cours ARO07–MSSD
Random Models of Dynamical Systems
Introduction to SDE’s
Itô formula and stochastic calculus

François Le Gland
INRIA Rennes + IRMAR

people.rennes.inria.fr/Francois.Le_Gland/insa-rennes/

November 8 and 15, 2021
Itô formula

for Brownian motion

for Itô processes
Itô formula for Brownian motion

**Theorem 1** let \( B \) be a standard Brownian motion and let \( f \) be a twice continuously differentiable function, then

\[
f(B(t)) = f(B(0)) + \int_0^t f'(B(s)) dB(s) + \frac{1}{2} \int_0^t f''(B(s)) ds
\]

almost surely

compare with change of variable formula in (ordinary) differential calculus

\[
\frac{d}{dt} f(X(t)) = f'(X(t)) \frac{d}{dt} X(t)
\]

**Remark** the process \( \phi \) defined by \( \phi(t) = f'(B(t)) \) for any \( t \geq 0 \) belongs to \( M^2_{loc} \): indeed, it is nonanticipating and introducing the r.v.

\[
M(T) = \sup_{0 \leq t \leq T} |B(t)| < \infty
\]

it holds

\[
\int_0^T |\phi(t)|^2 dt \leq (\sup_{|x| \leq M(T)} |f'(x)|^2) T < \infty \quad \text{a.s.}
\]
Example taking $f(x) = x^2$ so that $f'(x) = 2x$ and $\frac{1}{2} f''(x) = 1$, yields

$$B^2(t) = \int_0^t 2B(s) dB(s) + t \quad \text{a.s.}$$

recall (first lecture) that

$$\sum_{i=1}^n B(t^n_{i-1}) (B(t^n_i) - B(t^n_{i-1})) \to \frac{1}{2} (B^2(t) - t)$$

in $\mathbb{L}^2$ as $n \to \infty$, and the limit can also be identified as a stochastic integral, i.e.

$$\sum_{i=1}^n B(t^n_{i-1}) (B(t^n_i) - B(t^n_{i-1})) \to \int_0^t B(s) dB(s)$$

in $\mathbb{L}^2$ as $n \to \infty$
Proof let $0 = t_0^n < t_1^n < \cdots < t_n^n = t$ be a convergent subdivision of $[0, t]$, then (telescopic sum)

$$f(B(t)) = f(B(0)) + \sum_{i=1}^{n} (f(B(t_i^n)) - f(B(t_{i-1}^n)))$$

and Taylor expansion yields

$$f(v) - f(u) = f'(u)(v-u) + \frac{1}{2} f''(\theta(v,u)) (v-u)^2$$

$$= f'(u)(v-u) + \frac{1}{2} f''(u) q$$

$$+ \frac{1}{2} f''(u) [(v-u)^2 - q]$$

$$+ \frac{1}{2} [f''(\theta(v,u)) - f''(u)] (v-u)^2$$

for any $t \geq 0$ and any $0 \leq u \leq v$, and for some $u \leq \theta(v,u) \leq v$
taking \( v = B(t^n_i), u = B(t^n_{i-1}) \) and \( q = t^n_i - t^n_{i-1} \) yields

\[
f(B(t)) = f(B(0)) + \sum_{i=1}^{n} f'(B(t^n_i)) (B(t^n_i) - B(t^n_{i-1})) \\
+ \frac{1}{2} \sum_{i=1}^{n} f''(B(t^n_{i-1})) (t^n_i - t^n_{i-1}) \\
+ \frac{1}{2} \sum_{i=1}^{n} f''(B(t^n_{i-1})) [(B(t^n_i) - B(t^n_{i-1}))^2 - (t^n_i - t^n_{i-1})] \\
+ \frac{1}{2} \sum_{i=1}^{n} [f''(\theta(B(t^n_i), B(t^n_{i-1}))) - f''(B(t^n_{i-1}))] (B(t^n_i) - B(t^n_{i-1}))^2
\]

assume (for simplicity) that the second derivative \( f'' \) is bounded
dominating terms: by definition of the stochastic (Itô) integral

\[ \sum_{i=1}^{n} f'(B(t^n_i)) (B(t^n_i) - B(t^n_{i-1})) \to \int_0^t f'(B(s)) \, dB(s) \]

in \( L^2 \) (hence in probability) as \( n \to \infty \)

and by definition of the usual (Lebesgue) integral

\[ \sum_{i=1}^{n} f''(B(t^n_i)) (t^n_i - t^n_{i-1}) \to \int_0^t f''(B(s)) \, ds \]

almost surely (hence in probability) as \( n \to \infty \)
first error term:

\[ E_1 = \sum_{i=1}^{n} f''(B(t^n_{i-1})) \left[ (B(t^n_i) - B(t^n_{i-1}))^2 - (t^n_i - t^n_{i-1}) \right] \]

satisfies

\[ |E_1|^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} f''(B(t^n_{i-1})) \left[ (B(t^n_i) - B(t^n_{i-1}))^2 - (t^n_i - t^n_{i-1}) \right] \]

and recall that if \( X \) is a Gaussian r.v. with zero mean and variance \( \sigma^2 \) then \( \mathbb{E}|X^2 - \sigma^2|^2 = 2 \sigma^4 \), hence

\[ \mathbb{E}|E_1|^2 = \sum_{i=1}^{n} \mathbb{E}\left[ |f''(B(t^n_{i-1}))|^2 \left| (B(t^n_i) - B(t^n_{i-1}))^2 - (t^n_i - t^n_{i-1}) \right| \right] \]

\[ = 2 \sum_{i=1}^{n} \mathbb{E}|f''(B(t^n_{i-1}))|^2 (t^n_i - t^n_{i-1})^2 \]
second error term:

\[ E_2 = \sum_{i=1}^{n} \left[ f''(\theta(B(t_i^n), B(t_{i-1}^n))) - f''(B(t_{i-1}^n)) \right] (B(t_i^n) - B(t_{i-1}^n))^2 \]

satisfies

\[ |E_2| \leq \sum_{i=1}^{n} |f''(\theta(B(t_i^n), B(t_{i-1}^n))) - f''(B(t_{i-1}^n))| (B(t_i^n) - B(t_{i-1}^n))^2 \]

\[ \leq \max_{i=1 \cdots n} \max_{t_{i-1}^n \leq s \leq t_i^n} |f''(B(s)) - f''(B(t_{i-1}^n))| \sum_{i=1}^{n} (B(t_i^n) - B(t_{i-1}^n))^2 \]

\[ V_n(t) \]

the first factor goes to zero almost surely by continuity (hence in probability) as \( n \uparrow \infty \), and the second factor is bounded in \( L^1 \) (hence in probability)

\[ \square \]
multi–dimensional version

**Theorem 2** let $B$ be a $d$–dimensional Brownian motion with covariance matrix $\Sigma$ and let $f$ be a twice continuously differentiable function, then

$$f(B(t)) = f(B(0)) + \int_0^t f'(B(s)) dB(s) + \frac{1}{2} \int_0^t \text{trace}(\Sigma f''(B(s))) \, ds$$

almost surely

here the $d$–dimensional row vector $f'(x)$ denotes the Jacobian (first order derivative) vector, and the $d \times d$ matrix $f''(x)$ denotes the Hessian (second order derivative) matrix

in the special case of a standard Brownian motion, i.e. if $\Sigma = I$, then $\text{trace}(\Sigma f''(x)) = \text{trace}(f''(x)) = \Delta f(x)$ coincides with the Laplacian of the function $f$
multi–dimensional and time–dependent version

**Theorem 3** * let $B$ be a $d$–dimensional Brownian motion with covariance matrix $\Sigma$ and let $f$ be a differentiable function w.r.t. time and twice continuously differentiable w.r.t. space function, then

$$f(t, B(t)) = f(0, B(0)) + \int_0^t \frac{\partial f}{\partial t}(s, B(s)) \, ds + \int_0^t f'(s, B(s)) \, dB(s)$$

$$+ \frac{1}{2} \int_0^t \text{trace}(\Sigma f''(s, B(s))) \, ds$$

almost surely

compare with change of variable formula in (ordinary) differential calculus

$$\frac{d}{dt} f(t, X(t)) = \frac{\partial f}{\partial t}(t, X(t)) + f'(X(t)) \frac{d}{dt} X(t)$$
Itô processes

Definition $M^1([0, T])$ and $M^1_{loc}([0, T])$ denote the space of nonanticipating processes $\psi$ such that

$$\mathbb{E} \int_0^T |\psi(t)| \, dt < \infty \quad \text{and} \quad \int_0^T |\psi(t)| \, dt < \infty \quad \text{a.s.}$$

respectively, and let $M^1 = \bigcap_{T \geq 0} M^1([0, T])$ and $M^1_{loc} = \bigcap_{T \geq 0} M^1_{loc}([0, T])$ respectively.

Let $B = (B(t), t \geq 0)$ be a Brownian motion a process $X = (X(t), t \geq 0)$ that satisfies

$$X(t) = X(0) + \int_0^t \psi(s) \, ds + \int_0^t \phi(s) \, dB(s)$$

where $\psi \in M^1_{loc}$ and $\phi \in M^2_{loc}$ is called an Itô process.
almost surely the mappings

\[ t \mapsto D(t) = \int_0^t \psi(s) \, ds \quad \text{and} \quad t \mapsto M(t) = \int_0^t \phi(s) \, d\mathcal{B}(s) \]

are continuous on \([0, T]\) for any \(T \geq 0\).

**Proposition 4** [quadratic variation] let \(X\) be an Itô process and let \(0 = t_0^n < t_1^n < \cdots < t_n^n = t\) be a convergent subdivision of \([0, t]\), then the process \(D\) has a finite variation (hence its quadratic variation is zero)

\[
\sum_{i=1}^n |D(t_i^n) - D(t_{i-1}^n)| \leq \int_0^t |\psi(s)| \, ds < \infty
\]

and the process \(X\) has a finite quadratic variation that coincides with the quadratic variation of the martingale \(M\)

\[
\sum_{i=1}^n (X(t_i^n) - X(t_{i-1}^n))^2 \to \langle X \rangle(t) = \int_0^t |\phi(s)|^2 \, ds
\]

in probability as \(n \to \infty\).
Proof first part: note that

\[ |D(t_i) - D(t_{i-1})| = | \int_{t_{i-1}}^{t_i} \psi(s) \, ds | \leq \int_{t_{i-1}}^{t_i} |\psi(s)| \, ds \]

for any \( i = 1 \cdots n \), hence

\[ \sum_{i=1}^{n} |D(t_i) - D(t_{i-1})| \leq \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} |\psi(s)| \, ds = \int_{0}^{t} |\psi(s)| \, ds \]

mechanically

\[ \sum_{i=1}^{n} (D(t_i) - D(t_{i-1}))^2 \leq \max_{i=1 \cdots n} |D(t_i) - D(t_{i-1})| \sum_{i=1}^{n} |D(t_i) - D(t_{i-1})| \]

\[ \leq \max_{i=1 \cdots n} |D(t_i) - D(t_{i-1})| \int_{0}^{t} |\psi(s)| \to 0 \]

almost surely as \( n \uparrow \infty \)
second part: note that

\[
(X(t_i) - X(t_{i-1}))^2 = (D(t_i) - D(t_{i-1}))^2 + (M(t_i) - M(t_{i-1}))^2
\]

\[
+ 2(D(t_i) - D(t_{i-1}))(M(t_i) - M(t_{i-1}))
\]

for any \( i = 1 \cdots n \), hence

\[
\sum_{i=1}^{n} (X(t_i) - X(t_{i-1}))^2 - \langle M \rangle(t) = \sum_{i=1}^{n} (D(t_i) - D(t_{i-1}))^2
\]

\[
+ \sum_{i=1}^{n} (M(t_i) - M(t_{i-1}))^2 - \langle M \rangle(t)
\]

\[
+ 2 \sum_{i=1}^{n} (D(t_i) - D(t_{i-1}))(M(t_i) - M(t_{i-1}))
\]
using the triangle inequality and the Young inequality $2ab \leq \varepsilon a^2 + \frac{b^2}{\varepsilon}$ valid for any non–negative $a, b \geq 0$, yields

\[
\left| \sum_{i=1}^{n} (X(t_i) - X(t_{i-1}))^2 - \langle M \rangle(t) \right| \leq \sum_{i=1}^{n} (D(t_i) - D(t_{i-1}))^2
\]

\[
+ \left| \sum_{i=1}^{n} (M(t_i) - M(t_{i-1}))^2 - \langle M \rangle(t) \right|
\]

\[
+ 2 \sum_{i=1}^{n} |(D(t_i) - D(t_{i-1})) (M(t_i) - M(t_{i-1}))|
\]

\[
\leq \varepsilon \langle M \rangle(t) + \frac{\varepsilon + 1}{\varepsilon} \sum_{i=1}^{n} (D(t_i) - D(t_{i-1}))^2
\]

\[
+ (1 + \varepsilon) \left| \sum_{i=1}^{n} (M(t_i) - M(t_{i-1}))^2 - \langle M \rangle(t) \right|
\]
recall that
\[
| \sum_{i=1}^{n} (M(t_i) - M(t_{i-1}))^2 - \langle M \rangle(t) | \to 0
\]
in probability as \( n \uparrow \infty \), and from the first part
\[
\sum_{i=1}^{n} (D(t_i) - D(t_{i-1}))^2 \to 0
\]
almost surely (hence in probability) as \( n \uparrow \infty \)
therefore, for any positive \( \varepsilon, \delta > 0 \)
\[
\mathbb{P}[| \sum_{i=1}^{n} (X(t_i) - X(t_{i-1}))^2 - \langle M \rangle(t) | > \delta] \leq \mathbb{P}[\varepsilon \langle M \rangle(t) > \frac{1}{3} \delta]
\]
\[
+ \mathbb{P}[\frac{\varepsilon + 1}{\varepsilon} \sum_{i=1}^{n} (D(t_i) - D(t_{i-1}))^2 > \frac{1}{3} \delta]
\]
\[
+ \mathbb{P}[(1 + \varepsilon) | \sum_{i=1}^{n} (M(t_i) - M(t_{i-1}))^2 - \langle M \rangle(t) | > \frac{1}{3} \delta] \to 0
\]
as \( n \uparrow \infty \)
Proposition 5 let

\[ X(t) = X(0) + \int_0^t \psi(s) \, ds + \int_0^t \phi(s) \, dB(s) \]

be an Itô process where \( \psi \in M^1 \) and \( \phi \in M^2 \), then the decomposition is unique.

**Proof** assume that two decompositions exists

\[ X(t) = X(0) + \int_0^t \psi_1(s) \, ds + \int_0^t \phi_1(s) \, dB(s) \]

and

\[ X(t) = X(0) + \int_0^t \psi_2(s) \, ds + \int_0^t \phi_2(s) \, dB(s) \]

then by difference

\[ 0 = \int_0^t [\psi_1(s) - \psi_2(s)] \, ds + \int_0^t [\phi_1(s) - \phi_2(s)] \, dB(s) \]

for any \( t \geq 0 \).
introduce the process

\[ Z(t) = \int_0^t [\psi_1(s) - \psi_2(s)] \, ds = -\int_0^t [\phi_1(s) - \phi_2(s)] \, dB(s) \]

it is both a square integrable martingale, with quadratic variation

\[ \int_0^t |\phi_1(s) - \phi_2(s)|^2 \, ds \]

and a process with finite variation, hence its quadratic variation is zero, in other words

\[ \int_0^t |\phi_1(s) - \phi_2(s)|^2 \, ds = 0 \]

for any \( t \geq 0 \), hence \( \phi_1 \equiv \phi_2 \) and therefore

\[ Z(t) = \int_0^t [\psi_1(s) - \psi_2(s)] \, ds = 0 \]
This implies

\[ \int_0^t [\psi_1(s) - \psi_2(s)]^+ \, ds = \int_0^t [\psi_1(s) - \psi_2(s)]^- \, ds \]

for any \( t \geq 0 \), hence \([\psi_1 - \psi_2]^+ \equiv [\psi_1 - \psi_2]^-\) and finally

\[ \psi_1 - \psi_2 = [\psi_1 - \psi_2]^+ - [\psi_1 - \psi_2]^- \equiv 0 \]
multi–dimensional version

let $B = (B(t), t \geq 0)$ be a $d$–dimensional Brownian motion

a process $m$–dimensional process $X = (X(t), t \geq 0)$ that satisfies

$$X(t) = X(0) + \int_0^t \psi(s) \, ds + \int_0^t \phi(s) \, dB(s)$$

where $\psi \in M^1_{loc}$ and $\phi \in M^2_{loc}$ is called an Itô process

implicitly here, $\psi(s)$ is a $m$–dimensional column vector and $\phi(s)$ is a $m \times d$ matrix

Proposition 6 * [quadratic covariation] let $0 = t^n_0 < t^n_1 < \cdots < t^n_n = t$ be a convergent subdivision of $[0, t]$, then

$$\sum_{i=1}^n (X(t^n_i) - X(t^n_{i-1})) (X(t^n_i) - X(t^n_{i-1}))^* \to \langle X \rangle(t) = \int_0^t \phi(s) \phi^*(s) \, ds$$

in probability as $n \uparrow \infty$
Proposition 7 * any process $\psi$ in $M^1_{loc}([0, T])$ can be approximated by a sequence $\psi_n$ of step processes, in the following sense

$$\int_0^T |\psi(s) - \psi_n(s)| \, ds \to 0 \quad \text{a.s.}$$

any process $\phi$ in $M^2_{loc}([0, T])$ can be approximated by a sequence $\phi_n$ of step processes, in the following sense

$$\int_0^T |\phi(s) - \phi_n(s)|^2 \, ds \to 0 \quad \text{a.s.}$$
Proposition 8 * [integration by parts] let \( X_1 \) and \( X_2 \) be two Itô processes

\[
X_i(t) = X_i(0) + \int_0^t \psi_i(s) \, ds + \int_0^t \phi_i(s) \, dB(s)
\]

for \( i = 1, 2 \), then

\[
X_1(t) X_2(t) = X_1(0) X_2(0) + \int_0^t X_1(s) \, dX_2(s) + \int_0^t X_2(s) \, dX_1(s) + \int_0^t \phi_1(s) \phi_2(s) \, ds
\]

Remark the proposition is true if the processes \( \psi_i(s) \equiv \psi_i \) and \( \phi_i(s) \equiv \phi_i \) are constant for \( i = 1, 2 \), then it is true if these processes are step processes, and the general case follows by approximation.
Itô formula for an Itô process

Theorem 9 * let $X$ be an Itô process and let $f$ be a twice continuously differentiable function, then

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s)) \psi(s) \, ds + \int_0^t f'(X(s)) \phi(s) \, dB(s)$$

$$+ \frac{1}{2} \int_0^t f''(X(s)) |\phi(s)|^2 \, ds$$

almost surely, or equivalently

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s)) \, dX(s)$$

$$+ \frac{1}{2} \int_0^t f''(X(s)) \frac{d}{dt} \langle X \rangle(s) \, ds$$
Proof (sketch) the identity holds for the identity function $f(x) = x$

if the identity holds for the function $f(x)$, then it holds for the function $g(x) = x f(x)$: indeed, note that $g'(x) = f(x) + x f'(x)$ and $\frac{1}{2} g''(x) = f'(x) + \frac{1}{2} x f''(x)$, note also that $X(t)$ and

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s)) dX(s) + \frac{1}{2} \int_0^t f''(X(s)) d\langle X \rangle(s)$$

are two Itô processes, hence using the integration by parts formula yields

$$X(t) f(X(t)) = X(0) f(X(0)) + \int_0^t f(X(s)) dX(s) + \int_0^t X(s) \left[ f'(X(s)) dX(s) + \frac{1}{2} f''(X(s)) d\langle X \rangle(s) \right] + \int_0^t f'(X(s)) d\langle X \rangle(s)$$
collecting terms yields

\[ X(t) f(X(t)) = X(0) f(X(0)) + \int_0^t \left[ f(X(s)) + X(s) f'(X(s)) \right] dX(s) \]

\[ + \int_0^t \left[ f'(X(s)) + \frac{1}{2} f''(X(s)) \right] d\langle X \rangle(s) \]

or in other words

\[ g(X(t)) = g(X(0)) + \int_0^t g'(X(s)) dX(s) + \frac{1}{2} \int_0^t g''(X(s)) d\langle X \rangle(s) \]

i.e. the identity holds for the function \( g(x) = x f'(x) \)

by induction, the identity holds for any polynomial function, and by density, it holds for any twice continuously differentiable function.
multi–dimensional version

Theorem 10* let $B$ be a $d$–dimensional standard Brownian motion, let $X$ be a $m$–dimensional Itô process and let $f$ be a twice continuously differentiable function, then

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s)) \psi(s) \, ds + \int_0^t f'(X(s)) \phi(s) \, dB(s)$$

$$+ \frac{1}{2} \int_0^t \text{trace}[f''(X(s)) \phi(s) \phi^*(s)] \, ds$$

almost surely, or equivalently

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s)) \, dX(s)$$

$$+ \frac{1}{2} \int_0^t \text{trace}[f''(X(s)) \frac{d}{dt} \langle X \rangle(s)] \, ds$$
multi–dimensional and time–dependent version

**Theorem 11** * let $B$ be a $d$–dimensional standard Brownian motion, let $X$ be a $m$–dimensional Itô process and let the function $f$ be differentiable w.r.t. time and twice continuously differentiable w.r.t. space, then

$$f(t, X(t)) = f(0, X(0)) + \int_0^t \left[ \frac{\partial f}{\partial t}(s, X(s)) + f'(s, X(s)) \psi(s) \right] ds$$

$$+ \int_0^t f'(s, X(s)) \phi(s) dB(s)$$

$$+ \frac{1}{2} \int_0^t \text{trace}[f''(s, X(s)) \phi(s) \phi^*(s)] ds$$

almost surely, or equivalently

$$f(X(t)) = f(X(0)) + \int_0^t \frac{\partial f}{\partial t}(s, X(s)) ds + \int_0^t f'(s, X(s)) dX(s)$$

$$+ \frac{1}{2} \int_0^t \text{trace}[f''(s, X(s)) \frac{d}{dt} \langle X \rangle(s)] ds$$
Theorem 12 * [Burkholder–Davis–Gundy inequalities] let $B$ be a one–dimensional standard Brownian motion, and for any $\phi \in M^2([0, T])$ define

$$M(t) = \int_0^t \phi(s) dB(s) \quad M_*(t) = \max_{0 \leq s \leq t} |M(s)|$$

$$A(t) = \langle M \rangle(t) = \int_0^t |\phi(s)|^2 \, ds$$

for any $p \geq 2$, there exist positive constants $0 < c_p \leq C_p < \infty$ such that for any $0 \leq t \leq T$

$$c_p \, \mathbb{E}|A(t)|^{p/2} \leq \mathbb{E}|M_*(t)|^p \leq C_p \, \mathbb{E}|A(t)|^{p/2}$$

Remark uniform control in terms of the increasing quadratic variation process