16

Elements of Malliavin calculus

This chapter offers a brief introduction to Malliavin calculus and its applications to mathematical finance, in particular the computation of the Greeks by the Monte Carlo method. As we have seen in Section 12.4.2, the simplest way to compute sensitivities by the Monte Carlo method consists in approximating the derivatives by incremental ratios obtained by simulating the payoffs corresponding to close values of the underlying asset. If the payoff function is *not regular* (for example, in the case of a digital option with strike K and payoff function $\mathbb{1}_{[K,+\infty[}$) this technique is not efficient since the incremental ratio has typically a very large variance. In Section 12.4.2 we have seen that the problem can be solved by integrating by parts and differentiating the density function of the underlying asset, provided it is sufficiently regular: if the underlying asset follows a geometric Brownian motion, this is possible since the explicit expression of the density is known.

In a more general setting, the Malliavin calculus allows obtaining explicit integration-by-parts formulas even if the density of the underlying asset is not known and so it provides an effective tool to approximate the Greeks numerically (see, for example, the experiments in [137] where different methods of approximating the Greeks are compared).

The applications of Malliavin calculus to mathematical finance are relatively recent: Malliavin's results [244] initially attracted great interest in view of the proof and extension of Hörmander's hypoellipticity theorem [170] (cf. Section 9.5.2). From a theoretical point of view, a remarkable financial application is the Clark-Ocone formula [270], proved in Paragraph 16.2.1, that improves the martingale representation theorem and allows expressing the hedging strategy of an option in terms of the stochastic derivative of its price.

We also recall that Malliavin calculus was recently used to approximate numerically the price of American options by the Monte Carlo method: see, for instance, Fournié, Lasry, Lebuchoux, Lions and Touzi [138], Fournié, Lasry, Lebuchoux and Lions [137], Kohatsu-Higa and Pettersson [212], Bouchard, Ekeland and Touzi [53], Bally, Caramellino and Zanette [20].

Pascucci A.: PDE and Martingale Methods in Option Pricing

In this chapter we give some basic ideas of Malliavin calculus by analyzing several examples of the applications to the computation of the Greeks. We confine ourselves to the one-dimensional case, choosing simplicity instead of generality; furthermore, some proofs will only be sketched for the sake of brevity. For an organic presentation of the theory, we refer to the monographs by Nualart [267], Shigekawa [308], Sanz-Solé [296], Bell [37], Da Prato [83], Di Nunno, Oksendal and Proske [96]. We mention also some more coincise presentations, mainly application-oriented, that are available on the web: Kohatsu-Higa and Montero [211], Friz [144], Bally [19], Oksendal [272] and Zhang [343].

16.1 Stochastic derivative

In this paragraph we introduce the concept of stochastic (or Malliavin) derivative: the idea is to define the notion of differentiability within the family of random variables that are equal to (or can be approximated by) functions of independent increments of Brownian motion. Under suitable assumptions, we see that this family is wide enough to contain the solution of stochastic differential equations.

Unfortunately the notations that are necessary to introduce Malliavin calculus are a bit burdensome: at the beginning courage must not be lost and a little patience is needed to get acquainted with the notation. On first reading we advise the reader not to dwell too much on the details.

Let us consider a real Brownian motion W on the probability space $(\Omega, \mathcal{F}, P)$, endowed with the Brownian filtration $\mathcal{F}^W = \left(\mathcal{F}_t^W\right)_{t\in[0,T]}$. For the sake of simplicity, since this is not really restrictive, we suppose that $T = 1$ and, for $n \in \mathbb{N}$, let

$$t_n^k := \frac{k}{2^n}, \qquad k = 0, \dots, 2^n$$

be the $(k+1)$-th element of the n-th order dyadic partition of the interval $[0, T]$. Let

$$I_n^k :=]t_n^{k-1}, t_n^k], \qquad \Delta_n^k := W_{t_n^k} - W_{t_n^{k-1}},$$

be the k-th interval of the partition and the k-th increment of the Brownian motion, for $k = 1, \dots, 2^n$, respectively. Furthermore, we denote by

$$\Delta_n := \left(\Delta_n^1, \dots, \Delta_n^{2^n}\right)$$

the $\mathbb{R}^{2^n}$-vector of the n-th order Brownian increments and by C^∞_{pol} the family of smooth functions that, together with their derivatives of any order, have at most polynomial growth.

Definition 16.1 *Given $n \in \mathbb{N}$, the family of simple n-th order functionals is defined by*

$$\mathcal{S}_n := \{\varphi(\Delta_n) \mid \varphi \in C^\infty_{\text{pol}}(\mathbb{R}^{2^n}; \mathbb{R})\}.$$

We denote by

$$x_n = (x_n^1, \dots, x_n^{2^n}) \tag{16.1}$$

the point in $\mathbb{R}^{2^n}$. It is apparent that $W_T = \varphi(\Delta_n) \in \mathcal{S}_n$ for every $n \in \mathbb{N}$ with $\varphi(x_n^1, \dots, x_n^{2^n}) = x_n^1 + \dots + x_n^{2^n}$.

We also remark that

$$\mathcal{S}_n \subseteq \mathcal{S}_{n+1}, \qquad n \in \mathbb{N},$$

and we define

$$\mathcal{S} := \bigcup_{n \in \mathbb{N}} \mathcal{S}_n,$$

the family of simple functionals. By the growth assumption on φ, $\mathcal{S}$ is a subspace of $L^p(\Omega, \mathcal{F}_T^W)$ for every $p \geq 1$. Further, $\mathcal{S}$ is dense[1] in $L^p(\Omega, \mathcal{F}_T^W)$. We introduce now a very handy notation, that will be often used:

Notation 16.2 *For every $t \in]0, T]$, let $k_n(t)$ be the only element $k \in \{1, \dots, 2^n\}$ such that $t \in I_n^k$.*

Definition 16.3 *For every $X = \varphi(\Delta_n) \in \mathcal{S}$, the stochastic derivative of X at time t is defined by*

$$D_t X := \frac{\partial \varphi}{\partial x_n^{k_n(t)}}(\Delta_n).$$

Remark 16.4 Definition 16.3 is well-posed i.e. *it is independent of* n: indeed it is not difficult to see that, if we have for $n, m \in \mathbb{N}$

$$X = \varphi_n(\Delta_n) = \varphi_m(\Delta_m) \in \mathcal{S},$$

with $\varphi_n, \varphi_m \in C_{\text{pol}}^\infty$, then, for every $t \leq T$, we have

$$\frac{\partial \varphi_n}{\partial x_n^{k_n(t)}}(\Delta_n) = \frac{\partial \varphi_m}{\partial x_m^{k_m(t)}}(\Delta_m).$$

□

Now we endow $\mathcal{S}$ with the norm

$$\begin{aligned} \|X\|_{1,2} :=& E\left[X^2\right]^{\frac{1}{2}} + E\left[\int_0^T (D_s X)^2 ds\right]^{\frac{1}{2}} \\ =& \|X\|_{L^2(\Omega)} + \|DX\|_{L^2([0,T]\times\Omega)}. \end{aligned}$$

Definition 16.5 *The space $\mathbb{D}^{1,2}$ of the Malliavin-differentiable random variables is the closure of $\mathcal{S}$ with respect to the norm $\|\cdot\|_{1,2}$.*

In other terms, $X \in \mathbb{D}^{1,2}$ if and only if there exists a sequence (X_n) in $\mathcal{S}$ such that

[1] Since we are considering the Brownian filtration!

i) $X = \lim_{n\to\infty} X_n$ in $L^2(\Omega)$;
ii) the limit $\lim_{n\to\infty} DX_n$ exists in $L^2([0,T]\times\Omega)$.

In this case it seems natural to define the Malliavin derivative of X as

$$DX := \lim_{n\to\infty} DX_n, \qquad L^2([0,T]\times\Omega).$$

This definition is *well-posed* in view of the following:

Lemma 16.6 *Let (X_n) be a sequence in $\mathcal{S}$ such that*

i) $\lim_{n\to\infty} X_n = 0$ *in* $L^2(\Omega)$;
ii) there exists $U := \lim_{n\to\infty} DX_n$ *in* $L^2([0,T]\times\Omega)$.

Then $U = 0$ *a.e.*[2]

Remark 16.7 The proof of Lemma 16.6 is not obvious since the differentiation operator D is linear but *not bounded*, i.e.

$$\sup_{X\in\mathcal{S}} \frac{\|DX\|_{L^2}}{\|X\|_{L^2}} = +\infty.$$

Indeed it is quite simple to find an example of a sequence (X_n) bounded in $L^2(\Omega)$ and such that (DX_n) is not bounded in $L^2([0,T]\times\Omega)$: for fixed $\bar{n}\in\mathbb{N}$, it suffices to consider $X_n = \varphi_n(\Delta_{\bar{n}})$ with (φ_n) converging in $L^2(\mathbb{R}^{2^{\bar{n}}})$ to a suitable non-regular function. □

We defer the proof of Lemma 16.6 to Paragraph 16.2 and now we analyze some fundamental examples.

16.1.1 Examples

Example 16.8 For fixed t, let us prove that $W_t \in \mathbb{D}^{1,2}$ and[3]

$$D_s W_t = \mathbb{1}_{[0,t]}(s). \tag{16.2}$$

Indeed, recalling Notation 16.2, we consider the sequence

$$X_n = \sum_{k=1}^{k_n(t)} \Delta_n^k, \qquad n\in\mathbb{N}.$$

We have $X_n = W_{t_n^{k_n(t)}} \in \mathcal{S}_n$ and so

$$D_s X_n = \begin{cases} 1 & \text{if } s \le t_n^{k_n(t)}, \\ 0 & \text{if } s > t_n^{k_n(t)}, \end{cases}$$

i.e. $D_s X_n = \mathbb{1}_{[0,t_n^{k_n(t)}]}$. Then (16.2) follows from the fact that

[2] In $\mathscr{B}\otimes\mathcal{F}_T^W$.
[3] The stochastic derivative is defined as an L^2-limit, up to sets with null Lebesgue measure: thus, $D_s W_t$ is also equal to $\mathbb{1}_{]0,t[}(s)$ or to $\mathbb{1}_{[0,t]}(s)$.

i) $\lim_{n\to\infty} W_{t_n^{k_n(t)}} = W_t$ in $L^2(\Omega)$;
ii) $\lim_{n\to\infty} \mathbb{1}_{[0,t_n^{k_n(t)}]} = \mathbb{1}_{(0,t)}$ in $L^2([0,T]\times\Omega)$.

$\square$

Remark 16.9 If $X \in \mathbb{D}^{1,2}$ is $\mathcal{F}_t^W$-measurable, then

$$D_s X = 0, \qquad s > t.$$

Indeed, up to approximation, it suffices to consider the case $X = \varphi(\Delta_n) \in \mathcal{S}_n$ for some n: if X is $\mathcal{F}_t^W$-measurable, then it is independent[4] from Δ_n^k for $k > k_n(t)$. Therefore, for fixed $s > t$,

$$\frac{\partial \varphi}{\partial x_n^{k_n(s)}}(\Delta_n) = 0,$$

at least if n is large enough, in such a way that t and s belong to disjoint intervals of the n-th order dyadic partition. $\square$

Example 16.10 Let $u \in L^2(0,T)$ be a (deterministic) function and

$$X = \int_0^t u(r) dW_r.$$

Then $X \in \mathbb{D}^{1,2}$ and

$$D_s X = \begin{cases} u(s) & \text{for } s \le t, \\ 0 & \text{for } s > t. \end{cases}$$

Indeed the sequence defined by

$$X_n = \sum_{k=1}^{k_n(t)} u(t_n^{k-1}) \Delta_n^k$$

is such that

$$D_s X_n = \varphi(t_n^{k_n(s)})$$

if $s \le t_n^{k_n(t)}$ and $D_s X_n = 0$ for $s > t_n^{k_n(t)}$. Further, X_n and $D_s X_n$ approximate X and $u(s)\mathbb{1}_{[0,t]}(s)$ in $L^2(\Omega)$ and $L^2([0,T]\times\Omega)$ respectively. $\square$

---

[4] Recalling Remark A.43, since $t \in]t_n^{k_n(t)-1}, t_n^{k_n(t)}]$ we have:

i) if $t < t_n^{k_n(t)}$, then X is a function of $\Delta_n^1, \dots, \Delta_n^{k_n(t)-1}$ only;
ii) if $t = t_n^{k_n(t)}$, then X is a function of $\Delta_n^1, \dots, \Delta_n^{k_n(t)}$ only.

16.1.2 Chain rule

If $X, Y \in \mathbb{D}^{1,2}$, then the product XY in general is not square integrable and so it does not belong to $\mathbb{D}^{1,2}$. For this reason, sometimes it is worthwhile to use, instead of $\mathbb{D}^{1,2}$ the slightly smaller space (but closed under products):

$$\mathbb{D}^{1,\infty} = \bigcap_{p\geq 2} \mathbb{D}^{1,p}$$

where $\mathbb{D}^{1,p}$ is the closure of $\mathcal{S}$ with respect to the norm

$$\|X\|_{1,p} = \|X\|_{L^p(\Omega)} + \|DX\|_{L^p([0,T]\times\Omega)}.$$

We observe that $X \in \mathbb{D}^{1,p}$ if and only if there exists a sequence (X_n) in $\mathcal{S}$ such that

i) $X = \lim_{n\to\infty} X_n$ in $L^p(\Omega)$;
ii) the limit $\lim_{n\to\infty} DX_n$ exists in $L^p([0,T]\times\Omega)$.

If $p \leq q$, by Hölder's inequality we get

$$\|\cdot\|_{L^p([0,T]\times\Omega)} \leq T^{\frac{q-p}{pq}} \|\cdot\|_{L^p([0,T]\times\Omega)},$$

and so

$$\mathbb{D}^{1,p} \supseteq \mathbb{D}^{1,q}.$$

In particular, for every $X \in \mathbb{D}^{1,p}$, with $p \geq 2$, and an approximating sequence (X_n) in L^p, we have

$$\lim_{n\to\infty} DX_n = DX, \qquad \text{in } L^2([0,T]\times\Omega).$$

Example 16.11 By using the approximating sequence in Example 16.8, it is immediate to verify that $W_t \in \mathbb{D}^{1,\infty}$ for every t. □

Proposition 16.12 (Chain rule) *Let*[5] $\varphi \in C^\infty_{\text{pol}}(\mathbb{R})$. *Then:*

i) if $X \in \mathbb{D}^{1,\infty}$, *then* $\varphi(X) \in \mathbb{D}^{1,\infty}$ *and*

$$D\varphi(X) = \varphi'(X)DX; \tag{16.3}$$

ii) if $X \in \mathbb{D}^{1,2}$ *and* φ, φ' *are bounded, then* $\varphi(X) \in \mathbb{D}^{1,2}$ *and* (16.3) *holds.*

Further, if $\varphi \in C^\infty_{\text{pol}}(\mathbb{R}^N)$ *and* $X_1, \dots, X_N \in \mathbb{D}^{1,\infty}$, *then* $\varphi(X_1, \dots, X_N) \in \mathbb{D}^{1,\infty}$ *and we have*

$$D\varphi(X_1, \dots, X_N) = \sum_{i=1}^{N} \partial_{x_i}\varphi(X_1, \dots, X_N) DX_i.$$

[5] Actually it suffices that $\varphi \in C^1$ and that both φ and its first-order derivative have at most polynomial growth.

Proof. We prove only *ii)* since the other parts can be proved essentially in an analogous way. If $X \in \mathcal{S}$, $\varphi \in C^1$ and both φ and its first-order derivative are bounded, then $\varphi(X) \in \mathcal{S}$ and the claim is obvious.

If $X \in \mathbb{D}^{1,2}$, then there exists a sequence (X_n) in $\mathcal{S}$ converging to X in $L^2(\Omega)$ and such that (DX_n) converges to DX in $L^2([0,T] \times \Omega)$. Then, by the dominated convergence theorem, $\varphi(X_n)$ tends to $\varphi(X)$ in $L^2(\Omega)$. Further, $D\varphi(X_n) = \varphi'(X_n)DX_n$ and

$$\|\varphi'(X_n)DX_n - \varphi'(X)DX\|_{L^2} \le I_1 + I_2,$$

where

$$I_1 = \|(\varphi'(X_n) - \varphi'(X))DX\|_{L^2} \xrightarrow[n\to\infty]{} 0$$

by the dominated convergence theorem and

$$I_1 = \|\varphi'(X_n)(DX - DX_n)\|_{L^2} \xrightarrow[n\to\infty]{} 0$$

since (DX_n) converges to DX and φ' is bounded. □

Example 16.13 By the chain rule, $(W_t)^2 \in \mathbb{D}^{1,\infty}$ and

$$D_s W_t^2 = 2W_t \mathbb{1}_{[0,t]}(s).$$ □

Example 16.14 Let $u \in \mathbb{L}^2$ such that $u_t \in \mathbb{D}^{1,2}$ for every t. Then

$$X := \int_0^t u_r dW_r \in \mathbb{D}^{1,2}$$

and for $s \le t$

$$D_s \int_0^t u_r dW_r = u_s + \int_s^t D_s u_r dW_r.$$

Indeed, for fixed t, we consider the sequence defined by

$$X_n := \sum_{k=1}^{k_n(t)} u_{t_n^{k-1}} \Delta_n^k, \qquad n \in \mathbb{N},$$

approximating X in $L^2(\Omega)$. Then $X_n \in \mathbb{D}^{1,2}$ and, by the chain rule, we get

$$D_s X_n = u_{t_n^{k_n(s)-1}} + \sum_{k=1}^{k_n(t)} D_s u_{t_n^{k-1}} \Delta_n^k =$$

(since u is adapted and so, by Remark 16.9, $D_s u_{t_n^k} = 0$ if $s > t_n^k$)

$$= u_{t_n^{k_n(s)-1}} + \sum_{k=k_n(s)+1}^{k_n(t)} D_s u_{t_n^{k-1}} \Delta_n^k \xrightarrow[n\to\infty]{} u_s + \int_s^t D_s u_r dW_r$$

in $L^2([0,T] \times \Omega)$. □

Example 16.15 If $u \in \mathbb{D}^{1,2}$ for every t, then we have

$$D_s \int_0^t u_r dr = \int_s^t D_s u_r dr. \qquad \square$$

Example 16.16 Let us consider the solution (X_t) of the SDE

$$X_t = x + \int_0^t b(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r, \tag{16.4}$$

with $x \in \mathbb{R}$ and the coefficients $b, \sigma \in C_b^1$. Then $X_t \in \mathbb{D}^{1,2}$ for every t and we have

$$D_s X_t = \sigma(s, X_s) + \int_s^t \partial_x b(r, X_r) D_s X_r dr + \int_s^t \partial_x \sigma(r, X_r) D_s X_r dW_r. \tag{16.5}$$

We do not go into the details of the proof of the first claim. The idea is to use an approximation argument based on the Euler scheme (cf. Paragraph 12.2): more precisely, the claim follows from the fact that (X_t) is the limit of the sequence of piecewise constant processes defined by

$$X_t^n = X_{t_n^{k-1}}^n \mathbb{1}_{I_n^k}(t), \qquad t \in [0, T],$$

with $X_{t_n^k}^n$ defined recursively by

$$X_{t_n^k}^n = X_{t_n^{k-1}}^n + b(t_n^{k-1}, X_{t_n^{k-1}}^n)\frac{1}{2^n} + \sigma(t_n^{k-1}, X_{t_n^{k-1}}^n)\Delta_n^k,$$

for $k = 1, \dots, 2^n$. Once we have proved that $X_t \in \mathbb{D}^{1,2}$, (16.5) is an immediate consequence of Examples 16.14, 16.15 and of the chain rule. $\square$

Now we use the classical method of variation of constants to get an explicit expression of $D_s X_t$. Under the assumptions of Example 16.16, we consider the process

$$Y_t = \partial_x X_t, \tag{16.6}$$

solution of the SDE

$$Y_t = 1 + \int_0^t \partial_x b(r, X_r) Y_r dr + \int_0^t \partial_x \sigma(r, X_r) Y_r dW_r. \tag{16.7}$$

Lemma 16.17 *Let Y be as in (16.7) and Z be solution of the SDE*

$$Z_t = 1 + \int_0^t ((\partial_x \sigma)^2 - \partial_x b)(r, X_r) Z_r dr - \int_0^t \partial_x \sigma(r, X_r) Z_r dW_r. \tag{16.8}$$

Then $Y_t Z_t = 1$ for every t.

Proof. We have $Y_0 Z_0 = 1$ and, omitting the arguments, by the Itô formula we have

$$\begin{aligned} d(Y_t Z_t) =& Y_t dZ_t + Z_t dY_t + d\langle Y, Z\rangle_t \\ =& Y_t Z_t \big(\left((\partial_x \sigma)^2 - (\partial_x b)\right) dt - \partial_x \sigma dW_t \\ & + \partial_x b dt + \partial_x \sigma dW_t - (\partial_x \sigma)^2 dt \big) = 0, \end{aligned}$$

and the claim follows by the uniqueness of the representation for an Itô process, Proposition 5.3. □

Proposition 16.18 *Let X, Y, Z be the solutions of the SDEs* (16.4), (16.7) *and* (16.8), *respectively. Then*

$$D_s X_t = Y_t Z_s \sigma(s, X_s). \tag{16.9}$$

Proof. We recall that, for fixed s, the process $D_s X_t$ verifies the SDE (16.5) over $[s, T]$ and we prove that $A_t := Y_t Z_s \sigma(s, X_s)$ verifies the same equation: the claim will then follow from the uniqueness results for SDE.

By (16.7) we have

$$Y_t = Y_s + \int_s^t \partial_x b(r, X_r) Y_r dr + \int_s^t \partial_x \sigma(r, X_r) Y_r dW_r;$$

multiplying by $Z_s \sigma(s, X_s)$ and using Lemma 16.17

$$\begin{aligned} \underbrace{Y_t Z_s \sigma(s, X_s)}_{=A_t} = \underbrace{Y_s Z_s}_{=1} \sigma(s, X_s) + \int_s^t \partial_x b(r, X_r) \underbrace{Y_r Z_s \sigma(s, X_s)}_{=A_r} dr \\ + \int_s^t \partial_x \sigma(r, X_r) \underbrace{Y_r Z_s \sigma(s, X_s)}_{=A_r} dW_r, \end{aligned}$$

whence the claim. □

Remark 16.19 The concept of stochastic derivative and the results that we proved up to now can be extended to the multi-dimensional case without major difficulties, but for the heavy notation. If $W = (W^1, \dots, W^d)$ is a d-dimensional Brownian motion and we denote the derivative with respect to the i-th component of W by D^i, then we can prove that, for $s \le t$

$$D_s^i W_t^j = \delta_{ij}$$

where δ_{ij} is Kronecker's delta. More generally, if X is a random variable depending only on the increments of W^j, then $D^i X = 0$ for $i \neq j$. Further, for $u \in \mathbb{L}^2$

$$D_s^i \int_0^t u_r dW_r = u_s^i + \int_s^t D_s^i u_r dW_r.$$

□

16.2 Duality

In this paragraph we introduce the adjoint operator of the Malliavin derivative and we prove a duality result that is the core tool to demonstrate the stochastic integration-by-parts formula.

Definition 16.20 *For fixed $n \in \mathbb{N}$, the family $\mathcal{P}_n$ of the n-th order simple processes consists of the processes U of the form*

$$U_t = \sum_{k=1}^{2^n} \varphi_k(\Delta_n) \mathbb{1}_{I_n^k}(t), \tag{16.10}$$

with $\varphi_k \in C^\infty_{\text{pol}}(\mathbb{R}^{2^n}; \mathbb{R})$ for $k = 1, \dots, 2^n$.

Using Notation 16.2, formula (16.10) can be rewritten more simply as

$$U_t = \varphi_{k_n(t)}(\Delta_n).$$

We observe that

$$\mathcal{P}_n \subseteq \mathcal{P}_{n+1}, \qquad n \in \mathbb{N},$$

and we define

$$\mathcal{P} := \bigcup_{n \in \mathbb{N}} \mathcal{P}_n$$

the family of simple functionals. It is apparent that

$$D : \mathcal{S} \longrightarrow \mathcal{P}$$

i.e. $DX \in \mathcal{P}$ for $X \in \mathcal{S}$. By the growth assumption on the functions φ_k in (16.10), $\mathcal{P}$ is a subspace of $L^p([0,T] \times \Omega)$ for every $p \geq 1$ and furthermore $\mathcal{P}$ is dense in $L^p([0,T] \times \Omega, \mathscr{B} \otimes \mathcal{F}_T^W)$.

Now we recall notation (16.1) and we define the adjoint operator of D.

Definition 16.21 *Given a simple process $U \in \mathcal{P}$ of the form* (16.10), *we set*

$$D^* U = \sum_{k=1}^{2^n} \left(\varphi_k(\Delta_n) \Delta_n^k - \partial_{x_n^k} \varphi_k(\Delta_n) \frac{1}{2^n} \right). \tag{16.11}$$

$D^* U$ is called Skorohod integral [313] of U: in the sequel we also write

$$D^* U = \int_0^T U_t \diamond dW_t. \tag{16.12}$$

We observe that Definition (16.11) is well-posed since *it does not depend on n*. Further, we note that, differently from the Itô stochastic integral, for the Skorohod integral we do not require the process U to be adapted. For this reason D^* is also called *anticipative stochastic integral.*

Remark 16.22 If U is adapted, then φ_k in (16.10) is $\mathcal{F}^W_{t_n^{k-1}}$-measurable and so, by Remark 16.9, $\partial_{x_n^k}\varphi_k = 0$. Consequently we have

$$\int_0^T U_t \diamond dW_t = \sum_{k=1}^{2^n} \varphi_k(\Delta_n)\Delta_n^k = \int_0^T U_t dW_t.$$

In other terms, *for an adapted stochastic process, the Skorohod integral coincides with the Itô integral.* □

A central result in Malliavin calculus is the following:

Theorem 16.23 (Duality relation) *For every $X \in \mathcal{S}$ and $U \in \mathcal{P}$ we have*

$$E\left[\int_0^T (D_t X) U_t dt\right] = E\left[X \int_0^T U_t \diamond dW_t\right]. \tag{16.13}$$

Remark 16.24 (16.13) can be written equivalently in the form

$$\langle DX, U\rangle_{L^2([0,T]\times\Omega)} = \langle X, D^* U\rangle_{L^2(\Omega)}$$

that justifies calling the Skorohod integral the adjoint operator of D.

Proof. Let U be in the form (16.10) and let $X = \varphi_0(\Delta_m)$ with $\varphi \in C^\infty_{\text{pol}}(\mathbb{R}^{2^m};\mathbb{R})$: evidently it is not restrictive to assume $m = n$. We put $\delta = \frac{1}{2^n}$ and for every $j \in \{1,\dots,2^n\}$ and $k \in \{0,\dots,2^n\}$,

$$\varphi_k^{(j)}(x) = \varphi_k(\Delta_n^1,\dots,\Delta_n^{j-1}, x, \Delta_n^{j+1},\dots,\Delta_n^{2^n}), \qquad x \in \mathbb{R}.$$

Then we have

$$E\left[\int_0^T (D_t X) U_t dt\right] = \delta E\left[\sum_{k=1}^{2^n} \partial_{x_n^k}\varphi_0(\Delta_n)\varphi_k(\Delta_n)\right] =$$

(since the Brownian increments are independent and identically distributed, $\Delta_n^k \sim \mathcal{N}_{0,\delta}$)

$$= \delta \sum_{k=1}^{2^n} E\left[\int_{\mathbb{R}} \left(\frac{d}{dx}\varphi_0^{(k)}(x)\right)\varphi_k^{(k)}(x)\frac{e^{-\frac{x^2}{2\delta}}}{\sqrt{2\pi\delta}}dx\right] =$$

(integrating by parts)

$$= \delta \sum_{k=1}^{2^n} E\left[\int_{\mathbb{R}} \varphi_0^{(k)}(x)\left(\frac{x}{\delta}\varphi_k^{(k)}(x) - \frac{d}{dx}\varphi_k^{(k)}(x)\right)\frac{e^{-\frac{x^2}{2\delta}}}{\sqrt{2\pi\delta}}dx\right] =$$

$$= E\left[\varphi_0(\Delta_n)\sum_{k=1}^{2^n}\left(\varphi_k(\Delta_n)\Delta_n^k - \partial_{x_n^k}\varphi_k(\Delta_n)\delta\right)\right],$$

and this, in view of the definition of the Skorohod integral, concludes the proof. □

As a consequence of the duality relation, we prove Lemma 16.6.

Proof (of Lemma 16.6). Let (X_n) be a sequence in $\mathcal{S}$ such that

i) $\lim_{n\to\infty} X_n = 0$ in $L^2(\Omega)$;
ii) there exists $U := \lim_{n\to\infty} DX_n$ in $L^2([0,T]\times\Omega)$.

To prove that $U = 0$, we consider $V \in \mathcal{P}$: we have, by *ii)*,

$$E\left[\int_0^T U_t V_t dt\right] = \lim_{n\to\infty} E\left[\int_0^T (D_t X_n)\, V_t dt\right] =$$

(by the duality relation and then by *i)*)

$$= \lim_{n\to\infty} E\left[X_n \int_0^T V_t \diamond dW_t\right] = 0.$$

The claim follows from the density of $\mathcal{P}$ in $L^2([0,T]\times\Omega, \mathscr{B}\otimes\mathcal{F}_T^W)$. □

Remark 16.25 In an analogous way we prove that, if (U^n) is a sequence in $\mathcal{P}$ such that

i) $\lim_{n\to\infty} U^n = 0$ in $L^2([0,T]\times\Omega)$;
ii) there exists $X := \lim_{n\to\infty} D^* U^n$ in $L^2(\Omega)$,

then $X = 0$ a.s. Then, if $p \geq 2$ and U is such that there exists a sequence (U^n) in $\mathcal{P}$ such that

i) $U = \lim_{n\to\infty} U^n$ in $L^p([0,T]\times\Omega)$;
ii) the limit $\lim_{n\to\infty} D^* U^n$ exists in $L^p(\Omega)$,

we say that U *is p-th order Skorohod-integrable* and the following definition of Skorohod integral is well-posed:

$$D^* U = \int_0^T U_t \diamond dW_t := \lim_{n\to\infty} D^* U^n, \qquad \text{in } L^2(\Omega).$$

Further, the following duality relation

$$E\left[\int_0^T (D_t X) U_t dt\right] = E\left[X \int_0^T U_t \diamond dW_t\right]$$

holds, for every $X \in \mathbb{D}^{1,2}$ and U which is Skorohod-integrable of order two. □

16.2.1 Clark-Ocone formula

The martingale representation theorem asserts that, for every $X \in L^2(\Omega, \mathcal{F}_T^W)$, there exists $u \in \mathbb{L}^2$ such that

$$X = E[X] + \int_0^T u_s dW_s. \tag{16.14}$$

If X is Malliavin differentiable, using Example 16.14 we are able to obtain the expression of u: indeed, formally[6] we have

$$D_t X = u_t + \int_t^T D_t u_s dW_s$$

and so, taking conditional expectation, we can conclude that

$$E\left[D_t X \mid \mathcal{F}_t^W\right] = u_t. \tag{16.15}$$

(16.14)-(16.15) are known as Clark-Ocone formula. Now we proceed to prove it rigorously.

Theorem 16.26 (Clark-Ocone formula) *If $X \in \mathbb{D}^{1,2}$, then*

$$X = E[X] + \int_0^T E\left[D_t X \mid \mathcal{F}_t^W\right] dW_t.$$

Proof. It is not restrictive to suppose $E[X] = 0$. For every simple adapted process $U \in \mathcal{P}$ we have, by the duality relation of Theorem 16.23,

$$E[XD^*U] = E\left[\int_0^T (D_t X) U_t dt\right] =$$

(since U is adapted)

$$= E\left[\int_0^T E\left[D_t X \mid \mathcal{F}_t^W\right] U_t dt\right].$$

On the other hand, the Skorohod integral of the adapted process U coincides with the Itô integral and by (16.14) we get

$$E[XD^*U] = E\left[\int_0^T u_t dW_t \int_0^T U_t dW_t\right] =$$

(by Itô isometry)

$$= E\left[\int_0^T u_t U_t dt\right].$$

The claim follows by density, since U is arbitrary. □

Remark 16.27 As an interesting consequence of the Clark-Ocone formula we have that, if $X \in \mathbb{D}^{1,2}$ and $DX = 0$, then X is a.s. constant. □

[6] Assuming that $u_t \in \mathbb{D}^{1,2}$ for every t.

Now we dwell on the financial interpretation of the Clark-Ocone formula: we suppose that $X \in L^2(\Omega, \mathcal{F}_T^W)$ is the payoff of a European option on an asset S. We assume that the dynamics of the discounted price under the EMM is given by

$$d\widetilde{S}_t = \sigma_t \widetilde{S}_t dW_t.$$

Then, if (α, β) is a replicating strategy for the option, we have (cf. (10.57))

$$\widetilde{X} = E\left[\widetilde{X}\right] + \int_0^T \alpha_t d\widetilde{S}_t = E\left[\widetilde{X}\right] + \int_0^T \alpha_t \sigma_t \widetilde{S}_t dW_t.$$

On the other hand, by the Clark-Ocone formula we get

$$\widetilde{X} = E\left[\widetilde{X}\right] + \int_0^T E\left[D_t \widetilde{X} \mid \mathcal{F}_t^W\right] dW_t,$$

and so we obtain the expression of the replicating strategy:

$$\alpha_t = \frac{E\left[D_t \widetilde{X} \mid \mathcal{F}_t^W\right]}{\sigma_t \widetilde{S}_t}, \qquad t \in [0, T].$$

16.2.2 Integration by parts and computation of the Greeks

In this section we prove a stochastic integration-by-parts formula and by means of some remarkable examples, we illustrate its application to the computation of the Greeks by the Monte Carlo method. As we have already said in the introduction, the techniques based on Malliavin calculus can be effective also when *poor regularity properties* are assumed on the payoff function F, i.e. just where the direct application of the Monte Carlo method gives unsatisfactory results, even if the underlying asset follows a simple geometric Brownian motion.

The stochastic integration by parts allows removing the derivative of the payoff function, thus improving the numerical approximation: more precisely, let us suppose that we want to determine $\partial_\alpha E\left[F(S_T)Y\right]$ where S_T denotes the final price of the underlying asset depending on a parameter α (e.g. α is S_0 in the case of the Delta, α is the volatility in the case of the Vega) and Y is some random variable (e.g. a discount factor). The idea is to try to express $\partial_\alpha F(S_T)Y$ in the form

$$\int_0^T D_s F(S_T) Y U_s ds,$$

for some adapted integrable process U. By using the duality relation, formally we obtain

$$\partial_\alpha E\left[F(S_T)Y\right] = E\left[F(S_T) D^*(YU)\right],$$

that, as we shall see in the following examples, can be used to get a good numerical approximation.

In this section we want to show how to apply a technique, rather than dwelling on the mathematical details, so the presentation will be somewhat informal, starting already from the next statement.

Theorem 16.28 (Stochastic integration by parts) *Let $F \in C_b^1$ and let $X \in \mathbb{D}^{1,2}$. Then the following integration by parts holds:*

$$E\left[F'(X)Y\right] = E\left[F(X)\int_0^T \frac{u_t Y}{\int_0^T u_s D_s X ds} \diamond dW_t\right], \tag{16.16}$$

for every random variable Y and for every stochastic process u for which (16.16) *is well-defined.*

Sketch of the proof. By the chain rule we have

$$D_t F(X) = F'(X) D_t X;$$

multiplying by $u_t Y$ and integrating from 0 to T we get

$$\int_0^T u_t Y D_t F(X) dt = F'(X) Y \int_0^T u_t D_t X dt,$$

whence, provided that

$$\frac{1}{\int_0^T u_t D_t X dt}$$

has good integrability properties, we have

$$F'(X)Y = \int_0^T D_t F(X) \frac{u_t Y}{\int_0^T u_s D_s X ds} dt,$$

and, taking the mean

$$E\left[F'(X)Y\right] = E\left[\int_0^T D_t F(X) \frac{u_t Y}{\int_0^T u_s D_s X ds} dt\right] =$$

(by the duality relation)

$$= E\left[F(X)\int_0^T \frac{u_t Y}{\int_0^T u_s D_s X ds} \diamond dW_t\right].$$

□

Remark 16.29 The regularity assumptions on the function F can be greatly weakened: by using a standard regularization procedure, it is possible to prove the validity of the integration-by-parts formula for weakly differentiable (or even differentiable in a distributional sense) functions.

The process u in (16.16) can be often chosen in a suitable way in order to simplify the expression of the integral on the right-hand side (cf. Examples 16.36 and 16.37).

If $u = 1$ and $Y = \partial_\alpha X$, (16.16) becomes

$$E\left[\partial_\alpha F(X)\right] = E\left[F(X)\int_0^T \frac{\partial_\alpha X}{\int_0^T D_s X ds} \diamond dW_t\right]. \tag{16.17}$$

□

In the following Examples 16.30, 16.33 and 16.34, we consider the Black-Scholes dynamics for the underlying asset of an option under the EMM and we apply the integration-by-parts formula with $X = S_T$ where

$$S_T = x\exp\left(\sigma W_T + \left(r - \frac{\sigma^2}{2}\right)T\right). \tag{16.18}$$

Example 16.30 (Delta) We observe that $D_s S_T = \sigma S_T$ and $\partial_x S_T = \frac{S_T}{x}$. Then, by (16.17) we have the following expression for the Black-Scholes Delta

$$\begin{aligned}
\Delta &= e^{-rT}\partial_x E\left[F(S_T)\right] \\
&= e^{-rT}E\left[F(S_T)\int_0^T \frac{\partial_x S_T}{\int_0^T D_s S_T ds} \diamond dW_t\right] \\
&= e^{-rT}E\left[F(S_T)\int_0^T \frac{1}{\sigma T x} dW_t\right] \\
&= \frac{e^{-rT}}{\sigma T x}E\left[F(S_T)W_T\right].
\end{aligned} \tag{16.19}$$

□

We know that in general it is not allowed to “take out” a random variable from an Itô integral (cf. Section 4.3.2): let us see now how this can be made in the case of the anticipative stochastic integral.

Proposition 16.31 *Let $X \in \mathbb{D}^{1,2}$ and let U be a second-order Skorohod-integrable process. Then*

$$\int_0^T XU_t \diamond dW_t = X\int_0^T U_t \diamond dW_t - \int_0^T (D_t X)\, U_t dt. \tag{16.20}$$

Proof. For every $Y \in \mathcal{S}$, by the duality relation, we have

$$E\left[Y D^*(XU)\right] = E\left[\int_0^T (D_t Y) X U_t dt\right] =$$

(by the chain rule)

$$= E\left[\int_0^T (D_t(YX) - YD_t X)\, U_t dt\right] =$$

(by the duality relation)

$$= E\left[Y\left(XD^*U - \int_0^T D_t X U_t dt\right)\right],$$

and the claim follows by density. □

Formula (16.20) is crucial for the computation of Skorohod integrals. The typical case is when U is adapted: then (16.20) becomes

$$\int_0^T XU_t \diamond dW_t = X\int_0^T U_t dW_t - \int_0^T (D_t X)\, U_t dt,$$

and so it is possible to express the Skorohod integral as the sum of an Itô integral and of a Lebesgue integral.

Example 16.32 By a direct application of (16.20), we have

$$\int_0^T W_T \diamond dW_t = W_T^2 - T.$$

□

Example 16.33 (Vega) Let us compute the Vega of a European option with payoff function F in the Black-Scholes model: we first notice that

$$\partial_\sigma S_T = (W_T - 2\sigma T)S_T, \qquad D_s S_T \sigma S_T.$$

Then

$$\mathcal{V} = e^{-rT}\partial_\sigma E\left[F(S_T)\right] =$$

(by the integration-by-parts formula (16.17))

$$= e^{-rT}E\left[F(S_T)\int_0^T \frac{W_T - \sigma T}{\sigma T} \diamond dW_t\right] =$$

(by (16.20))

$$= e^{-rT}E\left[F(S_T)\left(\frac{W_T - \sigma T}{\sigma T}W_T - \frac{1}{\sigma}\right)\right].$$

□

Example 16.34 (Gamma) We compute the Gamma of a European option with payoff function F in the Black-Scholes model:

$$\Gamma = e^{-rT}\partial_{xx}E\left[F(S_T)\right] =$$

(by Example 16.30)

$$= \frac{e^{-rT}}{\sigma T} E\left[\partial_x \left(\frac{F(S_T)}{x}\right) W_T\right] = -\frac{e^{-rT}}{\sigma T x^2} E\left[F(S_T)W_T\right] + \frac{e^{-rT}}{\sigma T x} J,$$

where

$$J = E\left[\partial_x F(S_T)W_T\right] = E\left[F'(S_T)\partial_x S_T W_T\right] =$$

(applying (16.16) with $u = 1$ and $Y = (\partial_x S_T) W_T = \frac{S_T W_T}{x}$)

$$= E\left[F(S_T) \int_0^T \frac{W_T}{\sigma T x} \diamond dW_T\right] =$$

(by (16.20))

$$= \frac{1}{\sigma T x} E\left[F(S_T)(W_T^2 - T)\right].$$

In conclusion

$$\Gamma = \frac{e^{-rT}}{\sigma T x^2} E\left[F(S_T)\left(\frac{W_T^2 - T}{\sigma T} - W_T\right)\right].$$

□

16.2.3 Examples

Example 16.35 We give the expression of the Delta of an arithmetic Asian option with Black-Scholes dynamics (16.18) for the underlying asset. We denote the average by

$$X = \frac{1}{T} \int_0^T S_t dt$$

and we observe that $\partial_x X = \frac{X}{x}$ and

$$\int_0^T D_s X ds = \int_0^T \int_0^T D_s S_t dt ds = \sigma \int_0^T \int_0^t S_t ds dt = \sigma \int_0^T t S_t dt. \quad (16.21)$$

Then we have

$$\Delta = e^{-rT} \partial_x E\left[F(X)\right] = \frac{e^{-rT}}{x} E\left[F'(X)X\right] =$$

(by (16.17) and (16.21))

$$= \frac{e^{-rT}}{\sigma x} E\left[F(X) \int_0^T \frac{\int_0^T S_s ds}{\int_0^T s S_s ds} \diamond dW_t\right].$$

Now formula (16.20) can be used to compute the anticipative integral: some calculation leads to the following formula (cf., for example, [211]):

$$\Delta = \frac{e^{-rT}}{x} E\left[F(X)\left(\frac{1}{I_1}\left(\frac{W_T}{\sigma} + \frac{I_2}{I_1}\right) - 1\right)\right],$$

where

$$I_j = \frac{\int_0^T t^j S_t dt}{\int_0^T S_t dt}, \qquad j = 1, 2.$$

□

Example 16.36 (Bismut-Elworthy formula) We extend Example 16.30 to the case of a model with local volatility

$$S_t = x + \int_0^t b(s, S_s)ds + \int_0^t \sigma(s, S_s)dW_s.$$

Under suitable assumptions on the coefficients, we prove the following Bismut-Elworthy formula:

$$E\left[\partial_x F(S_T)G\right] = \frac{1}{T}E\left[F(S_T)\left(G\int_0^T \frac{\partial_x S_t}{\sigma(t, S_t)}dW_t - \int_0^T D_t G\frac{\partial_x S_t}{\sigma(t, S_t)}dt\right)\right], \tag{16.22}$$

for every $G \in \mathbb{D}^{1,\infty}$.

We recall that, by Proposition 16.18, we have

$$D_s S_T = Y_T Z_s \sigma(s, S_s), \tag{16.23}$$

since

$$Y_t := \partial_x S_t =: Z_t^{-1}.$$

Let us apply (16.16) after choosing

$$X = S_T, \qquad Y = GY_T, \qquad u_t = \frac{Y_t}{\sigma(t, S_t)},$$

to get

$$\begin{aligned} E\left[\partial_x F(S_T)G\right] &= E\left[F'(S_T)Y_T G\right] \\ &= E\left[F(S_T)\int_0^T \frac{GY_T Y_t}{\sigma(t, S_t)}\frac{1}{\int_0^T D_s S_T \frac{Y_s}{\sigma(s,S_s)}ds} \diamond dW_t\right] \end{aligned}$$

(by (16.23))

$$= E\left[F(S_T)\int_0^T \frac{GY_t}{\sigma(t, S_t)} \diamond dW_t\right]$$

and (16.22) follows from Proposition 16.31, since $\frac{Y_t}{\sigma(t,S_t)}$ is adapted. □

Example 16.37 In this example, taken from [19], we consider the Heston model

$$\begin{cases} dS_t = \sqrt{\nu_t} S_t dB_t^1, \\ d\nu_t = k(\bar{\nu} - \nu_t)dt + \eta\sqrt{\nu_t} dB_t^2, \end{cases}$$

where (B^1, B^2) is a correlated Brownian motion

$$B_t^1 = \sqrt{1-\varrho^2} W_t^1 + \varrho W_t^2, \qquad B_t^2 = W_t^2,$$

with W a standard 2-dimensional Brownian motion and $\varrho \in]-1,1[$. We want to compute the sensitivity of the price of an option with payoff F with respect to the correlation parameter ϱ.

First of all we observe that

$$S_T = S_0 \exp\left(\sqrt{1-\varrho^2} \int_0^T \sqrt{\nu_t} dW_t^1 + \varrho \int_0^T \sqrt{\nu_t} dW_t^2 - \frac{1}{2}\int_0^T \nu_t dt \right),$$

and so

$$\partial_\varrho S_T = S_T G, \qquad G := -\frac{\varrho}{\sqrt{1-\varrho^2}} \int_0^T \sqrt{\nu_t} dW_t^1 + \int_0^T \sqrt{\nu_t} dW_t^2. \tag{16.24}$$

Further, if we denote by D^1 the Malliavin derivative relative to the Brownian motion W^1, by Remark 16.19, we get $D_s^1 \nu_t = 0$ and

$$D_s^1 S_T = S_T \sqrt{1-\varrho^2}\sqrt{\nu_s}. \tag{16.25}$$

Then

$$\partial_\varrho E\left[F(S_T)\right] = E\left[F'(S_T)\partial_\varrho S_T\right] =$$

(by integrating by parts and choosing $X = S_T$, $Y = \partial_\varrho S_T$ and[7] $u_t = \frac{1}{\sqrt{\nu_t}}$ in (16.16))

$$= E\left[F(S_T) \int_0^T \frac{\partial_\varrho S_T}{\sqrt{\nu_t} \int_0^T \frac{D_s^1 S_T}{\sqrt{\nu_s}} ds} \diamond dW_t^1 \right] =$$

(by (16.24) and (16.25))

$$= \frac{1}{T\sqrt{1-\varrho^2}} E\left[F(S_T) \int_0^T \frac{G}{\sqrt{\nu_t}} \diamond dW_t^1 \right] =$$

(by Proposition 16.31 and since ν is adapted)

$$= \frac{1}{T\sqrt{1-\varrho^2}} E\left[F(S_T) \left(G \int_0^T \frac{1}{\sqrt{\nu_t}} dW_t^1 - \int_0^T \frac{D_t^1 G}{\sqrt{\nu_t}} dt \right) \right] =$$

(since $D_t^1 G = -\varrho\sqrt{\frac{\nu_t}{1-\varrho^2}}$)

$$= \frac{1}{T\sqrt{1-\varrho^2}} E\left[F(S_T)\left(G\int_0^T \frac{1}{\sqrt{\nu_t}} dW_t^1 + \frac{\varrho T}{\sqrt{1-\varrho^2}}\right)\right]. \qquad \square$$