Slides for MAD 5932-01, SCS-MATH FSU Tallahassee*

Prof. R. Tempone

Version: Thursday August 31, 2006

*Based on the lecture notes *Stochastic and Partial Differential Equations with Adapted Numerics,* by J. Goodman, K.-S. Moon, A. Szepessy, R. Tempone, G. Zouraris.

Aug 29, 2006 - Class contents:

1. Course Introduction, Admin details

2. Motivating examples (Chapter 1)

3. Brief Probability review

Admin details

- syllabus

- class location

- student groups, email list, order of groups for assignments

- HMW presentations by groups on Thursdays. /Hand in days are Tuesdays for the group that makes the presentation/

- Matlab access, SCS computer facilities, personal accounts.

- course webpage

Course goal: to understand numerical methods for problems formulated by stochastic or partial differential equations models in science, engineering and mathematical finance.

Motivating examples (Chapter 1)

Example 1 (Noisy Evolution of Stock Values)

Denote stock value by $S(t)$. Assume that $S(t)$ satisfies the differential equation

$$\frac{dS}{dt} = a(t)S(t),$$

which has the solution

$$S(t) = e^{\int_0^t a(u)du} S(0).$$

Since we do not know precisely how $S(t)$ evolves we would like to generalize the model to a stochastic setting

$$a(t) = r(t) + \text{"}noise\text{"}.$$

For instance, we will consider

$$dS(t) = r(t)S(t)dt + \sigma S(t)dW(t), \qquad (1)$$

where $dW(t)$ *will introduce noise in the evolution.*

What is the meaning of (1)? The answer is not as direct as in the deterministic ode case.

One way to give meaning to (1) is to use the Forward Euler discretization,

$$S_{n+1} - S_n = r_n S_n \Delta t_n + \sigma_n S_n \Delta W_n. \qquad (2)$$

Here ΔW_n *are independent normally distributed random variables with zero mean and variance* Δt_n*, i.e.*

$$E[\Delta W_n] = 0$$

and

$$Var[\Delta W_n] = \Delta t_n = t_{n+1} - t_n.$$

Then (1) *is understood as a limit of* (2) *when* $\max \Delta t \to 0$.

Applications to Option pricing

European call option: is a contract signed at time t which gives the right, but not the obligation, to buy a stock (or other financial instrument) for a fixed price K at a fixed future time $T > t$.

At time t the buyer pays the seller the amount $f(s, t; T)$ for the option contract.

What is a fair price for $f(s, t; T)$?

The Black-Scholes model for the value $f : (0,T) \times (0,\infty) \to \mathbb{R}$ of a European call option is the partial differential equation

$$\partial_t f + rs\partial_s f + \frac{\sigma^2 s^2}{2}\partial_s^2 f = rf, \; 0 < t < T,$$
$$f(s,T) = \max(s-K,0), \quad (3)$$

where the constants r and σ denote the riskless interest rate and the volatility, respectively.

Stochastic representation of $f(s,t)$ The Feynmann-Kač formula gives the alternative probability representation of the option price

$$f(s,t) = E[e^{-r(T-t)}\max(S(T)-K,0))|S(t)=s], \tag{4}$$

where the underlying stock value S is modeled by the stochastic differential equation (1) satisfying $S(t) = s$.
Thus, $f(s,t)$ is both the solution of a PDE (3) with the expected value of the solution of a SDE (4)!

Which one should we choose to discretize?

Example 2 (Porous media flow) *Consider an incompressible flow*

$$div(V) = 0, \tag{5}$$

and Darcy's law

$$V = -K\nabla P. \tag{6}$$

Here V is the flow velocity and P is the pressure field. The function K, the so called conductivity of the material, is the source of randomness, since in practical cases, it is not precisely known.

To study the concentration C of an inert pollutant carried by the flow V, we solve the convection equation

$$\partial_t C + V \cdot \nabla C = 0.$$

Observe: The variation of K is, via Darcy's law (6), determines the concentration C.

One way to determine the flow velocity is to solve the pressure equation

$$div(K \nabla P) = 0, \tag{7}$$

in a domain with given values of the pressure on the boundary of this domain.

Assume now that the flow is two dimensional with $V = (1, \widehat{V})$*, where* $\widehat{V}(x)$ *is stochastic with mean zero, i.e.* $E[\widehat{V}] = 0$*. Thus,*

$$\partial_t C + \partial_x C + \widehat{V} \partial_y C = 0.$$

Let us define $\bar{C}$ *as the solution of*

$$\partial_t \bar{C} + \partial_x \bar{C} = 0.$$

Is $\bar{C}$ *is the expected value of* C*, i.e. is*

$$\bar{C} \stackrel{?}{=} E[C]$$

true?

The answer is in general no. The difference comes from the expected value

$$E[\widehat{V}\partial_y C] \neq E[\widehat{V}]E[\partial_y C] = 0.$$

Can you see why?

The desired averaged quantity $\widetilde{C} = E[C]$ *is an example of turbulent diffusion and in the simple case* $\widehat{V}(x)dx = dW(x)$ *(cf. (1)) it will satisfy a convection diffusion equation of the form*

$$\partial_t \widetilde{C} + \partial_x \widetilde{C} = \frac{1}{2}\partial_{yy}\widetilde{C},$$

which is related to the Feynman-Kač formula (4).

Example 3 (Optimal Control of Investments)

Suppose that we invest in a risky asset, whose value $S(t)$ evolves according to the stochastic differential equation

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t),$$

and in a riskless asset $Q(t)$ that evolves with

$$dQ(t) = rQ(t)dt.$$

It is reasonable to assume $r < \mu$, why?

Our total wealth is then $X(t) = Q(t) + S(t)$.

Goal: *determine an optimal instantaneous policy of investment to maximize the expected value of our wealth at a given final time* T.

Let the time dependent proportion,

$$\alpha(t) \in [0, 1],$$

be defined by

$$\alpha(t)X(t) = S(t),$$

so that

$$(1 - \alpha(t))X(t) = Q(t).$$

Then our optimal control problem can be stated as

$$\max_{\alpha\in\mathcal{A}} E[g(X(T))|X(t)=x] \equiv u(t,x), \qquad (8)$$

where g is a given function.

How can we determine α?

The solution to (8) can be obtained by means of a Hamilton Jacobi equation, which is in general a nonlinear partial differential equation satisfied by $u(t, x)$ of the form

$$u_t + H(u, u_x, u_{xx}) = 0.$$

Part of our work is to study the theory of Hamilton Jacobi equations and numerical methods for control problems to determine the Hamiltonian H and the control α.

Class contents:

1. Probability background

2. Wiener process

3. Ito integral

Probability Background

A probability space is a triple $(\Omega, \mathcal{F}, P)$, where Ω is the set of outcomes, $\mathcal{F}$ is the set of events and $P : \mathcal{F} \to [0, 1]$ is a function that assigns probabilities to events satisfying certain rules.

Definition 1 (Measurable Space) *If* Ω *is a given non empty set, then a* σ-algebra $\mathcal{F}$ *on* Ω *is a collection* $\mathcal{F}$ *of subsets of* Ω *that satisfy:*

(1) $\Omega \in \mathcal{F}$;

(2) $F \in \mathcal{F} \Rightarrow F^c \in \mathcal{F}$, *where* $F^c = \Omega - F$ *is the complement set of* F *in* Ω; *and*

(3) $F_1, F_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{+\infty} F_i \in \mathcal{F}$.

Definition 2 (Probability Measure) *A probability measure on* $(\Omega, \mathcal{F})$ *is a set function* $P : \mathcal{F} \to [0, 1]$ *such that:*

(1) $P(\emptyset) = 0$, $P(\Omega) = 1$; *and*

(2) *If* $A_1, A_2, \ldots \in \mathcal{F}$ *are mutually disjoint sets then*

$$P\left(\bigcup_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} P(A_i).$$

Question 1: Give an example of a probability space and distinguish clearly the events $F \in \mathcal{F}$ from the outcomes $\omega \in \Omega$.

Question 2: Give an example of two different σ-algebras, $\mathcal{G} \subset \mathcal{F}$ for the same set of outcomes Ω. Can you give an intuitive interpretation of the relation $\mathcal{G} \subset \mathcal{F}$?

Question 3: Is the intersection of σ-algebras still a σ-algebra?

Question 4: What about the union of σ-algebras?

Definition 3 (generated σ-algebra) *Given a family of sets, $\{A_n\}$, there exists a unique σ-algebra, $\sigma(\{A_n\})$, s.t.*

1. $\{A_n\} \subset \sigma(\{A_n\})$,

2. *if* $\mathcal{F}$ *is a* $\sigma - algebra$,

$$\{A_n\} \subset \mathcal{F} \Rightarrow \sigma(\{A_n\}) \subset \mathcal{F}$$

Definition 4 *A random variable* X*, in the probability space* $(\Omega, \mathcal{F}, P)$*, is a function*

$$X : \Omega \to \mathbb{R}^d,$$

such that the inverse image

$$X^{-1}(A) \equiv \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F},$$

for all open subsets A *of* $\mathbb{R}^d$.

Equivalently, we may say that X is an $\mathcal{F}$-measurable function and write $X \in \mathcal{F}$.

Example: Consider a finite family of disjoint sets,

$$\{A_n\}_{n=1}^{N}$$

and let $\Omega \equiv \cup_{1 \le n \le N} A_n$, $\mathcal{F} \equiv \sigma(\{A_n\})$. What condition has to satisfy

$$X : \Omega \to \mathbb{R}$$

in order to be a random variable in $(\Omega, \mathcal{F})$?

Definition 5 (Independence) *Two sets* $A, B \in \mathcal{F}$ *are said to be independent if*

$$P(A \cap B) = P(A)P(B).$$

Two independent random variables X, Y *in* $\mathbb{R}^d$ *are independent if for all open sets* $A, B \subseteq \mathbb{R}^d$ *we have that the events*

$$X^{-1}(A) \text{ and } Y^{-1}(B) \text{ are independent }. \quad (9)$$

Definition 6 (Expected value) *Let* $X : \Omega \to \mathbb{R}$ *be a random variable and suppose that the density function*

$$p'(x) = \frac{P(X \in dx)}{dx}$$

is integrable. The expected value of X *is then defined by the integral*

$$E[X] = \int_{-\infty}^{\infty} xp'(x)dx, \tag{10}$$

which also can be written

$$E[X] = \int_{-\infty}^{\infty} xdp(x). \tag{11}$$

The last integral makes sense also in general when the density function is a measure, e.g. by successive approximation with random variables possessing integrable densities. A point mass, i.e. a Dirac delta measure, is an example of a measure.

Definition 7 (Stochastic Process) *A stochastic process* $X : [0,T] \times \Omega \to \mathbb{R}^d$ *in the probability space* $(\Omega, \mathcal{F}, P)$ *is a function such that* $X(t,\cdot)$ *is a random variable in* $(\Omega, \mathcal{F}, P)$ *for all* $t \in (0,T)$. *We will often write* $X(t) \equiv X(t,\cdot)$.

The t variable will usually be associated with the notion of time.

Definition 8 (Wiener process) *The one dimensional* Wiener process $W : [0, \infty) \times \Omega \to \mathbb{R}$, *also known as the Brownian motion, has the following properties:*

(1) *with probability 1, the mapping* $t \mapsto W(t)$ *is continuous and* $W(0) = 0$*;*

(2) *if* $0 = t_0 < t_1 < \ldots < t_N = T$, *then the Wiener increments*

$$W(t_N) - W(t_{N-1}), \ldots, W(t_1) - W(t_0)$$

are independent*; and*

(3) *for all* $t > s$ *the increment* $W(t) - W(s)$ *has the* normal *distribution, with* $E[W(t) - W(s)] = 0$ *and* $E[(W(t) - W(s))^2] = t - s$, *i.e. for real intervals* Γ *we have*

$$P(W(t) - W(s) \in \Gamma) = \frac{1}{\sqrt{2\pi(t-s)}} \int_\Gamma e^{\frac{-y^2}{2(t-s)}} dy.$$

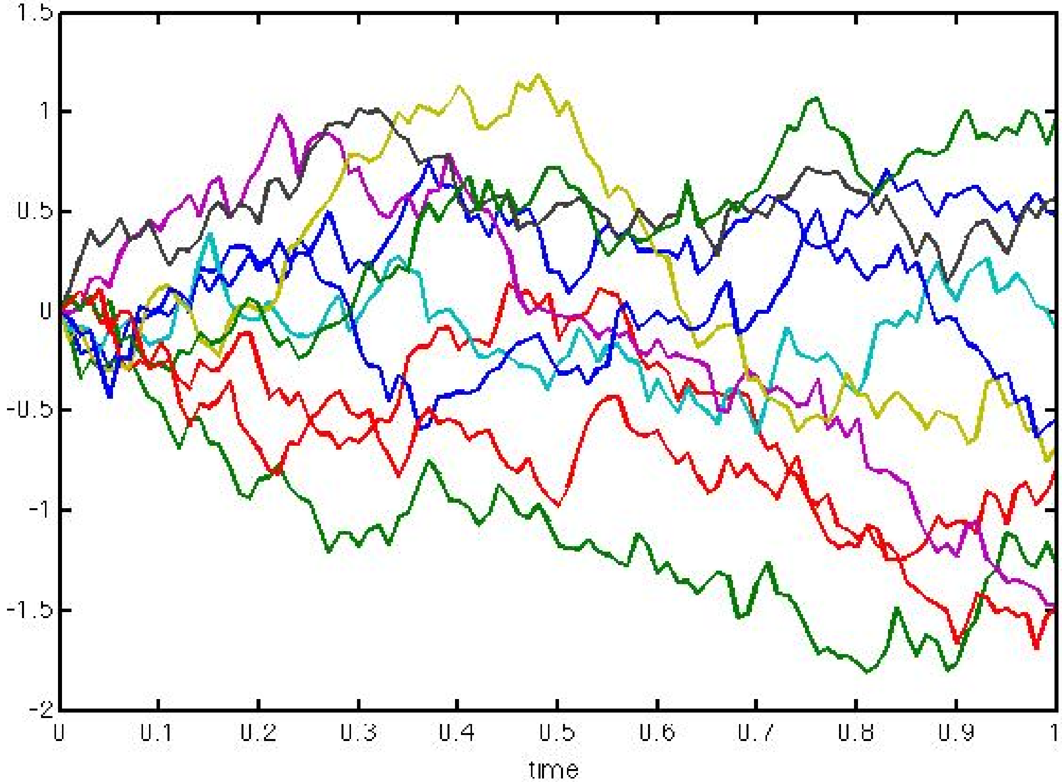
Problem: Given the times points

$$0 = t_0 < t_1 < \ldots < t_N = T.$$

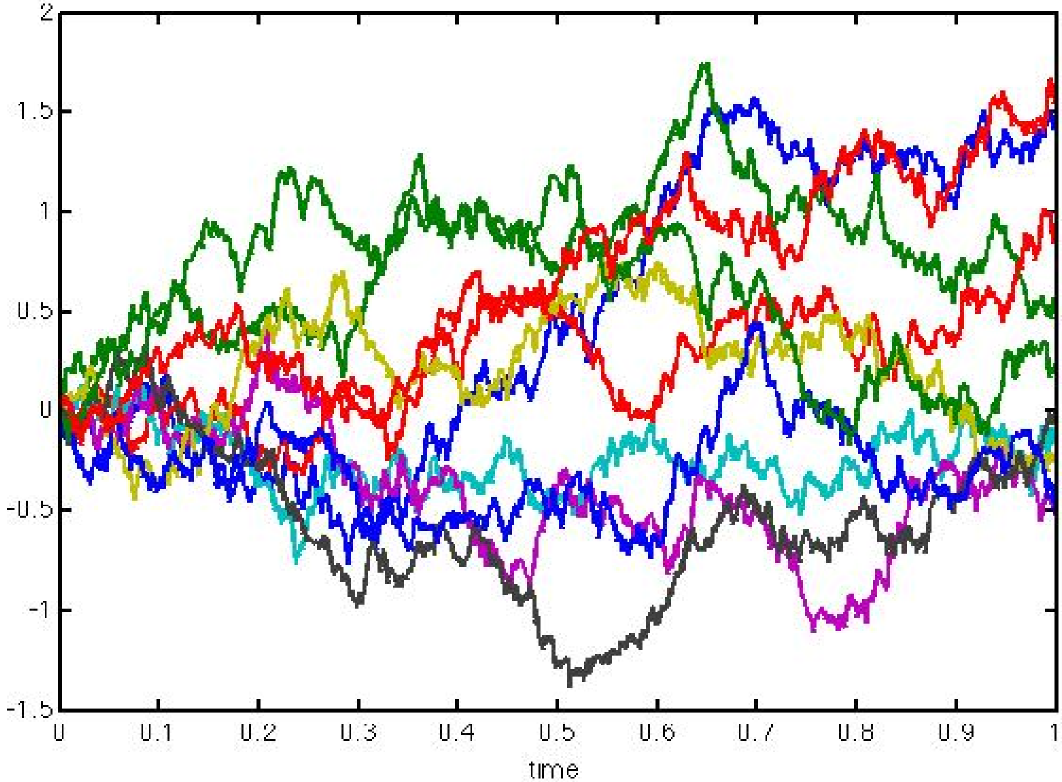
How do we sample realizations of $W(t_n)$, for $n = 0, \ldots, N$?

Sampling W at discrete times

```
>> N = 100;
>> dt = 1/N;
>> M = 10;
>> dW = sqrt(dt)*randn(N,M);
>> W = cumsum(dW);
>> W = [zeros(1,M);W];
>> t = linspace(0,1,N+1);
>> plot(t,W)
```



Sampling W on $[0,1]$, $N = 10^2$ uniform time steps, $M = 10$ realizations



Sampling W on $[0, 1]$, $N = 10^3$ uniform time steps, $M = 10$ realizations

Question: Which changes are needed to extend the previous code to variable step size, i.e. to sample W at given times, not necessarily evenly distributed,

$$0 = t_0 < \ldots < t_N?$$

Observe that the previous code is vectorized (no for loops) so it runs faster in MATLAB. What about its memory use?

Approximation and Definition of Stochastic Integrals

Questions on the definition of a stochastic integral

Remark 1 Problem: *How to define the stochastic integral* $\int_0^T W(t)dW(t)$, *where* $W(t)$ *is the Wiener process.*

Can we use the same approach as with Riemann integrals, taking sums

$$\sum_{n=0}^{N-1} W(\xi_n)(W(t_{n+1}) - W(t_n))$$

with $\xi_n \in [t_n, t_{n+1}]$*?*

As a first step, use the forward Euler *discretization*

$$\sum_{n=0}^{N-1} W(t_n) \underbrace{(W(t_{n+1}) - W(t_n))}_{=\Delta W_n}.$$

Taking expected values we obtain (why?)

$$\begin{aligned} E[\sum_{n=0}^{N-1} W(t_n)\Delta W_n] &= \sum_{n=0}^{N-1} E[W(t_n)\Delta W_n] \\ &= \sum_{n=0}^{N-1} E[W(t_n)] \underbrace{E[\Delta W_n]}_{=0} \\ &= 0. \end{aligned}$$

Now let us use instead the backward Euler *discretization*

$$\sum_{n=0}^{N-1} W(t_{n+1})\Delta W_n.$$

Taking expected values yields a different re-

sult:

$$\begin{aligned}
&\sum_{n=0}^{N-1} E[W(t_{n+1})\Delta W_n] \\
&\quad = \sum_{n=0}^{N-1} E[W(t_n)\Delta W_n] + E[(\Delta W_n)^2] \\
&\quad = \sum_{n=0}^{N-1} \Delta t \\
&\quad = T \neq 0.
\end{aligned}$$

Moreover, if we use the trapezoidal *method the result is*

$$\sum_{n=0}^{N-1} E\left[\frac{W(t_{n+1}) + W(t_n)}{2}\Delta W_n\right]$$
$$= \sum_{n=0}^{N-1} E[W(t_n)\Delta W_n] + E[(\Delta W_n)^2/2]$$
$$= \sum_{n=0}^{N-1} \frac{\Delta t}{2} = T/2 \neq 0.$$

$\square$

Conclusion: we need more information to define $\int_0^T W(s)dW(s)$ than to define a deterministic integral!

In fact, limits of the forward Euler define the so called *Itô integral*, while the trapezoidal method yields the so called *Stratonovich integral*.

Strong and weak convergence

Depending on the application, we focus either on

- *strong convergence*, where approximation of the outcomes of $X(T)$ is relevant,

- or *weak convergence*, where only the distribution (law) of $X(T)$ needs to be approximated.

Definition. The sequence of random variables $\{Y_n\}_{n\in\mathbb{N}}$ converges *strongly* to the random variable Y if

$$\|Y - Y_n\|_{L^2_P(\Omega)} \equiv \sqrt{E[(Y - Y_n)^2]} \to 0$$

Obs: By Chebychev we have

$$P(|Y - Y_n| \geq \epsilon) \leq \frac{E[(Y - Y_n)^2]}{\epsilon^2} \to 0$$

for ay fixed $\epsilon > 0$.

Definition. The sequence of random variables $\{Y_n\}_{n\in\mathbb{N}}$ converges *weakly* to the random variable Y if $E[g(Y)] - E[g(Y_n)] \to 0$, for all bounded continuous functions g.

Observe: strong convergence $\Rightarrow$ weak convergence, but the converse is in general not true.

Strong and weak convergence

Counterexample. Let random variables $\{Y_n\}_{n\in\mathbb{N}}$ be *iid* in $(\Omega, \mathcal{F}, P)$, and $Y_n \sim N(0,1), n = 1, \ldots$.
Verify that Y_n converges weakly but not strongly!

Proof of ($\Rightarrow$) for Lipschitz functions:

$$
\begin{aligned}
|E[g(X) - g(Y_n)]| &\leq E[|g(X) - g(Y_n)|] \\
&\leq C_g E[|X - Y_n|] \\
&\leq C_g \underbrace{\sqrt{E[|X - Y_n|^2]}}_{=\|X - Y_n\|_{L^2_P(\Omega)}} \to 0.
\end{aligned}
$$

Obs: The previous estimate may not be optimal. There are cases where the weak error goes to zero much faster than the strong one.

Ito Integrals

Theorem 1 *Suppose that there exists* $C > 0$ *s.t.* $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ *satisfies*

$$|f(t+\Delta t, W+\Delta W) - f(t,W)| \leq C(\Delta t + |\Delta W|)$$

then the forward Euler (left point quadrature) approximations

$$I_{\Delta t} = \sum_{n=0}^{N-1} f(t_n, W(t_n))\Delta W_n,$$

with $0 = t_0 < t_1 < \ldots < t_N = T$, *satisfy*

$$\|I_{\Delta t} - I_{\Delta t'}\|_{L^2_P(\Omega)} = E[I_{\Delta t} - I_{\Delta t'}]^{1/2} \leq \mathcal{O}\left(\sqrt{\Delta t_{\max}}\right) \tag{12}$$

Ito integrals

Remark 2 *The previous theorem implies that $I_{\Delta t}$ is Cauchy in $L^2_P(\Omega)$ and its limit defines the Ito integral*

$$\sum_{n=0}^{N-1} f(t_n, W(t_n))\Delta W_n$$
$$= \int_0^T f(s, W(s))dW(s) + \mathcal{O}\left(\sqrt{\Delta t_{\text{max}}}\right)$$

The previous estimate should be understood *as*

$$E\Big[\Big(\sum_{n=0}^{N-1} f(t_n, W(t_n))\Delta W_n - \int_0^T f(s, W(s))dW(s)\Big)^2\Big] \\ = \mathcal{O}\left(\Delta t_{\text{max}}\right)$$

Question: What is the computational work to reach an accuracy ϵ in $L^2_P(\Omega)$ sense using uniform time steps?

Information generated by a process.

Definition 9 *The symbol* $\mathcal{F}_t^W$ *denotes* the information generated by W on the interval $[0, t]$. *If, based on the observation of the trajectory* $\{W(s), 0 \leq s \leq t\}$ *it is possible to decide if an event* $A \in \mathcal{F}$ *has occurred or not, then we write* $A \in \mathcal{F}_t^W$.

If the value of a random variable Z can be completely determined by the observations $\{W(s), 0 \leq s \leq t\}$ then we write $Z \in \mathcal{F}_t^W$.

A stochastic process g is called **adapted to the filtration** $\{\mathcal{F}_t^W\}_{t\geq 0}$ if $g(t) \in \mathcal{F}_t^W$ for all $t \geq 0$.

Obs Math Grads: The filtration $\{\mathcal{F}_t^W\}_{t\geq 0}$ is actually an increasing family of σ-algebras. See Øksendal's book, Chapter 3, for precise definition.

Examples

1. $A = \{W(10) < 5\}$

2. $Z = \int_0^1 W(s)ds$

3. $f(t) = \sup_{s \leq t} W(s)$

4. $g(t) = \sup_{s \leq t+1} W(s)$

Remark 3 (Extension to adapted Itô integration) *Itô integrals can be extended to adapted processes. Assume* $f : [0,T] \times \Omega \to \mathbb{R}$ *is* adapted *to the* filtration $\{\mathcal{F}_t^W\}_{t \geq 0}$ *and that there is a constant* C *such that*

$$\sqrt{E[|f(t+\Delta t,\omega) - f(t,\omega)|^2]} \leq C\sqrt{\Delta t}. \quad (13)$$

Then the proof of Theorem 1 shows that (12) still holds.

Theorem 2 (Basic properties of Itô integrals)

Suppose that $f, g : [0, T] \times \Omega \to \mathbb{R}$ are Itô integrable, e.g. $\mathcal{F}_t^W$-adapted and satifying (13), and that c_1, c_2 are constants in $\mathbb{R}$. Then:

(1)

$$\int_0^T (c_1 f(s,\cdot) + c_2 g(s,\cdot))dW(s) \\ = c_1 \int_0^T f(s,\cdot)dW(s) + c_2 \int_0^T g(s,\cdot)dW(s).$$

(2) $E\left[\int_0^T f(s,\cdot)dW(s)\right] = 0.$

(3)

$$E\left[\left(\int_0^T f(s,\cdot)dW(s)\right)\left(\int_0^T g(s,\cdot)dW(s)\right)\right]$$
$$= \int_0^T E\left[f(s,\cdot)g(s,\cdot)\right] ds.$$

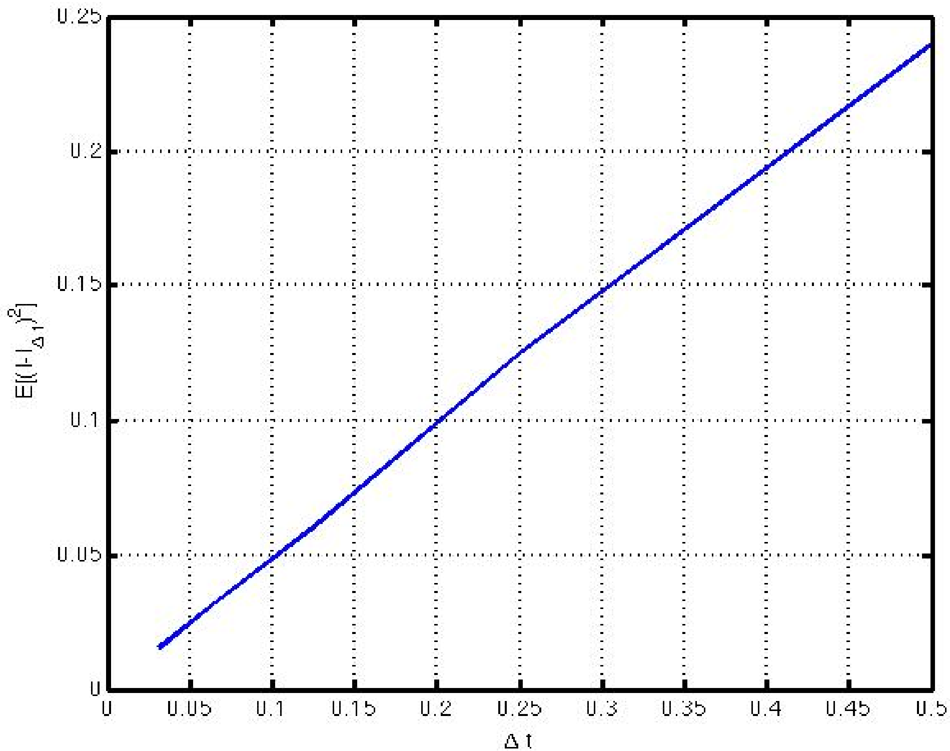
$\boxed{\textbf{Problem}}$ How can we approximate numerically the object $\int_0^T f(s, W(s))dW(s)$?

Approximation of $\int_0^1 W(s)dW(s)$:

```
dt = 1/N;
 t = linspace(0,1,N+1);
 dW = sqrt(dt)*randn(N,M);
 W = cumsum(dW);
 W = [zeros(1,M);W];
 I = 0.5*(W(N+1,:).^2-t(N+1));
 IFE = sum(W(1:N,:).*dW);
 Error(J,:) = I-IFE;
 mean_Square_Error = mean((Error.^2)')';
```

Question: Which changes are needed to extend the previous code to variable step size, i.e. to approximate I based on given points, not necessarily evenly distributed,

$$0 = t_0 < \ldots < t_N?$$



Strong approximation for
$\int_0^1 W(s)dW(s) = \frac{W^2(1)-1}{2}$ using F. Euler
with uniform time steps, $M = 10^3$.

Problem: Write a code to reproduce the previous results.

Question: Does the Wiener process really exist?

Answer: yes, see Example 2.18 in the notes, where we construct it as a limit of piecewise linear stochastic processes.

$\boxed{\textbf{Ito Stochastic Differential Equations}}$

$\boxed{\text{Goal}}$ Study existence, uniqueness and approximation for

$$\begin{aligned} dX_t =& a(t, X_t)dt + b(t, X_t)dW_t,\ 0 < t < T \\ X(0) =& x_0 \end{aligned}$$

which is understood as

$$\begin{aligned} X(t) = x_0 &+ \int_0^t a(s, X(s))ds \\ &+ \int_0^t b(s, X(s))dW_s,\ 0 < t < T \end{aligned} \tag{14}$$

$\boxed{\text{Idea}}$ As we did with Ito integrals, we define the solution to the SDE as the $L^2_P(\Omega)$ limit of forward Euler approximations.

Let $0 = t_0 < t_1 < \ldots < t_N = T$, then define the Euler Maruyama approximation by

$$\overline{X}(t_0) = x_0$$

and for $n = 0, \ldots, N-1$

$$\begin{aligned} \overline{X}(t_{n+1}) = \overline{X}(t_n) &+ a(t_n, \overline{X}(t_n))\Delta t_n \\ &+ b(t_n, \overline{X}(t_n))\Delta W_n, \end{aligned} \tag{15}$$

Remark 4 (Time continuous Forward Euler)

For theoretical purposes only we extend $\overline{X}$ *from* $0 = t_0 < t_1 < \ldots < t_N = T$*, to all* $t \in [0, T]$*.*
Let $t_n < t < t_{n+1}$*, then*

$$\begin{aligned}\overline{X}(t) = \overline{X}(t_n) &+ a(t_n, \overline{X}(t_n))(t - t_n) \\ &+ b(t_n, \overline{X}(t_n))(W(t) - W(t_n)),\end{aligned}$$

that is

$$\begin{aligned}\overline{X}(t) = \overline{X}(t_n) &+ \int_{t_n}^{t} a(t_n, \overline{X}(t_n))ds \\ &+ \int_{t_n}^{t} b(t_n, \overline{X}(t_n))dW(s)\end{aligned}$$

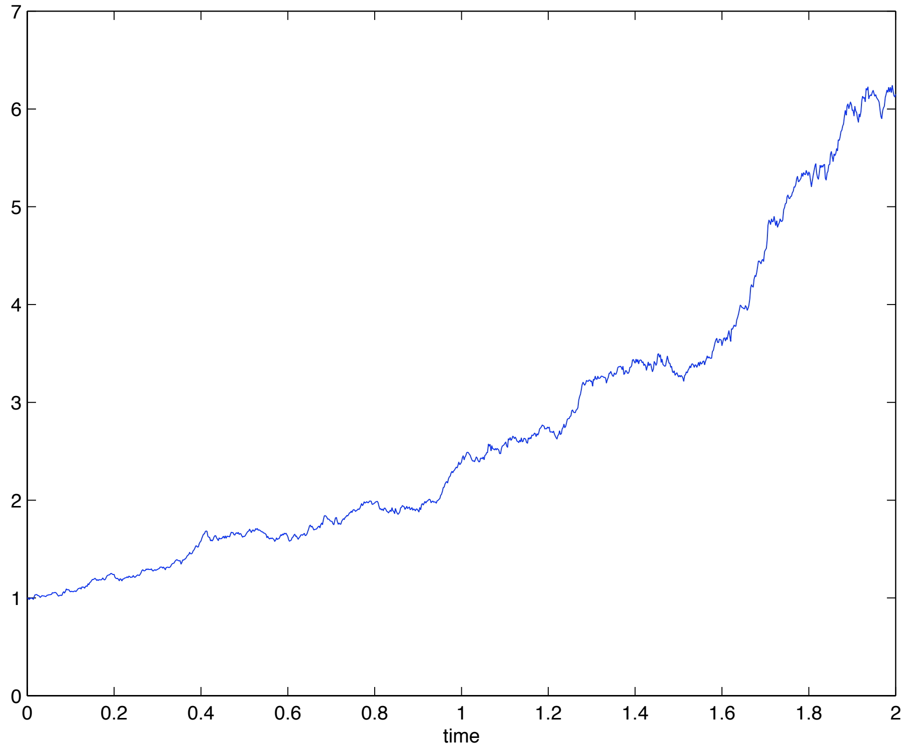
or, with the notations

$$\left.\begin{aligned} \overline{a}(s;\overline{X}) &= a(t_n,\overline{X}(t_n)),\\ \overline{b}(s;\overline{X}) &= b(t_n,\overline{X}(t_n)), \end{aligned}\right. \quad \text{for } t_n \leq s < t_{n+1}$$

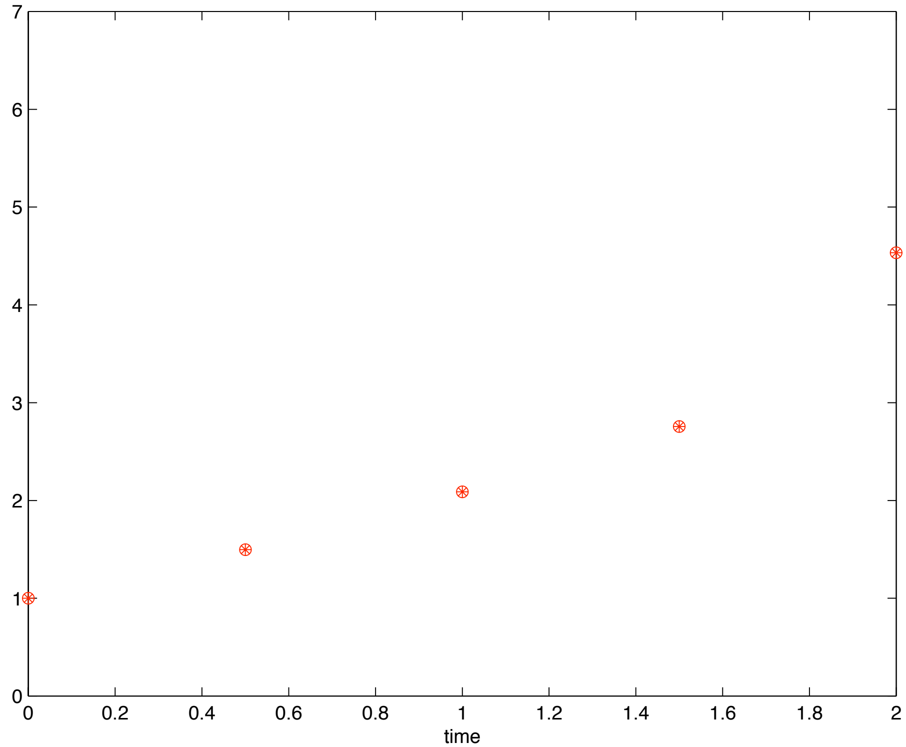
we can write

$$\begin{aligned} \overline{X}(t) = \overline{X}(t_n) &+ \int_{t_n}^{t} \overline{a}(s;\overline{X})ds \\ &+ \int_{t_n}^{t} \overline{b}(s;\overline{X})dW(s) \end{aligned}$$

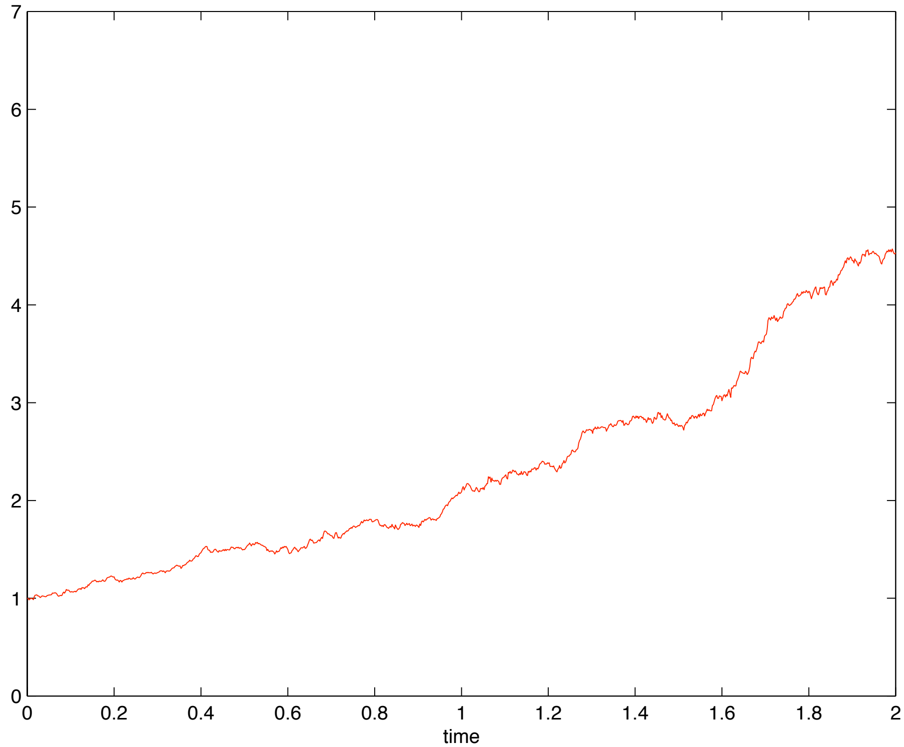
Observe that $\overline{a}$ *and* $\overline{b}$ *are simple functions and the integrals above are defined without limits.*



Exact X realization.



Discrete F. Euler approximation, $\overline{X}$.



Continuous time F. Euler approximation, $\overline{X}$.

Theorem 3 (Ito SDE strong approximation) *Let $\overline{X}_{\Delta t}$ and $\overline{X}_{\Delta t'}$ be Forward Euler approximations (15) to the Ito SDE (14). Assume that x_0 is deterministic and $\forall x, y \in \mathbb{R}$, $0 < s, t < T$ we have*

1.

$$\max\{|a(t,x) - a(t,y)|, |b(t,x) - b(t,y)|\} \leq C|x-y|, \quad (16)$$

2.

$$\max\{|a(t,x)-a(s,x)|,|b(t,x)-b(s,t)|\}$$
$$\leq C(1+|x|)\sqrt{|s-t|}. \tag{17}$$

Then, there exists a constant $K_T > 0$ *not depending on* $\Delta t, \Delta t'$ *s.t.*

$$\max_{t\in[0,T]} E[\left(\overline{X}_{\Delta t}(t)-\overline{X}_{\Delta t'}(t)\right)^2] \leq K_T \Delta t_{max}$$

and

$$\max_{t\in[0,T]} E[\left(\overline{X}_{\Delta t}(t)\right)^2] \leq K_T \tag{18}$$

$\boxed{\text{Proof of the thm.}}$

Step 1 Define the continuous Forward Euler and show (18)

Step 2 Consider the difference

$$
\begin{aligned}
(\overline{X}_{\Delta t}-\overline{X}_{\Delta t'})(s) = \\
&\int_0^s \Big(\overline{a}(s;\overline{X}_{\Delta t}(s)) - \overline{a}(s;\overline{X}_{\Delta t'}(s))\Big)\, ds \\
&+\int_0^s \Big(\overline{b}(s;\overline{X}_{\Delta t}(s)) - \overline{b}(s;\overline{X}_{\Delta t'}(s))\Big)\, dW(s)
\end{aligned}
$$

and estimate integrands using assumptions (16),(17).

Step 3 Use the Ito isometry

$$E[(\int f(t)dW_t)^2] = \int E[f^2(t)]dt$$

Step 4 Apply Grönwall's lemma to conclude.

Step 3

$$
\begin{aligned}
&E[(\overline{X}_{\triangle t} - \overline{X}_{\triangle t'})^2(s)] \\
&\le E\Big[\Big(\int_0^s \Big(\overline{a}(s; \overline{X}_{\triangle t}(s)) - \overline{a}(s; \overline{X}_{\triangle t'}(s))\Big)\, ds \\
&+ \int_0^s \Big(\overline{b}(s; \overline{X}_{\triangle t}(s)) - \overline{b}(s; \overline{X}_{\triangle t'}(s))\Big)\, dW(s)\Big)^2\Big] \\
&\le 2E\Big[\Big(\int_0^s \Big(\overline{a}(s; \overline{X}_{\triangle t}(s)) - \overline{a}(s; \overline{X}_{\triangle t'}(s))\Big)\, ds\Big)^2\Big] \\
&+2\, \underbrace{E\Big[\Big(\int_0^s \Big(\overline{b}(s; \overline{X}_{\triangle t}(s)) - \overline{b}(s; \overline{X}_{\triangle t'}(s))\Big)\, dW(s)\Big)^2\Big]}_{=\int_0^s E[(\overline{b}(s;\overline{X}_{\triangle t}(s)) - \overline{b}(s;\overline{X}_{\triangle t'}(s)))^2] ds}
\end{aligned}
$$

$$
\begin{aligned}
&E[(\overline{X}_{\triangle t} - \overline{X}_{\triangle t'})^2(s)] \\
&\leq 2E[s \int_0^s \left(\overline{a}(s; \overline{X}_{\triangle t}(s)) - \overline{a}(s; \overline{X}_{\triangle t'}(s))\right)^2 ds] \\
&\quad + 2 \int_0^s E[\left(\bar{b}(s; \overline{X}_{\triangle t}(s)) - \bar{b}(s; \overline{X}_{\triangle t'}(s))\right)^2] ds \\
&\leq 2s \int_0^s E[\left(\overline{a}(s; \overline{X}_{\triangle t}(s)) - \overline{a}(s; \overline{X}_{\triangle t'}(s))\right)^2] ds \\
&\quad + 2 \int_0^s E[\left(\bar{b}(s; \overline{X}_{\triangle t}(s)) - \bar{b}(s; \overline{X}_{\triangle t'}(s))\right)^2] ds
\end{aligned}
$$

Step 2 Recall that

$$E[\left(\bar{b}(s;\overline{X}_{\Delta t}(s)) - \bar{b}(s;\overline{X}_{\Delta t'}(s))\right)^2] =$$

$$E[\left(b(t_n,\overline{X}_{\Delta t}(t_n)) - b(t'_m,\overline{X}_{\Delta t'}(t'_m))\right)^2]$$

and estimate

$$\begin{aligned}
|b(t_n,\overline{X}_{\Delta t}(t_n)) - b(t'_m,\overline{X}_{\Delta t'}(t'_m)|
&\le |b(t_n,\overline{X}_{\Delta t}(t_n)) - b(t,\overline{X}_{\Delta t}(t)| \\
&+ \underbrace{|b(t,\overline{X}_{\Delta t}(t)) - b(t,\overline{X}_{\Delta t'}(t)|}_{\le C|\overline{X}_{\Delta t}(t) - \overline{X}_{\Delta t'}(t)|} \\
&+ |b(t,\overline{X}_{\Delta t'}(t)) - b(t'_m,\overline{X}_{\Delta t'}(t'_m)|
\end{aligned}$$

The first an last terms can be further estimated by

$$
\begin{aligned}
&|b(t_n, \overline{X}_{\Delta t}(t_n)) - b(t, \overline{X}_{\Delta t}(t)| \\
&\qquad \le \underbrace{|b(t_n, \overline{X}_{\Delta t}(t_n)) - b(t_n, \overline{X}_{\Delta t}(t)|}_{\le C|\overline{X}_{\Delta t}(t_n) - \overline{X}_{\Delta t}(t)|} \\
&\qquad + \underbrace{|b(t_n, \overline{X}_{\Delta t}(t)) - b(t, \overline{X}_{\Delta t}(t)|}_{\le C\sqrt{|t-t_n|}|\overline{X}_{\Delta t}(t)|}
\end{aligned}
$$

Finally, use assumptions to get

$$E[\left(\bar{b}(s;\overline{X}_{\Delta t}(s)) - \bar{b}(s;\overline{X}_{\Delta t'}(s))\right)^2] \\ \leq C\left(E[(\overline{X}_{\Delta t}(s) - \overline{X}_{\Delta t}(s))^2] + \Delta t_{max}\right)$$

and similarly for the drift term.

Combining **Step 2** and **Step 3** yields for $0 < s < T$

$$E[(\overline{X}_{\Delta t}(s) - \overline{X}_{\Delta t}(s))^2] \leq C\left(\int_0^s E[(\overline{X}_{\Delta t}(t) - \overline{X}_{\Delta t}(t))^2]dt + \Delta t_{max}\right),$$

From Grönwall's lemma then we arrive at

$$E[(\overline{X}_{\Delta t}(s) - \overline{X}_{\Delta t}(s))^2] \leq \Delta t_{max} C e^{Cs}$$

for $0 < s < T$ and this concludes the proof.

Remark 5 *The previous theorem shows that the Euler approximations are Cauchy in the norm*

$$\|X\|^2_{C^0([0,T];L^2_P(\Omega))} = \max_{t\in[0,T]} E[X^2(t)]$$

and in that sense the approximation error to the solution of the SDE is $\mathcal{O}\left(\Delta t_{max}^{1/2}\right)$

Question What is the corresponding weak error? Is it also $\mathcal{O}\left(\Delta t_{max}^{1/2}\right)$?

Ito's formula

This is a generalization of the deterministic chain rule. Let

$$dX_t = a(t, X_t)dt$$

then if $Y(t) = g(t, X(t))$ we have

$$dY(t) = (\partial_t g(t, X(t)) + \partial_x g(t, X(t))a(t, X_t))\, dt$$

Question: What is the corresponding formula for Ito SDEs?

Theorem 4 *Let the assumptions in* **Thm.** 3 *hold and let* $g : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$ *be bounded in* $C^2([0,T] \times \mathbb{R})$*. Then* $Y(t) \equiv g(t, X(t))$ *satisfies the SDE*

$$dY(t) = \left(\partial_t g + \partial_x g\, a + \partial_x^2 g \frac{b^2}{2}\right)(t, X(t))dt$$
$$+ \partial_x g(t, X(t)) b(t, X(t)) dW(t).$$

Proof: Assume g bounded in $C^3([0,T] \times \mathbb{R})$. We have to prove that

$$g(\tau, X(\tau)) - g(0, X(0)) = \int_0^\tau \left(\partial_t g + \partial_x g\, a + \partial_x^2 g \frac{b^2}{2}\right)(t, X(t))dt + \int_0^\tau \partial_x g(t, X(t)) b(t, X(t)) dW(t)$$

Consider a forward Euler approximation to X, $\overline{X}$. Then, in $L^2_P(\Omega)$, we have

$$g(\tau, X(\tau)) - g(0, X(0)) = g(\tau, \overline{X}(\tau)) - g(0, X(0)) + \mathcal{O}\left(\Delta t_{max}^{1/2}\right)$$

Now consider the telescopic sum

$$g(\tau, \overline{X}(\tau)) - g(0, X(0)) \\ = \sum_{n=0}^{N-1} \Big\{ g(t_{n+1}, \overline{X}(t_{n+1})) - g(t_n, \overline{X}(t_n)) \Big\}$$

and Taylor expand each term. To finish show that each obtained sum converges to the desired terms.

Taylor expansion: Let

$$\begin{aligned}\Delta \overline{X}_n =& \overline{X}(t_{n+1}) - \overline{X}(t_n) \\ =& a(t_n, \overline{X}(t_n))\Delta t_n + b(t_n, \overline{X}(t_n))\Delta W_n.\end{aligned}$$

Then

$$\begin{aligned} g(t_{n+1},&\overline{X}(t_{n+1})) - g(t_n, \overline{X}(t_n)) \\ &= \partial_t g \Delta t_n + \partial_x g \Delta \overline{X}_n \\ &+ \frac{\partial_{tt} g}{2}(\Delta t_n)^2 + \frac{\partial_{xx} g}{2}(\Delta \overline{X}_n)^2 + \partial_{tx} g \Delta t_n \Delta \overline{X}_n \\ &+ o((\Delta \overline{X}_n)^2) + o((\Delta t_n)^2)\end{aligned}$$

$$
\begin{aligned}
& g(\tau,\overline{X}(\tau)) - g(0,X(0)) \\
& \quad = \sum_{n=0}^{N-1} \Big\{ g(t_{n+1},\overline{X}(t_{n+1})) - g(t_n,\overline{X}(t_n)) \Big\} \\
& \quad = \underbrace{\sum_{n=0}^{N-1} \partial_t\, g \Delta t_n}_{\to \int_0^\tau \partial_t\, g(s,X(s))ds} + \underbrace{\sum_{n=0}^{N-1} \partial_x g\, b(t_n,\overline{X}(t_n)) \Delta W_n}_{\to \int_0^\tau (\partial_x\, gb)(s,X(s))dW(s)} \\
& \quad + \sum_{n=0}^{N-1} \frac{\partial_{xx} g\, b^2}{2} (\Delta W_n)^2 \\
& \quad + \sum_{n=0}^{N-1} \{ \frac{\partial_{tt} g}{2} (\Delta t_n)^2 + \frac{\partial_{xx} g}{2} (\Delta \overline{X}_n)^2 + \partial_{tx} g \Delta t_n \Delta \overline{X}_n \} \\
& \quad + \sum_{n=0}^{N-1} \{ o((\Delta \overline{X}_n)^2) + o((\Delta t_n)^2) \}
\end{aligned}
$$

We have

$$\sum_{n=0}^{N-1}\{\frac{\partial_{tt}g}{2}(\Delta t_n)^2+\frac{\partial_{xx}g}{2}(\Delta \overline{X}_n)^2+\partial_{tx}g\Delta t_n\Delta \overline{X}_n\} \to 0,$$

and that

$$\sum_{n=0}^{N-1}\frac{\partial_{xx}g\, b^2}{2}(\Delta W_n)^2 \to \int_0^\tau \frac{(\partial_{xx}g\, b^2)(s, X(s))}{2}ds.$$

By the assumption on the boundedness of the third derivatives (recall Taylor remainder form), we obtain

$$\sum_{n=0}^{N-1}\{o((\Delta \overline{X}_n)^2)+o((\Delta t_n)^2)\} \to 0.$$

To verify that

$$\sum_{n=0}^{N-1} \frac{\partial_{xx} g\, b^2}{2} (\Delta W_n)^2 \to \int_0^\tau \frac{(\partial_{xx} g\, b^2)(s, X(s))}{2} ds$$

we prove instead that

$$\sum_{n=0}^{N-1} \frac{\partial_{xx} g\, b^2}{2} (\Delta W_n)^2 - \underbrace{\sum_{n=0}^{N-1} \frac{\partial_{xx} g\, b^2}{2} \Delta t_n}_{\to \int_0^\tau \frac{(\partial_{xx} g\, b^2)(s, X(s))}{2} ds} \to 0.$$

i.e.

$$(I) \equiv E\left[\left(\sum_{n=0}^{N-1} \underbrace{\frac{(\partial_{xx} g\, b^2)(t_n, \overline{X}(t_n))}{2}}_{\equiv \alpha_n} \{ (\Delta W_n)^2 - \Delta t_n \} \right)^2 \right]$$

Now compute, for $0 \le n < m \le N-1$,

$$E[\alpha_n \alpha_m \{(\Delta W_n)^2 - \Delta t_n\}\{(\Delta W_m)^2 - \Delta t_m\}]$$
$$= E[\alpha_n \alpha_m \{(\Delta W_n)^2 - \Delta t_n\}] \underbrace{E[\{(\Delta W_m)^2 - \Delta t_m\}]}_{=0}$$
$$= 0$$

Therefore

$$(I) = \sum_{n=0}^{N-1} E[\alpha_n^2((\Delta W_n)^2 - \Delta t_n)^2]$$
$$= \sum_{n=0}^{N-1} E[\alpha_n^2] \underbrace{E[((\Delta W_n)^2 - \Delta t_n)^2]}_{=C\Delta t_n^2} \le C\Delta t_{\max} \to 0$$

Remark 6 *The regularity assumptions for g can be weakened. Let $g = g(x)$. To prove the result for g bounded in $C^2(\mathbb{R})$ use a mollifier $0 \le \phi_\delta \in C^\infty(\mathbb{R})$ satisfying $\int_{\mathbb{R}} \phi_\delta = 1$ and $|supp(\phi_\delta)| = \mathcal{O}(\delta)$. Then consider*

$$g_\delta(x) = \int_{\mathbb{R}} g(y)\phi_\delta(x-y)dy \in C^\infty(\mathbb{R})$$

which for fixed $\delta > 0$ is bounded in $C^3(\mathbb{R})$ and converges pointwise to g in $C^2(\mathbb{R})$ as $\delta \to 0$.

Write

$$
\begin{aligned}
g(X_\tau) - g(x_0) = & g_\delta(X_\tau) - g_\delta(x_0) \\
& + (g - g_\delta)(X_\tau) - (g - g_\delta)(x_0) \\
= & \int_0^\tau \left(\partial_x g_\delta \, a + \partial_x^2 g_\delta \frac{b^2}{2} \right) (t, X(t)) dt \\
& + \int_0^\tau \partial_x g_\delta(t, X(t)) b(t, X(t)) dW(t) \\
& + (g - g_\delta)(X_\tau) - (g - g_\delta)(x_0)
\end{aligned}
$$

To conclude, let $0 < \delta \to 0$ *and show, using dominated convergence, that*

- $E[(g - g_\delta)^2(X_\tau)] \to 0$,
- $E[(\int_0^\tau (L(g - g_\delta)) (t, X(t)) dt)^2] \to 0$

- $$E\left[\left(\int_0^\tau \partial_x(g-g_\delta)(t,X(t))b(t,X(t))dW(t)\right)^2\right]$$
  $$= \int_0^\tau E[(\partial_x(g-g_\delta)(t,X(t))b(t,X(t)))^2]dt \to 0$$

For $g = g(t,x)$ we need to mollify also in the t-direction.