Malliavin calculus in $\mathbb{R}^d$

Jason Swanson

August 6, 2009

Abstract

These notes represent my attempt to better understand the Malliavin calculus by describing how it works in $\mathbb{R}^d$, and comparing it to the ordinary multivariable calculus. These notes are not intended to be self-contained. Facts and theorems will be stated without proof, and the focus will be on understanding the concepts and the mechanics of Malliavin calculus. The main reference for these notes is [1].

1 Introduction

Let H be a real, separable Hilbert space, and let $\{X(h) : h \in H\}$ be a collection of random variables, all defined on the same probability space. We say that $\{X(h)\}$ is an isonormal Gaussian process on H if every $X(h)$ is normally distributed with mean zero, and $E[X(h)X(g)] = \langle h, g\rangle_H$ for all $h, g \in H$.

Given any real, separable Hilbert space H, it is always possible to construct an isonormal Gaussian process on H. (This can be proved using Kolmogorov's extension theorem.) It is also possible to show that an isonormal Gaussian process must be linear, in the sense that $X(\lambda h + \mu g) = \lambda X(h) + \mu X(g)$ for all $\lambda, \mu \in \mathbb{R}$ and all $h, g \in H$.

Example 1.1. Let Z be a standard normal random variable. For each $a \in \mathbb{R}$, define $X(a) = aZ$. Then $\{X(a) : a \in \mathbb{R}\}$ is an isonormal Gaussian process on $\mathbb{R}$. (Note that $\mathbb{R}$ is a one-dimensional Hilbert space, where the inner product is just ordinary multiplication.) The original normal random variable is embedded in this process as $Z = X(1)$.

Example 1.2. Let X be a nondegenerate multinormal random (column) vector in $\mathbb{R}^d$ with mean zero and covariance matrix $\Sigma = E[XX^T]$. For each $u \in \mathbb{R}^d$, define $X(u) = u^T X$. Let $H_\Sigma = \mathbb{R}^d$ with the inner product $\langle u, v\rangle_\Sigma = u^T \Sigma v$. Then $\{X(u)\}$ is an isonormal Gaussian process on H_Σ. The components of the original multinormal random vector are embedded in this process as $X(e_j) = e_j^T X$, where $\{e_j\}$ are the standard basis vectors in $\mathbb{R}^d$.

Example 1.3. Let $\{B(t) : 0 \le t \le 1\}$ be a standard Brownian motion. For each $h \in L^2([0,1])$, define $W(h) = \int_0^1 h(u)\,dB(u)$. It is easy to verify that $\{W(h) : h \in L^2([0,1])\}$ is an isonormal Gaussian process on $L^2([0,1])$. The original Brownian motion is embedded in this process as $B(t) = W(1_{[0,t]})$.

Suppose $\{X(h)\}$ is an isonormal Gaussian process on H, defined on some probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{G} = \sigma(X(h) : h \in H)$. Then any random variable $F : \Omega \to \mathbb{R}$ which is $\mathcal{G}$-measurable may be regarded as a functional of the Gaussian process $\{X(h)\}$. The Malliavin calculus gives us a way to differentiate and integrate these functionals.

Example 1.3 is of primary importance in Malliavin calculus, and can be used to extend the Itô integral to the so-called Skorohod integral. By changing the form of the inner product in Example 1.3, Malliavin calculus can used to study fractional Brownian motion, and other continuous-time Gaussian processes.

Note that the underlying Hilbert space in Example 1.3 is infinite-dimensional. Perhaps the most important applications of Malliavin calculus use infinite-dimensional Hilbert spaces. Example 1.2, however, involves a finite-dimensional Hilbert space. In this setting, the Malliavin calculus is much simpler, and in these notes, we will focus on this example only. We will stress the connection between Malliavin calculus and the ordinary multivariable calculus on $\mathbb{R}^d$.

2 The Hilbert spaces

Throughout these notes, vectors in $\mathbb{R}^d$ will be regarded as column vectors, so that $u^T v$ denotes the ordinary inner product in $\mathbb{R}^d$. The standard basis vectors in $\mathbb{R}^d$ will be denoted by $\{e_j\}$, so that the components of u in the standard basis are $e_j^T u$.

In these notes, we fix a nondegenerate multinormal random vector X in $\mathbb{R}^d$ with mean zero and covariance matrix $\Sigma = E[XX^T]$, defined on a probability space $(\Omega, \mathcal{F}, P)$. We have already seen the Hilbert space $H_\Sigma = \mathbb{R}^d$ with inner product $\langle u, v\rangle_\Sigma = u^T \Sigma v$ and norm $\|u\|_\Sigma^2 = \langle u, u\rangle_\Sigma$. The process $X(u) = u^T X$ is an isonormal Gaussian process on H_Σ.

To obtain an orthonormal basis for H_Σ, let M be the unique upper triangular matrix such that $\Sigma = M^T M$. (This is called the Cholesky decomposition.) Define $v_j = M^{-1} e_j$. Then

$$\langle v_i, v_j\rangle_\Sigma = v_i^T \Sigma v_j = v_i^T M^T M v_j = e_i^T e_j.$$

Therefore, $\{v_j\}$ is an orthonormal basis for H_Σ.

We will denote the components of a vector $u \in H_\Sigma$ in this basis by $u(j) = \langle u, v_j\rangle_\Sigma$. We then have the familiar properties $u = \sum_{j=1}^d u(j) v_j$ and $\langle u_1, u_2\rangle_\Sigma = \sum_{j=1}^d u_1(j) u_2(j)$.

2.1 Tensor products

Let H and $\widehat{H}$ be Hilbert spaces. Recall that any $x \in H$ can be identified with the linear functional $w \mapsto \langle x, w\rangle_H$. Given $x \in H$ and $\widehat{x} \in \widehat{H}$, let us define the tensor product of x and $\widehat{x}$, denoted by $x \otimes \widehat{x}$, as the bilinear functional on $H \times \widehat{H}$ given by

$$(x \otimes \widehat{x})(w, \widehat{w}) = \langle x, w\rangle_H \langle \widehat{x}, \widehat{w}\rangle_{\widehat{H}}.$$

Let $F = \{x \otimes \widehat{x} : x \in H, \widehat{x} \in \widehat{H}\}$ and E the set of all finite linear combinations of elements of F. We define an inner product on F by

$$\langle x \otimes \widehat{x}, y \otimes \widehat{y}\rangle = \langle x, y\rangle_H \langle \widehat{x}, \widehat{y}\rangle_{\widehat{H}},$$

and extend this inner product to E by linearity. The tensor product of H and $\widehat{H}$ is the completion of E under this inner product, and is denoted by $H \otimes \widehat{H}$. We may similarly define the tensor product of any finite number of Hilbert spaces. We write $H^{\otimes n}$ for the tensor product of H with itself n times, and similarly for $x^{\otimes n}$. It can be shown that if $\{e_j\}$ and $\{\widehat{e}_k\}$ are orthonormal bases for H and $\widehat{H}$ respectively, then $\{e_j \otimes \widehat{e}_k\}$ is an orthonormal basis for $H \otimes \widehat{H}$.

An orthonormal basis for the Hilbert space $H_\Sigma^{\otimes m}$ is therefore

$$\{v_{j_1} \otimes \cdots \otimes v_{j_m} : 1 \leq j_1, \ldots, j_m \leq d\}.$$

We will write the components of $u \in H_\Sigma^{\otimes m}$ as

$$u(j_1, \ldots, j_m) = \langle u, v_{j_1} \otimes \cdots \otimes v_{j_m} \rangle_{H_\Sigma^{\otimes m}},$$

so that

$$u = \sum_{j_1, \ldots, j_m = 1}^{d} u(j_1, \ldots, j_m) v_{j_1} \otimes \cdots \otimes v_{j_m},$$

and

$$\langle u, v \rangle_{H_\Sigma^{\otimes m}} = \sum_{j_1, \ldots, j_m = 1}^{d} u(j_1, \ldots, j_m) v(j_1, \ldots, j_m).$$

In particular, note that we can identify $H_\Sigma^{\otimes 2}$ with the set of $d \times d$ matrices, $H_\Sigma^{\otimes 3}$ with the set of $d \times d \times d$ arrays, and so on. More generally, we can identify $H_\Sigma^{\otimes m}$ with the L^2 space of real-valued functions on $\{1, \ldots, d\}^m$ equipped with counting measure.

If $u \in H_\Sigma^{\otimes p}$ and $v \in H_\Sigma^{\otimes q}$, then $u \otimes v \in H_\Sigma^{\otimes (p+q)}$, and

$$\begin{aligned} u \otimes v &= \sum_{i_1, \ldots, i_p = 1}^{d} \sum_{j_1, \ldots, j_q = 1}^{d} u(i_1, \ldots, i_p) v(j_1, \ldots, j_q) (v_{i_1} \otimes \ldots \otimes v_{i_p}) \otimes (v_{j_1} \otimes \ldots \otimes v_{j_q}) \\ &= \sum_{j_1, \ldots, j_{p+q} = 1}^{d} u(j_1, \ldots, j_p) v(j_{p+1}, \ldots, j_{p+q}) v_{j_1} \otimes \ldots \otimes v_{j_{p+q}}. \end{aligned}$$

In other words,

$$u \otimes v(j_1, \ldots, j_{p+q}) = u(j_1, \ldots, j_p) v(j_{p+1}, \ldots, j_{p+q}),$$

which gives a component-wise formula for the tensor product.

We say that $u \in H_\Sigma^{\otimes m}$ is symmetric if $u(j_1, \ldots, j_m) = u(j_{\sigma(1)}, \ldots, j_{\sigma(m)})$, for all $\sigma \in S_m$, where S_m is the group of permutations of $\{1, \ldots, m\}$. If $u \in H_\Sigma^{\otimes m}$, then we define the symmetrization of u as

$$\widetilde{u}(j_1, \ldots, j_m) = \frac{1}{m!} \sum_{\sigma \in S_m} u(j_{\sigma(1)}, \ldots, j_{\sigma(m)}).$$

If $u \in H_\Sigma^{\otimes p}$ and $v \in H_\Sigma^{\otimes q}$ are symmetric, and $r \in \{1, 2, \ldots, p \wedge q\}$, then the contraction of r indices of u and v is denoted by $u \otimes_r v$, and is the element of $H_\Sigma^{\otimes (p+q-2r)}$ whose components are

$$u \otimes_r v(j_1, \dots, j_{p+q-2r}) = \sum_{i_1,\dots,i_r=1}^{d} u(j_1, \dots, j_{p-r}, i_1, \dots, i_r) v(j_{p-r+1}, \dots, j_{p+q-2r}, i_1, \dots, i_r).$$

Note that $H_\Sigma^{\otimes 0}$ is just the space of scalars, $\mathbb{R}$; every $u \in H_\Sigma^{\otimes 1} = H_\Sigma$ is symmetric, and if $u, v \in H_\Sigma$, then $u \otimes_1 v = \langle u, v\rangle_\Sigma \in H_\Sigma^{\otimes 0}$. We will also adopt the convention that $u \otimes_0 v = u \otimes v$.

2.2 The function spaces

Recall that X is a multinormal random vector defined on $(\Omega, \mathcal{F}, P)$. Let $\mathcal{G} = \sigma(X)$. If H is a Hilbert space, we will write $L^2(\mathcal{G}; H)$ for the space of square-integrable functions mapping $(\Omega, \mathcal{G}, P)$ to H. Note that $\langle F, G\rangle_{L^2(\mathcal{G};H)} = E[\langle F, G\rangle_H]$ for all $F, G \in L^2(\mathcal{G}; H)$. For simplicity, we write $L^2(\mathcal{G}) = L^2(\mathcal{G}; \mathbb{R})$. The inner product and norm in $L^2(\mathcal{G})$ will be denoted by $\langle \cdot, \cdot \rangle_\mathcal{G}$ and $\| \cdot \|_\mathcal{G}$.

The space $L^2(\mathcal{G})$ is the primary setting for the Malliavin calculus. If a function $F : \Omega \to \mathbb{R}$ belongs to $L^2(\mathcal{G})$ and is "smooth enough" (in a sense to be made precise later), then we will see how to differentiate F. Recall that any function $F : \Omega \to \mathbb{R}$ which is $\mathcal{G}$-measurable can be written as $F = f(X)$ for some measurable function $f : \mathbb{R}^d \to \mathbb{R}$. It is natural, then, to suspect that there might be a connection between the Malliavin derivative of F and the ordinary derivative of f. More generally, we will explore the connection between the Malliavin calculus on $L^2(\mathcal{G})$ and the theory of L^2-derivatives on $\mathbb{R}^d$.

If ϖ is a Borel measure on $\mathbb{R}^d$ and H is a Hilbert space, then we will write $L^2(\varpi; H)$ for the space of measurable functions from $\mathbb{R}^d$ to H which are square-integrable with respect to the measure ϖ. Note that $\langle f, g\rangle_{L^2(\varpi;H)} = \int \langle f, g\rangle_H \, d\varpi$. For simplicity, we write $L^2(\varpi) = L^2(\varpi; \mathbb{R})$, and denote the inner product and norm on $L^2(\varpi)$ by $\langle \cdot, \cdot \rangle_\varpi$ and $\| \cdot \|_\varpi$. When necessary, we will use λ to denote Lebesgue measure on $\mathbb{R}^d$.

Let ν be the law of X, that is, $\nu(A) = P(X \in A)$ for all Borel sets A. Note that $d\nu = p \, d\lambda$, where

$$p(x) = \frac{1}{(2\pi)^{d/2}\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2} x^T \Sigma^{-1} x\right). \tag{2.1}$$

The map $f \mapsto F = f(X)$ provides a natural isomorphism from $L^2(\nu)$ to $L^2(\mathcal{G})$. As we will see, in the context of these notes, all of the main concepts and methods of Malliavin calculus can be phrased in terms of the ordinary calculus for functions in $L^2(\nu)$.

2.3 Hermite polynomials and basis vectors

For $n \geq 0$, let

$$H_n(x) = \frac{(-1)^n}{n!} e^{x^2/2} \frac{d^n}{dx^n}(e^{-x^2/2})$$

denote the n-th Hermite polynomial. The first few Hermite polynomials are $H_0(x) = 1$, $H_1(x) = x$, $H_2(x) = \frac{1}{2}(x^2 - 1)$, and $H_3(x) = \frac{1}{3!}(x^3 - 3x)$. When necessary, we adopt the convention that $H_{-1}(x) = 0$. It can be shown that

$$H_n'(x) = H_{n-1}(x), \tag{2.2}$$
$$(n+1)H_{n+1}(x) = xH_n(x) - H_{n-1}(x), \tag{2.3}$$
$$H_n(-x) = (-1)^n H_n(x), \tag{2.4}$$

for all $n \geq 0$. We will also sometimes use the normalization $h_n(x) = n!H_n(x)$.

The Hermite polynomials can be used to construct an orthonormal basis for $L^2(\mathcal{G})$ in the following way. Let $\Lambda = (\mathbb{N} \cup \{0\})^d$. For $a \in \Lambda$, we set $a! = \prod_{j=1}^d a_j!$, $|a| = \sum_{j=1}^d a_j$, and $x^a = \prod_{j=1}^d x_j^{a_j}$ for all $x \in \mathbb{R}^d$. Define $H_a : \mathbb{R}^d \to \mathbb{R}$ by

$$H_a(x) = \prod_{j=1}^d H_{a_j}(x_j).$$

If

$$\Phi_a = \sqrt{a!} \prod_{j=1}^d H_{a_j}(X(v_j)),$$

then it can be shown that $\{\Phi_a : a \in \Lambda\}$ is an orthonormal basis for $L^2(\mathcal{G})$.

The space spanned by $\{\Phi_a : a \in \Lambda, |a| = m\}$ is called the Wiener chaos of order m, and is denoted by $\mathcal{H}_m$. It can also be defined as the closed linear subspace of $L^2(\mathcal{G})$ generated by the random variables $\{H_m(X(v)) : v \in \mathbb{R}^d, \|v\|_\Sigma = 1\}$. Clearly, for $n \neq m$, $\mathcal{H}_n$ and $\mathcal{H}_m$ are orthogonal, and $L^2(\mathcal{G}) = \bigoplus_{m=0}^\infty \mathcal{H}_m$. Hence, any $F \in L^2(\mathcal{G})$ can be decomposed uniquely as $F = \sum_{m=0}^\infty Y_m$, where $Y_m \in \mathcal{H}_m$ and the sum converges in $L^2(\Omega)$. More specifically, $F = \sum_{m=0}^\infty \sum_{|a|=m} E[\Phi_a F]\Phi_a$. Since $\Phi_0 = 1$, this gives

$$F = EF + \sum_{m=1}^\infty Y_m,$$

where $Y_m = \sum_{|a|=m} E[\Phi_a F]\Phi_a \in \mathcal{H}_m$.

The space $\mathcal{H}_0$ is the set of constant random variables, and the space $\mathcal{H}_1$ consists of linear combinations of the components of X. In fact, according to the remark following the proof of Theorem 1.1.1 in [1], the space $\bigoplus_{j=0}^m \mathcal{H}_j$ is just the set of random variables of the form $p(X)$, where $p : \mathbb{R}^d \to \mathbb{R}$ is a polynomial of degree $k \leq m$. This implies, that in the context of these notes, each $\mathcal{H}_m$ is a finite-dimensional subspace of $L^2(\mathcal{G})$.

In order to transport all this machinery to $L^2(\nu)$, we define the map $\mathcal{I} : L^2(\nu) \to L^2(\mathcal{G})$ by $\mathcal{I}f = f(X)$, and observe that $\mathcal{I}$ is an isomorphism. Thus, if

$$\varphi_a(x) = (\mathcal{I}^{-1}\Phi_a)(x) = \sqrt{a!} \prod_{j=1}^d H_{a_j}(v_j^T x),$$

then $\{\varphi_a : a \in \Lambda\}$ is an orthonormal basis for $L^2(\nu)$.

Note that $\mathcal{I}^{-1}\mathcal{H}_m$ is the space spanned by $\{\varphi_a : a \in \Lambda, |a| = m\}$, and can also be defined as the closed linear subspace of $L^2(\nu)$ generated by the functions $\{H_m(v^T \cdot) : v \in \mathbb{R}^d, \|v\|_\Sigma =$

1}. We can decompose $L^2(\nu)$ into orthogonal subspace as $L^2(\nu) = \bigoplus_{m=0}^{\infty} \mathcal{I}^{-1}\mathcal{H}_m$. The space $\bigoplus_{j=0}^{m} \mathcal{I}^{-1}\mathcal{H}_m$ is simply the space of polynomials of degree $k \le m$, and any $f \in L^2(\nu)$ can be decomposed as

$$f = \langle 1, f \rangle_\nu + \sum_{m=1}^{\infty} f_m,$$

where $f_m = \mathcal{I}^{-1} Y_m = \sum_{|a|=m} \langle \varphi_a, f \rangle_\nu \varphi_a$.

3 Iterated "integrals"

The iterated integral operators $\{I_m\}_{m \ge 0}$ are a family of linear operators, $I_m : H_\Sigma^{\otimes m} \to L^2(\mathcal{G})$. The map I_0 is the identity, where we regard the scalars $\mathbb{R}$ as being embedded naturally in $L^2(\mathcal{G})$ as the constant random variables. To define the map I_m for $m \ge 1$, we must specify its action on the basis vectors. For this, let us define the map $\tau : \{1, \ldots, d\}^m \to \Lambda = (\mathbb{N} \cup \{0\})^d$ as follows. If $J = (j_1, \ldots, j_m)$, then $\tau(J)_i = \#\{k : j_k = i\}$. Note that $|\tau(J)| = \sum_i \tau(J)_i = m$. Moreover, τ maps $\{1, \ldots, d\}^m$ onto $\{a : |a| = m\}$, and for any fixed a with $|a| = m$, we have $\#\{J : \tau(J) = a\} = m!/a!$.

We now define

$$I_m(v_{j_1} \otimes \cdots \otimes v_{j_m}) = \prod_{i=1}^{d} h_{a_i}(X(v_i)) = \sqrt{a!}\, \Phi_a,$$

where $a = \tau(j_1, \ldots, j_m)$. For example, if $d = 4$, then

$$I_6(v_1 \otimes v_4 \otimes v_1 \otimes v_1 \otimes v_2 \otimes v_4) = h_3(X(v_1)) h_1(X(v_2)) h_0(X(v_3)) h_2(X(v_4)).$$

It follows immediately from this definition that $I_m(u) = I_m(\widetilde{u})$ for all $u \in H_\Sigma^{\otimes m}$.

Since $\{\Phi_a : a \in \Lambda\}$ is an orthonormal basis for $L^2(\mathcal{G})$, it follows that $E[I_p(u) I_q(v)] = 0$ whenever $p \ne q$. Suppose $u, v \in H_\Sigma^{\otimes m}$. Let us write

$$\widetilde{u} = \sum_{|a|=m} \sum_{J \in \tau^{-1}(a)} \widetilde{u}(j_1, \ldots, j_m) v_{j_1} \otimes \cdots \otimes v_{j_d}.$$

Since $\widetilde{u}$ is symmetric, $\widetilde{u}(j_1, \ldots, j_m)$ depends only on a. Hence, we have

$$\widetilde{u} = \sum_{|a|=m} \sum_{J \in \tau^{-1}(a)} \widetilde{u}(a) v_{j_1} \otimes \cdots \otimes v_{j_d},$$

which gives

$$\begin{aligned} I_m(u) = I_m(\widetilde{u}) &= \sum_{|a|=m} \sum_{J \in \tau^{-1}(a)} \widetilde{u}(a) I_m(v_{j_1} \otimes \cdots \otimes v_{j_d}) \\ &= \sum_{|a|=m} \sum_{J \in \tau^{-1}(a)} \widetilde{u}(a) \sqrt{a!}\, \Phi_a \\ &= \sum_{|a|=m} \frac{m!}{\sqrt{a!}} \widetilde{u}(a) \Phi_a. \end{aligned} \tag{3.1}$$

Since similar formulas holds for $\widetilde{v}$ and $I_m(\widetilde{v})$, we have

$$E[I_m(u)I_m(v)] = \langle I_m(\widetilde{u}), I_m(\widetilde{v})\rangle_{\mathcal{G}} = \sum_{|a|=m} \frac{(m!)^2}{a!}\widetilde{u}(a)\widetilde{v}(a).$$

On the other hand,

$$\langle \widetilde{u}, \widetilde{v}\rangle_{H_\Sigma^{\otimes m}} = \sum_{|a|=m} \sum_{J\in\tau^{-1}(a)} \widetilde{u}(a)\widetilde{v}(a) = \sum_{|a|=m} \frac{m!}{a!}\widetilde{u}(a)\widetilde{v}(a).$$

We have thus shown that

$$E[I_p(u)I_q(v)] = \begin{cases} 0 & \text{if } p \neq q, \\ m!\langle \widetilde{u}, \widetilde{v}\rangle_{H_\Sigma^{\otimes m}} & \text{if } p = q = m. \end{cases} \tag{3.2}$$

In particular, this shows that $EI_m(u) = 0$ for $m \geq 1$ and $\|I_m(u)\|_{\mathcal{G}}^2 = m!\|\widetilde{u}\|_{H_\Sigma^{\otimes m}}^2$ for all m. Note that by the triangle inequality, $\|\widetilde{u}\|_{H_\Sigma^{\otimes m}} \leq \|u\|_{H_\Sigma^{\otimes m}}$.

Since $\{\Phi_a : |a| = m\}$ is an orthonormal basis for $\mathcal{H}_m$, we see from (3.1) that I_m maps $H_\Sigma^{\otimes m}$ onto $\mathcal{H}_m$. Moreover, for each $Y \in \mathcal{H}_m$, there is a unique symmetric $\widetilde{u} \in H_\Sigma^{\otimes m}$ such that $Y = I_m(\widetilde{u})$. Thus, any $F \in L^2(\mathcal{G})$ can be written as

$$F = EF + \sum_{m=1}^{\infty} I_m(u_m).$$

If the u_m's are taken to be symmetric, then they are unique.

Another important identity satisfied by the iterated integral operators is

$$I_{p+1}(u \otimes v) = I_p(u)I_1(v) - pI_{p-1}(u \otimes_1 v), \tag{3.3}$$

whenever $u \in H_\Sigma^{\otimes p}$ is symmetric, and $v \in H_\Sigma$. This is a special case of the multiplication formula,

$$I_p(u)I_q(v) = \sum_{r=0}^{p\wedge q} r!\binom{p}{r}\binom{q}{r} I_{p+q-2r}(u \otimes_r v),$$

which holds whenever $u \in H_\Sigma^{\otimes p}$ and $v \in H_\Sigma^{\otimes q}$ are symmetric. For a proof of the multiplication formula, see [1]. The multiplication formula can be used to show that

$$I_m(u^{\otimes m}) = h_m(X(u)), \tag{3.4}$$

whenever $u \in H_\Sigma$ is a unit vector. In particular, $I(u) = I_1(u) = X(u)$ for all $u \in H_\Sigma$.

An alternative notation for the iterated integral operators is

$$I_m(u) = \int_{\{1,\ldots,d\}^m} u(j_1, \ldots, j_m)\, dX(j_1) \cdots dX(j_m).$$

In this notation, if $u \in H_\Sigma$, then

$$\int u(j)\, dX(j) = I(u) = X(u) = u^T X.$$

According to (3.3), if $u, v \in H_\Sigma$, then

$$\iint u(i)v(j)\, dX(i)\, dX(j) = \int u(i)\, dX(i) \int v(j)\, dX(j) - \langle u, v\rangle_\Sigma.$$

The final term in the above expression is a stochastic correction term, which ensures that the iterated integral has mean zero.

4 Ordinary multivariable calculus

4.1 Gradient

Recall that $\{e_j\}$ denotes the standard basis vectors in $\mathbb{R}^d$. If $f : \mathbb{R}^d \to \mathbb{R}$ is differentiable, then we define $\nabla f : \mathbb{R}^d \to \mathbb{R}^d$ by

$$\nabla f = \sum_{j=1}^{d} \frac{\partial f}{\partial x_j} e_j.$$

Similarly, if f is m times differentiable, we define

$$D^m f = \sum_{j_1,\ldots,j_m=1}^{d} \frac{d^m f}{dx_{j_1} \cdots dx_{j_m}} e_{j_1} \otimes \cdots \otimes e_{j_m}.$$

Note that $D^2 f$ is the $d \times d$ matrix of second order partial derivatives; $D^3 f$ is not a matrix, but a $d \times d \times d$ array of third order partial derivatives. In general, $D^m f$ is an m-fold tensor, and we will regard it as an element of $H_\Sigma^{\otimes m}$.

The directional derivative operator in the direction of a vector $u \in \mathbb{R}^d$ is denoted by ∂_u and is defined by $\partial_u f = u^T \nabla f$. The Schwartz space of test functions is denoted by $\mathcal{S}(\mathbb{R}^d)$. It consists of smooth functions, all of whose derivatives decay faster than any polynomial. Formally, $\mathcal{S}(\mathbb{R}^d)$ consists of all functions $\varphi \in C^\infty(\mathbb{R}^d)$ such that

$$\sup_{x \in \mathbb{R}^d} |x^a D^m \varphi(x)| < \infty,$$

for all $a \in \Lambda$ and $m \in \mathbb{N}_0$. Using integration by parts, we have $\langle \partial_u \varphi, \psi\rangle_\lambda = -\langle \varphi, \partial_u \psi\rangle_\lambda$ for all $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$.

We will also need the notion of a weak derivative. Let $f \in L^2(\nu)$ and let $U = \{u_1, \ldots, u_d\} \subset \mathbb{R}^d$ be linearly independent. If there exists $\widetilde{f} \in L^2(\nu; H_\Sigma)$ such that $\langle u^T \widetilde{f}, \varphi\rangle_\lambda = -\langle f, \partial_u \varphi\rangle_\lambda$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and all $u \in U$, then we will say that f is weakly differentiable. In this case, we define $\nabla f = \widetilde{f}$ and $\partial_u f = u^T \widetilde{f}$. The space $W^{1,2}(\nu)$ consists of all $f \in L^2(\nu)$ that are weakly differentiable. Note that, by linearity, if $f \in W^{1,2}(\nu)$, then $\langle \partial_u f, \varphi\rangle_\lambda = -\langle f, \partial_u \varphi\rangle_\lambda$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and all $u \in \mathbb{R}^d$. The space $W^{1,2}(\nu)$ is a normed vector space with norm $\|f\|_{W^{1,2}(\nu)}^2 = \|f\|_\nu^2 + \|\nabla f\|_{L^2(\nu; H_\Sigma)}^2$.

When checking if $f \in W^{1,2}(\nu)$, it is enough to verify that, for each $j \in \{1, \ldots, d\}$, there exists $\widetilde{f}_j \in L^2(\nu)$ such that $\langle \widetilde{f}_j, \varphi\rangle_\lambda = -\langle f, \partial_{u_j} \varphi\rangle_\lambda$ for all $g \in \mathcal{S}(\mathbb{R}^d)$. Indeed, if this is the case, then we may define $\widetilde{f} = \sum_{j=1}^d \widetilde{f}_j A^T e_j$, where A is an operator such that $Au_j = e_j$ for all j. It is then a straightforward exercise to check that $\widetilde{f} = \nabla f$.

For vector fields $f \in L^2(\nu; H_\Sigma)$, we will say that $f \in W^{1,2}(\nu; H_\Sigma)$ if $u^T f \in W^{1,2}(\nu)$ for all $u \in U$. Again, by linearity, if $f \in W^{1,2}(\nu; H_\Sigma)$, then $u^T f \in W^{1,2}(\nu)$ for all $u \in \mathbb{R}^d$.

4.2 Divergence

If $g : \mathbb{R}^d \to \mathbb{R}^d$ is a differentiable vector field, then the (classical) divergence of g is denoted by $\nabla^T g$ and is defined by

$$\nabla^T g = \sum_{j=1}^d \frac{\partial(e_j^T g)}{\partial x_j} = \sum_{j=1}^d \partial_{e_j}(e_j^T g). \tag{4.1}$$

The classical divergence operator obeys the product rule,

$$\nabla^T(fg) = (\nabla f)^T g + f\nabla^T g, \tag{4.2}$$

whenever $f : \mathbb{R}^d \to \mathbb{R}$ is differentiable and $g : \mathbb{R}^d \to \mathbb{R}^d$ is a differentiable vector field.

Suppose $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and $\Psi : \mathbb{R}^d \to \mathbb{R}^d$ satisfies $e_j^T \Psi \in \mathcal{S}(\mathbb{R}^d)$ for all j. Then

$$\langle \varphi, \nabla^T \Psi\rangle_\lambda = \sum_{j=1}^d \left\langle \varphi, \frac{\partial(e_j^T \Psi)}{\partial x_j} \right\rangle_\lambda = -\sum_{j=1}^d \left\langle \frac{\partial \varphi}{\partial x_j}, e_j^T \Psi \right\rangle_\lambda = -\langle \nabla\varphi, \Psi\rangle_{L^2(\lambda;H_I)}, \tag{4.3}$$

where H_I is $\mathbb{R}^d$, equipped with the ordinary inner product. Note that (4.1) still makes sense, even when g is only weakly differentiable. In that case, (4.2) and (4.3) are still valid. The result in (4.3) shows that, at least formally, $-\nabla^T$ is the adjoint of the gradient operator ∇ in the space $L^2(\lambda; H_I)$.

In our case, the $L^2(\nu)$-divergence operator (which we will denote by δ_ν) will look somewhat different, since we are working on $L^2(\nu; H_\Sigma)$. Let us define $\operatorname{Dom}\delta_\nu$ as the set of all vector fields $g \in L^2(\nu; H_\Sigma)$ such that

$$|\langle \nabla f, g\rangle_{L^2(\nu;H_\Sigma)}| \le c_g \|f\|_\nu,$$

for all $f \in W^{1,2}(\nu)$, where c_g is a constant depending only on g. If $g \in \operatorname{Dom}\delta_\nu$, then $\delta_\nu(g)$ is the unique element of $L^2(\nu)$ such that

$$\langle f, \delta_\nu(g)\rangle_\nu = \langle \nabla f, g\rangle_{L^2(\nu;H_\Sigma)},$$

for all $f \in W^{1,2}(\nu)$. In other words, the divergence operator δ_ν is the adjoint of the gradient operator ∇ in the space $L^2(\nu; H_\Sigma)$. (The existence of the adjoint follows from standard results in functional analysis, since the gradient operator is closed, and its domain, $W^{1,2}(\nu)$, is dense in $L^2(\nu)$.)

As we will see in Section 6, $W^{1,2}(\nu; H_\Sigma) \subset \operatorname{Dom}\delta_\nu$. Let $g \in W^{1,2}(\nu; H_\Sigma)$ and let $f \in W^{1,2}(\nu)$ be arbitrary. Recall that $d\nu = p\, d\lambda$, where p is given by (2.1). Note that

$$p(x) = \frac{1}{(2\pi)^{d/2}\sqrt{\det\Sigma}} e^{-\|y\|^2/2},$$

where $x = M^T y$. Using integration by parts, it follows that

$$\langle \nabla f, g\rangle_{L^2(\nu;H_\Sigma)} = \langle \nabla f, p\Sigma g\rangle_{L^2(\lambda;H_I)} = -\langle f, \nabla^T(p\Sigma g)\rangle_\lambda = -\langle f, (\nabla p)^T \Sigma g + p\nabla^T(\Sigma g)\rangle_\lambda,$$

Note that

$$(\nabla p)^T e_j = \frac{\partial p}{\partial x_j} = -p \sum_{i=1}^d y_i \frac{\partial y_i}{\partial x_j} = -p \sum_{i=1}^d y_i e_i^T (M^T)^{-1} e_j = -p y^T (M^T)^{-1} e_j = -p x^T \Sigma^{-1} e_j,$$

so that $(\nabla p)^T = -p x^T \Sigma^{-1}$. Thus,

$$\langle \nabla f, g \rangle_{L^2(\nu; H_\Sigma)} = \langle f, (-\nabla^T (\Sigma g) + x^T g) p \rangle_{L^2(\lambda)} = \langle f, -\nabla^T (\Sigma g) + x^T g \rangle_{L^2(\nu)}.$$

This shows that

$$\delta_\nu(g) = -\nabla^T(\Sigma g) + x^T g, \tag{4.4}$$

for all $g \in W^{1,2}(\nu; H_\Sigma)$.

4.3 Second order differential operators

If $f : \mathbb{R}^d \to \mathbb{R}$ is twice differentiable, then the Laplacian of f, denoted by Δf, is defined as the divergence of the gradient of f. That is,

$$\Delta f = \nabla^T(\nabla f) = \sum_{i=1}^d \frac{\partial}{\partial x_i}(e_i^T \nabla f) = \sum_{i=1}^d \frac{\partial}{\partial x_i}\left(e_i^T \sum_{j=1}^d \frac{\partial f}{\partial x_j} e_j \right) = \sum_{i,j=1}^d (e_i^T e_j) \frac{\partial^2 f}{\partial x_i \partial x_j} = \sum_{j=1}^d \frac{\partial^2 f}{\partial x_j^2}.$$

In our case, we can define an analogous operator. Let

$$\operatorname{Dom} L_\nu = \{ f \in W^{1,2}(\nu) : \nabla f \in \operatorname{Dom} \delta_\nu \},$$

and for $f \in \operatorname{Dom} L_\nu$, define $L_\nu f = -\delta_\nu(\nabla f)$. Note that

$$\{ f \in W^{1,2}(\nu) : \nabla f \in W^{1,2}(\nu; H_\Sigma) \} \subset \operatorname{Dom} L_\nu.$$

Suppose $f \in W^{1,2}(\nu)$ and $\nabla f \in W^{1,2}(\nu; H_\Sigma)$. Then by (4.4),

$$\begin{aligned}
L_\nu f = -\delta(\nabla f) &= \nabla^T(\Sigma \nabla f) - x^T \nabla f \\
&= \sum_{i=1}^d \frac{\partial}{\partial x_i}(e_i^T \Sigma \nabla f) - \sum_{j=1}^d x_j \frac{\partial f}{\partial x_j} \\
&= \sum_{i=1}^d \frac{\partial}{\partial x_i}\left(e_i^T \Sigma \sum_{j=1}^d \frac{\partial f}{\partial x_j} e_j \right) - \sum_{j=1}^d x_j \frac{\partial f}{\partial x_j} \\
&= \sum_{i,j=1}^d (e_i^T \Sigma e_j) \frac{\partial^2 f}{\partial x_i \partial x_j} - \sum_{j=1}^d x_j \frac{\partial f}{\partial x_j}.
\end{aligned}$$

In the special case that the components of X are independent standard normals, so that $\Sigma = I$ and $d\nu(x) = (2\pi)^{-d/2} e^{-\|x\|^2/2}\, d\lambda(x)$, we have $L_\nu f = \Delta f - x^T \nabla f$, which is the generator of the classical Ornstein-Uhlenbeck process on $\mathbb{R}^d$. Recall that the Ornstein-Uhlenbeck process on $\mathbb{R}^d$ is the solution Z to the stochastic differential equation $dZ = -Z\, dt + \sqrt{2}\, dB$, where B is a standard, d-dimensional standard Brownian motion.

5 The Malliavin derivative

Let $\mathcal{S} = \{f(X) : f \in C_c^\infty(\mathbb{R}^d)\} \subset L^2(\mathcal{G})$. If $F = f(X) \in \mathcal{S}$, then we define $DF = \nabla f(X) \in L^2(\mathcal{G}; H_\Sigma)$. Similarly, we define $D^m F = D^m f(X) \in L^2(\mathcal{G}; H_\Sigma^{\otimes m})$. For $F \in \mathcal{S}$, let

$$\|F\|_{m,2}^2 = \sum_{j=0}^m \|D^j F\|_{L^2(\mathcal{G}; H_\Sigma^{\otimes j})}^2 = \sum_{j=0}^m E\|D^j F\|_{H_\Sigma^{\otimes j}}^2.$$

We then define $\mathbb{D}^{m,2}$ as the closure of $\mathcal{S}$ in $L^2(\mathcal{G})$ with respect to $\|\cdot\|_{m,2}$. The operator D^m is closable, and extends to $\mathbb{D}^{m,2}$.

Theorem 5.1. *Suppose $F \in L^2(\mathcal{G})$ and write*

$$F = \sum_{m=0}^\infty I_m(u_m) = EF + u_1^T X + \sum_{m=2}^\infty I_m(u_m), \tag{5.1}$$

where each $u_m \in H_\Sigma^{\otimes m}$ is symmetric. Then $F \in \mathbb{D}^{1,2}$ if and only if

$$\sum_{m=1}^\infty m m! \|u_m\|_{H_\Sigma^{\otimes m}}^2 < \infty, \tag{5.2}$$

in which case,

$$D_j F := \langle DF, v_j \rangle_\Sigma = \sum_{m=1}^\infty m I_{m-1}(u_m(\cdot, j)) = u_1(j) + \sum_{m=2}^\infty m I_{m-1}(u_m(\cdot, j)), \tag{5.3}$$

and the left-hand side of (5.2) *equals* $\|DF\|_{L^2(\mathcal{G}; H_\Sigma)}^2$.

For a proof of this theorem, see [1]. Equation (5.3) is easy to verify in one particular special case. Suppose $F = f(X)$, where $f(x) = h_m(u^T x)$ and $u \in H_\Sigma$ is a unit vector. Then by (3.4), we have $F = I_m(u^{\otimes m})$ and

$$\begin{aligned} D_j F &= \langle \nabla f(X), v_j \rangle_\Sigma = \langle h_m'(u^T X) u, v_j \rangle_\Sigma = m h_{m-1}(u^T X) u(j) \\ &= m I_{m-1}(u^{\otimes(m-1)}) u(j) = m I_{m-1}(u^{\otimes(m-1)} u(j)) = m I_{m-1}(u^{\otimes m}(\cdot, j)), \end{aligned}$$

which agrees with (5.3).

Note that, by (5.1), we have

$$E[FX(v_j)] = E[FX^T v_j] = E[u_1^T X X^T v_j] = u_1^T \Sigma v_j = \langle u_1, v_j \rangle_\Sigma = u_1(j) = E[D_j F].$$

By linearity, it follows that

$$E[FX(u)] = E[\langle DF, u \rangle_\Sigma], \tag{5.4}$$

for all $F \in \mathbb{D}^{1,2}$ and all $u \in H_\Sigma$. This is commonly called the (Malliavin) integration by parts formula. This formula can also be derived by writing the expectations as integrals over $\mathbb{R}^d$, and using ordinary integration by parts.

Equation (5.3) can be generalized (see Exercise 1.2.5 in [1]) to

$$\begin{aligned}
D^N_{j_1,\dots,j_N} F &:= \langle D^N F, v_{j_1} \otimes \cdots \otimes v_{j_N} \rangle_{H_\Sigma^{\otimes N}} \\
&= \sum_{m=N}^{\infty} \frac{m!}{(m-N)!} I_{m-N}(u_m(\cdot, j_1, \dots, j_N)) \\
&= N! u_N(j_1, \dots, j_N) + \sum_{m=N+1}^{\infty} \frac{m!}{(m-N)!} I_{m-N}(u_m(\cdot, j_1, \dots, j_N)).
\end{aligned}$$

Consequently, if $u \in H_\Sigma^{\otimes N}$, then $E[\langle D^N F, u\rangle_{H_\Sigma^{\otimes N}}] = N! \langle u_N, u\rangle_{H_\Sigma^{\otimes N}}$. On the other hand, by (5.1), we have $E[F I_N(u)] = E[I_N(u_N) I_N(u)]$. Applying (3.2), this show that

$$E[F I_m(u)] = E[\langle D^m F, u\rangle_{H_\Sigma^{\otimes m}}],$$

for all $F \in \mathbb{D}^{m,2}$ and all symmetric $u \in H_\Sigma^{\otimes m}$, which is a generalization of (5.4) to higher derivatives.

It may be difficult to see from Theorem 5.1, but the Malliavin derivative is really just the weak derivative from Section 4.

Theorem 5.2. *Recall the isomorphism $\mathcal{I} : L^2(\nu) \to L^2(\mathcal{G})$ given by $\mathcal{I} f = f(X)$. By a slight abuse of notation, we may consider $\mathcal{I} : L^2(\nu; H_\Sigma) \to L^2(\mathcal{G}; H_\Sigma)$, also given by $\mathcal{I} f = f(X)$. In this case, $\mathcal{I} W^{1,2}(\nu) = \mathbb{D}^{1,2}$, $\|\mathcal{I} f\|_{1,2} = \|f\|_{W^{1,2}(\nu)}$, and $\mathcal{I} \nabla f = D \mathcal{I} f$ for all $f \in W^{1,2}(\nu)$.*

Proof. Let $F = f(X) \in \mathbb{D}^{1,2}$ and write $F = \sum_{m=0}^\infty I_m(u_m)$, where each $u_m \in H_\Sigma^{\otimes m}$ is symmetric. Recall that $I_m(u_m) \in \mathcal{H}_m$, the Wiener chaos of order m, and that $\mathcal{H}_m$ is finite-dimensional (at least in the context of these notes) and spanned by $\{H_m(X(v)) : v \in \mathbb{R}^d, \|v\|_\Sigma = 1\}$. Hence, we may write

$$F = \sum_{m=0}^{\infty} \sum_{k=1}^{N(m)} c_{m,k} H_m(X(v_{m,k})) = \sum_{m=0}^{\infty} \sum_{k=1}^{N(m)} \frac{c_{m,k}}{m!} I_m(v_{m,k}^{\otimes m}),$$

where each $v_{m,k} \in H_\Sigma$ is a unit vector. From this, we draw two conclusions. First,

$$f(x) = (\mathcal{I}^{-1} F)(x) = \sum_{m=0}^{\infty} \sum_{k=1}^{N(m)} c_{m,k} H_m(v_{m,k}^T x). \tag{5.5}$$

Second,

$$u_m = \sum_{k=1}^{N(m)} \frac{c_{m,k}}{m!} v_{m,k}^{\otimes m}.$$

Using (5.3) and (2.2), it follows that

$$\begin{aligned}
D_j F = \sum_{m=1}^{\infty} m I_{m-1}(u_m(\cdot, j)) &= \sum_{m=1}^{\infty} \sum_{k=1}^{N(m)} \frac{c_{m,k}}{(m-1)!} I_{m-1}(v_{m,k}^{\otimes(m-1)}) v_{m,k}(j) \\
&= \sum_{m=1}^{\infty} \sum_{k=1}^{N(m)} c_{m,k} H_{m-1}(v_{m,k}^T X) v_{m,k}(j) = \sum_{m=1}^{\infty} \sum_{k=1}^{N(m)} c_{m,k} H_m'(v_{m,k}^T X) \langle v_{m,k}, v_j \rangle_\Sigma.
\end{aligned}$$

In other words, $D_jF = \widetilde{f}_j(X)$, where

$$\widetilde{f}_j = \sum_{m=1}^{\infty} \sum_{k=1}^{N(m)} c_{m,k} \partial_{\Sigma v_j}(H_m(v_{m,k}^T \cdot)). \tag{5.6}$$

Comparing this with (5.5), it now follows, using integration by parts, that if $g \in \mathcal{S}(\mathbb{R}^d)$, then $\langle \widetilde{f}_j, g\rangle_\lambda = -\langle f, \partial_{\Sigma v_j} g\rangle_\lambda$. As in Section 4.1, this shows that $f \in W^{1,2}(\nu)$. We may use the method described at the end of Section 4.1 to check that $\nabla f = \sum_{j=1}^d \widetilde{f}_j v_j$, which implies that $DF = \nabla f(X)$. Moreover,

$$\|F\|_{1,2}^2 = \|F\|_{\mathcal{G}}^2 + \|DF\|_{L^2(\mathcal{G};H_\Sigma)}^2 = \|f\|_\nu^2 + \|\nabla f\|_{L^2(\nu;H_\Sigma)}^2 = \|f\|_{W^{1,2}(\nu)}^2.$$

In summary, we have shown that $\mathbb{D}^{1,2} \subset \mathcal{I}W^{1,2}(\nu)$, $\|\mathcal{I}f\|_{1,2} = \|f\|_{W^{1,2}(\nu)}$, and $\mathcal{I}\nabla f = D\mathcal{I}f$ for all $f \in \mathcal{I}^{-1}\mathbb{D}^{1,2}$.

It remains only to show that $\mathcal{I}W^{1,2}(\nu) \subset \mathbb{D}^{1,2}$. Suppose $f \in W^{1,2}(\nu)$ and let $F = f(X)$. As before, we may write f in the form (5.5). We may then deduce that $\widetilde{f}_j = \partial_{\Sigma v_j} f$ must have the form (5.6), which, as above, implies

$$\widetilde{f}_j(X) = \sum_{m=1}^{\infty} m I_{m-1}(u_m(\cdot, j)).$$

From here, we find that

$$\begin{aligned}\infty > \|\widetilde{f}\|_{L^2(\nu;H_\Sigma)}^2 = \sum_{j=1}^d \|\widetilde{f}_j\|_\nu^2 = \sum_{j=1}^d E|\widetilde{f}_j(X)|^2 = \sum_{j=1}^d \sum_{m=1}^{\infty} m\|I_{m-1}(u_m(\cdot, j))\|_{\mathcal{G}}^2 \\ = \sum_{m=1}^{\infty} mm! \sum_{j=1}^d \|u_m(\cdot, j)\|_{H_\Sigma^{\otimes(m-1)}}^2 = \sum_{m=1}^{\infty} mm! \|u_m\|_{H_\Sigma^{\otimes m}}^2,\end{aligned}$$

which implies $F \in \mathbb{D}^{1,2}$. □

For random vectors $u \in L^2(\mathcal{G}; H_\Sigma)$, we will say that $u \in \mathbb{D}^{1,2}(H_\Sigma)$ if $\langle u, v_j\rangle_\Sigma \in \mathbb{D}^{1,2}$ for all j. Note that this is equivalent to requiring that $\langle u, v\rangle_\Sigma \in \mathbb{D}^{1,2}$ for all $v \in H_\Sigma$. Also note that $\mathcal{I}W^{1,2}(\nu; H_\Sigma) = \mathbb{D}^{1,2}(H_\Sigma)$.

6 The Malliavin divergence operator

The domain of the divergence operator, denoted by $\operatorname{Dom}\delta$, is the subset of $L^2(\mathcal{G}; H_\Sigma)$ consisting of elements u such that

$$|\langle DF, u\rangle_{L^2(\mathcal{G};H_\Sigma)}| = |E\langle DF, u\rangle_\Sigma| \le c_u \|F\|_{\mathcal{G}},$$

for all $F \in \mathbb{D}^{1,2}$, where c_u is a constant depending only on u. If $u \in \operatorname{Dom}\delta$, then $\delta(u)$ is the unique element of $L^2(\mathcal{G})$ satisfying

$$\langle F, \delta(u)\rangle_{\mathcal{G}} = \langle DF, u\rangle_{L^2(\mathcal{G};H_\Sigma)},$$

for all $F \in \mathbb{D}^{1,2}$. In other words, the divergence operator δ is the adjoint of the derivative operator D.

Remark 6.1. The Malliavin divergence operator, in the context of these notes, is simply the $L^2(\nu)$-divergence operator from Section 4. In this sense, it is a differential operator. This fact flies in the face of convention, however, since the Malliavin divergence operator is sometimes called the "Skorohod integral."

One reason why we might think of δ as a (stochastic) integral operator is the following. Suppose $u \in L^2(\mathcal{G}; H_\Sigma)$ is constant almost surely, so that we may regard u as simply being an element of H_Σ. Then by (5.4), we have that $\delta(u) = X(u)$. In other words, δ extends the integral operator $I = I_1$ from deterministic integrands to stochastic integrands.

Another very convincing reason to think of δ as an integral operator occurs in the context of Example 1.3, where $H = L^2([0,1])$. In this case, elements $u \in L^2(\mathcal{G}; H)$ are stochastic processes $u(t)$ satisfying $E\int_0^1 |u(t)|^2\,dt < \infty$. If $u \in \operatorname{Dom}\delta$ and u is adapted to the filtration of the Brownian motion, then it can be shown that $\delta(u) = \int_0^1 u(t)\,dB(t)$, where this is the Itô integral. However, when $u \in \operatorname{Dom}\delta$, but is not adapted, then $\delta(u)$ is still well defined. In this sense, in the context of Example (1.3), $\delta(u)$ is an extension of the Itô integral to non-adapted integrands. It is called the Skorohod integral.

Returning to the finite-dimensional setting, note that if $u \in L^2(\mathcal{G}; H_\Sigma)$, then $u(j) = \langle u, v_j\rangle_\Sigma \in L^2(\mathcal{G})$. Hence, we may uniquely write $u(j) = \sum_{m=0}^\infty I_m(u_{m,j})$, where each $u_{m,j} \in H_\Sigma^{\otimes m}$ is symmetric. Note that $u_m := \sum_{k=1}^d u_{m,k} \otimes v_k \in H_\Sigma^{\otimes(m+1)}$ is symmetric in the first m variables. Also, $u_m(\cdot, j) = \sum_{k=1}^d u_{m,k} v_k(j) = u_{m,j}$. Thus,

$$u(j) = \sum_{m=0}^\infty I_m(u_m(\cdot, j)). \tag{6.1}$$

Theorem 6.2. *Let $u \in L^2(\mathcal{G}; H_\Sigma)$ have expansion* (6.1)*. Then $u \in \operatorname{Dom}\delta$ if and only if*

$$\sum_{m=0}^\infty (m+1)!\|\widetilde{u}_m\|^2_{H_\Sigma^{\otimes(m+1)}} < \infty, \tag{6.2}$$

in which case

$$\delta(u) = \sum_{m=0}^\infty I_{m+1}(\widetilde{u}_m),$$

and the left-hand side of (6.2) *is equal to $\|\delta(u)\|^2_{\mathcal{G}}$.*

Theorem 6.3. (Product rule for Malliavin divergence) *Let $F \in \mathbb{D}^{1,2}$ and $u \in \operatorname{Dom}\delta$. If $Fu \in L^2(\mathcal{G}; H_\Sigma)$, $F\delta(u) \in L^2(\mathcal{G})$, and $\langle DF, u\rangle_\Sigma \in L^2(\mathcal{G})$, then $Fu \in \operatorname{Dom}\delta$ and*

$$\delta(Fu) = F\delta(u) - \langle DF, u\rangle_\Sigma.$$

For proofs of these theorems, see [1]. With Theorem 6.2, it is not difficult to verify that $\mathbb{D}^{1,2}(H_\Sigma) \subset \operatorname{Dom}\delta$. Let $u \in \mathbb{D}^{1,2}(H_\Sigma)$. As above, let us write $u(j) = \sum_{m=0}^\infty I_m(u_m(\cdot, j))$. By (5.2), we have

$$\sum_{m=1}^\infty m m! \|u_{m,j}\|^2_{H_\Sigma^{\otimes m}} < \infty,$$

for each j. It follows that

$$\begin{aligned}
\sum_{m=0}^{\infty}(m+1)!\|\widetilde{u}_m\|^2_{H_\Sigma^{\otimes(m+1)}} &\le \sum_{m=0}^{\infty}(m+1)!\|u_m\|^2_{H_\Sigma^{\otimes(m+1)}}\\
&= \sum_{m=0}^{\infty}(m+1)!\sum_{j=1}^{d}\|u_{m,j}\|^2_{H_\Sigma^{\otimes m}}\\
&\le 2\sum_{j=1}^{d}\sum_{m=0}^{\infty} mm!\|u_{m,j}\|^2_{H_\Sigma^{\otimes m}} < \infty.
\end{aligned}$$

By (6.2), this shows that $\mathbb{D}^{1,2}(H_\Sigma) \subset \operatorname{Dom}\delta$.

Now recall the isomorphisms $\mathcal{I} : L^2(\nu) \to L^2(\mathcal{G})$ and $\mathcal{I} : L^2(\nu; H_\Sigma) \to L^2(\mathcal{G}; H_\Sigma)$ given by $\mathcal{I}f = f(X)$. It is a direct consequence of the definitions that $\mathcal{I}\operatorname{Dom}\delta_\nu = \operatorname{Dom}\delta$ and $\mathcal{I}\delta_\nu f = \delta\mathcal{I}f$ for every $f \in \operatorname{Dom}\delta_\nu$. In other words, the Malliavin divergence is just the $L^2(\nu)$-divergence.

Since $\mathcal{I}W^{1,2}(\nu; H_\Sigma) = \mathbb{D}^{1,2}(H_\Sigma)$, we may now conclude that $W^{1,2}(\nu; H_\Sigma) \subset \operatorname{Dom}\delta_\nu$. Also, by Theorem 6.3, the $L^2(\nu)$-divergence operator satisfies the product rule

$$\delta_\nu(fg) = f\delta_\nu(g) - (\nabla f)^T \Sigma g,$$

whenever $f \in W^{1,2}(\nu)$, $g \in \operatorname{Dom}\delta_\nu$, $fg \in L^2(\nu; H_\Sigma)$, $f\delta_\nu(g) \in L^2(\nu)$, and $(\nabla f)^T\Sigma g \in L^2(\nu)$. Note the similarity between this and (4.2).

7 The Ornstein-Uhlenbeck semigroup

Fix $t \ge 0$. The Ornstein-Uhlenbeck semigroup, $T_t : L^2(\mathcal{G}) \to L^2(\mathcal{G})$, is defined by

$$T_t(F) = \sum_{m=0}^{\infty} e^{-mt} I_m(u_m),$$

where $F = \sum_{m=0}^{\infty} I_m(u_m)$. (Recall that this decomposition is unique when $u_m \in H_\Sigma^{\otimes m}$ is symmetric.) Let

$$\operatorname{Dom} L = \left\{ F = \sum_{m=0}^{\infty} I_m(u_m) \in L^2(\mathcal{G}) : \sum_{m=0}^{\infty} m^2 \|I_m(u_m)\|^2_{\mathcal{G}} < \infty \right\}.$$

For $F \in \operatorname{Dom} L$, we define

$$LF = -\sum_{m=0}^{\infty} m I_m(u_m) \in L^2(\mathcal{G}).$$

One can check that the operator L satisfies

$$LF = \lim_{t\to 0} \frac{T_t(F) - F}{t}.$$

That is, L is the generator of the Ornstein-Uhlenbeck semigroup.

Theorem 7.1. *Let $F \in L^2(\mathcal{G})$. Then $F \in \operatorname{Dom} L$ if and only if $F \in \mathbb{D}^{1,2}$ and $DF \in \operatorname{Dom} \delta$, in which case $LF = -\delta(DF)$.*

For a proof of this theorem, see [1]. This theorem shows that, under the isomorphism $\mathcal{I} : L^2(\nu) \to L^2(\mathcal{G})$ given by $\mathcal{I}f = f(X)$, we have $\mathcal{I} \operatorname{Dom} L_\nu = \operatorname{Dom} L$ and $\mathcal{I}L_\nu f = L\mathcal{I}f$. In other words, we may identify the operator L on $L^2(\mathcal{G})$ with the operator L_ν on $L^2(\nu)$.

We remarked at the end of Section 4 that L_ν is the generator of the classical Ornstein-Uhlenbeck process on $\mathbb{R}^d$ when X is a vector of independent standard normals. More generally, however, L_ν is the generator of an $\mathbb{R}^d$-valued stochastic process Z satisfying the stochastic differential equation,

$$dZ = -Z\,dt + \sqrt{2}M^T\,dB,$$

where B is a standard, d-dimensional Brownian motion. Now suppose $u \in H_\Sigma$ is a unit vector. Then the process $Z(u) = u^T Z$ is a one-dimensional diffusion satisfying

$$dZ(u) = -Z(u)\,dt + \sqrt{2}\,d\widetilde{B},$$

where $\widetilde{B} = (Mu)^T B$. Since the squared Euclidean norm of Mu is $u^T M^T M u = \langle u, u \rangle_\Sigma = 1$, it follows that $\widetilde{B}$ is a standard, one-dimensional Brownian motion. Hence, $Z(u)$ is a classical, one-dimensional Ornstein-Uhlenbeck process, for every unit vector $u \in H_\Sigma$.

References

[1] D. Nualart. *The Malliavin calculus and related topics.* Probability and its Applications (New York). Springer-Verlag, New York, 1995.