for some $m \leq k - 1$:

$$\text{Span}\{V_{[\alpha]}(x) : \alpha \in \mathcal{A}(m)\} = \mathbb{R}^N. \quad (3.14)$$

Then the diffusion semigroup maps uniformly continuous functions to C^{k-m} -functions.

Proof : (2.30) may be restated, as $\exists \epsilon > 0$ s.t.:

$$\sum_{\alpha \in \mathcal{A}(m)} |\xi^T V_{[\alpha]}(x)|^2 \geq \epsilon |\xi|^2, \quad (3.15)$$

$\forall \xi \in \mathbb{R}^N$, or equivalently: the matrix $(VV^T)(x)$ is invertible $\forall x \in \mathbb{R}^N$, where $V(x) := (V_{[\alpha_1]}(x) \mid \dots \mid V_{[\alpha_{N_m}]}(x))$. Note: this matrix must have C_b^{k-m+1} entries, by the inverse function theorem, as this is the minimum level of differentiability for $V_{[\alpha]}$, $\alpha \in \mathcal{A}(m)$. It must be shown that an integration by parts formula can be obtained for the partial derivatives ∂_i . This may then be iterated to obtain the same result for ∂_α .

Claim: there exist $C_\alpha^i \in C_b^{k-m+1}$, which are **not** necessarily bounded, such that:

$$e_i = \sum_{\alpha \in \mathcal{A}(m)} C_\alpha^i(x) V_{[\alpha]}(x),$$

$\forall x \in \mathbb{R}$. This may be written in matrix form as

$$e_i = VC^i,$$

again $V(x) := (V_{[\alpha_1]}(x) \mid \dots \mid V_{[\alpha_{N_m}]}(x))$, and $C^i(x) = (C_{\alpha_1}^i(x) \dots C_{\alpha_{N_m}}^i(x))^T$. But, from the assumptions, $(VV^T)(x)$ is invertible $\forall x \in \mathbb{R}^N$. Therefore, choose

$$C^i = V^T(VV^T)^{-1}e_i.$$

That is, $C_\alpha^i(x) = (V^T(VV^T)^{-1}e_i)_\alpha(x)$. By the inverse function theorem, $C_\alpha^i \in C_b^{k-m+1}$. Note that $C_\alpha^i \in \mathcal{K}_0^{\text{loc}}(\mathbb{R}, k - m + 1)$, as it need not be the case that the function is uniformly bounded in x . Observe, for any compact K

$$\begin{aligned} \sup_{x \in K} |\partial_i P_t \varphi(x)| &= \sup_{x \in K} \left| \sum_{\alpha \in \mathcal{A}(m)} C_\alpha^i(x) V_{[\alpha]} P_t \varphi(x) \right| \\ &\leq \sup_{x \in K} \sum_{\alpha \in \mathcal{A}(m)} |\mathbb{E} [\Phi_\alpha^i(t, x) \varphi(X_t^x)]| \\ &\leq C_{K,i} \|\varphi\|_\infty. \end{aligned}$$

Using the integration by parts formulae derived in the previous sections. This may be done for any partial derivative and the procedure can be iterated for any multi-index α , satisfying $|\alpha| \leq k - m + 1$,

to deduce the gradient bounds:

$$\begin{aligned} \sup_{x \in K} |\partial_{i_1} \dots \partial_{i_M} P_t f(x)| &\leq C_K t^{-Mr/2} \|f\|_\infty \\ \sup_{x \in K} |\partial_{i_1} \dots \partial_{i_M} P_t f(x)| &\leq C_K \frac{t^{1/2}}{t^{Mr/2}} \|\nabla f\|_\infty, \end{aligned}$$

for any $i_1, \dots, i_N \in \{1, \dots, N\}$, $M \leq k - m$. This is sufficient to imply the required differentiability result, as showing uniform convergence on compacts in the proof of Proposition 2.31. ■

3.5. Application of the KLV method to non-smooth semigroups

In this section we seek to use the results of the previous one to deduce theoretical error bounds of the KLV method. As was demonstrated in the previous section, the one-parameter semigroup resulting from a diffusion with C^k coefficients is not smooth. This non-smoothness will affect the theoretical efficiency of the KLV method. **Note:** in this section, we shall still be assuming that the vector fields comprising the underlying SDE are uniformly bounded. It is not obvious that the work of this chapter can be extended to non-bounded coefficients. Firstly, we consider Stratonovich Taylor expansions for non-smooth functions.

Assume $f \in C^{n+2}$ along the vector fields $V_0, \dots, V_d$. That is: $V_{i_1} \dots V_{i_{n+2}} f$ exists everywhere and is continuous for any $i_1, \dots, i_{n+2} \in \{0, 1, \dots, d\}$. Assume further that $V_0 \in C_b^k(\mathbb{R}^N; \mathbb{R}^N)$ and $V_1, \dots, V_d \in C_b^{k+1}(\mathbb{R}^N; \mathbb{R}^N)$, where $k \geq n + 2$. Then

$$f(X_t^x) = \sum_{\alpha \in \mathcal{A}(d_1-1)} (V_\alpha f)(x) \int_{0 < t_1 < \dots < t_k < t} \circ dB_{t_1}^{\alpha_1} \dots \circ dB_{t_k}^{\alpha_k} + R_{d_1}(t, x, f), \quad (3.16)$$

where $d_1 \leq n$ and

$$R_{d_1}(t, x, f) = \sum_{\substack{\alpha \in \mathcal{A}(d_1-1) \\ i=0, \dots, d, \\ \text{s.t. } i * \alpha \notin \mathcal{A}(d_1-1)}} \int_{0 < t_0 < t_1 < \dots < t_k < t} (V_{i * \alpha} f)(X_{t_0}^x) \circ dB_{t_0}^i \circ dB_{t_1}^{\alpha_1} \dots \circ dB_{t_k}^{\alpha_k}.$$

Note that if $d_1 = n$ we may **not** expand the items appearing in the remainder term any further. Observe that $V_{i * \alpha} f \in C^2$ along $V_0, \dots, V_d$ for $\alpha \in \mathcal{A}(n-1)$ and $i = 0, \dots, d$. Since we need to express $V_{i * \alpha} f(X_t^x)$ as the integrand of a Stratonovich integral, and hence as a semimartingale², we need to be able to apply Itô's rule to it. Hence, this is the minimum differentiability requirement for this operation to be well-defined. If we were to expand the remainder term further, then we couldn't be certain that we could write $V_{i * \alpha} f(X_t^x)$ as a semi-martingale.

²as this semimartingale representation forms part of the definition

The following bound on the remainder is shown in Lyons and Victoir [29]:

$$\sup_{x \in \mathbb{R}^N} \left(\mathbb{E} R_{d_1}(t, x, f)^2 \right)^{1/2} \leq C \sum_{j=d_1+1}^{d_1+2} t^{j/2} \sup_{\substack{\alpha \in \mathcal{A}_j \setminus \mathcal{A}_{j-1} \\ |\alpha| \leq d_1}} \|V_\alpha f\|_\infty.$$

Indeed, the following bound for single applications of cubature with test functions of finite differentiability holds:

Proposition 3.21 *For any $d_1 \leq n$ there holds the following global error for a single application of the cubature measure:*

$$\left\| \mathbb{E} \left[f(X_t^{(\cdot)}) \right] - \mathbb{E}_{\mathbb{Q}} \left[f(X_t^{(\cdot)}) \right] \right\|_\infty \leq C \sum_{j=d_1+1}^{d_1+2} t^{j/2} \sup_{\substack{\alpha \in \mathcal{A}_j \setminus \mathcal{A}_{j-1} \\ |\alpha| \leq d_1}} \|V_\alpha f\|_\infty. \quad (3.17)$$

Proof : The proof is analogous to that in Lyons and Victoir [29]. ■

As we noted in the introduction to the KLV method, the term on the right hand side of the above inequality is small only if t and/or the driving noise of the SDE is small. The approach is refined by partitioning the interval and applying cubature iteratively over each subinterval of the partition.

THE KLV METHOD

Recall the definition of the algorithm which constitutes the KLV method³. The Markovian property of the approach allows one to bound the error over the global interval $[0, T]$, by the sum of the errors over the subintervals of the partition. In particular, for test functions of finite regularity:

Proposition 3.22 *The KLV approximation of degree d_1 satisfies, for $d_1 \leq k - m - 2$*

$$\left\| (P_T f)(\cdot) - \mathbb{E}_{KLV(\mathcal{D}, \cdot)} f \right\|_\infty \leq C \sum_{l=1}^k \sum_{j=d_1+1}^{d_1+2} s_l^{j/2} \sup_{\substack{\alpha \in \mathcal{A}_j \setminus \mathcal{A}_{j-1} \\ |\alpha| \leq k-m-2}} \|V_\alpha P_{T-t_l} f\|_\infty. \quad (3.18)$$

Proof : This follows from the same proof as in Lyons and Victoir [29], once one has considered (1.24). Namely that $P_{T-t_l} f \in C^{k-m}$, i.e. n is $k - m - 2$ in the above. ■

A PROBLEM WITH THE STRATONOVICH TAYLOR EXPANSION

The same problem exists for the finite differentiability case, as exists in the smooth case. When one seeks to use the gradient bounds to provide an upper bound for the terms $\|V_\alpha P_{T-t_l} f\|_\infty$, $\alpha \in \mathcal{A}_{m+2} \setminus \mathcal{A}_m$, one may not consider derivatives along the vector field V_0 . The issue was rectified in Crisan and Ghazali [8] by assuming an extra condition; the so-called **V0** condition.

³See (1.25) and (1.26).

Definition 3.23 (V0 condition) A family of vector fields $V_i, 0 \leq i \leq d$ is said to satisfy the **V0** condition if

$$V_0 = \sum_{\beta \in \mathcal{A}(2)} u_\beta V_{[\beta]},$$

for some $u_\beta \in C_b^{k+1}(\mathbb{R}^N)$.

Therefore, if the **V0** condition is assumed and one refers back to Corollary 2.28, it is easy to see how we apply this with the aim of obtaining an upper bound for the global error. Indeed, from Corollary 2.28, and Corollary 2.27 we get the following:

Corollary 3.24 There exists a constant C such that for all $\alpha \in \mathcal{A}(k-m)$

$$\|V_\alpha P_t f\|_\infty \leq C \frac{t^{1/2}}{t^{\|\alpha\|/2}} \|\nabla f\|_\infty.$$

We briefly mention the recent developments of Litterer [27], in which the author has shown that the **V0** condition on the drift can be relaxed and that one still obtains the same rate of convergence as shown above. It may be possible that these techniques may also be applied to cover the finite differentiability case. We have the following theorem:

Theorem 3.25 Suppose the vector fields $V_i, 0 \leq i \leq d$ satisfy the UFG condition, then

$$\|P_T f(\cdot) - \mathbb{E}_{KLV(\mathcal{D}, \cdot)} f\|_\infty \leq C_T \|\nabla f\|_\infty \left(s_k^{1/2} + \sum_{j=m}^{m+1} \sum_{i=1}^{k-1} \frac{s_i^{(j+1)/2}}{(T-t_i)^{j/2}} \right),$$

where the constant C_T is independent of the number of subintervals in the partition of $[0, T]$

Proof : This is proved in Lyons and Victoir [29], and is a simple combination of Proposition 3.22 and Corollary 3.24. ■

By taking uneven partitions of the interval $[0, T]$ we can derive high order convergence of the KLV method. Define, for $\gamma > 0$ and $0 \leq j \leq M$

$$t_j := T \left(1 - \left(1 - \frac{j}{M} \right)^\gamma \right).$$

Corollary 3.26 For $d_1 \leq k - m - 2$ we have the following global convergence rates:

$$\begin{aligned} \|P_T f(\cdot) - \mathbb{E}_{KLV(\mathcal{D}, \cdot)} f\|_\infty &\leq K n^{-\gamma/2} \|\nabla f\|_\infty, & \text{if } 0 < \gamma < d_1 - 1, \\ &\leq K n^{-(d_1-1)/2} \log(n) \|\nabla f\|_\infty, & \text{if } \gamma = d_1 - 1, \\ &\leq K n^{-(d_1-1)/2} \|\nabla f\|_\infty, & \text{if } \gamma > d_1 - 1. \end{aligned}$$

Remark 3.27 *It has been written in the above that d_1 , the degree of the applied cubature measure, satisfies $d_1 \leq k - m - 2$. It should be made clear that it is, of course, possible to implement a KLV scheme with a cubature formula of higher order. However, the theoretical gains from increasing the order are **limited** for the KLV method based on cubature formulae of order higher than $k - m - 2$. At this stage we cannot further expand the Stratonovich Taylor expansion. Hence, if one were implementing solely on the basis of the theoretical bounds, it would make little sense to implement the KLV method using a cubature formula of order higher than $k - m - 2$. It would undoubtedly involve greater computational effort, for no extra theoretical gain.*

4. The LFG condition and local differentiability of the semigroup

In this chapter we consider local differentiability of the diffusion semigroup under weak conditions. In particular, we seek to prove local differentiability of the semigroup along directions of the Lie algebra. The main tools for this exposé are the operators of the Malliavin Calculus, and the very general condition imposed is a local version of the condition used in Kusuoka and Stroock [25] and Kusuoka [22] known as the *UFG condition*. The work in this chapter replicates the acquisition of the precise rates of Kusuoka and Stroock.

A review of the literature on differentiability of the semigroup and related notions has already been given in the introduction and at the start of the first chapter. As far as the author is aware, this is the first time that a local notion of non-degeneracy has been used to derive local differentiability results, and is unaware of any further literature in this area. Since writing the thesis the author was made aware of a similar condition to LFG, which was discussed in Sussmann [39]. They are referred to by Sussmann as (LFT1) and (LFT2), and are used to prove integrability of distributions, rather than regularity of the diffusion semigroup. This paper is of more relevance in the next chapter, where it is duly discussed.

We note also several important techniques from the literature which have inspired the work in this chapter. In particular, the introduction of the truncating factor Y , which is used to develop the integration by parts formula of this chapter, is inspired by the work of Nualart in [33], where the author uses a similar factor to deduce existence and smoothness of the density for the law of the Brownian sheet. Several properties of the truncation factor owe much to the remarkable modulus of continuity inequality derived in Garsia, Rodemich and Rumsey [10], and the work of Airault and Malliavin in [1].

In this chapter we shall make some assumptions on the vector fields and the stochastic flow which results from these coefficients. Let $U \subset \mathbb{R}^N$ be some open subset of $\mathbb{R}^N$.

$$V_i \in C_b^\infty(U), \text{ and are globally Lipschitz on } \mathbb{R}^N, \quad i = 0, \dots, d. \quad (\text{C1})$$

$$X_t^{(\cdot)} \in C^\infty(U), \text{ and } X_t^x \in \mathbb{D}^\infty, \forall x \in U, t \geq 0. \quad (\text{C2})$$

The classical results on differentiability of solutions of SDEs require $V_i \in C_b^\infty(\mathbb{R}^N)$. Indeed, this is one potential limitation of the work in this chapter. One may be able to prove diffeomorphic prop-

erties for degenerate SDEs on a case-by-case basis, but one certainly lacks the formal machinery to deal in generalities.

To begin we state the main condition and theorem of the chapter:

Definition 4.1 (LFG Condition) Let $\tilde{U}$ be an open subset of $\mathbb{R}^N$. The vector fields $\{V_0, \dots, V_d\} \subset C_b^\infty(\tilde{U})$ are said to satisfy the LFG condition on $\tilde{U}$ if there exists $m \in \mathbb{N}$ such that for some coefficients $\varphi_{\alpha,\beta} \in C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$, we may write the following relationship on $\tilde{U}$, for all $\alpha \in \mathcal{A}$ such that for $\|\alpha\| \in \{m+1, m+2\}$:

$$V_{[\alpha]} = \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha,\beta} V_{[\beta]}. \quad (4.1)$$

This is essentially the satisfaction of the UFG condition, not uniformly over $\mathbb{R}^N$ but locally on some set $\tilde{U}$. Indeed, the definition allows the specification $\tilde{U} = \mathbb{R}^N$ and so this may be thought of as a generalisation of the UFG condition. Note also that we demand the coefficients, $\varphi_{\alpha,\beta}$, to be defined globally, but to only satisfy (4.1) on $\tilde{U}$. Indeed, it may be the case that coefficients - which are chosen in line with (4.1) - are identified but are not defined outside of $\tilde{U}$. Indeed, $V_{[\alpha]}$ may not be well-defined outside of $\tilde{U}$. In this case we can often extend them such that they satisfy $\varphi_{\alpha,\beta} \in C_b^\infty(\mathbb{R}^N)$. Conditions for which this may be done are given by *Whitney's extension theorem* in [41].

We make a quick note on the semantics of the LFG condition: as the 'UFG' in the UFG condition stands for '*uniformly finitely generated*', rather than running roughshod over this precedence by overloading the definition - and calling it, say, a 'local UFG condition' - we refer to it as 'the LFG condition' (for '*locally finitely generated*').

Theorem 4.2 Assume that $f \in C_b^\infty$. Let X_t^x be the solution of equation (1.9), whose coefficients satisfy **(C1)**. Assume further that **(C2)** holds, and that the LFG condition holds on $\tilde{U} \subseteq U$. Then for all $x \in \tilde{U}$

$$V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t f(x) = \mathbb{E}[f(X_t^x) \Phi_{\alpha_1, \dots, \alpha_N}(t, x)],$$

for some integration by parts factor $\Phi_{\alpha_1, \dots, \alpha_N}(t, x)$. Moreover, defining $\tilde{U}^\epsilon := \{x \in \tilde{U} : d(x, \partial\tilde{U}) \geq \epsilon\}$ the following gradient bounds hold:

$$\sup_{y \in \tilde{U}^\epsilon} |V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t f(y)| \leq C_\epsilon t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_N\|}{2}} \|f\|_\infty, \quad (4.2)$$

for each $\epsilon > 0$. Consequently, for each uniformly continuous function f , the diffusion semigroup is smooth along the vector fields of the Lie algebra on $\tilde{U}$.

Remark 4.3 In this chapter, we are assuming $V_i \in C_b^\infty(U)$. For many applications of interest, one would expect the LFG condition to be satisfied on some true subset of the interior of U , but for the vector fields and the LFG representation to degenerate at the boundary of U . It is for this

reason that we allow the LFG condition to be satisfied on some subset of U . If the LFG condition is satisfied on the entire set for which the vector fields are smooth, for example when $U = \mathbb{R}^N$, then the definition allows for that, but it also allows one to put a positive distance between the set on which the LFG condition is satisfied, and the boundary, which may preclude the satisfaction of the LFG on U due to degenerate behaviour. Moreover, the introduction of $\tilde{U}^\epsilon$ in the theorem is a result of some mollification techniques we shall need to apply to points near the ‘edge’ of $\tilde{U}$.

4.1. Adapting the integration by parts formula for local properties

We recall (1.21) from the first chapter, and remember the importance of using the UFG condition to derive the integration by parts formula. We noted that one could express the Malliavin derivative of $f(X_t^x)$ as follows:

$$Df(X_t^x) = \nabla(f \circ X_t)(x) \left(\int_0^{t \wedge \cdot} (J_u^x)^{-1} V_i(X_u^x) du \right)_{i=1, \dots, d}. \quad (4.3)$$

This relationship still holds in this situation, for all $x \in U$, as a result of assuming **(C2)**. Then, provided the UFG condition was satisfied, we were able to show

$$(J_u^x)^{-1} V_i(X_u^x) = (A(u, x) V(x))_i = \sum_{\beta \in \mathcal{A}(m)} a_{i, \beta}(u, x) V_{[\beta]}(x). \quad (4.4)$$

Although we cannot hope to replicate this relationship globally, it is easy to see that this relationship still holds for each $x \in \tilde{U}$, and for all SDE paths which have **not** left $\tilde{U}$. We assume in what follows that $x \in \tilde{U}^\epsilon$. Define $T^x := \inf\{t \geq 0 ; X_t^x \notin \tilde{U}\}$. Indeed, by assuming the satisfaction of the LFG condition and the same arguments that led to Lemma 2.2 in the first chapter, and it can be shown that:

$$(J_u^x)^{-1} V_i(X_u^x) = (A(u, x) V(x))_i = \sum_{\beta \in \mathcal{A}(m)} a_{i, \beta}(u, x) V_{[\beta]}(x), \quad (4.5)$$

holds $\mathbb{P}$ -a.s. on $\{u < T^x\}$. This simple observation is the basis of the analysis. Indeed, this enables us to write the following. Restricted to $\{s < T^x\}$:

$$\begin{aligned} D_s f(X_t^x) &= \nabla(f \circ X_t)(x) \left(\int_0^s (J_u^x)^{-1} V_i(X_u^x) du \right)_{i=1, \dots, d} \\ &= \sum_{\beta \in \mathcal{A}(m)} V_{[\beta]}(f \circ X_t)(x) \left[a_{\cdot, \beta}(s, x) \right]_{i=1, \dots, d}. \end{aligned} \quad (4.6)$$

To proceed from this point requires a clever trick. We need to introduce an extra factor into the equation, which vanishes on $\{s \geq T^x\}$, and which belongs to the space $\mathbb{D}^\infty$. Based on the observation that, for $x \in \tilde{U}^\epsilon$, there holds $\{s \geq T^x\} \subset \{\sup_{v \in [0, s]} |X_v^x - x| \geq \epsilon\}$ instead of seeking a factor which vanishes on the former set, we seek a one which vanishes on the latter¹. A truncation

¹Thus, it also vanishes on the former.

of the maximum process $M_s^x := \sup_{v \in [0, s]} |X_v^x - x|$ - although a seemingly prudent choice - is inadequate for this task, as the maximum process is merely once differentiable with respect to the Malliavin derivative. As has been constantly shown in the previous chapters, the elements of the integration by parts formula require stronger differentiability. Owing to the strong modulus of continuity bounds derived in Garsia, Rodemich and Rumsey [10], and the inspiration provided by Nualart through his application of a similar integration by parts factor to deduce smoothness of the density of the maximum of the Brownian sheet in [33], there is a suitable candidate for this factor. It is based on the observation that the paths of the SDE have a certain modulus of continuity. Define, for $\gamma \in (0, 1/2)$, $p > 2$ such that $\frac{1}{2p} < \gamma < \frac{1}{2} - \frac{1}{2p}$, the increasing process Y defined, for $s \leq t$, as

$$Y_s^x := \int_{[0, s]^2} \frac{|X_u^x - X_v^x|^{2p}}{|u - v|^{1+2p\gamma}} du dv.$$

Lemma 4.4 *There holds $Y_s^x \in \mathbb{D}^\infty(H)$ for all $x \in U$ and $s \geq 0$.*

Proof : The proof for Brownian motion is proved in Airault and Malliavin [1]. See the appendix for a general SDE solution. ■

By restricting the growth of Y we may restrict the growth of $\sup_{u \in [0, s]} |X_u^x - x|$. More precisely, we claim that for $r > 0$ there exists R_r such that

$$Y_s^x \leq R_r \quad \Rightarrow \quad \sup_{u \in [0, s]} |X_u^x - x| \leq r.$$

Indeed, by applying the lemma of Garsia, Rodemich and Rumsey [10] to Y (**see appendix for statement and discussion**), one may deduce that:

$$|X_u^x - X_v^x|^{2p} \leq C_{p, \gamma} Y_s^x |u - v|^{2p\gamma-1},$$

for all $u, v \in [0, s]$, and some constant $C_{p, \gamma}$. Hence, if $Y_s^x \leq R_r$, then

$$\begin{aligned} |X_u^x - X_v^x|^{2p} &\leq C_{p, \gamma} R_r |u - v|^{2p\gamma-1} \\ \Rightarrow |X_u^x - x| &\leq C_{p, \gamma}^{\frac{1}{2p}} R_r^{\frac{1}{2p}} u^{\gamma - \frac{1}{2p}} \\ \Rightarrow \sup_{u \in [0, s]} |X_u^x - x| &\leq C_{p, \gamma}^{\frac{1}{2p}} R_r^{\frac{1}{2p}} t^{\gamma - \frac{1}{2p}}. \end{aligned}$$

The claim can now be verified by choosing R_r such that $C_{p, \gamma}^{\frac{1}{2p}} R_r^{\frac{1}{2p}} t^{\gamma - \frac{1}{2p}} \leq r$. Define $r := \epsilon$. This is the minimum distance a path which originates in $\tilde{U}^\epsilon$ has to travel in order to leave $\tilde{U}$. We now proceed by truncating the function $s \mapsto Y_s^x$ on $[0, R_r]$. This will insure SDE paths which originate in $\tilde{U}^\epsilon$ will be truncated if they come close to the boundary of $\tilde{U}$. This permits use of the LFG condition, as has been discussed. Due to the cautious approach of the truncation, some paths - which originate in $\tilde{U}^\epsilon$ - may be truncated even if they do not leave $\tilde{U}$. This overcompensation shall

not pose a large problem. In particular, choose $\varphi \in C_b^\infty$ such that

1. $\varphi \equiv 1$ on $(-\infty, \frac{R_r}{2})$.
2. $\varphi \equiv 0$ on $(R_r, +\infty)$.
3. $\varphi(\mathbb{R}) \subset [0, 1]$.

We are now ready to further develop equation (4.6). As was done in the first chapter, we use this single equation to produce a square linear system of equations. The only difference this time is that we take the H -inner-product with $\tilde{k}_\alpha(t, x) := \left(\int_0^{t \wedge \cdot} a_{i,\alpha}(s, x) \varphi(Y_s^x) ds \right)_{i=1, \dots, d}$. This step removes the need to phrase equality in terms of occurring on the set $\{s < T^x\}$. Indeed, on $\{s \geq T^x\}$ it follows that $Y_s^x \geq R_r$ and hence $\varphi(Y_s^x) = 0$. So although equation (4.6) breaks down at this point, the fact that $\varphi(Y_s^x)$ vanishes here means that equality remains. Indeed, by taking $\alpha_1, \dots, \alpha_{N_m}$ to be an enumeration of the multi-indices of $\mathcal{A}(m)$, there holds $\mathbb{P}$ -a.s. *everywhere*

$$\begin{aligned} \left\langle Df(X_t^x), \tilde{k}_{\alpha_1}(t, x) \right\rangle_H &= \sum_{\beta \in \mathcal{A}(m)} V_{[\beta]}(f \circ X_t)(x) \left\langle k_\beta(t, x), \tilde{k}_{\alpha_1}(t, x) \right\rangle_H, \\ &\vdots \\ \left\langle Df(X_t^x), \tilde{k}_{\alpha_n}(t, x) \right\rangle_H &= \sum_{\beta \in \mathcal{A}(m)} V_{[\beta]}(f \circ X_t)(x) \left\langle k_\beta(t, x), \tilde{k}_{\alpha_n}(t, x) \right\rangle_H, \\ &\vdots \\ \left\langle Df(X_t^x), \tilde{k}_{\alpha_{N_m}}(t, x) \right\rangle_H &= \sum_{\beta \in \mathcal{A}(m)} V_{[\beta]}(f \circ X_t)(x) \left\langle k_\beta(t, x), \tilde{k}_{\alpha_{N_m}}(t, x) \right\rangle_H, \end{aligned}$$

and defining, for $\alpha \in \mathcal{A}(m)$, for notational simplicity:

$$\begin{aligned} D^{(\alpha)} f(X_t^x) &:= \left\langle Df(X_t^x), \tilde{k}_\alpha(t, x) \right\rangle_H \\ \tilde{M}_{\alpha, \beta}(t, x) &:= t^{-\frac{\|\alpha\| + \|\beta\|}{2}} \left\langle k_\alpha(t, x), \tilde{k}_\beta(t, x) \right\rangle_H \\ &= t^{-\frac{\|\alpha\| + \|\beta\|}{2}} \sum_{i=1}^d \int_0^t a_{i,\alpha}(s, x) a_{i,\beta}(s, x) \varphi(Y_s^x) ds, \end{aligned}$$

so that

$$D^{(\alpha)} f(X_t^x) = \sum_{\beta \in \mathcal{A}(m)} t^{\frac{\|\alpha\| + \|\beta\|}{2}} \tilde{M}_{\alpha, \beta}(t, x) V_{[\beta]}(f \circ X_t)(x).$$

Hence we have successfully circumvented the problems posed by non-uniformity of the UFG condition.

This was achieved by introducing an extra integration by parts factor, in the form of $\varphi(Y_s^x)$.

As in the first chapter, we must invert the matrix $\tilde{M}(t, x)$ to obtain:

$$V_{[\alpha]}(f \circ X_t)(x) = t^{-\frac{\|\alpha\|}{2}} \sum_{\beta \in \mathcal{A}(m)} t^{-\frac{\|\alpha\|}{2}} \tilde{M}_{\alpha, \beta}^{-1}(t, x) D^{(\beta)} f(X_t^x).$$

4.2. Robustness of the invertibility proof

We now attempt to show that the Malliavin covariance matrix is invertible, and in particular, that the inverse of its determinant is in L^p for all $p \geq 1$. As we shall see, the proof technique used by Kusuoka in [22] is very robust. It allows us to retain sharp asymptotic rates, and also allows the inclusion of a (potentially) quickly triggered stopping time, all without fuss. As was demonstrated in Chapter 1, it suffices to prove that for each $p \geq 1$, there exists $C = C(p) > 0$ such that

$$\mathbb{P} \left(\inf_{|\xi|=1} (\xi, \tilde{M}(t, x)\xi) \leq \epsilon_0 \right) \leq C\epsilon_0^p,$$

for all $\epsilon_0 > 0$, $t \in (0, 1]$. The upper bound is for all $x \in \tilde{U}^\epsilon$, and $C = C(n)$. Recalling (2.13), we see that for $y \geq 1$:

$$\begin{aligned} (\xi, \tilde{M}(t, x)\xi) &= \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha \xi_\beta \tilde{M}_{\alpha, \beta}(t, x) \\ &= \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha \xi_\beta t^{-(\frac{\|\alpha\|}{2} + \frac{\|\beta\|}{2})} \left\langle k_\alpha(t, x), \tilde{k}_\beta(t, x) \right\rangle_H \\ &= \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha t^{-\frac{\|\alpha\|}{2}} \int_0^{t \wedge \cdot} (a_{\cdot, \alpha}^0 + r_{\cdot, \alpha})(u, x) \varphi(Y_u^x) du \right\|_H^2 \\ &\geq \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha t^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge S_{R_r/2}^x} (a_{\cdot, \alpha}^0 + r_{\cdot, \alpha})(u, x) du \right\|_H^2, \end{aligned} \quad (4.7)$$

where $S_R^x := \inf\{t \geq 0 ; Y_s^x \geq R\}$. We have noted in the above that for $s < S_R^x$ there holds $\varphi(Y_s^x) = 1$. We now note that:

$$\begin{aligned} \mathbb{P} \left(\inf_{|\xi|=1} (\xi, \tilde{M}(t, x)) \leq \epsilon_0 \right) &\leq \mathbb{P} \left(\inf_{|\xi|=1} (\xi, \tilde{M}(t, x)\xi) \leq \epsilon_0, S_{\frac{R_r}{2}}^x > \frac{t}{y} \right) \\ &\quad + \mathbb{P} \left(S_{\frac{R_r}{2}}^x \leq \frac{t}{y} \right). \end{aligned} \quad (4.8)$$

First consider the term $\mathbb{P} \left(S_{\frac{R_r}{2}}^x \leq \frac{t}{y} \right) = \mathbb{P} \left(Y_{t/y}^x \geq \frac{R_r}{2} \right)$. For this term we use the Markov inequality to deduce:

$$\begin{aligned} \mathbb{P} \left(Y_{t/y}^x \geq \frac{R_r}{2} \right) &\leq \left(\frac{2}{R_r} \right)^p \mathbb{E} |Y_{t/y}^x|^p \\ &\leq \left(\frac{2}{R_r} \right)^p \sup_{x \in B_r} \mathbb{E} |Y_{t/y}^x|^p \\ &\leq C(n, p) \left(\frac{t}{y} \right)^p. \end{aligned} \quad (4.9)$$

The last inequality is proved in the appendix. We return to this later. Now consider the other term on the RHS of (4.8). Ideally, we would like to use the bound (4.7) to deduce that:

$$\begin{aligned} & \mathbb{P} \left(\inf_{|\xi|=1} (\xi, \tilde{M}(t, x) \xi) \leq \epsilon_0, S_{\frac{R_r}{2}}^x > \frac{t}{y} \right) \\ & \leq \mathbb{P} \left(\inf_{|\xi|=1} \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha t^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} (a_{\cdot, \alpha}^0 + r_{\cdot, \alpha})(u, x) du \right\|_H^2 \leq \epsilon_0 \right). \end{aligned}$$

This step is a valid one, but it deserves closer attention. In particular, the stopping time $S_{R_r/2}^x$ has been dropped from consideration, even though it serves an important purpose. It stops the consideration of $a_{\cdot, \alpha}(t, x)(s, \omega)$ for $(s, \omega) \notin \{(s, \omega) ; s < S_{R_r/2}^x(\omega)\}$. Indeed, by observing:

$$a_{\cdot, \alpha}(t, x) = a_{\cdot, \alpha}^0(t, x) + r_{\cdot, \alpha}(t, x),$$

one will see, from its definition (cf. 2.9) that the remainder term $r_{\cdot, \alpha}(t, x)$ is defined only because we required that the coefficients $\varphi_{\alpha, \beta}$ be defined globally. The reason care was taken to introduce the truncation function is to precisely avoid such (s, ω) . In fact, due to the mechanics of the proof, it is **irrelevant** how the remainder term, and hence the coefficients are defined at these points. Moreover, the fact that $a_{\cdot, \alpha}(t, x)$ is globally well-defined permits us to use the integrability result from Chapter 1 - namely Lemma 2.10.

As was noted in Chapter 1, we now remark that because $y \geq 1$ there holds:

$$\begin{aligned} & \inf_{|\xi|=1} \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha t^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} (a_{\cdot, \alpha}^0 + r_{\cdot, \alpha})(u, x) du \right\|_H^2 \\ & \geq \inf_{|\xi|=1} \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha \left[\frac{t}{y} \right]^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} (a_{\cdot, \alpha}^0 + r_{\cdot, \alpha})(u, x) du \right\|_H^2. \end{aligned}$$

At this stage, we may use the calculations in the first chapter to complete the argument. To do this we must use a variant of Lemma 2.10.

Lemma 4.5 *There holds, for all $p \in [1, \infty)$,*

$$\sup_{\substack{x \in \tilde{U}^\epsilon \\ t \in (0, 1]}} \mathbb{E} \left(\int_0^t \sum_{\alpha \in \mathcal{A}(m)} \sum_{i=1}^d t^{-\|\alpha\|-1} r_{i, \alpha}(u, x)^2 du \right)^p < \infty.$$

Proof : Follows immediately from Lemma 2.10. ■

Then by an analogous proof to the first chapter, we may deduce that

$$\begin{aligned} \mathbb{P} \left(\inf_{|\xi|=1} \left\| \sum_{\alpha \in \mathcal{A}(m)} \xi_\alpha \left[\frac{t}{y} \right]^{-\frac{\|\alpha\|}{2}} \int_0^{t/y \wedge \cdot} (a_{\cdot, \alpha}^0 + r_{\cdot, \alpha})(u, x) du \right\|_H^2 \leq \epsilon_0 \right) \\ \leq C_{m,p} (2\epsilon_0)^p + \tilde{C}_{m,p} \left(\frac{1}{\epsilon_0 y} \right)^p, \end{aligned} \quad (4.10)$$

where the upper bound holds for all $t \in (0, 1]$ and $x \in \tilde{U}$. Combining this upper bound with (A.18), we get that:

$$\begin{aligned} \mathbb{P} \left(\inf_{|\xi|=1} (\xi, \tilde{M}(t, x)) \leq \epsilon_0 \right) &\leq C_{m,p} (2\epsilon_0)^p + \tilde{C}_{m,p} \left(\frac{1}{\epsilon_0 y} \right)^p + \tilde{C}_{n,p} \left(\frac{1}{y} \right)^p \\ &\leq \hat{C}_{n,p,m} \epsilon_0^p, \end{aligned}$$

for all $t \in (0, 1]$ and $x \in \tilde{U}^\epsilon$, where we have chosen $y = \epsilon_0^{-2}$. This completes the proof.

4.3. Several local integration by parts formulae

The reader should, provided the first chapters have been digested, be familiar with the need for supplementary regularity results to be proved, to allow the iterative application of the integration by parts formula. In-keeping with the work of the first chapter, we shall refer to processes which appear in the integration by parts formulae as ‘local Kusuoka-Stroock processes’. This term was adopted in the previous chapter, and although the two definitions are similar, they do not directly coincide. We briefly apologise for any confusion which may result, but pledge to make clear the differences between the various notions of ‘Kusuoka-Stroock process’ shortly.

Definition 4.6 (Local Kusuoka-Stroock processes) *Let E be a separable Hilbert space and let $\rho \in \mathbb{R}$, $V \subset \mathbb{R}^N$ open. Denote by $\tilde{\mathcal{K}}_\rho^{loc}(V, E)$ the set of functions: $f : (0, 1] \times V \rightarrow \mathbb{D}^\infty(E)$ satisfying the following:*

1. *$f(t, \cdot)$ is smooth on V and $\frac{\partial^\alpha f}{\partial x^\alpha}(\cdot, \cdot)$ is continuous in $(t, x) \in (0, 1] \times V$ a.s. for any multi-index α .*
2. *$\sup_{t \in (0, 1], x \in V} t^{-\rho/2} \left\| \frac{\partial^\alpha f}{\partial x^\alpha} \right\|_{k,p,E} < \infty$, for all $k \in \mathbb{N}$, $p \in [1, \infty)$.*

Define $\tilde{\mathcal{K}}_\rho^{loc}(V) := \tilde{\mathcal{K}}_\rho^{loc}(V, \mathbb{R})$

We have been at least careful to introduce different notation for the Kusuoka-Stroock processes, which form an important part of the integration by parts formulae in Chapters 2, 3 and 4. This care did not however extend to the nomenclature in final two of the three chapters. In spite of this, it is thought that they are sufficiently similar to warrant this decision.

Lemma 4.7 (Relationships between the Kusuoka-Stroock processes) *The following hold for fixed $r \in \mathbb{R}$:*

1. $\bigcap_{k \in \mathbb{N}} \mathcal{K}_r^{loc}(E, k) \supset \mathcal{K}_r(E)$.
2. For $V \subsetneq \mathbb{R}^N$, $\tilde{\mathcal{K}}_r^{loc}(V, E) \supset \bigcap_{k \in \mathbb{N}} \mathcal{K}_r^{loc}(E, k)$
3. $\tilde{\mathcal{K}}_r^{loc}(\mathbb{R}^N, E) = \mathcal{K}_r(E)$

Lemma 4.8 (Properties of local Kusuoka-Stroock processes) *The following hold for each $n \in \mathbb{N}$:*

1. Suppose $f \in \tilde{\mathcal{K}}_\rho^{loc}(V, E)$, where $\rho \geq 0$. Then, for $i = 1, \dots, d$,

$$\int_0^\cdot f(s, x) dB_s^i \in \tilde{\mathcal{K}}_{\rho+1}^{loc}(V, E) \quad \text{and} \quad \int_0^\cdot f(s, x) ds \in \tilde{\mathcal{K}}_{\rho+2}^{loc}(V, E).$$

2. $a_{\alpha, \beta}, b_{\alpha, \beta} \in \tilde{\mathcal{K}}_{(\|\beta\| - \|\alpha\|) \vee 0}^{loc}(\tilde{U}^\epsilon)$ where $\alpha, \beta \in \mathcal{A}(m)$.
3. $Y, \varphi(Y) \in \tilde{\mathcal{K}}_0^{loc}(\tilde{U}^\epsilon)$.
4. $\tilde{k}_\alpha \in \tilde{\mathcal{K}}_{\|\alpha\|}^{loc}(\tilde{U}^\epsilon, H)$, where $\alpha \in \mathcal{A}(m)$.
5. $D^{(\alpha)}u := \langle Du(t, x), k_\alpha \rangle_H \in \tilde{\mathcal{K}}_{\rho + \|\alpha\|}^{loc}(V \cap \tilde{U}^\epsilon)$ where $u \in \tilde{\mathcal{K}}_\rho^{loc}(V)$ and $\alpha \in \mathcal{A}(m)$.
6. If $M^{-1}(t, x)$ is the inverse matrix of $M(t, x)$, then $M_{\alpha, \beta}^{-1} \in \tilde{\mathcal{K}}_0^{loc}(\tilde{U}^\epsilon)$, $\alpha, \beta \in \mathcal{A}(m)$,
7. If $f_i \in \tilde{\mathcal{K}}_{\rho_i}^{loc}(V_i)$ for $i = 1, \dots, N$, then

$$\prod_{i=1}^N f_i \in \tilde{\mathcal{K}}_{\rho_1 + \dots + \rho_N}^{loc}(\cap_{i=1}^N V_i) \quad \text{and} \quad \sum_{i=1}^N f_i \in \tilde{\mathcal{K}}_{\min(\rho_1, \dots, \rho_N)}^{loc}(\cap_{i=1}^N V_i).$$

Proof : The proof of all parts of this lemma has already been covered in previous versions of this lemma. The proof of part 3. is handled in the appendix. ■

In the remainder of this section we derive numerous integration by parts formulae, in a similar fashion to the work of the first chapter. Note that we are constantly assuming that the LFG condition holds in $\tilde{U}$. The dependence of the integration by parts factors on the minimum distance of $x \in \tilde{U}^\epsilon$ from the boundary of $\tilde{U}$, constructed to be $\frac{1}{n}$, is emphasised. The proofs are very similar to previous ones and are thus not included.

Theorem 4.9 (Integration by parts formula I) For all $\Phi \in \tilde{\mathcal{K}}_\rho^{loc}(\tilde{U}^\epsilon)$, $\rho \in \mathbb{R}$ and $\alpha \in \mathcal{A}(m)$ there exists $\Phi_\alpha \in \tilde{\mathcal{K}}_\rho^{loc}(\tilde{U}^\epsilon)$ such that

$$\mathbb{E}[\Phi(t, x) V_{[\alpha]}(f \circ X_t)(x)] = t^{-\|\alpha\|/2} \mathbb{E}[\Phi_\alpha^n(t, x) f(X_t^x)], \quad (4.11)$$

for any $f \in C_b^\infty(\mathbb{R}^N; \mathbb{R})$, $t > 0$, $x \in \tilde{U}^\epsilon$.

From this it is possible to prove other related integration by parts formulae; in particular (cf. Corollary 4.10). One should again take note of the non-trivial role the LFG condition plays in the other derivations.

Corollary 4.10 (Integration by Parts formula II) Under the assumptions of Theorem 4.9, and assuming $\Phi(t, x) \equiv 0$ on $\{T^x \leq t\}$, the following holds:

$$\mathbb{E}[\Phi(t, x) (V_{[\alpha]} f)(X_t^x)] = t^{-\|\alpha\|/2} \mathbb{E}[\tilde{\Phi}_\alpha^n(t, x) f(X_t^x)],$$

where $\tilde{\Phi}_\alpha \in \tilde{\mathcal{K}}_\rho^{loc}(\tilde{U}^\epsilon)$, for any $f \in C_b^\infty(\mathbb{R}^N; \mathbb{R})$, $t > 0$, $x \in \tilde{U}^\epsilon$.

Corollary 4.11 (Integration by Parts formula III) Under the assumptions of Theorem 2.20 the following holds:

$$V_{[\alpha]} \mathbb{E}[\Phi(t, x) f(X_t^x)] = t^{-\|\alpha\|/2} \mathbb{E}[\tilde{\tilde{\Phi}}_\alpha^n(t, x) f(X_t^x)],$$

where $\tilde{\tilde{\Phi}}_\alpha \in \tilde{\mathcal{K}}_\rho^{loc}(\tilde{U}^\epsilon)$, for any $f \in C_b^\infty(\mathbb{R}^N; \mathbb{R})$, $t > 0$, $x \in \tilde{U}^\epsilon$.

Proof of Corollary 4.10: : The first observation is the following relationship:

$$\begin{aligned} (V_{[\alpha]} f)(X_t^x) &= \nabla f(X_t^x) V_{[\alpha]}(X_t^x) \\ &= (J_t^x)^{-T} \nabla(f \circ X_t)(x) V_{[\alpha]}(X_t^x) \\ &= \nabla(f \circ X_t)(x) (J_t^x)^{-1} V_{[\alpha]}(X_t^x), \end{aligned}$$

where $(J_t^x)^{-T} := ((J_t^x)^{-1})^T$. At this point refer back to the closed linear system of equations, which induced the expression, on $\{T^x > t\}$:

$$(J_t^x)^{-1} V_{[\alpha]}(X_t^x) = \sum_{\beta \in \mathcal{A}(m)} a_{\alpha, \beta}(t, x) V_{[\beta]}(x).$$

Again, the central position of the LFG condition is emphasised. Note that this implies, on $\{T^x > t\}$

$$\begin{aligned} \nabla(f \circ X_t)(x) (J_t^x)^{-1} V_{[\alpha]}(X_t^x) &= \sum_{\beta \in \mathcal{A}(m)} a_{\alpha, \beta}(t, x) \nabla(f \circ X_t)(x) V_{[\beta]}(x) \\ &= \sum_{\beta \in \mathcal{A}(m)} a_{\alpha, \beta}(t, x) V_{[\beta]}(f \circ X_t)(x). \end{aligned}$$

As $\Phi(t, x) \equiv 0$ on $\{T^x \leq t\}$, this is equivalent to writing

$$\Phi(t, x) \nabla(f \circ X_t)(x) (J_t^x)^{-1} V_{[\alpha]}(X_t^x) = \sum_{\beta \in \mathcal{A}(m)} \Phi(t, x) a_{\alpha, \beta}(t, x) V_{[\beta]}(f \circ X_t)(x),$$

$\mathbb{P}$ -a.s. Now note that, from Lemma 4.8, $a_{\alpha, \beta} \in \tilde{\mathcal{K}}_{(\|\beta\| - \|\alpha\|) \vee 0}^{\text{loc}}(\tilde{U}^\epsilon)$. So it has been shown that:

$$\mathbb{E} [\Phi(t, x) V_{[\alpha]} f(X_t^x)] = \sum_{\beta \in \mathcal{A}(m)} \mathbb{E} [\Phi(t, x) a_{\alpha, \beta}(t, x) V_{[\beta]}(f \circ X_t)(x)].$$

The integration by parts formula (4.11) can then be applied N_m times, after noting that the product $\Phi a_{\alpha, \beta} \in \tilde{\mathcal{K}}_{\rho+[(\|\beta\| - \|\alpha\|) \vee 0]}^{\text{loc}}(\tilde{U}^\epsilon)$. And so,

$$\begin{aligned} \mathbb{E} [\Phi(t, x) V_{[\alpha]} f(X_t^x)] &= \sum_{\beta \in \mathcal{A}(m)} t^{-\frac{\|\beta\|}{2}} \mathbb{E} [\Psi_\beta(t, x) f(X_t^x)] \\ &= \sum_{\beta \in \mathcal{A}(m)} t^{-\frac{\|\beta\|}{2}} t^{-\frac{\|\alpha\| - \|\beta\|}{2}} \mathbb{E} \left[t^{\frac{\|\alpha\| - \|\beta\|}{2}} \Psi_\beta(t, x) f(X_t^x) \right] \\ &= t^{-\frac{\|\alpha\|}{2}} \mathbb{E} [\tilde{\Phi}_\alpha^n(t, x) f(X_t^x)], \end{aligned}$$

where $\tilde{\Phi}_\alpha^n = \sum_{\beta \in \mathcal{A}(m)} t^{\frac{\|\alpha\| - \|\beta\|}{2}} \Psi_\beta \in \tilde{\mathcal{K}}_\rho^{\text{loc}}(\tilde{U}^\epsilon)$.

■

4.4. The deduction of local gradient bounds

In this section the integration by parts formulae of the previous section are employed to obtain local gradient bounds. We consider semigroups with a potential term. That, is define $(P_t^c)_{t \geq 0}$ by

$$(P_t^c f)(x) = \mathbb{E} \left[\exp \left\{ \int_0^t c(X_s^x) ds \right\} f(X_t^x) \right].$$

This is the solution of a parabolic PDE which has been perturbed by a potential, c . We assume $c \in C_b^\infty(U)$. Then one can deduce the following:

Theorem 4.12 Assume $\alpha_1, \dots, \alpha_{n+m} \in \mathcal{A}(m)$. Then there is a constant C_n such that:

$$\|V_{[\alpha_1]} \dots V_{[\alpha_n]} P_t^c V_{[\alpha_{n+1}]} \dots V_{[\alpha_{n+m}]} f\|_{L^p(\tilde{U}^\epsilon)} \leq C_n t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n+m}\|}{2}} \|f\|_{L^p},$$

for any $f \in C_0^\infty(\tilde{U}^\epsilon)$, $t \in (0, 1]$ and $p \in [1, \infty]$.

Proof : As in the first chapter, the proof is based on an interpolation argument. Please consult Kusuoka [22, p268-270]. ■

RECOVERING A LOCAL VERSION OF HÖRMANDER'S THEOREM

In a similar fashion to the first section, we may consider the situation where Hörmander's criteria for hypoellipticity holds. In particular, we demonstrate local smoothness of the diffusion semigroup on $\tilde{U}$ under various assumptions. The integration by parts formulae are tailor-made to fit this situation.

Theorem 4.13 *A local version of Hörmander's theorem holds:*

1. Assume that the following holds for all $x_0 \in \tilde{U}$:

$$\text{Span}\{V_{[\alpha]}(x_0) : \alpha \in \mathcal{A}(m)\} = \mathbb{R}^N. \quad (4.12)$$

Then we have the following gradient bounds for iterated partial derivatives of the semigroup:

$$\sup_{y \in \tilde{U}^\epsilon} |\partial_{i_1} \dots \partial_{i_M} P_t f(y)| \leq C_n \frac{1}{t^{Mr/2}} \|f\|_\infty \quad (4.13)$$

$$\sup_{y \in \tilde{U}^\epsilon} |\partial_{i_1} \dots \partial_{i_M} P_t f(y)| \leq C_n \frac{t^{1/2}}{t^{Mr/2}} \|\nabla f\|_\infty, \quad (4.14)$$

for any $i_1, \dots, i_N \in \{1, \dots, N\}$, $M \in \mathbb{N}$ and $t \in (0, 1]$. Moreover, the diffusion semigroup is smooth in $\tilde{U}$, for any uniformly continuous function f .

2. Assume that (4.12) holds for some $x_0 \in \mathbb{R}^N$. Then, for some r_{x_0} , we have the following gradient bounds for iterated partial derivatives of the semigroup:

$$\begin{aligned} \sup_{y \in B_{r_{x_0}}(x_0)} |\partial_{i_1} \dots \partial_{i_K} P_t f| &\leq C \frac{1}{t^{Km/2}} \|f\|_\infty, \\ \sup_{y \in B_{r_{x_0}}(x_0)} |\partial_{i_1} \dots \partial_{i_K} P_t f| &\leq C \frac{t^{1/2}}{t^{Km/2}} \|\nabla f\|_\infty. \end{aligned}$$

Then the diffusion semigroup is a smooth in some ball of radius r_{x_0} with centre x_0 , for any uniformly continuous function f .

Proof : 1. Note that in this case we may restate (4.13) as: $\exists \epsilon > 0$ s.t.:

$$\sum_{\alpha \in \mathcal{A}(m)} |\xi^T V_{[\alpha]}(x)|^2 \geq \epsilon |\xi|^2, \quad (4.15)$$

$\forall \xi \in \mathbb{R}^N$ and $x \in \tilde{U}^\epsilon$, or equivalently: the matrix $(VV^T)(x)$ is invertible $\forall x \in \tilde{U}^\epsilon$, where $V(x) := (V_{[\alpha_1]}(x) \mid \dots \mid V_{[\alpha_{N_m}]}(x))$, and the determinant is non-zero even at the boundary of $\tilde{U}^\epsilon$.

To demonstrate smoothness we seek to deduce an integration by parts formula for iterated partial derivatives $\partial_{i_1} \dots \partial_{i_M}$, where $i_j \in \{1, \dots, N\}$ for $j = 1, \dots, M$. We claim that there exist $c_\alpha^i \in C_b^\infty(\tilde{U}^\epsilon)$ such that:

$$e_i = \sum_{\alpha \in \mathcal{A}(m)} c_\alpha^i V_{[\alpha]},$$

$\forall x \in \tilde{U}^\epsilon$. This can be re-written in matrix form as

$$e_i = VC^i,$$

again, where $V(x) := (V_{[\alpha_1]}(x) \mid \dots \mid V_{[\alpha_{N_m}]}(x))$, and $c^i(x) = (C_\alpha^i(x))_{\alpha \in \mathcal{A}(m)}$. But it holds that $(VV^T)(x)$ is invertible $\forall x \in \tilde{U}^\epsilon$, with non-zero determinant even at the boundary. Therefore, we may set

$$c^i = V^T(VV^T)^{-1}e_i.$$

That is, $c_\alpha^i(x) = (V^T(VV^T)^{-1}e_i)_\alpha(x)$. By the inverse function theorem, $c_\alpha^i \in C_b^\infty(\tilde{U}^\epsilon)$. Now observe that

$$\begin{aligned} |\mathbb{E}(\partial_i(f \circ X_t)(x))| &= \left| \sum_{\alpha \in \mathcal{A}(m)} c_\alpha^i(x) \mathbb{E}[V_{[\alpha]}(f \circ X_t)(x)] \right| \\ &\leq \sum_{\alpha \in \mathcal{A}(m)} \sup_{x \in \tilde{U}^\epsilon} |c_\alpha^i(x)| |\mathbb{E}[V_{[\alpha]}(f \circ X_t)(x)]| \\ &\leq K_{i,n} t^{-m/2} \|f\|_\infty. \end{aligned}$$

This upper bound is uniform over all $x \in \tilde{U}^\epsilon$. This may be done for any partial derivative and the procedure can be iterated for any multi-index α , to deduce the gradient bounds:

$$\begin{aligned} \sup_{x \in \tilde{U}^\epsilon} |\partial_{i_1} \dots \partial_{i_K} P_t f| &\leq C_n \frac{1}{t^{Km/2}} \|f\|_\infty \\ \sup_{x \in \tilde{U}^\epsilon} |\partial_{i_1} \dots \partial_{i_K} P_t f| &\leq C_n \frac{t^{1/2}}{t^{Km/2}} \|\nabla f\|_\infty, \end{aligned}$$

for any $i_1, \dots, i_N \in \{1, \dots, N\}$, $M \in \mathbb{N}$ and $t \in (0, 1]$. These gradient bounds may be applied in an analogous way to Proposition 2.31 to deduce smoothness on $\tilde{U}^\epsilon$.

To prove 2., we observe that as $x \mapsto \det A(x)$ is a continuous mapping for matrices A with continuous components. In particular, if

$$\det A(x_0) \neq 0,$$

then it must be that the above holds for all y in some neighbourhood of x_0 . By taking an even smaller neighbourhood of x_0 , we can also deduce that the determinant must also be non-zero on the boundary of the neighbourhood. From this observation, we may apply the result in 1. to complete the proof. ■

SOME EXAMPLES

It was mentioned in the introduction to this chapter that one seeks, almost paradoxically, to find examples with degenerate coefficients (whose coefficients are smooth and bounded only in some restricted subset of $\mathbb{R}^N$), but whose corresponding SDE solutions satisfy some very strong differentiability properties. Such examples, do exist and make interesting applications of this work.

Example 4.14 Consider the following one dimensional SDE

$$\begin{cases} dX_t^x &= dt + 2\sqrt{X_t^x}dB_t, \quad t > 0, \\ X_0^x &= x. \end{cases}$$

Then the equation has the explicit, unique, strong solution:

$$X_t^x = (B_t + \sqrt{x_0})^2$$

Hence it is clear that **(C1)** and **(C2)** both hold on $(0, \infty)$. Note also that the Lie algebra is of full rank on $(0, \infty)$. Thus, we may deduce that the semigroup $P_t f(\cdot)$ is smooth on $(0, \infty)$ for any uniformly continuous function f and that (4.13) and (4.14) hold.

Example 4.15 Consider the following one dimensional SDE

$$\begin{cases} dX_t^x &= -X_t^x(2\log X_t^x + 1)dt + 2X_t^x\sqrt{-\log X_t^x}dB_t, \quad t > 0, \\ X_0^x &= x. \end{cases}$$

Then for each $x_0 \in (0, 1)$ the equation has the explicit, unique, strong solution:

$$X_t^x = \exp \left\{ - \left(B_t + \sqrt{-\log x_0} \right)^2 \right\}$$

In particular, it is clear that **(C1)** and **(C2)** both hold on $(0, 1)$, and for uniformly continuous functions f we may deduce smoothness of the semigroup, $P_t f(\cdot)$, on this set. Moreover, (4.13) and (4.14) hold.

5. The generality of the LFG condition

The analysis in this thesis has rested on the assumption of either the UFG or the LFG condition. The reader has been assured that these conditions are very general. The purpose of this section is to delve further into their generality.

It has already been demonstrated that the UFG condition is more general than the uniform Hörmander condition, and that the LFG condition is more general than a local Hörmander condition, but is it the case, for instance, that the LFG condition holds *a priori* for smooth functions with bounded derivatives of all orders? The answer to this question is negative. Indeed, we shall construct examples where the UFG **and** the LFG conditions fail to hold, owing to degenerative behaviour at a single point. A more interesting question is whether the LFG condition holds *a priori* on compact sets in the case of analytic vector fields. Analytic vector fields have very favourable properties on compact sets, which can be harnessed to deduce behaviour of the Lie algebra. As we shall see, the quest for the satisfaction of the LFG condition can be separated into two distinct questions:

1. *Is the Lie algebra finitely generated?*
2. *If so, can the coefficients of the Lie algebra be chosen to be smooth with bounded derivatives?*

Both of the above questions are of fundamental importance to how the LFG condition is applied. The answer to the first question, as will be demonstrated, is ‘yes’ in the case of analytic coefficients and over compact sets. If the rank of the space generated by the vector fields is constant, then the answer is yes even if the coefficients are merely smooth. In the case of constant rank of the Lie algebra, the answer to the second question will also be ‘yes’.

Before tackling the problem in detail, we mention existing theory in this area. Sussmann [39] provides a significant contribution to the study of Lie algebras of non-full rank. The paper is mostly interested in demonstrating integrability of distributions. To this end, Sussmann discusses a very similar version of the LFG condition, cf (LFT1) and (LFT2). He uses these conditions to prove that the distribution associated to the Lie algebra has integral submanifolds. In particular, he shows that (LFT1) and (LFT2) hold under the assumption of analytic coefficients and (separately) under the assumption of uniform constant rank - the latter in Theorem 8.2.

Whereas Sussmann uses the conditions (LFT1,2) as tools for integrability, we are interested in the similar LFG condition in its own right.

5.1. Some pathologies disproving *a priori*-ness of the LFG condition

We begin this chapter by providing a counter-example to the incorrect assertion: "*The LFG condition is satisfied around all points of the state space, when the vector fields are smooth with bounded derivatives*".

Example 5.1 (Basic Example) Let $N = 1$, $d = 1$. Take $V_0 \equiv 1$,

$$V_1(x) := \begin{cases} \exp\{-1/x\}, & x > 0 \\ 0, & x \leq 0 \end{cases}.$$

We claim that this example does not satisfy the UFG condition. i.e. there is no finite uniform

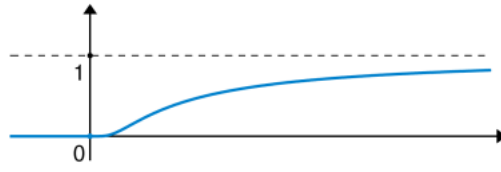


Figure 5.1.: A non-analytic smooth function

generating subset of the Lie algebra which generates the whole algebra. Our first observation is that the derivatives of V_1 can be given by

$$V_1^{(n)}(x) = \begin{cases} \frac{p_n(x)}{x^{2n}} V_1(x) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases},$$

where $p_n(x) = \sum_{i=0}^n a_n x^i$, $a_n \neq 0$, is some polynomial strictly of order n .

It can then be shown that we have the following general representation for the vector fields of the Lie algebra

$$V_{[\alpha]}(x) = \begin{cases} \frac{p_{n_\alpha}(x)}{x^{2n_\alpha}} \exp\{-\tilde{n}_\alpha/x\} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases},$$

where $n_\alpha := |\alpha|$ and $\tilde{n}_\alpha := |\alpha| - (||\alpha|| - |\alpha|)$. Our claim is that the LFG cannot hold in any neighbourhood $U \supset \{0\}$. If we assume it does, then for some fixed $m \in \mathbb{N}$ and all α such that $|\alpha| = m + 1$ there exist $\varphi_{\alpha,\beta} \in C_b^\infty$ such that:

$$V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha,\beta}(x) V_{[\beta]},$$

holds on U . Equivalently, for $x \in U$

$$\frac{p_{m+1}(x)}{x^{2(m+1)}} \exp\{-\tilde{n}_\alpha/x\} = \sum_{\beta \in \mathcal{A}(m)} \frac{\varphi_{\alpha,\beta}(x) p_{|\beta|}(x)}{x^{2|\beta|}} \exp\{-\tilde{n}_\beta/x\}.$$

i.e.

$$p_{m+1+2|\beta|}(x) \exp\{-\tilde{n}_\alpha/x\} = \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha,\beta}(x) p_{|\beta|}(x) x^{2(m+1-|\beta|)} \exp\{-\tilde{n}_\beta/x\}.$$

For example, if we now pick $\alpha = (\overbrace{1, 1, \dots, 1}^{m \text{ times}}, 0)$, so that $\tilde{n}_\alpha = m$ then

$$p_{m+1}(x) = \sum_{\beta \in \mathcal{A}(m)} \varphi_{\alpha,\beta}(x) \tilde{p}_{2(m+1)-|\beta|}(x) \exp\{(\tilde{n}_\alpha - \tilde{n}_\beta)/x\},$$

The only way we can get the RHS to be a polynomial of order $(m+1)$ is to choose either $\varphi_{\alpha,\beta} \equiv 0$ or $\varphi_{\alpha,\beta}(x) = x^{-2(m+1-|\beta|)} \tilde{p}_{m+1-|\beta|} \exp\{(\tilde{n}_\alpha - \tilde{n}_\beta)/x\}$ for some polynomial $\tilde{p}$. Moreover, we need at least one such choice of the latter. Notice further that $\tilde{n}_\alpha - \tilde{n}_\beta \geq m - (m-1) = 1$. This function explodes at $x = 0$, for all $\|\beta\| \leq m$. Hence UFG condition cannot hold, and the LFG condition does not hold in any neighbourhood of zero.

We can also construct examples in arbitrary dimensions based on the above.

Example 5.2 (Extension to multi-dimensional SDEs)

One may extend this to a multi-dimensional vector field: by again considering the function

$$\rho(x) := \begin{cases} \exp\{-1/x\}, & x > 0 \\ 0, & x \leq 0 \end{cases},$$

and setting for some fixed radius, $r > 0$,

$$\begin{aligned} V_0(x) &= (0, \dots, 1, \dots, 0), \\ V_1(x) &= (\varphi_1(x_1), \dots, \varphi_{k-1}(x_{k-1}), \rho(r^2 - |x|^2), \varphi_{k+1}(x_{k+1}), \dots, \varphi_N(x_N)), \end{aligned}$$

where $\varphi_1, \dots, \varphi_N \in C_b^\infty$. Then the UFG condition will not be satisfied due to the behaviour around

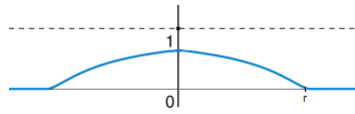


Figure 5.2.: Extension of non-analytic smooth function to general dimensions

$|x| = r$.

One saving grace of this catastrophic realisation is that although ρ is a smooth, bounded function it is not analytic. In particular, it is not equal to its Taylor expansion in any punctured neighbourhood around zero. Notice this was where any choice of LFG coefficient exploded. To see this fact one need only observe that ρ converges to zero faster than any polynomial hence all of its derivatives are zero at $x = 0$. But one of the main problems was that although ρ eventually surpasses polynomial

convergence, it takes an increasing amount of time to do this. Indeed, the derivatives of ρ have spikes around zero which grow in supremum norm **faster** than factorially in derivative order. It is this kind of behaviour which analytic functions avoid. In the next section we move past this pathology and focus on analytic coefficients. In particular, we demonstrate that on any compact set the Lie algebra is generated by a finite number of its elements.

PROPERTIES OF ANALYTIC FUNCTIONS

In what follows we assume that the infinitesimal generator $\mathcal{L}$ of the diffusion semigroup $\{P_t\}_{t \geq 0}$,

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

has the property $V_i^j \in C^\omega(U; \mathbb{R})$ for $0 \leq i \leq d$, $1 \leq j \leq N$, and some $U \subset \mathbb{R}^N$, open. Equivalently, that the coefficients of the stochastic flow

$$\begin{cases} dX_t^x &= \sum_{i=0}^d V_i(X_t^x) \circ dB_t^i, \\ X_0^x &= x, \end{cases}$$

have analytic components. We recall the definition of an analytic function on $\mathbb{R}^N$.

Definition 5.3 (Analytic Function) *The function $f : \mathcal{D} \mapsto \mathbb{R}$ is said to be analytic on the open set $\mathcal{D}$, if it is locally given by a convergent power series. i.e. if we can write for any $x_0 \in \mathcal{D} \subset \mathbb{R}^N$:*

$$f(x) = \sum_{|\alpha| \geq 0} C_\alpha (x - x_0)^\alpha,$$

for any x in some neighbourhood of x_0 . Alternatively, an analytic function is a smooth function $f \in C^\infty(\mathcal{D}; \mathbb{R})$ which is locally given by it's Taylor series, i.e.

$$f(x) = \sum_{|\alpha| \geq 0} \frac{\partial_\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha,$$

where $\alpha \in \mathbb{N}_0^N$ is a multiindex, $\partial_\alpha := \partial_1^{\alpha_1} \dots \partial_N^{\alpha_N}$, with $x^\alpha := x_1^{\alpha_1} \dots x_N^{\alpha_N}$ and $\alpha! := \alpha_1! \dots \alpha_N!$. Note also that $\alpha + \beta := (\alpha_1 + \beta_1, \dots, \alpha_N + \beta_N)$.

Moreover, we call a vector-valued function $W : \mathbb{R}^N \rightarrow \mathbb{R}^N$ analytic on $\mathcal{D} \subset \mathbb{R}^N$, if its components are analytic functions on $\mathcal{D}$.

Why analytic functions? The reason focus is placed on analytic functions is that their sets of zeroes satisfy special properties.

Lemma 5.4 (Zeroes of analytic functions) *Assume $f : \mathbb{R}^N \rightarrow \mathbb{R}$ is an analytic function on the open set $\mathcal{D} \subset \mathbb{R}^N$. Then, if f has an accumulation point of zeroes on $\mathcal{D}$ then $f \equiv 0$ on $\mathcal{D}$.*

Proof : We prove that the zeroes of a non-constant analytic function are isolated. Let f be an analytic function defined in some domain $\mathcal{D} \subset \mathbb{R}$ and let $f(z_0) = 0$ for $z_0 \in \mathcal{D}$. Because f is analytic, there is a Taylor series expansion for f around z_0 which converges on an open ball within $\mathcal{D}$. i.e. for $|z - z_0| < R$. We may express f around z_0 as $f(z) = (z - z_0)^\alpha \sum_{|\beta| \geq |\alpha|}^\infty a_\beta (z - z_0)^\beta$, with $a_\alpha \neq 0$ and $k \in \mathbb{N}$. i.e. a_α is the first non-zero term. Define $g(z) = \sum_{|\gamma|=0}^\infty a_{\alpha+\gamma} (z - z_0)^\gamma$ so that $f(z) = (z - z_0)^\alpha g(z)$. Observe that $g(z)$ is analytic on $|z - z_0| < R$. It is now sufficient to show that g is non-zero in some punctured neighbourhood around z_0 . Note that as $g(z)$ is analytic, it is continuous at z_0 . Notice that $g(z_0) = a_\alpha \neq 0$, so there exists an $\epsilon > 0$ so that for all z with $|z - z_0| < \epsilon$ it follows that $|g(z) - a_\alpha| < \frac{|a_\alpha|}{2}$. This implies that $g(z)$ is non-zero in this ball, as required. ■

In order to get a proper handle over the linear subspace spanned by an infinite number of vectors, we need to discuss determinants. No attempt to patronise the reader is made by giving a lengthy review of linear algebra, but a simple notation for determinants of matrices (and determinants of its submatrices) is introduced. Let $A \in \mathbb{R}^{m \times n}$ then for a given size of square submatrix: $1 \leq k \leq m \wedge n$ we enumerate the submatrices of said size and define

$$|A|_{l,k \times k}$$

to be the determinant of the l th $k \times k$ submatrix. Note that there are $\binom{m}{k} \binom{n}{k}$ submatrices of size $k \times k$.

Corollary 5.5 *Let $W_1, \dots, W_n$ be analytic vector fields on $\mathcal{D} \subset \mathbb{R}^N$. Then $|W_1, \dots, W_n|_{l,k \times k}$ analytic on $\mathcal{D}$, which, on any compact subset $K \subset \mathcal{D}$ is either identically zero, or zero at a finite number of points.*

Proof : The first statement about analyticity of the determinant is obvious once one has noted that it is a polynomial combination of analytic functions. It is a basic property of analytic functions that polynomial combinations inherit the analyticity. The proof of the second claim is also straightforward. Assume we have an infinite number of zeroes of $|W_1, \dots, W_n|_{l,k \times k}$ on the compact set K . Choose a sequence from this set of zeroes: $\{x_n\}_n \subset K$. Using the compactness of K this sequence must have a convergent subsequence $\{x_{n_k}\}_k$, i.e. $x_{n_k} \rightarrow x \in K$. This is saying no more than that $|W_1, \dots, W_n|_{l,k \times k}$ has an accumulation point of zeroes. Hence $|W_1, \dots, W_n|_{l,k \times k} \equiv 0$ and the result is proved. ■

5.2. Finite generation of the Lie algebra

In this section we consider one of the two properties linked to the LFG condition. We seek to show under two different assumptions, that the Lie algebra is finitely generated. The first is that the vector fields are analytic in an open subset of $\mathbb{R}^N$. The second situation considered is when the

Lie algebra has constant rank in some open set. In both situations we prove that the Lie algebra is finitely generated on every compact subset of the open set.

ANALYTIC COEFFICIENTS

In this subsection we deal exclusively with analytic coefficients. In particular, we show that the vector subspace spanned by the infinite number of vector fields of the Lie algebra, is, on compact sets, spanned by a finite number of these. Namely, that the Lie algebra is *finitely generated* on compact subsets of analytic regions. Throughout this section we assume that the vector fields are analytic in an open set $\mathcal{D} \subset \mathbb{R}^N$, and that $K \subset \mathcal{D}$ is a compact subset. In order to make this both aesthetically pleasing and readily understandable we have to introduce some notation. Define:

$$\mathcal{K}(x) := \text{Span} \{V_{[\alpha]}(x) ; \alpha \in \mathcal{A}\}.$$

Our first observation is that although the subspace $\mathcal{K}(x)$ of $\mathbb{R}^N$ is the span of an infinite number of vectors, for each $x \in \mathbb{R}^N$ we can choose precisely $r(x) := \text{rank}(\mathcal{K}(x))$ members of the Lie algebra: $V_{[\alpha_1]}, \dots, V_{[\alpha_{r(x)}]}$ which span $\mathcal{K}(x)$ at x . This is basic linear algebra. The problem we are considering is whether there exists $V_{[\alpha_1]}, \dots, V_{[\alpha_N]}$ such that:

$$\mathcal{K}(x) = \text{Span}\{V_{[\alpha_1]}(x), \dots, V_{[\alpha_N]}(x)\}, \quad \forall x \in K.$$

Define $\mathcal{K}_m(x) := \text{Span} \{V_{[\alpha]}(x) ; \|\alpha\| \leq m\}$. So that for each fixed $x \in \mathbb{R}^N$ there exists n_x such that:

$$\mathcal{K}(x) = \mathcal{K}_{n_x}(x).$$

This is obvious from the definition of both subspaces. Define also:

$$R(x) := \text{rank}(\mathcal{K}(x)),$$

$$R_m(x) := \text{rank}(\mathcal{K}_m(x)).$$

It is clear that for each $x \in \mathbb{R}^N$ there exists n_x such that $R(x) = R_{n_x}(x)$. Indeed, $R(x) = \lim_{n \rightarrow \infty} R_n(x)$. Rather than focus on the problem of finding a N_K such that $\mathcal{K}_{N_K} = \mathcal{K}$ on K , we study the equivalent problem of showing that the sequence $\{R_n\}_n$ is uniformly convergent to R on K .

It is clear that:

$$K = \bigcup_{r=0}^N \{R = r\} \cap K = \bigcup_{r=0}^N \{x \in \mathbb{R}^N : R(x) = r\} \cap K$$

If $\{R = r\} \cap K \neq \emptyset$ then there must exist $x \in K$ such that $R(x) = r$. Thus, there exists $n_x \in \mathbb{N}$ such that $R(x) = R_{n_x}(x) = r$. Hence, some $r \times r$ submatrix of $(V_{[\alpha]}(x))_{|\alpha| \leq n_x}$ must have non-zero determinant:

$$|(V_{[\alpha]}(x))_{\|\alpha\| \leq n_x}|_{l, r \times r} > 0, \quad \text{for some } l.$$

But, given that this analytic function is not zero everywhere on K it can only be zero for at most a finite number of points: $\{x_1, \dots, x_M\} \subset K$. This means that the vectors of this submatrix must be a finite generating set for all but a finite number of points of rank r on the compact set K . Alternatively, $\{R = r\} \cap K \setminus \{x_1, \dots, x_M\} = \{R_{n_x} = r\} \cap K$. This is promising. There are at most a finite number of more troublesome points. **But** these points are individually finitely generated. i.e. for each x_i , $i = 1, \dots, M$ there exists n_{x_i} such that $R(x_i) = R_{n_{x_i}}(x_i)$. So, if we then take $n_r := \max\{n_x, n_{x_1}, \dots, n_{x_M}\}$ then it is clear that:

$$\{R = r\} \cap K = \{R_{n_r} = r\} \cap K.$$

We then repeat this step for each possible rank of $\mathcal{K}$ for points in K , i.e. $0, 1, \dots, N$, and take $n_* := \max\{n_0, \dots, n_N\}$. It then follows that:

$$K = \{R = R_{n_*}\} \cap K.$$

Equivalently,

$$\mathcal{K} = \mathcal{K}_{n_*}, \quad \text{on } K.$$

Hence we have succeeded in demonstrating finite generation of the Lie algebra on compact sets.

Proposition 5.6 *Let K be a compact subset of a open region in which the vector fields $V_0, \dots, V_d$ are analytic. Then there exists $N_K \in \mathbb{N}$ s.t.*

$$\text{Span} \{V_{[\alpha]}(x) ; \alpha \in \mathcal{A}\} = \text{Span} \{V_{[\alpha]}(x) ; \alpha \in \mathcal{A}(N_K)\}, \quad \forall x \in K.$$

In particular, for each $V_{[\alpha]}$ such that $\alpha \notin \mathcal{A}(N_K)$, there exists $\varphi_{\alpha, \beta}^{(\cdot)} : \mathbb{R}^N \rightarrow \mathbb{R}$, $\beta \in \mathcal{A}(N_K)$, such that:

$$V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}(N_K)} \varphi_{\alpha, \beta}^x V_{[\beta]}(x), \quad (5.1)$$

for all $x \in K$.

This represents significant progress for proving the LFG condition holds on compact subsets of analytic regions. We are not yet finished though. In comparing (5.1) with the actual LFG condition, one will observe that the required smoothness property is yet to be demonstrated on the interior of K . Indeed, denoting the coefficients by $\varphi_{\alpha, \beta}^x$ is designed to highlight this difference. What remains is the daunting task of showing that one may express all elements of the Lie algebra as a smooth linear combination. This is a highly non-trivial problem. We shall restrict consideration to points across which the rank of the Lie algebra is constant.

Before proceeding to tackle this problem, we first take a slight deviation to prove a generalisation of the finite generation results, also under the assumption of constant rank. In particular we show that if the vector fields are smooth, and the Lie algebra has constant rank, then it can be finitely generated.

CONSTANT RANK

The biggest problem in the last section - had we not assumed analyticity properties of the coefficients - would have been how to deal with a change in rank of the Lie algebra. These points pose significant problems for general functions as there can be infinitely many of them and it is not obvious that finite generation would still hold. In the last section we showed that there would only be a finite number of isolated points across which the rank could change. Here, we will allow our functions to be smooth but we shall assume that there holds constant rank of Lie algebra across an open set $\mathcal{D}$. i.e. recalling the notation of the previous section

$$V_0, \dots, V_d \in C^\infty(\mathbb{R}^N; \mathbb{R}^N),$$

satisfy

$$R \equiv r, \quad \text{on } \mathcal{D},$$

where $r \in \{0, \dots, N\}$.

We now make a slight deviation to investigate the properties of the rank function. These will become useful when discussing finite generation.

Definition 5.7 (Lower Semi-continuous function) A function $f : \mathbb{R}^N \rightarrow \bar{\mathbb{R}}$ is said to be lower-semicontinuous if one of the following two equivalent conditions hold:

1. For all $x \in \mathbb{R}^N$ $\{y \in \mathbb{R}^N; f(y) > x\}$ is open.
2. $\liminf_{n \rightarrow \infty} f(x_n) \geq f(\liminf_{n \rightarrow \infty} x_n)$.

Or, in words: a lower semicontinuous functions is a piecewise continuous function which can make 'down-jumps' at a point, but not 'up-jumps'.

Lemma 5.8 The rank function $R_m : \mathbb{R}^N \rightarrow \{0, \dots, N\}$ is lower-semicontinuous.

Proof : This property seems intuitive after some thought, but assume $A(r) := \{y \in \mathbb{R}^N; R_m(y) > r\}$ is non-empty and take $y_0 \in A(r)$. Assume wlog that $y_0 \in A(r')$ where $r' > r$. Then there must hold:

$$|(V_{[\alpha]}(y_0))_{\alpha \in \mathcal{A}(i)}|_{k, r' \times r'} > 0,$$

for some $r' \times r'$ submatrix of $(V_{[\alpha]}(y_0))_{\alpha \in \mathcal{A}(m)}$. But since this is a smooth function of y_0 (which follows from the smooth columns of the submatrix) and hence also continuous, there is a neighbourhood U of y_0 such that

$$|V_{[\alpha]}(x); \alpha \in \mathcal{A}(m)|_{k, r' \times r'} > 0,$$

for all $x \in U$. This means that each point $x \in U$ is of rank **at least** $r' > r$. i.e. $R_m(x) > r$ for all $x \in U$. i.e. $A(r)$ is an open set. ■

Remark 5.9 For general lower-semicontinuity of rank, one need only assume continuity of the elements.

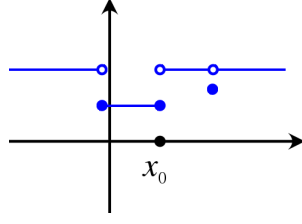


Figure 5.3.: A lower-semicontinuous function

We now prove a result for which a corollary is - under the assumptions of constant rank - finite generation.

Proposition 5.10 (Uniform convergence of lower-semicontinuous functions) *Let $K \subset \mathbb{R}^N$ be a compact set. Assume $\{f_i\}_{i \in \mathbb{N}} : K \rightarrow \mathbb{R}$ is an increasing sequence of lower semi-continuous functions such that $\sup_i f_i \equiv C$. Then for each $\epsilon > 0$ there exists $i_0 \in \mathbb{N}$ such that $\{f_{i_0} > C - \epsilon\} = K$.*

Proof : Because $f_i \uparrow_{\text{ptw}} C = \sup_j f_j$, for each $x \in K$ there exists $i_x \in \mathbb{N}$ such that:

$$f_{i_x}(x) > C - \epsilon.$$

Now, since each f_i is lower semi-continuous, there exists an open neighbourhood U_x of x such that

$$f_{i_x}(y) > C - \epsilon \quad \forall y \in U_x.$$

Now note that: $\bigcup_{x \in K} U_x$ is an open covering of K . From the compactness of K , there must be a finite subcovering:

$$K \subset \bigcup_{i=1}^M U_{x_i}.$$

Now if we take $i^* := \max\{i_1, \dots, i_M\}$, then it follows that $f_{i^*} > C - \epsilon$ for all $x \in K$. This completes the proof. ■

Corollary 5.11 *An increasing sequence of lower semi-continuous functions, which converges pointwise to some continuous function, converges uniformly on compact sets.*

Proof : This is a simple corollary of the above lemma once we note that if g is continuous and f_i is lower semicontinuous then $f_i - g$ is an increasing lower semicontinuous functions which converges pointwise to 0, hence this convergence is uniform on compacts. ■

Hopefully, the relevance of these facts for the task in hand should be clear.

Corollary 5.12 *If the rank function of the Lie algebra is constant on $\mathcal{D}$, then the Lie algebra is finitely generated on compact subsets of $\mathcal{D}$.*

Proof : Let $K \subset \mathcal{D}$ be a compact subset. Simply choose $\epsilon < 1$ in Proposition (5.10). Then it must be that for the increasing sequence $\{R_i\}_{i \in \mathbb{N}}$ of lower semi-continuous rank functions, that there exists n^* such that:

$$R_{n^*} = r, \quad \text{on } K.$$

i.e.

$$\text{Span } \{V_{[\alpha]}(x) : \alpha \in \mathcal{A}\} = \text{Span } \{V_{[\alpha]}(x) : \alpha \in \mathcal{A}(n^*)\}.$$

■

Before focussing on smoothness of the LFG coefficients under the constant rank assumptions, we consider the behaviour of the rank function where the vector fields are analytic on an open subset of $\mathbb{R}^N$.

5.3. Constant rank of the analytic Lie algebra

In this section we focus our attention on the rank of the Lie algebra. In particular, we consider the situation where the vector fields are analytic in some open subset of $\mathbb{R}^N$. It was proved in a previous section that the Lie algebra is finitely generated on compact subsets of analytic regions. For the rank function, this means that there exists N_K such that $R_{N_K} = R$ on K , where K is a compact subset.

Definition 5.13 (Primary and secondary Lie algebra ranks) Assume K is a compact subset of $\mathbb{R}^N$ satisfying $\text{Leb}(K) > 0$. We call $r \in \{0, \dots, N\}$ a **primary rank for the Lie algebra on K** if it holds that $\text{Leb} \{x \in K : R(x) = r\} > 0$. We call $r \in \{0, \dots, N\}$ a **secondary rank for the Lie algebra on the compact set K** , if it holds that $\text{Leb} \{x \in K : R(x) = r\} = 0$, but there exists $x \in K$ such that $R(x) = r$.

Proposition 5.14 If the vector fields $V_0, \dots, V_d$ are analytic on $\mathcal{D} \subset \mathbb{R}^N$, then for any compact subset $K \subset \mathcal{D}$, with $\text{Leb}(K) > 0$, the Lie algebra has only one primary rank, r , on K . Moreover, $\text{Leb}\{x \in K : R(x) = r\} = \text{Leb}(K)$ and this primary rank is the rank of the Lie algebra for all but a finite number of points.

Proof : We proved in the previous section that there exists $n(K)$ such that $R_{n(K)} = R$ on K . We consider the $N \times N_{n(K)}$ -valued matrix function, denoted A , and defined as:

$$A(x) := [V_{[\alpha]}(x)]_{\alpha \in \mathcal{A}(N_{n(K)})}.$$

On K , the rank of this matrix is the same as that of the Lie algebra. Hence it is sufficient to show the result for this matrix. Let us assume that A has two primary ranks: r_1 and r_2 on K . Without

loss of generality we may assume $r_1 > r_2$. For all $x \in \{y \in K : R_{n(K)} < r_1\} =: K_{r_1}^c$ it follows that:

$$\sum_{l=1}^{\binom{N}{r_2} \binom{N_{n(K)}}{r_1}} |A|_{l, r_1 \times r_1} = 0, \quad \text{on } K_{r_1}^c. \quad (5.2)$$

i.e. as the rank of A is strictly less than r_1 on $K_{r_1}^c$, it follows that every $r_1 \times r_1$ -submatrix of A must have zero determinant. But since $\text{Leb}(K_{r_1}^c) \geq \text{Leb}\{y \in K : R_{n(K)} = r_2\} > 0$, the set $K_{r_1}^c$ must have an accumulation point. This follows easily after noting that a set comprised of isolated points - a discrete set - is countable and hence has Lebesgue measure zero. Hence the analytic function in (5.2) has an accumulation point of zeroes. It must be zero on K . Hence r_1 is **not** a primary rank of the Lie algebra. This contradiction completes the proof. We note that the assertion: $\text{Leb}\{x \in K : R(x) = r\} = \text{Leb}(K)$ follows trivially from the definition of a primary rank, once we know that there is *only* one primary rank. The fact that this is the rank for all but a finite number of points is deduced from the analyticity of the function in (5.2). ■

5.4. Satisfaction of the LFG condition under constant rank

At the beginning of the section we discussed the two-fold nature of the problem of demonstrating LFG condition. One must first determine a finite number of vector fields which generate the whole Lie algebra. It is then essential to show that they can be combined as a linear combination with smooth coefficients. Whilst a seemingly simple problem, this is highly non-trivial. However, the result is fairly straightforward to prove when we assume that the linear span of the Lie algebra has ‘constant rank’ across a point. We used this same notion in the previous section, but we now give a precise definition of what this means.

Definition 5.15 (Point of constant rank) *A point $x \in \mathbb{R}^N$ is called a point of constant rank if there exists $\epsilon > 0$ and $r \in \{0, 1, \dots, N\}$ such that*

$$R(\cdot) = r, \quad \text{on } B_\epsilon(x). \quad (5.3)$$

The choice of LFG coefficients shall be made according to the prescription allowed through the implicit function theorem.

Theorem 5.16 (Implicit function theorem) *Let $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ be a C_b^∞ function with coordinates (x, y) . Fix $(x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^m$ and suppose $f(x_0, y_0) = c$, for $c \in \mathbb{R}^m$. Suppose further that the matrix $[\frac{\partial f_i}{\partial y_j}(x_0, y_0)]_{1 \leq i, j \leq m}$ is invertible. Then there exists an open set $U \subset \mathbb{R}^n$ containing x_0 and an open set $V \subset \mathbb{R}^m$ containing y_0 and a unique smooth function $g : U \rightarrow V$ with bounded derivatives such that*

$$\{(x, g(x)) : x \in U\} = \{(x, y) \in U \times V : f(x, y) = c\}. \quad (5.4)$$

This result is ideal for proving smoothness of the choice of LFG coefficients. We make a brief observation on the finite generation results. Assume that the $\{V_{[\beta]}, \beta \in \mathcal{A}(m^*)\}$ are a generating set for the whole Lie algebra. We consider an arbitrary member $V_{[\alpha]}$ of the Lie algebra which is not in this finite generating set. For example, assume $\|\alpha\| = m^* + 1$. Fix $x_0 \in \mathbb{R}^N$. Then, if $R(x_0) = r$, there exists $V_{[\beta_1]}, \dots, V_{[\beta_r]}$ and $y_1^{x_0}, \dots, y_r^{x_0}$ such that:

$$V_{[\alpha]}(x_0) = \sum_{i=1}^r y_i^{x_0} V_{[\beta_i]}(x_0). \quad (5.5)$$

Of course, the purpose of this section is to show that these may be chosen smoothly in some region around x_0 .

Proposition 5.17 *Let $x_0 \in \mathbb{R}^N$ be a point of constant rank r with $R(y) = r$ for all $y \in B_\epsilon(x_0)$. Then the LFG condition is satisfied in some neighbourhood of x_0 .*

Proof : As $x_0 \in \mathbb{R}^N$ is a point of constant rank r , there exists $V_{[\beta_1]}, \dots, V_{[\beta_r]}$, which are linearly independent such that, for some $l \leq \binom{N}{r}$:

$$\left| \begin{bmatrix} V_{[\beta_1]}(x_0) & \dots & V_{[\beta_r]}(x_0) \end{bmatrix} \right|_{l, r \times r} > 0,$$

where l corresponds to the l th $r \times r$ -submatrix of the matrix $[V_{[\beta_1]} \dots V_{[\beta_r]}]$. As the elements of this matrix are continuous (they are actually smooth), there exists some $\epsilon' > 0$ such that

$$\left| \begin{bmatrix} V_{[\beta_1]}(y) & \dots & V_{[\beta_r]}(y) \end{bmatrix} \right|_{l, r \times r} > 0,$$

for all $y \in B_{\epsilon'}(x_0)$. As the constant rank r holds on $B_\epsilon(x_0)$, and $V_{[\beta_1]}, \dots, V_{[\beta_r]}$ have at least rank r on $B_{\epsilon'}(x_0)$. It follows that $V_{[\beta_1]}, \dots, V_{[\beta_r]}$ have precisely rank r on $B_{\tilde{\epsilon}}(x_0)$ where $\tilde{\epsilon} := \epsilon \wedge \epsilon'$. Hence, they form a generating set for the Lie algebra on $B_{\tilde{\epsilon}}(x_0)$: both the Lie algebra and the set of vectors itself have the same rank, and the generating set is obviously contained in the Lie algebra, hence it must be a generating set for the Lie algebra in this region.

For reasons which shall become obvious, we now wish to choose $N - r$ smooth vectors, denoted as $W_{r+1}, \dots, W_N$, which, when combined with the linearly dependent vectors $V_{[\beta_1]}, \dots, V_{[\beta_r]}$, form a linearly independent set of vector fields on $B_{\tilde{\epsilon}}(x_0)$. If $r < N$, there will be many choices of $W_{r+1}, \dots, W_N$. For example, one could choose $N - r$ linearly independent vectors at x_0 and change their orientation so that the angle between the vectors and the hyperplane $\mathcal{K}(\cdot)$ remains constant. As this hyperplane is of constant rank and is itself comprised of smooth vectors fields, the resultant vectors $W_{r+1}, \dots, W_N$ will also be smooth. We write $W_i := V_{[\beta_i]}$ for $i = 1, \dots, r$, for notational convenience.

Consider the function $f : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}^N$:

$$f(x, y) := V_{[\alpha]}(x) - \sum_{j=1}^N y_j W_j(x).$$

This function is smooth, and is equal to zero at the point (x_0, y_0) where $y_0 = (y_1^{x_0}, \dots, y_r^{x_0}, 0, \dots, 0)$, by (5.5). Moreover,

$$\left[\frac{\partial f_i}{\partial y_j}(x_0, y_0) \right]_{1 \leq i, j \leq m} = [V_{[\alpha]}(x) - W_i(x)]_{i=1, \dots, r}.$$

This matrix is invertible at (x_0, y_0) by construction, as long as $V_{[\alpha]} - V_{[\beta_i]} \neq 0$ for $i = 1, \dots, r$. If this is the case, for $i = i_0$, we may instead apply the theorem on the function $\tilde{f}(x, y)$ with y_{i_0} replaced by $2y_{i_0}$. Then the Jacobian of this new matrix is invertible at the point $(x_0, \tilde{y}_0)$ where

$$\tilde{y}_0 = \left(y_1^{x_0}, \dots, \frac{1}{2} y_{i_0}^{x_0}, \dots, y_r^{x_0}, 0, \dots, 0 \right).$$

Hence, this pathology may be easily avoided and we may apply the machinery provided by the implicit function theorem. Namely, there exists an open neighbourhood U of x_0 , an open neighbourhood V of y_0 , and a unique smooth function $\varphi_\alpha : U \rightarrow V$ such that

$$f(x, \varphi(x)) = 0, \quad \text{on } U.$$

In particular,

$$V_{[\alpha]}(x) = \sum_{i=1}^N \varphi_{\alpha, i}(x) W_i(x).$$

But $W_{r+1}, \dots, W_N$ are linearly independent of the vectors in the subspace $\mathcal{K}(\cdot)$ on $U \cap B_{\tilde{\epsilon}}(x_0)$, so it must be that $\varphi_{r+1}, \dots, \varphi_N$ are identically zero on this set, too. Hence,

$$V_{[\alpha]}(x) = \sum_{i=1}^r \varphi_{\alpha, i}(x) V_{[\beta_i]}(x),$$

on $U \cap B_{\tilde{\epsilon}}(x_0)$, as required. ■

Theorem 5.18 (The generality of the LFG condition) *The following is a summary of the results from this section:*

1. Assume that $V_0, \dots, V_d \in C^\infty(\mathcal{D})$ are smooth vector fields on an open set $\mathcal{D}$. Then, if $x_0 \in \mathcal{D}$ is a point of constant rank, then the LFG condition holds in some neighbourhood of x_0 .
2. Assume that $V_0, \dots, V_d \in C^\omega(\mathcal{D})$ are analytic vector fields on an open set $\mathcal{D}$. Let K be a compact subset of $\mathcal{D}$. Then the LFG condition holds in some neighbourhood of all but a finite number of points in K . If $\mathcal{D} = \mathbb{R}^N$ then the points, which have **no** neighbourhood for which the LFG condition is satisfied, form a set of isolated points.

6. Sobolev inequalities, perturbed semigroups, the Lagrangian and relaxing the V_0 condition

We begin this section by seeking to extend the class of gradient bounds resulting from the integration by parts formulae. In particular, some weighted Sobolev-type inequalities are proved and significantly extend the work of previous sections. In the second part of this chapter we consider the case where the semigroup has been perturbed by a potential. This problem is common in PDE theory and although this has already been mentioned, the restrictive assumption that the potential is smooth is significantly relaxed. In the third and final part of this chapter, we consider another problem which is prevalent in PDE theory; namely the Lagrangian term. This problem has not been considered thus far (within the thesis), and can be seen to complete the gradient bound theory for solutions of the general Cauchy problem.

6.1. Weighted Sobolev spaces

In this section we develop extensively and systematically gradient bounds via the integration by parts formula.

In previous sections we were restricted to considering functions which were either uniformly continuous, Lipschitz continuous or in L^p . Whilst this covers a very broad spectrum of test functions, it does not include, for example, continuous functions of polynomial growth or functions whose derivatives may exist: in strong or weak sense. In this subsection we seek to address that deficiency.

We introduce a very general space of functions which may be viewed as a generalisation of Sobolev spaces. In this case we shall be seeking derivatives along vector fields rather than partial derivatives along the axis of the Euclidean space.

Let $\mathbf{W} = \{W_1, \dots, W_M\}$ be a set of M vector fields. Let $f \in C_0^\infty(\mathbb{R}^N)$ and define:

$$\begin{aligned} \|f\|_{\alpha, \beta, \infty}^{\mathbf{W}} &= \sup_{x \in \mathbb{R}^N} (1 + |x|^2)^{-\frac{\alpha}{2}} |W_{\beta_1} \dots W_{\beta_n} f(x)| \\ \|f\|_{\alpha, \beta, p}^{\mathbf{W}} &= \left[\int_{\mathbb{R}^N} (1 + |x|^2)^{-\frac{\alpha p}{2}} |W_{\beta_1} \dots W_{\beta_n} f(x)|^p dx \right]^{1/p}, \end{aligned}$$

where $\alpha \in \mathbb{R}$ and $\beta = (\beta_1, \dots, \beta_n) \in \{\emptyset\} \cup \bigcup_{k \in \mathbb{N}_0} \{1, \dots, M\}^k$, where no derivatives of f are taken

for $\beta = \emptyset$. Then define for $\tilde{k} \in \mathbb{N}_0$

$$\begin{aligned} \|f\|_{\alpha, \tilde{k}, \infty}^{\mathbf{W}} &= \sum_{|\beta| \leq \tilde{k}} \|f\|_{\alpha, \beta, \infty}^{\mathbf{W}} \\ \|f\|_{\alpha, \tilde{k}, p}^{\mathbf{W}} &= \left[\sum_{|\beta| \leq \tilde{k}} \left(\|f\|_{\alpha, \beta, p}^{\mathbf{W}} \right)^p \right]^{1/p}. \end{aligned}$$

Note that for $\tilde{k} = 0$, that the above norm is independent of $\mathbf{W}$. We omit the superscript in this case to emphasise this fact. We then define the associated normed linear spaces, for $\alpha \in \mathbb{R}$, $\tilde{k} \in \mathbb{N}_0$, and $p \in [1, \infty]$:

$$H_{\alpha, \tilde{k}, p}^{\mathbf{W}} := \overline{\{\varphi \in C_0^\infty(\mathbb{R}^N) : \|\varphi\|_{\alpha, \tilde{k}, p}^{\mathbf{W}} < \infty\}}.$$

i.e. $H_{\alpha, \tilde{k}, p}^{\mathbf{W}}$ is the closure of $C_0^\infty(\mathbb{R}^n)$ with respect to the norm $\|\cdot\|_{\alpha, \tilde{k}, p}^{\mathbf{W}}$. They may be thought of as being weighted Sobolev-type spaces. Indeed, for $\mathbf{W} = \{e_1, \dots, e_N\}$, and $\alpha = 1$, the closure $H_{0, \tilde{k}, p}^{\mathbf{W}}$ is the Sobolev space $W^{k, p}$. Observe that we have not stated what differentiability we require of the elements of $\mathbf{W}$. If we wish to consider vector fields of general differentiability, then we need to restrict the definition of the norms $\|\cdot\|_{\alpha, \tilde{k}, p}^{\mathbf{W}}$ so that they are well-defined. For instance, if $\mathbf{W} \subset C^l(\mathbb{R}^N)$, then the norms would be defined for each β such that $|\beta| \leq l + 1$, and consequently for $\tilde{k} \leq l + 1$.

For $\alpha \geq 0$, this denotes functions, which may be expressed as the limit of smooth functions, and which have weak derivatives along $\mathbf{W}$ up to order $\tilde{k}$ in L^p , which all have at most polynomial growth of order α . For $\alpha < 0$, this denotes functions with weak derivatives along $\mathbf{W}$ of order up to $\tilde{k}$ which decay at least of polynomial order α .

Some of the results of the previous section might be viewed as somewhat disappointing, given the necessity to phrase the norms over compact sets. The main reason for this problem is the unboundedness of the vector fields of the Lie algebra. As we shall show, we need to consider weighted spaces to rectify this problem and these new spaces are ideal for considering the action of the diffusion semigroup especially in the case of vector fields with linear growth. The bounds in this section rely heavily on the careful analysis we made of the explosion of the integration by parts factors in Chapter 3.

In what follows we assume the UFG condition of order m holds and define $\mathbf{V} := \{V_{[\alpha]} : \alpha \in \mathcal{A}(m)\}$. Unless otherwise stated, we assume $f \in C_0^\infty$, although this assumption may be relaxed by limit arguments. Note: we are still assuming that the vector fields $V_1, \dots, V_d \in C_b^{k+1}$ and that $V_0 \in C_b^k$. For compactness of notation, we shall write $\alpha(n_1) := \alpha + n_1$ and $\alpha(n_1, n_2) := \alpha + n_1 + n_2$.

Lemma 6.1 *For $n_1 + n_2 \leq k - m - 1$ the following hold for some constants $C, \tilde{C}$*

$$\begin{aligned} \left\| V_{[\alpha_1]} \dots V_{[\alpha_{n_1}]} P_t V_{[\alpha_{n_1+1}]} \dots V_{[\alpha_{n_1+n_2}]} f \right\|_{\alpha(n_1, n_2), 0, \infty} &\leq C t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \|f\|_{\alpha, 0, \infty} \\ \left\| V_{[\alpha_1]} \dots V_{[\alpha_{n_1}]} P_t V_{[\alpha_{n_1+1}]} \dots V_{[\alpha_{n_1+n_2}]} f \right\|_{\alpha(n_1, n_2), 0, 1} &\leq \tilde{C} t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \|f\|_{\alpha, 0, 1}, \end{aligned}$$

and as a result of employing the Riesz-Thorin interpolation theorem, for every $p \in [1, \infty]$:

$$\left\| V_{[\alpha_1]} \dots V_{[\alpha_{n_1}]} P_t V_{[\alpha_{n_1+1}]} \dots V_{[\alpha_{n_1+n_2}]} f \right\|_{\alpha(n_1, n_2), 0, p} \leq C_p t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \|f\|_{\alpha, 0, p}.$$

Proof : Observe that

$$\begin{aligned} & V_{[\alpha_1]} \dots V_{[\alpha_{n_1}]} P_t V_{[\alpha_{n_1+1}]} \dots V_{[\alpha_{n_1+n_2}]} f(x) \\ &= t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \mathbb{E} [f(X_t^x) \Phi(t, x)] \\ &\leq t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \mathbb{E} |f(X_t^x) \Phi(t, x)| \\ &\leq t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \|f\|_{\alpha, 0, \infty} \mathbb{E} \left| (1 + |X_t^x|^2)^{\frac{\alpha}{2}} \Phi(t, x) \right| \\ &\leq t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \|f\|_{\alpha, 0, \infty} \left\| (1 + |X_t^x|^2)^{\frac{\alpha}{2}} \right\|_{L^2(\Omega)} \|\Phi(t, x)\|_{L^2(\Omega)} \\ &\leq C t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \|f\|_{\alpha, 0, \infty} (1 + |x|^2)^{\frac{\alpha}{2}} (1 + |x|)^{n_1+n_2}. \end{aligned}$$

The last inequality follows from Corollary 3.13. Rearranging and taking the supremum over $x \in \mathbb{R}^N$ completes the proof for $p = \infty$. We now consider the case where $p = 1$. Observe that

$$\begin{aligned} & \int_{\mathbb{R}^N} (1 + |x|^2)^{-\alpha(n_1, n_2)/2} V_{[\alpha_1]} \dots V_{[\alpha_{n_1}]} P_t V_{[\alpha_{n_1+1}]} \dots V_{[\alpha_{n_1+n_2}]} f(x) g(x) dx \\ &\leq \int_{\mathbb{R}^N} f(x) V_{[\alpha_{n_1+n_2}]}^* \dots V_{[\alpha_{n_1+1}]}^* \tilde{P}_t V_{[\alpha_{n_1}]}^* \dots V_{[\alpha_1]}^* \tilde{g}(x) dx, \end{aligned} \quad (6.1)$$

where $\tilde{g}(x) = g(x)(1 + |x|^2)^{-\alpha(n_1, n_2)/2}$. Now we observe, using an analagous argument to the proof of the previous lemma:

$$\begin{aligned} & V_{[\alpha_{n_1+n_2}]}^* \dots V_{[\alpha_{n_1+1}]}^* \tilde{P}_t V_{[\alpha_{n_1}]}^* \dots V_{[\alpha_1]}^* \tilde{g}(x) \\ &= t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \mathbb{E} [\tilde{g}(X_t^x) \Phi(t, x)] \\ &\leq t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \|g\|_{0, 0, \infty} \mathbb{E} \left| (1 + |X_t^x|^2)^{-\alpha(n_1, n_2)/2} \Phi(t, x) \right| \\ &\leq t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \|g\|_{0, 0, \infty} \left\| (1 + |X_t^x|^2)^{-\alpha(n_1, n_2)/2} \right\|_{L^2(\Omega)} \|\Phi(t, x)\|_{L^2(\Omega)} \\ &\leq C t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \|g\|_{0, 0, \infty} (1 + |x|^2)^{-\alpha(n_1, n_2)/2} (1 + |x|^2)^{(n_1+n_2)/2} \\ &\leq C t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{n_1+n_2}\|}{2}} \|g\|_{0, 0, \infty} (1 + |x|^2)^{-\alpha/2}. \end{aligned}$$

The last inequality follows from Lemma 2.3 in Kunita [19]. Note that we have applied the first part of this lemma to $\tilde{P}_t$ and g . The result is completed by substituting the above inequality into (6.1) and taking the supremum over all $g \in C_0^\infty$ such that $\|g\|_\infty = \|g\|_{0, 0, \infty} \leq 1$. The Riesz Thorin interpolation argument is used to deduce the bounds for $1 < p < \infty$. ■

This simple corollary of the integration by parts formula has implications for P_t as a continuous linear operator between Banach spaces.

Proposition 6.2

For $n_1 \leq k - m - 1$, and $t \in (0, 1]$, it follows that $P_t \in B(H_{\alpha,0,p}, H_{\alpha(n_1),n_1,p}^{\mathbf{V}})$. Moreover,

$$\|P_t\|_{H_{\alpha,0,p}^{\mathbf{V}} \rightarrow H_{\alpha(n_1),n_1,p}^{\mathbf{V}}} \leq C_{\alpha,p} t^{-\frac{n_1 m}{2}}.$$

Proof : Observe that:

$$\left[\|P_t f\|_{\alpha(n_1),p}^{\mathbf{V}} \right]^p = \sum_{\substack{\beta=(\beta_1,\dots,\beta_n) \\ \text{s.t } \beta_i \in \mathcal{A}(m), n \leq n_1}} \left[\|P_t f\|_{\alpha,\beta,p}^{\mathbf{V}} \right]^p.$$

But it is clear from the definitions that

$$\|P_t f\|_{\alpha(n_1),\beta,p}^{\mathbf{V}} = \left\| V_{[\beta_1]} \dots V_{[\beta_{n_1}]} P_t f \right\|_{\alpha(n_1),0,p}.$$

And so, using Lemma 6.1, we get for $t \in (0, 1]$

$$\begin{aligned} \left[\|P_t f\|_{\alpha(n_1),n_1,p}^{\mathbf{V}} \right]^p &= \sum_{\substack{\beta=(\beta_1,\dots,\beta_n) \\ \text{s.t } \beta_i \in \mathcal{A}(m), n \leq n_1}} \left[\|P_t f\|_{\alpha(n_1),\beta,p}^{\mathbf{V}} \right]^p \\ &= \sum_{\substack{\beta=(\beta_1,\dots,\beta_n) \\ \text{s.t } \beta_i \in \mathcal{A}(m), n \leq n_1}} \left[\|V_{[\beta_1]} \dots V_{[\beta_{n_1}]} P_t f\|_{\alpha(n_1),0,p}^{\mathbf{V}} \right]^p \\ &\leq \sum_{\substack{\beta=(\beta_1,\dots,\beta_n) \\ \text{s.t } \beta_i \in \mathcal{A}(m), n \leq n_1}} \left[C_{p,\alpha,\beta} t^{-\frac{\|\beta_1\|+\dots+\|\beta_n\|}{2}} \|f\|_{\alpha,0,p} \right]^p \\ &\leq C_{\alpha,p}^p t^{-\frac{n_1 m p}{2}} \|f\|_{\alpha,0,p}^p. \end{aligned}$$

This completes the proof. ■

Although, it is reasonable to expect that when the test function is already endowed with some differentiability along $\mathbf{V}$, that the regularity results may be further strengthened, we were not able to prove results to this effect, and the following is a conjecture.

Conjecture 6.3 For each $n_2 + n_1 \leq k - m - 1$, and $t \in (0, 1]$, it follows that $P_t \in B(H_{\alpha,n_2,p}^{\mathbf{V}}, H_{\alpha(n_1,n_2),n_2+n_1,p}^{\mathbf{V}})$. Moreover,

$$\|P_t\|_{H_{\alpha,n_2,p}^{\mathbf{V}} \rightarrow H_{\alpha(n_1,n_2),n_2+n_1,p}^{\mathbf{V}}} \leq C_{\alpha,p} t^{-\frac{n_1 m}{2}}.$$

6.2. The perturbed semigroup

We have already provided gradient bounds for semigroups which have been perturbed by a potential term. Generally, the strategy thus far has been to assume that the semigroup term is "smooth enough" so as to apply the integration by parts formulae; by treating the potential as a Kusuoka-Stroock process. This treatment of the perturbed semigroup - as an afterthought of the integration by parts approach - is rectified here and in this section we study it more carefully. We analyse what happens when we relax the differentiability assumptions and allow c to be time-dependent. This question is highly non-trivial as the potential term acts as a memory for the semigroup, and captures short-term behaviour of the SDE. As we have already seen, this short-term behaviour can present singularities at time zero.

In this section, we consider the perturbed semigroup given by:

$$P_t^c f(x) := \mathbb{E} \left[f(X_t^x) \exp \left\{ \int_0^t c(s, X_s^x) ds \right\} \right].$$

If we assume that $c \in C_b^\infty((0, T) \times \mathbb{R}^N)$, there is no problem with applying the integration by parts formula, as we may treat the exponential term as a Kusuoka-Stroock process¹. The problem is more interesting when we relax these assumptions on c . Let us initially assume that $c \in C_b^\infty((0, T) \times \mathbb{R}^N)$ and that $f \in C_b^\infty(\mathbb{R}^N)$. Define $c_s(\cdot) := c(s, \cdot)$. In this section we shall assume that the vector fields: $V_0, \dots, V_d$ are **smooth and uniformly bounded with bounded derivatives**, for simplicity, although they can be extended to the non-uniformly bounded, general differentiability case. The following relationships are used.

Lemma 6.4

$$P_t^c f = P_t f + \int_0^t P_{t-s}(c_s P_s^c f) ds \quad (6.2)$$

$$P_t^c f = P_t f + \int_0^t c_s P_s^c(P_{t-s} f) ds. \quad (6.3)$$

Proof : In what follows, we use the fact that the semigroup and the perturbed semigroup are, under certain assumptions², solutions of PDEs (for a more detailed exposition, please consult the following chapter). In particular,

$$\begin{aligned} \frac{\partial}{\partial t} P_t f(x) &= \mathcal{L} P_t f(x) \\ \frac{\partial}{\partial t} P_t^c f(x) &= \mathcal{L} P_t^c f(x) + c(t, x) P_t^c f(x), \end{aligned}$$

¹This has already been discussed in previous chapters. Indeed, from this chapter we require only that $c \in C_b^{k-m-1}$.

²This is certainly the case when assuming that c and f are smooth and bounded

for $(t, x) \in (0, T] \times \mathbb{R}^N$. NB: $\mathcal{L}$ commutes with P_t on the space of smooth test functions.

$$\begin{aligned}
P_t^c f &= P_t f + [P_{t-s} P_s^c f]_{s=0}^{s=t} \\
&= P_t f + \int_0^t \frac{\partial}{\partial s} (P_{t-s} P_s^c f) ds \\
&= P_t f + \int_0^t [-\mathcal{L}(P_{t-s} P_s^c f) + P_{t-s} ((\mathcal{L} + c_s) P_s^c f)] ds \\
&= P_t f + \int_0^t P_{t-s} (c_s P_s^c f) ds,
\end{aligned}$$

The second part is proved in an analogous way

$$\begin{aligned}
P_t^c f &= P_t f + [P_s^c P_{t-s} f]_{s=0}^{s=t} \\
&= P_t f + \int_0^t \frac{\partial}{\partial s} (P_s^c P_{t-s} f) ds \\
&= P_t f + \int_0^t [(\mathcal{L} + c_s)(P_s^c P_{t-s} f) - P_s^c (\mathcal{L} P_{t-s} f)] ds \\
&= P_t f + \int_0^t c_s P_s^c (P_{t-s} f) ds.
\end{aligned}$$

■

This relationship shall form the basis of the analysis of the semigroup which has been perturbed by a potential. It is of particular use, as it allows us to use the known regularity results of the unperturbed semigroup.

FIRST DERIVATIVES

The representation of Lemma 6.4 allows a straightforward computation of first-order derivatives. Although, it is vital to note that derivatives along the diffusion part of the SDE (i.e. $V_1, \dots, V_d$) are the only ones we should expect to follow problem free. The reason for this is simple: those derivatives can be seen to explode at a rate $t^{-1/2}$, which is still integrable with respect to the Lebesgue measure on $[0, T]$. All other derivatives explode at a higher rate: t^{-1} or worse. For these derivatives we shall need to make some additional assumptions. In the meantime, we deal with derivatives along $V_1, \dots, V_d$:

$$\begin{aligned}
V_i P_t^c f &= V_i P_t f + \int_0^t V_i P_{t-s} (c_s P_s^c f) ds \\
&= t^{-1/2} \mathbb{E} [\Phi^i(t, x) f(X_t^x)] + \int_0^t (t-s)^{-1/2} \mathbb{E} [\Phi^i(t-s, x) c_s P_s^c f(X_{t-s}^x)] ds.
\end{aligned}$$

Analogously,

$$\begin{aligned} P_t^c V_i f &= P_t V_i f + \int_0^t c_s P_s^c (P_{t-s} V_i f) ds \\ &= t^{-1/2} \mathbb{E} \left[\tilde{\Phi}^i(t, x) f(X_t^x) \right] + \int_0^t (t-s)^{-1/2} c_s P_s^c \left(\mathbb{E}[\tilde{\Phi}^i(t-s, \cdot) f(X_{t-s}^{(\cdot)})] \right) ds. \end{aligned}$$

Hence we obtain the gradient bounds:

$$\begin{aligned} \|V_i P_t^c f\|_\infty &\leq C \frac{1}{\sqrt{t}} \|f\|_\infty \left[1 + t \|c\|_\infty e^{t\|c\|_\infty} \right] \\ \|P_t^c (V_i f)\|_\infty &\leq \tilde{C} \frac{1}{\sqrt{t}} \|f\|_\infty \left[1 + t \|c\|_\infty e^{t\|c\|_\infty} \right], \end{aligned}$$

where $\|c\|_\infty := \sup_{s \in [0, t]} \sup_{x \in \mathbb{R}^N} |c(s, x)|$. These $L^\infty(dx)$ bounds can be used to deduce corresponding $L^1(dx)$ bounds in a fashion we have now demonstrated numerous times. $L^p(dx)$ bounds for each $p \in (1, \infty)$ can then be deduced by employing the Riesz-Thorin interpolation theorem (again see for example Triebel [40]).

Proposition 6.5 For $p \in [1, \infty]$

$$\|V_i P_t^c f\|_{L^p(dx)} \leq C \frac{1}{\sqrt{t}} \|f\|_{L^p(dx)} \left[1 + t \|c\|_\infty e^{t\|c\|_\infty} \right] \quad (6.4)$$

$$\|P_t^c (V_i f)\|_{L^p(dx)} \leq \tilde{C} \frac{1}{\sqrt{t}} \|f\|_{L^p(dx)} \left[1 + t \|c\|_\infty e^{t\|c\|_\infty} \right]. \quad (6.5)$$

Although this gradient bound is interesting in itself, it may not be used to directly prove convergence of semigroup approximations, as the perturbed semigroup is not a linear function of c . We must relax the assumptions on f and c in a different way. We shall again use the relationships of Lemma 6.4. We shall make the following assumptions:

1. $f \in L^p(dx)$.
2. $c : [0, t] \times \mathbb{R}^N \rightarrow \mathbb{R}$ is continuous, bounded.

Our first task is to show that for $\{f_n\}_n \subset C_b^\infty(\mathbb{R}^N)$ and $\{c^n\}_n \subset C_b^\infty((0, t) \times \mathbb{R}^N)$ such that $\|f_n - f\|_{L^p(dx)} \rightarrow 0$ and $\|c^n - c\|_\infty \rightarrow 0$. It holds that $\{P_t^{c^n} f_n\}$ is Cauchy in L^p , with:

$$\|P_t^{c^n} f_n - P_t^c f\|_{L^p(dx)} \rightarrow 0.$$

Indeed, from Lemma 6.4

$$P_t^{c^n} f_n - P_t^{c^m} f_m = P_t(f_n - f_m) + \int_0^t P_{t-s} (c_s^n P_s^{c^n} f_n - c_s^m P_s^{c^m} f_m) ds.$$

Hence,

$$\begin{aligned}
\|P_t^{c^n} f_n - P_t^{c^m} f_m\|_{L^p} &\leq \|P_t(f_n - f_m)\|_{L^p} + \int_0^t \|P_{t-s}(c_s^n P^{c^n} f_n - c_s^m P^{c^m} f_m)\|_{L^p} ds \\
&\leq \|P_t(f_n - f_m)\|_{L^p} + \int_0^t \|P_{t-s}(c_s^n P^{c^n} f_n - c_s^m P^{c^m} f_m)\|_{L^p} ds \\
&\leq \|f_n - f_m\|_{L^p} + \int_0^t \|c_s^n P_s^{c^n} f_n - c_s^m P_s^{c^m} f_m\|_{L^p(dx)} ds \\
&\leq \|f_n - f_m\|_{L^p} + \int_0^t \|(c_s^n - c_s^m) P_s^{c^n} f_n\|_{L^p(dx)} ds \\
&\quad + \int_0^t \|c_s^m (P_s^{c^n} f_n - P_s^{c^m} f_m)\|_{L^p(dx)} ds \\
&\leq \|f_n - f_m\|_{L^p} + \int_0^t \|c^n - c^m\|_{L^\infty(dt \times dx)} \|P_s^{c^n} f_n\|_{L^p(dx)} ds \\
&\quad + \int_0^t \|c^m\|_{L^\infty(dt \times dx)} \|P_s^{c^n} f_n - P_s^{c^m} f_m\|_{L^p(dx)} ds \\
&\leq \|f_n - f_m\|_{L^p} + C_1 t \|c^n - c^m\|_{L^\infty(dt \times dx)} \\
&\quad + C_2 \int_0^t \|P_s^{c^n} f_n - P_s^{c^m} f_m\|_{L^p(dx)} ds,
\end{aligned}$$

where we have assumed, without loss of generality, that $\sup_{\substack{n \in \mathbb{N} \\ s \in [0, t]}} \|P_s^{c^n} f_n\|_{L^p} \leq C_1 < \infty$ and $\sup_{n \in \mathbb{N}} \|c^n\|_{L^\infty(dt \times dx)} \leq C_2 < \infty$. We may now use Gronwall's inequality, and the fact that $\exp(Ct) - 1 = O(t)$ for small times, to deduce that

$$\begin{aligned}
\|P_t^{c^n} f_n - P_t^{c^m} f_m\|_{L^p} &\leq \left[\|f_n - f_m\|_{L^p} + C_1 t \|c^n - c^m\|_{L^\infty(dt \times dx)} \right] \int_0^t \exp\{C_2 s\} ds \\
&\leq C_3 t \left[\|f_n - f_m\|_{L^p} + t \|c^n - c^m\|_{L^\infty(dt \times dx)} \right], \tag{6.6}
\end{aligned}$$

and hence the Cauchy property of the sequence in $L^p(dx)$. We can also make a similar calculation to deduce that the sequence of derivatives $\{V_i P_t^{c^n} f_n\}$ is Cauchy in L^p . We first consider the case for $p = \infty$ and then use this to prove for $p < \infty$:

$$\begin{aligned}
&\|V_i P_t^{c^n} f_n - V_i P_t^{c^m} f_m\|_\infty \\
&\leq \|V_i P_t(f_n - f_m)\|_\infty + \int_0^t \|V_i P_{t-s}(c_s^n P^{c^n} f_n - c_s^m P^{c^m} f_m)\|_\infty ds \\
&\leq C_1 \frac{1}{\sqrt{t}} \|f_n - f_m\|_\infty \\
&\quad + \int_0^t \frac{1}{\sqrt{t-s}} \|\mathbb{E}[\Phi^i(t-s, x)(c_s^n P^{c^n} f_n - c_s^m P^{c^m} f_m)(X_{t-s}^x)]\|_\infty ds \\
&\leq C_1 \frac{1}{\sqrt{t}} \|f_n - f_m\|_\infty + C_2 \int_0^t \frac{1}{\sqrt{t-s}} \|c_s^n P_s^{c^n} f_n - c_s^m P_s^{c^m} f_m\|_\infty ds
\end{aligned}$$

$$\begin{aligned}
&\leq C_1 \frac{1}{\sqrt{t}} \|f_n - f_m\|_\infty + C_2 \int_0^t \frac{1}{\sqrt{t-s}} \|(c_s^n - c_s^m) P_s^{c^n} f_n\|_\infty ds \\
&\quad + C_2 \int_0^t \frac{1}{\sqrt{t-s}} \|c_s^m (P_s^{c^n} f_n - P_s^{c^m} f_m)\|_\infty ds \\
&\leq C_1 \frac{1}{\sqrt{t}} \|f_n - f_m\|_\infty + C_2 \|c^n - c^m\|_{L^\infty(dt \times dx)} \int_0^t \frac{1}{\sqrt{t-s}} \|P_s^{c^n} f_n\|_\infty ds \\
&\quad + \|c^m\|_{L^\infty(dt \times dx)} \int_0^t \frac{1}{\sqrt{t-s}} \|P_s^{c^n} f_n - P_s^{c^m} f_m\|_\infty ds \\
&\leq C_1 \frac{1}{\sqrt{t}} \|f_n - f_m\|_\infty + \tilde{C}_1 \sqrt{t} \|c^n - c^m\|_{L^\infty(dt \times dx)} \\
&\quad + \tilde{C}_2 \int_0^t \frac{1}{\sqrt{t-s}} \|P_s^{c^n} f_n - P_s^{c^m} f_m\|_\infty ds.
\end{aligned}$$

Define $a_{n,m}(t) := \|V_i P_t^{c^n} f_n - V_i P_t^{c^m} f_m\|_\infty$. We use the bound (6.6) to deduce

$$\begin{aligned}
a_{n,m}(t) &\leq C_1 \frac{1}{\sqrt{t}} \|f_n - f_m\|_\infty + \tilde{C}_1 \sqrt{t} \|c^n - c^m\|_{L^\infty(dt \times dx)} \\
&\quad + \tilde{C}_2 \int_0^t \frac{1}{\sqrt{t-s}} \|P_s^{c^n} f_n - P_s^{c^m} f_m\|_\infty ds \\
&\leq C_1 \frac{1}{\sqrt{t}} \|f_n - f_m\|_\infty + \tilde{C}_1 \sqrt{t} \|c^n - c^m\|_{L^\infty(dt \times dx)} \\
&\quad + \tilde{C}_2 \int_0^t \frac{1}{\sqrt{t-s}} C_3 s \left[\|f_n - f_m\|_\infty + s \|c^n - c^m\|_{L^\infty(dt \times dx)} \right] ds \\
&\leq C_4 \frac{1}{\sqrt{t}} \|f_n - f_m\|_\infty + C \sqrt{t} \|c^n - c^m\|_{L^\infty(dt \times dx)},
\end{aligned}$$

which proves the Cauchy property. Note that we have used the following integral result (for $n = 1$):

$$\int_0^t \frac{s^{n-\frac{1}{2}}}{\sqrt{t-s}} ds = \frac{(2n-1)!!}{(2n)!!} \pi t^n.$$

Similar methods can be used to show the same type of bound for $b_{n,m}(t) := \|P_t^{c^n}(V_i f_n) - P_t^{c^m}(V_i f_m)\|_\infty$. This result may then be used to deduce a similar result for the $L^1(dx)$ norm, and hence, using a (Riesz-Thorin) interpolation argument, the bounds may be proved for the $L^p(dx)$ norm for each $p \in (1, \infty)$. These arguments are lengthy, and have already been exhibited numerous times, and are thus omitted.

Proposition 6.6

$$\begin{aligned}
\|P_t^{c^n}(V_i f_n) - P_t^{c^m}(V_i f_m)\|_{L^p} &\leq C_p \frac{1}{\sqrt{t}} \left[\|f_n - f_m\|_{L^p} + t \|c^n - c^m\|_{L^\infty(dt \times dx)} \right] \\
\|V_i P_t^{c^n} f_n - V_i P_t^{c^m} f_m\|_{L^p} &\leq \tilde{C}_p \frac{1}{\sqrt{t}} \left[\|f_n - f_m\|_{L^p} + t \|c^n - c^m\|_{L^\infty(dt \times dx)} \right].
\end{aligned}$$

SECOND ORDER DERIVATIVES

The situation for higher order derivatives is much more delicate. Indeed, the presence of the Lebesgue integral in the potential term means that integrability problems arise from the degenerate nature of the smoothness of the semigroup at time zero. In this case we shall need to make an extra assumption to deduce convergence of the second order derivatives. Namely, we shall **assume that the UFG condition holds for $m = 1$** . Although this condition can be compared to the uniform ellipticity condition, in as much as we require that the terms of diffusion matrix span the Lie algebra, one should note that **(UFG,1)** is still considerably more general. We shall further assume that $x \mapsto c(s, x)$ is globally Lipschitz continuous with Lipschitz constant, K_t , independent of $s \in [0, t]$. We begin by recalling the following calculation, which was first demonstrated in Chapter 1, and is a direct result of assuming the UFG condition:

$$V_{[\alpha]}(g \circ X_u)(x) = \sum_{\beta \in \mathcal{A}(m)} b_{\alpha, \beta}(u, x) V_{[\beta]} g(X_u^x).$$

In particular, under the assumption that the UFG condition holds with $m = 1$

$$\begin{aligned} V_i(P_s^c f \circ X_{t-s})(x) &= \sum_{\beta \in \mathcal{A}(1)} b_{i, \beta}(t-s, x) V_{[\beta]} P_s^c f(X_{t-s}^x) \\ &= \sum_{j=1}^d b_{i, j}(t-s, x) V_j P_s^c f(X_{t-s}^x). \end{aligned} \quad (6.7)$$

We now attempt to derive an integration by parts formula for $V_i V_j P_t^c f(x)$, for $1 \leq i, j \leq d$. Begin by assuming that $c \in C_b^\infty((0, t) \times \mathbb{R}^N)$ and $f \in C_b^\infty(\mathbb{R}^N)$, are both also uniformly bounded functions. These assumptions shall be relaxed later.

$$\begin{aligned} V_i V_j P_t^c f(x) &= V_i V_j P_t f + \int_0^t V_i V_j P_{t-s}(c_s P_s^c f) ds \\ &= \frac{1}{t} \mathbb{E}[f(X_t^x) \Phi_{i, j}(t, x)] + \int_0^t \frac{1}{\sqrt{t-s}} V_i \mathbb{E}\left[c_s P_s^c f(X_{t-s}^x) \tilde{\Phi}_j(t-s, x)\right] ds. \end{aligned}$$

At this point we are faced with a problem: another simple application of the integration by parts formula to the RHS leads to a term which is not integrable³. Instead, we seek to manually apply

³Namely, $\frac{1}{t-s}$

the operator V_i and use the differentiability of P_s^c . Observe, using (6.7):

$$\begin{aligned}
V_i V_j P_t^c f(x) &= \frac{1}{t} \mathbb{E} [f(X_t^x) \Phi_{i,j}(t, x)] + \int_0^t \frac{1}{\sqrt{t-s}} \mathbb{E} \left[V_i(c_s \circ X_{t-s})(x) P_s^c f(X_{t-s}^x) \tilde{\Phi}_j(t-s, x) \right] ds \\
&\quad + \int_0^t \frac{1}{\sqrt{t-s}} \mathbb{E} \left[c_s(X_{t-s}^x) V_i(P_s^c f \circ X_{t-s})(x) \tilde{\Phi}_j(t-s, x) \right] ds \\
&\quad + \int_0^t \frac{1}{\sqrt{t-s}} \mathbb{E} \left[c_s(X_{t-s}^x) P_s^c f(X_{t-s}^x) V_i \tilde{\Phi}_j(t-s, x) \right] ds \\
&= \frac{1}{t} \mathbb{E} [f(X_t^x) \Phi_{i,j}(t, x)] + \int_0^t \frac{1}{\sqrt{t-s}} \mathbb{E} \left[V_i(c_s \circ X_{t-s})(x) P_s^c f(X_{t-s}^x) \tilde{\Phi}_j(t-s, x) \right] ds \\
&\quad + \sum_{j=1}^d \int_0^t \frac{1}{\sqrt{t-s}} \mathbb{E} \left[c_s(X_{t-s}^x) b_{i,j}(t-s, x) V_j P_s^c f(X_{t-s}^x) \tilde{\Phi}_j(t-s, x) \right] ds \\
&\quad + \int_0^t \frac{1}{\sqrt{t-s}} \mathbb{E} \left[c_s(X_{t-s}^x) P_s^c f(X_{t-s}^x) V_i \tilde{\Phi}_j(t-s, x) \right] ds.
\end{aligned}$$

In particular,

$$V_i V_j P_t^c f(x) \leq C_1 \frac{1}{t} \|f\|_\infty + \int_0^t \frac{C_2}{\sqrt{t-s}} \|c\|_{W^{1,\infty}} \left[\|f\|_\infty + \sum_{j=1}^d \|V_j P_s^c f\|_\infty \right] ds. \quad (6.8)$$

And hence, recalling (6.4), we get that:

$$\begin{aligned}
V_i V_j P_t^c f(x) &\leq C_3 \frac{1}{t} \|f\|_\infty \left(1 + t^{\frac{3}{2}} \|c\|_{W^{1,\infty}} \right) \\
&\quad + \int_0^t \frac{C_4}{\sqrt{s(t-s)}} \|f\|_\infty \left[1 + s \|c\|_\infty e^{s\|c\|_\infty} \right] ds \\
&\leq C_3 \frac{1}{t} \|f\|_\infty \left(1 + t^{\frac{3}{2}} \|c\|_{W^{1,\infty}} \right) + C_4 \pi \|f\|_\infty \left[1 + t \|c\|_\infty e^{t\|c\|_\infty} \right] \\
&\leq C_5 \frac{1}{t} \|f\|_\infty \left(1 + t^{\frac{3}{2}} \|c\|_{W^{1,\infty}} + t^2 \|c\|_\infty e^{t\|c\|_\infty} \right) \\
&\leq C_6 \frac{1}{t} \|f\|_\infty \left(1 + t^{\frac{3}{2}} \|c\|_{W^{1,\infty}} \right).
\end{aligned}$$

Similar arguments may be employed to deduce gradient bounds for $P_t^c(V_i V_j f)(\cdot)$, and then usual arguments using the Riesz-Thorin interpolation theorem may be applied to deduce the following

Proposition 6.7

$$\|V_i V_j P_t^c f\|_{L^p(dx)} \leq C_p \frac{1}{t} \|f\|_{L^p(dx)} \left(1 + t^{\frac{3}{2}} \|c\|_{W^{1,\infty}} \right).$$

Again, we note that although this gradient bound is interesting in itself, it may not be used to directly prove convergence of semigroup approximations, as the perturbed semigroup is not a linear function of c . We seek to relax the assumptions on f and c in a different way. We make the following assumptions on f and c :

1. $f \in L^p(dx)$.
2. $c \in W^{1,\infty}((0, T) \times \mathbb{R}^N)$.

Now let $\{f_n\}_n \subset C_b^\infty$ such that $f_n \xrightarrow{\text{in } L^p} f$, and let $\{c^n\}_n \subset C_b^\infty((0, t) \times \mathbb{R}^N)$ such that $c_n \xrightarrow{\text{unif.}} c$. We aim to show that the sequence of second order derivatives, $\{V_i V_j P_t^{c^n} f_n\}_n$ is **Cauchy in $L^p(dx)$** , for each $t > 0$. In a similar way to the previous section, and using the calculations on the previous page as a guideline, it can be shown that this is the case for $p = \infty$:

$$\begin{aligned}
& \left\| V_i V_j P_t^{c^n} f_n - V_i V_j P_t^{c^m} f_m \right\|_\infty \\
& \leq \left\| V_i V_j P_t (f_n - f_m) \right\|_\infty + \int_0^t \left\| V_i V_j P_{t-s} (c_s^n P^{c^n} f_n - c_s^m P^{c^m} f_m) \right\|_\infty ds \\
& \leq C_1 \frac{1}{t} \|f_n - f_m\|_\infty \\
& \quad + \int_0^t \frac{1}{\sqrt{t-s}} \left\| V_i \mathbb{E} \left[\Phi_j(t-s, x) (c_s^n P^{c^n} f_n - c_s^m P^{c^m} f_m) (X_{t-s}^x) \right] \right\|_\infty ds \\
& \leq \dots \\
& \leq C_1 \frac{1}{t} \|f_n - f_m\|_\infty + C_2 \sqrt{t} [\|f_n - f_m\|_\infty + \|c^n - c^m\|_{W^{1,\infty}}] \\
& \quad + C_3 \int_0^t \frac{1}{\sqrt{t-s}} \left[\left\| V_i (P_s^{c^n} f_n - P_s^{c^m} f_m) \right\|_\infty + \sup_n \left\| V_i P_s^{c^n} f_n \right\|_\infty \|c^n - c^m\|_{L^\infty(dt \times dx)} \right] ds \\
& \leq C_1 \frac{1}{t} \|f_n - f_m\|_\infty + C_2 \sqrt{t} [\|f_n - f_m\|_\infty + \|c^n - c^m\|_{W^{1,\infty}}] \\
& \quad + C_4 \int_0^t \frac{1}{\sqrt{s(t-s)}} [\|f_n - f_m\|_{L^p} + \|c^n - c^m\|_{L^\infty(dt \times dx)}] ds, \\
& \leq C_5 \frac{1}{t} [\|f_n - f_m\|_\infty + t^{3/2} \|c^n - c^m\|_{W^{1,\infty}} + t \|c^n - c^m\|_{L^\infty(dt \times dx)}] \\
& \leq C_6 \frac{1}{t} [\|f_n - f_m\|_\infty + t \|c^n - c^m\|_{W^{1,\infty}}].
\end{aligned}$$

Note, we have assumed w.l.o.g. that $\sup_n \|V_i P_s^{c^n} f_n\|_\infty \leq C s^{-1/2}$. This establishes the Cauchy property of the sequence of second order derivatives. Similar results may be proved for the sequences $\{P_t^{c^n} (V_i V_j f_n)\}$, $\{V_i P_t^{c^n} (V_j f_n)\}$, which in turn can be used to the case $p = 1$; in the same way which has been demonstrated before.

If $f \in W^{1,\infty}((0, T) \times \mathbb{R}^N)$ then we can do much better than this. In particular, the rate at which the L^∞ -norm of derivatives explodes as $t \rightarrow 0$, is slower. Indeed, by using the gradient bound for $V_i V_j P_t f$ where f is globally Lipschitz continuous, we are able to deduce explosion is of order $t^{-1/2}$, rather than t^{-1} . To show this we shall use

$$\|V_i V_j P_t f\|_\infty \leq C \frac{1}{\sqrt{t}} \|\nabla f\|_\infty \quad \text{rather than} \quad \|V_i V_j P_t f\|_\infty \leq C \frac{1}{t} \|f\|_\infty$$

Then, by following identical steps as in (6.4), it may be shown that all of the above leads to the following result:

Proposition 6.8

$$\begin{aligned}\left\|V_i V_j (P_t^{c^n} f_n - P_t^{c^m} f_m)\right\|_{L^p} &\leq \frac{C_{1,p}}{t} [\|f_n - f_m\|_{L^p} + t \|c^n - c^m\|_{W^{1,\infty}}] \\ \left\|V_i (P_t^{c^n} V_j f_n - P_t^{c^m} V_j f_m)\right\|_{L^p} &\leq \frac{C_{2,p}}{t} [\|f_n - f_m\|_{L^p} + t \|c^n - c^m\|_{W^{1,\infty}}] \\ \left\|P_t^{c^n} (V_i V_j f_n) - P_t^{c^m} (V_i V_j f_m)\right\|_{L^p} &\leq \frac{C_{3,p}}{t} [\|f_n - f_m\|_{L^p} + t \|c^n - c^m\|_{W^{1,\infty}}].\end{aligned}$$

Moreover,

$$\begin{aligned}\left\|V_i V_j (P_t^{c^n} f_n - P_t^{c^m} f_m)\right\|_{L^\infty} &\leq \frac{\tilde{C}_{1,p}}{\sqrt{t}} [\|f_n - f_m\|_{W^{1,\infty}} + t \|c^n - c^m\|_{W^{1,\infty}}] \\ \left\|V_i (P_t^{c^n} V_j f_n - P_t^{c^m} V_j f_m)\right\|_{L^\infty} &\leq \frac{\tilde{C}_{2,p}}{\sqrt{t}} [\|f_n - f_m\|_{W^{1,\infty}} + t \|c^n - c^m\|_{W^{1,\infty}}] \\ \left\|P_t^{c^n} (V_i V_j f_n) - P_t^{c^m} (V_i V_j f_m)\right\|_{L^\infty} &\leq \frac{\tilde{C}_{3,p}}{\sqrt{t}} [\|f_n - f_m\|_{W^{1,\infty}} + t \|c^n - c^m\|_{W^{1,\infty}}].\end{aligned}$$

Remark 6.9 *In the absence of a bound of the type*

$$\|V_i V_j P_t f\|_{L^p} \leq C \frac{1}{\sqrt{t}} \|\nabla f\|_{L^p},$$

for $p < \infty$, it is not obvious to the author how to extend the gradient bounds for the second order derivatives of the potential term to L^p results, for $p < \infty$.

6.3. The Lagrangian term

In this section we extend the conversation to semigroups which not only involve a potential term, but also a Langrangian term. The methods we employ to cope with this term are heavily influenced by those of the previous section. We start by defining what is meant by a Langrangian term. Define:

$$L_g^c(t, x) := \mathbb{E} \left[\int_0^t g(t-s, X_s^x) \exp \left\{ \int_0^s c(u, X_u^x) du \right\} \right]. \quad (6.9)$$

We call L_g^c a Langrangian. To see why such a term is important, we make a brief reference to parabolic PDEs and the Feynman-Kac formula. Note that the Feynman-Kac formula and existence and uniqueness to solutions of certain parabolic PDEs shall be left to the next chapter. For now we note that, under certain conditions, we may express the solution to the initial value problem:

$$\begin{aligned}\frac{\partial u}{\partial t} + cu &= \mathcal{L}u + g, & \text{in } (0, T] \times \mathbb{R}^N, \\ u &= f, & \text{in } \{0\} \times \mathbb{R}^N,\end{aligned}$$

where $\mathcal{L} = \sum_{i=1}^d (V_i)^2 + V_0$ is a second-order differential operator, as

$$u(t, x) = P_t^c f(x) + L_g^c(t, x).$$

In this section we take the attitude that the study of L is interesting in its own right, but rest easy on the knowledge that L_g^c has a important role to play in PDE theory. In this section we shall again assume that the vector fields: $V_0, \dots, V_d$ are **smooth and uniformly bounded with bounded derivatives**, for simplicity, although they can be extended to the non-uniformly bounded, general differentiability case.

Our first observation, through a simple application of Fubini's theorem⁴, is the following:

$$L_g^c(t, x) = \int_0^t P_s^c g_{t-s}(x) ds,$$

where $g_u(\cdot) := g(u, \cdot)$. This simple observation makes the resultant analysis far easier. Indeed, we may control the behaviour of L through the behaviour of the perturbed semigroup P^c .

FIRST ORDER DERIVATIVES

We start by assuming that c and g are smooth, bounded functions. These conditions are subsequently relaxed by limit arguments.

Proposition 6.10 *Let $p > 2$ and $i \in \{1, \dots, d\}$. Then*

$$\|V_i L_g^c(t, \cdot)\|_{L^p(dx)} \leq C t^{p^*} \|g\|_{L^p(dt \times dx)} \left[1 + t \|c\|_\infty e^{t\|c\|_\infty}\right] \quad (6.10)$$

$$\|L_{V_i g}^c(t, x)\|_{L^p(dx)} \leq \tilde{C} t^{p^*} \|g\|_{L^p(dt \times dx)} \left[1 + t \|c\|_\infty e^{t\|c\|_\infty}\right] \quad (6.11)$$

for each $t \in [0, T]$, where $p^* = \frac{1}{2} - \frac{1}{p}$.

Proof :

$$\begin{aligned} \|V_i L_g^c(t, \cdot)\|_{L^p(dx)} &= \left\| V_i \int_0^t P_s^c g_{t-s}(x) ds \right\|_{L^p} \\ &= \left\| \int_0^t V_i P_s^c g_{t-s}(x) ds \right\|_{L^p} \\ &\leq \int_0^t \|V_i P_s^c g_{t-s}(x)\|_{L^p} ds \\ &\leq \int_0^t C \frac{1}{\sqrt{s}} \|g_{t-s}\|_{L^p(dx)} \left(1 + s \|c\|_\infty e^{s\|c\|_\infty}\right) ds \\ &\leq \left(\int_0^t \|g_{t-s}\|_{L^p}^p ds \right)^{1/p} \left(\int_0^t \left[C \frac{1}{\sqrt{s}} \left(1 + s \|c\|_\infty e^{s\|c\|_\infty}\right) \right]^q ds \right)^{1/q} \\ &\leq C t^{p^*} \|g\|_{L^p(dt \times dx)} \left[1 + t \|c\|_\infty e^{t\|c\|_\infty}\right], \end{aligned}$$

where we have used Hölder's inequality, and inequality (6.4) where $p^* = \frac{1}{q} - \frac{1}{2} = \frac{1}{2} - \frac{1}{p}$. The second inequality is proved analogously using (6.5). ■

⁴Indeed, at this early stage we assume enough regularity on c , g , etc, to make such an application possible.

Again, we observe that the non-linearity of the perturbed semigroup, as a function of c and g , means that we need to work harder to deduce regularity properties than we had to in previous chapters.

Proposition 6.11 *Let $p > 2$ and $i \in \{1, \dots, d\}$. Let $\{c^n\}, \{g^n\} \subset C^\infty((0, T) \times \mathbb{R}^N)$. Then*

$$\begin{aligned} \|(L_{V_i g^n}^{c^n} - L_{V_i g^m}^{c^m})(t, \cdot)\|_{L^p(dx)} &\leq C_p t^{p^*} [\|g^n - g^m\|_{L^p} + t \|c^n - c^m\|_{L^\infty}] \\ \|(V_i L_{g^n}^{c^n} - V_i L_{g^m}^{c^m})(t, \cdot)\|_{L^p(dx)} &\leq \tilde{C}_p t^{p^*} [\|g^n - g^m\|_{L^p} + t \|c^n - c^m\|_{L^\infty}]. \end{aligned}$$

Moreover, the sequences $\{L_{V_i g^n}^{c^n}(t, \cdot)\}, \{V_i L_{g^n}^{c^n}(t, \cdot)\}$ are Cauchy in $L^p(dx)$ for each $t \in (0, T]$, when both $\{c^n\}$ and $\{g^n\}$ are. In this case they have limits which are themselves $L^p(dx)$ functions.

NB: Proposition 6.11 allows us to make better use of Proposition 6.10, as we may now make sense of $V_i L_g^c$ and $L_{V_i g}^c$ as limits. In particular, Proposition 6.10 may be extended to cope with

1. $g \in L^p([0, T] \times \mathbb{R}^N)$, for $p > 2$.
2. $c \in C([0, T] \times \mathbb{R}^N)$, bounded.

SECOND ORDER DERIVATIVES

If one attempts to take second order derivatives, then one runs up against similar problems to those in the previous section. In particular, the singularity for the second order of derivatives is of the form t^{-1} , which is not integrable. In this case, the problems are even worse, and as well as making the assumption that **the UFG condition holds for $m = 1$** , we must make stronger assumptions on the function g . Let us begin by assuming that g is a smooth, bounded function.

Then

$$\begin{aligned} \|V_i V_j L_g^c(t, \cdot)\|_{L^p(dx)} &= \left\| V_i V_j \int_0^t P_s^c g_{t-s}(x) ds \right\|_{L^p} \\ &= \left\| \int_0^t V_i V_j P_s^c g_{t-s}(x) ds \right\|_{L^p} \\ &\leq \int_0^t \|V_i V_j P_s^c g_{t-s}(x)\|_{L^p} ds. \end{aligned} \tag{6.12}$$

At this point we recall (6.8), and note that by adapting the calculations immediately preceding this inequality, we may show a similar result for g

$$V_i V_j P_s^c g_{t-s}(x) \leq C_1 \frac{1}{\sqrt{s}} \|\nabla g\|_\infty + \int_0^s \frac{C_2}{\sqrt{s-u}} \|c\|_{W^{1,\infty}} \left[\|g\|_\infty + \sum_{j=1}^d \|V_j P_u^c g_{t-s}\|_\infty \right] du. \tag{6.13}$$

The only difference between (6.13) and (6.8) is that we have applied the gradient bound for $V_i V_j P_t f$ where f is globally Lipschitz continuous. That is, we have used

$$\|V_i V_j P_t f\|_\infty \leq C \frac{1}{\sqrt{t}} \|\nabla f\|_\infty \quad \text{rather than} \quad \|V_i V_j P_t f\|_\infty \leq C \frac{1}{t} \|f\|_\infty.$$

This step uses the integrability of $s^{-1/2}$ and avoids the fact the non-integrability of s^{-1} , at the cost of assuming g is globally Lipschitz continuous. Indeed, by following similar steps to (6.4), it may be shown that

$$\begin{aligned}
V_i V_j P_s^c g_{t-s}(x) &\leq C_3 \frac{1}{\sqrt{s}} \|\nabla g\|_\infty \left(1 + s \|c\|_{W^{1,\infty}}\right) \\
&\quad + \int_0^s \frac{C_4}{\sqrt{u(s-u)}} \|g\|_\infty \left[1 + u \|c\|_\infty e^{u\|c\|_\infty}\right] du \\
&\leq C_3 \frac{1}{\sqrt{s}} \|\nabla g\|_\infty \left(1 + s \|c\|_{W^{1,\infty}}\right) \\
&\quad + C_4 \pi \|g\|_\infty \left[1 + s \|c\|_\infty e^{s\|c\|_\infty}\right] \\
&\leq C_5 \frac{1}{\sqrt{s}} (\|g\|_\infty + \|\nabla g\|_\infty) \left(1 + s \|c\|_{W^{1,\infty}} + s^{3/2} \|c\|_\infty e^{s\|c\|_\infty}\right) \\
&\leq C_6 \frac{1}{\sqrt{s}} \|g\|_{W^{1,\infty}} \left(1 + s \|c\|_{W^{1,\infty}}\right).
\end{aligned}$$

Substituting the above into (6.12) gives

$$\begin{aligned}
\|V_i V_j L_g^c(t, \cdot)\|_{L^\infty(dx)} &= \left\| V_i V_j \int_0^t P_s^c g_{t-s}(x) ds \right\|_{L^\infty} \\
&\leq \int_0^t \|V_i V_j P_s^c g_{t-s}(x)\|_{L^\infty} ds \\
&\leq C \sqrt{t} \|g\|_{W^{1,\infty}} \left(1 + t \|c\|_{W^{1,\infty}}\right).
\end{aligned}$$

The non-linearity of the perturbed semigroup, as a function of c and g , means that regularity properties do not follow immediately from these bounds. Let $\{g^n\}_n, \{c^n\}_n \subset C_b^\infty((0, T) \times \mathbb{R}^N)$ such that $g^n \xrightarrow{\text{unif.}} g$, and $c^n \xrightarrow{\text{unif.}} c$, respectively. We aim to show that the sequence of second order derivatives, $\{V_i V_j L_{g^n}^{c^n}(t, \cdot)\}_n$ is Cauchy in $L^\infty(dx)$ for each $t > 0$.

$$\begin{aligned}
\|V_i V_j L_{g^n}^{c^n}(t, \cdot) - V_i V_j L_{g^m}^{c^m}(t, \cdot)\|_\infty &= \left\| V_i V_j \int_0^t P_s^{c^n} g_{t-s}^n(\cdot) ds - V_i V_j \int_0^t P_s^{c^m} g_{t-s}^m(\cdot) ds \right\|_\infty \\
&= \left\| \int_0^t V_i V_j \left(P_s^{c^n} g_{t-s}^n(\cdot) - P_s^{c^m} g_{t-s}^m(\cdot) \right) ds \right\|_\infty \\
&= \int_0^t \left\| V_i V_j \left(P_s^{c^n} g_{t-s}^n(\cdot) - P_s^{c^m} g_{t-s}^m(\cdot) \right) \right\|_\infty ds \\
&\leq \int_0^t \frac{C_p}{\sqrt{s}} [\|g^n - g^m\|_{W^{1,\infty}} + s \|c^n - c^m\|_{W^{1,\infty}}] ds \\
&\leq C_{1,p} \sqrt{t} [\|g^n - g^m\|_{W^{1,\infty}} + t \|c^n - c^m\|_{W^{1,\infty}}],
\end{aligned}$$

where we have used the gradient bounds of Proposition 6.8. Similar bounds may be proved for $\{V_i L_{V_j g^n}^{c^n}(t, \cdot)\}_n$ and $\{L_{V_i V_j g^n}^{c^n}(t, \cdot)\}_n$, to give the following result.

Proposition 6.12

$$\begin{aligned}\left\|V_i V_j L_{g^n}^{c^n}(t, \cdot) - V_i V_j L_{g^m}^{c^m}(t, \cdot)\right\|_\infty &\leq C_{1,p} \sqrt{t} [\|g^n - g^m\|_{W^{1,\infty}} + t \|c^n - c^m\|_{W^{1,\infty}}] \\ \left\|V_i L_{V_j g^n}^{c^n}(t, \cdot) - V_i L_{V_j g^m}^{c^m}(t, \cdot)\right\|_\infty &\leq C_{2,p} \sqrt{t} [\|g^n - g^m\|_{W^{1,\infty}} + t \|c^n - c^m\|_{W^{1,\infty}}] \\ \left\|L_{V_i V_j g^n}^{c^n}(t, \cdot) - L_{V_i V_j g^m}^{c^m}(t, \cdot)\right\|_\infty &\leq C_{3,p} \sqrt{t} [\|g^n - g^m\|_{W^{1,\infty}} + t \|c^n - c^m\|_{W^{1,\infty}}].\end{aligned}$$

Remark 6.9 is also relevant for the second derivatives of the Lagrangian term. We are not able to show that the above Proposition holds for general L^p norms.

6.4. Relaxing the V_0 condition

In the introductory section it was outlined how ‘ultracontractivity’ of the diffusion semigroup was used to prove fast convergence of the so-called KLV method. The proof of this convergence was based upon an iterated Stratonovich-type Taylor expansion on a test function $f(X_t^x)$. Eventually this Taylor expansion was applied to the diffusion semigroup $P_{T-t}f(X_t^x)$. This naïve expansion necessitated an extra assumption: the ‘ V_0 condition’, was first introduced in Crisan and Ghazali [8].

In this section we demonstrate how, given the fact that the test function is the diffusion semigroup, which is the solution of the Cauchy problem, a different expansion is more prudent. We then show how this new expansion fits better the regularity of the semigroup and completely relax the UFG condition.

Assume $g \in C_b^\infty([0, T] \times \mathbb{R}^N)$. Then, by applying Itô’s lemma for Stratonovich integrals, it is easy to see:

$$g(T-t, X_t^x) = g(T, x) + \int_0^t (V_0 - \partial_t) g(T-s, X_s^x) ds + \sum_{i=1}^d \int_0^t V_i g(T-s, X_s^x) \circ dB_s^i$$

This equation may be iterated to obtain an expansion. To this end, define the following vector fields on $[0, T] \times \mathbb{R}^N$:

$$\begin{aligned}\tilde{V}_i &:= V_i, \quad i = 1, \dots, d. \\ \tilde{V}_0 &:= V_0 - \partial_t\end{aligned}$$

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

$$g(T-t, X_t^x) = \sum_{\alpha \in \mathcal{A}(m)} (\tilde{V}_\alpha g)(T, x) \hat{B}_t^{\circ \alpha} + R_m(t, x, g),$$

where $\tilde{V}_\alpha = \tilde{V}_{\alpha_n} \dots \tilde{V}_{\alpha_1}$, and

$$R_m(t, x, g) = \sum_{\substack{\alpha \in \mathcal{A}(m) \\ i=1, \dots, d \\ \text{s.t. } \alpha * i \notin \mathcal{A}(m)}} \int_{0 < t_1 < \dots < t_{k+1} < t} \tilde{V}_{\alpha * i} g(T - t_0, X_{t_0}^x) \circ dB_{t_1}^{\alpha_1} \dots \circ dB_{t_k}^{\alpha_k} \circ dB_{t_{k+1}}^i.$$

It is then straightforward to show that:

$$\left\| \sqrt{\mathbb{E}[R_m(t, \cdot, g)^2]} \right\|_\infty \leq C \sum_{j=m+1}^{m+2} t^{j/2} \sup_{\substack{\alpha \in \mathcal{A}(j) \setminus \mathcal{A}(j-1) \\ t_0 \in [0, t]}} \left\| \tilde{V}_\alpha g(T - t_0, \cdot) \right\|_\infty.$$

As a result of this, and a similar bound for the expectation of the remainder taken with respect to the cubature measure, the following is easy to show (cf Proposition 1.24)

Proposition 6.13

$$\left\| (\mathbb{E} - \mathbb{E}_{\mathbb{Q}_t}) \left[g(T - t, X_t^{(\cdot)}) \right] \right\|_\infty \leq C \sum_{j=m+1}^{m+2} t^{j/2} \sup_{\substack{\alpha \in \mathcal{A}(j) \setminus \mathcal{A}(j-1) \\ t_0 \in [0, t]}} \left\| \tilde{V}_\alpha g(T - t_0, \cdot) \right\|_\infty. \quad (6.14)$$

Proof : This is a simple adaptation of the proof in Lyons and Victoir [29] for the error of the cubature measure applied to test functions which are time-dependent. ■

The above is an upper bound for the error of a finite measure based on a single application of the cubature formula. We already know from the introduction, that iterated applications of the cubature over a partition proves to be a far more efficient. The Markovian property of the cubature formula and the semigroup property of the diffusion, allow us to deduce far tighter upper bounds than (6.14). The difference between what is done here and what was used in Lyons and Victoir [29], is that we shall take advantage of the fact that the diffusion semigroup is a function of space **and** time. It is smooth along the vector fields $V_1, \dots, V_d$, but is also smooth along the vector field $\partial_t - V_0$, as it is the solution of the Cauchy problem.

Proposition 6.14 *The KLV approximation satisfies*

$$\begin{aligned} & \sup_{x \in \mathbb{R}^N} \left| P_T f(x) - \mathbb{E}_{KLV(\mathcal{D}, x)} f \right| \\ & \leq C \sum_{i=1}^k \sum_{j=m+1}^{m+2} s_i^{j/2} \sup_{\substack{\alpha \in \mathcal{A}(j) \setminus \mathcal{A}(j-1) \\ t_0 \in [0, s_i]}} \left\| \tilde{V}_\alpha P_{T-t_{i-1}-t_0} f(\cdot) \right\|_\infty. \end{aligned}$$

Proof : The proof is the same as that of Lyons and Victoir [29] or Litterer and Lyons [28]). The only difference is in the application of Proposition 6.13, rather than the usual (time-independent) Stratonovich Taylor expansion. ■

Corollary 6.15 *Under the assumption of the UFG condition, there holds the following gradient*

bounds:

$$\left\| \tilde{V}_\alpha P_{T-t_{i-1}-t_0} f(\cdot) \right\|_\infty \leq C \frac{(T-t_{i-1}-t_0)^{1/2}}{(T-t_{i-1}-t_0)^{\|\alpha\|/2}} \|\nabla f\|_\infty. \quad (6.15)$$

And hence

$$\sup_{t_0 \in [0, s_i]} \left\| \tilde{V}_\alpha P_{T-t_{i-1}-t_0} f(\cdot) \right\|_\infty \leq C \frac{1}{(T-t_i)^{(\|\alpha\|-1)/2}} \|\nabla f\|_\infty. \quad (6.16)$$

Proof : The idea is to use Theorem 7.4 - the fact that $P_t f(x)$ provides the solution to the Cauchy problem - to deduce gradient bounds for $P_t f(x)$ along $\tilde{V}_0 = V - 0 - \partial_t$. Hence, for each $(t, x) \in (0, T] \times \mathbb{R}^N$

$$(\partial_t - V_0)P_t f(x) = \left(\sum_{i=1}^d V_i^2 \right) P_t f(x) = Ct^{-1} \mathbb{E} [f(X_t^x) \Phi(t, x)],$$

by using the gradient bounds we obtained in the second chapter for $P_t f$, along the directions $V_1, \dots, V_d$. This means that we obtain the following gradient bounds. This fact may be used iteratively to obtain (6.15) and (6.16). ■

This lemma may be applied together with Proposition 6.13 to obtain the known convergence rates for the KLV method.

7. The Feynman-Kac connection revisited

In this chapter we bring the derived regularity results to bear on the Cauchy problem. A great motivation for studying the diffusion semigroup is that it offers a representation of the solution of the Cauchy problem, when such a solution exists. The latter point emphasises that, implicitly one is traditionally forced to rely on existence criteria for the solution of the Cauchy problem (see Karatzas and Shreve [17]). In this chapter, we attempt to change that. The results in the section are a result of interesting discussions with Marta Sanz-Solé and, in particular, Francois Delarue and Dan Crisan.

It would verge on the negligent to discuss the diffusion semigroup without providing a full context and motivation of its place within the theory of parabolic PDEs. Surely one should motivate the study of the diffusion semigroup through its representation as the solution of a PDE, via the Feynman-Kac connection? However, for most of this thesis, we didn't so much as state the Feynman-Kac formula, and discussed neither existence nor uniqueness to the corresponding partial differential equation. *Why have we been so negligent until now?* These issues have been deliberately ignored as the regularity results **can be brought to bear** on the Feynman-Kac formula, and the question of existence and uniqueness of the solution of a parabolic PDE. In this chapter we give sufficient conditions for existence of a solution to the Cauchy problem, and then show that these conditions also provide uniqueness. We begin by discussing what is meant by saying a function u a solution of the Cauchy problem.

Let $V_0, \dots, V_d$ be vector fields on $\mathbb{R}^N$. Define the parabolic operator $\frac{\partial}{\partial t} - \mathcal{L}^c$ by its action on the function u ,

$$\left(\frac{\partial}{\partial t} - \mathcal{L}^c \right) u := \frac{\partial u}{\partial t} - \frac{1}{2} \sum_{i=1}^d V_i^2 u - V_0 u - cu.$$

Note that we have implicitly assumed that u is *sufficiently differentiable* for this action by $\frac{\partial}{\partial t} - \mathcal{L}^c$ to make sense. Assume $c, g : (0, T] \times \mathbb{R}^N \rightarrow \mathbb{R}$. Then we may define the parabolic PDE, known as 'the Cauchy problem'

Definition 7.1 (The Cauchy problem) *The Cauchy problem is the task of finding a function u :*

$[0, T] \times \mathbb{R}^N \rightarrow \mathbb{R}$, such that u 'satisfies' the parabolic partial differential equation:

$$\frac{\partial u}{\partial t} = \mathcal{L}u + cu + g, \quad \text{in } (0, T] \times \mathbb{R}^N, \quad (7.1)$$

$$u = f, \quad \text{in } \{0\} \times \mathbb{R}^N. \quad (7.2)$$

We shall often write $\mathcal{L}^c := \mathcal{L} + c$, for compactness of notation. We consider two different notions of a solution to the Cauchy problem.

Strong solution Let $u : [0, T] \times \mathbb{R}^N \rightarrow \mathbb{R}$ be such that for some $p \in [1, \infty)$

1. $u \in C([0, T] \times \mathbb{R}^N)$.
2. $V_i u, V_i^2 u \in C((0, T] \times \mathbb{R}^N)$ for $i = 1, \dots, d$.
3. u satisfies (7.1) and (7.2).

Weak solution Let $u : [0, T] \times \mathbb{R}^N \rightarrow \mathbb{R}$ be such that

1. $u \in L^1_{loc}([0, T] \times \mathbb{R}^N)$.
2. For each $\epsilon > 0$ there holds $V_i u, V_i^2 u \in L^1_{loc}((\epsilon, T] \times \mathbb{R}^N)$ for $i = 1, \dots, d$.
3. u satisfies (7.2), and for each $\varphi \in C_c^\infty((0, T) \times \mathbb{R}^N)$ there holds

$$\int_0^\infty \int_{\mathbb{R}^N} [(-\partial_t - (\mathcal{L}^c)^*) \varphi](t, x) u(t, x) dx dt = \int_0^\infty \int_{\mathbb{R}^N} \varphi(t, x) g(t, x) dx dt, \quad (7.3)$$

where $(\mathcal{L}^c)^* = \frac{1}{2} \sum_{i=1}^d V_i^2 - \tilde{V}_0 + \tilde{c}$, and

$$\begin{aligned} \tilde{V}_0 &= V_0 - \frac{1}{2} \sum_{i=1}^d \operatorname{div} (V_i) V_i, \\ \tilde{c} &= c - \operatorname{div} (V_0) + \frac{1}{2} \sum_{i=1}^d V_i \operatorname{div} (V_i) + \frac{1}{2} \sum_{i=1}^d (\operatorname{div} V_i)^2, \end{aligned}$$

is the formal adjoint of $\mathcal{L}^c$.

4. There holds weak continuity at zero. i.e. for each $\rho \in C_0^\infty(\mathbb{R}^N)$

$$\lim_{t \downarrow 0} \int_{\mathbb{R}^N} \rho(x) (u(t, x) - f(x)) dx = 0.$$

We may write $\mathbf{C}(c, f, g)$ as a shorthand for this problem.

Remark 7.2

1. In the traditional notion of a strong solution one requires $u \in C^{1,2}((0, T) \times \mathbb{R}^n)$. If there has been one particular underlying theme in this thesis, it is that this cannot be expected in general and so the definition has been generalised.
2. Note that we require only the derivatives $V_i u$ and $V_i V_j u$ to be, in differing senses, well-defined. The fact that u satisfies either (7.1), (7.2), if a strong solution, and (7.2), (7.3), if a weak solution, means that all required derivatives (including $\partial_t - V_0$) are also well-defined¹.
3. Although the two notions of solution in the above may look very different, they are in fact very similar. For a strong solution we merely have the extra assumption of continuity of derivatives. Indeed, it is easy to verify that a strong solution is also a weak solution.
4. The notion of weak continuity at zero is called such, as it is a weaker version of strong continuity of a process u . Indeed, if there were to be a uniform bound for the continuity of 4. over ρ such that $\|\rho\|_\infty \leq 1$, then one would have the usual notion of strong continuity.

7.1. Existence of solutions

Most of the studies of this operator focus on the infinitesimal generator of the diffusion semigroup, $\mathcal{L}$. In particular, on sufficient conditions to guarantee differentiability with respect to this operator. In most situations this is a prudent decision. However, as we have alluded to in the first remark above, a problem arises outside a Hörmander setting. Indeed, the work in this thesis has shown that one should not, even under very general assumptions on the Lie algebra, expect to get differentiability along V_0 . This is no coincidence. However, we can expect to get differentiability along:

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

It thus makes more sense to base the search for a solution around this fact. Indeed, we have the following

Theorem 7.3 (Existence of a strong solution)

Define

$$\begin{aligned} v(t, x) := & \mathbb{E} \left[f(X_t^x) \exp \left\{ \int_0^t c(s, X_s^x) ds \right\} \right] \\ & + \mathbb{E} \left[\int_0^t g(t-s, X_s^x) \exp \left\{ \int_0^s c(u, X_u^x) du \right\} ds \right]. \end{aligned} \quad (7.4)$$

Then

¹Provided we assume certain conditions on c and g

1. If $f \in C(\mathbb{R}^N)$ is globally Lipschitz continuous, $c \in C([0, T] \times \mathbb{R}^N)$ is bounded, and $g \in C([0, T] \times \mathbb{R}^N)$ is also Lipschitz continuous, and $(UFG, 1)$ is satisfied, then v is a strong solution to $\mathbf{C}(c, f, g)$.
2. If $g \equiv 0$, $c \equiv 0$, $f \in C(\mathbb{R}^N)$ is constant at infinity, and (UFG, m) is satisfied for some $m \in \mathbb{N}$, then v is a strong solution to $\mathbf{C}(c, f, g)$.

Theorem 7.4 (Existence of a weak solution)

Define v as in Theorem 7.3. Then

1. If $g \equiv 0$, $f \in L^1(\mathbb{R}^N)$, $c \in C([0, T] \times \mathbb{R}^N)$ is globally Lipschitz continuous, and $(UFG, 1)$ is satisfied, then v is a weak solution to $\mathbf{C}(c, f, g)$.
2. If $g \equiv 0$, $c \equiv 0$, $f \in L^1(\mathbb{R}^N)$, (UFG, m) is satisfied for some $m \in \mathbb{N}$, then v is a weak solution.

Proof of Theorem 7.4 : We note that in the notation introduced in the previous chapters: $v(t, x) = P_t^c f(x) + L_g^c(t, x)$. We prove Theorem 7.4 and then deduce Theorem 7.3 via a few straightforward observations. Note that we technically do not require to consider L_g^c for the proof of Theorem 7.3, as $g \equiv 0$, but we include it as we require it for the proof of Theorem 7.3.

Let $\varphi \in C_0^\infty((0, T) \times \mathbb{R}^N)$. Then

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^N} (-\partial_t - (\mathcal{L}^c)^*) \varphi(t, x) v(t, x) dx dt \\ &= \int_0^T \int_{\mathbb{R}^N} (-\partial_t - (\mathcal{L}^c)^*) \varphi(t, x) (P_t^c f(x) + L_g^c(t, x)) dx dt \\ &= I_1 + I_2, \end{aligned}$$

where

$$\begin{aligned} I_1 &= \int_0^T \int_{\mathbb{R}^N} (-\partial_t - (\mathcal{L}^c)^*) \varphi(t, x) P_t^c f(x) dx dt, \\ I_2 &= \int_0^T \int_{\mathbb{R}^N} (-\partial_t - (\mathcal{L}^c)^*) \varphi(t, x) L_g^c(t, x) dx dt. \end{aligned}$$

We shall show that $I_1 \equiv 0$ and

$$I_2 = \int_0^T \int_{\mathbb{R}^N} \varphi(t, x) g(t, x) dx dt. \quad (7.5)$$

Recall the following from Kusuoka and Stroock [24]:

$$\int_{\mathbb{R}^N} P_t f(x) g(x) dx = \int_{\mathbb{R}^N} f(x) (P_t^c)^* g(x) dx, \quad (7.6)$$

where $(P_t^c)^*$ is the semigroup with infinitesimal generator $(\mathcal{L}^c)^*$. This result holds for $f, g \in C_0^\infty(\mathbb{R}^N)$, but if $g \in C_0^\infty(\mathbb{R}^N)$, then it may be extended to include the case where $f \in L^1(\mathbb{R}^N)$.

Hence

$$I_1 = \int_0^T \int_{\mathbb{R}^N} (P_t^c)^* (-\partial_t - (\mathcal{L}^c)^*) \varphi(t, x) f(x) dx dt.$$

It may be shown that $(P_t^c)^*$ and $(\mathcal{L}^c)^*$ commute on C_0^∞ and so

$$\begin{aligned} I_1 &= \int_0^T \int_{\mathbb{R}^N} -\partial_t (P_t^c)^* (\varphi(t, x) f(x)) dx dt \\ &= - \int_{\mathbb{R}^N} (P_t^c)^* (\varphi(t, x) f(x)) dx \Big|_{t=0}^T \\ &= 0, \end{aligned}$$

as φ has compact support in $(0, T) \times \mathbb{R}^N$. We now attempt to show (7.5). The first important step is to show the following for $f, g \in C_0^\infty((0, T) \times \mathbb{R}^N)$:

$$\int_0^T \int_{\mathbb{R}^N} \tilde{L}_\varphi^c(t, x) \psi(t, x) dx dt = \int_0^T \int_{\mathbb{R}^N} \varphi(t, x) L_\psi^c(t, x) dx dt,$$

where $\tilde{L}_\varphi^c(t, x) = \int_0^{T-t} (P_u^c)^* \varphi_{t+u}(x) du$.

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^N} \varphi(t, x) L_\psi^c(t, x) dx dt &= \int_0^T \int_{\mathbb{R}^N} \int_0^t \varphi(t, x) P_s^c \psi_{t-s}(x) ds dx dt \\ &= \int_{\mathbb{R}^N} \int_0^T \int_0^t (P_s^c)^* \varphi_t(x) \psi(t-s, x) ds dt dx \\ &= \int_{\mathbb{R}^N} \int_0^T \int_u^T (P_{v-u}^c)^* \varphi_v(x) \psi(v, x) du dv dx \\ &= \int_{\mathbb{R}^N} \int_0^T \int_0^{T-v} P_u^c \varphi_{v+u}(x) du \psi(v, x) dv dx \\ &= \int_0^T \int_{\mathbb{R}^N} \tilde{L}_\varphi^c(v, x) \psi(v, x) dx dv, \end{aligned}$$

as required. As was the case with (7.6) this result may be extended to deal with the case where $\psi \in C([0, T] \times \mathbb{R}^N)$, for $\varphi \in C_0^\infty((0, T) \times \mathbb{R}^N)$. It remains to compute the $\tilde{L}_{(-\partial_t - (\mathcal{L}^c)^*)\varphi}^c(v, x)$:

$$\begin{aligned} \tilde{L}_{(-\partial_t - (\mathcal{L}^c)^*)\varphi}^c(v, x) &= \int_0^{T-t} (P_u^c)^* (-\partial_{t+u} - (\mathcal{L}^c)^*) \varphi_{t+u}(x) du \\ &= \int_0^{T-t} (P_u^c)^* (-\partial_{t+u} - (\mathcal{L}^c)^*) \varphi_{t+u}(x) du \\ &= \int_0^{T-t} [\partial_u - (\mathcal{L}^c)^*] (P_u^c)^* \varphi_{t+u}(x) - \partial_u [(P_u^c)^* \varphi_{t+u}](x) du \\ &= -(P_u^c)^* \varphi_{t+u}(x) \Big|_{u=0}^{T-t} = \varphi(t, x), \end{aligned}$$

as required. We have demonstrated part 3. of the definition. Parts 1. and 2. have been demonstrated at various points in the thesis, but were the particular focus of Chapter 6. The various claims of this theorem can be verified by the gradient bounds of this chapter. It remains to show part 4. Observe

$$\begin{aligned}
& \int_{\mathbb{R}^N} \rho(x)(v(t, x) - f(x))dx \\
&= \int_{\mathbb{R}^N} \rho(x)(P_t^c f(x) + L_g^c(t, x) - f(x))dx \\
&= \int_{\mathbb{R}^N} ((P_t^c)^* \rho(x) - \rho(x)) f(x)dx + \int_{\mathbb{R}^N} \rho(x) L_g^c(t, x)dx \\
&\leq \|(P_t^c)^* \rho - \rho\|_\infty \|f\|_{L^1(dx)} + t \int_{\mathbb{R}^N} \rho(x)dx \sup_{\substack{s \in [0, t] \\ x \in \text{Supp} \rho}} P_s^c g_{t-s}(x)
\end{aligned}$$

It is clear that the second term converges to 0 as $t \rightarrow 0$. We show that $\|(P_t^c)^* \rho - \rho\|_\infty \rightarrow 0$:

$$\begin{aligned}
(P_t^c)^* \rho(x) - \rho(x) &= \int_0^t \partial_s (P_s^c)^* \rho(x) ds \\
&= \int_0^t (\mathcal{L}^c)^* (P_s^c)^* \rho(x) ds \\
&= \int_0^t (P_s^c)^* [(\mathcal{L}^c)^* \rho](x) ds \\
&\leq Ct \|(\mathcal{L}^c)^* \rho\|_\infty \\
&\xrightarrow{\text{as } t \rightarrow 0} 0.
\end{aligned}$$

This completes the proof of weak continuity. ■

Proof of Theorem 7.3 : We have already proved that v is a weak solution in the two situations provided in Theorem 7.4. Moreover, we know from earlier work in Chapter 6 - if we assume either condition of Theorem 7.3 - then $v \in C([0, T] \times \mathbb{R}^N)$ and the prescribed derivatives: $V_i v, V_i^2 v \in C([0, T] \times \mathbb{R}^N)$. The latter observation means that taking the adjoint operation in (7.3) is well-defined. i.e.

$$\begin{aligned}
\int_{[0, T] \times \mathbb{R}^N} (-\partial_t - (\mathcal{L}^c)^*) \varphi(t, x) v(t, x) dx dt &= \int_{[0, T] \times \mathbb{R}^N} \varphi(t, x) (\partial_t - \mathcal{L}^c) v(t, x) dx dt \\
&= \int_{[0, T] \times \mathbb{R}^N} \varphi(t, x) g(t, x) dx dt.
\end{aligned}$$

The denseness of $C_0^\infty((0, T) \times \mathbb{R}^N)$ and the continuity of $g, V_i v, V_i^2 v$ implies that (7.1) must also hold. This completes the proof. ■

7.2. Uniqueness of solutions

We must now define what it means for a solution of the Cauchy problem to be unique. We shall restrict consideration to problems for which the diffusion semigroup **is** a solution to the Cauchy problem.

Definition 7.5 (Uniqueness to the Cauchy problem)

A **strong solution**, u , to the Cauchy problem is unique if

$$u = v, \quad \text{on } [0, T] \times \mathbb{R}^N$$

A **weak solution**, u , to the Cauchy problem is said to be unique if

$$\int_{[0, T] \times \mathbb{R}^N} |u(t, x) - v(t, x)| dt dx = 0. \quad (7.7)$$

where v is given by (7.4).

It is clear that (a) strong uniqueness implies weak uniqueness (b) weak uniqueness + continuity of u and v implies strong uniqueness.

Theorem 7.6 A weak solution, u , to the Cauchy problem is unique if for almost every $t \in [0, T]$ and each $\rho \in C_0^\infty(\mathbb{R}^N)$, we have that the process $(V_{t, \rho}(s))_{0 \leq s \leq t}$

$$\begin{aligned} V_{t, \rho}(s) := \int_{\mathbb{R}^N} \rho(x) \left\{ u(t-s, X_s^x) \exp \left(\int_0^s c(u, X_u^x) du \right) \right. \\ \left. + \int_0^s g(t-u, X_u^x) \exp \left(\int_0^u c(\theta, X_\theta^x) d\theta \right) du \right\} dx, \end{aligned}$$

is a martingale.

Proof : If, for $t \in [0, T]$, $(V_{t, \rho}(s))_{0 \leq s \leq t}$ is a martingale, then:

$$\mathbb{E} [V_{t, \rho}(0)] = \mathbb{E} [V_{t, \rho}(t)].$$

That is to say

$$\begin{aligned} \int_{\mathbb{R}^N} \rho(x) u(t, x) dx &= \int_{\mathbb{R}^N} \rho(x) \mathbb{E} \left[u(0, X_t^x) \exp \left(\int_0^t c(u, X_u^x) du \right) \right. \\ &\quad \left. + \int_0^t g(t-u, X_u^x) \exp \left(\int_0^u c(\theta, X_\theta^x) d\theta \right) du \right] dx, \\ &= \int_{\mathbb{R}^N} \rho(x) v(t, x) dx, \end{aligned}$$

where v is given by (7.4) Hence

$$\int_0^T \int_{\mathbb{R}^N} \rho(x) (u(t, x) - v(t, x)) dx = 0.$$

This holds for all $\rho \in C_0^\infty(\mathbb{R}^N)$. Moreover, we can choose $t \mapsto \rho_t(x)$, so that $\rho = \rho_t(x)$ satisfies $\rho \in C_0^\infty((0, T) \times \mathbb{R}^N)$. By the denseness of such functions in $L^1([0, T] \times \mathbb{R}^N)$, it holds that $u = v$ a.e. on $[0, T] \times \mathbb{R}^N$, as required. ■

Theorem 7.7 (Uniqueness of solutions to the Cauchy Problem) *Under the assumptions of Theorems 7.3 and 7.4 the prescribed solutions are also unique.*

Proof : We first consider the problem of showing that the strong solution of Theorem 7.3 is also unique in a strong sense. We consider $\{u_n\}_{n \in \mathbb{N}}, \{c_n\}_{n \in \mathbb{N}}$ such that

$$\begin{aligned} \|u^n - u\|_{L^1([0, T] \times \mathbb{R}^N)} &\rightarrow 0, \\ \|c^n - c\|_{W^{1, \infty}([0, T] \times \mathbb{R}^N)} &\rightarrow 0, \end{aligned}$$

That such a sequence exists is guaranteed in both cases 1. and 2. of Theorem 7.3. Note that we may also assume that the sequences are bounded within their respective spaces. Let $\rho \in C_0^\infty(\mathbb{R}^N)$ and define:

$$\begin{aligned} V_t^{(n)}(s) := \int_{\mathbb{R}^N} \rho_t(x) &\left\{ u^n(t - s, X_s^x) \exp \left(\int_0^s c^n(u, X_u^x) du \right) \right. \\ &\left. + \int_0^s g(t - u, X_u^x) \exp \left(\int_0^u c(\theta, X_\theta^x) d\theta \right) du \right\} dx, \end{aligned}$$

First we will show that $\{V_t^{(n)}\}_n$ is Cauchy in $L^1([0, t] \times \mathbb{R}^N)$.

$$\mathbb{E} \int_0^t \left| V_t^{(n)}(s) - V_t^{(m)}(s) \right| ds \leq \mathbb{E} I_1(t),$$

where

$$\begin{aligned} I_1(t) := \int_{[0, t] \times \mathbb{R}^N} &|\rho(x)| \left| u^n(t - s, X_s^x) \exp \left(\int_0^s c^n(u, X_u^x) du \right) \right. \\ &\left. - u^m(t - s, X_s^x) \exp \left(\int_0^s c^m(u, X_u^x) du \right) \right| dx ds. \end{aligned}$$

We now note that $|a_n b_n - a_m b_m| \leq |a_n| |b_n - b_m| + [\sup_n |b_n|] |a_n - a_m|$. Hence,

$$\begin{aligned} I_1(t) &\leq \int |\rho| \left| u^n \exp \left(\int c^n du \right) - u^m \exp \left(\int c^m du \right) \right| dx ds \\ &\leq \int |\rho| |u^n| \left| \exp \left(\int c^n du \right) - \exp \left(\int c^m du \right) \right| \\ &\quad + \sup_n \left| \exp \left(\int c^n du \right) \right| |u^n - u^m| dx ds. \end{aligned}$$

It can be shown that

$$\begin{aligned} \left| \exp \left(\int c^n du \right) - \exp \left(\int c^m du \right) \right| &\leq \sup_n \left| \exp \left(\int \|c^n\|_\infty du \right) \right| \left(\int c^n du - \int c^m du \right) \\ &\leq \exp \left(\sup_n \|c\|_\infty s \right) \|c^n - c^m\|_\infty s. \end{aligned}$$

Hence,

$$\begin{aligned} \mathbb{E} I_1(t) &\leq \mathbb{E} \int_{[0,t] \times \mathbb{R}^N} |\rho(x)| \left(|u^n(t-s, X_s^x)| \exp(\|c\|_\infty s) \|c^n - c^m\|_\infty \right. \\ &\quad \left. + \exp(s \sup_n \|c^n\|_\infty) |u^n(t-s, X_s^x) - u^m(t-s, X_s^x)| \right) dx ds \\ &\leq \mathbb{E} \int_{[0,t] \times \mathbb{R}^N} |\rho((X_s^y)^{-1})| |J_s^y|^{-1} \left(|u^n(t-s, y)| \exp(\|c\|_\infty s) \|c^n - c^m\|_\infty \right. \\ &\quad \left. + \exp(s \sup_n \|c^n\|_\infty) |u^n(t-s, y) - u^m(t-s, y)| \right) dy ds \\ &\leq \sup_{(s,y) \in [0,t] \times \mathbb{R}^N} \mathbb{E} |\rho((X_s^y)^{-1})| |J_s^y|^{-1} \left(\sup_n \|u^n\|_{L^1} \exp(\|c\|_\infty t) \|c^n - c^m\|_\infty \right. \\ &\quad \left. + \exp(t \sup_n \|c^n\|_\infty) \|u^n - u^m\|_{L^1} \right) \\ &\leq C \left(\sup_n \|u^n\|_{L^1} \exp(\|c\|_\infty t) \|c^n - c^m\|_\infty + \exp(t \sup_n \|c^n\|_\infty) \|u^n - u^m\|_{L^1} \right) \end{aligned}$$

as required. Hence, $V_t^{(n)}$ is Cauchy in $L^1([0, T] \times \mathbb{R}^N)$, and is also convergent. We now seek to show that the limit $\{V_t(s)\}_{0 \leq s \leq t-\epsilon}$ is a martingale for each $\epsilon > 0$. Observe, by applying Itô's lemma (cf Karatzas and Shreve [17])

$$\begin{aligned} dV_t^{(n)}(s) &:= \int_{\mathbb{R}^N} \rho(x) \exp \left(\int_0^s c^n(u, X_u^x) du \right) \{ -\partial_t + \mathcal{L} + c \} u^n(t-s, X_s^x) ds dx \\ &\quad + \int_{\mathbb{R}^N} \rho(x) \exp \left(\int_0^s c^n(u, X_u^x) du \right) \sum_{i=1}^d V_i u^n(t-s, X_s^x) dB_s^i dx \\ &\quad + \int_{\mathbb{R}^N} \rho(x) g(t-s, X_s^x) \exp \left(\int_0^s c(\theta, X_\theta^x) d\theta \right) ds dx, \end{aligned}$$

Therefore, for each $s \leq t - \epsilon$:

$$\begin{aligned}
\mathbb{E}V_t(s) &= \lim_{n \rightarrow \infty} \mathbb{E}V_t^{(n)}(s) \\
&= \lim_{n \rightarrow \infty} \int_{\mathbb{R}^N} \rho(x) u^n(t, x) dx \\
&+ \int_{\mathbb{R}^N} \int_0^s \rho(x) \lim_{n \rightarrow \infty} \mathbb{E} \exp \left(\int_0^u c^n(\theta, X_\theta^x) d\theta \right) \{-\partial_t + \mathcal{L} + c\} u^n(t-u, X_u^x) du dx \\
&\quad + \int_{\mathbb{R}^N} \int_0^s \rho(x) \mathbb{E} \exp \left(\int_0^u c(\theta, X_\theta^x) d\theta \right) g(t-u, X_u^x) du dx \\
&= \int_{\mathbb{R}^N} \rho(x) u(t, x) dx,
\end{aligned}$$

by assumption 2. in the existence argument. This means that $\{V_t(s)\}_{0 \leq s \leq t-\epsilon}$ is a martingale. The argument may be completed by using the weak continuity of the solution at 0. We use it to show that $V_t(s) \xrightarrow{L^1(\Omega)} V_t(t)$ as $s \uparrow t$. This proves that $\{V_t(s)\}_{0 \leq s \leq t}$ is a martingale, and hence the result. Observe:

$$\begin{aligned}
\mathbb{E} |V_t(s) - V_t(t)| &= \mathbb{E} \left| \int_{\mathbb{R}^N} \rho(x) \left\{ f(X_t^x) - u(t-s, X_s^x) \exp \left(\int_0^s c(u, X_u^x) du \right) \right\} dx \right| \\
&\leq \mathbb{E} \left| \int_{\mathbb{R}^N} \rho(x) \{f(X_t^x) - f(X_s^x)\} dx \right| \\
&\quad + \mathbb{E} \left| \int_{\mathbb{R}^N} \rho(x) \{f(X_s^x) - u(t-s, X_s^x)\} dx \right| \\
&\quad + \mathbb{E} \left| \int_{\mathbb{R}^N} \rho(x) u(t-s, X_s^x) \left\{ \exp \left(\int_0^s c(u, X_u^x) du \right) - 1 \right\} dx \right| \\
&\leq \mathbb{E} \left| \int_{\mathbb{R}^N} \left(\rho((X_t^y)^{-1}) |J_t^y|^{-1} - \rho((X_s^y)^{-1}) |J_s^y|^{-1} \right) f(y) dy \right| \\
&\quad + \mathbb{E} \left| \int_{\mathbb{R}^N} \rho((X_s^y)^{-1}) |J_s^y|^{-1} \{f(y) - u(t-s, y)\} dy \right| \\
&\quad + \mathbb{E} \left| \int_{\mathbb{R}^N} \rho((X_s^y)^{-1}) |J_s^y|^{-1} u(t-s, y) \left\{ \exp \left(\int_0^s c(u, X_{u-s}^y) du \right) - 1 \right\} dy \right| \\
&\leq \sup_{x \in \mathbb{R}^N} \mathbb{E} \left| \rho((X_t^y)^{-1}) |J_t^y|^{-1} - \rho((X_s^y)^{-1}) |J_s^y|^{-1} \right| \|f\|_{L^1(dx)} \\
&\quad + \mathbb{E} \left| \int_{\mathbb{R}^N} \rho((X_s^y)^{-1}) |J_s^y|^{-1} \{f(y) - u(t-s, y)\} dy \right| \\
&\quad + \mathbb{E} \left| \int_{\mathbb{R}^N} \rho((X_s^y)^{-1}) |J_s^y|^{-1} |u(t-s, y)| \left\{ \exp \left(\int_0^s c(u, X_{u-s}^y) du \right) - 1 \right\} dy \right| \\
&\rightarrow 0.
\end{aligned}$$

Note that the weak continuity at zero is used to prove convergence of the second and third terms. That the first term converges to 0 can be shown by the Lipschitz continuity of ρ and by using upper bounds of the sort:

$$\sup_{y \in \mathbb{R}^N} \mathbb{E} |(X_t^y)^{-1} - (X_s^y)^{-1}| \leq C |t - s|.$$

Similar bounds hold for the determinant of $(J_s^y)^{-1}$. This shows that the weak solution to the Cauchy problem is unique. We now assume u is a strong solution to the Cauchy problem. In particular, it holds that $u \in C([0, T] \times \mathbb{R}^N)$. Under the assumptions of Theorem 7.3 we have already shown that $v \in C([0, T] \times \mathbb{R}^N)$. Hence, by this continuity

$$\int_{\mathbb{R}^N} |u(t, x) - v(t, x)| dx dt = 0, \quad \implies \quad u = v, \quad \text{on } [0, T] \times \mathbb{R}^N,$$

and strong uniqueness holds. ■

Remark 7.8 (Strong continuity of the semigroup)

Deciding when the semigroup is strongly continuous is of much interest, and has been completely characterised through the Hille-Yosida theorem. We briefly discuss this property and why we seemingly cannot use the methods presented in this thesis to analyse it. By strong continuity it is meant,

$$\lim_{t \rightarrow 0} \|P_t^c f - f\|_{L^p(dx)} = 0,$$

for each $f \in L^p(dx)$. Although our integration by parts formulae could potentially shed light on this property, it would seem as if this problem is highly non-trivial and cannot be expected to hold in general. To see this consider $\varphi \in C_0^\infty$:

$$\begin{aligned} P_t^c \varphi(x) - \varphi(x) &= \int_0^t \partial_s P_s^c \varphi(x) ds \\ &= \int_0^t \mathcal{L}^c P_s^c \varphi(x) ds \\ &= \int_0^t s^{-1} \mathbb{E} [\Phi(s, x) \varphi(X_s^x)] ds, \end{aligned}$$

where we have applied the integration by parts formulae to obtain the third equality. Note: we have also been forced to assume that the drift vector, V_0 , may be expressed as a linear combination of $V_i, V_{[i,j]}$ for $1 \leq i, j \leq d$. However, there is no obvious way to proceed from here, as it would seem that this integral could explode.

A. Appendix

SELECTED RESULTS

Lemma A.1 (Gronwall's Inequality) *Suppose that a continuous function $g(t)$ satisfies*

$$0 \leq g(t) \leq \alpha(t) + \beta \int_0^t g(s) ds ; \quad 0 \leq t \leq T,$$

with $\beta \leq 0$ and $\alpha : [0, T] \rightarrow \mathbb{R}$ integrable. Then

$$g(t) \leq \alpha(t) + \beta \int_0^t \alpha(s) e^{\beta(t-s)} ds ; \quad 0 \leq t \leq T.$$

Proof : See Karatzas and Shreve [17, p387-388]. ■

Proposition A.2 (The Burkholder-Davis-Gundy Inequality) *Let assume that M is a continuous local martingale. Then for every $p > 0$ there is a universal constant K_p such that:*

$$\mathbb{E} \left[\sup_{t \leq T} |M_t| \right]^p \leq K_p \mathbb{E} \left[\langle M \rangle_T^{p/2} \right],$$

for any stopping time T . Note that $\langle M \rangle_t$ denotes the quadratic variation of the martingale at time t .

Proof : See, for example, Karatzas and Shreve [17, p166]. ■

Proposition A.3 (Jensen's inequality for definite integrals) *There holds the following, for integrable $u \in L^p([0, T], \mathcal{L})$, for $p \geq 1$.*

$$\left(\int_0^t u_s ds \right)^p \leq t^{p-1} \int_0^t u_s^p ds.$$

A.1. 'Introduction and background material'

Theorem 1.5 [Basic Integration by Parts Formula] *Assume F, G are smooth random variables, and let $h' \in H$. Then the following equality holds.*

$$\mathbb{E}(D_h F.G) = \mathbb{E}(FG \int_0^\infty h'(u)dB_u - F.D_h G).$$

Proof : One has already seen that

$$D_h F = \left. \frac{d}{d\epsilon} F(\omega + \epsilon h) \right|_{\epsilon=0}.$$

One proceeds by manipulating the RHS and using the *Cameron-Martin Theorem*.

$$\begin{aligned} & \mathbb{E}[D_h F.G] \\ &= \mathbb{E} \left[\left. \frac{d}{d\epsilon} F(\cdot + \epsilon h) \right|_{\epsilon=0} G \right] \\ &= \mathbb{E} \left[\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [F(\cdot + \epsilon h) - F(\cdot)] G(\cdot) \right] \\ &\stackrel{(i)}{=} \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \mathbb{E}[(F(\cdot + \epsilon h) - F(\cdot)) G(\cdot)] \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \mathbb{E} \left[F(\cdot) G(\cdot - \epsilon h) \exp \left(\epsilon \int_0^\infty h'(u)dB_u - \frac{1}{2} \epsilon^2 \int_0^\infty h'(u)du \right) - FG \right] \\ &= \mathbb{E} \left[F(\cdot) \frac{d}{d\epsilon} \left(G(\cdot - \epsilon h) \exp \left(\epsilon \int_0^\infty h'(u)dB_u - \frac{1}{2} \epsilon^2 \int_0^\infty h'(u)du \right) \right) \Big|_{\epsilon=0} \right] \\ &= \mathbb{E} \left[F(\cdot) G(\cdot) \int_0^\infty h'(u)dB_u - F(\cdot) \frac{d}{d\epsilon} G(\cdot + \epsilon h) \Big|_{\epsilon=0} \right] \\ &= \mathbb{E} \left[FG \int_0^\infty h'(u)dB_u - F.D_h G \right]. \end{aligned}$$

(i): The Dominated Convergence Theorem is used here.

N.B. It has been used that $\frac{d}{d\epsilon} G(\cdot - \epsilon h)|_{\epsilon=0} = -\frac{d}{d\epsilon} G(\cdot + \epsilon h)|_{\epsilon=0}$. ■

A.2. Appendix for 'Kusuoka's gradient bounds'

Proposition 2.5 *For any $T \in (0, 1]$, $p \in [1, \infty)$, $\alpha, \beta \in \mathcal{A}(m)$ and $\gamma \in \mathcal{A}$, the following hold*

$$\sup_{t \in (0, T]} \mathbb{E} \left[t^{-\|\gamma\|/2} \left| \hat{B}_t^{\circ \gamma} \right| \right]^p < \infty, \quad (2.10)$$

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in (0, T]}} \mathbb{E} \left[t^{-(m+1-\|\alpha\|)/2} |r_{\alpha, \beta}(t, x)| \right]^p < \infty. \quad (2.11)$$

Proof : As we are considering Stratonovich integrals throughout this report, one needs to take care in handling the semimartingales which result from this choice. The proof is done as follows; we first show how the result holds for a general semimartingale. We already proved (2.10) during the main body of the text; using a simple argument. We prove (2.11) via an inductive argument. Assume that

$$\xi_t^x = \int_0^t u(s, x) dW_s + \int_0^t v(s, x) ds,$$

where

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in [0, T]}} \mathbb{E} (t^{-r_u} |u(t, x)|)^p < \infty, \quad \text{and} \quad \sup_{\substack{x \in \mathbb{R}^N \\ t \in [0, T]}} \mathbb{E} (t^{-r_v} |v(t, x)|)^p < \infty,$$

for some constants, $r_u, r_v \in [0, \infty)$. Then, wlog for $p \geq 2$:

$$\begin{aligned} \mathbb{E} |\xi_t^x|^p &= \mathbb{E} \left| \int_0^t u(s, x) dW_s + \int_0^t v(s, x) ds \right|^p \\ &\stackrel{(1)}{\leq} 2^{p-1} \left\{ \mathbb{E} \left| \int_0^t u(s, x) dW_s \right|^p + \mathbb{E} \left| \int_0^t v(s, x) ds \right|^p \right\} \\ &\stackrel{(2)}{\leq} 2^{p-1} \left\{ C_p \mathbb{E} \left(\int_0^t |u(s, x)|^2 ds \right)^{\frac{p}{2}} + t^{p-1} \mathbb{E} \int_0^t |v(s, x)|^p ds \right\} \\ &\stackrel{(3)}{\leq} 2^{p-1} \left\{ C_p t^{\frac{p}{2}-1} \mathbb{E} \int_0^t |u(s, x)|^p ds + t^{p-1} \mathbb{E} \int_0^t |v(s, x)|^p ds \right\} \\ &\stackrel{(4)}{\leq} 2^{p-1} \left\{ C_p t^{\frac{1}{2}(p-1)} \mathbb{E} |u(s, x)|^p ds + t^{p-1} \int_0^t \mathbb{E} |v(s, x)|^p ds \right\}. \end{aligned}$$

1. Hölder's inequality for finite sums.
2. Burkholder's inequality, Jensen's inequality resp.
3. Jensen's inequality for definite integrals.
4. Tonelli's theorem.

Now we observe that;

$$\begin{aligned} \mathbb{E} |u(s, x)|^p &\leq \left(\sup_{\substack{x \in \mathbb{R}^N \\ s \in [0, T]}} \mathbb{E} [s^{-r_u} |u(s, x)|]^p \right) s^{pr_u} \\ \mathbb{E} |v(s, x)|^p &\leq \left(\sup_{\substack{x \in \mathbb{R}^N \\ s \in [0, T]}} \mathbb{E} [s^{-r_v} |v(s, x)|]^p \right) s^{pr_v}. \end{aligned}$$

And so,

$$\begin{aligned}
\mathbb{E} |\xi_t^x|^p &\leq \tilde{C}_p \left\{ t^{\frac{1}{2}p-1} \left(\sup_{\substack{x \in \mathbb{R}^N \\ s \in [0, T]}} \mathbb{E} [s^{-r_u} |u(s, x)|]^p \right) \left(\int_0^t s^{pr_u} ds \right) \right. \\
&\quad \left. + t^{p-1} \left(\sup_{\substack{x \in \mathbb{R}^N \\ s \in [0, T]}} \mathbb{E} [s^{-r_v} |v(s, x)|]^p \right) \left(\int_0^t s^{pr_v} ds \right) \right\} \\
&\leq \hat{C}_p \left\{ t^{\frac{1}{2}p-1} t^{pr_u+1} \left(\sup_{\substack{x \in \mathbb{R}^N \\ s \in [0, T]}} \mathbb{E} [s^{-r_u} |u(s, x)|]^p \right) \right. \\
&\quad \left. + t^{p-1} t^{pr_v+1} \left(\sup_{\substack{x \in \mathbb{R}^N \\ s \in [0, T]}} \mathbb{E} (s^{-r_v} |v(s, x)|)^p \right) \right\} \\
&\leq \hat{C}_{p,u,v} \left\{ t^{p(r_u + \frac{1}{2})} + t^{p(r_v + 1)} \right\}.
\end{aligned}$$

That is, if we take $r_\xi = \min\{r_u + \frac{1}{2}, r_v + 1\}$, then for all $p \in [1, \infty)$,

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in (0, T]}} \mathbb{E} [t^{-r_\xi} |\xi_t^x|]^p < \infty. \quad (\text{A.1})$$

Proof of (2.11):

The proof of this result is similar to the induction carried out above. We notice that the remainder term, as defined, is the sum of numerous iterated Stratonovich integrals. We prove that the result holds for one element of the sum. This may then be easily extended to the sum of multiple such objects.

We have already seen (cf Proposition 2.3) that, for any $\alpha, \beta \in \mathcal{A}(m)$, $p \in [1, \infty)$, $T > 0$:

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in [0, T]}} \mathbb{E} |a_{\alpha, \beta}(t, x)|^p < \infty. \quad (\text{A.2})$$

Moreover, since $c_{\alpha * \gamma, \beta}^j \in C_b^\infty$, it follows that

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in [0, T]}} \mathbb{E} \left| c_{\alpha * \gamma, \beta}^j(X_t^x) \right|^p < \infty. \quad (\text{A.3})$$

Moreover, its partial derivatives are also in C_b^∞ , so $V_i c_{\alpha * \gamma, \beta}^j$, $V_i^2 c_{\alpha * \gamma, \beta}^j$, etc.. belong to C_b^∞ , and must also satisfy (A.3).

We again prove the result by induction on $|\gamma|$. Assume $|\gamma| = 1$. Using the fact that both $c_{\alpha * \gamma, \delta}^j(X_t^x)$ and $a_{\delta, \beta}(t, x)$ have semimartingale representations, given by:

$$c_{\alpha * \gamma, \delta}^j(X_t^x) = c_{\alpha * \gamma, \delta}^j(x) + \sum_{k=1}^d \int_0^t (V_k c_{\alpha * \gamma, \delta}^j)(X_s^x) dB_s^k + \int_0^t \mathcal{L} c_{\alpha * \gamma, \delta}^j(X_s^x) ds,$$

where $\mathcal{L} := \frac{1}{2} \sum_{k=0}^d V_k^2 + V_0$, and

$$\begin{aligned} a_{\delta,\beta}(t, x) &= \delta_{\delta,\beta} + \sum_{i=0}^d \sum_{\xi \in \mathcal{A}(m)} \int_0^t c_{\delta,\xi}^i(X_s^x) a_{\xi,\beta}(s, x) dB_s^i \\ &\quad + \frac{1}{2} \sum_{i=1}^d \sum_{\xi \in \mathcal{A}(m)} \int_0^t (V_i c_{\delta,\xi}^i)(X_s^x) a_{\xi,\beta}(s, x) ds \\ &\quad + \frac{1}{2} \int_0^t \sum_{i=1}^d \sum_{\xi, \tilde{\xi} \in \mathcal{A}(m)} (c_{\delta,\xi}^i c_{\tilde{\xi},\xi}^i)(X_s^x) a_{\xi,\beta}(s, x) ds. \end{aligned}$$

Hence we may express the product as a semimartingale. Note that this semimartingale will be comprised of integrands which are sums and products of objects like those in (A.2), (A.3), and hence if $\gamma \in \{1, \dots, d\}$ we may apply the inductive step for some u, v with $r_u = r_v = 0$, to get

$$\mathbb{E} \left[t^{-r_\gamma} \int_0^t c_{\alpha*\gamma,\delta}^j(X_s^x) a_{\delta,\beta}(s, x) \circ dB_t^\gamma \right]^p < \infty, \quad (\text{A.4})$$

where $r_\gamma = \min\{\frac{1}{2}, 1\} = \frac{1}{2}$. Now if $\gamma = 0$, then we apply the inductive step with $u \equiv 0$ and $v(t, x) = c_{\alpha*\gamma,\delta}^j(X_t^x) a_{\delta,\beta}(t, x)$. That is, $0 = r_v \ll r_u$, to obtain

$$\mathbb{E} \left[t^{-r_\gamma} \int_0^t c_{\alpha*\gamma,\delta}^j(X_s^x) a_{\delta,\beta}(s, x) ds \right]^p < \infty, \quad (\text{A.5})$$

where $r_\gamma = 1$.

We now assume the result holds for some $k \in \mathbb{N}$. i.e. we have the following for all $\gamma \in \mathcal{A}$ satisfying $|\gamma| = k$:

$$\begin{aligned} \sup_{\substack{x \in \mathbb{R}^N \\ t \in (0, T]}} \mathbb{E} \left[t^{-\|\gamma\|/2} \left| \int_0^t \int_0^{s_k} \dots \int_0^{s_2} (-1)^{|\gamma|} c_{\alpha*\gamma,\delta}^j(X_{s_1}^x) a_{\delta,\beta}(s_1, x) \circ dB_{s_1}^{\gamma_1} \dots \circ dB_{s_k}^{\gamma_k} \right| \right]^p \\ < \infty. \end{aligned} \quad (\text{A.6})$$

To ease the notational burden, we write,

$$Z(t, x, \gamma) := \int_0^t \int_0^{s_k} \dots \int_0^{s_2} (-1)^{|\gamma|} c_{\alpha*\gamma,\delta}^j(X_{s_1}^x) a_{\delta,\beta}(s_1, x) \circ dB_{s_1}^{\gamma_1} \dots \circ dB_{s_k}^{\gamma_k},$$

for $\gamma = (\gamma_1, \dots, \gamma_k)$. Observe, that for $i \in \{1, \dots, d\}$

$$\begin{aligned} Z(t, x, \gamma * i) &= \int_0^t Z(s, x, \gamma) \circ dB_s^i \\ &= \int_0^t Z(s, x, \gamma) dB_s^i + \frac{1}{2} \langle Z(\cdot, x, \gamma), B^i \rangle_t \\ &= \int_0^t Z(s, x, \gamma) dB_s^i + \frac{1}{2} \delta_{\gamma_{k-1}, \gamma_k} \int_0^t Z(s, x, \gamma') dt \\ &= \int_0^t Z(s, x, \gamma) dB_s^i + \frac{1}{2} \delta_{\gamma_{k-1}, \gamma_k} Z(t, x, \gamma' * 0). \end{aligned}$$

By the inductive hypothesis, $Z(t, x, \gamma' * 0)$ satisfies (A.6) with $r_{\gamma' * 0} = (\|\gamma'\| + 2)/2$, and we also use the inductive step on the right-hand term with $u(t, x) = Z(t, x, \gamma)$ and $v \equiv 0$, so that $r_v \gg r_u = \|\gamma\|/2$, with

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in (0, T]}} \mathbb{E} [t^{-r_{\gamma * i}} |Z(t, x, \gamma * i)|]^p < \infty,$$

where $r_{\gamma * i} = \min \left\{ \frac{\|\gamma\| + 1}{2}, \frac{\|\gamma'\| + 2}{2} \right\} = \frac{\|\gamma\| + 1}{2}$. If $i = 0$ then we may apply the inductive step with $u \equiv 0$ and $v(t, x) = Z(t, x, \gamma)$, so that with $\|\gamma\|/2 = r_v \ll r_u$ we get

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in [0, T]}} \mathbb{E} [t^{-r_{\gamma * 0}} |Z(t, x, \gamma * 0)|]^p < \infty,$$

where $r_{\gamma * 0} = \frac{\|\gamma\| + 2}{2}$. Hence the result is proved.

Finally, note that a finite sum of these would also satisfy a similar inequality with $r_{\text{sum}} = \min\{r_k ; r_k \text{ is optimal (i.e. (A.1) holds) for } k\text{th sum member}\}$. i.e.

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in (0, T]}} \mathbb{E} \left[t^{-(m+1-\|\alpha\|)/2} |r_{\alpha, \beta}(t, x)| \right]^p < \infty,$$

as required. ■

Lemma 2.7 *The statement of Proposition 2.6 holds, providing the following can be shown for each $p \in [1, \infty)$: there exists $C > 0$ s.t.*

$$\mathbb{P} \left(\inf_{|\xi|=1} (\xi, M(t, x)\xi) < \frac{1}{n} \right) < Cn^{-p},$$

for all $n \geq 1$, $t \in (0, 1]$, $x \in \mathbb{R}^N$.

Proof : We note that

$$\Lambda(t, x) := \inf_{\xi \in S_{N_m-1}} \{(\xi, M(t, x)\xi)\} = \min\{\lambda; \lambda \text{ is an eigenvalue of } M(t, x)\}.$$

We also observe that since $M(t, x)$ is a real, symmetric matrix, it is non-negative definite and hence $\Lambda(t, x) \geq 0$. Moreover, $\det M(t, x) = \lambda_1 \dots \lambda_{N_m}$, i.e. the determinant is the product of its eigenvalues. Hence,

$$\begin{aligned} [\det M(t, x)]^{-p} &\leq \Lambda(t, x)^{-N_m p}, \\ \Rightarrow \sup_{t \in (0, 1], x \in \mathbb{R}^N} \mathbb{E}[\det M(t, x)]^{-p} &\leq \sup_{t \in (0, 1], x \in \mathbb{R}^N} \mathbb{E}[\Lambda(t, x)]^{-N_m p}. \end{aligned}$$

Now note that

$$\begin{aligned} \mathbb{E}(\Lambda(t, x)^{-N_m p}) &= \mathbb{E} \left[\int_0^\infty \mathbf{1}_{\{y \leq \Lambda(t, x)^{-N_m p}\}} dy \right] \\ &= \int_0^\infty \mathbb{P}(y \leq \Lambda(t, x)^{-N_m p}) dy \\ &= \int_0^\infty \mathbb{P}(\Lambda(t, x) \leq y^{-1/N_m p}) dy. \end{aligned}$$

It follows that

$$\begin{aligned} \sup_{t \in (0, 1], x \in \mathbb{R}^N} \mathbb{E}[\det M(t, x)]^{-p} &\leq \sup_{t \in (0, 1], x \in \mathbb{R}^N} \int_0^\infty \mathbb{P}(\Lambda(t, x) \leq y^{-1/N_m p}) dy \\ &\leq 1 + \int_1^\infty \sup_{t \in (0, 1], x \in \mathbb{R}^N} \mathbb{P}(\Lambda(t, x) \leq y^{-1/N_m p}) dy \\ &\leq 1 + C \int_1^\infty y^{-q/N_m p} dy < \infty, \end{aligned}$$

where q is picked so that $q > N_m p$. ■

Lemma 2.10 *There holds, for all $p \in [1, \infty)$,*

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in (0, 1]}} \mathbb{E} \left[\int_0^t \sum_{\alpha \in \mathcal{A}(m)} \sum_{i=1}^d t^{-\|\alpha\|-1} r_{i,\alpha}(u, x)^2 du \right]^p < \infty.$$

Proof : We may apply the semimartingale rate bound obtained in the proof of Proposition 2.5.

Indeed, we observe that:

$$\begin{aligned}\xi_t &:= \int_0^t u(s, x) dB_s + \int_0^t v(s, x) ds, \\ u(s, x) &\equiv 0, \\ v(s, x) &= \sum_{\alpha \in \mathcal{A}(m)} \sum_{i=1}^d t^{-\|\alpha\|} r_{i,\alpha}(s, x)^2.\end{aligned}$$

Observe from Proposition 2.11, noting $\|\alpha\| \leq m$,

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in (0, T]}} \mathbb{E} (t^{-r_u} |u(t, x)|)^p < \infty, \quad \sup_{\substack{x \in \mathbb{R}^N \\ t \in (0, T]}} \mathbb{E} (t^{-r_v} |v(t, x)|)^p < \infty,$$

where $r_v = 0$ and r_u is arbitrarily large. Hence it follows that:

$$\sup_{\substack{x \in \mathbb{R}^N \\ t \in (0, T]}} \mathbb{E} (t^{-r} |\xi_t^x|)^p < \infty,$$

where $r_\xi = r_v + 1 = 1$, as required. ■

Lemma 2.19 [Properties of Kusuoka-Stroock processes]

The following hold

1. Suppose $f \in \mathcal{K}_r(E)$, where $r \geq 0$. Then, for $i = 1, \dots, d$,

$$\int_0^\cdot f(s, x) dB_s^i \in \mathcal{K}_{r+1}(E) \quad \text{and} \quad \int_0^\cdot f(s, x) ds \in \mathcal{K}_{r+2}(E).$$

2. $a_{\alpha,\beta}, b_{\alpha,\beta} \in \mathcal{K}_{(\|\beta\| - \|\alpha\|) \vee 0}$ where $\alpha, \beta \in \mathcal{A}(m)$.

3. $k_\alpha \in \mathcal{K}_{\|\alpha\|}(H)$, where $\alpha \in \mathcal{A}(m)$.

4. $D^{(\alpha)}u := \langle Du(t, x), k_\alpha \rangle_H \in \mathcal{K}_{r+\|\alpha\|}$ where $u \in \mathcal{K}_r$ and $\alpha \in \mathcal{A}(m)$.

5. If $M^{-1}(t, x)$ is the inverse matrix of $M(t, x)$, then $M_{\alpha,\beta}^{-1} \in \mathcal{K}_0$, $\alpha, \beta \in \mathcal{A}(m)$.

6. If $f_i \in \mathcal{K}_{r_i}$ for $i = 1, \dots, N$, then

$$\prod_{i=1}^N f_i \in \mathcal{K}_{r_1+\dots+r_N} \quad \text{and} \quad \sum_{i=1}^N f_i \in \mathcal{K}_{\min(r_1, \dots, r_N)}.$$

Proof :

(1): It is clear that if $f(t, \cdot)$ is smooth and $\partial_\alpha f(\cdot, \cdot)$ is continuous then the same is true of $\int_0^\cdot f(s, x) dB_s^i$ for $i = 0, \dots, d$, with

$$\partial_\alpha \int_0^\cdot f(s, x) dB_s^i = \int_0^\cdot \partial_\alpha f(s, x) dB_s^i.$$

For $k \geq 1$, $p \in [1, \infty)$, $i = 1, \dots, d$, we have (note that the dependence of the norms on the Hilbert space E has been suppressed):

$$\left\| \int_0^t \partial_\alpha f(s, x) dB_s^i \right\|_{k,p}^p = \mathbb{E} \left\| \int_0^t \partial_\alpha f(s, x) dB_s^i \right\|^p + \sum_{j=1}^k \mathbb{E} \left\| D^j \int_0^t \partial_\alpha f(s, x) dB_s^i \right\|_{H^{\otimes j}}^p. \quad (\text{A.7})$$

Focussing for a moment of the LHS, and assuming w.l.o.g. $p \geq 2$ (as there holds monotonicity of norms in p), we see that for $j = 0, \dots, k$, there holds

$$\begin{aligned} & \mathbb{E} \left\| D^j \left[\int_0^t \partial_\alpha f(s, x) dB_s^i \right] \right\|_{H^{\otimes j}}^p \\ &= \mathbb{E} \left\| \int_0^t D^j \partial_\alpha f(s, x) dB_s^i + \int_0^t D^{j-1} \partial_\alpha f(s, x) \otimes e_i ds \right\|_{H^{\otimes j} \otimes E}^p \\ &\leq 2^{p-1} \left[\mathbb{E} \left\| \int_0^t D^j \partial_\alpha f(s, x) dB_s^i \right\|_{H^{\otimes j}}^p + \mathbb{E} \left\| \int_0^t D^{j-1} \partial_\alpha f(s, x) \otimes e_i ds \right\|_{H^{\otimes j}}^p \right] \\ &\leq \tilde{C}_p \left[\mathbb{E} \int_0^t t^{\frac{1}{2}(p-1)} \|D^j \partial_\alpha f(s, x)\|_{H^{\otimes j}}^p ds + t^{p-1} \|D^{j-1} \partial_\alpha f(s, x)\|_{H^{\otimes(j-1)}}^p ds \right] \\ &\leq \tilde{C}_p t^{\frac{1}{2}(p-1)} \left[\int_0^t \mathbb{E} \|D^j \partial_\alpha f(s, x)\|_{H^{\otimes j}}^p ds + \int_0^t \mathbb{E} \|D^{j-1} \partial_\alpha f(s, x)\|_{H^{\otimes(j-1)}}^p ds \right] \\ &\leq \tilde{C}_p t^{\frac{1}{2}(p-1)} \int_0^t \|\partial_\alpha f(s, x)\|_{k,p}^p ds \\ &\leq \tilde{C}_p t^{\frac{1}{2}(p-1)} \int_0^t s^{rp/2} \sup_{\substack{x \in \mathbb{R}^N \\ v \in (0,1]}} v^{-rp/2} \|\partial_\alpha f(v, x)\|_{k,p}^p ds \\ &\leq \tilde{\tilde{C}}_p t^{\frac{1}{2}(p[r+1])}, \end{aligned}$$

where we have used Burkholder-Davis-Gundy inequality, Jensen's inequality and Hölder's inequality for finite sums. Note that the above holds for $j = 0$ by taking D^{j-1} to be the zero map. The upper bound is independent of $x \in \mathbb{R}^N$ and by a simple rearrangement, and combining with (A.7), the result follows. Note that the result for $\int_0^\cdot f(s, x) ds$ is proved similarly.

2: The fact that $a_{\alpha,\beta}(t, \cdot), b_{\alpha,\beta}(t, \cdot)$ are smooth with partial derivatives which are jointly continuous in $(t, x) \in (0, 1] \times \mathbb{R}^N$ follows from Theorem 1.12. The fact that $a_{\alpha,\beta}, b_{\alpha,\beta} : [0, T] \times \mathbb{R}^N \rightarrow \mathbb{D}^\infty$ follows from Theorem 1.13. The fact that the appropriate bound holds for $a_{\alpha,\beta}$ with rate $r = (\|\beta\| - \|\alpha\|) \wedge 0$ follows from applying the expression for $a_{\alpha,\beta}$, given in (2.9), and Proposition 2.5. The corresponding result for $b_{\alpha,\beta}$ is derived in an analogous way to $a_{\alpha,\beta}$.

3: This follows easily from **1** and **2**.

4: From Nualart [33][Prop 1.3.3] we have the following:

$$\langle Du, k_\alpha \rangle_H = u \delta(k_\alpha) - \delta(u k_\alpha)$$

Moreover, we know that $u, k_\alpha \in \mathbb{D}^\infty$, and that $\delta : \mathbb{D}^\infty \rightarrow \mathbb{D}^{\infty-1}$, hence it is clear that $\langle Du, k_\alpha \rangle_H \in \mathbb{D}^\infty$. The existence of regular derivatives of all orders follows from direct differentiation. The required bounds follows easily from 6.

5: Our first observation is that if $f \in \mathcal{K}_r(E)$, where $r \geq 0$, then $g(t, x) := t^{-s/2} f(t, x)$ satisfies $g \in \mathcal{K}_{r-s}(E)$. This is obvious, and from this basic observation we note that $M_{\alpha, \beta}(t, x) := t^{-(\|\alpha\| + \|\beta\|)/2} \langle k_\alpha(t, x), k_\beta(t, x) \rangle_H$ must satisfy $M_{\alpha, \beta} \in \mathcal{K}_0$. This comes from applying the above observation, along with 3. and 4. of this Lemma. To prove the same about elements of the inverse of $M(t, x)$ we first note that smoothness (in x) and joint continuity (in (t, x)) follows from the inverse function theorem. To prove Malliavin differentiability and the corresponding bounds, we use the ideas of the proof of Nualart [33, Lemma 2.1.6]. That is, we seek to prove the following:

Lemma A.4 *Let $A(., .)$ be a square random matrix, which is invertible almost surely and such that $|\det A(t, x)|^{-1} \in L^p$ for all $p \geq 1$. Assume further that the elements of $A_{\alpha, \beta}(t, x) \in \mathbb{D}^\infty$ and satisfy:*

$$\sup_{\substack{x \in \mathbb{R}^N \\ s \in [0, t]}} \|A_{\alpha, \beta}(s, x)\|_{k, p} < \infty.$$

Then $A_{\alpha, \beta}^{-1}(t, x) \in \mathbb{D}^\infty$ and the elements satisfy:

$$\sup_{\substack{x \in \mathbb{R}^N \\ s \in [0, t]}} \|A_{\alpha, \beta}^{-1}(s, x)\|_{k, p} < \infty. \quad (\text{A.8})$$

The proof of this lemma is almost identical to the proof of Nualart [33, Lemma 2.1.6]. One merely needs to take care in showing (A.8). This is done easily by using a Hölder-type inequality for the seminorms $\|\cdot\|_{k, p}$ (cf Nualart [33, Proposition 1.5.6]).

Remark A.5 *If we hadn't chosen to multiply and divide the elements of the matrix $\hat{M}(t, x) := (\langle k_\alpha(t, x), k_\beta(t, x) \rangle)$ by $t^{\frac{\|\alpha\| + \|\beta\|}{2}}$, when forming the matrix M , then more care would have been required to ensure that the rate of decay of the inverse (as a Kusuoka Stroock process) is independent of the dimension of the matrix. Indeed, it can be shown the inverse of the determinant of $\hat{M}$ is bounded above by a rate which is dimension dependent. However, this dimensionality dependence disappears when one considers the product with the matrix of cofactors, which has the equal and opposite dimensionality dependence.*

¹cf Proposition 1.5.4 in Nualart [33]

6: It is clear that smoothness, joint continuity and Malliavin differentiability are inherited from the constituent functions. The second property remains to be shown. Consider $\prod_{i=1}^N f_i$. It may be shown that for the k th Malliavin derivative the following Leibniz-type rule holds:

$$D^k \prod_{i=1}^N f_i = \sum_{i_1 + \dots + i_N = k} \binom{k}{i_1, \dots, i_N} D^{i_1} f_1 \otimes \dots \otimes D^{i_N} f_N.$$

Now noting that, if $i_1 + \dots + i_N = k$, we have

$$\|D^{i_1} f_1 \otimes \dots \otimes D^{i_N} f_N\|_{H^{\otimes k}} = \prod_{j=1}^N \|D^{i_j} f_j\|_{H^{\otimes i_j}},$$

so that

$$\begin{aligned} & \left\| \prod_{i=1}^N f_i(t, x) \right\|_{k,p}^p \\ &= \mathbb{E} \left| \prod_{i=1}^N f_i(t, x) \right|^p + \sum_{j=1}^k \left\| D^j \prod_{i=1}^N f_i(t, x) \right\|_{H^{\otimes j}}^p \\ &= \mathbb{E} \left| \prod_{i=1}^N f_i(t, x) \right|^p + \sum_{j=1}^k \left\| \sum_{i_1 + \dots + i_N = j} \binom{j}{i_1, \dots, i_N} \bigotimes_{m=1}^N D^{i_m} f_m(t, x) \right\|_{H^{\otimes j}}^p \\ &\leq \mathbb{E} \left| \prod_{i=1}^N f_i(t, x) \right|^p + \sum' C(p, j) \left\| \bigotimes_{m=1}^N D^{i_m} f_m(t, x) \right\|_{H^{\otimes j}}^p \\ &\leq \prod_{i=1}^N \|f_i(t, x)\|_{L^{p_i}(\Omega)}^p + \sum' C(p, j) \prod_{m=1}^N \|D^{i_m} f_m(t, x)\|_{H^{\otimes i_m}}^p, \end{aligned}$$

where $p^{-1} = p_1^{-1} + \dots + p_N^{-1}$, applying Hölder's Generalised Inequality. Whence, letting $r = \sum_{i=1}^N r(i)$ we see that

$$\begin{aligned} \sup_{t \in (0,1], x \in \mathbb{R}^N} t^{-r/2} \left\| \prod_{i=1}^N f_i(t, x) \right\|_{k,p}^p &\leq \prod_{i=1}^N \sup_{t \in (0,1], x \in \mathbb{R}^N} t^{-\frac{r_i}{2}} \|f_i(t, x)\|_{L^{p_i}(\Omega)}^p \\ &\quad + \sum' C(p, j) \prod_{m=1}^N \sup_{t \in (0,1], x \in \mathbb{R}^N} t^{-\frac{r_i}{2}} \|D^{i_m} f_m(t, x)\|_{H^{\otimes i_m}}^p \\ &< \infty. \end{aligned}$$

To see that $\sum_{i=1}^N f_i \in \mathcal{K}_{\min(r_1, \dots, r_N)}$. We note that $\mathcal{K}_r \subset \mathcal{K}_s$ for $r \leq s$. Hence, it should be clear that the sum is contained in that $\mathcal{K}_r$ in which **all** of its terms are contained, namely, $\mathcal{K}_{\min(r_1, \dots, r_N)}$. ■

A.3. 'Regularity of PDEs with C_b^k coefficients'

Lemma 3.9 (Properties of Kusuoka-Stroock processes)

1. Suppose $f \in \mathcal{K}_r(E, n)$, where $r \geq 0$. Then, for $i = 1, \dots, d$,

$$\int_s^\cdot f(u, x) dB_u^i \in \mathcal{K}_{r+1}(E, n) \quad \text{and} \quad \int_s^\cdot f(u, x) du \in \mathcal{K}_{r+2}(E, n).$$

2. $a_{\alpha, \beta}, b_{\alpha, \beta} \in \mathcal{K}_{(\|\beta\| - \|\alpha\|) \vee 0}(\mathbb{R}, k - m)$ where $\alpha, \beta \in \mathcal{A}(m)$.

3. $k_\alpha \in \mathcal{K}_{\|\alpha\|}(H, k - m)$, where $\alpha \in \mathcal{A}(m)$.

4. $D^{(\alpha)}u := \langle Du(t, x), k_\alpha \rangle_H \in \mathcal{K}_{r+\|\alpha\|}(\mathbb{R}, n \wedge (k - m))$ for $u \in \mathcal{K}_r(\mathbb{R}, n)$ and $\alpha \in \mathcal{A}(m)$.

5. If $M^{-1}(t, x)$ is the inverse $M(t, x)$, then $M_{\alpha, \beta}^{-1} \in \mathcal{K}_0(\mathbb{R}, n - m)$, $\alpha, \beta \in \mathcal{A}(m)$.

6. If $f_i \in \mathcal{K}_{r_i}(\mathbb{R}, n_i)$ for $i = 1, \dots, N$, then

$$\prod_{i=1}^N f_i \in \mathcal{K}_{r_1 + \dots + r_N}(\mathbb{R}, \min_i n_i) \quad \text{and} \quad \sum_{i=1}^N f_i \in \mathcal{K}_{\min_i r_i}(\mathbb{R}, \min_i n_i).$$

Proof : The proof of this lemma is very similar to the corresponding lemma in the first chapter. Notes are made on where the proof differs, rather than providing a full and extensive reproof, to avoid repetition.

1. It is clear that if $f(t, \cdot)$ n -times differentiable and $\partial_\alpha f(\cdot, \cdot)$ is continuous then the same is true of $\int_s^\cdot f(u, x) dB_u^i$ for $i = 0, \dots, d$, with

$$\partial_\alpha \int_s^\cdot f(u, x) dB_u^i = \int_s^\cdot \partial_\alpha f(u, x) dB_u^i.$$

The remainder of the proof is analogous.

Proof of 2: The fact that $a_{\alpha, \beta}(t, \cdot), b_{\alpha, \beta}(t, \cdot)$ are k -times differentiable with partial derivatives of order $|\gamma|$, which are jointly continuous in (t, x) , and which are in $\mathbb{D}^{k-|\gamma|, p}$ for all $p \geq 1$ follows from Theorem 0.6.2 and Theorem 0.6.3. The appropriate bounds can be seen to hold by observing the expression for $a_{\alpha, \beta}$ and applying Proposition 2.5. The corresponding result for $b_{\alpha, \beta}$ is derived in an analogous way.

Proof of 3: This follows easily from **1, 2**.

Proof of 4: From Nualart [33][Prop 1.3.3] we have the following:

$$\langle Du, k_\alpha \rangle_H = u \delta(k_\alpha) - \delta(u k_\alpha)$$

Moreover, we know that for each $p \geq 1$ there holds $u \in \mathbb{D}^{n,p}$, $k_\alpha \in \mathbb{D}^{(k-m-1),p}$, and that $\delta : \mathbb{D}^{k,p} \rightarrow \mathbb{D}^{k-1,p}$ ², hence it is clear that $\langle Du, k_\alpha \rangle_H \in \mathbb{D}^{n \wedge (k-m-1),q}$ for any $q \geq 1$. The existence of regular derivatives of orders less than $n \wedge (k-m-1)$ follows from direct differentiation, and the required bounds follow from 6.

Proof of 5: The k -times differentiability of the inverse (in x) and joint continuity (in (t, x)) is a result of the inverse function theorem. The Malliavin differentiability of the matrix inverse can be deduced by extending Lemma A.4, for square matrices with elements of a general Malliavin differentiability.

Proof of 6: It is clear and straightforward to demonstrate that the differentiability and joint continuity are inherited from the constituent functions. The level of differentiability is a result of the product rule for differentiation. The second property of a K-S-process can be shown in a similar way, making sure to take care with the finite level of differentiability. ■

In what follows and unless otherwise stated we assume $f \in C_b^\infty$.

Theorem 3.10 [Integration by Parts formula I] *Under the conditions (C, UFG) the following integration by parts formula holds for $\Phi \in \mathcal{K}_r^{loc}(\mathbb{R}, n)$, and for any $\alpha \in \mathcal{A}(m)$:*

$$\mathbb{E} [\Phi(t, x) V_{[\alpha]}(f \circ X_t)(x)] = t^{-\|\alpha\|/2} \mathbb{E} [\Phi_\alpha(t, x) f(X_t^x)], \quad (\text{A.9})$$

where $\Phi_\alpha \in \mathcal{K}_r^{loc}(\mathbb{R}, (n-1) \wedge (k-m-1))$. Moreover,

$$\sup_{t \in (0,1]} \mathbb{E} |\Phi_\alpha(t, x)|^p \leq C_p (1 + |x|)^p \sup_{t \in (0,1]} \mathbb{E} |\Phi(t, x)|^p$$

Corollary 3.11 [Integration by Parts formula II] *Under the same conditions as Theorem 3.10, the following integration by parts formula holds for $\Phi \in \mathcal{K}_r^{loc}(\mathbb{R}, n)$, and for any $\alpha \in \mathcal{A}(m)$:*

$$\mathbb{E} [\Phi(t, x) V_{[\alpha]} f(X_t^x)] = t^{-\|\alpha\|/2} \mathbb{E} [\Phi'_\alpha(t, x) f(X_t^x)], \quad (\text{A.10})$$

where $\Phi'_\alpha \in \mathcal{K}_r^{loc}(\mathbb{R}, (n-1) \wedge (k-m-1))$. Moreover,

$$\sup_{t \in (0,1]} \mathbb{E} |\Phi'_\alpha(t, x)|^p \leq C_p (1 + |x|)^p \sup_{t \in (0,1]} \mathbb{E} |\Phi(t, x)|^p.$$

²cf Proposition 1.5.4 in Nualart [33]

Corollary 3.12 [Integration by Parts formula III] *Under the same conditions as Theorem 3.10, the following integration by parts formula holds for $\Phi \in \mathcal{K}_r^{loc}(\mathbb{R}, n)$, and for any $\alpha \in \mathcal{A}(m)$:*

$$V_{[\alpha]} \mathbb{E} [\Phi(t, x) f(X_t^x)] = t^{-\|\alpha\|/2} \mathbb{E} [\Phi_\alpha''(t, x) f(X_t^x)], \quad (\text{A.11})$$

where $\Phi_\alpha'' \in \mathcal{K}_r^{loc}(\mathbb{R}, (n-1) \wedge (k-m-1))$. Moreover,

$$\sup_{t \in (0,1]} \mathbb{E} |\Phi_\alpha''(t, x)|^p \leq C_p (1 + |x|)^p \sup_{t \in (0,1]} \mathbb{E} |\Phi(t, x)|^p.$$

Corollary 3.13 [Integration by Parts Formula IV] *Under the same conditions as Theorem 3.10, the following integration by parts formula holds for $N + M \leq k - m$ and $\alpha_1, \dots, \alpha_N \in \mathcal{A}(m)$:*

$$V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t (V_{[\alpha_{N+1}]} \dots V_{[\alpha_{N+M}]} f)(x) = t^{-\frac{\|\alpha_1\| + \dots + \|\alpha_{N+M}\|}{2}} \mathbb{E} [\Phi_\alpha(t, x) f(X_t^x)],$$

where $\Phi_\alpha \in \mathcal{K}_0^{loc}(\mathbb{R}, (k-m-N))$. Moreover,

$$\sup_{t \in (0,1]} \mathbb{E} |\Phi_\alpha(t, x)|^p \leq C_p (1 + |x|)^{(N+M)p}$$

Corollary 3.14 [IBPF for a Semigroup with a Potential] *Under the conditions (C, UFG) the following integration by parts formula holds for $N \leq k - m$ and $\alpha_1, \dots, \alpha_N \in \mathcal{A}(m)$:*

$$V_{[\alpha_1]} \dots V_{[\alpha_N]} P_t^c f(x) = t^{-(\|\alpha_1\| + \dots + \|\alpha_N\|)/2} \mathbb{E} [\Phi_{\alpha_1, \dots, \alpha_N}^c(t, x) f(X_t^x)], \quad (\text{A.12})$$

where $\Phi_{\alpha_1, \dots, \alpha_N}^c \in \mathcal{K}_0^{loc}(\mathbb{R}, (k-m-N))$. Moreover,

$$\sup_{t \in (0,1]} \mathbb{E} |\Phi_{\alpha_1, \dots, \alpha_N}^c(t, x)|^p \leq C_p (1 + |x|)^{Np}$$

Proof of Theorem 3.10 : : It was shown in the last section that:

$$V_{[\alpha]} (f \circ X_t)(x) = t^{-\|\alpha\|/2} \sum_{\beta \in \mathcal{A}(m)} t^{-\|\beta\|/2} M_{\alpha, \beta}^{-1}(t, x) D^{(\beta)} f(X_t^x)$$

$\mathbb{P}$ -a.s. Hence, using:

$$\begin{aligned} D^\beta(\Phi(t, x) M_{\beta, \gamma}^{-1}(t, x) f(X_t^x)) &= D^{(\gamma)}\Phi(t, x) M_{\beta, \gamma}^{-1}(t, x) f(X_t^x) \\ &\quad + \Phi(t, x) D^{(\gamma)}M_{\beta, \gamma}^{-1}(t, x) f(X_t^x) \\ &\quad + \Phi(t, x) M_{\beta, \gamma}^{-1}(t, x) D^{(\beta)}f(X_t^x) \end{aligned}$$

We get that:

$$\begin{aligned} \mathbb{E}[\Phi(t, x) V_{[\alpha]}(f \circ X_t)(x)] &= t^{-\|\beta\|/2} \sum_{\gamma \in \mathcal{A}(m)} t^{-\|\gamma\|/2} \mathbb{E}[\Phi(t, x) M_{\beta, \gamma}^{-1}(t, x) D^{(\gamma)}f(X_t^x)] \\ &= t^{-\|\beta\|/2} \sum_{\gamma \in \mathcal{A}(m)} t^{-\|\gamma\|/2} \mathbb{E}[f(X_t^x) \{ \Phi(t, x) M_{\beta, \gamma}^{-1}(t, x) \delta(k_\gamma(t, x)) \\ &\quad - \Phi(t, x) D^{(\gamma)}M_{\beta, \gamma}^{-1}(t, x) - D^{(\gamma)}\Phi(t, x) M_{\beta, \gamma}^{-1}(t, x) \}] \end{aligned}$$

And so

$$\begin{aligned} \Phi_\alpha(t, x) &= \sum_{\gamma \in \mathcal{A}(m)} t^{-\|\gamma\|/2} \left\{ \Phi(t, x) M_{\beta, \gamma}^{-1}(t, x) \delta(k_\gamma(t, x)) - \Phi(t, x) D^{(\gamma)}M_{\beta, \gamma}^{-1}(t, x) \right. \\ &\quad \left. - D^{(\gamma)}\Phi(t, x) M_{\beta, \gamma}^{-1}(t, x) \right\} \end{aligned}$$

That $\Phi_\alpha \in \mathcal{K}_r^{\text{loc}}([n-1] \wedge [k-m-1])$ follows from Lemma 2.19 parts 3,4,5,6. N.B. Observe that as the process $k_\gamma(t, x)(\cdot)$ is $\mathcal{F}_u$ -adapted, the adjoint $\delta(k_\gamma(t, x))$, is nothing more than the Itô integral of $k_\gamma(t, x)$ with respect to the d -dimensional Brownian motion $B_t = (B_t^1, \dots, B_t^d)$, i.e.

$$\delta(k_\gamma(t, x)) = \sum_{i=1}^d \int_0^t k_\gamma(t, x)(s) dW_s^i.$$

Thus, it follows that for processes $f \in K_{r,s}(H, n)$ which are a.e. adapted as stochastic processes in H , that $\delta(f) := \delta(f(\cdot, \cdot)) \in \mathcal{K}_{r+1}^{\text{loc}}(n)$. Finally, we observe that, due to (3.4)

$$\sup_{t \in (0,1]} \mathbb{E} \left| D^{(\gamma)}M_{\beta, \gamma}^{-1}(t, x) \right|^p \leq C(1 + |x|)^p \quad (\text{A.13})$$

Hence, the bound

$$\sup_{t \in (0,1]} \mathbb{E} |\Phi_\alpha(t, x)|^p \leq C_p(1 + |x|)^p \sup_{t \in (0,1]} \|\Phi(t, x)\|_{2,q}^p$$

follows by applying the following to the expression for $\Phi_\alpha(t, x)$: (A.14), Hölder's inequality, and the uniform boundedness of the L^r norm of M^{-1} and k_γ over $(t, x) \in (0, 1] \times \mathbb{R}^N$ for each $r \geq 1$.

Proof of Corollary 2.21: The first observation is the following relationship:

$$\begin{aligned} (V_{[\alpha]}f)(X_t^x) &= \nabla f(X_t^x) V_{[\alpha]}(X_t^x) \\ &= (J_t^x)^{-T} \nabla(f \circ X_t)(x) V_{[\alpha]}(X_t^x) \\ &= \nabla(f \circ X_t)(x) (J_t^x)^{-1} V_{[\alpha]}(X_t^x), \end{aligned}$$

where $(J_t^x)^{-T} := ((J_t^x)^{-1})^T$. At this point refer back to the closed linear system of equations, which induced the expression:

$$(J_t^x)^{-1} V_{[\alpha]}(X_t^x) = \sum_{\beta \in \mathcal{A}(m)} a_{\alpha, \beta}(t, x) V_{[\beta]}(x).$$

Again, the central position of the UFG condition is emphasised.

$$\begin{aligned} \nabla(f \circ X_t)(x) (J_t^x)^{-1} V_{[\alpha]}(X_t^x) &= \sum_{\beta \in \mathcal{A}(m)} a_{\alpha, \beta}(t, x) \nabla(f \circ X_t)(x) V_{[\beta]}(x) \\ &= \sum_{\beta \in \mathcal{A}(m)} a_{\alpha, \beta}(t, x) V_{[\beta]}(f \circ X_t)(x). \end{aligned}$$

From Lemma 2.19, $a_{\alpha, \beta} \in \mathcal{K}_{[\|\beta\| - \|\alpha\|] \vee 0}^{\text{loc}}(k - m - 1)$. Hence, it has been shown that:

$$\mathbb{E} [\Phi(t, x) V_{[\alpha]} f(X_t^x)] = \sum_{\beta \in \mathcal{A}(m)} \mathbb{E} [\Phi(t, x) a_{\alpha, \beta}(t, x) V_{[\beta]}(f \circ X_t)(x)].$$

The integration by parts formula (2.26) can then be applied N_m times, after noting that the product $\Phi a_{\alpha, \beta} \in \mathcal{K}_{r + [\|\beta\| - \|\alpha\|] \vee 0}^{\text{loc}}([n - 1] \wedge [k - m - 1])$. And so,

$$\begin{aligned} \mathbb{E} [\Phi(t, x) V_{[\alpha]} f(X_t^x)] &= \sum_{\beta \in \mathcal{A}(m)} t^{-\frac{\|\beta\|}{2}} \mathbb{E} [\Psi_{\beta}(t, x) f(X_t^x)] \\ &= \sum_{\beta \in \mathcal{A}(m)} t^{-\frac{\|\beta\|}{2}} t^{-\frac{\|\alpha\| - \|\beta\|}{2}} \mathbb{E} \left[t^{\frac{\|\alpha\| - \|\beta\|}{2}} \Psi_{\beta}(t, x) f(X_t^x) \right] \\ &= t^{-\frac{\|\alpha\|}{2}} \mathbb{E} [\Phi_{\alpha}(t, x) f(X_t^x)], \end{aligned}$$

where $\Phi_{\alpha} = \sum_{\beta \in \mathcal{A}(m)} t^{\frac{\|\alpha\| - \|\beta\|}{2}} \Psi_{\beta} \in \mathcal{K}_r^{\text{loc}}((n - 1) \wedge (k - m - 1))$. Finally, we again observe that, due to (3.4)

$$\sup_{t \in (0, 1]} \mathbb{E} \left| D^{(\gamma)} M_{\beta, \gamma}^{-1}(t, x) \right|^p \leq C(1 + |x|)^p \quad (\text{A.14})$$

Hence, the bound

$$\sup_{t \in (0, 1]} \mathbb{E} |\Phi'_{\alpha}(t, x)|^p \leq C_p(1 + |x|)^p \sup_{t \in (0, 1]} \|\Phi(t, x)\|_{2, q}^p$$

follows by applying the following to the expression for $\Phi'_\alpha(t, x)$: (A.14), Hölder's inequality, and the uniform boundedness of the L^r norm of M^{-1} and k_γ over $(t, x) \in (0, 1] \times \mathbb{R}^N$ for each $r \geq 1$.

Proof of Corollary 3.12 :

Observe that

$$\begin{aligned} V_{[\alpha]} \mathbb{E}[\Phi(t, x) f(X_t^x)] &= \mathbb{E}[V_{[\alpha]} \Phi(t, x) f(X_t^x) + \Phi(t, x) V_{[\alpha]}(f \circ X_t)(x)] \\ &\stackrel{(i)}{=} \mathbb{E}[V_{[\alpha]} \Phi(t, x) f(X_t^x) + \Phi_\alpha(t, x) f(X_t^x)] \\ &= t^{\|\alpha\|/2} \mathbb{E}[\Phi''_\alpha(t, x) f(X_t^x)], \end{aligned}$$

where $\Phi''_\alpha(t, x) = t^{\|\alpha\|/2} V_{[\alpha]} \Phi(t, x) + \Phi_\alpha(t, x)$. To prove that $\Phi_\alpha \in \mathcal{K}_r^{\text{loc}}([n-1] \wedge [k-m-1])$, we first observe that

$$|V_{[\alpha]}(x)| \leq C_\alpha(1 + |x|).$$

This follows from the global Lipschitz continuity of $V_0, \dots, V_d$, (and hence boundedness of their derivatives) and the form of $V_{[\alpha]}$. Hence, $V_{[\alpha]}$ is bounded on compact subsets of $\mathbb{R}^N$, and $t^{\|\alpha\|/2} V_{[\alpha]} \Phi \in \mathcal{K}_{r+\|\alpha\|}^{\text{loc}}(n-1)$. We already know from Theorem 3.10 that $\Phi_\alpha \in \mathcal{K}_r^{\text{loc}}([n-1] \wedge [k-m-1])$, thus $\Phi''_\alpha \in \mathcal{K}_r^{\text{loc}}([n-1] \wedge [k-m-1])$. Moreover, it is clear that:

$$V_{[\alpha]} \Phi(t, x) \leq |V_{[\alpha]}| |\nabla \Phi(t, x)| \leq C_\alpha(1 + |x|) |\nabla \Phi(t, x)|.$$

This, combined with the same arguments for Φ_α in the previous proof, completes the argument.

Proof of Corollary 3.13 :

Once it is noted that $\text{Id} \in \mathcal{K}_0$, the proof follows from N applications of Corollary 3.12 and M applications of Corollary 3.11. The bound follows by recursively applying (3.9).

Proof of Corollary 3.14 :

The potential term need not provide any extra complications as long as we treat it as a Kusuoka-Stroock process. That is, we put:

$$\Phi(t, x) := \exp \left(\int_0^t c(X_s^x) ds \right).$$

It is straightforward to show that $\Phi \in \mathcal{K}_0^{\text{loc}}(k-m)$. The proof is completed by applying Corollary 3.13. ■

A.4. 'The LFG condition and local differentiability of the semigroup'

Lemma 4.4 Define, for $\gamma \in (0, 1/2)$ and $p > 2$, $Y : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ by:

$$Y_s^x := \int \int_{[0,s]^2} \frac{|X_u^x - X_v^x|^{2p}}{|u - v|^{1+2p\gamma}} du dv.$$

Then $Y_s^x \in \mathbb{D}^\infty$ for each $s \in [0, t]$, and for all $\gamma \in (\frac{1}{2p}, \frac{1}{2} - \frac{1}{2p})$. Moreover, for any $U' \subset U$, open, such that $d(U', \partial U) > 0$ and any $(k, p) \in \mathbb{N}_0 \times [1, \infty)$ there holds

$$\sup_{s \in (0, t] x \in U'} \|Y_s^x\|_{k, p} < \infty. \quad (\text{A.15})$$

Proof : We show first that $Y_s \in L^p(\Omega)$ for each $s \in \mathbb{R}$. Our first observation comes from Kunita [19][Thm 2.1] and is that for some constant depending only on p :

$$\mathbb{E} |X_v^x - X_u^x|^p \leq C_p(1 + |x|^p) |v - u|^{\frac{p}{2}} \quad (\text{A.16})$$

This implies that $\sup_{x \in U'} \mathbb{E} |X_v^x - X_u^x|^{2p} \leq C_{p, U'} |u - v|^p$. Hence,

$$\begin{aligned} \mathbb{E} |Y_s^x|^q &\leq \int \int_{[0,s]^2} \frac{\mathbb{E} |X_u^x - X_v^x|^{2pq}}{|u - v|^{(1+2p\gamma)q}} du dv \\ &\leq C_{p, q, d} \int \int_{[0,s]^2} |u - v|^{pq - (1+2p\gamma)q} du dv. \end{aligned} \quad (\text{A.17})$$

So choosing γ such that $pq - (1 + 2p\gamma)q > 0$ - that is, $\gamma < \frac{1}{2} - \frac{1}{2p}$ - results in

$$\sup_{\substack{s \in (0, t] \\ x \in U'}} \mathbb{E} |Y_s^x|^q < \infty.$$

Now we show $Y_s^x \in \mathbb{D}^{k, q}$ for each $k, q \geq 1$. Our first observation is that analogous bounds to (A.16) exist for (Malliavin and standard) derivatives of SDE solutions. The bound for the standard derivatives comes from Kunita [19][Thm 3.3]. The bound for the Malliavin derivatives holds from applying analogous techniques as in Kunita [19].

$$\begin{aligned} \mathbb{E} \left\| D^k [X_v^x - X_u^x] \right\|_{H^{\otimes k}}^p &\leq C_{k, p}(1 + |x|^p) |v - u|^{\frac{p}{2}} \\ \mathbb{E} \left| \frac{\partial^{|\alpha|}}{\partial x^\alpha} [X_v^x - X_u^x] \right|^p &\leq \tilde{C}_{\alpha, p}(1 + |x|^p) |v - u|^{\frac{p}{2}} \end{aligned}$$

We now turn to the Faà di Bruno formula (see for example Craik [7] which generalises the chain rule to higher derivatives to see that, for $p \in \mathbb{Z}$ with $p \geq 1$:

$$D^k |X_u^x - X_v^x|^{2p} = \sum_{\pi \in \Pi} \frac{(2p)!}{(2p - |\pi|)!} |X_u^x - X_v^x|^{2p - |\pi|} \prod_{B \in \pi} D^{|B|} (X_u^x - X_v^x)$$

where Π is the set of all partitions of $\{1, \dots, k\}$, $|\pi|$ represents the number of blocks in the partition, π , and $|B|$ is the size of the block B . Then it follows that, with $\frac{1}{r_1} + \sum_{B \in \pi} \frac{1}{r(B)} = 1$ for each $\pi \in \Pi$,

$$\begin{aligned} & \mathbb{E} \left\| D^k |X_u^x - X_v^x|^{2p} \right\|_{H^{\otimes k}}^q \\ & \leq C(q) \sum_{\pi \in \Pi} \frac{(2p)!}{(2p - |\pi|)!} \mathbb{E} \left\| |X_u^x - X_v^x|^{2p - |\pi|} \prod_{B \in \pi} D^{|B|} (X_u^x - X_v^x) \right\|_{H^{\otimes k}}^q \\ & \leq C(p, q) \sum_{\pi \in \Pi} \left\| |X_u^x - X_v^x|^{2p - |\pi|} \right\|_{L^{r_1}(\Omega)}^q \prod_{B \in \pi} \left\| D^{|B|} (X_u^x - X_v^x) \right\|_{L^{r(B)}}^q \\ & \leq C(p, q, t, U') \sum_{\pi \in \Pi} |u - v|^{(p - \frac{|\pi|}{2})q} \prod_{B \in \pi} |u - v|^{\frac{q}{2}} \\ & \leq \tilde{C}(p, q, k, t, U') |u - v|^{pq}. \end{aligned}$$

This may be used in a similar fashion to prove, for all $x \in U' \subset \mathbb{R}^N$

$$\mathbb{E} \left\| D^k Y_s^x \right\|_{H^{\otimes k}}^q \leq \tilde{C}(p, q, k, t, U'),$$

The same techniques can also be applied to prove a similar result for $\mathbb{E} \left| \frac{\partial^{|\alpha|}}{\partial x^\alpha} Y_s^x \right|^q$, and combinations of the two different types of derivative. Last but not least, we state and proof a result which was used in (4.9). Namely,

$$\mathbb{E} \left| Y_{t/y}^x \right|^p \leq C(n, p) \left(\frac{t}{y} \right)^p. \quad (\text{A.18})$$

Indeed, we may prove something even stronger (the full strength was not needed for our requirements). Proceeding as in (A.17), by instead using Jensen's inequality (twice) we see

$$\begin{aligned} \mathbb{E} |Y_t^x|^q & \leq C t^{2(q-1)} \int \int_{[0, t]^2} |u - v|^{pq - (1+2p\gamma)q} du dv \\ & \leq C t^{2(q-1)} \int \int_{[0, t]^2} 2^{pq - (1+2p\gamma)q} du dv \\ & \leq \tilde{C} t^{2q} \\ & \leq \tilde{C} t^q \end{aligned}$$

Note we have used that $t \leq 1$ and that $|u - v| < 2$ for $u, v \in [0, t]$. ■

Bibliography

- [1] Hélène Airault and Paul Malliavin, *Intégration géométrique sur l'espace de Wiener*, Bull. Sci. Math. (2) **112** (1988), no. 1, 3–52.
- [2] Dominique Bakry, *L'hypercontractivité et son utilisation en théorie des semigroupes*, Lectures on probability theory (Saint-Flour, 1992), Lecture Notes in Math., vol. 1581, Springer, Berlin, 1994, pp. 1–114.
- [3] Dominique Bakry and Michel Ledoux, *Lévy-Gromov's isoperimetric inequality for an infinite-dimensional diffusion generator*, Invent. Math. **123** (1996), no. 2, 259–281.
- [4] J.-M. Bismut, *Mécanique aléatoire*, Tenth Saint Flour Probability Summer School—1980 (Saint Flour, 1980), Lecture Notes in Math., vol. 929, Springer, Berlin, 1982, pp. 1–100.
- [5] Ju. N. Blagoveščenskij and M. I. Freĭdlin, *Some properties of diffusion processes depending on a parameter*, Dokl. Akad. Nauk SSSR **138** (1961), 508–511.
- [6] Patrick Cattiaux and Laurent Mesnager, *Hypoelliptic non-homogeneous diffusions*, Probab. Theory Related Fields **123** (2002), no. 4, 453–483.
- [7] Alex D. D. Craik, *Prehistory of Faà di Bruno's formula*, Amer. Math. Monthly **112** (2005), no. 2, 119–130.
- [8] Dan Crisan and Saadia Ghazali, *On the convergence rates of a general class of weak approximations of SDEs*, Stochastic differential equations: theory and applications, Interdiscip. Math. Sci., vol. 2, World Sci. Publ., Hackensack, NJ, 2007, pp. 221–248.
- [9] K. D. Elworthy, *Stochastic dynamical systems and their flows*, Stochastic analysis (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978), Academic Press, New York, 1978, pp. 79–95.
- [10] Adriano M. Garsia, Eugene R. Rodemich, and Howard Rumsey, Jr., *A real variable lemma and the continuity of paths of some Gaussian processes*, Indiana Univ. Math. J. **20** (1970/1971), 565–578.
- [11] Saadia Ghazali, *The global error in weak approximations of stochastic differential equations*, Ph.D. thesis, University of London, 2006.

- [12] I. I. Gihman, *On a scheme of formation of random processes*, Doklady Akad. Nauk SSSR (N. S.) **58** (1947), 961–964.
- [13] I. L. Gihman, *On some differential equations with random functions*, Ukrain. Mat. Žurnal **2** (1950), no. 3, 45–69.
- [14] Lars Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171.
- [15] Nobuyuki Ikeda and Shinzo Watanabe, *Stochastic differential equations and diffusion processes*, second ed., North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam, 1989. MR 1011252 (90m:60069)
- [16] Jürgen Jost, *Postmodern analysis*, third ed., Universitext, Springer-Verlag, Berlin, 2005.
- [17] Ioannis Karatzas and Steven E. Shreve, *Brownian motion and stochastic calculus*, second ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991.
- [18] Peter E. Kloeden and Eckhard Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics (New York), vol. 23, Springer-Verlag, Berlin, 1992.
- [19] Hiroshi Kunita, *Stochastic differential equations and stochastic flows of diffeomorphisms*, École d'été de probabilités de Saint-Flour, XII—1982, Lecture Notes in Math., vol. 1097, Springer, Berlin, 1984, pp. 143–303.
- [20] ———, *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1990.
- [21] Shigeo Kusuoka, *Approximation of expectation of diffusion process and mathematical finance*, Taniguchi Conference on Mathematics Nara '98, Adv. Stud. Pure Math., vol. 31, Math. Soc. Japan, Tokyo, 2001, pp. 147–165.
- [22] ———, *Malliavin calculus revisited*, J. Math. Sci. Univ. Tokyo **10** (2003), no. 2, 261–277.
- [23] Shigeo Kusuoka and Daniel W. Stroock, *Applications of the Malliavin calculus. I*, Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. 32, North-Holland, Amsterdam, 1984, pp. 271–306.
- [24] ———, *Applications of the Malliavin calculus. II*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **32** (1985), no. 1, 1–76.
- [25] ———, *Applications of the Malliavin calculus. III*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **34** (1987), no. 2, 391–442.
- [26] Michel Ledoux, *The geometry of Markov diffusion generators*, Ann. Fac. Sci. Toulouse Math. (6) **9** (2000), no. 2, 305–366, Probability theory.

- [27] Christian Litterer, *The signature in numerical algorithms*, Ph.D. thesis, Univ. of Oxford, 2008.
- [28] Christian Litterer and Terry Lyons, *Introducing cubature to filtering*, Unpublished (2009).
- [29] Terry Lyons and Nicolas Victoir, *Cubature on Wiener space*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **460** (2004), no. 2041, 169–198, Stochastic analysis with applications to mathematical finance.
- [30] Paul Malliavin, *Stochastic calculus of variation and hypoelliptic operators*, Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976) (New York), Wiley, 1978, pp. 195–263.
- [31] Paul-André Meyer, *Démonstration probabiliste de certaines inégalités de Littlewood-Paley. II. L'opérateur carré du champ*, Séminaire de Probabilités, X (Première partie, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975), Springer, Berlin, 1976, pp. 142–161. Lecture Notes in Math., Vol. 511.
- [32] James Norris, *Simplified Malliavin calculus*, Séminaire de Probabilités, XX, 1984/85, Lecture Notes in Math., vol. 1204, Springer, Berlin, 1986, pp. 101–130.
- [33] David Nualart, *The Malliavin calculus and related topics*, second ed., Probability and its Applications (New York), Springer-Verlag, Berlin, 2006.
- [34] Bernt Øksendal, *An introduction to malliavin calculus with applications to economics*, (1997), Lecture Notes, Norwegian School of Economics and Business Administration.
- [35] L. C. G. Rogers and David Williams, *Diffusions, Markov processes, and martingales. Vol. 2*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2000, Itô calculus, Reprint of the second (1994) edition.
- [36] Jean-Pierre Roth, *Opérateurs carré du champ et formule de Lévy-Kinchine sur les espaces localement compacts*, C. R. Acad. Sci. Paris Sér. A **278** (1974), 1103–1106.
- [37] Marta Sanz-Solé, *Applications of Malliavin calculus to SPDE's*, Stochastic partial differential equations and applications (Trento, 2002), Lecture Notes in Pure and Appl. Math., vol. 227, Dekker, New York, 2002, pp. 429–442.
- [38] A. H. Stroud, *Approximate calculation of multiple integrals*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1971, Prentice-Hall Series in Automatic Computation.
- [39] Héctor J. Sussmann, *Orbits of families of vector fields and integrability of distributions*, Trans. Amer. Math. Soc. **180** (1973), 171–188.
- [40] Hans Triebel, *Interpolation theory, function spaces, differential operators*, North-Holland Mathematical Library, vol. 18, North-Holland Publishing Co., Amsterdam, 1978.

- [41] Hassler Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc. **36** (1934), no. 1, 63–89.
- [42] David Williams, *To begin at the beginning: ...*, Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980), Lecture Notes in Math., vol. 851, Springer, Berlin, 1981, pp. 1–55.