A REVIEW LECTURE ON STOCHASTIC CALCULUS

Marco DOZZI and Bijan ZANGENEH

Départment de Mathématiques

Université de Nancy I

B.P. 239

F-54506 Vandoeuvre-les-Nancy

France

marco.dozzi@iecn.u-nancy.fr

and

Sharif University of Technology

Department of Mathematical Sciences

P.O. Box 11365-9415

Tehran, Iran

I. MOTIVATION

Let us start with some examples, which will be studied as illustrations in more detail in this text.

1) Population growth : Let N_t be the size of a population at time t. A very simple model for the evolution of a population is given by the differential equation $\frac{d}{dt}N_t = aN_t$ $(a \in \mathbb{R})$. Now this evolution may be influenced by factors which are not or not completely under control. The analytic form of the function which describes these factors is therefore not known. In such a case the evolution can be described by means of a stochastic process (or a random function) which is the solution of a stochastic differential equation, i.e.

a differential equation (e.g. $\frac{d}{dt}N_t = aN_t$) where a random term (e.g. a white noise $b\ \dot{B}_t$) is added. Since these random functions often exist in the sense of distributions only, the differential equation is usually written in integral form as follows :

$$N_t = N_0 + \int_0^t aN_s ds + \int_0^t bN_s\ \dot{B}_s\ ds, \tag{1}$$

where the last integral is understood as the stochastic integral $\int_0^t bN_s dB_s$. We will see that the solution of (1) is a continuous, but not differentiable random function. This is due to the fact that the trajectories of B are not of finite variation on compact (time) intervals, and the integral can therefore not be defined as a Stieltjes integral in general.

2) Portfolio management : Suppose that the capital at time t of some company is divided into a safe and a risky part, the latter being invested at a higher interest rate than the former. Again the evolution of the capital in time can be described by differential e quations, the equation for the risky part containing a stochastic term. The company would like to maximize the expected value of the capital (or some increasing functional of it) at some fixed future time t_1. The problem is now to optimize at each time t the proportion of the capital under risk. As in the deterministic case this kind of problem can be solved by means of Hamilton-Jacobi equations which have been extended to the stochastic case by R. Bellmann. The operators appearing in these equations are now of second order since the quadratic variation associated to the stochastic integrals which describe the randomness do not vanish.

3) Direct current engine : Consider an electric engine driven by a input tension which can be modified at each time t. As a simple model let us consider the case where the state of engine is described by the angle of a point on its turning axis and by the velocity by which this angle changes. The state of the engine is then described by means of a system of differential equations, as a function of the input tension. Again a stochastic term can be added if there is some perturbation from outside which influences the state of the engine. The question is now how to choose, at each time t, the input tension for an optimal working of the engine.

II. BROWNIAN MOTION

II.1. Definition of Brownian motion

Let us first give two equivalent definitions of Brownian motion.

Definition 1. A Brownian motion $(B_t\,;\, t \geq 0)$ with $B_0 = 0$ P-a.s. is a stochastic process such that, for all $n \in \mathbb{N}$ and all $t_1, \ldots, t_n \geq 0$ with $0 = t_0 < t_1 < \cdots < t_n$, the increments $B_{t_k} - B_{t_{k-1}}$ $(k = 1, \ldots, n)$ are independent and normally distributed with $EB_{t_k} = 0$ and $\operatorname{var}(B_{t_k} - B_{t_{k-1}}) = t_k - t_{k-1}$.

Definition 2. A Brownian motion $(B_t\,;\, t \geq 0)$ with $B_0 = 0$ P-a.s. is a Gaussian process with $EB_t = 0$ and $EB_sB_t = \min\{s, t\}$ for all $s, t \geq 0$.

It is easily shown that a Brownian motion in the sense of Definition 2 has independent and normally distributed increments. It is sufficient to show that the increments (being Gaussian) are uncorrelated and have the variance required in Definition 1. In order to show the converse, let $Z_k = B_{t_k} - B_{t_{k-1}}$, $k = 1, \ldots, n$. Let $B = (B_{t_1}, \ldots, B_{t_n})$, $Z = (Z_1, \ldots, Z_n)$ and let

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}$$

Then $B = AZ$. Let φ_B be the characteristic function of B. Then, with $u = (u_1, \ldots, u_n)^T \in \mathbb{R}^n$, $\varphi_B(u) = E[\exp(iu^TB)] = E[\exp(iu^TAZ)] = \varphi_Z(u^TA) = \exp\Big(-\dfrac{1}{2}u^TCu\Big)$, where $C = ADA$ and $D = \begin{pmatrix} t_1 & 0 & \cdots & 0 \\ 0 & t_2 - t_1 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & & t_n - t_{n-1} \end{pmatrix}$.

C is symmetric and positive definite, and it is the covariance of B. Moreover c_{ij} (the (i, j) th element of C) is given by

$$c_{ij} = E[B_{t_i}B_{t_j}] = \begin{cases} E[(B_{t_j} - B_{t_i})B_{t_i}] + EB_{t_i}^2 = t_i \text{ if } t_i \leq t_j, \\ E[(B_{t_i} - B_{t_j})B_{t_j}] + EB_{t_j}^2 = t_j \text{ if } t_j \leq t_i. \end{cases}$$

Therefore $c_{ij} = \min\{t_i, t_j\}$ as is required in Definition 2.

Let us now give a construction of Brownian motion by means of the Haar basis in $L^2([0,1])$. More precisely let us show

Theorem 1. Let $\{X_n\}_{n\geq 0}$ be a sequence of independent random variables each normally distributed with mean 0 and variance 1. Then, with probability 1, the series $\sum_{n=0}^{\infty} X_n \int_0^t H_n(s)ds$ converges uniformly in $t \in [0,1]$ and defines a Brownian motion $(B_t\,;\, 0 \leq t \leq 1)$.

Remark : In order to define a Brownian motion on $[0,\infty[$, one can choose a countable number of independent sequences $\{X_n^m\}_{n\geq 0}$ $(m \in \mathbb{N})$ as in Theorem 1 and construct Brownian motions $(B_t^m\,;\, 0 \leq t \leq 1)$ $(m \in \mathbb{N})$, which can be pieced together as follows : if $m-1 \leq t < m$, define B_t inductively by

$$B_t = B^m_{t-(m-1)} + B_{m-1}.$$

Proof of Theorem 1. Let $H_k(t)$ $(k \geq 0, 0 \leq t \leq 1)$ be the Haar functions, i.e.

$$H_k(t) = \left\{ \begin{array}{ll} 2^{n/2} & \text{if } k2^{-n} - 1 \leq t \leq k2^{-n} - 1 + 2^{-(n+1)} \\ -2^{n/2} & \text{if } k2^{-n} - 1 + 2^{-(n+1)} \leq t \leq k2^{-n} - 1 + 2^{-n} \\ 0 & \text{otherwise.} \end{array} \right\} \text{ and } 2^n \leq k < 2^{n+1}$$

a) $\sum_{k=0}^{\infty} a_k \int_0^t H_k(s)ds$ converges uniformly in $t \in [0,1]$ provided that $|a_k| = 0(k^{\varepsilon})$ for some $\varepsilon \in \left]0, \frac{1}{2}\right[$ as $k \to \infty$.

In fact, the functions $\int_0^t H_k(s)ds$ are nonnegative and attain a maximum value of $\frac{1}{4}2^{-n/2}$ provided that $2^n \leq k < 2^{n+1}$. Furthermore, as k varies, these functions have disjoint support. Writing $b_n = \max\{|a_k|\,;\, 2^n \leq k < 2^{n+1}\}$, we have $|b_n| < C2^{\varepsilon n}$ for some constant $C > 0$ as $n \to \infty$, and, therefore, $\sum_{k=0}^{\infty} a_k \int_0^t H_k(s)ds \leq \sum_{n=0}^{\infty} b_n 2^{-n/2} < \infty$.

b) $P\big(|X_n| = O(\sqrt{\log n})\big) = 1$.

In fact, for $x > 0$, $P(|x_n| > x) \sim \sqrt{\frac{2}{\pi}}\frac{1}{x}\exp\Big(-\frac{x^2}{2}\Big)$ as $x \to +\infty$. Therefore,

$\sum_{n=2}^{\infty} P(|X_n| \geq c\sqrt{\log n}) \leq K \sum_{n=2}^{\infty} (\log n)^{-1/2} n^{-c^2/2}$ for some constant K. The assertion follows by choosing $c > \sqrt{2}$ and by the lemma of Borel-Cantelli.

c) Let us now show that the series in Theorem 1 is a Brownian motion. By a) and b) above it converges P-a.s. and uniformly in $t \in [0,1]$. The sum is therefore a continuous function of t. In order to show that the increments are normally distributed we calculate their characteristic function. Let $X_t = \sum_{k=0}^{\infty} X_n \int_0^t H_k(s)ds$ and let $S_k(t) = \int_0^t H_k(s)ds$. Then, for $0 = t_0 < t_1 < t_2$ and $\lambda_1, \lambda_2 \in \mathbb{R}$,

$$\mathrm{E}\left[\exp\left(i\lambda_1 X_{t_1} + i\lambda_2 (X_{t_2} - X_{t_1})\right)\right] = \mathrm{E}\left[\exp\left(\sum_{k=0}^{\infty} X_k \{i(\lambda_1 - \lambda_2) S_k(t_1) + i\lambda_2 S_k(t_2)\}\right)\right].$$

Since the X_k are independent, this equals to

$$\prod_{k=0}^{\infty} \exp\left(-\frac{1}{2}\{(\lambda_1 - \lambda_2)S_k(t_1) + \lambda_2 S_k(t_2)\}^2\right). \qquad (1)$$

Let now $\varphi_s(t) = 1_{[0,s]}(t)$. By Parseval's equality (the Haar functions are a complete orthonormal system in $L^2[0,1]$), we get for $s_1 < s_2$:

$$s_1 = (\varphi_{s_1}, \varphi_{s_2}) \equiv \int_0^1 \varphi_{s_1}(t)\varphi_{s_2}(t)dt = \sum_{k=0}^{\infty} (s_1, H_k)(s_2, H_k) = \sum_{k=0}^{\infty} S_k(s_1) S_k(s_2).$$

Therefore (1) is equal to

$$\exp\left(-\frac{1}{2}\{(\lambda_1 - \lambda_2)^2 t_1 + 2(\lambda_1 - \lambda_2)\lambda_2 t_1 + \lambda_2^2 t_2\}\right) = \exp\left(-\frac{1}{2}\{\lambda_1^2 t_1 + \lambda_2^2 (t_2 - t_1)\}\right).$$

This shows that $(X_{t_1}, X_{t_2} - X_{t_1})$ satisfies the Definition 1. The proof for $(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_k} - X_{t_{k-1}})$ is similar. □

Remark : The proof shows that the trajectories of Brownian motion are continuous P-a.s. They are, however, not differentiable at any point in $\mathbb{R}_+$ P-a.s. As we will see in the next chapter, the total variation of the trajectories on any interval is infinite P-a.s. The variation of the trajectories can be characterized by their modulus of continuity : P. Lvy showed that P-a.s.

$$\limsup_{\delta \to 0} \left(2\delta \log \frac{1}{\delta}\right)^{-1/2} \max\{|B_t - B_s| \,;\, 0 \leq s < t \leq 1, t - s < \delta\} = 1. \qquad (2)$$

II.2. The Hausdorff dimension of the trajectories

The fact that the total variation of trajectories of Brownian motion is infinite on any interval gives raise to the question of the dimension of its graph. Let us consider its Hausdorff dimension more generally in $\mathbb{R}^n$.

Let $A \subset \mathbb{R}^n$. The α-dimensional outer measure $S_\alpha(A)$ is defined for any $\alpha > 0$ as follows. Consider the coverings of A by means of open balls $B_i \subset \mathbb{R}^n$ such that the diameter of the B_i's is not greater than ε. Let $S_{\alpha,\varepsilon}(A)$ be defined by $S_{\alpha,\varepsilon}(A) = \inf \sum_i$ (diameter $B_i)^\alpha$, where the infimum is taken over all the coverings of A with the properties mentionel above. As ε decreases, the infimum will extend over smaller and smaller sets and is therefore not decreasing. So there exists $S_\alpha(A) := \lim_{\varepsilon\to 0} S_{\alpha,\varepsilon}(A) \in [0,\infty]$. Clearly, S_α is monotone in the sense that $S_\alpha(A) \leq S_\alpha(B)$ if $A \subset B$. Furthermore S_α is subadditive, i.e. for any sequence of sets $A_1, A_2, \dots, A_n, \dots$ in $\mathbb{R}^n$ we have $S_\alpha\Big(\cup_n A_n\Big) \leq \sum_n S_\alpha(A_n)$.

Lemma. For any set $A \subset \mathbb{R}^n$ there exists a unique number α^*, such that

$$\alpha < \alpha^* \Rightarrow S_\alpha(A) = \infty \; , \; \alpha > \alpha^* \Rightarrow S_\alpha(A) = 0. \tag{3}$$

Proof. Let us show that

$$S_\alpha(A) < \infty \; , \; \beta > \alpha \Rightarrow S_\beta(A) = 0.$$

In fact, let $\{B_i\}$ be a covering of A by balls with diameter not greater than ε, such that

$$\sum_i (\text{diameter}\, B_i)^\alpha \leq S_{\alpha,\varepsilon}(A) + 1 \leq S_\alpha(A) + 1 = K < \infty.$$

Then

$$S_{\beta,\varepsilon}(A) \leq \sum_i (\text{diameter}\, B_i)^\beta \leq \varepsilon^{\beta-\alpha} \sum_i (\text{diameter}\, B_i)^\alpha \leq \varepsilon^{\beta-\alpha} K.$$

Since $\beta > \alpha$ we have, by letting $\varepsilon \to 0$, $S_\beta(A) = 0$. Therefore, there exists a transition point α^* which satisfies (2). $\square$

Definition. The Hausdorff dimension of $A \subset \mathbb{R}^n$ is the number α^* such that

$$\alpha^* = \sup\{\alpha\,;\, S_\alpha(A) = \infty\} = \inf\{\alpha\,;\, S_\alpha(A) = 0\}.$$

Let $G(B_t\,;\, 0 \le t \le 1) = \{\big(t, B_t(\omega)\big)\,;\, 0 \le t \le 1\}$ be the graph of $B_t(\omega), 0 \le t \le 1$. Then one can show that the Hausdorff dimension of $G(B_t\,;\, 0 \le t \le 1) = \dfrac{3}{2}$ (see R. Adler (1981)).

Theorem 2. The Hausdorff dimension of $G(B_t; 0 \le t \le 1)$ is equal to $\frac{3}{2}$.

Proof. a) Let us first show that it is $\le \frac{3}{2}$. Split $[0,1]$ into 2^n intervals of length 2^{-n}. By (2) the graph is contained in the rectangles with edge length 2^{-n} and $2^{-n\beta}$ for all $\beta > \frac{1}{2}$. These rectangles can also be covered by $2^{n(1-\beta)}$ squares with edge length 2^{-n}. By the definition of the Hausdorff dimension of the graph, it is no more than the infinum of all λ such that

$$2^n 2^{n(1-\beta)} 2^{-\lambda n} \to 0 \qquad \text{as } n \to \infty.$$

This means $\lambda < \frac{3}{2}$ and the assertion follows.

b) It remains to show that $\dim G = \dim G(B_t; 0 \le t \le 1) \ge \frac{3}{2}$. Let us first show that

$$\int_\eta^1 \int_\eta^1 E\big[|B_t - B_s|^2 + |t-s|^2\big]^{-\frac{\alpha}{2}} ds\, dt < \infty \tag{4}$$

for all $\eta > 0$ and all $1 < \alpha < \frac{3}{2}$. In fact, $E[\cdots]^{-\frac{\alpha}{2}}$ transforms to

$(2\pi)^{-1/2} \int_{-\infty}^\infty \big[u^2|t-s| + |t-s|^2\big]^{-\alpha/2} \exp\big(-\frac{u^2}{2}\big) du \le 2(2\pi)^{-1/2}|t-s|^{-\alpha/2}$
$\int_0^\infty \big(u^2 + |t-s|\big)^{-\alpha/2} du \le c|t-s|^{-\alpha+1/2}.$

Therefore the left side of (4) is finite whenever $-\alpha + \frac{1}{2} < -1$ i.e. $\alpha < \frac{3}{2}$. Let us deduce now the assertion from (4). Let μ the image measure of Lebesgue measure under the transformation $t \to x = (t, B_t)$, $0 \le t \le 1$. Let $\delta > 0$ such that $\gamma := \frac{3}{2} - 2\delta > 1$. Then

$$\int_G \int_G |x-y|^{-(\gamma+\delta)} dx\, dy < \infty.$$

Therefore there exists a closed set $G_1 \subset G$ with positive μ-measure such that

$$\int_{G_1} |x-y|^{-(\gamma+\delta)} dy \le M$$

for some $M > 0$ and all $x \in G_1$. Let $\{\epsilon_n\}_{n\in\mathbb{N}}$ be a sequence decreasing to zero, and, for each n, let $S_{n1}, \cdots, S_{nm_n}$ be open spheres with diameters at most ϵ_n which cover G_1. Choose a point X_{ni} from each $S_{ni} \cap G_1$. Then

$$| \text{ diameter } S_{ni}|^{-(\gamma+\delta)}\mu(S_{ni}) \leq \int_{G_1} |X_{ni} - y|^{-(\gamma+\delta)} dy \leq M.$$

Thus

$$0 \leq \mu(G_1) \leq \sum_{i=1}^{m_n} \mu(S_{ni}) \leq M \sum_{i=1}^{m_n} |\text{diameter } S_{ni}|^{\gamma+\delta},$$

and therefore

$$\sum_{i=1}^{m_n} |\text{diameter } S_{ni}|^{\gamma} \geq \epsilon_n^{-\delta} \sum_{i=1}^{m_n} |\text{diameter} S_{ni}|^{\gamma+\delta} \geq \frac{1}{M}\epsilon_n^{-\delta}\mu(G_1)$$

which diverges as $n \to \infty$. Hence $\dim(G_1) \geq \gamma = \frac{3}{2} - 2\delta$. Therefore $\dim(G) \geq \dim(G_1) \geq \frac{3}{2}$.

III. MARTINGALES

III.1. Conditional expectations

Let $(\Omega, \mathcal{F}, P)$ be a probability space. We first consider conditional expectations with respect to a σ-field $\mathcal{B} \subset \mathcal{F}$ which is generated by a finite partition of Ω. Let $B_1, \ldots, B_n \in \mathcal{F}$ such that $\cup_{i=1}^n B_i = \Omega$ and $B_i \cap B_j = \emptyset$ if $i \neq j$ and let $\mathcal{B} = \sigma(B_1, \ldots, B_n)$. Let X be a random variable on $(\Omega, \mathcal{F}, P)$ taking the values $x_1, x_2, \ldots, x_j, \ldots$ Then

$$P(X = x_j | B_i) = P(B_i \cap \{X = x_j\})/P(B_i).$$

For $x_j \neq 0$, $B_i \cap \{X = x_j\} = \{I_{B_i} X = x_j\}$, where I_{B_i} is the indicator function of B_i. Therefore

$$\begin{aligned} &\mathrm{E}[X \mid B_i] = \textstyle\sum_{x_j \neq 0} x_j P(X = x_j \mid B_i) \\ &= \tfrac{1}{P(B_i)} \textstyle\sum_{x_j \neq 0} x_j P(I_{B_i} X = x_j) = \mathrm{E}[X I_{B_i}]/P(B_i). \end{aligned} \tag{1}$$

We define now the conditional expectation of X given $\mathcal{B}$ as the random variable which takes the value $\mathrm{E}[X \mid B_i]$ on B_i i.e.

$$Y = \mathrm{E}[X \mid \mathcal{B}] \Leftrightarrow Y(\omega) = \mathrm{E}[X \mid B_i] \text{ if } \omega \in B_i.$$

Since YI_{B_i} takes only the values 0 and 1, the latter with probability $P(B_i)$, we have $\mathrm{E}[YI_{B_i}] = P(B_i)\,\mathrm{E}[X \mid B_i]$. By (1) and the linearity of the expected value, we have, for any linear combination $Z = \sum_{i=1}^{n} \lambda_i I_{B_i}$, $\mathrm{E}[YZ] = \mathrm{E}[XZ]$. This leads us to the following general definition of conditional expectation :

Definition. Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\mathcal{B} \subset \mathcal{F}$ be a sub-σ-field of $\mathcal{F}$. Let X be a random variable on $(\Omega, \mathcal{F}, P)$ with $\mathrm{E}\,|X| < \infty$. A random variable Y is called the conditional expectation of X given $\mathcal{B}$ if (i) Y is $\mathcal{B}$-measurable, (ii) $\mathrm{E}[ZY] = \mathrm{E}[ZX]$ for all $\mathcal{B}$-measurable and bounded random variables Z.

Existence and uniqueness of $\mathrm{E}[X \mid \mathcal{B}]$ follows by a measure-theoretic result as follows: Let $\nu(A) = \int_A X dP$. Then ν is a measure on $\mathcal{B}$, which is absolutely continuous with respect to P. By the theorem of Radon-Nikodym there exists a random variable Y which is $\mathcal{B}$-**measurable** and determined uniquely up to negligeable sets such that $\nu(A) = \int_A Y dP$ for all $A \in \mathcal{B}$. We set $Y = \mathrm{E}[X \mid \mathcal{B}]$.

Let us mention now some important **properties** of conditional expectation :

a) Linearity : $\mathrm{E}[\lambda_1 X_1 + \lambda_2 X_2 \mid \mathcal{B}] = \lambda_1 \mathrm{E}[X_1 \mid \mathcal{B}] + \lambda_2 \mathrm{E}[X_2 \mid \mathcal{B}]$ for all $\lambda_1, \lambda_2 \in \mathbb{R}$.

b) Projection property : Let $\mathcal{C}$ be a sub-σ-field of $\mathcal{B}$. Then $\mathrm{E}[X \mid \mathcal{C}] = \mathrm{E}[\mathrm{E}[X \mid \mathcal{B}] \mid \mathcal{C}]$.

c) W is a $\mathcal{B}$-measurable random variable $\Rightarrow \mathrm{E}[XW \mid \mathcal{B}] = W\,\mathrm{E}[X \mid \mathcal{B}]$.

d) $\varphi : \mathbb{R} \to \mathbb{R}$ convex $\Rightarrow \mathrm{E}[\varphi(X) \mid \mathcal{B}] \geq \varphi(\mathrm{E}[X \mid \mathcal{B}])$ (Inequality of Jensen).

For the proofs of a)-c), it suffices to verify conditions (i) and (ii) of the definition :

a) $Y_i = \mathrm{E}[X_i \mid \mathcal{B}]$ $(i = 1, 2)$ is $\mathcal{B}$-measurable by definition ; therefore $\lambda_1 Y_1 + \lambda_2 Y_2$ is $\mathcal{B}$-measurable. Moreover, for all $\mathcal{B}$-measurable and bounded random variables Z,

$$\mathrm{E}[(\lambda_1 Y_1 + \lambda_2 Y_2)Z] = \lambda_1 \mathrm{E}[Y_1 Z] + \lambda_2 \mathrm{E}[Y_2 Z] = \lambda_1 \mathrm{E}[X_1 Z] + \lambda_2 \mathrm{E}[X_2 Z] = \mathrm{E}[(\lambda_1 X_1 + \lambda_2 X_2)Z].$$

b) Let us give a geometrical argumentation. Let $Y = \mathrm{E}[X \mid \mathcal{B}]$. Then, by (ii) in the definition above, $\mathrm{E}[(X - Y)Z] = 0$ and therefore $X - Y$ is orthogonal to Z. Let U (resp. V) be the vector space of all $\mathcal{B}$-(resp. $\mathcal{C}$-)mesurable random variables. Then $U \supset V$. Since $X - Y$ is orthogonal to U, it is also orthogonal to V. Therefore $X - \mathrm{E}[Y \mid \mathcal{C}] = X - Y + Y - \mathrm{E}[Y \mid \mathcal{C}]$ is orthogonal to V, which shows that $\mathrm{E}\left[(X - \mathrm{E}[Y \mid \mathcal{C}])Z\right] = 0$ for all bounded and $\mathcal{C}$-measurable random variables Z.

c) $W\,\mathrm{E}[X \mid \mathcal{B}]$ is $\mathcal{B}$-measurable. For all bounded and $\mathcal{B}$-measurable random variables Z we have

$$\mathrm{E}\left[W\,\mathrm{E}[X \mid \mathcal{B}]Z\right] = \mathrm{E}\left[\,\mathrm{E}[X \mid \mathcal{B}]WZ\right] = \mathrm{E}[X.WZ] = \mathrm{E}[WX.Z].$$

d) Since φ is convex

$$\varphi(X) - \varphi(\mathrm{E}[X \mid \mathcal{B}] \geq \varphi'(X)(X - \mathrm{E}[X \mid \mathcal{B}]) \geq 0.$$

The assertion follows by taking conditional expectations with respect to $\mathcal{B}$.

Example : Urn of Polya.

Suppose we have an urn with R_0 red and B_0 black balls. One ball is chosen at random, and $a \in \{-1\} \cup \mathbb{N}_0$ balls of the same colour as the chosen ball are put back into the urn. Notice that $a = -1$ (resp. $a = 0$) corresponds to choosing a ball without (resp. with) replacing it into the urn. After having chosen n balls at random there are R_n red and B_n black balls in the urn, where $R_n + B_n = R_0 + B_0 + na$. Let $X_n = R_n(R_n + B_n)^{-1}$ be the proportion of red balls and let $\mathcal{F}_n = \sigma(R_0, B_0, R_1, \ldots, R_n)$ be the σ-field generated by $R_0, B_0, R_1, \ldots, R_n$. Since $B_n = R_0 + B_0 + na - R_n$, X_n is $\mathcal{F}_n$-measurable. Moreover, for $r_0 + b_0 + (n+1)a > 0$,

$$\begin{aligned}
&\mathrm{E}[X_{n+1} \mid R_n = r, R_0 = r_0, B_0 = b_0] \\
&= \tfrac{r}{r_0+b_0+na} \cdot \tfrac{r+a}{r_0+b_0+(n+1)a} + \tfrac{r_0+b_0+na-r}{r_0+b_0+na} \cdot \tfrac{r}{r_0+b_0+(n+1)a} \\
&= \tfrac{r}{r_0+b_0+na} = X_n.
\end{aligned}$$

Therefore $\mathrm{E}[X_{n+1} \mid \mathcal{F}_n] = X_n$. This is the so-called martingale property which will be considered in more detail in the next paragraph.

III.2. Maximal inequalities for martingales

Definition. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_t\,;\, t \geq 0)$ an increasing family of sub-σ-fields of $\mathcal{F}$. A stochastic process $(X_t\,;\, t \geq 0)$ such that X_t is $\mathcal{F}_t$-measurable and $\mathrm{E}\,|X_t| < \infty$ for all $t \geq 0$ is called a) a ***martingale***, if $\mathrm{E}[X_t \mid \mathcal{F}_s] = X_s$ for all $s \leq t$, b) a ***submartingale***, if $\mathrm{E}[X_t \mid \mathcal{F}_s] \geq X_s$ for all $s \leq t$, c) a ***supermartingale***, if $\mathrm{E}[X_t \mid \mathcal{F}_s] \leq X_s$ for all $s \leq t$.

Remarks : Let $(X_t\,;\, t \geq 0)$ be a martingale, and let $\varphi : \mathbb{R} \to \mathbb{R}$ be convex (for example : $\varphi(x) = |x|^\alpha$ $(\alpha \geq 1)$, $\varphi(x) = x \vee 0$, $\varphi(x) = -(x \wedge 0)$) ; then $Y_t = \varphi(X_t)$ is a submartingale. If $(X_t\,;\, t \geq 0)$ is a submartingale, and if φ is convex and non-decreasing, then $\varphi(X_t)$ is a submartingale ; in fact :

$$\mathrm{E}[\varphi(X_t) \mid \mathcal{F}_s] \underset{\text{III.1.d)}}{\geq} \varphi(\mathrm{E}[X_t \mid \mathcal{F}_s]) \underset{(\geq)}{=} \varphi(X_s) \text{ for all } s \leq t.$$

Definition. A ***stopping time*** T is a mapping from Ω to $[0, \infty]$ such that $T < \infty$ P-a.s. and $\{\omega \in \Omega\,;\, T(\omega) \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$. The σ-field associated to T is defined by $\mathcal{F}_T = \big\{A \in \mathcal{F}\,;\, A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0\big\}$.

Let us now give a maximal inequality for martingales.

Theorem 1. Let $(M_t\,;\, t \geq 0)$ be a martingale. Let $[r, s]$ be a compact interval in $\mathbb{R}_+$ and let S be a countable dense subset of $\mathbb{R}_+$. Then, for any $\lambda > 0$,

$$P(\sup\{|M_t|\,;\, t \in S \cap [r, s]\} \geq \lambda) \leq \frac{1}{\lambda}\,\mathrm{E}\left[|M_s \mid I_{\left\{\sup\{|M_t|\,;\, t \in [r,s]\} > \lambda\right\}}\right].$$

Remark : If $(M_t\,;\, t \geq 0)$ is continuous (or right-continuous only), the above inequality implies that $P\big(\sup_{[r,s]} |M_t| \geq \lambda\big) \leq \frac{1}{\lambda}\,\mathrm{E}\,|M_s|$. The proof below shows that (1) remains true for any positive submartingale.

Proof. Let us first show the inequality for a finite parameter set : Let $\{M_k\}_{1 \leq k \leq n}$ be a martingale with respect to $\{\mathcal{F}_k\}_{1 \leq k \leq n}$ and let T be the stopping time given by $T(\omega) = \min\{k \geq 1\,;\, M_k(\omega) \geq \lambda\}$. Then $T \wedge n$ is a bounded stopping time and therefore

$$\begin{aligned}\mathrm{E}\, M_{T\wedge n} &= \sum_{k=1}^n \mathrm{E}[M_k I_{\{T\wedge n=k\}}] = \sum_{k=1}^n \mathrm{E}\left[\, \mathrm{E}[M_n \mid \mathcal{F}_k] I_{\{T\wedge n=k\}}\right] = \sum_{k=1}^n \mathrm{E}[M_n I_{\{T\wedge n=k\}}] \\ &= \mathrm{E}\, M_n = \mathrm{E}\, M_0.\end{aligned}$$

Now

$$\mathrm{E}[M_{T\wedge n}] = \mathrm{E}[M_T I_{\{T\le n\}} + M_n I_{\{T>n\}}].$$

However

$$\mathrm{E}[M_T I_{\{T\le n\}}] \ge \lambda\, \mathrm{E}[I_{\{T\le n\}}] = \lambda P(T\le n)$$

and therefore

$$\mathrm{E}\, M_n = \mathrm{E}\, M_{T\wedge n} \ge \lambda P(T\le n) + \mathrm{E}[M_n I_{\{T>n\}}]$$

whence

$$P(T\le n) \le \frac{1}{\lambda}\, \mathrm{E}[M_n I_{\{T\le n\}}] \le \frac{1}{\lambda}\, \mathrm{E}[M_n^+ I_{\left\{\,\sup\{M_k\,;\,1\le k\le n\}\ge\lambda\right\}}]. \tag{2}$$

By the same development, applied to the martingale $\{-M_k\}_{1\le k\le n}$, we get

$$P(\min\{k\ge 1\,;\, M_k\le -\lambda\}\le n) \le \frac{1}{\lambda}\, \mathrm{E}[M_n^- I_{\left\{\,\inf\{M_k\,;\,1\le k\le n\}\le-\lambda\right\}}]. \tag{3}$$

These two inequalities imply

$$P(\max\{|M_k|\,;\,1\le k\le n\}\ge\lambda) \le \frac{1}{\lambda}\, \mathrm{E}[|M_n| I_{\{\sup|M_k|\ge\lambda\}}].$$

The assertion follows by choosing an increasing sequence of finite subsets of $S\cap[r,s]$, the union of which is equal to $S\cap[r,s]$. □

Corollary. Suppose that there exists $p>1$ such that $\mathrm{E}\,|M_s|^p<\infty$. Then

$$\mathrm{E}\left[\,\sup\{|M_t|^p\,;\, t\in S\cap[r,s]\}\right] \le \Big(\frac{p}{p-1}\Big)^p\, \mathrm{E}[|M_s|^p]. \tag{4}$$

Remark : As for Theorem 1, (4) implies that

$$\mathrm{E}\left[\,\sup_{[r,s]}|M_t|^p\right] \le \Big(\frac{p}{p-1}\Big)^p\, \mathrm{E}[|M_s|^p]$$

if $(M_t\,;\, t\ge 0)$ is right-continuous.

Proof. Let $Y = \sup\{|M_t|\,;\, t \in S \cap [r,s]\}$. Then, by Theorem 1,

$$\begin{aligned} \mathrm{E}\,Y^p &= \int_0^\infty P(Y^p > y)dy = \int_0^\infty P(Y > y^{1/p})dy = \int_0^\infty P(Y > z)pz^{p-1}dz \\ &\le p\int_0^\infty z^{p-2}\int_{\{Y>z\}} X_s dP = p\int_\Omega X_s\Big[\int_0^{Y(\omega)} z^{p-2}dz\Big]dP \\ &= \tfrac{p}{p-1}\int_\Omega X_s Y^{p-1}dP. \end{aligned} \tag{5}$$

By Hölder's inequality ($\mathrm{E}\,|XZ| \le \|X\|_\alpha \|Z\|_\beta$ if $\dfrac{1}{\alpha}+\dfrac{1}{\beta} = 1$, $\|X\|_\alpha = \{\mathrm{E}\,|X|^\alpha\}^{1/\alpha}$), applied to X_s and Y^{p-1} and

$$\alpha = p\ ,\ \beta = \frac{p}{p-1}\ ,\ \|Y^{p-1}\|_\beta = \{\mathrm{E}\,Y^p\}^{1/\beta}$$

we get

$$\mathrm{E}\,|X_s Y^{p-1}| \le \|X_s\|_p \|Y^{p-1}\|_{p/(p-1)} = \|X_s\|_p \|Y\|_p^{p-1}. \tag{6}$$

From (5) and (6) we get $\|Y\|_p^p \le \dfrac{p}{p-1}\|X_s\|_p\|Y\|_p^{p-1}$. This implies (4). □

Let us now look at the variation of the trajectories of continuous martingales. Let $\{\pi_m[0,t]\}_{m\in\mathbb{N}}$ be a sequence of partitions of $[0,t]$, i.e. $0 = t_0^m < \cdots < t_{n(m)}^m = t$ such that $\max\{|t_{i+1}^m - t_i^m|\,;\, i = 0,1,\ldots,n(m)-1\} \to 0$ as $m \to \infty$. Let us denote this by mesh $\pi_m[0,t] \underset{m\to\infty}{\longrightarrow} 0$.

Proposition 1. Let $(M_t\,;\, t \ge 0)$ be a continuous martingale such that $EM_t^2 < \infty$ for all $t \ge 0$. If

$$\limsup_{m\to\infty} \sum_{j=0}^{2^m t-1} |M_{(j+1)2^{-m}} - M_{j2^{-m}}| < \infty \quad P\text{-a.s.},$$

then $M_t = M_0$ P-a.s.

This means that the constant martingales are the only continuous martingales whose trajectories are of finite variation.

Proof. Without restriction of generality let us suppose $M_0 = 0$ (otherwise consider the martingale $M_t' = M_t - M_0$). Let $t \in \mathbb{N}$. Then, by the martingale property,

$$\mathrm{E}\,M_t^2 = \mathrm{E}\left[\sum_{j=0}^{2^m t-1}(M_{(j+1)2^{-m}}^2 - M_{j2^{-m}}^2)\right] = \mathrm{E}\left[\sum_{j=0}^{2^m t-1}(M_{(j+1)2^{-m}} - M_{j2^{-m}})^2\right]$$

for all $m \in \mathbb{N}$. Therefore

$$\mathrm{E}\, M_t^2 = \limsup_{m\to\infty} \mathrm{E}\left[\sum_{j=0}^{2^m t-1} (M_{(j+1)2^{-m}} - M_{j2^{-m}})^2\right]$$
$$\leq \limsup_{m\to\infty} \mathrm{E}\left[\sup\{|M_{(j+1)2^{-m}} - M_{j2^{-m}}|\,;\, 0 \leq j \leq 2^m t - 1\} \sum_{j=0}^{2^m t-1} |M_{(j+1)2^{-m}} - M_{j2^{-m}}|\right]$$
$$\leq \mathrm{E}\left[\limsup_{m\to\infty} \sup\{|M_{(j+1)2^{-m}} - M_{j2^{-m}}|\,;\, 0 \leq j \leq 2^m t - 1\} \sum_{j=0}^{2^m t-1} |M_{(j+1)2^{-m}} - M_{j2^{-m}}|\right]$$

by the lemma of Fatou. Since M is continuous and $[0,t]$ is compact, $|M_{(j+1)2^{-m}} - M_{j2^{-m}}| \underset{m\to\infty}{\longrightarrow} 0$ uniformly in j. If the variation of M on $[0,t]$ is finite, then $\mathrm{E}\, M_t^2 = 0$ i.e. $M_t = 0 = M_0$ P-a.s.

It is therefore necessary to look at higher order variation. Let $(B_t\,;\, t \geq 0)$ be a Brownian motion. Since it has independent increments it is a martingale with respect to its natural filtration. In fact, $\mathrm{E}[B_t - B_s \mid \sigma(B_u\,;\, 0 \leq u \leq s)] = \mathrm{E}[B_t - B_s] = 0$ for all $s < t$. One can show that

$$\limsup_{n\to\infty} \sum_{j=0}^{2^m t-1} (B_{(j+1)2^{-m}} - B_{j2^{-m}})^2 = t. \tag{7}$$

The quadratic variation of Brownian motion is therefore finite and nonzero. This is true for any continuous square integrable martingale $(M_t\,;\, t \geq 0)$. In fact, one can show that there exists a continuous increasing process, usually denoted by $\langle M \rangle_t$, which is $\mathcal{F}_t$-measurable for all t and such that

$$\lim_{m\to\infty} \sum_{j=0}^{2^m t-1} (M_{(j+1)2^{-m}} - M_{j2^{-m}})^2 = \langle M \rangle_t \tag{8}$$

$(\langle M \rangle_t\,;\, t \geq 0)$ is called the quadratic variation process.
Let us finally show a generalization of (7).

Proposition 2. Let $\{\pi_m[0,t]\}_{m\in\mathbb{N}}$ be a sequence of partitions of $[0,t]$ with mesh $\pi_m[0,t] \to 0$ as $m \to \infty$, and let $(H_t\,;\, t \geq 0)$ be an $\mathcal{F}_t$-adapted process such that $\int_0^t \mathrm{E}\, H_s^2 ds < \infty$. Then $\left\{\sum_{j=0}^{n(m)} H_{t_{j-1}^m}(B_{t_j^m} - B_{t_{j-1}^m})^2\right\}_{m\in\mathbb{N}}$ converges, as $m \to \infty$, to $\int_0^t \mathrm{E}\, H_s ds$ in $L^2(P)$.

Proof. The write t_j instead of t_j^m to simplify the notations. We have

$$\begin{aligned}
& \mathrm{E}\left[\left\{\sum_{j=0}^{n(m)} H_{t_{j-1}}[(B_{t_j}-B_{t_{j-1}})^2-(t_j-t_{j-1})]\right\}^2\right] \\
= & \mathrm{E}\left[\sum_{j,k=0}^{n(m)} H_{t_{j-1}} H_{t_{k-1}}[(B_{t_j}-B_{t_{j-1}})^2-(t_j-t_{j-1})][(B_{t_k}-B_{t_{k-1}})^2-(t_k-t_{k-1})]\right] \\
= & \mathrm{E}\left[\left\{\sum_{j=0}^{n(m)} H_{t_{j-1}}^2\{(B_{t_j}-B_{t_{j-1}})^2-(t_j-t_{j-1})\}^2\right]\right. \\
= & \sum_{j=0}^{n(m)} \mathrm{E}\, H_{t_{j-1}}^2\left\{\mathrm{E}[(B_{t_j}-B_{t_{j-1}})^4]-2(t_j-t_{j-1})\,\mathrm{E}[(B_{t_j}-B_{t_{j-1}})^2]+(t_j-t_{j-1})^2\right\} \\
= & 2\sum_{j=0}^{n(m)} \mathrm{E}\, H_{t_{j-1}}^2 (t_j-t_{j-1})^2 \underset{m\to\infty}{\longrightarrow} 0.
\end{aligned}$$

IV. STOCHASTIC INTEGRALS AND ITO'S FORMULA

IV.1. Stochastic integrals

We have seen in the last paragraph that the trajectories of Brownian motion are P-a.s. not of finite variation on compact intervals. Therefore, integrals with respect to Brownian motion, e.g. $\int_0^t B_s dB_s$ can not be defined as Stieltjes-integrals in general.

In this section we define $\int_0^t H_s dB_s$ for a suitable class of processes $(H_t\,;\,t\geq 0)$. Let $(\Omega,\mathcal{F},P)$ be a probability space, and let $(\mathcal{F}_t\,;\,t\geq 0)$ be an increasing and right-continuous filtration. Let $\mathcal{L}_0$ be the class of processes of the type $H_t = \sum_{i=0}^{\infty} h_i I_{]t_i,t_{i+1}]}(t) + h_0 I_{\{0\}}(t)$, where $\{t_n\}_{n\geq 0}$ is an increasing sequence such that $t_0=0$ and $\lim_{n\to\infty} t_n = \infty$ and where h_i is $\mathcal{F}_{t_i}$-measurable and $\sup_i \|h_i\| := \sup_i \sup_\omega |h_i(\omega)| < \infty$. Let $\mathcal{L}_2$ be the class of measurable and $\mathcal{F}_t$-adapted processes $(H_t\,;\,t\geq 0)$ such that $\mathrm{E}\left[\int_0^t H_s^2(\omega)ds\right] < \infty$ for all $t>0$.

Lemma 1. $\mathcal{L}_0$ is dense in $\mathcal{L}_2$ with respect to the metric d given by

$$d(H,K)=\sum_{n=0}^{\infty} 2^{-n}\inf\{1,\|H-K\|_n),\ \text{where}\ \|H\|_n^2=\mathrm{E}\left[\int_0^n H_s^2 ds\right].$$

Proof. $(\mathcal{L}_2,d)$ is in fact a complete metric space. Let $H\in\mathcal{L}_2$ and let us show that there exists a sequence $\{H_n\}_{n\geq 0}\subset\mathcal{L}_0$ such that $\|H_n-H\|_t\to 0$ as $n\to\infty$ for all $t>0$. First we notice that we may suppose H bounded, i.e. $|H|\leq c$

for some constant $c > 0$. In fact, let $\overline{H}_n(t) = H_t I_{\{|H_t| \leq n\}}$; then $\overline{H}_n$ is bounded by n and, by the dominated convergence theorem, $\lim_{n\to\infty} \|H - \overline{H}_n\|_t = 0$ for all $t > 0$. Second, let $\tilde{H}_\epsilon(t) = \frac{1}{\epsilon}\int_{t-\epsilon}^t H_u I_{R_+}(u)du$ for some fixed $\epsilon > 0$. $(\tilde{H}_\epsilon(t); t \geq 0)$ is still $\mathcal{F}_t$-adapted and is moreover continuous. By choosing H bounded and noticing that $\tilde{H}_\epsilon(t) \to H(t)$ as $\epsilon \to 0$ a.e. on $\mathbb{R}_+$, we conclude again by the dominated convergence theorem that $\lim_{n\to\infty} \|\tilde{H}_{1/n} - H\|_t = 0$ for all $t > 0$. The may therefore choose H continuous and bounded. Let now $H_n(t) = H(\frac{k}{2^n})$ if $t \in]\frac{k}{2^n}, \frac{k+1}{2^n}]$. H_n is $\mathcal{F}_t$-adapted and converges uniformily on $[0, t]$ to H as $n \to \infty$ P-a.s. Since H is bounded $\lim_{n\to\infty} \|H_n - H\|_t = 0$ for all $t > 0$. □

Let us now define the stochastic integral $(I(H)_t; t \geq 0)$ for $H \in \mathcal{L}_0$ as follows :

$$I(H)_t := \sum_{i=0}^{\infty} h_i (B_{t_{i+1}\wedge t} - B_{t_i \wedge t}) \tag{1}$$

We have

$$\begin{aligned} E\big[I(H)_t^2\big] &= E\big[\textstyle\sum_{i=0}^\infty h_i^2 (B_{t_{i+1}\wedge t} - B_{t_i\wedge t})^2\big] \text{ (since } B \text{ has independant increments)} \\ &= \textstyle\sum_{i=0}^\infty E\big[h_i^2 E[(B_{t_{i+1}\wedge t} - B_{t_i\wedge t})^2]\mathcal{F}_{t_i\wedge t}]\big] \text{ (by property c) of} \\ &\qquad\qquad\qquad\qquad \text{(conditional expectations)} \\ &= \textstyle\sum_{i=0}^\infty E h_i^2 (t_{i+1} - t_i) = E\big[\int_0^t H_s^2 ds\big]. \end{aligned} \tag{2}$$

Notice that $(I(H)_t; t \geq 0)$ is a continuous martingale such that $E(I(H)_t^2] < \infty$ for all $t \geq 0$. Let us denote by $\mathcal{M}^2$ the space of these martingales. (2) shows therefore that $I : \mathcal{L}_0 \to \mathcal{M}^2$ is an isometry between $\mathcal{L}_0$ (endowed with the norm $\|.\|_t$) and $\mathcal{M}^2$ (endowed with the norm $\|I(H)\|_2 = \{E[I(H)_t^2]\}^{1/2}$). Since $\mathcal{L}_0$ is dense in $\mathcal{L}_2$ and $\mathcal{L}_2$ is a complete metric space (endowed with d), I admits a unique extension to $\mathcal{L}_2$ and (2) holds for all $H \in \mathcal{L}_2$. Let us denote this extension by $I(H)_t = \int_0^t H_s dB_s$. This is the so-called **Itô-integral**. Notice that h_i is supposed to be $\mathcal{F}_{t_i}$-measurable, i.e. $H \in \mathcal{L}_2$ must be evalued at t_i, the starting points of the subintervals $]t_i, t_{i+1}]$. As a consequence it will be shown in theorem 2 below that the integral $I(H)$ is a martingale. This is no longer true if we replace H_{t_i} by $\frac{1}{2}(H_{t_i} + H_{t_{i+1}})$, which gives the so-called **Stratonovich integral** ; in fact, the Stratonovich integral, usually denoted by $\int_0^t H_u \circ dB_u$ is different from the Itô integral and is not a martingale in general. Let us now summarize the result on the Itô-integral.

Theorem 2. Let $(B_t; t \geq 0)$ be a Brownian motion and let $(H_t; t \geq 0) \in \mathcal{L}_2$. Then the Itô integral $(\int_0^t H_u dB_u, t \geq 0)$ is in $\mathcal{M}^2$ and has the following properties :

$$\begin{array}{lll} a) & H, K \in \mathcal{L}_2 \Longrightarrow \int_0^t (H_u + K_u) dB_u = \int_0^t H_u dB_u + \int_0^t K_u dB_u & (3) \\ b) & H, K \in \mathcal{L}_2 \Longrightarrow E\big[\int_s^t H_u dB_u \int_s^t K_u dB_u | \mathcal{F}_s\big] = E\big[\int_s^t H_u K_u du | \mathcal{F}_s\big] & (4) \\ & \text{in particular : } E\big[(\int_s^t H_u dB_u)^2 | \mathcal{F}_s\big] = E\big[\int_s^t H_u^2 du | \mathcal{F}_s\big] & \\ c) & P\big(\sup_{[0,t]} |\int_0^s H_u dB_u| \geq \lambda\big) \leq \frac{1}{\lambda^2} \int_0^t E H_u^2 du \text{ for all } \lambda > 0. & (5) \end{array}$$

Proof. $a)$ and $b)$ are proved by means of approximations by elements of $\mathcal{L}_1$, where (3) and (4) hold trivially. The extension to $\mathcal{L}_2$ follows from (2). To prove $c)$, we apply Theorem (3.1) to the positive submartingale $I(H)_t^2$ to get the upper bound $\frac{1}{\lambda^2} E[I(H)_t^2]$, which is equal to the right side of (5) by the isometry (2). □

Remark. Part $b)$ means in fact that the quadratic variation process of the Itô integral is given by

$$< \int_0^{\cdot} H_u dB_u >_t = \int_0^t H_u^2 du \quad \text{for all} \quad t \geq 0.$$

Let us finally consider as an example the Itô integral $\int_0^t B_s dB_s$ and compare it to the Stratonovich integral. Let $0 = t_0 < t_1 < \cdots < t_n = t$ $(n \in \mathbb{N})$ be partitions of $[0, t]$ with mesh converging to zero as $n \to \infty$. By a direct computation or by means of the Itô formula which we are going to prove in the next section it is easy to see that $\sum_{i=0}^{n-1} B_{t_i}(B_{t_{i+1}} - B_{t_i})$ converges in $\mathcal{M}^2$ to $\frac{1}{2}(B_t^2 - t)$ which is therefore the Itô integral $\int_0^t B_s dB_s$. In order to determine the Stratonovich integral, notice that

$$\sum_{i=0}^{n-1} \frac{1}{2}(B_{t_i} + B_{t_{i+1}})(B_{t_{i+1}} - B_{t_i}) = \frac{1}{2} \sum_{i=0}^{n-1} (B_{t_{i+1}}^2 - B_{t_i}^2) = \frac{1}{2} B_t^2.$$

Therefore :

$$\int_0^t B_s \circ dB_s = \frac{1}{2} B_t^2 = \int_0^t B_s dB_s + \frac{1}{2} t.$$

Let us mention that the Itô integral defined in this section can be extended to the class of measurable and adapted processes such that $\int_0^t H_s^2 ds < \infty$ P-a.s. for all $t \geq 0$. In fact, there exists a sequence $\{T_n\}_{n \geq 1}$ of stopping times T_n

such that $T_n \to \infty$ P-a.s. as $n \to \infty$ and such that $(H_{t\wedge T_n}; t \geq 0) \in \mathcal{L}_2$. Take e.g. $T_n = \inf\{t > 0; \int_0^t H_s^2 ds > n\}$. The integral $\int_0^t H_s dB_s$ is then defined by

$$\int_0^t H_s dB_s = \lim_{n\to\infty} \int_0^{t\wedge T_n} H_n(s) dB_n(s),$$

where $H_n(t) = H_{t\wedge T_n}$ and $B_n(t) = B_{t\wedge T_n}$.

IV.2 The Itô formula.

In this section we look at processes $(X_t; t \geq 0)$ of the form

$$X_t = x_0 + \int_0^t u_s ds + \int_0^t v_s dB_s \tag{6}$$

where $(B_t; t \geq 0)$ is Brownian motion, $(u_t; t \geq 0)$ and $(v_t; t \geq 0)$ are measurable and $\mathcal{F}_t$-adapted processes such that

$$\int_0^t |u_s| ds < \infty, \quad \int_0^t v_s^2 ds < \infty \text{ for all } t \geq 0 \quad P\text{-a.s.}$$

Theorem 3. Let $g(t,x) \in C^2([0,\infty[\times\mathbb{R})$ and let $(X_t; t \geq 0)$ be given by (6). Then

$$g(t,X_t) = g(0,X_0) + \int_0^t \frac{\partial g}{\partial s}(s,X_s)ds + \int_0^t \frac{\partial g}{\partial x}(s,X_s)[u_s ds + v_s dB_s] + \frac{1}{2}\int_0^t \frac{\partial^2 g}{\partial x^2}(s,X_s) v_s^2 ds. \tag{7}$$

Proof. The follow Oksendal (1992). First we notice that we may assume that g and its partial derivatives appearing in (7) are bounded. In fact if the theorem is proved under this condition, we approximate g by C^2-functions g_n such that g_n and its partial derivatives are bounded for each m and converge uniformly on compact subsets of $[0,\infty[\times\mathbb{R}$ to g and the corresponding partial derivatives.

By Taylor's formula

$$\begin{aligned} g(t,X_t) &= g(0,X_0) + \textstyle\sum_j \Delta g(t_j, X_{t_j}) \\ &= g(0,X_0) + \textstyle\sum_j \frac{\partial g}{\partial t}(t_j, X_{t_j})\Delta t_j + \sum_j \frac{\partial g}{\partial x}(t_j, X_{t_j})\Delta X_j + \frac{1}{2}\sum_j \frac{\partial^2 g}{\partial t^2}(t_j, X_{t_j})(\Delta t_j)^2 \\ &+ \textstyle\sum_j \frac{\partial^2 g}{\partial t \partial x}(t_j, X_{t_j})\Delta t_j \Delta X_j + \frac{1}{2}\sum_j \frac{\partial^2 g}{\partial x^2}(t_j, X_{t_j})(\Delta X_j)^2 + \sum_j R_j, \end{aligned} \tag{8}$$

where $\Delta t_j = t_{j+1} - t_j$, $\Delta X_j = X_{t_{j+1}} - X_{t_j}$, $\Delta g(t_j, X_{t_j}) = g(t_{j+1}, X_{t_{j+1}}) - g(t_j, X_{t_j})$ and $0 = t_0 < t_1 < \cdots < t_n = t$ is a partition of $[0,t]$. Moreover $R_j = \sigma(|\Delta t_j|^2 + |\Delta X_j|^2)$ for all j as $n \to \infty$.

We look now at the asymptotic behaviour of each term on the right hand side of (8) as the mesh of the partitions converges to zero. The 4 th. and 5 th. term are easily seen to converge to zero since Δt_j is upper bounded by the mesh, which can be taken in front of the sums while the remainder of the sums stay bounded in $L^1(P)$ or $L^2(P)$. Moreover, as $n \to \infty$,

$$\begin{array}{lll} \sum_j \frac{\partial g}{\partial t}(t_j, X_{t_j})\Delta t_j & \to \int_0^t \frac{\partial g}{\partial s}(s, X_s)ds & \text{(Lebesgue integral)} \\ \sum_j \frac{\partial g}{\partial x}(t_j, X_{t_j})\Delta X_j & \to \int_0^t \frac{\partial g}{\partial x}(s, X_s)dX_s & \text{(by Theorem 2)} \end{array}$$

Now $(\Delta X_j)^2 = u_{t_j}^2(\Delta t_j)^2 + 2u_{t_j}v_{t_j}(\Delta t_j)(\Delta B_j) + v_{t_j}^2(\Delta B_j)^2$. By similar arguments as above, the sums with respect to the first and second term will converge to zero, while in the last term we have the quadratic variation of the integral $\int_0^t v_s dB_s$, i.e. the sum with respect to this term converges to

$$\frac{1}{2}\int_0^t \frac{\partial^2 g}{\partial x^2}(s, X_s)d <X>_s = \frac{1}{2}\int_0^t \frac{\partial^2 g}{\partial x^2}(s, X_s)v_s^2 ds.$$

This follows from Proposition 3.2. By the same arguments one shows that $\sum_j R_j \to 0$ as $n \to \infty$. □

Example. Let $u_s = 0$ and $v_s = 1$ and $g(t,x) = x^2$. Then $X_t = B_t$ and $g(t, B_t) = B_t^2 = B_0^2 + 2\int_0^t B_s dB_s + \frac{1}{2}\int_0^t 2ds$. Therefore we get $\int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t)$.

We generalize now the Itô formula to the multidimensional case.

Theorem 4. Let $B = (B^1, \cdots, B^m)$ be an m-dimensional Brownian motion (*i.e.* $B^1, \cdots, B^m$ are independant Brownian motions in $\mathbb{R}$), let $u = (u^1, \cdots, u^n)$ and $v = (v^{ij}; i = 1, \cdots, n; j = 1, \cdots, m)$, where the u^i and v^{ij} satisfy the same assumptions as u and v in Theorem 3. Let $X = (X^1, \cdots, X^n)$ be given by

$$X = X_0 + \int_0^t u_s ds + \int_0^t v_s dB_s.$$

Let $g(t,x) \in C^2([0,\infty[\times\mathbb{R}^n)$. Then

$$\begin{aligned} g(t, X_t) &= g(0, X_0) + \int_0^t \frac{\partial g}{\partial s}(s, X_s)ds + \sum_{i=1}^n \int_0^t \frac{\partial g}{\partial x_i}(s, X_s)dX_s^i \\ &+ \frac{1}{2}\sum_{i,j=1}^n \int_0^t \frac{\partial^2 g}{\partial x_i \partial x_j}(s, X_s)dX_s^i dX_s^j, \end{aligned} \tag{9}$$

where $dX_s^i dX_s^j = \sum_{k=1}^m (v_s^{ij})^2 ds$.

The proof is similar to Theorem 3 and is therefore omitted.

IV.3. Applications of the Itô formula.

a) Integration by parts. Let $X_t^1 = \int_0^t u_s ds$ and $X_t^2 = B_t$ and let $g(t, x_1, x_2) = x_1 x_2$. Then, by (9),

$$\begin{aligned} g(t, X_t^1, X_t^2) = X_t^1 B_t &= \int_0^t B_s dX_s^1 + \int_0^t X_s^1 dB_s + \int_0^t dX_s^1 dB_s \\ &= \int_0^t B_s u_s ds + \int_0^t X_s^1 dB_s \end{aligned}$$

The stochastic integral of a process of finite variation can therefore be calculated as a Lebesgue integral.

b) Exponential. Let $X_t^1 = B_t$, $X_t^2 = t =< B >_t$ and let $g(t, x_1, x_2) = \exp(x_1 - \frac{1}{2}x_2)$. Then, by (9),

$$\exp(B_t - \frac{1}{2}t) = 1 + \int_0^t \exp(B_s - \frac{1}{2}s) dB_s. \quad (10)$$

This process is called the stochastic exponential ; it satisfies the stochastic differential equation $X_t = 1 + \int_0^t X_s dB_s$. Furthermore, for $\lambda > 0$, $\exp(\lambda X - \frac{1}{2}\lambda^2 u) = \sum_{n=0}^{\infty} \lambda^n H_n(u, x)$, where $H_n(u, x) = \frac{1}{n!}(-u)^n \exp(\frac{x^2}{2u}) \frac{\partial^n}{\partial x^n} \exp(-\frac{x^2}{2u})$ are the Hermite polynomials. It is well known that $\{H_n(u, x)\}_{n \geq 0}$ froms an orthogonal basis of $L^2(\exp(-\frac{x^2}{2u}))$. Now,one can show that $\exp(B_t - \frac{1}{2}t) = \sum_{n=0}^{\infty} H_n(t, B_t)$, where, by iterating (10), we get $H_n(t, B_t) = \int_0^t dB_{t_1} \int_0^{t_1} dB_{t_2} \cdots \int_0^{t_{n-1}} dB_{t_n}$. One can show that $\{H_n(., B_.)\}_{n \geq 0}$ is orthogonal in $L^2(P)$. H_n is therefore called the nth stochastic Hermite polynomial, and the series expansion is called the chaos expansion of the stochastic exponential.

c) Harmonic functions. Let $B = (B^1, \cdots, B^n)$ be a Brownian motion in $\mathbb{R}^n$, i.e. $B^1, \cdots, B^n$ are independent Brownian motions in $\mathbb{R}$. Suppose that $B_0 = x \neq 0$, let

$$\begin{aligned} f(x) &= \log(|x|^2) \text{ if } n = 2, \\ f(x) &= |x|^{-\frac{n-1}{2}} \text{ if } n \geq 3. \end{aligned}$$

By (9) we get

$$\begin{aligned} &\text{for} \quad n = 2 \quad : \log|B_t| = \log|x| + \int_0^t \frac{B_s^1}{|B_s|^2} dB_s^1 + \int_0^t \frac{B_s^2}{|B_s|^2} dB_s^2, \\ &\text{for} \quad n = 3 \quad : |B_t|^{-1} = |x|^{-1} - \sum_{i=1}^3 \int_0^t \frac{B_s^i}{|B_s|^2} dB_s^i. \end{aligned}$$

We see that $f(B_t)$ is a martingale. This us due to the fact that the second order term in the Itô formula vanishes ; in fact, f is harmonic and therefore $\sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} f(X) = 0$.

V. Stochastic differential equations

V.1. Existence and uniqueness of strong solutions.

Having defined integrals with respect to Brownian motion we are now in position to consider stochastic differential equations. In the introduction we mentioned several problems, including population growth, which lead to stochastic differential equations. As in the classical theory we start with an existence and uniqueness theorem for solutions under Lipschitz conditions.

Theorem 1. Let $T > 0$. Let $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^m$ and $\sigma : [0,T] \times \mathbb{R}^n \times \mathbb{R}^{n\times m}$ be measurable functions satisfying

$$|b(t,x)| + |\sigma(t,x)| \leq C(1 + |x|) \qquad x \in \mathbb{R}^n, t \in [0,T] \tag{6}$$

for some constant C, where $|\sigma(x)|^2 = \sum_{i=1}^n \sum_{j=1}^m |\sigma^{ij}(x)|^2$ and

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq D|x - y| \quad x, y \in \mathbb{R}^n, t \in [0,T] \tag{L}$$

for some constant D. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, let $(\mathcal{F}_t; t \geq 0)$ be an increasing and rightcontinuous filtration and let $B = (B^1, \cdots, B^m)$ be an $\mathcal{F}_t$-Brownian motion in $\mathbb{R}^m$. Moreover let Z be an $\mathcal{F}_0$-measurable random variable in $\mathbb{R}^n$ with $E|Z|^2 < \infty$. Then the system of stochastic differential equations

$$X_t = Z + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s \qquad t \in [0,T], \tag{1}$$

has a unique (up to indistinguishability) $\mathcal{F}_t$-adapted, continuous solution $(X_t; t \in [0,T])$.

Remark. The solution $(X_t; t \geq 0)$ of (1) is called a strong solution since it is defined on the given probability space $(\Omega, \mathcal{F}, P)$.

Proof. The follow Oksendal (1992). Let us first show uniqueness. It will follow from the isometry (4.2) and the Lipschitz property (L). Let $(X_t; t \in [0,T])$ and $(\hat{X}_t; t \in [0,T])$ be solutions of (1) satisfying the conditions of $(X_t; t \in [0,T])$ stated in the theorem and with initial conditions $X_0 = Z$ and $\hat{X}_0 = \hat{Z}$.

Put $a(s,\omega) = b(s, X_s) - b(s, \hat{X}_s)$ and $\gamma(s,\omega) = \sigma(s, X_s) - \sigma(s, \hat{X}_s)$. Then

$$\begin{aligned}
&E[|X_t - \hat{X}_t|^2] = E[(Z - \hat{Z} + \textstyle\int_0^t \text{ ads } + \int_0^t \gamma dB_s)^2] \\
&\leq 3E[|Z - \hat{Z}|^2] + 3E[(\textstyle\int_0^t \text{ ads})^2] + 3E[(\int_0^t \gamma dB_s)^2] \\
&\leq 3E[|Z - \hat{Z}|^2] + 3tE[\textstyle\int_0^t a^2 ds] + 3E[\int_0^t \gamma^2 ds] \\
&\leq 3E[|Z - \hat{Z}|^2] + 3(1+t)D^2 \textstyle\int_0^t E[|X_s - \hat{X}_s|^2] ds.
\end{aligned}$$

So the function

$$v(t) = E[|X_t - \hat{X}_t|^2]; 0 \leq t \leq T$$

satisifies

$$v(t) \leq F + A \int_0^t v(s) ds, \text{ where } F = 3E[|Z - \hat{Z}|^2] \text{ and } A = 3(1+T)D^2. \quad (2)$$

Let $w(t) = \int_0^t v(s)ds$. Then $w'(t) \leq F + Aw(t)$. Hence, since $w(0) = 0$, $w(t) \leq \frac{F}{A}(\exp(At) - 1)$. (Consider $f(t) = w(t)\exp(-At)$.) Therefore

$$v(t) \leq F \exp(At).$$

N ow assume that $Z = \hat{Z}$. Then $F = 0$ and so $v(t) = 0$ for all $t \geq 0$. Hence

$$P[|X_t - \hat{X}_t| = 0 \quad \text{for all } t \in \mathbb{Q} \cap [0,T]] = 1,$$

where $\mathbb{Q}$ denotes the rational numbers.
By continuity of $t \to |X_t - \hat{X}_t|$ it follows that

$$P[|X(t,\omega) - \hat{X}(t,\omega)| = 0 \text{ for all } t \in [0,T]] = 1,$$

and the uniqueness is proved.
The proof of the existence is similar to the familiar existence proof for ordinary differential equations:
Define $Y_t^{(0)} = X_0$ and $Y_t^{(k)} = Y_t^{(k)}(\omega)$ inductively as follows

$$Y_t^{(k+1)} = X_0 + \int_0^t b(s, Y_s^{(k)}) ds + \int_0^t \sigma(s, Y_s^{(k)}) dB_s. \quad (5)$$

Then, similar computation as for the uniqueness above gives

$$E\left[|Y_t^{(k+1)} - Y_t^{(k)}|^2\right] \leq (1+T)3D^2. \int_0^t E\left[|Y_s^{(k)} - Y_s^{(k-1)}|^2\right] ds,$$

for $k \geq 1$, $t \leq T$ and

$$E\big[|Y_t^{(1)} - Y_t^{(0)}|^2\big] \leq 2C^2t^2\big(1 + E[|X_0|^2]\big) + 2C^2t\big(1 + E[|X_0|^2]\big) \leq A_1 t$$

where the constant A_1 only depends on C, T and $E[|X_0|^2]$. So by induction on k we obtain

$$E\big[|Y_t^{(k+1)} - Y_t^{(k)}|^2\big] \leq \frac{A_2^{k+1}t^{k+1}}{(k+1)!} \quad ; \quad k \geq 0,\ t \in [0,T] \tag{6}$$

for some suitable constant A_2 depending only on C, D, T and $E[|X_0|^2]$. Now

$$\begin{aligned} &\sup_{0\leq t\leq T} |Y_t^{(k+1)} - Y_t^{(k)}| \leq \int_0^T \Big|b(s, Y_s^{(k)}) - b(s, Y_s^{(k-1)})\Big| ds \\ &+ \sup_{0\leq t\leq T} \Big| \int_0^t \big(\sigma(s, Y_s^{(k)}) - b(s, Y_s^{(k-1)})\big) dB_s \Big|. \end{aligned}$$

By the martingale inequality (4.5) we obtain

$$\begin{aligned} P\ &\big[\sup_{0\leq t\leq T} |Y_t^{(k+1)} - Y_t^{(k)}| > 2^{-k}\big] \leq P\Big[\big(\int_0^T |b(s, Y_s^{(k)}) - b(s, Y_s^{(k-1)})| ds\big)^2 > 2^{-2k-2}\Big] \\ &+ P\big[\sup_{0\leq t\leq T} |\int_0^t \big(\sigma(s, Y_s^{(k)}) - \sigma(s, Y_s^{(k-1)})\big) dB_s| > 2^{-k-1}\big] \\ &\leq 2^{2k+2} T \int_0^T E\big(|b(s, Y_s^{(k)}) - b(s, Y_s^{(k-1)})|^2\big) dt \\ &+ 2^{k+2} \int_0^T E\big[|\sigma(s, Y_s^{(k)}) - \sigma(s, Y_s^{(k-1)})|^2\big] ds \\ &\leq 2^{2k+2} D^2 (T+1) \int_0^T \tfrac{A_2^k t^k}{k!} dt \leq \tfrac{(4A_2T)^{k+1}}{(k+1)!}, \text{ if } A_2 \geq 4D^2(T+1). \end{aligned}$$

Therefore, by the Borel-Cantelli lemma,

$$P\big[\sup_{0\leq t\leq T} |Y_t^{(k+1)} - Y_t^{(k)}| > 2^{-k} \text{ for infinitely many } k\big] = 0.$$

Thus, for a.a. ω there exists $k_0 = k_0(\omega)$ such that

$$\sup_{0\leq t\leq T} |Y_t^{(k+1)} - Y_t^{(k)}| < 2^{-k} \text{ for } k \geq k_0.$$

Therefore the sequence

$$Y_t^{(n)}(\omega) = Y_t^{(0)}(\omega) + \sum_{k=0}^{n-1} \big(Y_t^{k+1}(\omega) - Y_t^{(k)}(\omega)\big)$$

is uniformly convergent in $[0,T]$, for a.a. ω.
Denote the limit by $X_t = X_t(\omega)$. Then X_t is t-continuous for a.a. ω since $Y_t^{(n)}$ is t-continuous for all n. Moreover, $X_t(.)$ is $\mathcal{F}_t$-measurable for all t, since $Y_t^{(n)}(.)$ has this property for all n.

Next, note that for $m > n \geq 0$ we have by (6)

$$E\big[|Y_t^{(m)} - Y_t^{(n)}|^2\big]^{1/2} = ||Y_t^{(m)} - Y_t^{(n)}||_{L^2(P)} = \| \sum_{k=n}^{m-1}(Y_t^{(k+1)} - Y_t^{(k)}||_{L^2(P)}$$
$$\leq \sum_{k=n}^{m-1} ||Y_t^{(k+1)} - Y_t^{(k)}||_{L^2(P)} \leq \sum_{k=n}^{\infty} \big[\frac{(A_2T)^{k+1}}{(k+1)!}\big]^{1/2} \to 0 \text{ as } n \to \infty. \tag{7}$$

So $\{Y_t^{(n)}\}$ converges in $L^2(P)$ to a limit Y_t, say. A subsequence of $Y_t^{(n)}(\omega)$ will then converge ω-pointwise to $Y_t(\omega)$ and therefore we must have $Y_t = X_t$ a.s. In particular, $E[|X_t|^2] < \infty$.

It remains to shows that X_t satisfies (1). For all n we have

$$Y_t^{(n+1)} = X_0 + \int_0^t b(s, Y_s^{(n)})ds + \int_0^t \sigma(s, Y_s^{(n)})dB_s. \tag{8}$$

Now $Y_t^{(n+1)} \to X_t$ as $n \to \infty$, uniformly in $t \in [0,T]$ for a.a. ω. By (7) and the Fatou lemma we have

$$E\big[\int_0^T |X_t - Y_t^{(n)}|^2 dt\big] \leq \limsup_{m\to\infty} E\big[\int_0^T |Y_t^{(m)} - Y_t^{(n)}|^2 dt\big] \to 0$$

as $n \to \infty$. It follows be the Itô isometry that

$$\int_0^t \sigma(s, Y_s^{(n)})dB_s \to \int_0^t \sigma(s, X_s)dB_s$$

and by the Hölder inequality that

$$\int_0^t b(s, Y_s^{(n)})ds \to \int_0^t b(s, X_s)ds$$

in $L^2(\Omega)$. Therefore, taking the limit of (8) as $n \to \infty$ we obtain (1) for X_t. □

The importance of the above theorem lies in the fact that it holds for more general driving processes than Brownian motion and Lebesgue measure in (1). Uniqueness of strong solutions of (1) for $m = n = 1$ can be proven under more general conditions on σ, as the following theorem shows.

Theorem 2. Let $h : \mathbb{R}_+ \to \mathbb{R}_+$ be non-decreasing with $h(0) = 0$ and $\int_{0+}^{\epsilon} \frac{du}{h(u)^2} = \infty$ for some $\epsilon > 0$. Suppose $m = n = 1$ in (1) and

$$\begin{aligned} |\sigma(t,x) - \sigma(t,y)| &\leq h(|x-y|) \text{ for all } t \geq 0 \text{ and all } x, y \in \mathbb{R}, \\ |b(t,x) - b(t,y)| &\leq L|x-y| \text{ for some constant } D > 0 \text{ and all } x, y \in \mathbb{R}. \end{aligned}$$

Then there is uniqueness of solutions of (1).

Remark. The proof of existence of a solution requires other results and is too long to include into this text. The refer to the book of N. IKEDA and S. WATANABE. Let us notice that Theorem 2 is applicable to the case where σ is Hölder continuous with exponent $\frac{1}{2}$, i.e. $|\sigma(t,x)-\sigma(t,y)| \leq C|x-y|^{\frac{1}{2}}$ for all $t>0$ and $x,y \in \mathbb{R}$ (Choose $h(u) = Cu^{\frac{1}{2}}$.)

Proof. Let $\{a_n\}_{n\in\mathbb{N}}$ be defined by $1 > a_1 > a_2 > \cdots > a_n > \cdots > 0$ and

$$\int_{a_1}^{1} h(u)^{-2}du = 1, \int_{a_2}^{a_1} h(u)^{-2}du = 2, \cdots, \int_{a_n}^{a_{n-1}} h(u)^{-2}du = n, \cdots$$

Then $a_n \to 0$ as $n \to \infty$. Let ψ be a continuous function with support in $]a_n, a_{n-1}[$ and such that $0 \leq \psi(u) \leq \frac{2}{n}h^{-2}(u)$ and $\int_{a_n}^{a_{n-1}} \psi(u)du = 1$ for all $n \in \mathbb{N}$. Let $\varphi_n(x) = \int_0^{|x|} dy \int_0^y \psi(u)du$. Then, by continuity of ψ, φ_n is contimously differentiable and $|\varphi_n'(x)| \leq 1$, $\varphi_n(x) \to |x|$ as $n \to \infty$.
Let now X and $\hat{X}$ be two solutions of (1) with $X_0 = \hat{X}_0$ P-a.s. Then

$$X_t - \hat{X}_t = \int_0^t \big[b(s,X_s) - b(s,\hat{X}_s)\big]ds + \int_0^t \big[\sigma(s,X_s) - \sigma(s,\hat{X}_s)\big]dB_s.$$

By Itô's formula (4.7)

$$\begin{aligned} \varphi_n(X_t - \hat{X}_t) &= \varphi_n(X_0 - \hat{X}_0) \\ &+ \textstyle\int_0^t \varphi_n'(X_s - \hat{X}_s)\big\{\big[b(s,X_s) - b(s,\hat{X}_s)\big]ds + \big[\sigma(s,X_s) - \sigma(s,\hat{X}_s)\big]dB_s\big\} \\ &+ \tfrac{1}{2}\textstyle\int_0^t \varphi_n''(X_s - \hat{X}_s)\big[\sigma(s,X_s) - \sigma(s,\hat{X}_s)\big]^2 ds. \end{aligned} \tag{9}$$

Let us denote by I_1 and I_2 the expectations of the integrals on the right side of (9).

$$\begin{aligned} |I_1| &\leq L\textstyle\int_0^t E\big[|\varphi_n'(X_s - \hat{X}_s)|\,|X_s - \hat{X}_s|\big]ds \leq L\int_0^t E|X_s - \hat{X}_s|ds \\ |I_2| &\leq \textstyle\int_0^t \frac{2}{n}h(|X_s - \hat{X}_s|)^{-2}h(|X_s - \hat{X}_s|)^2 ds = \frac{2t}{n} \to 0 \text{ as } n \to \infty. \end{aligned}$$

Since $E[\varphi_n(X_t - \hat{X}_t)] \to E[X_t - \hat{X}_t|$ as $n \to \infty$, it follows that

$$E|X_t - \hat{X}_t| \leq K\int_0^t E|X_s - \hat{X}_s|ds.$$

By the lemma of Growall (see below) we conclude that $E|X_t - \hat{X}_t| = 0$ and therefore $X_t = \hat{X}_t$ P-a.s. for all $t \geq 0$. □

Lemma of Growall. Let $\varphi : [0,t] \to \mathbb{R}$ and $\psi : [0,t] \to \mathbb{R}$ be bounded functions. Then the following implication holds

$$\varphi(t) \le \psi(t) + \lambda \int_0^t \varphi(s)ds \Longrightarrow \varphi(t) \le \psi(t) + \lambda \int_0^t e^{\lambda(t-s)}\psi(s)ds.$$

Proof.

$$\begin{aligned}\varphi(t) &\le \psi(t) + \lambda \textstyle\int_0^t \big(\psi(s) + \lambda \int_0^s \varphi(u)du\big)ds \\ &= \psi(t) + \lambda \textstyle\int_0^t \psi(s)ds + \lambda^2 \int_0^t \int_0^s \varphi(u)du\, ds \\ &\le \cdots \le \psi(t) + \lambda \textstyle\int_0^t \psi(s). \sum_{k=0}^{n-1} \frac{1}{k!}\big(\lambda(t-s)\big)^k ds + \lambda^{n+1} \int_0^t ds_1 \cdots \int_0^t ds_{n+1}\varphi(s_{n+1}).\end{aligned}$$

The first integral converges to $\lambda \int_0^t \psi(s)e^{\lambda(t-s)}ds$, while the second converges to zero since it is upper bounded by $\sup_{[0,t]} |\varphi(s)| \frac{1}{(n+1)!}t^{n+1}$. □

V.2 Solution of some explicit stochastic differential equations.

a) Linear differential equations. Suppose $n = 1$ and consider

$$X_t = X_0 + \int_0^t (A_s X_s + B_s)ds + \int_0^t \sum_{i=1}^m (C_s^i X_s + D_s^i)dB_s^i. \tag{10}$$

We suppose that X_0 is $\mathcal{F}_0$-measurable and $(A_t; t \ge 0)$, $(B_t; t \ge 0)$, $(C_t^i; t \ge 0)$, $(D_t^i; t \ge 0)$ are bounded, measurable and $\mathcal{F}_t$-adapted functions. $B = (B^1, \cdots, B^m)$ is an m-dimensional Brownian motion.

Proposition 3. The solution of (10) is given by

$$X_t = \Phi_t\big\{X_0 + \int_0^t \Phi_s^{-1}(B_s - \sum_{i=1}^m C_s^i D_s^i)ds + \sum_{i=1}^m \int_0^t \Phi_s^{-1} D_s^i dB_s^i\big\}, \tag{11}$$

where

$$\Phi_t = \exp\big\{ \int_0^t (A_s - \frac{1}{2}\sum_{i=1}^m (C_s^i)^2)ds + \sum_{i=1}^m \int_0^t C_s^i dB_s^i\big\} \tag{12}$$

is the solution of the associated homogenous equation

$$\Phi_t = 1 + \int_0^t A_s \Phi_s ds + \int_0^t \sum_{i=1}^m \Phi_s C_s^i dB_s^i.$$

Proof. Let $(U_t; t \ge 0)$ resp. $(V_t; t \ge 0)$ be the quantities in $\{.\}$ in (11) resp (12). Then $X_t = U_t \exp V_t$. An application of Itô's formula (Theorem 4.4) shows that $(X_t; t \ge 0)$ satisfies (10).

b) Next let us consider the following stochastic differential equation :

$$X_t = X_0 + \int_0^t \big[\alpha b(X_s)h(X_s) + \frac{1}{2} b(X_s)b'(X_s)\big]ds + \int_0^t b(X_s)dB_s \qquad (13)$$

where α is a constant and $h(x) = \int \frac{dx}{b(x)}$.

Proposition 4. (13) has the solution $X_t = h^{-1}(Y_t)$, where

$$Y_t = \exp(\alpha t)h(X_0) + \exp(\alpha t)\int_0^t \exp(-\alpha s)dB_s. \qquad (14)$$

Proof. By Itô's formula (Theorem 4.3)

$$h^{-1}(Y_t) = h^{-1}(Y_0) + \int_0^t (h^{-1})'(Y_s)dY_s + \frac{1}{2}\int_0^t (h^{-1})''(Y_s)ds.$$

Moreover,

$$\begin{aligned} (h^{-1})'(Y_s) &= \big[h'(h^{-1}(Y_s))\big]^{-1} = \big[h'(X_s)\big]^{-1} = b(X_s) \\ (h^{-1})''(Y_s) &= (-1)\big[h'(h^{-1}(Y_s))\big]^{-2}\big[h'(h^{-1}(Y_s))\big]' \\ &= (-1)b(X_s)^2 h''(X_s)\big[h^{-1}(Y_s)\big]' = b'(X_s)b(X_s). \end{aligned}$$

Therefore

$$h^{-1}(Y_t) = X_0 + \int_0^t b(X_s)\big[\alpha h(X_s)ds + dB_s\big] + \frac{1}{2}\int_0^t b'(X_s)b(X_s)ds. \qquad \square$$

Example.

$$X_t = X_0 - \int_0^t \big(\sin(2X_s) + \frac{1}{4}\sin(4X_s)\big)ds + \sqrt{2}\int_0^t \cos^2 X_s dB_s.$$

Choose $b(x) = \sqrt{2}\cos^2 x$ (and therefore $h(x) = \frac{1}{\sqrt{2}} tg\, x$) to get the solution

$$X_t = \text{ arctg } (e^{-t} tg\, X_0 + \sqrt{2}e^{-t}\int_0^t e^s dB_s).$$

Let us finally mention two examples of stochastic differential equations, where the hypotheses of Theorem 1 and 2 are not satisfied.

c) $X_t = \int_0^t$ sign $(X_s)dB_s$. If $(X_t; t \geq 0)$ is a solution, $(-X_t; t \geq 0)$ is a solution too. Notice that X is a continuous martingale with $< X >_t =$

$\int_0^t (\text{sign } X_s)^2 ds = t$. One can show that this implies that $(X_t; t \geq 0)$ is a Brownian motion (different from $(B_t; t \geq 0)$ in general).

d)

$$X_t = c - \frac{1}{2}\int_{t_0}^t \exp(-2X_s)ds + \int_{t_0}^t \exp(-X_s)dB_s. \tag{15}$$

By means of Itô's formula (Theorem 4.3) one can show that $X_t = \log(e^c + B_t)$ is a solution of (15). But it only exists up to $T := \inf\{t > t_0; B_t = -e^c\}$. In fact $\lim_{t\to T}|X_t| = \infty$ and X_t is only defined on the interval $[t_0, T[$. Itô's formula must therefore be applied to $X_t^n := \log(e^c + B_{t\wedge T_n})$, where $T_n = \inf\{t > t_0; e^c + B_t < \frac{1}{n}\}$. This gives the solution of (15) on $[t_0, T_n]$ which converges P-a.s. to $[t_0, T[$ as $n \to \infty$.

Appendix : The Fisk-Stratonovich integral.

In defining the stochastic integral in chapter IV, we mentioned that a change of the evaluation point of the subinterval $]t_i, t_{i+1}]$ in a partition of $[0, t]$ changes the value of the integral in general. In order to have an integral which contains both, the Itô -and the Stratonovich- integral, we consider the case where the evaluation point is chosen following a distribution function F with support in $[0, 1]$. We set

$$S_n^F(t) := \sum_{i=0}^{n-1}\int_0^1 f\big(X_{t_{i-1}} + \lambda(X_{t_i} - X_{t_{i-1}})\big)dF(\lambda)(B_{t_i} - B_{t_{i-1}}). \tag{16}$$

If $F = 1_{[0,\infty[}$ (resp. $F = 1_{[\frac{1}{2},\infty[}$), then, under appropriate condtions on F and X, S_n^F will converge to the Itô -(resp. Stratonovich)- integral.

Proposition 5. Let $f : \mathbb{R} \to \mathbb{R}$ be continuously differentiable, let F be a distribution function on $[0, 1]$ and let $\overline{F} = \int_0^1 \lambda dF(\lambda)$. Let $(X_t; t \geq 0)$ be the stochastic process given by (4.6). Then the Fisk-Stratonovich-integral $F - \int_0^t f(X_s)dB_s$, defined as the limit in probability of (16) as $n \to \infty$, is equal to

$$F - \int_0^t f(X_s)dB_s = \int_0^t f(X_s)dB_s + \overline{F}\int_0^t f'(X_s)v_s ds, \tag{17}$$

where the first integral on the right side of (17) is the Itô-integral.

Remark. From (17) we deduce the following relation between the Itô and the Stratonovich integral :

$$\int_0^t f(X_s) \circ dB_s = \int_0^t f(X_s) dB_s + \frac{1}{2} \int_0^t f'(X_s) v_s ds.$$

In particular, the stochastic differential equation $Y_t = \int_0^t f(Y_s) \circ dB_s$ is equivalent to

$$Y_t = \int_0^t f(Y_s) dB_s + \frac{1}{2} \int_0^t f'(Y_s) f(Y_s) ds.$$

Proof of Proposition 5. $S_n^F(t) = I_1 + I_2$, where

$$\begin{aligned} I_1 &= \textstyle\sum_{i=0}^{n-1} f(X_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) \text{ converges to the Itô integral } n \to \infty, \\ I_2 &= \textstyle\sum_{i=0}^{n-1} \int_0^1 \{ f(X_{t_{i-1}} + \lambda(X_{t_i} - X_{t_{i-1}})) - f(X_{t_{i-1}}) \} dF(\lambda)(B_{t_i} - B_{t_{i-1}}) \\ &= \textstyle\int_0^1 \lambda dF(\lambda) \sum_{i=0}^{n-1} f'(X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) \\ &+ \textstyle\sum_{i=0}^{n-1} \int_0^1 \int_0^1 \lambda \{ f'(X_{t_{i-1}} + \lambda s(X_{t_i} - X_{t_{t-1}})) - f'(X_{t_{i_1}}) \} ds dF(\lambda)(X_{t_i} - X_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) \end{aligned} \tag{18}$$

by the mean value theorem. The first sum on the right side of (18) converges to the second integral on the right side of (17), white the second sum converges to zero by continuity of f' and X. □

Example. Population growth. Let $(N_t; t \geq 0)$ the size of a population al time t. Consider

$$\begin{aligned} N_t &= N_0 + \textstyle\int_0^t r N_s ds + \int_0^t \alpha N_s dB_s && \text{(Itô)}, \\ \overline{N}_t &= N_0 + \textstyle\int_0^t r \overline{N}_s ds + \int_0^t \alpha \overline{N}_s \circ dB_s && \text{(Stratonovich)}. \end{aligned}$$

The solutions of these equations are

$$N_t = N_0 \exp\left((r - \frac{1}{2}\alpha^2)t + \alpha B_t\right) \quad ; \quad \overline{N}_t = N_0 \exp(rt + \alpha B_t).$$

The asymptotic behaviour of the solutions is totally different :

$$\lim_{t\to\infty} N_t = \begin{cases} \infty \text{ if } r > \frac{1}{2}\alpha^2 \\ 0 \text{ if } r < \frac{1}{2}\alpha^2 \end{cases} \quad ; \quad \lim_{t\to\infty} \overline{N}_t = \begin{cases} \infty \text{ if } r > 0 \\ 0 \text{ if } r < 0 \end{cases} \quad P\text{-a.s.}$$

This follows by the law of the iterated logarithm for Brownian motions.

VI STOCHASTIC DYNAMICAL SYSTEMS

VI.1 The Markov property and the infinitesimal generator.

Definition. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and let $(\mathcal{F}_t; t \geq 0)$ be an increasing filtration. An $\mathcal{F}_t$-adapted process $(X_t; t \geq 0)$ is markovian if $E[f(X_u)|\mathcal{F}_t] = E[f(X_u)|\sigma(X_t)]$ for all $t, u \geq 0$ with $t < u$ and all bounded and measurable functions $f : \mathbb{R} \to \mathbb{R}$.

Remark. Let $(X_t; t \geq 0)$ be a process with independent increments such that $X_0 = 0$ P-a.s. Then it is markovian with respect to its own filtration. In fact,

$$E\big[f(X_u)|\sigma(X_s; 0 \leq s \leq t)\big] = E\big[g(X_u - X_t, X_t)|\sigma(X_s; 0 \leq s \leq t)\big] \quad (1)$$

where $g(x, X_t) = f(x + X_t)$. Since $\sigma(X_s; 0 \leq s \leq t) = \sigma(X_s - X_r; 0 \leq r < s \leq t)$, the right side of (1) equals to

$$E\big[g(X_u - X_t, X_t)|\sigma(X_t)\big] = E\big[f(X_u)|\sigma(X_t)\big] \text{ for all } t, u \geq 0 \text{ with } t < u.$$

Here we have used the following lemma on conditional expectations :

Lemma 5.1. Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be σ-fields on Ω such that $\mathcal{A} \subset \mathcal{C}$ and $\mathcal{B}$ and $\mathcal{C}$ are independent. Let Z be an $\mathcal{A} \vee \mathcal{B}$-measurable and bounded random variable. Then $E[Z|\mathcal{C}] = E[Z|\mathcal{A}]$.

Proof. We have to show that $\int_C Z\, dP = \int_C E[Z|\mathcal{A}]dP$ for all $C \in \mathcal{C}$. Let first $Z = 1_{A \cap B}$ where $A \in \mathcal{A}$, $B \in \mathcal{B}$. Then

$$\begin{aligned} \int_C 1_{A\cap B} dP &= \int_{C\cap A} 1_B dP = \int_{C \cap A} E[1_B|\mathcal{C}]dP = \int_{C\cap A} E[1_B]dP = \int_{C\cap A} E[1_B|\mathcal{A}]dP \\ &= \int_C E[1_{A\cap B}|\mathcal{A}]dP. \end{aligned}$$

Since $\mathcal{A} \vee \mathcal{B} = \sigma(A \cap B; A \in \mathcal{A}, B \in \mathcal{B})$ the above equation holds for $Z = 1_D$ for all $D \in \mathcal{A} \vee \mathcal{B}$, and, by monotone limits, far all Z satisfying the conditions of Lemma 1. □

Next, we show that the solutions of the stochastic differential equations considered so for are markovian.

Theorem 1. Under the conditions of Theorem 5.1. the solutions of (5.1) is markovian.

Proof. The proof of Theorem 5.1 shows that the sequence $\{(Y_t^{(k)}; t \in [s, T])\}_{k \in \mathbb{N}}$ defined by (5.5) with $[0, t]$ replaced by $[s, t]$ for some fixed $s > 0$ converges to the solution of (5.1) in $L^2(P)$. By construction $Y_t^{(k)}$ is measurable with respect to $\sigma(X_s) \vee \sigma(B_{s+h}; h > 0, s + h \leq t)$. By choosing a subsequence $\{Y_\cdot^{(k_n)}\}_{n\in\mathbb{N}}$ of $\{Y_\cdot^{(k)}\}_{k\in\mathbb{N}}$ which converges P-a.s. to $X_\cdot$, we see that X_t is measurable with respect to $\sigma(X_s) \vee \sigma(B_{s+h} - B_s; h > 0, s + h \leq t)$. The Markov property of $(X_t; t \geq 0)$ follows now from Lemma 1 with $\mathcal{A} = \sigma(X_s)$, $\mathcal{B} = \sigma(B_{s+h} - B_s; h > 0, s + h \leq t)$ and $\mathcal{C} = \mathcal{F}_s$. □

Let us now consider the case $m = n = 1$ in equation (5.1), and let $k : \mathbb{R} \to \mathbb{R}$ be twice continuously differentiable. By Itô's formula (Theorem 4.3) we get

$$k(X_t) = k(X_s)+\int_s^t \left\{k'(X_s)b(s, X_s)+\frac{1}{2}k''(X_s)\sigma(s, X_s)^2\right\}ds+\int_s^1 k'(X_s)\sigma(s, X_s)dB_s.$$

Therefore, for $s < t$,

$$E\left[k(Y_t)|\mathcal{F}_s\right] = k(X_s) + \int_s^t E\left[A_u k(X_u)|\mathcal{F}_s\right]du,$$

where $A_u k(X_u) = k'(X_u)b(u, X_u)+\frac{1}{2}k''(X_u)\sigma(u, X_u)^2$. By the Markov property of X we have

$$E\left[k(X_t)|\mathcal{F}_s\right] = E\left[k(X_t)|\sigma(X_s)\right]$$

and therefore,

$$\lim_{t\to s} \frac{1}{t-s}\left(E\left[k(X_t)|X_s = x\right]-k(x)\right) = A_s k(x) = k'(x)b(s, x)+\frac{1}{2}k''(x)\sigma(s, x)^2, x \in \mathbb{R}. \tag{2}$$

$(A_s; s \geq 0)$ is a real functional on the set of twice continuously differentiable functions and is called the infinitesimal generator of $(X_t; t \geq 0)$. If we choose $k(x) = P(X_u \leq z|X_t = x)$ for some fixed $u > t$ and $z \in \mathbb{R}$ and write $F(t, x; u, z)$ instead of $k(x)$, we get from (2)

$$-\frac{\partial}{\partial s}F(s, x; u, z) = \frac{\partial}{\partial x}F(s, x; u, z)b(s, x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}F(s, x; u, z)\sigma(s, x)^2. \tag{3}$$

This is a partial differential equation for the transition probabilities of $(X_t; t \geq 0)$. Together with the initial value i.e. the law of X_0, (7) determines the law of the process $(X_t; t \geq 0)$. Equation (3) is called the backward equation of Kolmogorov.

Examples :

a) Brownian motion : $A = \frac{1}{2}\frac{d^2}{dx^2}$.

b) Ornstein-Uhlenbeck process : $A = -\alpha x \frac{d}{dx} + \frac{\sigma^2}{2}\frac{d^2}{dx^2}$. The Ornstein-Uhlenbeck process satisfies the stochastic differential equation $X_t = X_0 - \alpha \int_0^t X_s ds + \sigma B_t$.

c) Bessel process : $A = \frac{1}{2x}\frac{d}{dx} + \frac{1}{2}\frac{d^2}{dx^2}$. The Bessel process is the radial component of a Brownian motion in the plane $B = (B^1, B^2) : X_t = ((B_t^1)^2 + (B_t^2)^2)^{1/2}$. It satisifies the stochastic differential equation $X_t = X_0 + \int_0^t \frac{B_s^1}{R_s} dB_s^1 + \int_0^t \frac{B_s^2}{R_s} dB_s^2 + \frac{1}{2}\int_0^t \frac{ds}{R_s}$.

Remark. The infinitesimal generator associated to (5.1) (without the assumption $m = n = 1$ made above) reads as follows :

$$A_s k(X) = \sum_{i=1}^{n} b_i(s,x) \frac{\partial}{\partial x_i} k(x) + \frac{1}{2}\sum_{i,j=1}^{n} (\sigma\sigma^T)_{ij}(s,x) \frac{\partial^2}{\partial x_i \partial x_j} k(x), \quad x \in \mathbb{R}^n,$$

where σ^T is the transposed of σ.

VI.2 Dynamic system's equations.

Let us now consider the system of stochastic differential equations

$$X_t^i = x_i + \int_{t_0}^t b_i(s, X_s, U_s) ds + \int_{t_0}^t \sum_{j=1}^n \sigma_{ij}(s, X_s, U_s) dB_s^j \quad t_0 < t < t_1, i = 1, \cdots, n,$$

where $(U_s; s \geq t_0)$ is the socalled control process, i.e. an $\mathcal{F}_t$-adapted process with values in $\mathbb{R}^k$, by means of which the process $X = (X^1, \cdots, X^n)$ can be influenced. Again, $B = (B^1, \cdots, B^n)$ is an $\mathcal{F}_t$-Brownian motion and $b = (b_1, \cdots, b_n)$ and $\sigma = (\sigma_{ij}; i,j = 1, \cdots, n)$ are supposed to satisfy conditions which imply existence and uniqueness of the solution $X = (X^1, \cdots, X^n)$ for a given initial condition $x = (x_1, \cdots, x_n)$ (e.g. assume that $b(s,x,u)$ and $\sigma(s,x,u)$ satisfy the assumptions of Theorem 5.1, uniformly in u).
Moreover, we assume that a cost function J is given by

$$J(t,x,u) = E_{t,x}\Big[\int_t^{t_1} F(s, X_s, u) ds + K(t_1, X_{t_1})\Big],$$

where $E_{t,x}$ means the expectation with respect to the probability measure of the solution X which starts at time t in the point x and where t_1 is some fixed or random time (e.g. the first time (s, X_s) leaves some domain $G \subset \mathbb{R}^{n+1}$, supposing $(t,x) \in G$). F and K are supposed to be bounded, F in addition continuous. (The meaning of F, K and t_1 will become clear in the examples below).

The problem is now to find a control process $(U_s^*; s \geq t_0)$ which minimizes the cost function. We restrict ourselves to the socalled Markov controls, i.e. the control process for which there exists a function $u_0(s,x)$, such that $U_s = u_0(s, X_s)$ $(s \geq t_0)$. This means that the value of U at time s depends only on the (present) value X_s and not on the past values $(X_r; r \leq s)$.

In the deterministic case, i.e. with $\sigma = 0$, the optimization problem is usually solved by means of the Hamilton-Jacobi equations. This method has been extended by R. BELLMANN to the case of stochastic diffential equations and tells us that

$$\inf_v \big\{F(s,x,v)+A^v h^*(s,x)\big\} = F(s,x,u_0^*(s,x))+A^{u_0^*}h^*(s,x) = 0 \text{ and } h^*(t_1, X_{t_1}) = K(t_1, X_{t_1}),$$

where u_0^* is the function which determines the control process minimizing the cost function and $h^*(s,x) = \inf\{J(s,x,u_0)\ ;\ u_0 \text{ defines a Markov control }\}$. A^v is the infinitesimal generator of $Y_s := (s, X_s)$, parameterized by $v \in \mathbb{R}^k$, i.e.

$$A^v k(s,x) = \frac{\partial}{\partial s} k(s,x) + \sum_{i=1}^n b_i(s,x,v)\frac{\partial}{\partial x_i} k(s,x,v) + \frac{1}{2}\sum_{i,j=1}^n (\sigma\sigma^T)_{ij}(s,x,v)\frac{\partial^2}{\partial x_i \partial x_j} k(s,x).$$

VI.3. Examples of control problems.

a) A portfolio management problem (B. OKSENDAL (1992)). Let $(X_t; t \geq 0)$ be the capital at time t, let X_t^1 (resp. X_t^2) the safe resp. the risky part of X_t and assume that

$$dX_t^1 = aX_t^1 dt, \quad dX_t^2 = bX_t^2 dt + \beta X_t^2 dB_t \quad a, b > 0, a < b, \beta \in \mathbb{R},$$

where $(B_t, t \geq 0)$ is a Brownian motion. let $U_t = \frac{X_t^1}{X_t}$. Then X_t satisfies the stochastic differential equation

$$X_t = x_0 + \int_0^t \big(aU_s X_s + b(1-U_s)X_s\big)ds + \int_0^t \beta(1-U_s)X_s dB_s,$$

where x_0 is the capital at $t = 0$. We assume $x_0 > 0$. The problem is now to choose $(U_t; t \geq 0)$ is such a way that $E|X_{t_1}|^r$ is maximal for some fixed $t_1 > 0$ and $r \in]0,1[$. Therefore $J(s,x,u) = E_{s,x}|X_{t_1}|^r$ (i.e. $F = 0$ and

$K(t_1, X_{t_1}) = |X_{t_1}|^r$). Notice that, here, we have to maximize J ; in order to get the minimization problem of 6.2 we replace J by $-J$ and look for a Markov control which minimizes $-J$ (resp. maximizes J). The infinitesimal generator of (t, X_t) is given by

$$A^v k(t,x) = \frac{\partial}{\partial t} k(t,x) + \big(av + b(1-v)\big) x \frac{\partial}{\partial x} k(t,x) + \frac{1}{2}\beta^2 (1-v)^2 x^2 \frac{\partial^2}{\partial x^2} k(t,x), \quad v \in \mathbb{R}. \tag{4'}$$

We consider now $A^v h^*(t,x)$ as a function of v. The value $\overline{v}$ for which $A^v h^*(t,x)$ is maximal is given by

$$\overline{v} = 1 - \frac{(a-b)\partial/\partial x h^*}{\beta^2 x \partial^2/\partial x^2 h^*.} \tag{5}$$

We see that $\overline{v} \in [0,1]$ is satisfied if $\frac{\partial^2}{\partial x^2} h^* > 0$ and $\frac{\partial}{\partial x} h^* < 0$. If we substitute (5) into (4′) we get

$$\begin{aligned} A^{\overline{v}} h^*(t,x) &= \tfrac{\partial}{\partial t} h^*(t,x) + ax \tfrac{\partial}{\partial x} h^*(t,x) - \tfrac{1}{2} \tfrac{(a-b)^2 (\partial/\partial x h^*(t,x))^2}{\beta^2 \partial^2/\partial x^2 h^*(t,x)} = 0, \\ h^*(t_1, x) &= |x|^r. \end{aligned}$$

Solving this partial differential equation for h^* (by means of separation of variables : $h^*(t,x) = h_1(t) x^r$) gives

$$h^*(t,x) = x^r e^{\lambda(t_1 - t)}, \text{ where } \lambda = ar - \frac{1}{2}\frac{(a-b)^2}{\beta^2}\frac{r}{r-1}.$$

Therefore $U_t^* = 1 - \frac{a-b}{\beta^2(r-1)}$. Notice, that U_t^* is in fact constant. Let us therefore write u^* instead of U_t^*. The solution of the stochastic differential equation we started with is given by

$$X_t = x \exp\big\{\beta(1-u^*)B_t - \frac{1}{2}\beta^2(1-u^*)^2 t + (au^* - b(1-u^*)t\big\}.$$

b) Direct current engine (H.W KNOBLOCH, H. KWAKERNAAK (1985)). Let us consider a simplified model where the engine is driven by the input vollage $(U_t; t \geq 0)$ and where the state of the engine at time t is described by the angle θ_t of a point on the axis of the engine and by the velocity $\dot{\theta}_t = \omega_t$. Then $J\dot{\omega}_t = -B\omega_t + kU_t$, where J, B and k are positive constants (J is the moment of inertia, B is the coefficient of friction, k is a coefficient of proportionality). The system's equation then reads

$$dX_t = LX_t + MU_t dt, \text{ where } X_t = \begin{pmatrix} \theta_t \\ \dot{\theta}_t \end{pmatrix}, L = \begin{pmatrix} 0 & 1 \\ 0 & -\frac{B}{J} \end{pmatrix}, M = \begin{pmatrix} 0 \\ \frac{K}{J} \end{pmatrix}, t_0 < t < t_1.$$

Let us assume that the state of the axis is disturbed by an Ornstein-Uhlenbeck process given by $dZ_t = -\gamma Z_t dt + \sigma dB_t$. Then we have

$$d\begin{pmatrix} X_t \\ Z_t \end{pmatrix} = \begin{pmatrix} & L & \overset{0}{J} \\ 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} X_t \\ Z_t \end{pmatrix} dt + \begin{pmatrix} M \\ 0 \end{pmatrix} U_t dt + \sigma \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} dB_t \equiv \overline{L} \begin{pmatrix} X_t \\ Z_t \end{pmatrix} dt + \overline{M} U_t dt + \overline{\sigma} dB_t.$$

Moreover, let the cost function be given by

$$J(t,x,U) = E_{t,x}\big[X_{t_1}^T R X_{t_1} + \int_t^{t_1} [\theta_t^2 + \rho U_t^2] dt\big], \quad t_0 < t,$$

where $R \in \mathbb{R}^{3\times 3}$, $\rho > 0$, i.e. the costs are proportional to the mean square deviation of the angle and the voltage during the interval $[t, t_1]$ and to the square of the angle at the end of the interval. Again, the problem is to find $(U_t^*; t_0 < t < t_1)$ such that $J(t_0, x, U^*)$ is minimal.
We proceed as in example a) and find $\overline{v}(t,x) = -\frac{1}{2\rho} M^T \nabla_x h^*(t,x)$. In order to determine $h^*(t,x)$ we set $h^*(t,x) = x^T S_t x + f(t)$, where $S_t \in \mathbb{R}^{3\times 3}$ is supposed to be symmetric nonnegative definite, and both, $(S_t; t \geq t_0)$ and $(f_t; t \geq t_0)$ are supposed to be coninously differentiable with respect to t. We set $S_{t_1} = R$ and $f(t_1) = 0$. Then $\overline{v}(t,x) = -\frac{1}{\rho} \overline{M}^T S_t x$, where S_t satisifies the matrix Riccati differential equation

$$\frac{d}{dt} S_t = -2\overline{L}^T S_t + \frac{1}{\rho} S_t \overline{M}\,\overline{M}^T S_t - I_3, \quad t_0 < t < t_1, S_{t_1} = R,$$

where I_3 is the identity in $\mathbb{R}^{3\times 3}$.
Moreover, $f(t)$ is given by $\frac{d}{dt} f(t) = -$ trace $(\sigma \sigma^T S_t)$, $f(t_1) = 0$. Notice that the first term in $h^*(t,x)$, namely $x^T S_t x$, is the solution for the derterministic problem, whereas the second term, namely $f(t)$, is due to the stochastic component.

References

[1] R. ADLER (1981). The geometry of random fields. Wiley.

[2] B. OKSENDAL (1992). Stochastic differential equations. 3ed edition. Springer.

[3] N. IKEDA, S. WATANABE (1988). Stochastic differential equations and diffusion processes. North Holland.

[4] H. W. KNOBLOCH, H. KWAKERNAAK (1985). Lineare Kontrolltheorie. Springer.