PROBABILITY THEORY

- <u>frequency limits</u>

$$p_i = \lim_{N\to\infty} \frac{N_i}{N}$$

<u>eg</u> face i occurs N_i times out of N tosses of a die

<u>intuitive</u> — but difficult to develop theoretically

- <u>probability spaces</u> contain all necessary probabilistic information — assumed given

$$(\Omega, \mathcal{A}, P)$$

<u>sample space</u>
of basic outcomes
· nonempty set

<u>probability measure</u>
$P : \mathcal{A} \to [0,1]$

① $P(\phi) = 0$

② $P(\Omega) = 1$

③ $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ if $A_i \cap A_j = \phi$, $i \neq j$

<u>σ-algebra of events</u>
class of subsets of Ω

① $\phi, \Omega \in \mathcal{A}$

② $A^c = \Omega \setminus A \in \mathcal{A}$ if $A \in \mathcal{A}$

③ $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ if $A_1, A_2, \dots \in \mathcal{A}$

RANDOM VARIABLES

$$X : \Omega \longrightarrow \mathbb{R}$$ on probability space $(\Omega, \mathcal{A}, P)$

- realization
  sample $X(\omega)$ provides numerical information about outcome $\omega \in \Omega$

- X must be measurable — information content is compatible with that in the probability space

$$X^{-1}(a) := \{\omega \in \Omega;\ X(\omega) \leq a\} \in \mathcal{A} \quad \text{for all } a \in \mathbb{R}$$

- distribution function of a random variable X $\quad F_X : \mathbb{R} \longrightarrow \mathbb{R}$

$$F_X(a) = P\left(X^{-1}(a)\right) = P\left(\{\omega \in \Omega;\ X(\omega) \leq a\}\right)$$

for all $a \in \mathbb{R}$

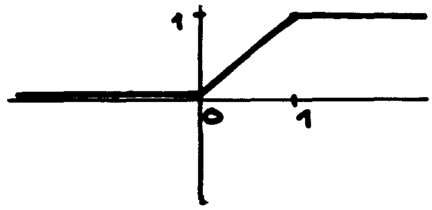
- more convenient than measures
- usually the distribution function is given directly and the probability space is either not known or not stated explicitly

$$P\left(\{\omega \in \Omega;\ a < X(\omega) \leq b\}\right) = F_X(b) - F_X(a)$$

• <u>continuous random variables</u> continuum of possible values

① X uniformly distributed on $[0,1]$

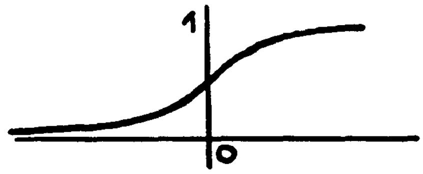
$$F_X(a) = \begin{cases} 0 & \text{if } a \leq 0 \\ a & \text{if } 0 \leq a \leq 1 \\ 1 & \text{if } 1 \leq a \end{cases}$$



Hence $X \in [a,b] \subseteq [0,1]$ with probability $b-a$

② $X \sim N(0;1)$ standard Gaussian

$$F_X(a) = \int_{-\infty}^{a} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-x^2/2}}_{} dx$$



<u>density function</u> $p(x) = \frac{dF_X}{dx}(x)$

• <u>discrete random variables</u> only finitely or countably many possible values $x_1, x_2, \dots, x_n, \dots$

easier to work directly with probabilities

$$p_i = P(\{\omega \in \Omega;\ X(\omega) = x_i\}) \qquad \sum_{i \geq 1} p_i = 1$$

③ X is 2-point distributed with values ± 1

$$P(X = -1) = 1/2, \quad P(X = +1) = 1/2$$

Moments of Random Variables

• characterise salient features of the variability of a R.V.

<u>expectation</u>
<u>expected value</u>
<u>mean</u>

$$\mu = E(X) = \begin{cases} \sum_{i\geq 1} x_i p_i & \text{discrete case} \\ \int_{-\infty}^{\infty} x\, p(x)\, dx & \text{continuous case with density} \end{cases}$$

• higher moments indicate the scatter about the mean

<u>pth moment</u>

$$E(X^p)$$

<u>pth centered moment</u>

$$E\left((X-\mu)^p\right)$$

<u>Variance</u> $\quad Var(X) = E\left((X-\mu)^2\right) = \sigma^2$

σ = standard deviation

$Var(X) = E(X^2) - \mu^2$
too, better for calculations

<u>Examples</u>

1. uniform $[0,1]$ $\quad \mu = 1/2, \quad \sigma^2 = 1/12$
2. $N(0;1)$ Gaussian $\quad \mu = 0, \quad \sigma^2 = 1$
3. 2-point $X = \pm 1$ $\quad \mu = 0, \quad \sigma^2 = 1$

<u>Note</u> – general Gaussian $X \sim N(\mu;\sigma^2)$ with mean μ and variance σ^2 has higher centered moments

<u>odd</u> $\quad E\left((X-\mu)^{2n+1}\right) = 0$

<u>even</u> $\quad E\left((X-\mu)^{2n}\right) = 1\cdot 3 \cdots (2n-1)\cdot \sigma^{2n}$

If X_1, X_2 are random variables on same $(\Omega, \mathcal{A}, P)$

<u>in general</u> $\begin{cases} E(X_1+X_2) = E(X_1) + E(X_2) \\ Var(X_1+X_2) \neq Var(X_1) + Var(X_2) \end{cases}$

$\uparrow =$ only in special cases

<u>covariance</u> $\quad cov(X_1, X_2) = E\big((X_1-\mu_1)(X_2-\mu_2)\big)$

$$= E(X_1 X_2) - \mu_1\mu_2$$

* X_1 and X_2 are <u>independent</u> if for all $a, b \in \mathbb{R}$

$F_{X_1,X_2}(a,b) := P(\{\omega\in\Omega;\ X_1(\omega)\leq a$ <u>and</u> $X_2(\omega)\leq b\})$

joint distribution function

$$= P(\{\omega\in\Omega;\ X_1(\omega)\leq a\})\cdot P(\{\omega\in\Omega;\ X_2(\omega)\leq b\})$$

$$= F_{X_1}(a)\cdot F_{X_2}(b)$$

individual distribution functions

<u>Hence</u> for independent random variables

$$E(X_1 X_2) = E(X_1)\, E(X_2)$$

$$Var(X_1+X_2) = Var(X_1) + Var(X_2)$$

<u>ie</u> means multiply, variances add

Convergence of Random Variables

- Let $X_1, X_2, \ldots, X_n, \ldots$ and $\overline{X}$ be random variables on a common probability space $(\Omega, \mathcal{A}, P)$

<u>Convergence</u> $X_n \longrightarrow \overline{X}$ as $n \longrightarrow \infty$?

There are several useful types of convergence

① <u>convergence with probability 1</u> (w.p.1)

$$X_n(\omega) \to \overline{X}(\omega) \text{ in } \mathbb{R} \text{ for all } \omega \in A, \quad P(A) = 1$$

② <u>convergence in pth-mean</u>

$$E\left(|X_n - \overline{X}|^p\right) \to 0$$

$p=1$, <u>convergence in mean</u>
$p=2$, <u>mean-square convergence</u>

(assumes pth moments exist)

③ <u>convergence in probability</u>

$$P\left(\{\omega \in \Omega ;\ |X_n(\omega) - \overline{X}(\omega)| \geq \varepsilon\}\right) \longrightarrow 0 \quad \text{for all } \varepsilon > 0$$

④ <u>convergence in distribution</u>

$$F_{X_n}(a) \longrightarrow F_{\overline{X}}(a)$$ for all $a \in \mathbb{R}$ where $F_{\overline{X}}$ is continuous.

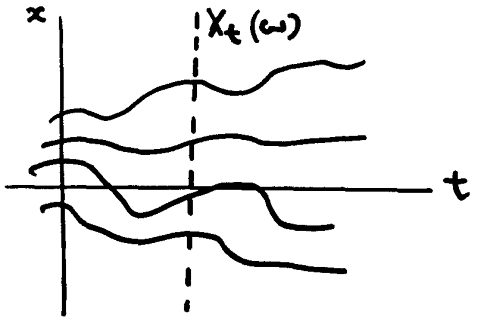
Very roughly + under special assumptions — convergences are progressively weaker

STOCHASTIC PROCESSES

time set $\mathbb{T} \subseteq \mathbb{R}$ probability space $(\Omega, \mathcal{A}, P)$

$X : \mathbb{T} \times \Omega \longrightarrow \mathbb{R}$ write $X_t(\omega)$ for $X(t,\omega)$

X is a stochastic process if $X_t : \Omega \to \mathbb{R}$ is a random variable for each $t \in \mathbb{T}$



sample path realization $X_\cdot(\omega) : \mathbb{T} \longrightarrow \mathbb{R}$

ω fixed

There are many possible types of time dependence eg

1. independent X_s, X_t indept. if $s \neq t$
2. identically distributed $F_{X_t}(x) \equiv F(x)$ for all $t \in \mathbb{T}$
3. independent increments $X_{\tau_2} - X_{\tau_1}, X_{\tau_4} - X_{\tau_3}$ indept. $\tau_1 \le \tau_2 \le \tau_3 \le \tau_4$
4. Markovian   future depends only on present, not both present and past

discrete time stochastic processes   eg $\mathbb{T} = \{t_1, t_2, \ldots, t_n, \ldots\}$
sequence of RV's $X_{t_1}, X_{t_2}, \ldots, X_{t_n}, \ldots$

continuous time stochastic processes   eg $\mathbb{T} = [0,T]$, $[0,\infty)$

- generally complicated, subtle measure theoretic arguments required unless regularity of sample paths is assumed sample path continuous stochastic processes

sample paths $X_\cdot(\omega) : \mathbb{T} \to \mathbb{R}$ are continuous functions of $t \in \mathbb{T}$ for all $\omega \in A$, $P(A) = 1$

transition probabilities $\quad s \leq t$, measurable $B \subseteq \mathbb{R}$

$$P(s, x; t, B) = \int_B \underbrace{p(s,x;t,y)}_{\text{transition density}} \, dy = P\left(\{\omega \in \Omega ; X_t(\omega) \in B\} \,\middle|\, X_s = x\right)$$

measure on a σ-algebra of $\mathbb{R}$ subsets for each $s \leq t$, $x \in \mathbb{R}$

conditional probability

A stochastic process $X : [0,T] \times \Omega \to \mathbb{R}$ is a diffusion process with drift $a(s,x)$ and diffusion coefficient $b(s,x)$ if its transition densities satisfy for all $s \in [0,T]$, $x \in \mathbb{R}$ and $\varepsilon > 0$

① $$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|>\varepsilon} p(s,x;t,y)\,dy = 0$$ no instantaneous jumps

② $$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|<\varepsilon} (y-x)\, p(s,x;t,y)\,dy = a(s,x)$$ drift

③ $$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|<\varepsilon} (y-x)^2\, p(s,x;t,y)\,dy = b^2(s,x)$$ squared diffusion coefficient

diffusion processes — Markovian, sample path continuous, transition densities satisfy Kolmogorov PDEs

forward $$\frac{\partial p}{\partial t} + \frac{\partial}{\partial y}\{a(t,y)\,p\} - \frac{1}{2}\frac{\partial^2}{\partial y^2}\{b^2(t,y)\,p\} = 0$$ (s,x) fixed

backward $$\frac{\partial p}{\partial s} + a(s,x)\frac{\partial p}{\partial x} + \frac{1}{2} b^2(s,x) \frac{\partial^2 p}{\partial x^2} = 0$$ (t,y) fixed

Fokker-Planck Equation

Wiener Processes

- simplest, prototype diffusion process describing physically observed <u>Brownian motion</u>

- <u>drift</u> $a(s,x) \equiv 0$ <u>diffusion coefficient</u> $b(s,x) \equiv 1$

- <u>transition densities</u> $$p(s,x;t,y) = \frac{1}{\sqrt{2\pi(t-s)}}\, e^{-\frac{(y-x)^2}{2(t-s)}}$$

heat equatio

<u>Standard Wiener process</u> $W : [0,\infty) \times \Omega \to \mathbb{R}$

1. $W_0 = 0$ w.p.1

2. $W_t - W_s \sim N(0; t-s)$ $\left\{ \begin{array}{ll} E(W_t - W_s) = 0 & E(W_t) = 0 \\ E\left((W_t - W_s)^2\right) = t-s & E(W_t^2) = t \end{array} \right.$

3. independent increments

   $W_{t_2} - W_{t_1}$, $W_{t_4} - W_{t_3}$ independent $0 \le t_1 \le t_2 \le t_3 \le t$

4. sample path continuous

<u>BUT</u> sample paths are <u>NOT</u> <u>differentiable</u> anywhere

<u>hint</u> mean-square convergence $$E\left(\left(\frac{W_t - W_s}{t-s}\right)^2\right) = \frac{E\left((W_t - W_s)^2\right)}{(t-s)^2} = \frac{t-s}{(t-s)^2} = \frac{1}{t-s}$$

They are <u>not</u> even of <u>bounded variation</u> on $[0,T]$

$0 \qquad T$

$t_0 \; t_1 \cdots t_j \cdots t_N$

<u>any partition</u>

$$\sup_N \sum_{j=0}^{N-1} \left| W_{t_{j+1}}(\omega) - W_{t_j}(\omega) \right| = +\infty$$

w.p.1

General case: the **stochastic differential equation**

$$dX_t(\omega) = a\big(t, X_t(\omega)\big)\,dt + b\big(t, X_t(\omega)\big)\,dW_t(\omega)$$

is symbolic for the **stochastic integral equation**

$$X_t(\omega) = X_{t_0}(\omega) + \underbrace{\int_{t_0}^{t} a\big(s, X_s(\omega)\big)\,ds}_{\substack{\text{deterministic}\\ \text{integral for}\\ \text{each } \omega\in\Omega}} + \underbrace{\int_{t_0}^{t} b\big(s, X_s(\omega)\big)\,dW_s(\omega)}_{\substack{\text{stochastic}\\ \text{integral ?}}}$$

deterministic Riemann-Stieltjes integral

$$\int_0^T f(s)\,dR(s) = \lim_{N\to\infty} \sum_{j=0}^{N-1} f(\tau_j)\big\{R(t_{j+1}) - R(t_j)\big\}$$

exists if and only if R has bounded variation on $[0,T]$

arbitary $\tau_j \in [t_j, t_{j+1}]$

robust !

A stochastic integral cannot be a Riemann-Stieltjes integral for each ω

Ito stochastic integral

$$\int_0^T f(s,\omega)\,dW_s(\omega) = \underset{N\to\infty}{\text{m.s.-lim}} \sum_{j=0}^{N-1} f(t_j,\omega)\big\{W_{t_{j+1}}(\omega) - W_{t_j}(\omega)\big\}$$

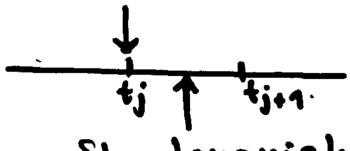
always evaluated at the beginning of each partion subinterval

admissible integrands

- $E\big(f^2(t,\cdot)\big) < \infty$
- $f(t,\cdot)$ **nonanticipative**

$\updownarrow$

independent of $W_\tau - W_t$ for all $\tau > t$.



Stratonovich integral uses mid-point

Sample Path Dynamics of a Diffusion Process

- **Langevin** ~1910
  - linear drift $a(t,x) = -a \cdot x$
  - constant diffusion coefficient $b(t,x) \equiv b$

$$\frac{d}{dt} X_t(\omega) = -a X_t(\omega) + b \, \xi_t(\omega)$$

$\uparrow$ Gaussian white noise

- $\xi_t \sim N(0;1)$ with ξ_s, ξ_t independent for $s \neq t$
- $\operatorname{cov}(\xi_s \xi_t) = c \cdot \delta(t-s)$ Dirac delta function
- flat spectral density "white"

ie $\xi_t(\omega) = \frac{dW_t}{dt}(\omega)$! ξ_t cannot exist as a normal function

Symbolically with $\xi_t \, dt = dW_t$ Langevin's equation

$$dX_t(\omega) = -a X_t(\omega) \, dt + b \, dW_t(\omega)$$

is meaningful as a stochastic integral equation

$$X_t(\omega) = X_0(\omega) - \underbrace{\int_0^t a X_s(\omega) \, ds}_{\substack{\text{deterministic} \\ \text{integral for} \\ \text{each } \omega \in \Omega}} + b \, W_t(\omega)$$

$\uparrow \int_0^t b \, dW_s(\omega)$

STOCHASTIC DIFFERENTIAL EQUATIONS

$$dX_t = a(t, X_t)\,dt + b(t, X_t)\,dW_t$$

interpreted mathematically as stochastic integral equation

$$X_t = X_{t_0} + \underbrace{\int_{t_0}^{t} a(s, X_s)\,ds}_{\text{Riemann integral for each path}} + \underbrace{\int_{t_0}^{t} b(s, X_s)\,dW_s}_{\text{Ito stochastic integral}}$$

Assumptions

Ⓐ coefficent functions $a, b : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$

• continuous

• lipschitz condition
$$|a(t,x) - a(t,y)| \leq K|x-y|$$
$$|b(t,x) - b(t,y)| \leq K|x-y|$$

• linear growth condition
$$|a(t,x)| + |b(t,x)| \leq K\sqrt{1+|x|^2}$$

Ⓑ initial value X_{t_0}, $0 \leq t_0 < T$

• nonanticipative • $E(X_{t_0}^2) < \infty$

Conclusions there exists $X : [t_0, T] \times \Omega \longrightarrow \mathbb{R}$

ⓐ unique w.p.1

ⓑ X_t nonanticipative

ⓒ $E(X_t^2) < \infty$

ⓓ sample path continuous

ⓔ diffusion process with coefficients a, b

Proof — method of successive approximations

STOCHASTIC EULER SCHEME

SDE $$dX_t = a(t, X_t)\,dt + b(t, X_t)\,dW_t \qquad t_0 \le t \le T$$

time partition 

time step $\Delta_n = t_{n+1} - t_n$

$$X_{t_{n+1}} = X_{t_n} + \int_{t_n}^{t_{n+1}} a(s, X_s)\,ds + \int_{t_n}^{t_{n+1}} b(s, X_s)\,dW_s$$

assume $X_{t_n} = Y_n$ and write

$$Y_{n+1} = Y_n + a(t_n, Y_n)\underbrace{\int_{t_n}^{t_{n+1}} ds}_{\boxed{\Delta_n = t_{n+1} - t_n}} + b(t_n, Y_n)\underbrace{\int_{t_n}^{t_{n+1}} dW_s}_{\boxed{\Delta W_n = W_{t_{n+1}} - W_{t_n}}}$$

stochastic Euler scheme

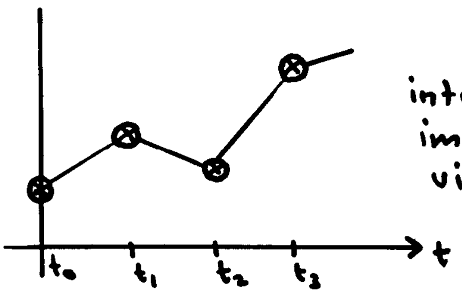
$$\boxed{Y_{n+1} = Y_n + a(t_n, Y_n)\,\Delta_n + b(t_n, Y_n)\,\Delta W_n}$$

consistent with Ito integral $\qquad n = 0, 1, 2, \dots$

For a given noise sample path $W_t(\omega)$ and initial $Y_0(\omega)$

$$Y_0(\omega) \xrightarrow{\Delta W_0(\omega)} Y_1(\omega) \xrightarrow{\Delta W_1(\omega)} Y_2(\omega) \longrightarrow \dots$$

ie generates the corresponding sample path of the sequence of random variables Y_n, $n = 1, 2, \dots$



interpolate to improve visualization

$Y_n(\omega)$ is supposed to approximate $X_{t_n}(\omega)$ here

Generating noise increments

$$\Delta W_n = W_{t_{n+1}} - W_{t_n} \sim N(0;\Delta_n) \qquad \Delta_n = t_{n+1} - t_n$$

ie $$\Delta W_n = G_n \sqrt{\Delta_n} \quad \text{where } G_n \sim N(0;1)$$

standard Gaussian

Box-Muller transformation

U_n, U_n' independent, uniformly distributed on $[0,1]$

$$G_n = \sqrt{-2\ln(U_n)} \cdot \cos(2\pi U_n')$$

$$G_n' = \sqrt{-2\ln(U_n)} \cdot \sin(2\pi U_n')$$

$\Rightarrow$ G_n, G_n' independent, $N(0;1)$ distributed

(**Polar-Marsaglia method** avoids trig functions, but discards 25% of values — still very efficient)

Pseudo-random number generators

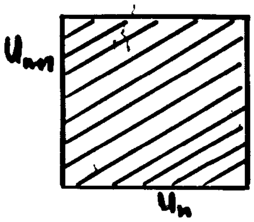
- provide "independent" U_n uniformly distributed on $[0,1]$

linear congruential generator $$X_{n+1} = aX_n + b \pmod c$$

eg $a = 7^5$, $b = 0$, $c = 2^{31}-1$

$$U_n = X_n / c \in [0,1]$$

Note successive (U_n, U_{n+1}) lie on lines of slope a/c in unit square $[0,1]^2$



For appropriate parameters a,b,c these fill square quite densely and uniformly

U_n — not perfect, but good for most purposes
— reproducible

linear SDE $\qquad dX_t = aX_t\,dt + bX_t\,dW_t \qquad 0 \le t \le T$

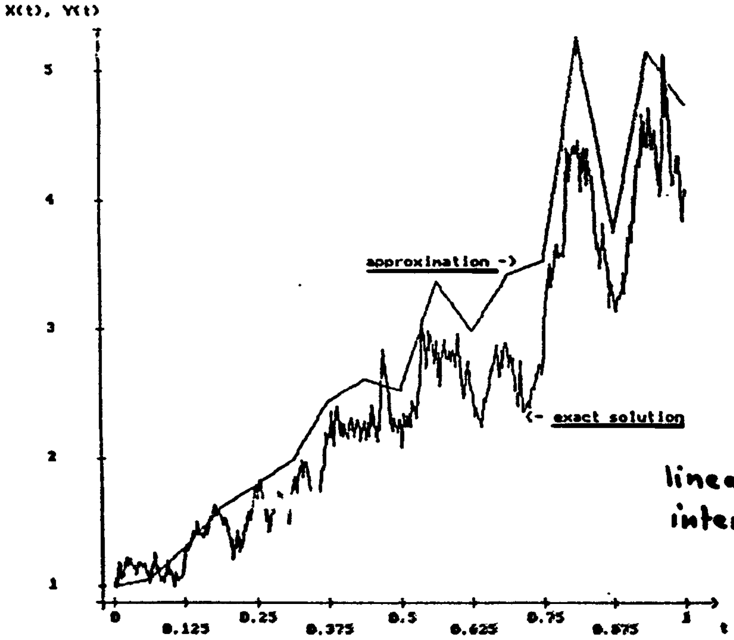
Euler scheme $\qquad Y_{n+1} = Y_n + aY_n\Delta_n + bY_n\Delta W_n$

$a = 1.5 \qquad b = 1.0$

$X_0 = Y_0 = 1.0$

equal $\Delta_n \equiv 2^{-4}$

$t_n = n\cdot 2^{-n} \qquad T = 1.0$



linearly interpolated

exact solution $\qquad X_t = X_0 \exp\left\{\left(a - \frac{1}{2}b^2\right)t + bW_t\right\}$

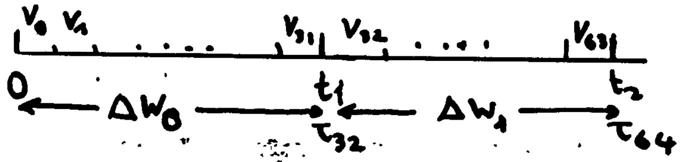
• picture shows X_{τ_j} for $\tau_j = j\cdot 2^{-9}$ — smallest timestep for resolution of PC-screen

$$W_{\tau_j} = \sum_{i=0}^{j-1} V_i \qquad V_i \sim N(0; 2^{-9})$$

• for **same sample path** in Euler scheme note $W_{t_n} = W_{\tau_{n\cdot 2^5}}$

$$\Delta W_n = \sum_{i=n\cdot 2^5}^{(n+1)2^5-1} V_i$$

independent variances add !



Accuracy and Convergence

$0 \longmapsto T$ equal time steps $\Delta_n \equiv \delta = T/N_T$

exact solution X_T Euler iterate $Y^{\delta}_{N_T}$

do the $Y^{\delta}_{N_T} \longrightarrow X_T$ as $\delta \to 0$?   in what sense?
how fast?

- there are several different useful types of convergence
  — appropriate choice depends on purpose for which an approximation is required.

good sample path approximations

$$\varepsilon(\delta) = E\left(\left|Y^{\delta}_{N_T} - X_T\right|\right) \longrightarrow 0$$

strong convergence

good distributional approximations
eg moments

$$\mu_1(\delta) = \left|E\left(Y^{\delta}_{N_T}\right) - E\left(X_T\right)\right| \longrightarrow 0$$

weak convergence
more general definition later

$$\mu_1(\delta) \leq \varepsilon(\delta)$$

order γ strong convergence $\varepsilon(\delta) \leq K\,\delta^{\gamma}$

order β weak convergence $\mu_1(\delta) \leq K_1 \delta^{\beta}$

* usually the order is for theoretical discretization error
  - roundoff error
  - pseudo-random number error
  - finite sampling error

dominant

Finite sampling and confidence intervals

$\nu = E(Z)$ expectation

$\hat{\nu} = \frac{1}{L} \sum_{j=1}^{L} Z(\omega_j)$ arithmetic average of finite sample $Z(\omega_1), \dots, Z(\omega_L)$

how well does the random variable $\hat{\nu}$ estimate ν ?

Take M batches of N samples each

$$Z(\omega_{k,j}) \qquad \begin{array}{l} k = 1, \dots, M \\ j = 1, \dots, N \end{array}$$

batch average $\hat{\nu}_k = \frac{1}{N} \sum_{j=1}^{N} Z(\omega_{k,j})$ approximately Gaussian if $N \gg 1$ (LLN)

overall average $\hat{\nu} = \frac{1}{M} \sum_{k=1}^{M} \hat{\nu}_k = \frac{1}{MN} \sum_{k=1}^{M} \sum_{j=1}^{N} Z(\omega_{k,j})$

sample variance of batch averages $\hat{\sigma}_\nu^2 = \frac{1}{M-1} \sum_{k=1}^{M} (\hat{\nu}_k - \hat{\nu})^2$

unbiased

Student t-distribution $t_{1-\alpha, M-1}$ $\alpha \in (0, 1)$, M-1 degrees of freedom

$$\Delta \hat{\nu} = t_{1-\alpha, M-1} \sqrt{\frac{\hat{\sigma}_\nu^2}{M}}$$

100 $(1-\alpha)$% confidence interval

$$(\hat{\nu} - \Delta\hat{\nu}, \hat{\nu} + \Delta\hat{\nu})$$

random interval depending on sample used

true expectation ν lies in this confidence interval **with at least probability** $1-\alpha$

$\boxed{\text{sampling error} \sim 1/\sqrt{M}}$ need $M \gg 1$

Strong convergence

linear SDE $\quad dX_t = aX_t\,dt + bX_t\,dW_t$

Euler scheme $\quad Y_{n+1} = Y_n + aY_n\,\Delta_n + bY_n\,\Delta W_n$

$0 \longmapsto T$ equal steps $\Delta_n \equiv \delta = T/N_T$

strong error $\quad \varepsilon(\delta) = E\left(\left|Y^{\delta}_{N_T} - X_T\right|\right)$

error estimate $\quad \hat{\varepsilon}(\delta) = \frac{1}{MN}\sum_{k=1}^{M}\sum_{j=1}^{N}\left|Y^{\delta}_{N_T}(\omega_{k,j}) - X_T(\omega_{k,j})\right|$

M batches
N samples each

exact solution $\quad X_T = X_0 \exp\left\{\left(a - \frac{1}{2}b^2\right)T + bW_T\right\}$

$a = 1.5, \quad b = 1.0$

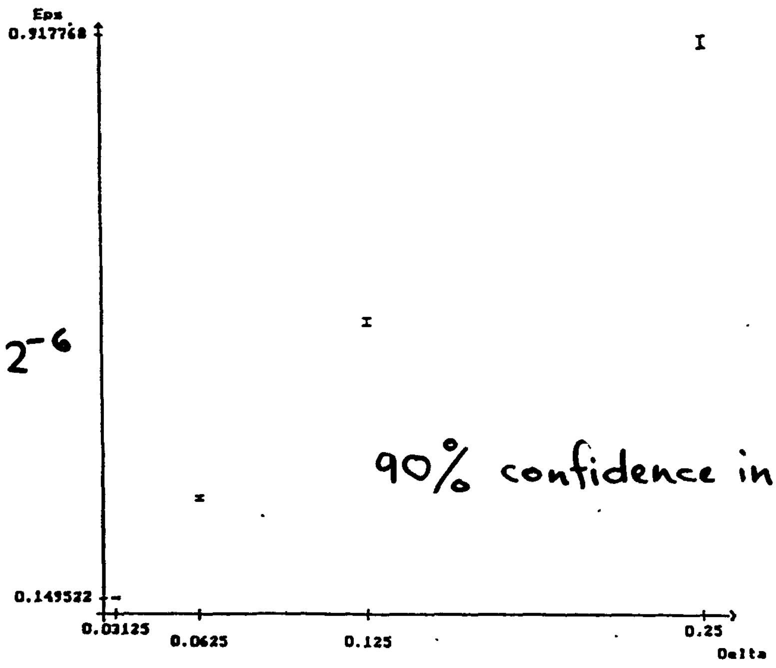
$X_0 = Y_0 = 1.0$

$T = 1, \quad \delta = 2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}$

$M = 20, \quad N = 100$

$\alpha = 0.1$

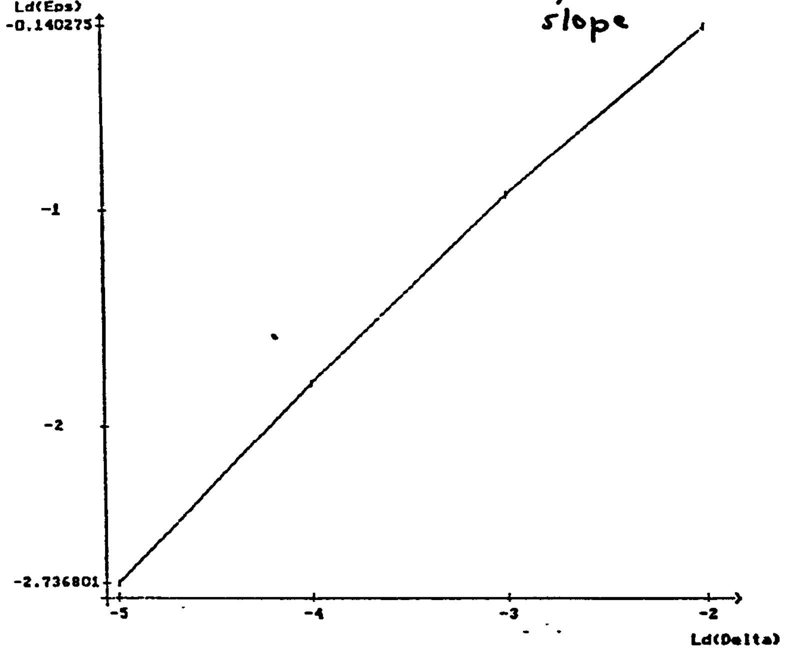
$t_{0.9,19} = 1.73$



90% confidence interva

$$\varepsilon = K\delta^{\gamma} \Rightarrow \log\varepsilon = \log K + \gamma \log\delta$$

slope



"slope" $\gamma \simeq 0.8$

includes <u>all</u> errors not just theoretical discretization error

Theoretically for a general SDE the Euler scheme has strong order

$$\boxed{\gamma = 0.5}$$

Can do better in special cases

$$dX_t = aX_t\,dt + b\,dW_t$$

additive noise

$\gamma = 1.0$ here

<u>no noise</u> $b \equiv 0$

* reduces to deterministic Euler scheme for ODEs with global discretization error.

$$|Y^{\delta}_{N_T} - X_T| \leq K\delta^1$$

<u>ie</u> order $\gamma = 1.0$

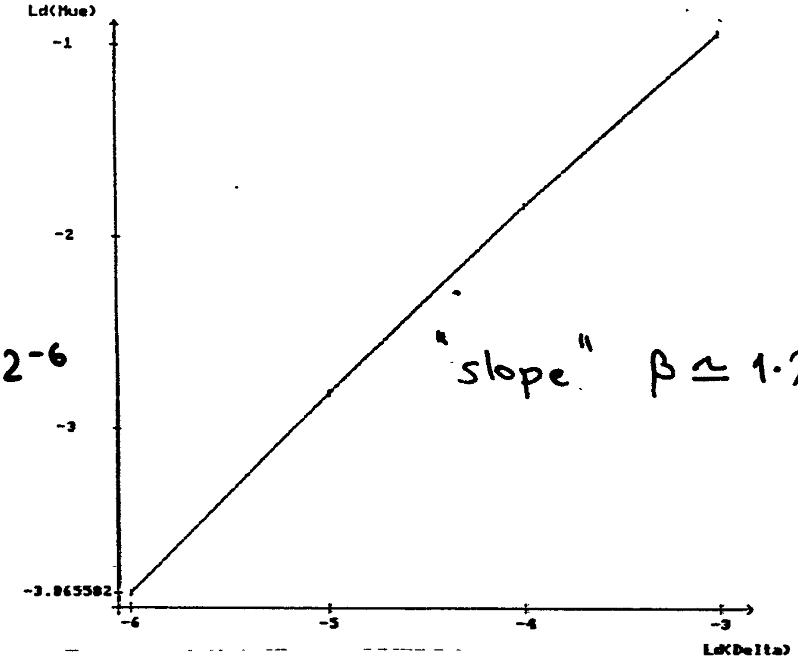
Weak convergence

weak error $\mu_1(\delta) = \left| E(Y^{\delta}_{N_T}) - E(X_T) \right|$

error estimate $\hat{\mu}_1(\delta) = \left| \frac{1}{MN} \sum_{k=1}^{M} \sum_{j=1}^{N} Y^{\delta}_{N_T}(\omega_{k,j}) - E(X_T) \right|$

$$E(X_T) = E(X_0)\, e^{(a - 1/2\, b^2)T}$$

$a = 1.5, \quad b = 1.0$
$X_0 = Y_0 = 1.0$
$T = 1; \quad \delta = 2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}$
$M = 20, \quad N = 100$



"slope" $\beta \simeq 1.20$

general weak convergence criterion $\mu_g(\delta) = \left| E(g(Y^{\delta}_{N_T})) - E(g(X_T)) \right|$

includes all moments

g · continuous
· polynomial growth

order β weak convergence $\mu_g(\delta) \leq K_g\, \delta^{\beta}$

same β for all g!

For a general SDE the stochastic Euler scheme has weak order $\boxed{\beta = 1.0}$

Higher order schemes

heuristic adaptations of higher order deterministic schemes are usually inconsistent or only low order

linear SDE $\quad dX_t = a X_t\, dt + b X_t\, dW_t$

Heun scheme $\quad Y_{n+1} = Y_n + \frac{1}{2}\left(a Y_n + a \Upsilon_n\right)\Delta_n + \frac{1}{2}\left(b Y_n + b \Upsilon_n\right)\Delta W_n$

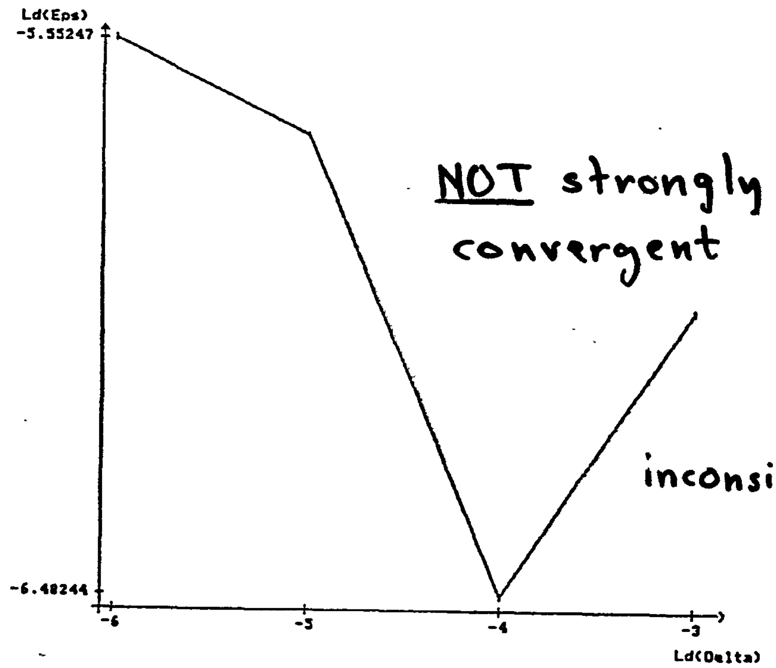
$$\Upsilon_n = Y_n + a Y_n \Delta_n + b Y_n \Delta W_n$$

$a = 1.5\,, \quad b = 1.0$

$X_0 = Y_0 = 1.0$

$T = 1$

$\delta = 2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}$



NOT strongly convergent

inconsistent

- stochastic calculus is less robust than deterministic calculus, so more care is needed in deriving schemes

- higher order convergence requires more information about the noise changes <u>within</u> the discretization subintervals than is contained in the simple noise increments ΔW_n

Stratonovich Stochastic Differential Equations

$$dX_t = a(t, X_t)\,dt + b(t, X_t) \circ dW_t$$

"$\circ$" denotes Stratonovich

$$X_t = X_{t_0} + \int_{t_0}^{t} a(s, X_s)\,ds + \underbrace{\int_{t_0}^{t} b(s, X_s) \circ dW_s}_{\text{Stratonovich stochastic integral}}$$

evaluate at midpoint

Generally Ito and Stratonovich SDEs with same coefficient do <u>not</u> have same solutions

<u>Ito</u> $\quad dX_t = aX_t\,dt + bX_t\,dW_t \qquad X_t = X_0\, e^{(a - \frac{1}{2}b^2)t + bW_t}$

<u>Strat.</u> $\quad dX_t = aX_t\,dt + bX_t \circ dW_t \qquad X_t = X_0\, e^{at + bW_t}$

<u>drift correction</u> $\qquad \underline{a}(t,x) = a(t,x) - \frac{1}{2} b(t,x) \frac{\partial b}{\partial x}(t,x)$

with drift correction Ito and Stratonovich SDEs have the same solution

<u>equivalent SDEs</u>

$$dX_t = a(t, X_t)\,dt + b(t, X_t)\,dW_t \qquad \text{Ito}$$

$$dX_t = \underline{a}(t, X_t)\,dt + b(t, X_t) \circ dW_t \qquad \text{Stratonovich}$$

<u>example</u> $\qquad a(t,x) = ax, \quad b(t,x) = bx, \quad \underline{a}(t,x) = (a - \frac{1}{2}b^2)x$.

$$\left.\begin{aligned} &\underline{\text{Ito}} \quad dX_t = aX_t\,dt + bX_t\,dW_t \\ &\underline{\text{Strat.}} \quad dX_t = (a - \tfrac{1}{2}b^2)X_t\,dt + bX_t \circ dW_t \end{aligned}\right\} \quad X_t = X_0\, e^{(a - \frac{1}{2}b^2)t + bW_t}$$

equivalent SDEs

Stratonovich stochastic calculus

- same chain rule as deterministic calculus, so can solve Stratonovich SDEs by same methods as for deterministic ODEs

$$dX_t = aX_t\,dt + bX_t \circ dW_t$$

$$\frac{dX}{X} = a\,dt + b\,dW$$

$$\ln\frac{X_t}{X_0} = \int_{X_0}^{X_t}\frac{dX}{X} = \int_0^t a\,ds + \int_0^t b\,dW_s = at + bW_t$$

$$\Rightarrow X_t = X_0\, e^{at+bW_t}$$

BUT Stratonovich calculus does not have the same direct link to diffusion process theory or to martingale theory as Ito calculus

mathematical proofs thus much harder

Both Ito and Stratonovich calculi are correct mathematically

easy to switch from one to the other using drift correction

Which one to use is a modelling issue

- what is true nature of noise in model
- rough rules of thumb available, but ...

STOCHASTIC CALCULUS

Ito SDE $\quad dX_t = a(t,X_t)\,dt + b(t,X_t)\,dW_t$

$$Y_t = f(t,X_t)$$

stochastic chain rule

Ito formula $\quad dY_t = L^0 f(t,X_t)\,dt + L^1 f(t,X_t)\,dW_t$

$$L^0 = \frac{\partial}{\partial t} + a\frac{\partial}{\partial x} + \underbrace{\frac{1}{2}b^2\frac{\partial^2}{\partial x^2}}_{\text{extra term due to } E(\Delta W)^2 = \Delta} \qquad L^1 = b\frac{\partial}{\partial x}$$

Stochastic Taylor expansions

iterated application of Ito formula

Ito SDE $\quad X_t = X_{t_0} + \int_{t_0}^{t} \underbrace{a(s,X_s)}\,ds + \int_{t_0}^{t} \underbrace{b(s,X_s)}\,dW_s$

$f = a \qquad a(s,X_s) = a(t_0,X_{t_0}) + \int_{t_0}^{s} L^0 a(\tau,X_\tau)\,d\tau + \int_{t_0}^{s} L^1 a(\tau,X_\tau)\,dW_\tau$

$f = b \qquad b(s,X_s) = b(t_0,X_{t_0}) + \int_{t_0}^{s} L^0 b(\tau,X_\tau)\,d\tau + \int_{t_0}^{s} L^1 b(\tau,X_\tau)\,dW_\tau$

$$X_t = X_{t_0} + a(t_0,X_{t_0})\int_{t_0}^{t} ds + b(t_0,X_{t_0})\int_{t_0}^{t} dW_s$$

remainder terms
$$\left\{\begin{aligned} &+ \int_{t_0}^{t}\int_{t_0}^{s} L^0 a(\tau,X_\tau)\,d\tau\,ds + \int_{t_0}^{t}\int_{t_0}^{s} L^1 a(\tau,X_\tau)\,dW_\tau\,ds \\ &+ \int_{t_0}^{t}\int_{t_0}^{s} L^0 b(\tau,X_\tau)\,d\tau\,dW_s + \int_{t_0}^{t}\int_{t_0}^{s} \underbrace{L^1 b(\tau,X_\tau)}_{\text{expand next}}\,dW_\tau\,dW_s \end{aligned}\right.$$

continue expanding varying integrands – many possibilities

$f = L^1 b \qquad L^1 b(\tau, X_\tau) = L^1 b(t_0, X_{t_0}) + \int_{t_0}^{\tau} L^0 L^1 b(u, X_u)\,du + \int_{t_0}^{\tau} L^1 L^1 b(u, X_u)\,dW_u +$

$$X_t = X_{t_0} + a(t_0, X_{t_0}) \int_{t_0}^{t} ds + b(t_0, X_{t_0}) \int_{t_0}^{t} dW_s$$
$$+ L^1 b(t_0, X_{t_0}) \int_{t_0}^{t} \int_{t_0}^{s} dW_\tau\, dW_s$$
$$+ \text{remainder terms}$$

<u>typically</u>
- constant coefficient terms in main expansion
- time varying terms in remainder
- successively higher multiple stochastic integrals provide more and more information about Wiener process within time interval.

apply on $[t_n, t_{n+1}]$ and truncate to obtain consistent numerical schemes of increasingly higher order of convergence

<u>Euler scheme</u> $\quad X_{t_{n+1}} \simeq X_{t_n} + a(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} ds + b(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} dW_s$

<u>Milstein scheme</u> $\quad X_{t_{n+1}} \simeq X_{t_n} + a(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} ds + b(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} dW_s$
$$+ L^1 b(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW_\tau\, dW_s$$

Euler $\gamma = 0.5$

Milstein $\gamma = 1.0$

MULTI-DIMENSIONAL SDEs

m-dimensional Wiener process

$$W_t = \left(W_t^1, W_t^2, \ldots, W_t^m\right)$$

components are pairwise independent standard scalar Wiener processes

d-dimensional Ito diffusion process

$$X_t = \left(X_t^1, X_t^2, \ldots, X_t^d\right)$$

satisfying the d-dimensional Ito SDE

$$(1) \qquad dX_t = a(t, X_t)\,dt + \sum_{j=1}^{m} b^j(t, X_t)\,dW_t^j$$

vector form

$$\boxed{dX_t^i = a^i(t, X_t)\,dt + \sum_{j=1}^{m} b^{i,j}(t, X_t)\,dW_t^j \qquad i = 1, 2, \ldots, d}$$

component form

equivalent Stratonovich SDE

$$(1^*) \qquad dX_t = \underline{a}(t, X_t)\,dt + \sum_{j=1}^{m} b^j(t, X_t) \circ dW_t^j$$

with corrected drift

$$\underline{a}^i(t, x) = a^i(t, x) - \frac{1}{2} \sum_{j=1}^{m} \sum_{k=1}^{d} b^{k,j}(t, x) \frac{\partial b^{i,j}}{\partial x^k}(t, x)$$

$$i = 1, 2, \ldots, d$$

ITO FORMULA

$$Y_t = f(t, X_t) \qquad f : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^1$$

$$\boxed{dY_t = L^0 f(t,X_t)\, dt + \sum_{j=1}^{m} L^j f(t,X_t)\, dW_t^j}$$

integral form $\quad Y_t = f(t_0, X_{t_0}) + \int_{t_0}^{t} L^0 f(s,X_s)\, ds + \sum_{j=1}^{m} \int_{t_0}^{t} L^j f(s,X_s)\, dW_s^j$

with **operators**

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{d} a^k \frac{\partial}{\partial x^k} + \frac{1}{2} \sum_{k,l=1}^{d} \sum_{j=1}^{m} b^{k,j} b^{l,j} \frac{\partial^2}{\partial x^k \partial x^l}$$

$$L^j = \sum_{k=1}^{d} b^{k,j} \frac{\partial}{\partial x^k} \qquad j = 1, 2, \ldots, m$$

EXAMPLE $\quad dX_t = a X_t\, dt + b^1 X_t\, dW_t^1 + b^2 X_t\, dW_t^2$

$d=1$, $m=2$ $\quad Y_t = f(t, X_t) = (X_t)^2$

$$a(t,x) = ax, \quad b^1(t,x) = b^1 x, \quad b^2(t,x) = b^2 x, \quad f(t,x) = x^2$$

$$L^0 f(t,x) = \left\{ 2a + (b^1)^2 + (b^2)^2 \right\} x^2, \quad L^j f(t,x) = 2 b^j x^2$$

Ito Formula for $Y_t = (X_t)^2$ gives

$$dY_t = \left\{ 2a + (b^1)^2 + (b^2)^2 \right\} Y_t\, dt + 2 b^1 Y_t\, dW_t^1 + 2 b^2 Y_t\, dW_t^2$$

Stochastic Taylor expansions for $f(t, X_t)$ about (t_0, X_{t_0})

- apply the **Ito formula iteratively** to the successive integrand functions $L^0 f$, $L^j f$, $L^0 L^0 f$, $L^0 L^j f$, $L^j L^0 f$, $L^{j_1} L^{j_2} f$, ... and substitute into the ongoing expansion, thus introducing successive **multiple stochastic integrals**

$$\int_{t_0}^{t} ds, \quad \int_{t_0}^{t} dW_s^j, \quad \int_{t_0}^{t}\int_{t_0}^{s} d\tau\, ds, \quad \int_{t_0}^{t}\int_{t_0}^{s} dW_\tau^j\, ds, \quad \int_{t_0}^{t}\int_{t_0}^{s} d\tau\, dW_s^j, \ \ldots$$

Multi-indices provide a **succinct terminology**

$\{0, 1, \ldots, m\}$ **index set**

multi-index α of **length** $l(\alpha) = l \geq 1$ $\qquad \alpha = (j_1, j_2, \ldots, j_l)$

where $j_1, j_2, \ldots, j_l \in \{0, 1, \ldots, m\}$

- define $\alpha = \phi$, **empty multi-index**, length $l(\phi) = 0$
- denote by $\underline{\mathcal{M}}$ the set of all multi-indices, $l \geq 0$

operations on a multi-index $\alpha = (j_1, j_2, \ldots, j_l) \in \mathcal{M} \setminus \{\phi\}$

$l = 1 \quad \therefore \quad -\alpha = \alpha- = \phi$

$l \geq 2 \qquad \underline{-\alpha} = -(j_1, \ldots, j_l) = (j_2, \ldots, j_l)$ **drop first component**

$\qquad\qquad \underline{\alpha-} = (j_1, \ldots, j_l)- = (j_1, \ldots, j_{l-1})$ **drop last component**

EXAMPLE $\alpha = (1, 0, 2)$, $-\alpha = (0, 2)$, $\alpha- = (1, 0)$

- **Write** $I_{\alpha,t_0,t} = I_\alpha[1]_{t_0,t}$ for $g \equiv 1$

- Define **coefficient functions**

$$f_\alpha : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^1$$

for multi-indices $\alpha \in \mathcal{M}$ and function $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ recursively

$$f_\phi = f, \qquad f_\alpha = \underset{\substack{\uparrow \\ \text{first component} \\ \text{of } \alpha}}{L^{j_1}} f_{-\alpha} \quad \text{for } l(\alpha) \geq 1$$

EXAMPLE $f_{(1)} = L^1 f$, $f_{(1,0)} = L^1 f_{(0)} = L^1 L^0 f$

ie $\boxed{f_{(j_1, j_2, \dots, j_l)} = L^{j_1} L^{j_2} \dots L^{j_l} f}$

EXAMPLE $dX_t = aX_t\,dt + bX_t\,dW_t$ $\quad$ $\underline{d = m = 1}$

$$L^0 = \frac{\partial}{\partial t} + ax\frac{\partial}{\partial x} - \frac{1}{2} b^2 x^2 \frac{\partial^2}{\partial x^2} \qquad L^1 = bx\frac{\partial}{\partial x}$$

$\underline{f(t,x) \equiv x}$

$$f_{(0)}(x) = L^0 f(x) = ax$$
$$f_{(1)}(x) = L^1 f(x) = bx$$
$$f_{(0,0)}(x) = L^0 f_{(0)}(x) = a^2 x$$
$$f_{(1,0)}(x) = L^1 f_{(0)}(x) = bax$$
$$f_{(1,1)}(x) = L^1 f_{(1)}(x) = b^2 x$$

Write $W_t^0 = t$ so $dW_t^0 = dt$ deterministic integration

$$(dt, dW_t) = (dW_t^0, dW_t^1, \dots, dW_t^m)$$
$$\underline{j=0}, \quad \underline{j=1}, \dots, \quad \underline{j=m}$$

Integrands $g : [0,T] \times \Omega \to \mathbb{R}^1$ nonanticipative, etc

Define Ito multiple stochastic integrals $I_\alpha[g]_{t_0,t}$ recursively

$\alpha = \phi$ $\qquad I_\phi[g]_{t_0,t} = g(t)$

$l(\alpha) \geq 1$ $\qquad I_\alpha[g]_{t_0,t} = \int_{t_0}^t I_{\alpha-}[g]_{t_0,s}\, dW_s^{j_l}$ last component of α

$$\begin{cases} j_l = 0 & \text{deterministic integral} \\ j_l \neq 0 & \text{Ito stochastic integral} \end{cases}$$

EXAMPLES

$$I_\phi[g]_{0,t} = g(t)$$

$$I_{(0)}[g]_{0,t} = \int_0^t g(s)\, ds \qquad (0)- = \phi$$

$$I_{(1)}[g]_{0,t} = \int_0^t g(s)\, dW_s^1 \qquad (1)- = \phi$$

$$I_{(1,0)}[g]_{0,t} = \int_0^t I_{(1)}[g]_{0,s}\, ds = \int_0^t \int_0^s g(\tau)\, dW_\tau^1\, ds$$

$$I_{(0,2)}[g]_{0,t} = \int_0^t I_{(0)}[g]_{0,s}\, dW_s^2 = \int_0^t \int_0^s g(\tau)\, d\tau\, dW_s^2$$

$$I_{(0,2,1)}[g]_{0,t} = \int_0^t I_{(0,2)}[g]_{0,s}\, dW_s^1 = \int_0^t \int_0^s \int_0^\tau g(u)\, du\, dW_\tau^2\, dW_s^1$$

Which multi-indices occur in Taylor expansions?

hierarchical subset $\mathcal{A} \subset \mathcal{M}$

① nonempty

② uniformly bounded $\quad \sup_{\alpha \in \mathcal{A}} l(\alpha) < \infty$

③ hierarchical $\quad -\alpha \in \mathcal{A}$ if $\alpha \in \mathcal{A} \setminus \{\phi\}$

EXAMPLES index set $\{0, 1\}$

$$\mathcal{A}_1 = \{\phi\}, \quad \mathcal{A}_2 = \{\phi, (0), (1)\}, \quad \mathcal{A}_3 = \{\phi, (0), (1), (1,1)\}$$

remainder set $\mathcal{B}(\mathcal{A}) \subset \mathcal{M}$ for a hierarchical set $\mathcal{A}$

$$\boxed{\mathcal{B}(\mathcal{A}) = \{\alpha \in \underbrace{\mathcal{M} \setminus \mathcal{A}}_{\text{complement of } \mathcal{A}} : -\alpha \in \mathcal{A}\}}$$

EXAMPLES index set $\{0, 1\}$, $\mathcal{A}_i$ as above

$$\mathcal{B}(\mathcal{A}_1) = \{(0), (1)\}$$

$$\mathcal{B}(\mathcal{A}_2) = \{(0,0), (1,0), (0,1), (1,1)\}$$

$$\mathcal{B}(\mathcal{A}_3) = \{(0,0), (1,0), (0,1), (0,1,1), (1,1,1)\}$$

General Ito-Taylor Stochastic Taylor Expansion

X_t d-dimensional $\qquad W_t$ m-dimensional

$\mathcal{A}$ hierarchical set
$\mathcal{B}(\mathcal{A})$ remainder set $\Big\}$ indices $\{0, 1, \ldots, m\}$

$f : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^1 \qquad a, b^j, f$ regular enough

$$\boxed{\begin{aligned} f(t, X_t) &= \sum_{\alpha \in \mathcal{A}} f_\alpha(t_0, X_{t_0})\, I_{\alpha, t_0, t} && \underline{\text{main terms}} \\ &+ \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha \left[f_\alpha(\cdot, X_\cdot) \right]_{t_0, t} && \underline{\text{remainder terms}} \end{aligned}}$$

EXAMPLE $\quad \mathcal{A} = \{\phi, (0), (1)\} \quad \mathcal{B}(\mathcal{A}) = \{(0,0), (0,1), (1,0), (1,1)\}$

$$f(t, X_t) = \underbrace{f(t_0, X_{t_0})}_{\alpha = \phi} + \underbrace{L^0 f(t_0, X_{t_0}) \int_{t_0}^{t} ds}_{\alpha = (0)} + \underbrace{L^1 f(t_0, X_{t_0}) \int_{t_0}^{t} dW_s^1}_{\alpha = (1)}$$

<u>remainder terms have variable integrands</u>

$$\left\{ \begin{aligned} &+ \int_{t_0}^{t} \int_{t_0}^{s} L^0 L^0 f(\tau, X_\tau)\, d\tau\, ds && \alpha = (0,0) \\ &+ \int_{t_0}^{t} \int_{t_0}^{s} L^0 L^1 f(\tau, X_\tau)\, d\tau\, dW_s^1 && \alpha = (0,1) \\ &+ \int_{t_0}^{t} \int_{t_0}^{s} L^1 L^0 f(\tau, X_\tau)\, dW_\tau^1\, ds && \alpha = (1,0) \\ &+ \int_{t_0}^{t} \int_{t_0}^{s} L^1 L^1 f(\tau, X_\tau)\, dW_\tau^1\, dW_s^1 && \alpha = (1,1) \end{aligned} \right.$$

Truncated Ito-Taylor Expansions

$f(t,x) \equiv x$, $\mathcal{A}$ hierarchical set, $\mathcal{B}(\mathcal{A})$ remainder set

$$X_t = \sum_{\alpha \in \mathcal{A}} f_\alpha(X_{t_0})\, I_{\alpha,t_0,t} + \underbrace{\sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha[f_\alpha(X_\cdot)]_{t_0,t}}_{\text{discarded in approximations}}$$

Strong Approximations

$$\mathcal{A} = \Lambda_k := \{\alpha \in \mathcal{M};\ \underset{\substack{\uparrow \\ \text{length of } \alpha}}{l(\alpha)} + \underset{\substack{\uparrow \\ \text{number of zero components in } \alpha}}{n(\alpha)} \le 2k\} \qquad k = 1,2,3,\ldots$$

$$X_t^{(k)} = \sum_{\alpha \in \Lambda_k} f_\alpha(X_{t_0})\, I_{\alpha,t_0,t} \to X_t \quad \text{w.p.1 uniformly in } t \in [t_0,T]$$

Weak Approximations

$$\mathcal{A} = \Gamma_k := \{\alpha \in \mathcal{M};\ l(\alpha) \le k\} \qquad k = 1,2,3,\ldots$$

$$X_t^{(k)} = \sum_{\alpha \in \Gamma_k} f_\alpha(X_{t_0})\, I_{\alpha,t_0,t} \longrightarrow X_t \quad \text{weakly in } t \in [t_0,T]$$

$$\Leftrightarrow \quad \sup_{t_0 \le t \le T} \left| E\left(g(X_t^{(k)})\right) - E\left(g(X_t)\right) \right| \to 0$$

g continuous polynomial growth

STRONG SCHEMES

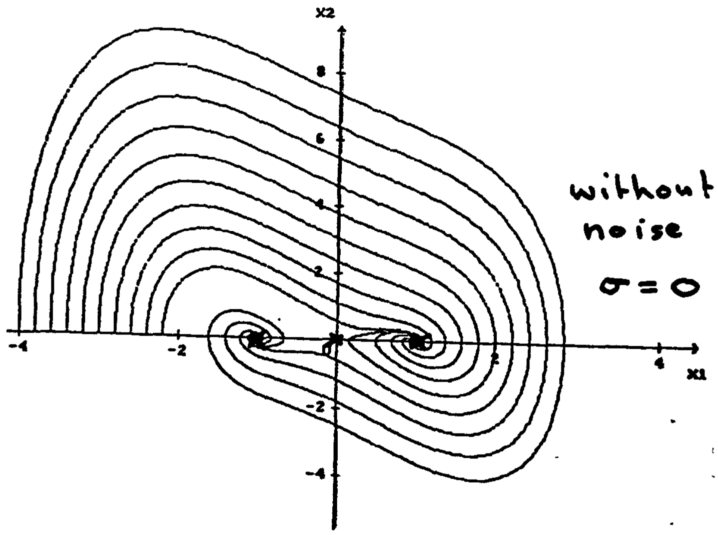
good sample path approximations

- dynamics
- filtering
- control
- finance
- parametric estimation

Noisy Duffing-van der Pol oscillator

$$\ddot{x} + \dot{x} - (\alpha - x^2)\, x = \sigma x \xi_t$$

white noise

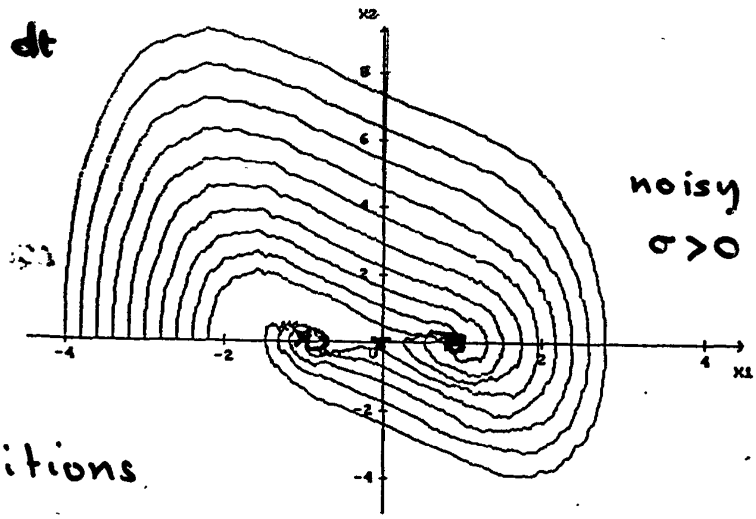


without noise $\sigma = 0$

2-dim. system of Ito SDEs

$$dX_t^1 = X_t^2\, dt$$

$$dX_t^2 = \left\{ X_t^1 \left(\alpha - (X_t^1)^2 \right) - X_t^2 \right\} dt + \sigma X_t^1\, dW_t$$



noisy $\sigma > 0$

Milstein scheme

$0 \leq t \leq 8 \quad$ step $\Delta = 2^{-7}$

same noise sample path for different initial conditions

Strong convergence criterion

$$E\left(\left| Y_{N_T}^{\delta} - X_T \right| \right) \leq K_T \cdot \delta^{\gamma}$$

order γ strong convergence

constant depends on T, SDE, scheme

uniform in $t \in [0, T]$ too

1^{st}- dimensional autonomous case $\qquad d = m = 1$

$$dX_t = a(X_t)\,dt + b(X_t)\,dW_t = \sum_{j=0}^{1} b^j(X_t)\,dW_t^j \qquad b^0 = a,\ b^1 = b \quad W_t^0 = t,\ W_t^1 = W_t$$

index set $\{0,1\}$ $\qquad L^0 = \frac{\partial}{\partial t} + a\frac{\partial}{\partial x} + \frac{1}{2}b^2\frac{\partial^2}{\partial x^2}, \quad L^1 = b\frac{\partial}{\partial x}$

$f(t,x) \equiv x \qquad f_\phi \equiv x,\ f_{(0)} = a,\ f_{(1)} = b,\ f_{(1,1)} = L^1 b,\ \ldots$

order $\gamma = 0.5$ strong Taylor scheme $\qquad \mathcal{A}^s_{0.5} = \{\phi, (0), (1)\}$ hierarchical set

$$Y_{n+1} = \sum_{\alpha \in \mathcal{A}^s_{0.5}} f_\alpha(t_n, Y_n)\, I_{\alpha, t_n, t_{n+1}}$$

$$= \underbrace{f_\phi(t_n,Y_n)}_{Y_n}\, \underbrace{I_{\phi,t_n,t_{n+1}}}_{1} + \underbrace{f_{(0)}(t_n,Y_n)}_{a(Y_n)}\, \underbrace{I_{(0),t_n,t_{n+1}}}_{\int_{t_n}^{t_{n+1}} dW_t^0} + \underbrace{f_{(1)}(t_n,Y_n)}_{b(Y_n)}\, \underbrace{I_{(1),t_n,t_{n+1}}}_{\int_{t_n}^{t_{n+1}} dW_t^1}$$

stochastic Euler scheme

$$\boxed{Y_{n+1} = Y_n + a(Y_n)\Delta_n + b(Y_n)\Delta W_n}$$

(Maruyama 1955)

$$\Delta_n = \int_{t_n}^{t_{n+1}} dt = \int_{t_n}^{t_{n+1}} dW_t^0$$
$$\Delta W_n = \int_{t_n}^{t_{n+1}} dW_t = \int_{t_n}^{t_{n+1}} dW_t^1$$

order $\gamma = 1.0$ strong Taylor scheme $\qquad \mathcal{A}^s_{1.0} = \{\phi, (0), (1), (1,1)\}$

additional term $\qquad f_{(1,1)}\, I_{(1,1)} = \underbrace{L^1 b(Y_n)}_{b(Y_n)b'(Y_n)}\ \underbrace{\int_{t_n}^{t_{n+1}}\int_{t_n}^{t} dW_s^1\, dW_t^1}_{\frac{1}{2}\{(\Delta W_n)^2 - \Delta_n\}}$

Milstein scheme

$$\boxed{\begin{aligned} Y_{n+1} = {} & Y_n + a(Y_n)\Delta_n + b(Y_n)\Delta W_n \\ & + \tfrac{1}{2} b(Y_n)b'(Y_n)\left\{(\Delta W_n)^2 - \Delta_n\right\} \end{aligned}}$$

Milstein 1974

order $\gamma = 1.5$ strong Taylor scheme

$$\mathcal{A}^s_{1.5} = \left\{ \begin{array}{l} \phi, (0), (1), (1,1) \\ (0,0), (1,0), (0,1), (1,1,1) \end{array} \right\}$$

new terms

$(0,0)$ $\quad f_{(0,0)} = L^0 a = a a' + \frac{1}{2} b^2 a'' \qquad I_{(0,0),t_n,t_{n+1}} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} ds\, dt = \frac{1}{2} \Delta_n^2$

$(1,0)$ $\quad f_{(1,0)} = L^1 a = b a' \qquad I_{(1,0),t_n,t_{n+1}} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW_s\, dt = \Delta Z_n$

$(0,1)$ $\quad f_{(0,1)} = L^0 b = a b' + \frac{1}{2} b^2 b'' \qquad I_{(0,1),t_n,t_{n+1}} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} ds\, dW_t$

$(1,1,1)$ $\quad f_{(1,1,1)} = L^1 L^1 b = b(bb')' = b(b')^2 + b^2 b''$

$$I_{(1,1,1),t_n,t_{n+1}} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} \int_{t_n}^{s} dW_u\, dW_s\, dW_t = \frac{1}{2}\left\{\frac{1}{3}(\Delta W_n)^2 - \Delta_n\right\}\Delta W_n$$

correlated Gaussian RVs

$\Delta Z_n \sim N\left(0; \frac{1}{3}\Delta_n^3\right) \qquad \Delta W_n = G_n \Delta_n^{1/2} \qquad G_n, G_n' \sim N(0;1)$ independent

$E(\Delta W_n \cdot \Delta Z_n) = \frac{1}{2}\Delta_n^2 \qquad \Delta Z_n = \frac{1}{2}\Delta_n^{3/2}\left\{G_n + \frac{1}{\sqrt{3}} G_n'\right\}$

identity $\quad I_{(1,0)} + I_{(0,1)} = I_{(0)} . I_{(1)} \qquad I_{(0,1)} = \Delta_n \cdot \Delta W_n - \Delta Z_n$

$$\boxed{\begin{aligned} Y_{n+1} = {} & Y_n + a\Delta_n + b\Delta W_n + \frac{1}{2} b b' \{(\Delta W_n)^2 - \Delta_n\} \\ & + \frac{1}{2}\left(aa' + \frac{1}{2}b^2 a''\right)\Delta_n^2 + b a' \Delta Z_n \\ & + \left(ab' + \frac{1}{2}b^2 b''\right)\{\Delta_n \cdot \Delta W_n - \Delta Z_n\} + \frac{1}{2}\left(b(b')^2 + b^2 b''\right)\left\{\frac{1}{3}(\Delta W_n)^2 - \Delta_n\right\}\Delta W_n \end{aligned}}$$

SIMPLIFICATIONS ① **additive noise** $\quad b = \text{const.}, \; b' = 0, \; b'' = 0$ etc

Euler ≡ Milstein order $\gamma = 1.0$

② **special coefficient relationships** eg $ba' \equiv ab' + \frac{1}{2}b^2 b''$

$\Rightarrow f_{(1,0)} I_{(1,0)} + f_{(0,1)} I_{(0,1)}$

$\equiv ba'\left(I_{(1,0)} + I_{(0,1)}\right) \equiv ba' \, I_{(0)} . I_{(1)}$

no need to generate ΔZ_n random variable in this case

STRONG TAYLOR SCHEMES

consistent, high order convergence

based on truncated Taylor expansions

Ito SDE $$dX_t^i = a^i(t, X_t)\,dt + \sum_{j=1}^{m} b^{i,j}(t, X_t)\,dW_t^j$$

$i = 1, \dots, d$

Write $X_t = (X_t^1, \dots, X_t^d)$, $W_t^0 = t$, $b^{i,0} = a^i$

Multi-indices $\alpha = (j_1, \dots, j_l)$, $j_i \in \{0, 1, \dots, m\}$ empty index $\alpha = \phi$

admissible hierarchical sets
$$\mathcal{A}_\gamma^s = \left\{ \alpha \in \mathcal{M} ;\ \begin{array}{l} l(\alpha) + n(\alpha) \le 2\gamma \\ \text{or } \ l(\alpha) = n(\alpha) = \gamma + \frac{1}{2} \end{array} \right\}$$

$l(\alpha) =$ length of α $\quad l(\phi) = 0$

$n(\alpha) =$ number of $j_i = 0$ in α

Take $f(t, X) = X^i$ $(i = 1, \dots, d)$ on $t_n \le t \le t_{n+1}$ $\quad f_\alpha^i = L^{j_1} \cdots L^{j_l} f^i$, $\alpha = (j_1, \dots, j_l)$

order γ strong-Taylor scheme
$$\boxed{Y_{n+1}^i = \sum_{\alpha \in \mathcal{A}_\gamma^s} f_\alpha^i(t_n, Y_n)\, I_{\alpha, t_n, t_{n+1}}}$$

$\gamma = 0.5, 1.0, 1.5, 2.0, 2.5, \dots$

- converges with strong order γ if coefficients $a^i, b^{i,j}$ are regular enough to ensure that f_α coefficients are Lipschitz and have linear growth, etc.

Multi-dimensional noise $\quad m > 1, \quad d \geq 1$

e.g. $m=2$, $d \geq 1$

$$dX_t^i = a^i(t, X_t)\,dt + b^{i,1}(t, X_t)\,dW_t^1 + b^{i,2}(t, X_t)\,dW_t^2 \qquad i = 1, \ldots, d$$

index set $\{0, 1, 2\}$ $\qquad W_t^1, W_t^2$ independent Wiener processes

stochastic Euler scheme

$$\mathcal{A}^s_{0.5} = \{\phi, (0), (1), (2)\} \qquad \text{order 0.5 strong convergence}$$

$$\boxed{Y_{n+1}^i = Y_n^i + a^i(t_n, Y_n)\Delta_n + b^{i,1}(t_n, Y_n)\Delta W_n^1 + b^{i,2}(t_n, Y_n)\Delta W_n^2}$$

$$\mathcal{A}^s_{1.0} = \{\phi, (0), (1), (2), (1,1), (2,2), (1,2), (2,1)\}$$

order $\gamma = 1.0$ Milstein scheme also includes

$$I_{(j,j)), t_n, t_{n+1}} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW_s^j\, dW_t^j = \frac{1}{2}\left\{(\Delta W_n^j)^2 - \Delta_n\right\} \qquad j = 1, 2$$

$$I_{(j_1, j_2)), t_n, t_{n+1}} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW_s^{j_1}\, dW_t^{j_2} \qquad (j_1, j_2) = (1,2) \text{ or } (2,1)$$

identity $\qquad I_{(1,2)), t_n, t_{n+1}} + I_{(2,1)), t_n, t_{n+1}} = \Delta W_n^1 \cdot \Delta W_n^2$

But there is no simple formula for $I_{(1,2)}$ or $I_{(2,1)}$ alone solely in terms of ΔW_n^1, ΔW_n^2 and Δ_n

Clark & Cameron example

$$dX_t^1 = dW_t^1$$
$$dX_t^2 = X_t^1\,dW_t^2 \qquad X_0^1 = X_0^2 = 0$$

Milstein scheme

$$Y_{n+1}^1 = Y_n^1 + \Delta W_n^1$$
$$Y_{n+1}^2 = Y_n^2 + Y_n^1 \Delta W_n^2 + \cancel{I_{(1,2)), t_n, t_{n+1}}} \leftarrow \text{substitute } \tfrac{1}{2}\Delta W_n^1 \Delta W_n^2$$

Clark & Cameron showed that the 2nd component satisfies

$$\sqrt{E\left(|Y_{N_T}^2 - X_T^2|^2\right)} = K \cdot T^{1/2} \cdot \Delta^{1/2}$$

order $1/2$

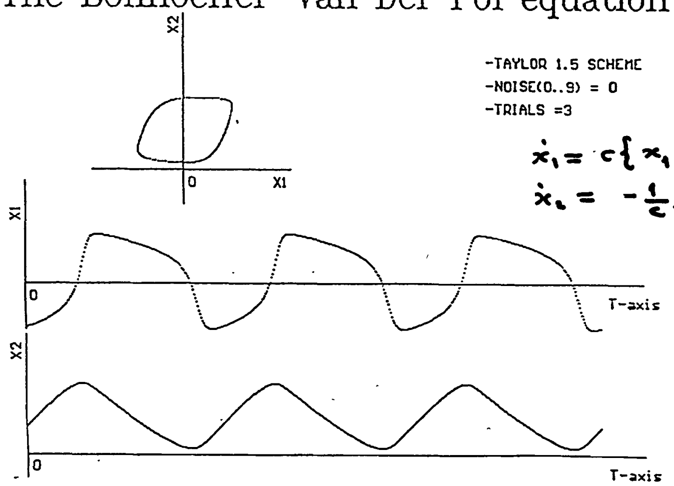
(equi time steps)

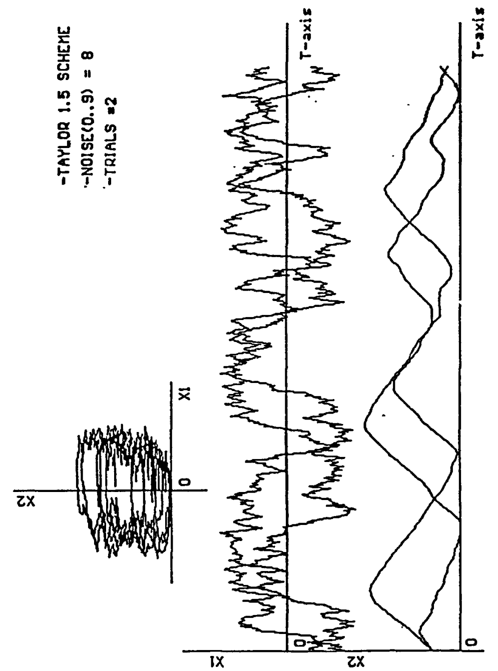
The Bonhoeffer–Van Der Pol equation

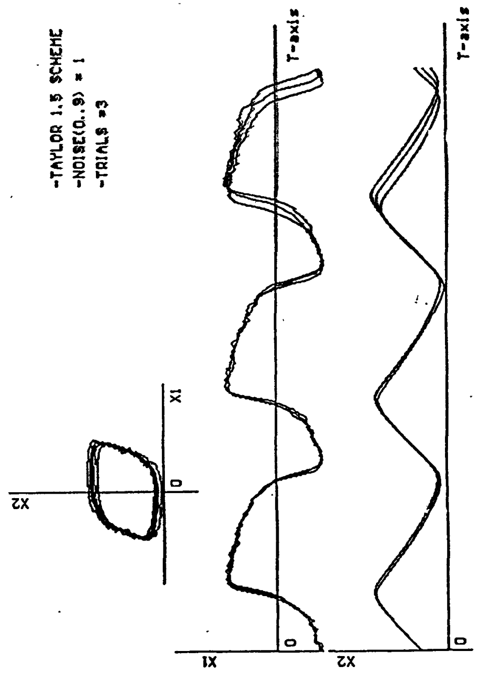
-TAYLOR 1.5 SCHEME
-NOISE(0..9) = 0
-TRIALS =3

$$\dot{x}_1 = c\{x_1 + x_2 - \tfrac{1}{3}x_1^3 + z\} + \sigma\xi_t$$
$$\dot{x}_2 = -\tfrac{1}{c}\{x_1 + bx_2 - a\}$$

additive noise





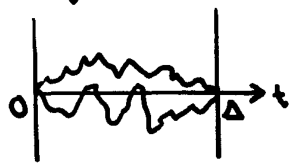


Approximation of multiple stochastic integrals

Mixed multiple stochastic integrals like $I_{(1,2),t_n,t_{n+1}}$ must be generated separately, but usually their distribution law is not known. They can be approximated using random Fourier series, ie <u>Karhunen-Loève expansions</u>

<u>Brownian bridge process</u>

$$W_t^j - \frac{t}{\Delta} W_\Delta^j = \frac{1}{2} a_{j,0} + \sum_{r=1}^{\infty} \left\{ a_{j,r} \cos\left(\frac{2\pi r t}{\Delta}\right) + b_{j,r} \sin\left(\frac{2\pi r t}{\Delta}\right) \right\}$$



$$a_{j,r}\ ,\ b_{j,r} \sim N\left(0; \frac{\Delta}{2\pi^2 r^2}\right) \quad \text{independent}$$

Differentiate, multiply and integrate series term by term

$$\int_0^\Delta W_t^1 \frac{dW_t^2}{dt}\, dt = \boxed{I_{(1,2),0,\Delta} = \frac{1}{2} W_\Delta^1 . W_\Delta^2 - \frac{1}{2}\left\{a_{2,0}.W_\Delta^1 - a_{1,0}.W_\Delta^2\right\} + \pi \sum_{r=1}^{\infty} \left\{ a_{1,r}\cdot b_{2,r} - b_{1,r}\cdot a_{2,r} \right\}}$$

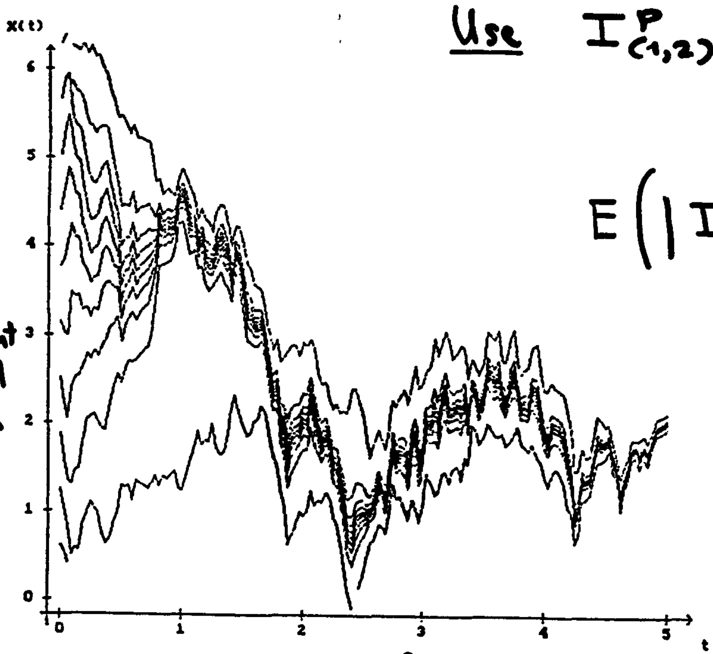
series expansions

<u>Use</u> $I^p_{(1,2),0,\Delta}$ truncated at $r = p$

needs $4p+4$ R.Vs

$$E\left(\left| I^p_{(1,2)} - I_{(1,2)} \right|^2\right) \leq \frac{\Delta^2}{2\pi^2 p}$$

same sample path for different initial values



<u>Take</u> $p \geq \frac{K}{\Delta}$ to ensure order $\gamma = 1.0$ strong convergence in the Milstein scheme

Stochastic gradient flow on circle

$$dX_t = \sin X_t\, dW_t^1 + \cos X_t\, dW_t^2 \quad (\text{mod } 2\pi)$$

Wong-Zakai Theorem

<u>Milstein scheme</u> $0 \leq t \leq 5$, $\Delta = 2^{-7}$, $I^{20}_{(1,2)}$ used

MAJOR PRACTICAL ISSUE

- the effective use of higher order strong Taylor schemes requires efficient methods for generating or approximating the necessary multiple stochastic integrals
  - general schemes of strong $\gamma \geq 4.0$ are impractical
- software for general SDEs is likely to be inefficient unless it takes into account special structural features of specific SDEs which can avoid the need to generate all multiple stochastic integrals in the scheme, eg

① additive noise $\quad b^{i,j}(t,x) \equiv \text{const.} \Rightarrow$ derivatives vanish
$\Rightarrow$ many 0 coefficient

② commutative noise $\quad L^{j_1} b^{i,j_2}(t,x) \equiv L^{j_2} b^{i,j_1}(t,x) \quad j_1 \neq j_2$

$\Rightarrow$ terms $j_1 \neq j_2$ $\quad L^{j_1} b^{i,j_2} I_{(j_1,j_2)} + L^{j_2} b^{i,j_1} I_{(j_2,j_1)}$

$$= L^{j_1} b^{i,j_2} \left\{ \underbrace{I_{(j_1,j_2)} + I_{(j_2,j_1)}} \right\} = L^{j_1} b^{i,j_2} \underbrace{I_{(j_1)} I_{(j_2)}}_{\text{identity}}$$

there is no need to generate the integrals $I_{(j_1,j_2)}, I_{(j_2,j_1)}$

eg $\quad dX_t = a(X_t)dt + b^1(X_t)\,dW_t^1 + b^2(X_t)\,dW_t^2$

assume $L^1 b^2 \equiv L^2 b^1$ ie $b^1 b^{2\prime} \equiv b^2 b^{1\prime} := B$

Milstein scheme

$$
\begin{aligned}
Y_{n+1} = Y_n &+ a(Y_n)\Delta_n + b^1(Y_n)\Delta W_n^1 + b^2(Y_n)\Delta W_n \\
&+ \tfrac{1}{2} b^1(Y_n) b^{1\prime}(Y_n) \left\{ (\Delta W_n^1)^2 - \Delta_n \right\} \\
&+ \tfrac{1}{2} b^2(Y_n) b^{2\prime}(Y_n) \left\{ (\Delta W_n^2)^2 - \Delta_n \right\} \\
&+ B(Y_n).\,\Delta W_n^1 \cdot \Delta W_n^2 \leftarrow \text{instead of } I_{(1,2)} \text{ and } I_{(2,1)}
\end{aligned}
$$

MORAL — LOOK FOR SPECIAL STRUCTURE TO SIMPLIFY THE SCHEME

ANOTHER PRACTICAL ISSUE

Taylor scheme coefficients

$$f_{(j_1, j_2, \ldots, j_l)} = L^{j_1} L^{j_2} \ldots L^{j_l} f \qquad f(t,x) = x$$

involve mixed higher order partial derivatives of drift and diffusion coefficients

complicated if $l \geq 2$, $d > 1$ and $m > 1$

symbolic manipulators like MAPLE can be used to determine f_α automatically

```
> with(student):
>
>  L :=
>
>  proc(F)
>  local ip,v;
>      ip := op(procname);
>      v := op(F);
>      if ip=0 then;
>          seq(a.k(v)*Diff(F,v[k]),k=1..d);
>          seq(seq(seq(b.k.j(v)*b.l.j(v)*Diff(Diff(F,v[k]),v[l]),j=1..m),l=1..
>          ),k=1..d);
>          Diff(F,v[d+1]) + convert({""},'+') + 1/2*convert({"},'+');
>      else; seq(b.k.ip*Diff(F,v[k]),k=1..d);
>     convert({"},'+');
>   fi;
>  end:
```

Alternatively stochastic Taylor schemes can be modified to avoid the use of derivatives in their coefficients as in deterministic Runge-Kutta schemes

$$dX_t = a(t, X_t)\,dt + b(t, X_t)\,dW_t, \qquad d = m = 1$$

Milstein scheme includes $L^1 b = b(t,x) \frac{\partial b}{\partial x}(t,x)$ term

$$b(t_n, Y_n) \frac{\partial b}{\partial x}(t_n, Y_n) \simeq \frac{b\left(t_n, Y_n + a(Y_n)\Delta_n + b(Y_n)\sqrt{\Delta_n}\right) - b(t_n, Y_n)}{\sqrt{\Delta_n}}$$

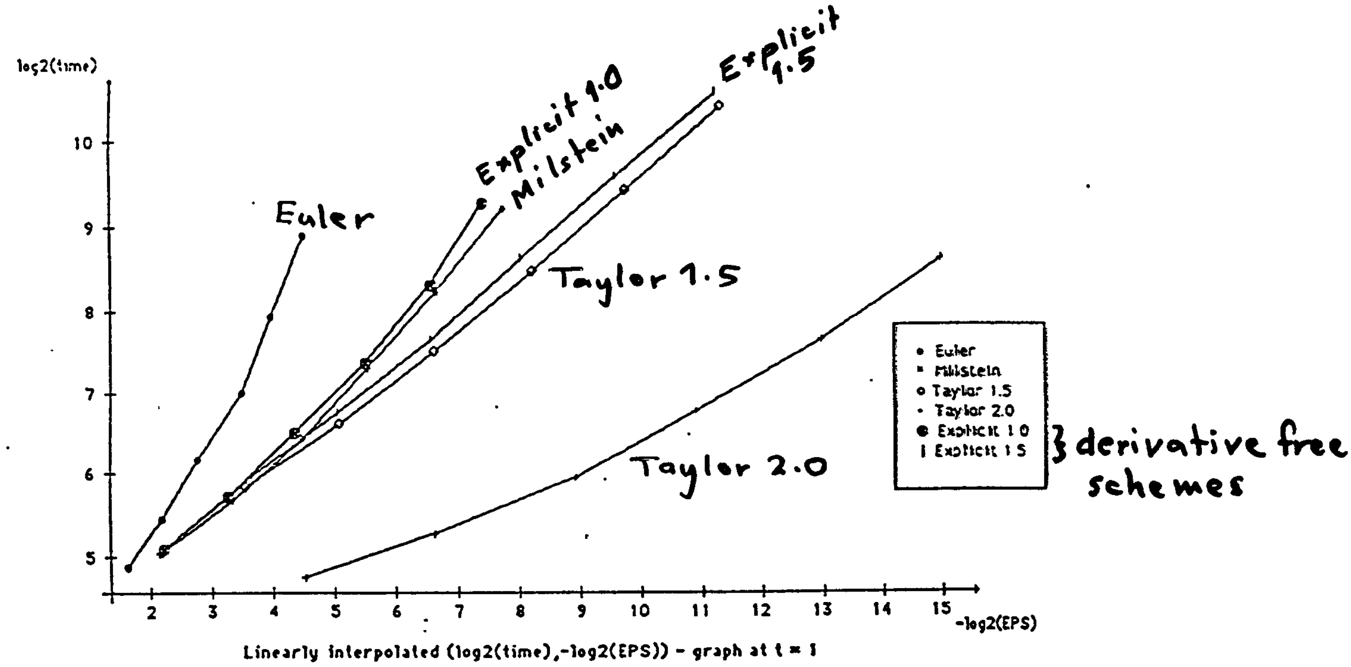
Platen's order $\gamma = 1.0$ derivative free scheme

"stochastic Runge-Kutta scheme"

$$\boxed{\begin{aligned} Y_{n+1} &= Y_n + a(t_n, Y_n)\,\Delta_n + b(t_n, Y_n)\,\Delta W_n \\ &\quad + \frac{1}{2\sqrt{\Delta_n}} \left\{ b(t_n, \Upsilon_n) - b(t_n, Y_n) \right\} \left\{ (\Delta W_n)^2 - \Delta_n \right\} \\ \text{with} \quad \Upsilon_n &= Y_n + \underbrace{a(Y_n)\Delta_n}_{\text{could be omitted}} + b(Y_n)\sqrt{\Delta_n} \end{aligned}}$$

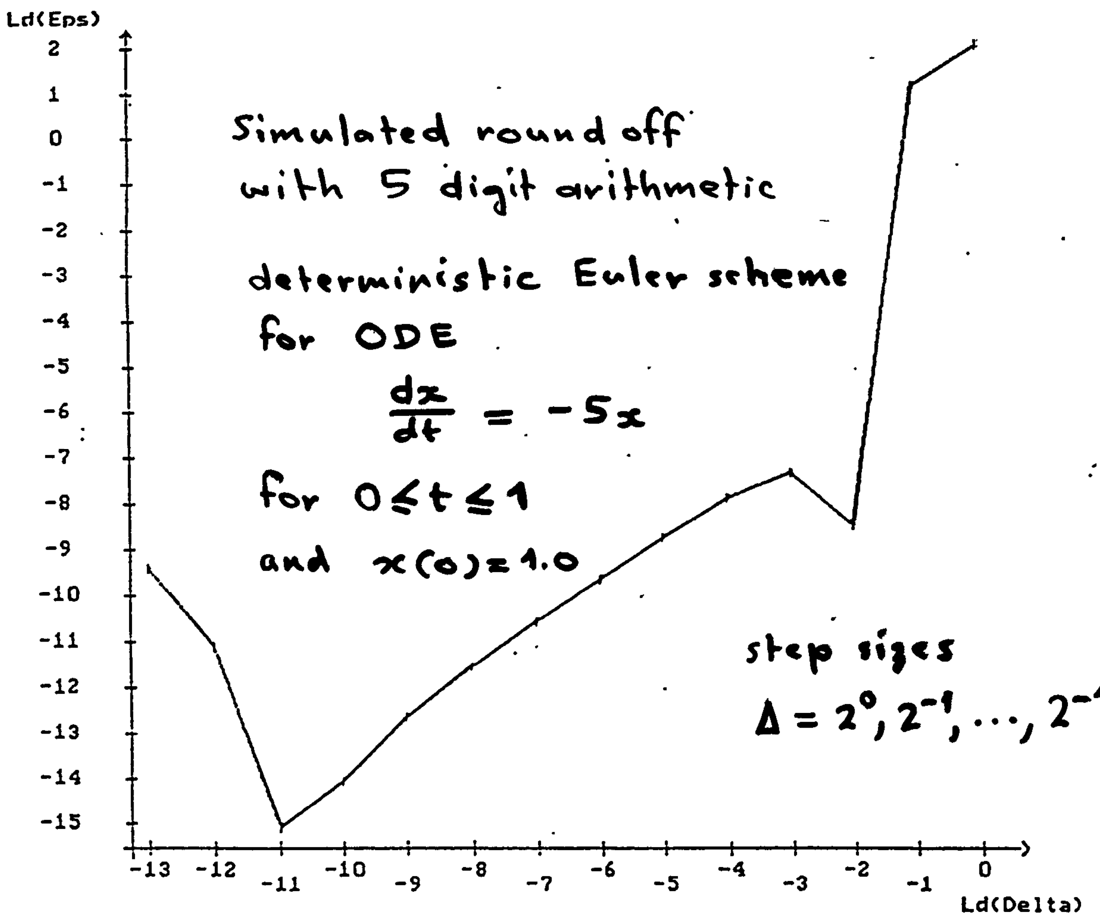
Higher order strong schemes of this kind are also available

Comparison of CPU times for given error



$$dX_t = \frac{1}{2} X_t \, dt + X_t \, dW_t \qquad EPS = E\left(\left| Y_{n_T} - X_T \right|\right) \qquad T = 1$$

$$X_0 = 1, \quad 0 \leq t \leq 1 \qquad X_t = X_0 e^{W_t}$$



Simulated round off with 5 digit arithmetic

deterministic Euler scheme for ODE

$$\frac{dx}{dt} = -5x$$

for $0 \leq t \leq 1$

and $x(0) = 1.0$

step sizes

$\Delta = 2^0, 2^{-1}, \ldots, 2^{-13}$

NUMERICAL INSTABILITY

- although convergent in theory, a numerical scheme will not be useful in practice unless it can control the propagatio[n] of initially small errors

numerical stability

- very rapid time changes can trigger numerical instabilities

stiff differential equations

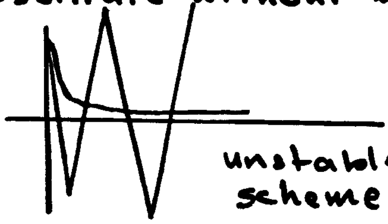
ODE $\quad \dfrac{dx}{dt} = -10^N x \qquad x(t) = x_0 e^{-10^N t}$

Euler scheme $\Delta_n \equiv \Delta$

$$x_{n+1} = x_n - 10^N x_n \Delta = (1 - 10^N \Delta) x_n$$

$$x_n = (1 - 10^N \Delta)^n x_0$$

$\Delta > 2.10^{-N}$ iterates oscillate without bound



unstable scheme

$x_n \longrightarrow 0$ monotonically if $\Delta < 10^{-N}$

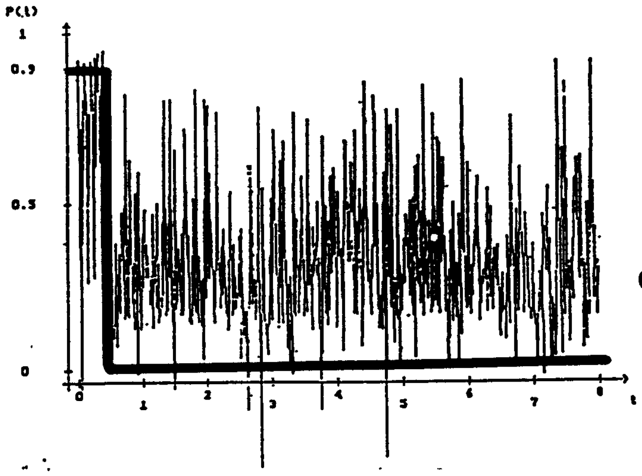
BUT for $N \gg 1$ such time steps may be smaller than the roundoff resolution of the computer used

Implicit Euler scheme $\Delta_n \equiv \Delta$

$$\underline{x_{n+1}} = x_n - 10^N \underline{x_{n+1}} \Delta \quad \text{ie} \quad (1 + 10^N \Delta) x_{n+1} = x_n$$

$$x_n = \left(\frac{1}{1 + 10^N \Delta} \right)^n x_0 \longrightarrow 0$$

monotonically as $n \to \infty$ for all step sizes $\Delta > 0$



Milstein scheme with $\Delta_n \equiv 2^{-7}$ for Zakai filtering SDE

$$d\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{bmatrix} -50 & 50 \\ 50 & -50 \end{bmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} dt + \begin{bmatrix} 15 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} dW_t$$

for noisy observations of a 2-state Markov chain

Implicit strong schemes

- to avoid inconsistencies just mak[e] purely deterministic integral $I_{(0)}$, $I_{(0,0)}, \cdots$ coefficients implicit

$$dX_t^i = a^i(t, X_t)\,dt + b^i(t, X_t)\,dW_t$$

explicit Euler scheme $\quad Y_{n+1}^i = Y_n^i + a^i(t_n, Y_n)\Delta_n + b^i(t_n, Y_n)\Delta W_n$ $\quad i = 1, \ldots, d$

implicit Euler scheme

$$Y_{n+1}^i = Y_n^i + a^i(t_{n+1}, Y_{n+1})\Delta_n + b^i(t_n, Y_n)\Delta W_n$$

Y_{n+1} solve algebraically or numerically

stochastic term remains explicit

family of implicit Euler schemes

$$\boxed{Y_{n+1}^i = Y_n^i + \left\{\alpha_i\, a^i(t_{n+1}, Y_{n+1}) + (1-\alpha_i)\, a^i(t_n, Y_n)\right\}\Delta_n + b^i(t_n, Y_n)\,\Delta W_n}$$

$0 \le \alpha_i \le 1$, $i = 1, \ldots, d$

order $\gamma = 0.5$ strong convergence

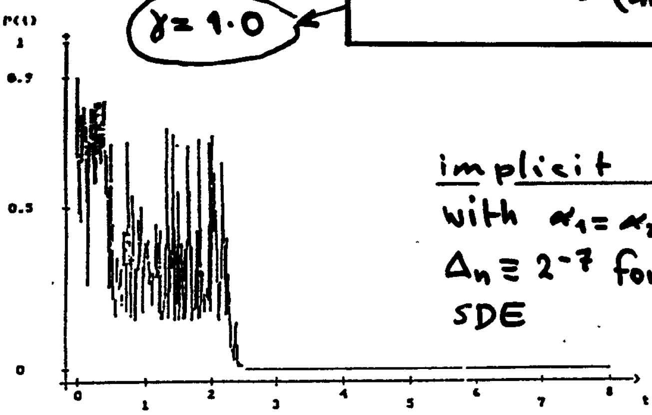
implicitness parameter α_i can be different in each component

family of implicit Milstein schemes

$$\boxed{Y_{n+1}^i = Y_n^i + \left\{\alpha_i\, a^i(t_{n+1}, Y_{n+1}) + (1-\alpha_i)\, a^i(t_n, Y_n)\right\}\Delta_n + b^i(t_n, Y_n)\Delta W_n + \tfrac{1}{2} b^i(t_n, Y_n)\frac{\partial b}{\partial x}(t_n, Y_n)\left\{(\Delta W_n)^2 - \Delta_n\right\}}$$

$\gamma = 1.0$

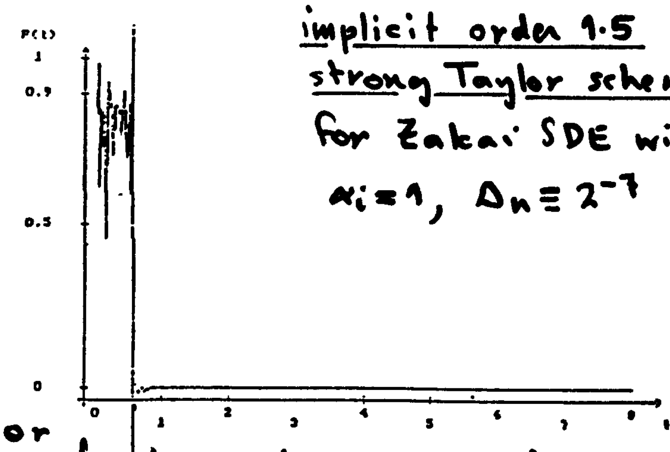
$i = 1, \ldots, d$



implicit Milstein scheme with $\alpha_1 = \alpha_2 = 1$ and stepsize $\Delta_n \equiv 2^{-7}$ for Zakai filtering SDE

can rearrange algebraically to make explicit

implicit order 1.5 strong Taylor scheme for Zakai SDE with $\alpha_i = 1$, $\Delta_n \equiv 2^{-7}$



① $L^0 a$ coefficient of $I_{(0,0)}$ integral also implicit here

② scheme involves

$$\Delta Z_n = I_{(1,0),t_n,t_{n+1}} = \int_{t_n}^{t_{n+1}} W_t\,dt$$

as well as ΔW_n.

continuous or rapid sampling of observation process W_t

STOCHASTIC FLOW ON A TORUS

$$T^2 \cong S^1 \times S^1 \cong [0,2\pi) \times [0,2\pi)$$

$W_t^1, W_t^2, W_t^3, W_t^4$ pairwise indept. Wiener processes

$$b^1(\underset{\sim}{x}) = b^1(x_1,x_2) = \begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}\sin(x_1)$$

$$b^2(\underset{\sim}{x}) = b^2(x_1,x_2) = \begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}\cos(x_1)$$

$$b^3(\underset{\sim}{x}) = b^3(x_1,x_2) = \begin{pmatrix}-\sin\alpha\\ \cos\alpha\end{pmatrix}\sin(x_2)$$

$$b^4(\underset{\sim}{x}) = b^4(x_1,x_2) = \begin{pmatrix}-\sin\alpha\\ \cos\alpha\end{pmatrix}\cos(x_2)$$

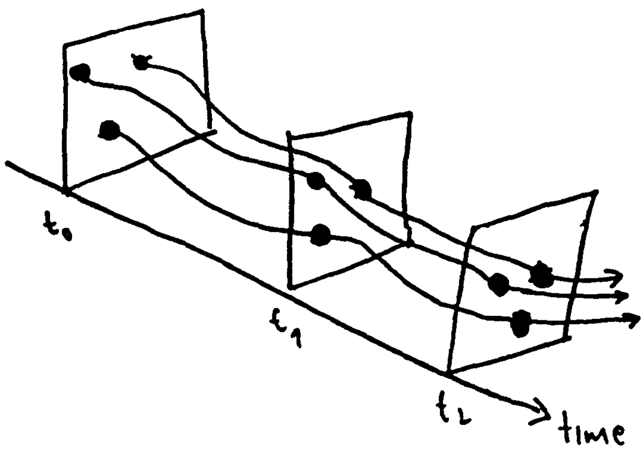
α = coupling parameter

Baxendale's 2-dim Stratonovich SDE

$$\boxed{d\underset{\wedge}{X}_t = \sum_{j=1}^{4} b^j(\underset{\wedge}{X}_t)\circ dW_t^j} \quad (\text{mod } 2\pi)$$

Milstein scheme • $\Delta = 0.01$, $\alpha = 1.0$

• uniform grid of initial values



note how paths coalesce in a blob that moves around randomly

WEAK APPROXIMATIONS

exact X_T

approximate $Y^{\delta}_{N_T}$

the actual realizations of the random variables are <u>not</u> always important

Sometimes only want or need a good approximation of the probability distribution or density

exact $F_{X_T}(a) = \int_{-\infty}^{a} p_T(x)\,dx = P\left(\{\omega\in\Omega;\ X_T(\omega)\le a\}\right)$

approx. $F_{Y^{\delta}_{N_T}}(a) = \int_{-\infty}^{a} p^{\delta}_T(y)\,dy = P\left(\{\omega\in\Omega;\ Y^{\delta}_{N_T}(\omega)\le a\}\right)$

① <u>probability</u>

$$P\left(\{\omega\in\Omega;\ X_T(\omega)\in[a,b]\}\right) = \int_a^b p_T(x)\,dx = E\left(\mathbb{1}_{[a,b]}(X_T)\right)$$

$$\mathbb{1}_{[a,b]}(x) = \begin{cases} 0 & \text{if } x\notin[a,b] \\ 1 & \text{if } x\in[a,b] \end{cases}$$

② <u>moments</u>

$$E\left((X_T)^p\right) = \int_{-\infty}^{\infty} x^p\, p_T(x)\,dx$$

typically $p=2$

③ <u>PDE solutions</u>

$$\frac{\partial u}{\partial s} + a(s,x)\frac{\partial u}{\partial x} + \frac{1}{2}b^2(s,x)\frac{\partial^2 u}{\partial x^2}$$

$$u(T,x) = g(x) \qquad 0\le s\le T$$

Kolmogorov formula

$$\boxed{u(s,x) = E\left(g\left(X_T^{s,x}\right)\right)}$$

$X_t^{s,x}$ is solution of SDE

$$dX_t = a(t,X_t)\,dt + b(t,X_t)\,dW_t$$

on $s\le t\le T$ with $X_s = x$

<u>general weak approximation criterion</u>

$$\left|E\left(g\left(Y^{\delta}_{N_T}\right)\right) - E\left(g(X_T)\right)\right| \longrightarrow 0$$

as $\delta\to 0$ for continuous, polynomially bounded $g:\mathbb{R}\to\mathbb{R}$

WEAK TAYLOR SCHEMES

- consistent
- higher order
- based on truncated Taylor expansions

order $\beta > 0$ weak convergence
$$\left| E\left(g(Y^{\delta}_{N_T})\right) - E\left(g(X_T)\right) \right| \leq K_g \cdot \delta^{\beta}$$

<u>Ito SDE</u>
$$dX^i_t = a^i(t, X_t)\,dt + \sum_{j=1}^{m} b^{i,j}(t, X_t)\,dW^j_t \qquad (i = 1, \ldots, d)$$

<u>Write</u> $X_t = (X^1_t, \ldots, X^d_t)$, $W^0_t = t$, $b^{i,0} = a^i$

<u>Multi-indices</u> $\alpha = (j_1, \ldots, j_\ell)$, $j_i \in \{0, 1, \ldots, m\}$

empty index $\alpha = \phi$

<u>admissible hierarchical sets</u>
$$\mathcal{A}^{w}_{\beta} = \left\{ \alpha \in \mathcal{M} ;\ \ell(\alpha) \leq \beta \right\}$$

$\ell(\alpha) =$ length of α

$\ell(\phi) = 0$

<u>Take</u> $f(t, X) = X^i$ $(i = 1, \ldots, d)$ on $t_n \leq t \leq t_{n+1}$

$f^i_\alpha = L^{j_1} \ldots L^{j_\ell} f^i$

$\alpha = (j_1, \ldots, j_\ell)$

<u>order β weak Taylor scheme</u>
$$\boxed{Y^i_{n+1} = \sum_{\alpha \in \mathcal{A}^w_\beta} f^i_\alpha(t_n, Y_n)\, I_{\alpha, t_n, t_{n+1}}}$$

$\beta = 1.0, 2.0, 3.0, \ldots$ <u>integers only</u>

- converges with weak order β if the coeffiients $a^i, b^{i,j}$ are regular enough to ensure that the f_α are Lipschitz and have bounded linear growth, etc

$$\text{eg} \quad a^i, b^{i,j} \in C^{2(\beta+1)}_{\text{poly. growth}}(\mathbb{R}^n, \mathbb{R})$$

if indept. of t

1-dimensional autonomous case $\quad d = m = 1$

$$dX_t = a(X_t)\,dt + b(X_t)\,dW_t = \sum_{j=0}^{2} b^j(X_t)\,dW_t^j \qquad b^0 = a,\ b^1 = b,\quad W_t^0 = t,\ W_t^1 = W_t$$

index set $\quad \{0,1\} \qquad L^0 = \frac{\partial}{\partial t} + a\frac{\partial}{\partial x} + \frac{1}{2}b^2\frac{\partial^2}{\partial x^2}, \quad L^1 = b\frac{\partial}{\partial x}$

$f(t,x) \equiv x \qquad f_\phi = x,\ f_{(0)} = a,\ f_{(1)} = b,\ f_{(1,1)} = L^1 b, \ldots$

order $\beta = 1.0$ weak Taylor scheme

$$\mathcal{A}^w_{1.0} = \{\phi, (0), (1)\} = \mathcal{A}^s_{0.5}$$

same hierarchical set as order $\gamma = 0.5$ strong Taylor scheme

$$\boxed{Y_{n+1} = Y_n + a(Y_n)\Delta_n + b(Y_n)\Delta W_n}$$

stochastic Euler scheme

order $\beta = 2.0$ weak Taylor scheme

$$\mathcal{A}^w_{2.0} = \{\phi, (0), (1), \underbrace{(1,1), (1,0), (0,1), (0,0)}_{\text{new}}\}$$

new terms

$(0,0) \qquad f_{(0,0)} = L^0 a = aa' + \frac{1}{2}b^2 a'' \qquad I_{(0,0),t_n,t_{n+1}} = \int_{t_n}^{t_{n+1}}\int_{t_n}^{t} ds\,dt = \frac{1}{2}\Delta_n^2$

$(1,0) \qquad f_{(1,0)} = L^1 a = ba' \qquad I_{(1,0),t_n,t_{n+1}} = \int_{t_n}^{t_{n+1}}\int_{t_n}^{t} dW_s\,dt = \Delta Z_n$

$(0,1) \qquad f_{(0,1)} = L^0 b = ab' + \frac{1}{2}b^2 b'' \qquad I_{(0,1),t_n,t_{n+1}} = \Delta_n \cdot \Delta W_n - \Delta Z_n$ identity

$(1,1) \qquad f_{(1,1)} = L^1 b = bb' \qquad I_{(1,1),t_n,t_{n+1}} = \frac{1}{2}\{(\Delta W_n)^2 - \Delta_n\}$

$$\boxed{\begin{aligned} Y_{n+1} = Y_n &+ a\Delta_n + b\Delta W_n + \tfrac{1}{2}bb'\{(\Delta W_n)^2 - \Delta_n\} \\ &+ \tfrac{1}{2}\left(aa' + \tfrac{1}{2}b^2 a''\right)\Delta_n^2 + ba'\,\Delta Z_n \\ &+ \left(ab' + \tfrac{1}{2}b^2 b''\right)\{\Delta_n \cdot \Delta W_n - \Delta Z_n\} \end{aligned}}$$

$\Delta W_n, \Delta Z_n$ correlated Gaussian RVs

Comparison of strong and weak Taylor schemes

indices $\{0,1\}$

$\gamma = 0.5$ $\quad \mathcal{A}^s_{0.5} = \{\phi, (0), (1)\}$

Euler scheme

$\beta = 1.0$ $\quad \mathcal{A}^w_{1.0} = \{\phi, (0), (1)\}$

plus $(1,1,1)$ in $\mathcal{A}^s_{1.5}$

$\gamma = 1.0$ $\quad \mathcal{A}^s_{1.0} = \{\phi, (0), (1), (1,1)\}$ Milstein scheme

$\beta = 2.0$ $\quad \mathcal{A}^w_{2.0} = \{\phi, (0), (1), (1,1), (1,0), (0,1), (0,0)\}$ order 2.0 weak Taylor

$\gamma = 2.0$ $\quad \mathcal{A}^s_{2.0} = \left\{\begin{matrix} \phi, (0), (1), (1,1), (1,0), (0,1), (0,0) \\ (1,1,1), (0,1,1), (1,0,1), (1,1,0), (1,1,1,1) \end{matrix}\right\}$ order 2.0 strong Taylor

<u>In general</u> — the weak Taylor scheme of a given order $\beta = k$ has <u>fewer terms</u> than the corresponding strong Taylor scheme of order $\gamma = k$

<u>MORAL</u> — decide first if the task on hand really needs a strong approximation or if a weak approximation would suffice

Multi-dimensional noise $\quad m > 1$

index set $\{0, 1, \ldots, m\}$

for $\beta \geq 2.0$ weak Taylor schemes involve mixed multiple stochastic integrals such as

$$I_{(j_1,j_2),t_n,t_{n+1}} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW_s^{j_1}\, dW_t^{j_2} \qquad j_1, j_2 \in \{1, \ldots, m\},\ j_1 \neq j_2$$

which can be approximated by random fourier series

simplifications due to special structure should be exploited

- additive noise
- commutative noise

SIMPLIFIED WEAK TAYLOR SCHEMES

- the multiple stochastic integrals I_α in a weak Taylor scheme of order β can be replaced by random variables $\hat{I}_\alpha$ with sufficiently many lower moments approximating those of I_α to $O(\Delta_n^{\beta+1})$ without changing the weak order β of the scheme

Ito SDE $\qquad dX_t = a(X_t)\,dt + b(X_t)\,dW_t \qquad$ autonomous $d=m=1$

simplified weak Euler scheme

$$\boxed{Y_{n+1} = Y_n + a(Y_n)\,\Delta_n + b(Y_n)\,\widehat{\Delta W}_n}$$

retains weak order $\beta = 1.0$ if random variables $\widehat{\Delta W}_n$ satisfy

$$\left|E\left(\widehat{\Delta W}_n\right)\right| + \left|E\left((\widehat{\Delta W}_n)^3\right)\right| + \left|E\left((\widehat{\Delta W}_n)^2\right) - \Delta_n\right| \leq K.\Delta_n^2$$

eg. $\widehat{\Delta W}_n$ is 2-point distributed. $\quad P\left(\{\omega \in \Omega;\ \widehat{\Delta W}_n(\omega) = \pm\sqrt{\Delta_n}\}\right) = \frac{1}{2}$

simplified order 2.0 weak Taylor scheme

$$\boxed{\begin{aligned} Y_{n+1} = \ & Y_n + a.\Delta_n + b.\,\widehat{\Delta W}_n \\ & + \tfrac{1}{2} b b' \left\{(\widehat{\Delta W}_n)^2 - \Delta_n\right\} \\ & + \tfrac{1}{2}\left(b a' + a b' + \tfrac{1}{2} b^2 b''\right) \Delta_n \cdot \widehat{\Delta W}_n \\ & + \tfrac{1}{2}\left(a a' + \tfrac{1}{2} b^2 a''\right) \Delta_n^2 \end{aligned}}$$

, here ΔZ_n is replaced by $\Delta_n \cdot \widehat{\Delta W}_n$ and ΔW_n by $\widehat{\Delta W}_n$ with

$$\left|E\left(\widehat{\Delta W}_n\right)\right| + \left|E\left((\widehat{\Delta W}_n)^3\right)\right| + \left|E\left((\widehat{\Delta W}_n)^5\right)\right|$$
$$+ \left|E\left((\widehat{\Delta W}_n)^2 - \Delta_n\right)\right| + \left|E\left((\widehat{\Delta W}_n)^4 - 3\Delta_n^2\right)\right| \leq K.\Delta_n^3$$

eg. $\widehat{\Delta W}_n$ is 3-point distributed with

$$P\left(\{\omega \in \Omega;\ \widehat{\Delta W}_n(\omega) = \pm\sqrt{3\Delta_n}\}\right) = \frac{1}{6}\ , \quad P\left(\{\omega \in \Omega;\ \widehat{\Delta W}_n(\omega) = 0\}\right) = \frac{2}{3}$$

multi-dimensional case $d, m \geq 1$

$$dX_t^i = a^i(t, X_t)\, dt + \sum_{j=1}^m b^{i,j}(t, X_t)\, dW_t^j$$

simplified order 2.0 weak Taylor scheme (Talay)

$$\boxed{\begin{aligned} Y_{n+1}^i = Y_n^i + a^i \Delta_n + \tfrac{1}{2} L^0 a^i \Delta_n^2 \\ + \sum_{j=1}^m \left\{ b^{i,j} + \tfrac{1}{2}\left(L^0 b^{i,j} + L^j a^i\right)\Delta_n \right\} \widetilde{\Delta W}_n^j \\ + \frac{1}{2}\sum_{j_1,j_2=1}^m L^{j_1} b^{i,j_2} \left\{ \widetilde{\Delta W}_n^{j_1} \widetilde{\Delta W}_n^{j_2} + V_{j_1,j_2} \right\} \end{aligned}}$$

where
- $\widetilde{\Delta W}_n$ is 3-point distributed (as above)
- V_{j_1,j_2} are independent, 2-point distributed with

$$P\left(V_{j_1,j_2} = \pm \Delta_n\right) = \tfrac{1}{2}, \quad V_{j_1,j_1} = -\Delta_n, \quad V_{j_1,j_2} = -V_{j_2,j_1}$$

$j_2 = 1, \ldots, j_1 - 1$ ← $j_1 = 1, \ldots, m$ → $j_2 = j_1 + 1, \ldots, m$

$$\Delta Z_n^j = I_{(j,0)} \sim \Delta_n \widetilde{\Delta W}_n^j, \quad I_{(j_1,j_2)} \sim \widetilde{\Delta W}_n^{j_1} \widetilde{\Delta W}_n^{j_2} \pm \Delta_n$$

$j_1 \neq j_2, \quad j_i \neq 0$

DERIVATIVE-FREE WEAK TAYLOR SCHEMES

- weak Taylor schemes can also be simplified without loss of convergence order by using derivative-free counterparts of the coefficients $f_\alpha = f_{(j_1,\ldots,j_\ell)} = L^{j_1} \cdots L^{j_\ell} f$

derivative-free simplified order 2.0 weak scheme ($d = m = 1$ case) (Platen)

$$Y_{n+1} = Y_n + \tfrac{1}{2}\left\{a(\Upsilon_n) + a(Y_n)\right\}\Delta_n$$
$$+ \tfrac{1}{4}\left\{b(\Upsilon_n^+) + b(\Upsilon_n^-) + 2b(Y_n)\right\}\widetilde{\Delta W}_n$$
$$+ \frac{1}{4\sqrt{\Delta_n}}\left\{b(\Upsilon_n^+) - b(\Upsilon_n^-)\right\}\left\{\left(\widetilde{\Delta W}_n\right)^2 - \Delta_n\right\}$$

$\widetilde{\Delta W}_n$ 3-point distributed

with supporting values

$$\Upsilon_n = Y_n + a(Y_n)\Delta_n + b(Y_n)\widetilde{\Delta W}_n, \quad \Upsilon_n^\pm = Y_n + a(Y_n)\Delta_n \pm b(Y_n)\sqrt{\Delta_n}$$

VARIANCE REDUCTION

- in actual calculations only an <u>arithmetic average</u> of finitely many samples

$$\hat{E}_N^\delta = \frac{1}{N}\sum_{j=1}^{N} g\left(Y_{N_T}^\delta(\omega_j)\right)$$

to estimate $E\left(g\left(Y_{N_T}^\delta\right)\right)$.

- $\hat{E}_N^\delta$ is a <u>random variable</u> $\therefore$ need confidence intervals

width $\sim \frac{1}{\sqrt{N}}$ $\qquad$ $N \gg 1$ for accuracy

- <u>variance reduction</u> methods attempt to reduce the number of samples required by <u>estimating a different quantity</u> with the same mean, but smaller variance

based on the Girsanov measure transformation

original SDE

$$X_t^{s,x} = x + \int_s^t a\left(X_\tau^{s,x}\right)d\tau + \int_s^t b\left(X_\tau^{s,x}\right)dW_\tau$$

W is Wiener process w.r.t. measure P

transformed SDE

$$\tilde{X}_t^{s,x} = x + \int_s^t a\left(\tilde{X}_\tau^{s,x}\right)d\tau + \int_s^t b\left(\tilde{X}_\tau^{s,x}\right)\underbrace{\left\{dW_\tau - d\left(\tau,\tilde{X}_\tau^{s,x}\right)d\tau\right\}}_{d\tilde{W}_\tau}$$

$\tilde{W}$ is Wiener process w.r.t. transformed measure $\tilde{P}$

$$\frac{d\tilde{P}}{dP} = \frac{\theta_t}{\theta_s} \qquad \theta_t = \theta_s + \int_s^t \theta_\tau\, d\left(\tau,\tilde{X}_\tau^{s,x}\right)dW_\tau$$

Radon-Nikodym derivative

$$u(s,x) = E\left(g\left(X_T^{s,x}\right)\right) = \tilde{E}\left(g\left(\tilde{X}_T^{s,x}\right)\right) \qquad \text{mean wrt } \tilde{P}$$

$$= E\left(g\left(\tilde{X}_T^{s,x}\right)\frac{\theta_T}{\theta_s}\right) \qquad \text{mean wrt } P$$

$\leftarrow$ estimate this expectation

eg pick $\bar{u}$ similar to u

$$d(t,x) = \frac{-1}{\bar{u}(t,x)}\, b(x)\frac{\partial}{\partial x}\bar{u}(t,x)$$

EXTRAPOLATION METHODS

- higher order weak convergence can be obtained from a low order weak scheme by extrapolation due to the special form of leading error coefficients of weak approximations

Analogous to deterministic Richardson extrapolation

$$\mu(\delta) = \mu_{true} + K.\delta^2 + O(\delta^4)$$

$$\mu(2\delta) = \mu_{true} + 4K\delta^2 + O(\delta^4)$$

$$\mu_{true} = \frac{4\mu(\delta) - \mu(2\delta)}{3} + O(\delta^4)$$

Talay & Tubaro (1989) with stochastic Euler, $E(g(Y^{\delta}_{N_T}))$ is a weak order $\beta = 1.0$ approximation of $E(g(X_T$

order $\beta = 2.0$ extrapolation

$$\boxed{V_2^{\delta}(T) = 2E\left(g(Y^{\delta}_{N_T})\right) - E\left(g(Y^{2\delta}_{N_T})\right)}$$

General case (Kloeden & Platen)

order β weak scheme $\quad Y^{\delta_1}_{N_T}, Y^{\delta_2}_{N_T}, \ldots, Y^{\delta_{\beta+1}}_{N_T}$

$\delta_\ell = \delta . d_\ell \qquad 0 < d_1 < \cdots < d_{\beta+1} < \infty \qquad \sum_{l=1}^{\beta+1} a_\ell = 1$

$\sum_{\ell=1}^{\beta+1} a_\ell (d_\ell)^{\gamma} = 0$

$\gamma = \beta, \ldots, 2\beta - 1$

general order 2β extrapolation

$$\boxed{V^{\delta}_{2\beta} = \sum_{l=1}^{\beta+1} a_l . E\left(g(Y^{\delta_l}_{N_T})\right)}$$

eg $\beta = 2.0$, $d_\ell = \ell$

$$V^{\delta}_{4.0} = \frac{18}{11} E\left(g(Y^{\delta}_{N_T})\right) - \frac{9}{11} E\left(g(Y^{2\delta}_{N_T})\right) + \frac{2}{11} E\left(g(Y^{3\delta}_{N_T})\right)$$

also $\frac{32}{21}$ $\quad -\frac{12}{21}$ $\quad \frac{1}{21}$

ie can double convergence order $\qquad$ comparative complexity ?

MOMENT EQUATIONS

X d-dim

linear SDE $\quad dX_t = A(t)X_t\,dt + \sum_{j=1}^{m} B^j(t)X_t\,dW_t^j$

moment ordinary DEs

$m(t) = E(X_t) \qquad \frac{dm}{dt} = A(t)m$

$P(t) = E\left(X_t X_t^T\right) \qquad \frac{dP}{dt} = A(t)P + PA(t) + \sum_{j=1}^{m} B(t)PB(t$

$d \times d$ symmetric matrix

alternatively

estimate $E\left(g(Y_{N_T}^{\delta})\right)$ using a weak scheme for $Y_{N_T}^{\delta}$

- confidence intervals only

BUT can use any function g once the $Y_{N_T}^{\delta}$ have been calculated

Nonlinear SDE $\quad dX_t = aX_t(1-X_t)\,dt + bX_t\,dW_t$

$$\frac{d}{dt}E(X_t) = aE(X_t) - aE(X_t^2)$$

$$\frac{d}{dt}E(X_t^2) = (2a+b^2)E(X_t^2) - \underbrace{\cancel{2aE(X_t^3)}}_{\text{discard}}$$

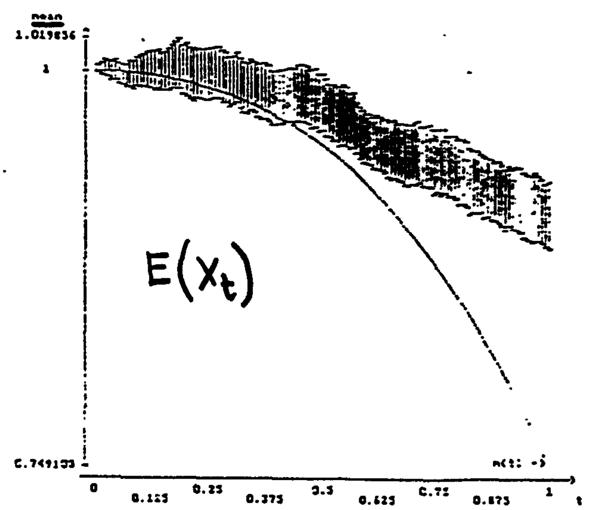
- usually a closed system of ODEs does not exis
- closure proceedure eg truncation

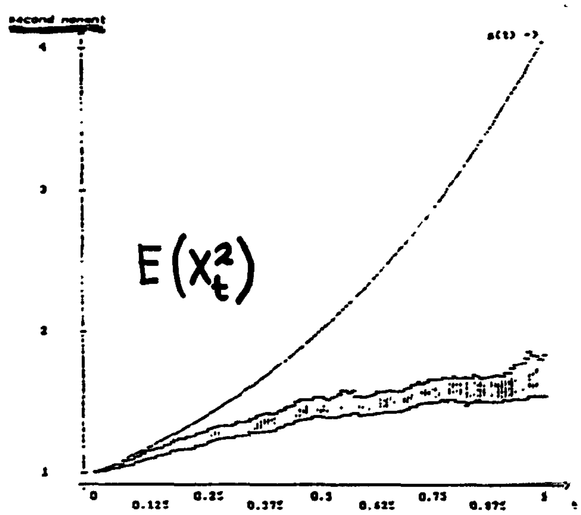
exact solution of truncated system

$$E(X_t) = E(X_0)e^{at} - \frac{a}{a+b^2}E(X_0^2)e^{(2a+b^2)t}$$

$$E(X_t^2) = E(X_0^2)e^{(2a+b^2)t}$$

- order 2.0 weak Taylor scheme, $\Delta_n \equiv 2^{-7}$
- 99% confidence intervals with $M = 20$ batches, $N = 200$ samples





FREQUENCY HISTOGRAMS

- often an SDE has a statistically stationary solution (ie with time <u>invariant measure</u> or density) which all other solutions approach asymptotically
- a weak scheme can be used to construct a <u>frequency histogram</u> of the measure's density
- ie simply count the number of realizations $X_T(\omega)$ for some large T falling into given partition subsets

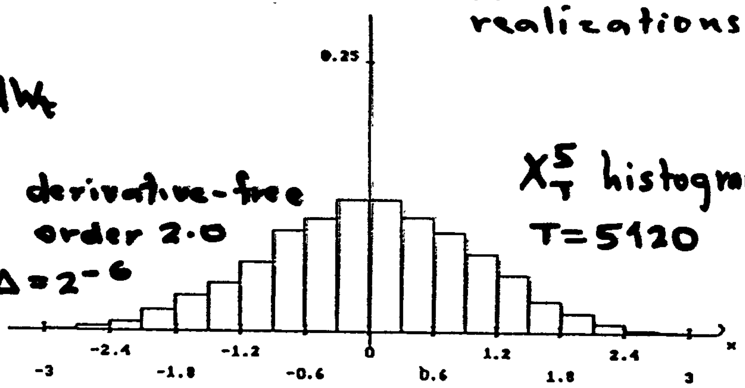
$$P\left(\{\omega\in\Omega;\ X_T(\omega)\in[x_i,x_{i+1}]\}\right) = E\left(\mathbb{1}_{[x_i,x_{i+1}]}(X_T)\right)$$

$$\simeq \frac{N_i}{N} \begin{array}{l}\leftarrow \text{realizations in } [x_i,x_{i+1}] \\ \leftarrow \text{total number of realizations}\end{array}$$

<u>Ornstein-Uhlenbeck Process</u>

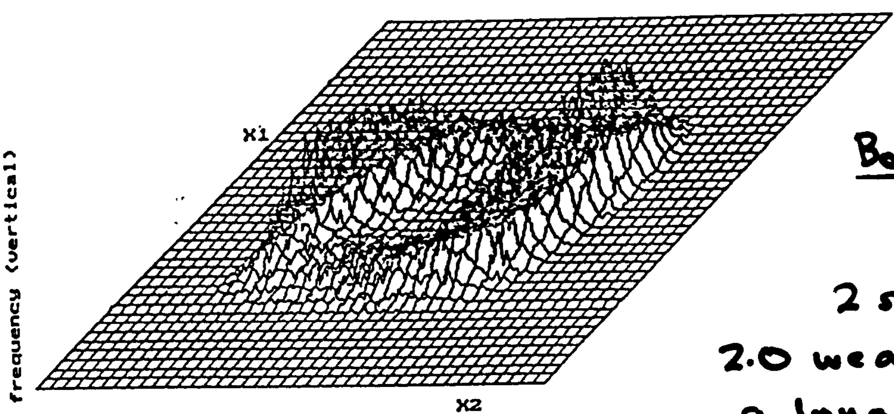
5-dim $\quad dX_t = AX_t\,dt + b\,dW_t$

$$A = \begin{bmatrix} 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \qquad b = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$



derivative-free order 2.0 $\Delta = 2^{-6}$

X_T^5 histogram $T = 5120$



<u>Bonhöffer - van der Pol system</u>

2 sample paths of order 2.0 weak Taylor scheme over a long-time interval

Ergodic process: typical sample path over a long-time interval behaves similarly to an ensemble of paths

LYAPUNOV EXPONENTS

Stratonovich

linear SDE $\qquad dX_t = AX_t\,dt + BX_t \circ dW_t \qquad d \geq 1,\ m = 1$

Lyapunov exponents measure asymptotic exponential rate of expansion or contraction

$$\lambda(x_0,\omega) = \overline{\lim_{t\to\infty}}\ \frac{1}{t} \ln \|X_t^{x_0}(\omega)\|$$

$\sim$ real parts of eigenvalues

Oseledec's Multiplicative Ergodic Theorem

① nonrandom $\quad \lambda_d \leq \cdots \leq \lambda_2 \leq \lambda_1$

② random $\quad E_d(\omega) \oplus \cdots \oplus E_1(\omega) = \mathbb{R}^d$

③ $\lambda(x_0,\omega) \leq \lambda_j \quad$ if $\quad x_0 \in E_j(\omega) \oplus \cdots \oplus E_1(\omega)$

- $\lambda_1 < 0 \qquad$ null solution $X_t \equiv 0$ asymptotically stable
- $\lambda_d \ll \lambda_1 \qquad$ stiff SDE, vastly different time scales

Polar coordinates $\qquad R_t = \|X_t\|, \quad S_t = X_t / \|X_t\| \in \mathcal{S}^{d-1}$ d-dim. sphere

$$\boxed{\begin{aligned} dR_t &= q(S_t)R_t\,dt + q_1(S_t)R_t \circ dW_t \\ dS_t &= h(S_t,A)\,dt + h(S_t,B) \circ dW_t \end{aligned}}$$

- $q_1(s) = s^T B s$
- $h(s,A) = \{A - (s^T A s)I\}s$
- $q(s) = s^T A s - (s^T B s)^2 + \frac{1}{2} s^T (B + B^T) s$

$$\ln R_t = \ln R_0 + \int_0^t q(S_\tau)\,d\tau + \int_0^t q_1(S_\tau) \circ dW_\tau$$

$$\lim_{t\to\infty} \frac{1}{t} \ln R_t = \lim_{t\to\infty} \frac{1}{t} \int_0^t q(S_\tau)\,d\tau$$

ergodicity

$$= \int_{\mathcal{S}^{d-1}} q(s)\,\mu(ds) = \lambda_1$$

invariant measure on $\mathcal{S}^{d-1}$ $\qquad$ top Lyapunov exponent

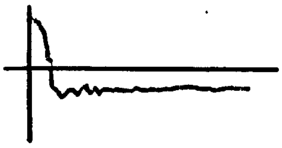
. invariant measure μ on S^{d-1}

- satisfies PDE on S^{d-1}
- difficult to solve numerically

- direct calculation of $\frac{1}{t}\ln\|x_t\|$

use weak scheme Y_n^δ, estimate $\frac{1}{n_T\delta}\ln\|Y_{n_T}^\delta\|$

avoid large logarithms
$$L_T^\delta = \frac{1}{n_T\delta}\sum_{n=1}^{n_T}\ln\left(\frac{\|Y_n^\delta\|}{\|Y_{n-1}^\delta\|}\right)$$



- take $T\to\infty$ until values stabilise

Stochastic Bifurcation (λ_1 changes sign)

noisy Brusselator equations

$$dX_t^1 = \left\{(\alpha-1)X_t^1 + \alpha (X_t^1)^2 + (X_t^1+1)^2 X_t^2\right\}dt + \sigma X_t^1(1+X_t^1)\,dW_t$$

$$dX_t^2 = \left\{-\alpha X_t^1 - \alpha (X_t^1)^2 - (X_t^1+1)^2 X_t^2\right\}dt - \sigma X_t^1(1+X_t^1)\,dW_t$$

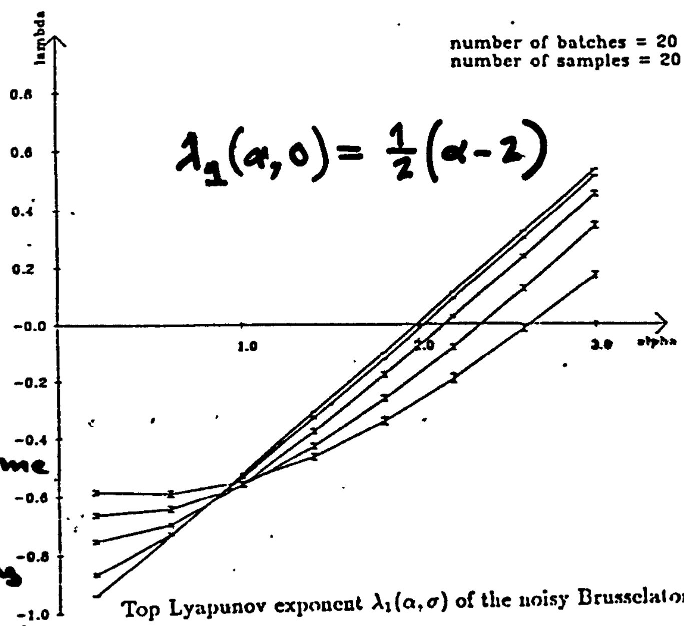
(0,0) undergoes Hopf bifurcation at $\alpha=2$ in the noise-free system

linearized system (Ito)

$$dX_t^1 = \left\{(\alpha-1)X_t^1 + X_t^2\right\}dt + \sigma X_t^1\,dW_t$$

$$dX_t^2 = \left\{-\alpha X_t^1 - X_t^2\right\}dt - \sigma X_t^1\,dW_t$$

order 2.0 weak Taylor scheme



Top Lyapunov exponent $\lambda_1(\alpha,\sigma)$ of the noisy Brusselator.

noisy Brusselator equations have "noisy" Hopf bifurcation at $\alpha \simeq 2$ (depending on σ)

EXAMPLE 1 – WEAK APPROXIMATION

5-dimensional Ornstein-Uhlenbeck process

$$d\begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \\ X_t^4 \\ X_t^5 \end{pmatrix} = \begin{bmatrix} 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \\ X_t^4 \\ X_t^5 \end{pmatrix} dt + \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} dW_t}_{\text{additive noise}}$$

scalar Wiener process

Order 2.0 weak Taylor scheme

$0 \leq t \leq T = 5120$

equidistant steps $\Delta = 2^{-6}$

Plotted histogram for $X_T^{(5)}$ using 5000 different sample paths