Stochastic differential equations and interpretation of parabolic partial differential equations

Benjamin Jourdain*

January 25, 2006

1 Basic probability theory

Let Ω be endowed with a probability measure $\mathbb{P}$.

Definition 1.1. *We say that a random vector $X : \Omega \to \mathbb{R}^d$ has the density p where $p : \mathbb{R}^d \to \mathbb{R}_+$ is such that $\int_{\mathbb{R}^d} p(x)dx = 1$ if*

$$\forall A \subset \mathbb{R}^d,\ \mathbb{P}(X \in A) = \int_A p(x)dx \Leftrightarrow \forall f : \mathbb{R}^d \to \mathbb{R} \text{ bounded },\ \mathbb{E}(f(X)) = \int_{\mathbb{R}^d} f(x)p(x)dx.$$

Intuitively $\mathbb{P}(X \in [x, x + dx]) = p(x)dx$.

Example 1.2. The real random variable U is uniformly distributed on the interval $[a, b]$ where $a < b$ if it has the density $p(u) = \frac{1_{\{a<u<b\}}}{b-a}$.

Definition 1.3. *A real random variable X with density p is*

1. *integrable if $\int_{\mathbb{R}} |x|p(x)dx < +\infty$ and then its expectation is*
$$\mathbb{E}(X) = \int_{\mathbb{R}} xp(x)dx,$$
2. *square integrable if $\mathbb{E}(X^2) = \int_{\mathbb{R}} x^2 p(x)dx < +\infty$ and then its variance is*
$$\text{Var }(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] = \mathbb{E}(X^2) - (\mathbb{E}(X))^2.$$
The standard deviation of X is $\sigma(X) = \sqrt{\text{Var }(X)}$.

Example 1.4. We say that the real variable X is Gaussian (or normal) with parameter (μ, σ^2) with $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*$ and denote $X \sim \mathcal{N}(\mu, \sigma^2)$ if X possesses the density

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

*ENPC-CERMICS, 6-8 avenue Blaise Pascal, Cité Descartes, Champs sur Marne, 77455 Marne la Vallée Cedex 2, e-mail:jourdain@cermics.enpc.fr

Then X is square integrable and

$$\mathbb{E}(X) = \int_{\mathbb{R}} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{dx}{\sigma\sqrt{2\pi}} = \mu + \int_{\mathbb{R}} (x-\mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{dx}{\sigma\sqrt{2\pi}} = \mu - \left[\frac{\sigma}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right]_{-\infty}^{+\infty} = \mu.$$

Using this result, then the change of variable $y = (x-\mu)/\sigma$, one obtains

$$\begin{aligned} \mathrm{Var}\,(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] &= \int_{\mathbb{R}} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{dx}{\sigma\sqrt{2\pi}} = \sigma^2 \int_{\mathbb{R}} y \times y e^{-\frac{y^2}{2}} \frac{dy}{\sqrt{2\pi}} \\ &= \sigma^2 \left(\left[-\frac{y e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \right]_{-\infty}^{+\infty} + \int_{\mathbb{R}} e^{-\frac{y^2}{2}} \frac{dy}{\sqrt{2\pi}} \right) = \sigma^2. \end{aligned}$$

Hence the expectation and the variance of X are respectively μ and σ^2. By a computation similar to the variance, one obtains

$$\mathbb{E}\left[(X - \mathbb{E}(X))^4\right] = \sigma^4 \int_{\mathbb{R}} y^3 \times y e^{-\frac{y^2}{2}} \frac{dy}{\sqrt{2\pi}} = 3\sigma^4 \int_{\mathbb{R}} y^2 \times y e^{-\frac{y^2}{2}} \frac{dy}{\sqrt{2\pi}} = 3\sigma^4. \tag{1}$$

Proposition 1.5. *The expectation is linear : if X and Y are integrable real random variables and $\lambda \in \mathbb{R}$ then $X + \lambda Y$ is integrable and*

$$\mathbb{E}(X + \lambda Y) = \mathbb{E}(X) + \lambda \mathbb{E}(Y).$$

The concept of independence plays a key role in probability theory.

Definition 1.6. *We say that the real random variables $X_1, \ldots, X_n$ are independent if*

$$\forall A_1, \ldots, A_n \subset \mathbb{R},\ \mathbb{P}(X_1 \in A_1, \ldots, X_n \in A_n) = \mathbb{P}(X_1 \in A_1) \ldots \mathbb{P}(X_n \in A_n).$$

Independence of $X_1, \ldots, X_n$ implies that for all functions $f_i : \mathbb{R} \to \mathbb{R}$ such that $f_i(X_i)$ is integrable for $i \in \{1, \ldots, n\}$, then $\prod_{i=1}^n f_i(X_i)$ is integrable and

$$\mathbb{E}\left[\prod_{i=1}^n f_i(X_i)\right] = \prod_{i=1}^n \mathbb{E}\left[f_i(X_i)\right].$$

In case each X_i has the density p_i, this means that the random vector $(X_1, \ldots, X_n)$ has the density $p(x_1, \ldots, x_n) = \prod_{i=1}^n p_i(x_i)$.

Proposition 1.7. *If $X_1, \ldots, X_n$ are* **independent** *square integrable real variables, then $X_1 + \ldots + X_n$ is square integrable and*

$$\mathrm{Var}\,(X_1 + \ldots + X_n) = \sum_{i=1}^n \mathrm{Var}\,(X_i).$$

Proof :

$$\begin{aligned}
\text{Var } (X_1+\ldots+X_n) &= \mathbb{E}\Bigg[\Bigg(\sum_{i=1}^{n}(X_i-\mathbb{E}(X_i))\Bigg)^2\Bigg] \\
&= \sum_{i,j=1}^{n}\mathbb{E}\left[(X_i-\mathbb{E}(X_i))(X_j-\mathbb{E}(X_j))\right] \text{ by linearity of } \mathbb{E}, \\
&= \sum_{i=1}^{n}\mathbb{E}\left[(X_i-\mathbb{E}(X_i))^2\right] + \sum_{1\le i\ne j\le n}\mathbb{E}[X_i-\mathbb{E}(X_i)]\mathbb{E}[X_j-\mathbb{E}(X_j)] \text{ by independence,} \\
&= \sum_{i=1}^{n}\text{Var } (X_i) + 0.
\end{aligned}$$

□

Let us recall the two fundamental convergence theorems in probability theory : the strong law of large numbers and the central limit theorem. We assume that $(X_i)_{i\geq 1}$ is a sequence of real random variables such that for each $n \geq 1$,

$$\forall A_1,\ldots,A_n \subset \mathbb{R},\ \mathbb{P}(X_1\in A_1\ldots,X_n\in A_n)=\mathbb{P}(X_1\in A_1)\mathbb{P}(X_1\in A_2)\ldots\mathbb{P}(X_1\in A_n).$$

In other words, the variables $(X_i)_{i\geq 1}$ are independent and identically distributed. When X_1 has the density $p(x_1)$ this means that $(X_1,\ldots,X_n)$ has the density $q(x_1,\ldots,x_n) = p(x_1)\ldots p(x_n)$.
The strong law of large numbers states the convergence of the empirical mean

$$\boxed{\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n}X_i}$$

to the expectation $\mathbb{E}(X_1)$ common to the variables X_i as $n \to +\infty$.

Theorem 1.8. *If the random variables X_i are integrable then*

$$\mathbb{P}\left(\lim_{n\to+\infty}\bar{X}_n=\mathbb{E}(X_1)\right)=1.$$

The strong law of large numbers justifies the Monte-Carlo method which consists in approximating $\mathbb{E}(X_1)=\int xp(x)dx$ by the empirical mean $\bar{X}_n$. The central limit theorem gives the precision of this approximation.

Theorem 1.9. *Assume that the random variables X_i are square integrable. Then, as $n\to+\infty$, the distribution of $\sqrt{n}(\mathbb{E}(X_1)-\bar{X}_n)$ converges to $\mathcal{N}(0, \text{Var }(X_1))$.*
When Var $(X_1) > 0$ *this means that*

$$\forall a<b\in\mathbb{R},\ \mathbb{P}(a\le\sqrt{n}(\mathbb{E}(X_1)-\bar{X}_n)\le b)\to\int_a^b \exp\left(-\frac{y^2}{2\text{Var }(X_1)}\right)\frac{dy}{\sqrt{2\pi\text{Var }(X_1)}}.$$

For each $n \geq 1$, the expectation of the random variable $\sqrt{n}(\mathbb{E}(X_1)-\bar{X}_n)$ is 0 and according to Proposition 1.7

$$\begin{aligned}
\text{Var } (\sqrt{n}(\mathbb{E}(X_1)-\bar{X}_n)) = \mathbb{E}\left[n(\bar{X}_n-\mathbb{E}(X_1))^2\right] &= \frac{1}{n}\mathbb{E}\left[\left(\sum_{i=1}^{n}X_i-n\mathbb{E}(X_1)\right)^2\right] \\
&= \frac{1}{n}\text{Var }\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}\text{Var }(X_i)=\text{Var }(X_1).
\end{aligned}$$

Hence the expectation and the variance of the limit distribution $\mathcal{N}(0, \text{Var } (X_1))$ are not surprising.

Remark 1.10. Choosing $b = -a = 1.96\sqrt{\text{Var } (X_1)}$ gives the following confidence interval for the Monte-Carlo approximation of $\mathbb{E}(X_1)$:

$$\mathbb{P}\left(\mathbb{E}(X_1) \in \left[\bar{X}_n - \frac{1.96\sqrt{\text{Var } (X_1)}}{\sqrt{n}}, \bar{X}_n + \frac{1.96\sqrt{\text{Var } (X_1)}}{\sqrt{n}}\right]\right) \simeq \int_{-1.96}^{1.96} e^{-z^2/2}\frac{dz}{\sqrt{2\pi}} = 0.95.$$

Of course in general when one wants to compute $\mathbb{E}(X_1)$, one does not know Var (X_1). But Var (X_1) can be replaced by the following approximation $V_n = \frac{1}{n}\sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$ which is convergent according to the strong law of large numbers. The derived confidence interval gives the precision of the Monte-Carlo computation with no supplementary cost. In order to obtain a better precision for fixed n, one can compute the empirical mean $\bar{Y}_n$ of random variables Y_i with the same expectation as X_1 but smaller variance. This is the principle of variance reduction techniques.

2 Brownian motion

Let us first consider the simple symmetric random walk on the set of integers $\mathbb{Z}$. The walker starts at time 0 at the origin. From time n to time $n+1$, it moves from its present location to one of its 2 nearest neighbours with equal probabilities 1/2. The mathematical formalization of this walk is based on a sequence $(\xi_i)_{i\geq 1}$ of independent and identically distributed variables such that $\mathbb{P}(\xi_i = 1) = \mathbb{P}(\xi_i = -1) = \frac{1}{2}$. The variable ξ_{n+1} represents the move of the walker from time n to time $n+1$. By induction, the position of the walker at time n is

$$S_n = \sum_{i=1}^n \xi_i \text{ (Convention : } S_0 = 0).$$

The Brownian motion is obtained by a suitable renormalization of this random walk :

$$\forall t \in [0, +\infty[,\ B_t^k = \frac{1}{\sqrt{k}} S_{\lfloor kt \rfloor}$$

where $k \in \mathbb{N}^*$ and for $x \in \mathbb{R}$, $\lfloor x \rfloor$ denotes the integer part of x.
Let $t > 0$. Setting $X_i = \sqrt{t}\xi_i$ and $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, one obtains

$$B_t^k = \frac{1}{\sqrt{kt}}\left(\lfloor kt \rfloor \bar{X}_{\lfloor kt \rfloor}\right) = \sqrt{\frac{\lfloor kt \rfloor}{kt}}\left(\sqrt{\lfloor kt \rfloor}\bar{X}_{\lfloor kt \rfloor}\right).$$

By symmetry $\mathbb{E}(X_1) = \mathbb{E}(\xi_1) = 0$ and since $\xi_1^2 = 1$, Var $(X_1) = \mathbb{E}(X_1^2) = \mathbb{E}(t\xi_1^2) = t$. Hence the central limit theorem 1.9 implies that as k tends to ∞, $\sqrt{\lfloor kt \rfloor}\bar{X}_{\lfloor kt \rfloor}$ converges in distribution to a random variable $B_t \sim \mathcal{N}(0, t)$. As $\sqrt{\frac{\lfloor kt \rfloor}{kt}}$ converges to 1, one deduces that B_t^k converges in distribution to B_t.
Now if $s > t$, by a similar reasoning one obtains the distribution of $B_s^k - B_t^k$ converges to $\mathcal{N}(0, s-t)$. Since $B_s^k - B_t^k = \frac{1}{\sqrt{k}}\sum_{i=\lfloor kt \rfloor+1}^{\lfloor ks \rfloor} \xi_i$, this random variable is independent from B_t^k. In the limit $k \to +\infty$, independence is preserved. Hence $(B_t^k, B_s^k - B_t^k)$ converges in distribution to $(B_t, B_s - B_t)$ where $B_t \sim \mathcal{N}(0, t)$ and $B_s - B_t \sim \mathcal{N}(0, s-t)$ are independent random variables.

More generally when $0 = t_0 \leq t_1 \leq \ldots \leq t_n$, the vector $(B^k_{t_1}, B^k_{t_2} - B^k_{t_1}, \ldots, B^k_{t_n} - B^k_{t_{n-1}})$ converges in distribution to $(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}})$ where for $i \in \{0, \ldots n-1\}$, the random variables $B_{t_{i+1}} - B_{t_i}$ are independent and respectively distributed according to $\mathcal{N}(0, t_{i+1} - t_i)$.
In fact the whole family $(B^k_t)_{t\geq 0}$ converges in distribution to the standard Brownian motion.

Definition 2.1. *We say that a family $(B_t)_{t\geq 0}$ is a standard Brownian motion if*

1. $B_0 = 0$,
2. *For $0 = t_0 \leq t_1 \leq \ldots \leq t_n$, the random variables $B_{t_{i+1}} - B_{t_i}$, $i \in \{0, \ldots n-1\}$ are independent and respectively distributed according to $\mathcal{N}(0, t_{i+1} - t_i)$,*
3. *the sample-paths $t \in [0, +\infty[\to B_t$ are continuous :*

$$\mathbb{P}(t \in [0, +\infty[\to B_t \textit{ continuous}) = 1.$$

Remark 2.2.

- The Brownian motion is a **continuous** (see 3) process with **independent** and **stationary** increments (see 2).
- Continuity of the sample-paths is of course not obtained by the renormalization of the symmetric random walk described above. In fact, one can prove that the sample-paths of the Brownian motion are locally Hölder continuous with exponent $\alpha < 1/2$ but nowhere differentiable.

The following property concerning the quadratic variation of the Brownian path is the key to understandind the specificity of stochastic differential calculus.

Proposition 2.3. *Let $t > 0$, $n \in \mathbb{N}^*$ and $t_k = \frac{kt}{n}$ for $k \in \{0, \ldots, n\}$.*

$$\mathbb{E}\left[\left(\sum_{k=0}^{n-1}(B_{t_{k+1}} - B_{t_k})^2 - t\right)^2\right] = \frac{2t^2}{n}.$$

Proof : Since $B_{t_{k+1}} - B_{t_k} \sim \mathcal{N}(0, \frac{t}{n})$,

$$\mathbb{E}((B_{t_{k+1}} - B_{t_k})^2) = \mathrm{Var}\,(B_{t_{k+1}} - B_{t_k}) = \frac{t}{n}.$$

By linearity of the expectation one deduces that $\mathbb{E}\left[\sum_{k=0}^{n-1}(B_{t_{k+1}} - B_{t_k})^2\right] = t$. Hence

$$\mathbb{E}\left[\left(\sum_{k=0}^{n-1}(B_{t_{k+1}} - B_{t_k})^2 - t\right)^2\right] = \mathrm{Var}\left(\sum_{k=0}^{n-1}(B_{t_{k+1}} - B_{t_k})^2\right) = \sum_{k=0}^{n-1}\mathrm{Var}\,((B_{t_{k+1}} - B_{t_k})^2),$$

by independence of the Brownian increments and Proposition 1.7. Since according to (1),

$$\mathrm{Var}\,((B_{t_{k+1}} - B_{t_k})^2) = \mathbb{E}\left[(B_{t_{k+1}} - B_{t_k})^4\right] - \left(\mathbb{E}\left[(B_{t_{k+1}} - B_{t_k})^2\right]\right)^2 = \frac{3t^2}{n^2} - \frac{t^2}{n^2} = \frac{2t^2}{n^2},$$

one easily concludes. $\square$

One has

$$B_{t_k}(B_{t_{k+1}} - B_{t_k}) = \frac{1}{2}\left(B_{t_{k+1}}^2 - B_{t_k}^2 - (B_{t_{k+1}} - B_{t_k})^2\right).$$

Summing this equality for $k \in \{0, \ldots, n-1\}$, one obtains

$$\sum_{k=0}^{n-1} B_{t_k}(B_{t_{k+1}} - B_{t_k}) = \frac{1}{2}\left(B_t^2 - \sum_{k=0}^{n-1}(B_{t_{k+1}} - B_{t_k})^2\right).$$

According to Proposition 2.3, the right-hand-side converges to $\frac{1}{2}(B_t^2 - t)$, as $n \to +\infty$. So does the Riemann sum in the left-hand-side. Since we are going to define the stochastic integral $\int_0^t B_s dB_s$ as the limit of this sum, we obtain

$$\int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t). \tag{2}$$

Notice that when g is a C^1 function on $[0, t]$ then

$$\sum_{k=0}^{n-1}(g(t_{k+1}) - g(t_k))^2 \leq \sum_{k=0}^{n-1} \sup_{s\in[0,t]} (g'(s))^2 \frac{t^2}{n^2} = \frac{\sup_{s\in[0,t]}(g'(s))^2 t^2}{n} \overset{n\to+\infty}{\longrightarrow} 0.$$

Therefore $\int_0^t g(s)dg(s) = \frac{1}{2}(g^2(t) - g^2(0))$.
On this simple example, we can see the main difference between usual and stochastic integral calculus. The lack of regularity of the Brownian path with respect to the time variable implies that the quadratic variation of this path does not vanishes as $n \to +\infty$. As a consequence a supplementary term ($-\frac{1}{2}t$ in the above example) appears.

3 Stochastic integrals and Itô's formula

3.1 Construction of stochastic integrals with respect to $(B_t)_{t\geq 0}$

In order to define Itô's stochastic integral $\int_0^t H_s dB_s$ we have to assume that for $s \geq 0$, H_s does not depend on the increments of the Brownian motion posterior to s.

Definition 3.1. *We say that the process $(H_t)_{t\geq 0}$ is adapted if for each $t \geq 0$, H_t is independent from the future increments $(B_s - B_t)_{s\geq t}$ of the Brownian motion.*

Example 3.2. The process $(B_t)_{t\geq 0}$ is adapted but $(B_{t+1})_{t\geq 0}$ is not.
If X_0 is independent from the Brownian motion and for each $t \geq 0$, $H_t = f_t(X_0, (B_s)_{s\leq t})$ for a determinic function f_t, then $(H_t)_{t\geq 0}$ is adapted. For instance, $(g(X_0, \max_{s\in[0,t]} B_s))_{t\geq 0}$ is an adapted process.

To construct Itô's stochastic integral we fix a time-horizon $T > 0$ and proceed with three steps.

First step : Let $(H_t)_{t\in[0,T]}$ be a simple process of the form

$$H_t = \sum_{k=0}^{n-1} Z_k 1_{]t_k, t_{k+1}]}(t)$$

where $0 = t_0 \leq t_1 \leq \ldots \leq t_n = T$ and for $k \in \{0, \ldots, n-1\}$, Z_k is a bounded random variable independent from $(B_s - B_{t_k})_{s \geq t_k}$.
For $t \in]t_k, t_{k+1}]$ with $k \in \{0, \ldots, n-1\}$, we set

$$\int_0^t H_s dB_s = \sum_{j=0}^{k-1} Z_j (B_{t_{j+1}} - B_{t_j}) + Z_k (B_t - B_{t_k}).$$

Now if for $s \geq 0$, $s \wedge t = \min(s, t)$, one has

$$\mathbb{E}\left(\int_0^t H_s dB_s\right) = \sum_{j=0}^{k-1} \mathbb{E}(Z_j)\mathbb{E}(B_{t_{j+1}} - B_{t_j}) + \mathbb{E}(Z_k)\mathbb{E}(B_t - B_{t_k}) = 0,$$

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^t H_s dB_s\right)^2\right] &= \mathbb{E}\left[\left(\sum_{j=0}^{k} Z_j (B_{t_{j+1}\wedge t} - B_{t_j \wedge t})\right)^2\right] \\
&= \sum_{j=0}^{k} \mathbb{E}(Z_j^2)\mathbb{E}\left[(B_{t_{j+1}\wedge t} - B_{t_j\wedge t})^2)\right] \\
&\quad + 2 \sum_{1 \leq j < l \leq k} \mathbb{E}\left[Z_j (B_{t_{j+1}\wedge t} - B_{t_j \wedge t}) Z_l\right] \mathbb{E}(B_{t_{l+1}\wedge t} - B_{t_l \wedge t}) \\
&= \sum_{j=0}^{k} \mathbb{E}(Z_j^2)(t_{j+1} \wedge t - t_j \wedge t) = \sum_{j=0}^{k-1} \mathbb{E}(Z_j^2)(t_{j+1} - t_j) + \mathbb{E}(Z_k^2)(t - t_k)
\end{aligned}$$

The two equalities write

$$\forall t \in [0, T],\ \mathbb{E}\left[\int_0^t H_s dB_s\right] = 0 \text{ and } \mathbb{E}\left[\left(\int_0^t H_s dB_s\right)^2\right] = \mathbb{E}\left[\int_0^t H_s^2 ds\right]. \quad (3)$$

The isometry equality which follows from the adaptation of the process $(H_t)_{t\in[0,T]}$ is the key property which permits to generalize the construction.

Second step : Let $(H_t)_{t\in[0,T]}$ be an adapted process such that $\mathbb{E}\left(\int_0^T H_s^2 ds\right) < +\infty$. Then there is a sequence of simple processes $(H_t^p)_{t\in[0,T]}$, $p \geq 1$ such that $\lim_{p\to+\infty} \mathbb{E}\left(\int_0^T (H_s - H_s^p)^2 ds\right) = 0$. By (3), for any $t \in [0, T]$, $\int_0^t H_s^p dB_s$ converges to a limit satisfying (3) and the limit does not depend of the approximating sequence of simple processes. We define $\int_0^t H_s dB_s$ as this limit.

Last step : Let $(H_t)_{t\in[0,T]}$ be an adapted process such that $\mathbb{P}\left(\int_0^T H_s^2 ds < +\infty\right) = 1$[1]. For $m \in \mathbb{N}^*$ we set

$$\tau_m = \inf\left\{t \in [0, T] : \int_0^t H_s^2 ds \geq m\right\} \text{ (convention } \inf\emptyset = T).$$

Now $(H_t^m = 1_{\{t\leq\tau_m\}} H_t)_{t\in[0,T]}$ is an adapted process such that $\mathbb{E}\left(\int_0^T (H_s^m)^2 ds\right) \leq m$. Moreover the sequence $(\tau_m)_{m\geq1}$ is non-decreasing and for m large enough τ_m is equal to T. Then for $t \in [0, T]$, we define

$$\int_0^t H_s dB_s = \begin{cases} \int_0^t H_s^1 dB_s & \text{if } t \leq \tau_1 \\ \int_0^t H_s^m dB_s & \text{where } m \geq 2 \text{ is such that } \tau_{m-1} < t \leq \tau_m \text{ otherwise} \end{cases}.$$

[1]Notice that $\mathbb{E}\left(\int_0^T H_s^2 ds\right) < +\infty \Rightarrow \mathbb{P}\left(\int_0^T H_s^2 ds < +\infty\right) = 1$ but the converse is not true.

Proposition 3.3. *Let $(H_t)_{t\in[0,T]}$ be an adapted process such that $\mathbb{P}\left(\int_0^T H_s^2 ds < +\infty\right) = 1$. Then $(\int_0^t H_s dB_s)_{t\in[0,T]}$ is an adapted continuous process. Moreover if $\mathbb{E}\left(\int_0^T H_s^2 ds\right) < +\infty$, then (3) holds.*

Remark 3.4. When $(H_s)_{s\in[0,T]}$ is a simple process and for $0 \le r \le t \le T$, Y is a bounded random variable independent from $(B_s - B_r)_{s\ge r}$, then one easily checks[2] that

$$\mathbb{E}\left(Y\int_0^t H_s dB_s\right) = \mathbb{E}\left(Y\int_0^r H_s dB_s\right).$$

This equality is a simple formulation of the so-called martingale property of the stochastic integral. It is preserved for the adapted processes $(H_s)_{s\in[0,T]}$ such that $\mathbb{E}\left(\int_0^T H_s^2 ds\right) < +\infty$, which have been considered in the second step of the construction.

3.2 Itô's formula

The contruction of the stochastic integral by an approximating procedure based on the isometry property (3) does not really give intuition about this integral. The main tool which permits to understand stochastic integration is Itô's formula.

Proposition 3.5. *Let $f(t,x)$ be a $C^{1,2}$ function (globally C^1 and C^2 with respect to x) on $[0,+\infty[\times\mathbb{R}$. Then*

$$\forall t \ge 0,\ f(t,B_t) = f(0,0) + \int_0^t \frac{\partial f}{\partial s}(s,B_s)ds + \int_0^t \frac{\partial f}{\partial x}(s,B_s)dB_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s,B_s)ds.$$

Proof : The proof is based on Taylor expansions. Let $n \in \mathbb{N}^*$ and $t_k = kt/n$ for $k \in \{0,\dots,n\}$.

$$\begin{aligned} f(t,B_t) - f(0,0) &= \sum_{k=0}^{n-1}\left(f(t_{k+1},B_{t_{k+1}}) - f(t_k,B_{t_k})\right) \\ &= \sum_{k=0}^{n-1}\frac{\partial f}{\partial s}(t_k,B_{t_k})(t_{k+1}-t_k) + \sum_{k=0}^{n-1}\frac{\partial f}{\partial x}(t_k,B_{t_k})(B_{t_{k+1}} - B_{t_k}) \\ &\quad + \frac{1}{2}\sum_{k=0}^{n-1}\frac{\partial^2 f}{\partial x^2}(t_k,B_{t_k})(B_{t_{k+1}} - B_{t_k})^2 + R_t^n \end{aligned}$$

The usual Riemann sum $\sum_{k=0}^{n-1}\frac{\partial f}{\partial s}(t_k,B_{t_k})(t_{k+1}-t_k)$ converges to $\int_0^t \frac{\partial f}{\partial s}(s,B_s)ds$ as n tends to $+\infty$. The second term of the right-hand-side tends to the stochastic integral $\int_0^t \frac{\partial f}{\partial x}(s,B_s)dB_s$. By a generalization of Proposition 2.3, one checks that the third term converges to $\frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s,B_s)ds$. The remainder R_t^n vanishes when $n \to +\infty$. $\square$

We see that the term $\frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s,B_s)ds$ which would not appear in usual differential calculus comes from the non-zero quadratic variation of the Brownian path.
For the choice $f(t,x) = \frac{1}{2}x^2$, since $\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial x} = x$ and $\frac{\partial^2 f}{\partial x^2} = 1$, one recovers (2).

In order to generalize Itô's formula, we introduce Itô's processes.

[2] We assume that $t \in]t_k, t_{k+1}]$ and that $r = t_l$ with $l \le k$ (if necessary, we add r to $\{t_0,\dots,t_n\}$). Then $\mathbb{E}(Y\int_0^t H_s dB_s) = \mathbb{E}(Y\sum_{j=0}^{l-1} Z_j(B_{t_{j+1}} - B_{t_j})) + \sum_{j=l-1}^{k-1}\mathbb{E}(YZ_j)\mathbb{E}(B_{t_{j+1}} - B_{t_j}) + \mathbb{E}(YZ_j)\mathbb{E}(B_t - B_{t_k}) = \mathbb{E}(Y\int_0^r H_s dB_s) + 0 + 0$.

Definition 3.6. *We say that $(X_t)_{t\geq 0}$ is an Itô's process if*

$$\forall t \geq 0,\ X_t = X_0 + \int_0^t H_s dB_s + \int_0^t K_s ds \text{ where}$$

1. X_0 *is independent from* $(B_t)_{t\geq 0}$,
2. $(H_t)_{t\geq 0}$ *is an adapted process such that* $\forall t \geq 0$, $\mathbb{P}\left(\int_0^t H_s^2 ds < +\infty\right) = 1$,
3. $(K_t)_{t\geq 0}$ *is an adapted process such that* $\forall t \geq 0$, $\mathbb{P}\left(\int_0^t |K_s| ds < +\infty\right) = 1$.

An Itô's process is a continuous and adapted process. In addition since the quadratic variation of the usual integral with respect to ds vanishes whereas the one of the stochastic integral does not, one can check uniqueness of the decomposition of Itô's processes (i.e. uniqueness of $(H_t, K_t)_{t\geq 0}$).

By generalizing the proof of Proposition 3.5, one obtains

Theorem 3.7. *Let* $X_t = X_0 + \int_0^t H_s dB_s + \int_0^t K_s ds$ *be an Itô's process and* f *a* $C^{1,2}$ *function on* $[0, +\infty[\times\mathbb{R}$. *Then*

$$\forall t \geq 0,\ f(t, X_t) = f(0, X_0) + \int_0^t \frac{\partial f}{\partial s}(s, X_s) ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s) dX_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s) H_s^2 ds, \tag{4}$$

where of course $dX_s = H_s dB_s + K_s ds$.

One easily deduces the following integration by parts formula :

Corollary 3.8. *Let* $X_t = X_0 + \int_0^t H_s^X dB_s + \int_0^t K_s^X ds$ *and* $Y_t = Y_0 + \int_0^t H_s^Y dB_s + \int_0^t K_s^Y ds$ *denote two Itô's processes. Then*

$$\forall t \geq 0,\ X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \int_0^t H_s^X H_s^Y ds.$$

Proof : Applying Itô's formula (4) to the processes $Z_t = X_t + Y_t$, X_t and Y_t for the choice $f(t, x) = \frac{1}{2}x^2$ ($\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial x} = x$ and $\frac{\partial^2 f}{\partial x^2} = 1$), one obtains

$$\begin{aligned}
\frac{1}{2}(X_t + Y_t)^2 &= \frac{1}{2}(X_0 + Y_0)^2 + \int_0^t (X_s + Y_s)(dX_s + dY_s) + \frac{1}{2}\int_0^t (H_s^X + H_s^Y)^2 ds \\
\frac{1}{2}X_t^2 &= \frac{1}{2}X_0^2 + \int_0^t X_s dX_s + \frac{1}{2}\int_0^t (H_s^X)^2 ds \\
\frac{1}{2}Y_t^2 &= \frac{1}{2}Y_0^2 + \int_0^t Y_s dY_s + \frac{1}{2}\int_0^t (H_s^Y)^2 ds.
\end{aligned}$$

One concludes by subtracting the last two equalities to the first one. □

Let now $(B_t^1)_{t\geq 0}, \ldots, (B_t^d)_{t\geq 0}$ denote d independent standard real Brownian motions. Then $W_t = (B_t^1, \ldots, B_t^d)$ is a so-called d-dimensional Brownian motion. For $(H_t)_{t\geq 0}$ and $(K_t)_{t\geq 0}$ adapted processes with respective values in $\mathbb{R}^{n\times d}$ and $\mathbb{R}^n$ such that $\forall t \geq 0$,

$\mathbb{P}\left(\int_0^t \|H_s\|^2 + \|K_s\| ds < +\infty\right) = 1$ and X_0 an $\mathbb{R}^n$-valued random vector independent of $(W_t)_{t\geq 0}$, then

$$X_t = X_0 + \int_0^t H_s dW_s + \int_0^t K_s ds$$

is an $\mathbb{R}^n$-valued Itô's process. The previous equality means

$$\forall t \geq 0,\ \forall i \in \{1, \dots, n\},\ X_t^i = X_0^i + \sum_{k=1}^{d} \int_0^t H_s^{ik} dB_s^k + \int_0^t K_s^i ds.$$

One can generalize Itô's formula to such multidimensional processes :

Theorem 3.9. *Let $f(t,x)$ be a $C^{1,2}$ function on $[0,+\infty[\times\mathbb{R}^n$. Then*

$$\begin{aligned}\forall t \geq 0,\ f(t,X_t) =& f(0,X_0) + \int_0^t \frac{\partial f}{\partial s}(s,X_s) ds + \int_0^t \nabla_x f(s,X_s) dX_s \\ &+ \frac{1}{2}\int_0^t \sum_{i,j=1}^{n} \left(\sum_{k=1}^{d} H_s^{ik} H_s^{jk} \frac{\partial^2 f}{\partial x_i \partial x_j}(s,X_s)\right) ds. \qquad (5)\end{aligned}$$

4 Stochastic differential equations and partial differential equations

4.1 Stochastic differential equations

The following theorem states existence and uniqueness for stochastic differential equations of the form

$$dX_t = \sigma(t,X_t) dW_t + b(t,X_t) dt. \qquad (6)$$

It assumes a Lipschitz regularity property on the coefficients σ and b. Like the Cauchy-Lipschitz theorem for ordinary differential equations, this result is obtained by a fixed-point approach.

Theorem 4.1. *Let $T > 0$, X_0 be an $\mathbb{R}^n$-valued random vector independent from the Brownian motion $W_t = (B_t^1, \dots, B_t^d)$ and $\sigma : [0,T]\times\mathbb{R}^n \to \mathbb{R}^{n\times d}$ and $b : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ denote coefficients such that*

$$\exists K > 0,\ \forall t \in [0,T] \begin{cases} \forall x \in \mathbb{R}^n,\ \|\sigma(t,x)\| + \|b(t,x)\| \leq K(1+\|x\|) \\ \forall x,y \in \mathbb{R}^n,\ \|\sigma(t,x)-\sigma(t,y)\| + \|b(t,x)-b(t,y)\| \leq K\|x-y\| \end{cases}. \qquad (7)$$

Then there exists a unique $\mathbb{R}^n$-valued Itô's process $(X_t)_{t\in[0,T]}$ such that

$$\forall t \in [0,T],\ X_t = X_0 + \int_0^t \sigma(s,X_s) dW_s + \int_0^t b(s,X_s) ds.$$

In addition, when $\mathbb{E}(\|X_0\|^2) < +\infty$, then $\mathbb{E}\left(\sup_{t\in[0,T]} \|X_t\|^2\right) < +\infty$.

Example 4.2. The unique solution of the stochastic differential equation

$$dX_t = dW_t - cX_t dt \quad \text{where} \quad c \in \mathbb{R}$$

is the Ornstein-Ulhenbeck process $X_t = e^{-ct}\left(X_0 + \int_0^t e^{cs} dW_s\right)$.

4.2 Link with parabolic partial differential equations

From now on, we assume that (7) holds. Let us introduce the infinitesimal generator $\mathcal{A}_t$ of the stochastic differential equation at time t that is the differential operator defined for smooth functions $\varphi : \mathbb{R}^n \to \mathbb{R}$ by

$$\mathcal{A}_t\varphi(x) = \frac{1}{2}\sum_{i,j=1}^n a_{ij}(t,x)\frac{\partial^2\varphi}{\partial x_i \partial x_j}(x) + \sum_{i=1}^n b_i(t,x)\frac{\partial \varphi}{\partial x_i}(x) \text{ where } a_{ij}(t,x) = \sum_{k=1}^d \sigma_{ik}\sigma_{jk}(t,x).$$

The following probabilistic representation of the solution of the partial differential equation

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) + \mathcal{A}_t u(t,x) + v(x)u(t,x) = 0, \ (t,x) \in [0,T]\times\mathbb{R}^n \\ u(T,x) = f(x), x \in \mathbb{R}^n \end{cases} \tag{8}$$

is known as the Feynman-Kac formula.

Proposition 4.3. *Assume that $v : \mathbb{R}^n \to \mathbb{R}$ is bounded from above and that u is a $C^{1,2}$ solution of* (8) *such that $\nabla_x u$ is bounded on $[0,T]\times\mathbb{R}^n$. Then for any $(t,x) \in [0,T]\times\mathbb{R}^n$,*

$$u(t,x) = \mathbb{E}\left(e^{\int_t^T v(X_s^{t,x})ds} f(X_T^{t,x})\right) \text{ where } \begin{cases} dX_s^{t,x} = \sigma(s,X_s^{t,x})dW_s + b(s,X_s^{t,x})ds, \ s\in[t,T] \\ X_t^{t,x} = x \end{cases}.$$

Proof : Appling Itô's formula (5), to the Itô's process $(X_s^{t,x}, \int_t^s v(X_r^{t,x})dr)_{t\le s\le T}$ with the function $f : (s,x,y) \in [t,T]\times\mathbb{R}^n\times\mathbb{R} \to e^y u(s,x)$, one has

$$\begin{aligned} e^{\int_t^T v(X_r^{t,x})dr}u(T,X_T^{t,x}) - u(t,X_t^{t,x}) = &\int_t^T e^{\int_t^s v(X_r^{t,x})dr}\left(\frac{\partial u}{\partial s} + \mathcal{A}_s u + vu\right)(s,X_s^{t,x})ds \\ &+ \int_t^T e^{\int_t^s v(X_r^{t,x})dr}\nabla_x u(s,X_s^{t,x}).\sigma(s,X_s^{t,x})dW_s \end{aligned}$$

Using (8), $X_t^{t,x} = x$ and taking expectations, one deduces that

$$\mathbb{E}\left(e^{\int_t^T v(X_r^{t,x})dr} f(X_T^{t,x})\right) - u(t,x) = \mathbb{E}\left(\int_t^T e^{\int_t^s v(X_r^{t,x})dr}\nabla_x u(s,X_s^{t,x}).\sigma(s,X_s^{t,x})dW_s\right).$$

Now using (7) and the hypotheses on functions u and v, one has

$$\mathbb{E}\left(\int_t^T e^{2\int_t^s v(X_r^{t,x})dr}\|\sigma^*(s,X_s^{t,x})\nabla_x u(s,X_s^{t,x})\|^2 ds\right) \le Ce^{2T\sup_{\mathbb{R}^n} v}\|\nabla_x u\|_\infty^2\left(1 + \sup_{s\in[t,T]}\mathbb{E}(\|X_s^{t,x}\|^2)\right).$$

The right-hand-side is finite according to the last assertion in Theorem 4.1. Therefore by (3), the expectation of the stochastic integral is 0. □

Remark 4.4. When σ and b and therefore $\mathcal{A}$ do not depend on the time variable t and u is a smooth solution of

$$\frac{\partial u}{\partial t}(t,x) = \mathcal{A}u(t,x) + v(x)u(t,x), \ (t,x)\in[0,T]\times\mathbb{R}^n \text{ with } u(0,x) = f(x), x\in\mathbb{R}^n$$

then $u(t,x) = \mathbb{E}\left(f(X_t^{0,x})e^{\int_0^t v(X_s^{0,x})ds}\right)$. This equality is proved by applying Itô's formula to $(X_s^{0,x}, \int_0^s v(X_r^{0,x})dr)_{s\in[0,t]}$ with $f(s,x,y) = e^y u(t-s,x)$ and taking expectations.

In case $d=n$, $\sigma=I_n$ where I_n denotes the identity matrix and $b=\frac{\nabla_x\psi_I}{\psi_I}$ where $\psi_I: \mathbb{R}^n\to\mathbb{R}$, one has

$$\frac{\partial u}{\partial t}=\frac{1}{2}\Delta_x u+\frac{\nabla_x\psi_I}{\psi_I}.\nabla_x u+vu.$$

As

$$\begin{aligned}\psi_I^2\left(\frac{1}{2}\Delta u+\frac{\nabla_x\psi_I}{\psi_I}.\nabla_x u\right)&=\frac{1}{2}\left(\psi_I^2\nabla_x.\nabla_x u+\nabla_x\psi_I^2.\nabla_x u\right)=\frac{1}{2}\nabla_x.(\psi_I^2\nabla_x u)\\&=\frac{1}{2}\Delta_x(\psi_I^2 u)-\nabla_x.(\psi_I\nabla_x\psi_I u),\end{aligned}$$

one has formally

$$\frac{\partial(\psi_I^2 u)}{\partial t}=\frac{1}{2}\Delta_x(\psi_I^2 u)-\nabla_x.\left(\frac{\nabla_x\psi_I}{\psi_I}(\psi_I^2 u)\right)+v(\psi_I^2 u).$$

Therefore $q(t,x)=\psi_I^2(x)\mathbb{E}\left(e^{\int_0^t v(X_s^{0,x})ds}\right)$ solves

$$\frac{\partial q}{\partial t}=\frac{1}{2}\Delta_x q-\nabla_x.\left(bq\right)+vq,\ (t,x)\in\mathbb{R}_+\times\mathbb{R}^n,\ q(0,x)=\psi_I^2(x),\ x\in\mathbb{R}^n.$$

This remark is the key to understand the Diffusion Monte-Carlo method in quantum chemistry.

Let us now introduce the adjoint $\mathcal{A}_t^*$ of the infinitesimal generator :

$$\mathcal{A}_t^*\varphi(x)=\frac{1}{2}\sum_{i,j=1}^n\frac{\partial^2}{\partial x_i\partial x_j}\left(a_{ij}(t,x)\varphi(x)\right)-\sum_{i=1}^n\frac{\partial}{\partial x_i}\left(b_i(t,x)\varphi(x)\right).$$

Proposition 4.5. *Assume that for any $t\in[0,T]$, the solution X_t of (6) has the density $p(t,x)$. If p, σ and b are respectively $C^{1,2}$, $C^{0,2}$ and $C^{0,1}$ on $[0,T]\times\mathbb{R}^n$, then p is a classical solution to the Fokker-Planck partial differential equation*

$$\frac{\partial p}{\partial t}(t,x)=\mathcal{A}_t^*p(t,x),\ (t,x)\in[0,T]\times\mathbb{R}^n. \tag{9}$$

Proof : Let $\varphi(t,x)$ denote a $C^{1,2}$ function with compact support on $[0,T]\times\mathbb{R}^n$. By Itô's formula (5), one has

$$\varphi(T,X_T)=\varphi(0,X_0)+\int_0^T\left(\frac{\partial\varphi}{\partial t}+\mathcal{A}_t\varphi\right)(t,X_t)dt+\int_0^T\nabla_x\varphi(t,X_t).\sigma(t,X_t)dW_t.$$

Since φ is compactly supported, $\mathbb{E}\left(\int_0^T\|\sigma^*(t,X_t)\nabla_x\varphi(t,X_t)\|^2dt\right)<+\infty$. Hence by (3), the expectation of the stochastic integral is zero. Therefore taking expectations, one has

$$\mathbb{E}(\varphi(T,X_T))=\mathbb{E}(\varphi(0,X_0))+\mathbb{E}\left[\int_0^T\left(\frac{\partial\varphi}{\partial t}+\mathcal{A}_t\varphi\right)(t,X_t)dt\right].$$

Since for any $t \in [0,T]$, X_t has the density $p(t,x)$, one deduces

$$\int_{\mathbb{R}^n} \left(\varphi(T,x)p(T,x) - \varphi(0,x)p(0,x) - \int_0^T \frac{\partial \varphi}{\partial t}(t,x)p(t,x)dt \right) dx$$
$$= \int_0^T \left(\int_{\mathbb{R}^n} \mathcal{A}_t \varphi(t,x)p(t,x)dx \right) dt. \qquad (10)$$

One makes integration by parts with respect to the time variable in the left-hand-side and to the spatial variables in the right-hand-side to deduce

$$\int_{[0,T]\times\mathbb{R}^n} \varphi(t,x)\frac{\partial p}{\partial t}(t,x)dtdx = \int_{[0,T]\times\mathbb{R}^n} \varphi(t,x)\mathcal{A}_t^* p(t,x)dtdx.$$

One concludes since φ is arbitrary and both sides of (9) are continuous functions. $\square$

Remark 4.6. If for any $t \in [0,T]$, X_t has the density $p(t,x)$ but the functions p, σ and b do not meet the regularity assumptions in Proposition 4.5, (10) still means that p is a weak solution to (9).
In fact, to obtain a weak solution it is not even necessary to assume the existence of densities. Indeed, denoting by $P_t(dx)$ the probability law of X_t ($\forall f : \mathbb{R}^n \to \mathbb{R}$ bounded, $\mathbb{E}(f(X_t)) = \int_{\mathbb{R}^n} f(x)P_t(dx)$), the reasoning made to obtain (10) always ensures that $P_t(dx)dt$ is a weak solution to (9).

5 Bibliography

For a simple introduction to stochastic integration with respect to the Brownian motion we refer to [2] in french and [3] in english. The books [1], [5] and [6] present more advanced properties and deal with stochastic integration with respect to continuous semi-martingales. A notion of solutions to stochastic differential equations weaker than the one presented above is developped in [7]. Last, [4] deals with the very general theory of stochastic integrals with respect to possibly discontinuous semi-martingales.

References

[1] Karatzas, I. and Shreve, S.E. Brownian motion and stochastic calculus. Second edition. *Graduate Texts in Mathematics, 113. Springer-Verlag, New York*, 1997.

[2] Lamberton, D. and Lapeyre, B. Introduction au calcul stochastique appliqué à la finance. Second edition.

[3] Lamberton, D. and Lapeyre, B. Introduction to stochastic calculus applied to finance. *Chapman & Hall*, 1996.

[4] Protter, P.E. Stochastic integration and differential equations. Second edition. *Applications of Mathematics (New York), 21. Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin*, 2004.

[5] Revuz, D. and Yor, M. Continuous martingales and Brownian motion. Third edition. *Grundlehren der Mathematischen Wissenschaften, 293. Springer-Verlag, Berlin*, 1999.

[6] Rogers, L.C.G. and Williams, D. Diffusions, Markov processes, and martingales. Vol. 1 and 2. Second edition. *John Wiley & Sons*, 1994.

[7] Stroock, D.W. and Varadhan, S.R.S. Multidimensional diffusion processes *Grundlehren der Mathematischen Wissenschaften, 233. Springer-Verlag, Berlin*, 1979.