Math Finance Seminar: Numerical Simulation of SDEs

T. Shores

October 26, 2006
Outline

1. Brownian Motion
2. Stochastic Integrals
3. Stochastic Differential Equations
4. Euler-Maruyama Method
5. Convergence of EM Method
6. Milstein’s Higher Order Method
7. Linear Stability
8. Stochastic Chain Rule
9. Parting Shots
References


Standard (continuous) Brownian motion, or Wiener process, over $[0, T]$: a random variable $W(t)$ depending continuously on $t \in [0, T]$ such that

1. $W(0) = 0$ with probability 1.
2. For $0 \leq s < t \leq T$ the random variable
   \[ W(t) - W(s) \sim N(0, t - s). \]
3. For $0 \leq s < t < u < v \leq T$ the random variables $W(t) - W(s)$ and $W(v) - W(u)$ are independent.
Definition

Standard (continuous) Brownian motion, or Wiener process, over \([0, T]\): a random variable \(W(t)\) depending continuously on \(t \in [0, T]\) such that

1. \(W(0) = 0\) with probability 1.
2. For \(0 \leq s < t \leq T\) the random variable
   \[W(t) - W(s) \sim N(0, t - s).\]
3. For \(0 \leq s < t < u < v \leq T\) the random variables
   \(W(t) - W(s)\) and \(W(v) - W(u)\) are independent.
Standard (continuous) Brownian motion, or Wiener process, over \([0, T]\): a random variable \(W(t)\) depending continuously on \(t \in [0, T]\) such that

1. \(W(0) = 0\) with probability 1.
2. For \(0 \leq s < t \leq T\) the random variable
   \[
   W(t) - W(s) \sim N(0, t - s).
   \]
3. For \(0 \leq s < t < u < v \leq T\) the random variables
   \(W(t) - W(s)\) and \(W(v) - W(u)\) are independent.
Discretized Brownian Motion

Discretized Brownian motion over \([0, T]\) in \(N\) steps: a sequence of random variable \(W_j = W(t_j)\), where \(\delta t = T/N\) and \(t_j = j \delta t\), such that

1. \(W(0) = 0\) with probability 1.
2. For \(j = 1, 2, \ldots, N\), \(W_j = W_j + dW_j\).
3. For \(j = 1, 2, \ldots, N\), \(dW_j \sim N(0, \delta t)\).

Notice that items (1)–(3) of continuous Brownian motion follow from these conditions. In fact, thanks to independence and identical distributions,

\[
W_{j+k} - W_j = \sum_{i=1}^{k} dW_i \sim N(0, k \delta t).
\]

(Recall, for independent \(X, Y\), \(E[aX + bY] = aE[X] + bE[Y]\) and \(\text{var}(aX + bY) = a^2 \text{var}(X) + b^2 \text{var}(Y)\).)
Discretized Brownian Motion

Discretized Brownian motion over \([0, T]\) in \(N\) steps: a sequence of random variable \(W_j = W(t_j)\), where \(\delta t = T/N\) and \(t_j = j \delta t\), such that

1. \(W(0) = 0\) with probability 1.
2. For \(j = 1, 2, \ldots, N\), \(W_j = W_j + dW_j\).
3. For \(j = 1, 2, \ldots, N\), \(dW_j \sim N(0, \delta t)\).

Notice that items (1)–(3) of continuous Brownian motion follow from these conditions. In fact, thanks to independence and identical distributions,

\[
W_{j+k} - W_j = \sum_{i=1}^{k} dW_i \sim N(0, k \delta t).
\]

(Recall, for independent \(X, Y\), \(E[aX + bY] = aE[X] + bE[Y]\) and \(\text{var}(aX + bY) = a^2 \text{var}(X) + b^2 \text{var}(Y)\).)
Discretized Brownian motion over $[0, T]$ in $N$ steps: a sequence of random variable $W_j = W(t_j)$, where $\delta t = T/N$ and $t_j = j \delta t$, such that

1. $W(0) = 0$ with probability 1.
2. For $j = 1, 2, \ldots, N$, $W_j = W_j + dW_j$.
3. For $j = 1, 2, \ldots, N$, $dW_j \sim N(0, \delta t)$.

Notice that items (1)–(3) of continuous Brownian motion follow from these conditions. In fact, thanks to independence and identical distributions,

$$W_{j+k} - W_j = \sum_{i=1}^{k} dW_i \sim N(0, k \delta t).$$

(Recall, for independent $X, Y$, $E[aX + bY] = aE[X] + bE[Y]$ and $\text{var}(aX + bY) = a^2 \text{var}(X) + b^2 \text{var}(Y)$. )
Discretized Brownian Motion

Discretized Brownian motion over \([0, T]\) in \(N\) steps: a sequence of random variable \(W_j = W(t_j)\), where \(\delta t = T/N\) and \(t_j = j\delta t\), such that

1. \(W(0) = 0\) with probability 1.
2. For \(j = 1, 2, \ldots, N\), \(W_j = W_j + dW_j\).
3. For \(j = 1, 2, \ldots, N\), \(dW_j \sim N(0, \delta t)\).

Notice that items (1)–(3) of continuous Brownian motion follow from these conditions. In fact, thanks to independence and identical distributions,

\[
W_{j+k} - W_j = \sum_{i=1}^{k} dW_i \sim N(0, k\delta t).
\]

(Recall, for independent \(X, Y\), \(E[aX + bY] = aE[X] + bE[Y]\) and \(\text{var}(aX + bY) = a^2\text{var}(X) + b^2\text{var}(Y)\).)
Here is the file used by Higham. Let’s run it and play with the parameters. In particular, rem out the resetting of the random number generator:

```matlab
% BPATH2 Brownian path simulation: vectorized
randn('state',100) % set the state of randn
T = 1; N = 500; dt = T/N;
dW = sqrt(dt)*randn(1,N); % increments
W = cumsum(dW); % cumulative sum
plot([0:dt:T],[0,W],'r-') % plot W against t
xlabel('t','FontSize',16)
ylabel('W(t)','FontSize',16,'Rotation',0)
```
We can also simulate random walks that are functions of Brownian motion. Here is the example of

\[ X(t) = u(W(t), t) = e^{(t + \frac{1}{2}W(t))} \]

% BPATH3 Function along a Brownian path
randn('state',100) % set the state of randn
T = 1; N = 500; dt = T/N; t = [dt:dt:1];
M = 1000; % M paths simultaneously
dW = sqrt(dt)*randn(M,N); % increments
W = cumsum(dW,2); % cumulative sum
U = exp(repmat(t,[M 1]) + 0.5*W);
Umean = mean(U);
plot([0,t],[1,Umean],'b-'), hold on % plot mean over M paths
plot([0,t],[ones(5,1),U(1:5,:)],'r--'), hold off % plot 5 individual paths
xlabel('t','FontSize',16)
ylabel('U(t)','FontSize',16,'Rotation',0,'HorizontalAlignment','right')
legend('mean of 1000 paths','5 individual paths',2)
averr = norm((Umean - exp(9*t/8)),'inf') % sample error
Let $W(t)$ be a Wiener process and $h(t)$ a function of $t$. Then we define

$$X(t) - X(0) = \int_0^t h(\tau) dW(\tau)$$

provided that $X(t)$ is a random process such that

$$X(t) - X(0) = \lim_{m \to \infty} \sum_{j=0}^{N-1} h(t_j) (W(t_{j+1}) - W(t_j))$$

where $0 = t_0 < t_1 < \cdots < t_N = t$ and $\max_j (t_{j+1} - t_j) \to 0$ as $N \to \infty$. 
Let $W(t)$ be a Wiener process and $h(t)$ a function of $t$. Then we define

$$X(t) - X(0) = \int_0^t h(\tau) \, dW(\tau)$$

provided that $X(t)$ is a random process such that

$$X(t) - X(0) = \lim_{m \to \infty} \sum_{j=0}^{N-1} h\left(\frac{t_j + t_{j+1}}{2}\right) (W(t_{j+1}) - W(t_j))$$

where $0 = t_0 < t_1 < \cdots < t_N = t$ and $\max_j (t_{j+1} - t_j) \to 0$ as $N \to \infty$. 

---

**Stratonovich**

Let $W(t)$ be a Wiener process and $h(t)$ a function of $t$. Then we define

$$X(t) - X(0) = \int_0^t h(\tau) \, dW(\tau)$$

provided that $X(t)$ is a random process such that

$$X(t) - X(0) = \lim_{m \to \infty} \sum_{j=0}^{N-1} h\left(\frac{t_j + t_{j+1}}{2}\right) (W(t_{j+1}) - W(t_j))$$

where $0 = t_0 < t_1 < \cdots < t_N = t$ and $\max_j (t_{j+1} - t_j) \to 0$ as $N \to \infty$. 

---

**Brownian Motion**  
**Stochastic Integrals**  
**Stochastic Differential Equations**  
**Euler-Maruyama Method**  
**Convergence of EM Method**  
**Milstein’s Higher Order Method**  
**Linear Stability**  
**Stochastic Chain Rule**  
**Parting Shots**
A Special Case: Ito

Take \( h(t) = W(t) \). Some shorthand: \( W_j = W(t_j) \),
\( W_{j+1/2} = W(t_j + \frac{\delta t}{2}) = W(t_j + t_{j+1}) \) and \( dW_j = W_{j+1} - W_j \).

Thus \( W_N = W(T) \) and \( W_0 = W(0) \). For the Ito integral:

- Note the identity
  \[
  b(a - b) = \frac{1}{2} (a^2 - b^2 - (a - b)^2)
  \]

- Hence
  \[
  \sum_{j=0}^{N-1} W_j (W_{j+1} - W_j) = \frac{1}{2} \sum_{j=0}^{N-1} \left( W_{j+1}^2 - W_j^2 - (dW_j)^2 \right)
  \]
  \[
  = \frac{1}{2} \left( W(T)^2 - W(0)^2 - \sum_{j=0}^{N-1} (dW_j)^2 \right)
  \]
A Special Case: Ito

Take \( h(t) = W(t) \). Some shorthand: \( W_j = W(t_j) \),
\( W_{j+1/2} = W(t_j + \frac{\delta t}{2}) = W(t_j + t_{j+1}) \) and \( dW_j = W_{j+1} - W_j \).
Thus \( W_N = W(T) \) and \( W_0 = W(0) \). For the Ito integral:

- Note the identity

\[
b(a - b) = \frac{1}{2} \left( a^2 - b^2 - (a - b)^2 \right)
\]

- Hence

\[
\sum_{j=0}^{N-1} W_j (W_{j+1} - W_j) = \frac{1}{2} \sum_{j=0}^{N-1} \left( W_{j+1}^2 - W_j^2 - (dW_j)^2 \right)
\]

\[
= \frac{1}{2} \left( W(T)^2 - W(0)^2 - \sum_{j=0}^{N-1} (dW_j)^2 \right)
\]
A Special Case: Ito

Now recall from statistics that i.i.d. r.v.’s $X_1, \ldots, X_N \sim N(\mu, \sigma^2)$, then

$$Y = \sum_{j=1}^{N} \left( \frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(N),$$

which has mean $N$ and variance $2N$.

Hence, since $\delta W_j \sim N(0, \delta t)$, we have that

$$\sum_{j=0}^{N-1} (dW_j)^2 = \sum_{j=0}^{N-1} (W_{j+1} - W_j)^2 \sim \delta t \chi^2(N).$$

Thus, this sum has mean $N \delta t = T$ and variance $\delta t^2 2N = 2T \delta t$.

So it is reasonable that the sum approaches $T$ as $\delta t \to 0$.

Hence

$$\int_{0}^{T} W(t) dW(t) = \frac{1}{2} W(T)^2 - \frac{1}{2} T.$$
A Special Case: Ito

- Now recall from statistics that i.i.d. r.v.’s $X_1, \ldots, X_N \sim N(\mu, \sigma^2)$, then

$$Y = \sum_{j=1}^{N} \left( \frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(N),$$

which has mean $N$ and variance $2N$.

- Hence, since $\delta W_j \sim N(0, \delta t)$, we have that

$$\sum_{j=0}^{N-1} (dW_j)^2 = \sum_{j=0}^{N-1} (W_{j+1} - W_j)^2 \sim \delta t \chi^2(N).$$

- Thus, this sum has mean $N \delta t = T$ and variance $\delta t^2 2N = 2T \delta t$.

- So it is reasonable that the sum approaches $T$ as $\delta t \to 0$.

- Hence

$$\int_0^T W(t) \, dW(t) = \frac{1}{2} W(T)^2 - \frac{1}{2} T.$$
A Special Case: Ito

Now recall from statistics that i.i.d. r.v.'s $X_1, \ldots, X_N \sim N(\mu, \sigma^2)$, then

$$Y = \sum_{j=1}^{N} \left( \frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(N),$$

which has mean $N$ and variance $2N$.

Hence, since $\delta W_j \sim N(0, \delta t)$, we have that

$$\sum_{j=0}^{N-1} (dW_j)^2 = \sum_{j=0}^{N-1} (W_{j+1} - W_j)^2 \sim \delta t \chi^2(N).$$

Thus, this sum has mean $N \delta t = T$ and variance $\delta t^2 2N = 2T \delta t$.

So it is reasonable that the sum approaches $T$ as $\delta t \to 0$.

Hence

$$\int_{0}^{T} W(t) \, dW(t) = \frac{1}{2} W(T)^2 - \frac{1}{2} T.$$
A Special Case: Ito

- Now recall from statistics that i.i.d. r.v.'s $X_1, \ldots, X_N \sim N (\mu, \sigma^2)$, then
  \[
  Y = \sum_{j=1}^{N} \left( \frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2 (N),
  \]
  which has mean $N$ and variance $2N$.

- Hence, since $\delta W_j \sim N (0, \delta t)$, we have that
  \[
  \sum_{j=0}^{N-1} (dW_j)^2 = \sum_{j=0}^{N-1} (W_{j+1} - W_j)^2 \sim \delta t \chi^2 (N).
  \]

- Thus, this sum has mean $N \delta t = T$ and variance $\delta t^2 2N = 2T \delta t$.

- So it is reasonable that the sum approaches $T$ as $\delta t \to 0$.

- Hence
  \[
  \int_{0}^{T} W (t) dW (t) = \frac{1}{2} W (T)^2 - \frac{1}{2} T.
  \]
A Special Case: Ito

Now recall from statistics that i.i.d. r.v.'s $X_1, \ldots, X_N \sim N(\mu, \sigma^2)$, then

$$Y = \sum_{j=1}^{N} \left( \frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(N),$$

which has mean $N$ and variance $2N$.

Hence, since $\delta W_j \sim N(0, \delta t)$, we have that

$$\sum_{j=0}^{N-1} (dW_j)^2 = \sum_{j=0}^{N-1} (W_{j+1} - W_j)^2 \sim \delta t \chi^2(N).$$

Thus, this sum has mean $N \delta t = T$ and variance $\delta t^2 2N = 2T \delta t$.

So it is reasonable that the sum approaches $T$ as $\delta t \to 0$.

Hence

$$\int_0^T W(t) \, dW(t) = \frac{1}{2} W(T)^2 - \frac{1}{2} T.$$
A Special Case: Stratonovich

Take $h(t) = W(t)$. For the Stratonovich integral:

- Note the identity

$$W_{j+1/2} = \frac{W_j + W_{j+1}}{2} + \frac{1}{2} (W_{j+1/2} - W_{j+1}) + \frac{1}{2} (W_{j+1/2} - W_j)$$

$$= \frac{W_j + W_{j+1}}{2} + \frac{1}{2} (-U_j) + \frac{1}{2} (V_j),$$

where $U_j, V_j \sim N(0, \frac{\delta t}{2})$ are independent r.v.'s.

- Note $W_{j+1} - W_j = U_j + V_j$ and set $\Delta Z_j = \frac{1}{2} (-U_j + V_j).$
A Special Case: Stratonovich

Take \( h(t) = \mathcal{W}(t) \). For the Stratonovich integral:

- Note the identity

\[
\mathcal{W}_{j+1/2} = \frac{\mathcal{W}_j + \mathcal{W}_{j+1}}{2} + \frac{1}{2} (\mathcal{W}_{j+1/2} - \mathcal{W}_{j+1}) + \frac{1}{2} (\mathcal{W}_{j+1/2} - \mathcal{W}_j)
\]

\[
= \frac{\mathcal{W}_j + \mathcal{W}_{j+1}}{2} + \frac{1}{2} (-U_j) + \frac{1}{2} (V_j),
\]

where \( U_j, V_j \sim N(0, \frac{\delta t}{2}) \) are independent r.v.’s.

- Note \( \mathcal{W}_{j+1} - \mathcal{W}_j = U_j + V_j \) and set \( \Delta Z_j = \frac{1}{2} (-U_j + V_j) \).
Now expand (at the board) and really get the telescoping effect, so the sum becomes

\[ \sum_{j=0}^{N-1} W_{j+1/2} (W_{j+1} - W_j) = \sum_{j=0}^{N-1} \left( \frac{W_j + W_{j+1}}{2} + \Delta Z_j \right) (W_{j+1} - W_j) = \frac{1}{2} \left( W(T)^2 - W(0)^2 \right) + \sum_{j=0}^{N-1} \Delta Z_j (W_{j+1} - W_j) \]

Each term in the latter sum is a \( \frac{1}{2} \left( V_j^2 - U_j^2 \right) \), so has mean zero and variance \( \frac{\delta t^2}{4} \), since \( U_j^2, V_j^2 \sim \frac{\delta t}{2} \chi^2(1) \) are independent.

Hence, the sum of these independent variables is a random variable of mean zero and variance \( N \delta t \frac{\delta t}{4} = \frac{T}{4} \delta t \).

Hence, \( \int_0^T W(t) \, dW(t) = \frac{1}{2} W(T)^2 \).
A Special Case: Stratonovich

Now expand (at the board) and really get the telescoping effect, so the sum becomes

$$\sum_{j=0}^{N-1} W_{j+1/2} (W_{j+1} - W_j) = \sum_{j=0}^{N-1} \left( \frac{W_j + W_{j+1}}{2} + \Delta Z_j \right) (W_{j+1} - W_j)$$

$$= \frac{1}{2} \left( W(T)^2 - W(0)^2 \right) + \sum_{j=0}^{N-1} \Delta Z_j (W_{j+1} - W_j)$$

Each term in the latter sum is a $\frac{1}{2} \left( V_j^2 - U_j^2 \right)$, so has mean zero and variance $\frac{\delta t^2}{4}$, since $U_j^2, V_j^2 \sim \frac{\delta t}{2} \chi^2 (1)$ are independent.

Hence, the sum of these independent variables is a random variable of mean zero and variance $N \delta t \frac{\delta t}{4} = \frac{T}{4} \delta t$.

Hence, $\int_0^T W(t) \, dW(t) = \frac{1}{2} W(T)^2$. 
A Special Case: Stratonovich

Now expand (at the board) and really get the telescoping effect, so the sum becomes

$$\sum_{j=0}^{N-1} W_{j+1/2} (W_{j+1} - W_j) = \sum_{j=0}^{N-1} \left( \frac{W_j + W_{j+1}}{2} + \Delta Z_j \right) (W_{j+1} - W_j)$$

$$= \frac{1}{2} \left( W(T)^2 - W(0)^2 \right) + \sum_{j=0}^{N-1} \Delta Z_j (W_{j+1} - W_j)$$

Each term in the latter sum is a $$\frac{1}{2} \left( V_j^2 - U_j^2 \right)$$, so has mean zero and variance $$\frac{\delta t^2}{4}$$, since $$U_j^2, V_j^2 \sim \frac{\delta t}{2} \chi^2 (1)$$ are independent.

Hence, the sum of these independent variables is a random variable of mean zero and variance $$N \delta t \frac{\delta t}{4} = \frac{T}{4} \delta t$$.

Hence, $$\int_0^T W(t) \, dW(t) = \frac{1}{2} W(T)^2$$.
Now expand (at the board) and really get the telescoping effect, so the sum becomes

$$\sum_{j=0}^{N-1} W_{j+1/2} (W_{j+1} - W_j) = \sum_{j=0}^{N-1} \left( \frac{W_j + W_{j+1}}{2} + \Delta Z_j \right) (W_{j+1} - W_j)$$

$$= \frac{1}{2} \left( W (T)^2 - W (0)^2 \right) + \sum_{j=0}^{N-1} \Delta Z_j (W_{j+1} - W_j)$$

Each term in the latter sum is a \( \frac{1}{2} \left( V_j^2 - U_j^2 \right) \), so has mean zero and variance \( \frac{\delta t^2}{4} \), since \( U_j^2, V_j^2 \sim \frac{\delta t}{2} \chi^2 (1) \) are independent.

Hence, the sum of these independent variables is a random variable of mean zero and variance \( N \delta t \frac{\delta t}{4} = \frac{T}{4} \delta t \).

Hence, \( \int_0^T W (t) \, dW (t) = \frac{1}{2} W (T)^2 \).
The file stint.m:

% Ito and Stratonovich integrals of W dW
randn('state',100) % set the state of randn
T = 1; N = 500; dt = T/N;
dW = sqrt(dt)*randn(1,N); % increments
W = cumsum(dW); % cumulative sum
ito = sum([0,W(1:end-1)]).*dW
strat = sum((0.5*([0,W(1:end-1)]+W) +
0.5*sqrt(dt)*randn(1,N)).*dW)
itoerr = abs(ito - 0.5*(W(end)^2-T))
straterr = abs(strat - 0.5*W(end)^2)
Deterministic Definitions

Deterministic Differential Equation:

To compute a function $x(t)$, $0 \leq t \leq T$, such that on the interval $[0, T]$, given $x(0)$ (this is an IVP, really):

- Derivative form: $\frac{dx}{dt} = f(x, t)$.
- Differential form: $dx = f(x, t) \, dt$.
- Integral form: $x(t) = x(0) + \int_{0}^{t} f(x(s), s) \, ds$.

Each has a point of view about the ODE, but these are all equivalent definitions involving deterministic variability $f(x, t)$. 
Deterministic Definitions

Deterministic Differential Equation:

To compute a function $x(t)$, $0 \leq t \leq T$, such that on the interval $[0, T]$, given $x(0)$ (this is an IVP, really):

- **Derivative form:** $\frac{dx}{dt} = f(x, t)$.
- **Differential form:** $dx = f(x, t)\, dt$.
- **Integral form:** $x(t) = x(0) + \int_0^t f(x(s), s)\, ds$.

Each has a point of view about the ODE, but these are all equivalent definitions involving deterministic variability $f(x, t)$. 
Deterministic Definitions

**Deterministic Differential Equation:**

To compute a function \( x(t) \), \( 0 \leq t \leq T \), such that on the interval \([0, T]\), given \( x(0) \) (this is an IVP, really):

- **Derivative form:** \( \frac{dx}{dt} = f(x, t) \).
- **Differential form:** \( dx = f(x, t) \, dt \).
- **Integral form:** \( x(t) = x(0) + \int_0^t f(x(s), s) \, ds \).

Each has a point of view about the ODE, but these are all equivalent definitions involving deterministic variability \( f(x, t) \).
Deterministic Differential Equation:

To compute a function $x(t)$, $0 \leq t \leq T$, such that on the interval $[0, T]$, given $x(0)$ (this is an IVP, really):

- **Derivative form**: $\frac{dx}{dt} = f(x, t)$.
- **Differential form**: $dx = f(x, t) \, dt$.
- **Integral form**: $x(t) = x(0) + \int_0^t f(x(s), s) \, ds$.

Each has a point of view about the ODE, but these are all equivalent definitions involving deterministic variability $f(x, t)$. 
Deterministic Definitions

Deterministic Differential Equation:

To compute a function \( x(t) \), \( 0 \leq t \leq T \), such that on the interval \([0, T]\), given \( x(0) \) (this is an IVP, really):

- Derivative form: \( \frac{dx}{dt} = f(x, t) \).
- Differential form: \( dx = f(x, t) \, dt \).
- Integral form: \( x(t) = x(0) + \int_0^t f(x(s), s) \, ds \).

Each has a point of view about the ODE, but these are all equivalent definitions involving deterministic variability \( f(x, t) \).
Stochastic Definitions

Stochastic Differential Equation:

To compute a stochastic process $X(t)$, $0 \leq t \leq T$, such that on the interval $[0, T]$, given $X(0)$ (this is an IVP, really):

- We not only want to account for deterministic variability, $f(X(t), t)$, but also stochastic variability:
- Differential form:
  \[ dX(t) = f(X(t), t)\, dt + g(X(t), t)\, dW(t). \]
- Integral form:
  \[ X(t) = X(0) + \int_0^t f(X(s), s)\, ds + \int_0^t g(X(s), s)\, dW(s). \]
- Caution: either form forces us to make a choice about which is the appropriate stochastic integral to use.
Stochastic Definitions

Stochastic Differential Equation:

To compute a stochastic process $X(t)$, $0 \leq t \leq T$, such that on the interval $[0, T]$, given $X(0)$ (this is an IVP, really):

- We not only want to account for deterministic variability, $f(X(t), t)$, but also stochastic variability:

- Differential form:

  \[ dX(t) = f(X(t), t)\, dt + g(X(t), t)\, dW(t). \]

- Integral form:

  \[ X(t) = X(0) + \int_0^t f(X(s), s)\, ds + \int_0^t g(X(s), s)\, dW(s). \]

- Caution: either form forces us to make a choice about which is the appropriate stochastic integral to use.
Stochastic Definitions

Stochastic Differential Equation:

To compute a stochastic process $X(t)$, $0 \leq t \leq T$, such that on the interval $[0, T]$, given $X(0)$ (this is an IVP, really):

- We not only want to account for deterministic variability, $f(X(t), t)$, but also stochastic variability:

- Differential form:
  
  $dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t)$.

- Integral form:

  $X(t) = X(0) + \int_0^t f(X(s), s) \, ds + \int_0^t g(X(s), s) \, dW(s)$.

- Caution: either form forces us to make a choice about which is the appropriate stochastic integral to use.
Stochastic Definitions

Stochastic Differential Equation:

To compute a stochastic process $X(t)$, $0 \leq t \leq T$, such that on the interval $[0, T]$, given $X(0)$ (this is an IVP, really):

- We not only want to account for deterministic variability, $f(X(t), t)$, but also stochastic variability:

- Differential form:
  $$dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t).$$

- Integral form:
  $$X(t) = X(0) + \int_0^t f(X(s), s) \, ds + \int_0^t g(X(s), s) \, dW(s).$$

- Caution: either form forces us to make a choice about which is the appropriate stochastic integral to use.
Stochastic Definitions

Stochastic Differential Equation:

To compute a stochastic process $X(t)$, $0 \leq t \leq T$, such that on the interval $[0, T]$, given $X(0)$ (this is an IVP, really):

- We not only want to account for deterministic variability, $f(X(t), t)$, but also stochastic variability:

- Differential form:
  
  $dX(t) = f(X(t), t)\, dt + g(X(t), t)\, dW(t)$.  

- Integral form:
  
  $X(t) = X(0) + \int_0^t f(X(s), s)\, ds + \int_0^t g(X(s), s)\, dW(s)$.  

- Caution: either form forces us to make a choice about which is the appropriate stochastic integral to use.
Example (Risky Asset Pricing, a.k.a., Geometric Brownian Motion):

An asset price $X(t)$ can be viewed as a random process. The relative change in price, $dX(t)/X(t)$ can be viewed as having two (additive) components:

- A deterministic factor: $\lambda dt$. If there were no risk, we could think of $\lambda$ as the growth rate over time. In the simplest case, $\lambda$ is constant.

- A random factor: $\mu dW(t)$, where $dW = \sqrt{dt} Z$, $Z \sim N(0,1)$ and $W(t)$ is Brownian motion. In the simplest case, $\mu$ is constant.

- So the stochastic differential equation that results is the linear differential equation

$$\frac{dX(t)}{X(t)} = \lambda dt + \mu dW(t)$$

or $dX(t) = \lambda X(t) dt + \mu X(t) dW(t)$ (multiplicative noise).

- Exact solution: $X(t) = X(0) e^{\left(\lambda - \frac{1}{2} \mu^2\right) t + \mu W(t)}$. 
Example (Risky Asset Pricing, a.k.a., Geometric Brownian Motion):

An asset price $X(t)$ can be viewed as a random process. The relative change in price, $dX(t)/X(t)$ can be viewed as having two (additive) components:

- A deterministic factor: $\lambda dt$. If there were no risk, we could think of $\lambda$ as the growth rate over time. In the simplest case, $\lambda$ is constant.

- A random factor: $\mu dW(t)$, where $dW = \sqrt{dt}Z$, $Z \sim N(0, 1)$ and $W(t)$ is Brownian motion. In the simplest case, $\mu$ is constant.

So the stochastic differential equation that results is the linear differential equation

$$\frac{dX(t)}{X(t)} = \lambda dt + \mu dW(t)$$

or $dX(t) = \lambda X(t) dt + \mu X(t) dW(t)$ (multiplicative noise).

- Exact solution: $X(t) = X(0) e^{(\lambda - \frac{1}{2} \mu^2) t + \mu W(t)}$. 

Example (Risky Asset Pricing, a.k.a., Geometric Brownian Motion):

An asset price $X(t)$ can be viewed as a random process. The relative change in price, $dX(t)/X(t)$, can be viewed as having two (additive) components:

- A deterministic factor: $\lambda dt$. If there were no risk, we could think of $\lambda$ as the growth rate over time. In the simplest case, $\lambda$ is constant.

- A random factor: $\mu dW(t)$, where $dW = \sqrt{dt}Z$, $Z \sim N(0,1)$ and $W(t)$ is Brownian motion. In the simplest case, $\mu$ is constant.

- So the stochastic differential equation that results is the linear differential equation

$$\frac{dX(t)}{X(t)} = \lambda dt + \mu dW(t)$$

or

$$dX(t) = \lambda X(t) dt + \mu X(t) dW(t)$$

(multiplicative noise).

- Exact solution: $X(t) = X(0)e^{\left(\lambda - \frac{1}{2}\mu^2\right)t + \mu W(t)}$. 
An asset price $X(t)$ can be viewed as a random process. The relative change in price, $dX(t)/X(t)$ can be viewed as having two (additive) components:

- A deterministic factor: $\lambda dt$. If there were no risk, we could think of $\lambda$ as the growth rate over time. In the simplest case, $\lambda$ is constant.

- A random factor: $\mu dW(t)$, where $dW = \sqrt{dt}Z$, $Z \sim N(0, 1)$ and $W(t)$ is Brownian motion. In the simplest case, $\mu$ is constant.

So the stochastic differential equation that results is the linear differential equation

$$
\frac{dX(t)}{X(t)} = \lambda dt + \mu dW(t)
$$

or $dX(t) = \lambda X(t) dt + \mu X(t) dW(t)$ (multiplicative noise).

Exact solution: $X(t) = X(0) e^{(\lambda - \frac{1}{2} \mu^2)t + \mu W(t)}$. 

Example (Risky Asset Pricing, a.k.a., Geometric Brownian Motion):
Example (Risky Asset Pricing, a.k.a., Geometric Brownian Motion):

An asset price $X(t)$ can be viewed as a random process. The relative change in price, $dX(t)/X(t)$ can be viewed as having two (additive) components:

- A deterministic factor: $\lambda dt$. If there were no risk, we could think of $\lambda$ as the growth rate over time. In the simplest case, $\lambda$ is constant.

- A random factor: $\mu dW(t)$, where $dW = \sqrt{dt}Z$, $Z \sim N(0, 1)$ and $W(t)$ is Brownian motion. In the simplest case, $\mu$ is constant.

So the stochastic differential equation that results is the linear differential equation

$$\frac{dX(t)}{X(t)} = \lambda dt + \mu dW(t)$$

or $dX(t) = \lambda X(t) dt + \mu X(t) dW(t)$ (multiplicative noise).

Exact solution: $X(t) = X(0) e^{(\lambda-\frac{1}{2}\mu^2)t+\mu W(t)}$. 
Deterministic Case

Numerical Solutions:

- Discretize time 0 = \( t_0 < t_1 < \cdots < t_N = T \), \( t_{j+1} - t_j = \Delta t \).
- March forward in time to compute \( x_j \approx x(t_j) \) using the identity
  \[
  x(t_{j+1}) = x(t_j) + \int_{t_j}^{t_{j+1}} f(x(s), s) \, ds.
  \]

- Explicit Euler (left Riemann sums): \( x_{j+1} = x_j + f(x_j, t_j) \Delta t \)
- Implicit Euler (right Riemann sums):
  \[
  x_{j+1} = x_j + f(x_{j+1}, t_{j+1}) \Delta t, \quad j = 0, 1, \ldots, N - 1.
  \]
- Applied to the model problem \( x(0) = 1, \ dx = \lambda x \, dt \), these
give respectively \( x_j = (1 + \lambda \Delta t)^j \) and \( x_j = 1 / (1 - \lambda \Delta t)^j \),
resp. (Exact solution: \( x(t) = e^{\lambda t} \).)
Deterministic Case

Numerical Solutions:

- Discretize time $0 = t_0 < t_1 < \cdots < t_N = T$, $t_{j+1} - t_j = \Delta t$.
- March forward in time to compute $x_j \approx x(t_j)$ using the identity
  \[ x(t_{j+1}) = x(t_j) + \int_{t_j}^{t_{j+1}} f(x(s), s) \, ds. \]

- Explicit Euler (left Riemann sums): $x_{j+1} = x_j + f(x_j, t_j) \Delta t$
- Implicit Euler (right Riemann sums):
  \[ x_{j+1} = x_j + f(x_{j+1}, t_{j+1}) \Delta t, \quad j = 0, 1, \ldots, N - 1. \]
- Applied to the model problem $x(0) = 1$, $dx = \lambda x \, dt$, these give respectively $x_j = (1 + \lambda \Delta t)^j$ and $x_j = 1 / (1 - \lambda \Delta t)^j$, resp. (Exact solution: $x(t) = e^{\lambda t}$.)
Deterministic Case

Numerical Solutions:

- Discretize time $0 = t_0 < t_1 < \cdots < t_N = T$, $t_{j+1} - t_j = \Delta t$.
- March forward in time to compute $x_j \approx x(t_j)$ using the identity

$$x(t_{j+1}) = x(t_j) + \int_{t_j}^{t_{j+1}} f(x(s), s) \, ds.$$ 

- Explicit Euler (left Riemann sums): $x_{j+1} = x_j + f(x_j, t_j) \Delta t$
- Implicit Euler (right Riemann sums):
  $$x_{j+1} = x_j + f(x_{j+1}, t_{j+1}) \Delta t, \quad j = 0, 1, \ldots, N - 1.$$ 

- Applied to the model problem $x(0) = 1$, $dx = \lambda x \, dt$, these give respectively $x_j = (1 + \lambda \Delta t)^j$ and $x_j = 1 / (1 - \lambda \Delta t)^j$, resp. (Exact solution: $x(t) = e^{\lambda t}$.)
Deterministic Case

**Numerical Solutions:**

- Discretize time $0 = t_0 < t_1 < \cdots < t_N = T$, $t_{j+1} - t_j = \Delta t$.
- March forward in time to compute $x_j \approx x(t_j)$ using the identity

$$x(t_{j+1}) = x(t_j) + \int_{t_j}^{t_{j+1}} f(x(s), s) \, ds.$$

- **Explicit Euler (left Riemann sums):** $x_{j+1} = x_j + f(x_j, t_j) \Delta t$
- **Implicit Euler (right Riemann sums):**
  $$x_{j+1} = x_j + f(x_{j+1}, t_{j+1}) \Delta t, \quad j = 0, 1, \ldots, N - 1.$$
- Applied to the model problem $x(0) = 1$, $dx = \lambda x \, dt$, these give respectively $x_j = (1 + \lambda \Delta t)^j$ and $x_j = 1 / (1 - \lambda \Delta t)^j$, resp. (Exact solution: $x(t) = e^{\lambda t}$.)
Deterministic Case

**Numerical Solutions:**

- Discretize time $0 = t_0 < t_1 < \cdots < t_N = T$, $t_{j+1} - t_j = \Delta t$.
- March forward in time to compute $x_j \approx x(t_j)$ using the identity
  \[
  x(t_{j+1}) = x(t_j) + \int_{t_j}^{t_{j+1}} f(x(s), s) \, ds.
  \]
- Explicit Euler (left Riemann sums): $x_{j+1} = x_j + f(x_j, t_j) \Delta t$
- Implicit Euler (right Riemann sums):
  \[
  x_{j+1} = x_j + f(x_{j+1}, t_{j+1}) \Delta t, \quad j = 0, 1, \ldots, N - 1.
  \]
- Applied to the model problem $x(0) = 1, \ dx = \lambda x \ dt$, these give respectively $x_j = (1 + \lambda \Delta t)^j$ and $x_j = 1 / (1 - \lambda \Delta t)^j$, resp. (Exact solution: $x(t) = e^{\lambda t}$.)
Deterministic Case

Numerical Solutions:

- Discretize time $0 = t_0 < t_1 < \cdots < t_N = T$, $t_{j+1} - t_j = \Delta t$.
- March forward in time to compute $x_j \approx x(t_j)$ using the identity

  $$x(t_{j+1}) = x(t_j) + \int_{t_j}^{t_{j+1}} f(x(s), s) \, ds.$$

- Explicit Euler (left Riemann sums): $x_{j+1} = x_j + f(x_j, t_j) \Delta t$.
- Implicit Euler (right Riemann sums):
  $$x_{j+1} = x_j + f(x_{j+1}, t_{j+1}) \Delta t, \quad j = 0, 1, \ldots, N - 1.$$
- Applied to the model problem $x(0) = 1$, $dx = \lambda x \, dt$, these give respectively $x_j = (1 + \lambda \Delta t)^j$ and $x_j = 1 / (1 - \lambda \Delta t)^j$, resp. (Exact solution: $x(t) = e^{\lambda t}$.)
Convergence and stability of the Euler methods:

- Classical analysis shows that under reasonable conditions, the methods are convergent of order one in $\Delta t$, i.e.,
  \[ \| [x_j - x(t_j)] \| = \mathcal{O}(\Delta t), \quad \delta t \to 0. \]
- The methods are stable in this sense: There is a positive $h_0$ such that for $h \in (0, h_0)$, if $|x_0 - x(0)| \leq \epsilon$, then
  \[ \| [x_j - x(t_j)] \| = \mathcal{O}(\epsilon), \quad \epsilon \to 0. \]
- NB: this version of stability only applies to finite interval problems on $[0, T]$. What about long term behavior?

- Define the linear stability domain of a method to be the subset $D = \{ z = \lambda \Delta t \mid \lim_{j \to \infty} x_j = 0 \}$ where the sequence $\{x_j\}$ is produced by applying the method to the model problem $dx/dt = \lambda x$, $x(0) = 1$.
- The method is **A-stable** if $D$ contains the open left half-plane. Reason: negative $\Re(\lambda)$ and positive $\Delta t$ are main parameters of interest for asymptotic (or absolute) stability since then the solution $x(t) = x(0) e^{\lambda t}$ to the ODE is “asymptotically stable.”
- Now examine the stability of both Euler methods.
Convergence and stability of the Euler methods:

- Classical analysis shows that under reasonable conditions, the methods are convergent of order one in $\Delta t$, i.e.,
  \[ \| [x_j - x(t_j)] \| = O(\Delta t), \delta t \to 0. \]
- The methods are stable in this sense: There is a positive $h_0$ such that for $h \in (0, h_0)$, if $|x_0 - x(0)| \leq \epsilon$, then
  \[ \| [x_j - x(t_j)] \| = O(\epsilon), \epsilon \to 0. \]
- NB: this version of stability only applies to finite interval problems on $[0, T]$. What about long term behavior?
- Define the linear stability domain of a method to be the subset $D = \{ z = \lambda \Delta t | \lim_{j \to \infty} x_j = 0 \}$ where the sequence $\{x_j\}$ is produced by applying the method to the model problem $dx/dt = \lambda x, x(0) = 1$.
- The method is A-stable if $D$ contains the open left half-plane. Reason: negative $\Re(\lambda)$ and positive $\Delta t$ are main parameters of interest for asymptotic (or absolute) stability since then the solution $x(t) = x(0) e^{\lambda t}$ to the ODE is “asymptotically stable.”
- Now examine the stability of both Euler methods.
Convergence and stability of the Euler methods:

- Classical analysis shows that under reasonable conditions, the methods are convergent of order one in $\Delta t$, i.e.,
  \[ \| x_j - x(t_j) \| = O(\Delta t), \quad \delta t \to 0. \]

- The methods are stable in this sense: There is a positive $h_0$ such that for $h \in (0, h_0)$, if $|x_0 - x(0)| \leq \epsilon$, then
  \[ \| x_j - x(t_j) \| = O(\epsilon), \quad \epsilon \to 0. \]

- NB: this version of stability only applies to finite interval problems on $[0, T]$. What about long term behavior?

- Define the linear stability domain of a method to be the subset $D = \{ z = \lambda \Delta t \mid \lim_{j \to \infty} x_j = 0 \}$ where the sequence $\{x_j\}$ is produced by applying the method to the model problem $dx/dt = \lambda x$, $x(0) = 1$.

- The method is A-stable if $D$ contains the open left half-plane. Reason: negative $\Re(\lambda)$ and positive $\Delta t$ are main parameters of interest for asymptotic (or absolute) stability since then the solution $x(t) = x(0)e^{\lambda t}$ to the ODE is “asymptotically stable.”

- Now examine the stability of both Euler methods.
Convergence and stability of the Euler methods:

- Classical analysis shows that under reasonable conditions, the methods are convergent of order one in $\Delta t$, i.e.,
  \[ \| [x_j - x(t_j)] \| = O(\Delta t), \delta t \to 0. \]

- The methods are stable in this sense: There is a positive $h_0$ such that for $h \in (0, h_0)$, if $|x_0 - x(0)| \leq \epsilon$, then
  \[ \| [x_j - x(t_j)] \| = O(\epsilon), \epsilon \to 0. \]

- NB: this version of stability only applies to finite interval problems on $[0, T]$. What about long term behavior?

- Define the linear stability domain of a method to be the subset
  \[ D = \{ z = \lambda \Delta t \mid \lim_{j \to \infty} x_j = 0 \} \]
  where the sequence $\{x_j\}$ is produced by applying the method to the model problem
  \[ \frac{dx}{dt} = \lambda x, \quad x(0) = 1. \]

- The method is A-stable if $D$ contains the open left half-plane. Reason: negative $\Re(\lambda)$ and positive $\Delta t$ are main parameters of interest for asymptotic (or absolute) stability since then the solution $x(t) = x(0)e^{\lambda t}$ to the ODE is “asymptotically stable.”

- Now examine the stability of both Euler methods.
Convergence and stability of the Euler methods:

- Classical analysis shows that under reasonable conditions, the methods are convergent of order one in $\Delta t$, i.e.,
  $$\| [x_j - x(t_j)] \| = O(\Delta t), \, \delta t \to 0.$$ 

- The methods are stable in this sense: There is a positive $h_0$ such that for $h \in (0, h_0)$, if $|x_0 - x(0)| \leq \epsilon$, then
  $$\| [x_j - x(t_j)] \| = O(\epsilon), \, \epsilon \to 0.$$ 

- NB: this version of stability only applies to finite interval problems on $[0, T]$. What about long term behavior?

- Define the linear stability domain of a method to be the subset
  $$D = \{ z = \lambda \Delta t \mid \lim_{j \to \infty} x_j = 0 \}$$
  where the sequence $\{x_j\}$ is produced by applying the method to the model problem
  $$dx/dt = \lambda x, \, x(0) = 1.$$ 

- The method is A-stable if $D$ contains the open left half-plane.
  Reason: negative $\Re(\lambda)$ and positive $\Delta t$ are main parameters of interest for asymptotic (or absolute) stability since then the solution $x(t) = x(0)e^{\lambda t}$ to the ODE is “asymptotically stable.”

- Now examine the stability of both Euler methods.
Convergence and stability of the Euler methods:

- Classical analysis shows that under reasonable conditions, the methods are convergent of order one in $\Delta t$, i.e.,
  \[ \| [x_j - x(t_j)] \| = O(\Delta t), \delta t \to 0. \]
- The methods are stable in this sense: There is a positive $h_0$ such that for $h \in (0, h_0)$, if $|x_0 - x(0)| \leq \epsilon$, then
  \[ \| [x_j - x(t_j)] \| = O(\epsilon), \epsilon \to 0. \]
- NB: this version of stability only applies to finite interval problems on $[0, T]$. What about long term behavior?
- Define the **linear stability domain** of a method to be the subset $D = \{ z = \lambda \Delta t \mid \lim_{j \to \infty} x_j = 0 \}$ where the sequence $\{x_j\}$ is produced by applying the method to the model problem
  \[ \frac{dx}{dt} = \lambda x, \quad x(0) = 1. \]
- The method is **A-stable** if $D$ contains the open left half-plane. Reason: negative $\Re(\lambda)$ and positive $\Delta t$ are main parameters of interest for asymptotic (or absolute) stability since then the solution $x(t) = x(0)e^{\lambda t}$ to the ODE is “asymptotically stable.”

Now examine the stability of both Euler methods.
Convergence and stability of the Euler methods:

- Classical analysis shows that under reasonable conditions, the methods are convergent of order one in $\Delta t$, i.e.,
  \[ \| [x_j - x(t_j)] \| = O(\Delta t), \quad \delta t \to 0. \]

- The methods are stable in this sense: There is a positive $h_0$ such that for $h \in (0, h_0)$, if $|x_0 - x(0)| \leq \epsilon$, then
  \[ \| [x_j - x(t_j)] \| = O(\epsilon), \quad \epsilon \to 0. \]

- NB: this version of stability only applies to finite interval problems on $[0, T]$. What about long term behavior?

Define the linear stability domain of a method to be the subset $D = \{ z = \lambda \Delta t \mid \lim_{j \to \infty} x_j = 0 \}$ where the sequence $\{x_j\}$ is produced by applying the method to the model problem $dx/dt = \lambda x$, $x(0) = 1$.

- The method is A-stable if $D$ contains the open left half-plane. Reason: negative $\Re(\lambda)$ and positive $\Delta t$ are main parameters of interest for asymptotic (or absolute) stability since then the solution $x(t) = x(0)e^{\lambda t}$ to the ODE is “asymptotically stable.”

- Now examine the stability of both Euler methods.
Numerical Solutions to
\[ dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t) \]:

- Discretize time \(0 = \tau_0 < \tau_1 < \cdots < \tau_N = T, \tau_{j+1} - \tau_j = \Delta t\).
- March forward in time to compute \(X_j \approx X(\tau_j)\) using the identity

\[
X(\tau_{j+1}) = X(\tau_j) + \int_{\tau_j}^{\tau_{j+1}} f(X(s), s) \, ds \\
= + \int_{\tau_j}^{\tau_{j+1}} g(X(s), s) \, dW(s).
\]

- Euler-Maruyama (EM) method:
  \[X_{j+1} = X_j + f(X_j, \tau_j) \Delta t + g(X_j, \tau_j)(W(\tau_{j+1}) - W(\tau_j))\]
- Convergence and stability need re-interpretation here.
### Stochastic Case: Euler-Maruyama Method

#### Numerical Solutions to
\[ dX(t) = f(X(t), t)\, dt + g(X(t), t)\, dW(t): \]

- Discretize time \(0 = \tau_0 < \tau_1 < \cdots < \tau_N = T, \tau_{j+1} - \tau_j = \Delta t.\)
- March forward in time to compute \(X_j \approx X(\tau_j)\) using the identity

\[
X(\tau_{j+1}) = X(\tau_j) + \int_{\tau_j}^{\tau_{j+1}} f(X(s), s)\, ds
= + \int_{\tau_j}^{\tau_{j+1}} g(X(s), s)\, dW(s).
\]

- **Euler-Maruyama (EM) method:**
  \[X_{j+1} = X_j + f(X_j, \tau_j)\Delta t + g(X_j, \tau_j)(W(\tau_{j+1}) - W(\tau_j))\]
- Convergence and stability need re-interpretation here.
Numerical Solutions to
\(dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t)\):

- Discretize time 0 = \(\tau_0 < \tau_1 < \cdots < \tau_N = T\), \(\tau_{j+1} - \tau_j = \Delta t\).
- March forward in time to compute \(X_j \approx X(\tau_j)\) using the identity

\[
X(\tau_{j+1}) = X(\tau_j) + \int_{\tau_j}^{\tau_{j+1}} f(X(s), s) \, ds
\]

\[
= + \int_{\tau_j}^{\tau_{j+1}} g(X(s), s) \, dW(s).
\]

- Euler-Maruyama (EM) method:
  \(X_{j+1} = X_j + f(X_j, \tau_j) \Delta t + g(X_j, \tau_j)(W(\tau_{j+1}) - W(\tau_j))\)

- Convergence and stability need re-interpretation here.
Stochastic Case: Euler-Maruyama Method

Numerical Solutions to
\[ dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t) \]:

- Discretize time \(0 = \tau_0 < \tau_1 < \cdots < \tau_N = T, \tau_{j+1} - \tau_j = \Delta t\).
- March forward in time to compute \(X_j \approx X(\tau_j)\) using the identity
  \[
  X(\tau_{j+1}) = X(\tau_j) + \int_{\tau_j}^{\tau_{j+1}} f(X(s), s) \, ds \tag{2}
  
  = + \int_{\tau_j}^{\tau_{j+1}} g(X(s), s) \, dW(s). \tag{3}
  \]

- Euler-Maruyama (EM) method:
  \[X_{j+1} = X_j + f(X_j, \tau_j) \Delta t + g(X_j, \tau_j)(W(\tau_{j+1}) - W(\tau_j))\]
- Convergence and stability need re-interpretation here.
Stochastic Case: Euler-Maruyama Method

Numerical Solutions to
\[ dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t) \]:

- Discretize time \(0 = \tau_0 < \tau_1 < \cdots < \tau_N = T\), \(\tau_{j+1} - \tau_j = \Delta t\).
- March forward in time to compute \(X_j \approx X(\tau_j)\) using the identity

\[
X(\tau_{j+1}) = X(\tau_j) + \int_{\tau_j}^{\tau_{j+1}} f(X(s), s) \, ds \\
= \int_{\tau_j}^{\tau_{j+1}} g(X(s), s) \, dW(s).
\]

- Euler-Maruyama (EM) method:
  \[X_{j+1} = X_j + f(X_j, \tau_j) \Delta t + g(X_j, \tau_j)(W(\tau_{j+1}) - W(\tau_j))\]
- Convergence and stability need re-interpretation here.
Implementation Convention: A discrete Brownian path is generated using $dt$. Then the Euler-Maruyama time step is a multiple of $dt$, say $R \cdot dt = \Delta t$.

%EM Euler-Maruyama method on linear SDE
%
% SDE is $dX = \lambda X \, dt + \mu X \, dW$, $X(0) = X_{\text{zero}}$,
% where $\lambda = 2$, $\mu = 1$ and $X_{\text{zero}} = 1$.
%
% Discretized Brownian path over $[0,1]$ has $dt = 2^{-8}$.
% Euler-Maruyama uses timestep $R \cdot dt$.
randn('state',100)
lambda = 2; mu = 1; Xzero = 1; % problem parameters
T = 1; N = 2^8; dt = T/N;
dW = sqrt(dt)*randn(1,N); % Brownian increments
W = cumsum(dW); % discretized Brownian path
\[
X_{true} = X_{zero} \exp((\lambda - 0.5 \mu^2) * ([dt:dt:T]) + \mu W);
\]

```matlab
plot([0:dt:T],[Xzero,Xtrue],'m-'), hold on
R = 4; Dt = R*dt; L = N/R; % L EM steps of size Dt = R*dt
Xem = zeros(1,L); % preallocate for efficiency
Xtemp = Xzero;
for j = 1:L
    Winc = sum(dW(R*(j-1)+1:R*j));
    Xtemp = Xtemp + Dt*lambda*Xtemp + mu*Xtemp*Winc;
    Xem(j) = Xtemp;
end
plot([0:Dt:T],[Xzero,Xem],'r--*'), hold off
xlabel('t','FontSize',12)
ylabel('X','FontSize',16,'Rotation',0,'HorizontalAlignment','right')
emerr = abs(Xem(end)-Xtrue(end))
```
Strong Convergence

Numerical Method for $dX = f(X, t)\, dt + g(X, t)\, dW$ on $[0, T]$: 

- Converges strongly if mean of the error converges to zero, i.e., 
  $$\lim_{n\to\infty} E[|X_n - X(\tau)|] = 0,$$

- and with order of convergence $\gamma$ if there exists $C > 0$ such that for any fixed $\tau = n \Delta t \in [0, T]$,
  $$E[|X_n - X(\tau)|] \leq C\Delta t^\gamma$$

for all $\Delta t$ sufficiently small. Put another way, the expected value of the error is $O(\Delta t)$, $\Delta t \to 0$.

- Uniform order convergence does follow for EM, but this isn’t obvious, nor is it the form of the definition of strong convergence in Kloeden-Platen, as is apparently the case here.
Strong Convergence

Numerical Method for $dX = f(X, t)\, dt + g(X, t)\, dW$ on $[0, T]$:

- Converges strongly if mean of the error converges to zero, i.e.,

$$\lim_{n \to \infty} E[|X_n - X(\tau)|] = 0,$$

- and with order of convergence $\gamma$ if there exists $C > 0$ such that for any fixed $\tau = n \Delta t \in [0, T]$,

$$E[|X_n - X(\tau)|] \leq C \Delta t^\gamma$$

for all $\Delta t$ sufficiently small. Put another way, the expected value of the error is $O(\Delta t)$, $\Delta t \to 0$.

- Uniform order convergence does follow for EM, but this isn’t obvious, nor is it the form of the definition of strong convergence in Kloeden-Platen, as is apparently the case here.
Numerical Method for $dX = f(X, t)\, dt + g(X, t)\, dW$ on $[0, T]$:

- Converges **strongly** if mean of the error converges to zero, i.e.,
  
  $$\lim_{n \to \infty} E[|X_n - X(\tau)|] = 0,$$

- and with **order of convergence** $\gamma$ if there exists $C > 0$ such that for any fixed $\tau = n \Delta t \in [0, T]$,
  
  $$E[|X_n - X(\tau)|] \leq C\Delta t^\gamma$$

for all $\Delta t$ sufficiently small. Put another way, the expected value of the error is $O(\Delta t)$, $\Delta t \to 0$.

- Uniform order convergence does follow for EM, but this isn’t obvious, nor is it the form of the definition of strong convergence in Kloeden-Platen, as is apparently the case here.
Strong Convergence

Numerical Method for $dX = f(X, t)\, dt + g(X, t)\, dW$ on $[0, T]$

- Converges **strongly** if mean of the error converges to zero, i.e.,
  $$\lim_{n \to \infty} E[|X_n - X(\tau)|] = 0,$$

- and with **order of convergence** $\gamma$ if there exists $C > 0$ such that for any fixed $\tau = n \Delta t \in [0, T]$,
  $$E[|X_n - X(\tau)|] \leq C\Delta t^{\gamma}$$

for all $\Delta t$ sufficiently small. Put another way, the expected value of the error is $O(\Delta t)$, $\Delta t \to 0$.

- Uniform order convergence does follow for EM, but this isn’t obvious, nor is it the form of the definition of strong convergence in Kloeden-Platen, as is apparently the case here.
Idea Behind the Experiment:

- If you think that there is a valid order condition
  \[ E_{\Delta t} \leq C \Delta t^\gamma, \]
- assume that the inequality is sharp and replace it by
  \[ E_{\Delta t} \approx C \Delta t^\gamma. \]
- Take logs of both sides and get
  \[ Y_{\Delta t} = \log E_{\Delta t} \approx \log C + \gamma \log \Delta t. \]
- Do a log-log plot of \( E_{\Delta t} \) against \( \Delta t \).
- A graph that resembles a straight line of slope \( \gamma \) and intercept \( \log C \) supports your suspicion.
Idea Behind the Experiment:

- If you think that there is a valid order condition

\[ E_{\Delta t} \leq C \Delta t^\gamma, \]

- Assume that the inequality is sharp and replace it by

\[ E_{\Delta t} \approx C \Delta t^\gamma. \]

- Take logs of both sides and get

\[ Y_{\Delta t} = \log E_{\Delta t} \approx \log C + \gamma \log \Delta t. \]

- Do a log-log plot of \( E_{\Delta t} \) against \( \Delta t \).
- A graph that resembles a straight line of slope \( \gamma \) and intercept \( \log C \) supports your suspicion.
Idea Behind the Experiment:

- If you think that there is a valid order condition
  \[ E_{\Delta t} \leq C \Delta t^\gamma, \]
  assume that the inequality is sharp and replace it by
  \[ E_{\Delta t} \approx C \Delta t^\gamma. \]

- Take logs of both sides and get
  \[ Y_{\Delta t} = \log E_{\Delta t} \approx \log C + \gamma \log \Delta t. \]

- Do a log-log plot of \( E_{\Delta t} \) against \( \Delta t \).
- A graph that resembles a straight line of slope \( \gamma \) and intercept \( \log C \) supports your suspicion.
An Experiment

Idea Behind the Experiment:

- If you think that there is a valid order condition

\[ E_{\Delta t} \leq C \Delta t^\gamma, \]

- assume that the inequality is sharp and replace it by

\[ E_{\Delta t} \approx C \Delta t^\gamma. \]

- Take logs of both sides and get

\[ Y_{\Delta t} = \log E_{\Delta t} \approx \log C + \gamma \log \Delta t. \]

- Do a log-log plot of \( E_{\Delta t} \) against \( \Delta t \).
- A graph that resembles a straight line of slope \( \gamma \) and intercept \( \log C \) supports your suspicion.
Idea Behind the Experiment:

- If you think that there is a valid order condition
  \[ E_{\Delta t} \leq C \Delta t^\gamma, \]

- assume that the inequality is sharp and replace it by
  \[ E_{\Delta t} \approx C \Delta t^\gamma. \]

- Take logs of both sides and get
  \[ Y_{\Delta t} = \log E_{\Delta t} \approx \log C + \gamma \log \Delta t. \]

- Do a log-log plot of \( E_{\Delta t} \) against \( \Delta t \).

- A graph that resembles a straight line of slope \( \gamma \) and intercept \( \log C \) supports your suspicion.
An Experiment

Idea Behind the Experiment:

- If you think that there is a valid order condition
  \[ E_{\Delta t} \leq C \Delta t^\gamma, \]

- assume that the inequality is sharp and replace it by
  \[ E_{\Delta t} \approx C \Delta t^\gamma. \]

- Take logs of both sides and get
  \[ Y_{\Delta t} = \log E_{\Delta t} \approx \log C + \gamma \log \Delta t. \]

- Do a log-log plot of \( E_{\Delta t} \) against \( \Delta t \).
- A graph that resembles a straight line of slope \( \gamma \) and intercept \( \log C \) supports your suspicion.
An Experiment Continued

Compute geometric Brownian motion by taking the mean of 1000 different Brownian paths on $[0, 1]$ at $T = \tau = 1$. Use $\delta t = 2^{-9}$ and $\Delta t = 2^{p-1}\delta t$, $1 \leq p \leq 5$. Then do a log-log plot, linear regression to estimate $\gamma$ (q in the program), and the norm of the residual:

% EMSTRONG Test strong convergence of Euler-Maruyama
% Solves $dX = \lambda X \, dt + \mu X \, dW$, $X(0) = X_0$,
% where $\lambda = 2$, $\mu = 1$ and $X_0 = 1$.
% Discretized Brownian path over $[0,1]$ has $dt = 2^{-9}$.
% E-M uses 5 different timesteps: 16dt, 8dt, 4dt, 2dt, dt.
% Examine strong convergence at $T=1$: $E | X_L - X(T) |$.

randn('state',100)
lambda = 2; mu = 1; Xzero = 1; % problem parameters
T = 1; N = 2^9; dt = T/N; %
M = 1000; % number of paths sampled
Xerr = zeros(M,5); % preallocate array
for s = 1:M, % sample over discrete Brownian paths
dW = sqrt(dt)*randn(1,N); % Brownian increments
W = cumsum(dW); % discrete Brownian path
Xtrue = Xzero*exp((lambda-0.5*mu^2)+mu*W(end));
for p = 1:5
R = 2^(p-1); Dt = R*dt; L = N/R; % L Euler steps of size Dt = R*dt
Xtemp = Xzero;
for j = 1:L
Winc = sum(dW(R*(j-1)+1:R*j));
Xtemp = Xtemp + Dt*lambda*Xtemp + mu*Xtemp*Winc;
end
Xerr(s,p) = abs(Xtemp - Xtrue); % store the error at t = 1
end
end
Dtvals = dt*(2.^(0:4));
subplot(221) % top LH picture
loglog(Dtvals,mean(Xerr),'b*-'), hold on
loglog(Dtvals,(Dtvals.^(.5)),'r--'), hold off % reference slope of 1/2
axis([1e-3 1e-1 1e-4 1])
xlabel('\Delta t'), ylabel('Sample average of | X(T) - X_L |')
title('emstrong.m','FontSize',10)
%%%% Least squares fit of error = C * Dt^q %%%%
A = [ones(5,1), log(Dtvals)']; rhs = log(mean(Xerr)');
sol = A\rhs; q = sol(2)
resid = norm(A*sol - rhs)
Weak Convergence

Numerical Method for
\[ dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t) \] on \([0, T]::

- Converges **weakly** if mean of functions of the error taken from some set of test functions (like polynomials, which would give moments) converges to zero, i.e.,
  \[ \lim_{n \to \infty} |E[p(X_n)] - E[p(X(\tau))]| = 0, \]

  and with **order of convergence** \(\gamma\) if there exists \(C > 0\) such that for any fixed \(\tau = n \Delta t \in [0, T],\)
  \[ |E[p(X_n)] - E[p(X(\tau))]| \leq C\Delta t^\gamma \]

  for all \(\Delta t\) sufficiently small.
Weak Convergence

Numerical Method for  
\[ dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t) \] on \([0, T]\):

- Converges \textbf{weakly} if mean of functions of the error taken from some set of test functions (like polynomials, which would give moments) converges to zero, i.e.,

  \[
  \lim_{n \to \infty} \left| E[p(X_n)] - E[p(X(\tau))] \right| = 0,
  \]

- and with order of convergence \(\gamma\) if there exists \(C > 0\) such that for any fixed \(\tau = n \Delta t \in [0, T]\),

  \[
  \left| E[p(X_n)] - E[p(X(\tau))] \right| \leq C \Delta t^\gamma
  \]

for all \(\Delta t\) sufficiently small.
Weak Convergence

Numerical Method for
\[ dX(t) = f(X(t), t)\, dt + g(X(t), t)\, dW(t) \] on \([0, T]\):

- Converges **weakly** if mean of functions of the error taken from some set of test functions (like polynomials, which would give moments) converges to zero, i.e.,

\[
\lim_{n \to \infty} |E[p(X_n)] - E[p(X(\tau))]| = 0,
\]

- and with **order of convergence** \(\gamma\) if there exists \(C > 0\) such that for any fixed \(\tau = n \Delta t \in [0, T]\),

\[
|E[p(X_n)] - E[p(X(\tau))]| \leq C\Delta t^{\gamma}
\]

for all \(\Delta t\) sufficiently small.
Note: We have assumed that errors other than sampling error like floating point error and sampling bias are negligible compared to sampling error. This is reasonable in relatively small experiments.

% EMWEAK Test weak convergence of Euler-Maruyama
% Solves dX = lambda*X dt + mu*X dW, X(0) = Xzero,
% where lambda = 2, mu = 1 and Xzero = 1.
% E-M uses 5 different timesteps: 2^(p-10), p = 1,2,3,4,5.
% Examine weak convergence at T=1: | E (X_L) - E (X(T)) |.
% Different paths are used for each E-M timestep.
% Code is vectorized over paths.
% Uncommenting the line indicated below gives the weak E-M method.

randn('state',100);
lambda = 2; mu = 0.1; Xzero = 1; T = 1; % problem parameters
M = 50000; % number of paths sampled
Xem = zeros(5,1); % preallocate arrays
for p = 1:5 % take various Euler timesteps
  Dt = 2^(p-10); L = T/Dt; % L Euler steps of size Dt
  Xtemp = Xzero*ones(M,1);
for j = 1:L
    Winc = sqrt(Dt)*randn(M,1);
    Xtemp = Xtemp + Dt*lambda*Xtemp + mu*Xtemp.*Winc;
end
Xem(p) = mean(Xtemp);
end
Xerr = abs(Xem - exp(lambda));
Dtvals = 2.^([1:5]-10);
subplot(222) % top RH picture
loglog(Dtvals,Xerr,'b*--'), hold on
loglog(Dtvals,Dtvals,'r--'), hold off % reference slope of 1
axis([1e-3 1e-1 1e-4 1])
xlabel('\Delta t'), ylabel('\mid E(X(T)) - Sample average of X_L \mid')
title('emweak.m','FontSize',10)
%%%% Least squares fit of error = C * dt^q %%%%
A = [ones(p,1), log(Dtvals)']; rhs = log(Xerr);
sol = A\rhs; q = sol(2)
resid = norm(A*sol - rhs)
A careful study of Ito-Taylor expansions leads to a higher order method (Milstein’s method):

\[ X_{j+1} = X_j + f(X_j, \tau_j) \Delta t + g(X_j, \tau_j)(W(\tau_{j+1}) - W(\tau_j)) + \frac{1}{2}g(X_j)g_x(X_j, \tau_j)\left((W(\tau_{j+1}) - W(\tau_j))^2 - \Delta t\right) \]
An Experiment

Now run the experiment milstrong.m to solve the population dynamics stochastic differential equation (the stochastic Verhulst equation)

\[ dX(t) = rX(t)(K - X(t)) \, dt + \beta X(t) \, dW(t) \]

which is simply a stochastic logistic equation.

One interesting aspect of the program: the exact (strong) solution is well known, but involves another stochastic integral. Hence, the most accurate solution (smallest \( \Delta t \)) is used as a “reference” solution.
The Deterministic Case

Long Term Stability of the Euler Methods:

- Does not mean stability on finite intervals, which would require that perturbations in initial conditions cause perturbations in the computed solution that remain bounded as $\delta t \to 0$.

- Define the **linear stability domain** of a method to be the subset $D = \{ z = \lambda \Delta t \mid \lim_{j \to \infty} x_j = 0 \}$ where the sequence $\{x_j\}$ is produced by applying the method to the model problem $dx/dt = \lambda x$, $x(0) = 1$.

- The method is **A-stable** if $D$ contains the open left half-plane. Reason: negative $\Re(\lambda)$ and positive $\Delta t$ are main parameters of interest for this asymptotic (or absolute) stability.
The Deterministic Case

Long Term Stability of the Euler Methods:

- Does *not* mean stability on finite intervals, which would require that perturbations in initial conditions cause perturbations in the computed solution that remain bounded as $\delta t \to 0$.

- Define the linear stability domain of a method to be the subset $D = \{ z = \lambda \Delta t \mid \lim_{j \to \infty} x_j = 0 \}$ where the sequence $\{x_j\}$ is produced by applying the method to the model problem $dx/dt = \lambda x$, $x(0) = 1$.

- The method is **A-stable** if $D$ contains the open left half-plane. Reason: negative $\Re(\lambda)$ and positive $\Delta t$ are main parameters of interest for this asymptotic (or absolute) stability.
The Deterministic Case

Long Term Stability of the Euler Methods:

- Does not mean stability on finite intervals, which would require that perturbations in initial conditions cause perturbations in the computed solution that remain bounded as $\delta t \to 0$.
- Define the **linear stability domain** of a method to be the subset $D = \{ z = \lambda \Delta t \mid \lim_{j \to \infty} x_j = 0 \}$ where the sequence $\{x_j\}$ is produced by applying the method to the model problem $dx/dt = \lambda x$, $x(0) = 1$.
- The method is **A-stable** if $D$ contains the open left half-plane. Reason: negative $\Re(\lambda)$ and positive $\Delta t$ are main parameters of interest for this asymptotic (or absolute) stability.
The Deterministic Case

Long Term Stability of the Euler Methods:

- Does \textit{not} mean stability on finite intervals, which would require that perturbations in initial conditions cause perturbations in the computed solution that remain bounded as $\delta t \to 0$.

- Define the \textbf{linear stability domain} of a method to be the subset $D = \{ z = \lambda \Delta t \mid \lim_{j \to \infty} x_j = 0 \}$ where the sequence $\{x_j\}$ is produced by applying the method to the model problem $dx/dt = \lambda x$, $x(0) = 1$.

- The method is \textbf{A-stable} if $D$ contains the open left half-plane. Reason: negative $\Re(\lambda)$ and positive $\Delta t$ are main parameters of interest for this asymptotic (or absolute) stability.
The Stochastic Case

Long Term Stability in Stochastic Setting:

- The model problem is
  \[ dX(t) = \lambda X(t) \, dt + \mu X(t) \, dW(t). \]

Solution:
  \[ X(t) = X(0) e^{(\lambda - \frac{1}{2} \mu^2) t + \mu W(t)}. \]

- The mathematical stability of a solution comes in two flavors, assuming that \( X(0) \neq 0 \) with probability 1.

- Mean-square stability:
  \[ \lim_{t \to \infty} E\left[ X(t)^2 \right] = 0 \iff \Re(\lambda) + \frac{1}{2} |\mu|^2 < 0. \]

- Stochastic asymptotic stability:
  \[ \lim_{t \to \infty} \left| X(t)^2 \right| = 0, \text{ with probability } 1 \iff \Re\left( \lambda - \frac{1}{2} \mu^2 \right) < 0. \]
The Stochastic Case

Long Term Stability in Stochastic Setting:

- The model problem is

\[ dX(t) = \lambda X(t) \, dt + \mu X(t) \, dW(t). \]

Solution:

\[ X(t) = X(0) e^{(\lambda - \frac{1}{2} \mu^2) t} + \mu W(t) \]

- The mathematical stability of a solution comes in two flavors, assuming that \( X(0) \neq 0 \) with probability 1.

- Mean-square stability:

\[ \lim_{t \to \infty} E \left[ X(t)^2 \right] = 0 \iff \Re(\lambda) + \frac{1}{2} |\mu|^2 < 0. \]

- Stochastic asymptotic stability:

\[ \lim_{t \to \infty} \left| X(t)^2 \right| = 0, \text{ with probability 1} \iff \Re \left( \lambda - \frac{1}{2} \mu^2 \right) < 0. \]
The Stochastic Case

Long Term Stability in Stochastic Setting:

- The model problem is

\[ dX(t) = \lambda X(t) \, dt + \mu X(t) \, dW(t). \]

Solution:

\[ X(t) = X(0) e^{(\lambda - \frac{1}{2} \mu^2) t + \mu W(t)}. \]

- The mathematical stability of a solution comes in two flavors, assuming that \( X(0) \neq 0 \) with probability 1.

  - Mean-square stability:

  \[ \lim_{t \to \infty} E \left[ X(t)^2 \right] = 0 \iff \Re(\lambda) + \frac{1}{2} |\mu|^2 < 0. \]

  - Stochastic asymptotic stability:

  \[ \lim_{t \to \infty} \left| X(t)^2 \right| = 0, \text{ with probability } 1 \iff \Re \left( \lambda - \frac{1}{2} \mu^2 \right) < 0. \]
The Stochastic Case

Long Term Stability in Stochastic Setting:

- The model problem is

\[ dX(t) = \lambda X(t) \, dt + \mu X(t) \, dW(t). \]

Solution:

\[ X(t) = X(0) e^{(\lambda - \frac{1}{2} \mu^2) t + \mu W(t)} \]

- The mathematical stability of a solution comes in two flavors, assuming that \( X(0) \neq 0 \) with probability 1.

- Mean-square stability:

\[
\lim_{t \to \infty} E\left[ X(t)^2 \right] = 0 \iff \Re(\lambda) + \frac{1}{2} |\mu|^2 < 0.
\]

- Stochastic asymptotic stability:

\[
\lim_{t \to \infty} \left| X(t) \right|^2 = 0, \text{ with probability 1} \iff \Re \left( \lambda - \frac{1}{2} \mu^2 \right) < 0.
\]
The Stochastic Case

Long Term Stability in Stochastic Setting:

- The model problem is

\[ dX(t) = \lambda X(t) \, dt + \mu X(t) \, dW(t). \]

Solution:

\[ X(t) = X(0) e^{(\lambda - \frac{1}{2} \mu^2)t + \mu W(t)} \]

- The mathematical stability of a solution comes in two flavors, assuming that \( X(0) \neq 0 \) with probability 1.

  - Mean-square stability:

    \[ \lim_{t \to \infty} E\left[ X(t)^2 \right] = 0 \iff \Re(\lambda) + \frac{1}{2} |\mu|^2 < 0. \]

  - Stochastic asymptotic stability:

    \[ \lim_{t \to \infty} \left| X(t)^2 \right| = 0, \text{ with probability 1} \iff \Re \left( \lambda - \frac{1}{2} \mu^2 \right) < 0. \]
The Stochastic Case

Long Term Stability in Stochastic Setting:

- The model problem is
  \[ dX(t) = \lambda X(t) \, dt + \mu X(t) \, dW(t). \]

  Solution:
  \[ X(t) = X(0) e^{(\lambda - \frac{1}{2} \mu^2) t + \mu W(t)} \]

- The mathematical stability of a solution comes in two flavors, assuming that \( X(0) \neq 0 \) with probability 1.

- Mean-square stability:
  \[ \lim_{t \to \infty} E\left[ X(t)^2 \right] = 0 \iff \Re(\lambda) + \frac{1}{2} |\mu|^2 < 0. \]

- Stochastic asymptotic stability:
  \[ \lim_{t \to \infty} \left| X(t)^2 \right| = 0, \text{ with probability 1} \iff \Re \left( \lambda - \frac{1}{2} \mu^2 \right) < 0. \]
The Numerical Stochastic Case

Long Term Stability of Numerical Method:

One can show:

- Mean-square stability of a numerical method:

\[
\lim_{j \to \infty} \mathbb{E} \left[ X_j^2 \right] = 0 \iff |1 + \Delta t \lambda|^2 + \frac{1}{2} \Delta t |\mu|^2 < 0.
\]

- Stochastic asymptotic stability of a numerical method:

\[
\lim_{j \to \infty} |X_j^2| = 0, \text{ with probability } 1,
\]

\[
\iff \mathbb{E} \left[ \log \left| 1 + \Delta t \lambda + \sqrt{\Delta t} \mu N(0, 1) \right| \right] < 0.
\]
The Numerical Stochastic Case

Long Term Stability of Numerical Method:

One can show:

- Mean-square stability of a numerical method:

\[
\lim_{j \to \infty} E \left[ X_j^2 \right] = 0 \iff |1 + \Delta t \lambda|^2 + \frac{1}{2} \Delta t |\mu|^2 < 0.
\]

- Stochastic asymptotic stability of a numerical method:

\[
\lim_{j \to \infty} |X_j^2| = 0, \text{ with probability } 1,
\]

\[
\iff E \left[ \log \left| 1 + \Delta t \lambda + \sqrt{\Delta t} \mu \mathcal{N}(0,1) \right| \right] < 0.
\]
Long Term Stability of Numerical Method:

One can show:

- Mean-square stability of a numerical method:

\[ \lim_{j \to \infty} E \left[ X_j^2 \right] = 0 \iff |1 + \Delta t \lambda|^2 + \frac{1}{2} \Delta t |\mu|^2 < 0. \]

- Stochastic asymptotic stability of a numerical method:

\[ \lim_{j \to \infty} |X_j^2| = 0, \text{ with probability 1,} \]

\[ \iff E \left[ \log |1 + \Delta t \lambda + \sqrt{\Delta t} \mu N(0, 1)| \right] < 0. \]
Run the script stab.m. Settings are $\Delta t = 1, 1/2, 1/4, \lambda = 1/2,$ and $\mu = \sqrt{6}$. For asymptotic stability, run over a single path, while for mean-square stability, an average of paths. Note, ideally in mean-square case we should have straight line graphs, since we calculate logy graphs.
Deterministic Case

Let’s start with the deterministic chain rule: given a function $F(x, t)$, the first order differential is given by

$$df = \frac{\partial F(x, t)}{\partial x} dx + \frac{\partial F(x, t)}{\partial t} dt,$$

which gives first order (linear) approximations by the Taylor formula. Of course, if $x = x(t)$, we simply plug that into the formula for the one variable differential. We might reason accordingly that if $X = X(t)$, is a stochastic process, then we should be able to plug $X$ into $x$ and get the correct differential. Wrong! Well, at least if you use Ito integrals. (With Stratonovich integrals you would be right.)
Stochastic Chain Rule

For a function \( F(X, t) \) of a stochastic process \( X(t) \):

- Start over with a Taylor expansion

\[
\begin{align*}
\frac{dF}{dt} & = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} dx^2 \\
& \quad + \frac{\partial^2 F}{\partial x \partial t} dx dt + \frac{1}{2} \frac{\partial^2 F}{\partial t^2} dt^2.
\end{align*}
\]

- Now make the substitutions \( x = X(t) \) and \( dx = dX(t) = f(X(t), t) dt + g(X(t), t) dW(t) \).

- For a first order (linear) approximation, we have no problem in discarding the higher order \( dt^2 \) term.

- Nor does the mixed term present a problem:

\[
dX dt = (f(X(t), t) dt + g(X(t), t) dW(t)) dt.
\]

But \( dW(t) \sim \sqrt{dt} N(0, 1) \), so \( dX \) is of order \( dt^{3/2} \)
For a function $F(X, t)$ of a stochastic process $X(t)$:

- Start over with a Taylor expansion

$$dF = \frac{\partial F(x, t)}{\partial x} dx + \frac{\partial F(x, t)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F(x, t)}{\partial x^2} dx^2$$

$$+ \frac{\partial^2 F(x, t)}{\partial x \partial t} dx \, dt + \frac{1}{2} \frac{\partial^2 F(x, t)}{\partial t^2} dt^2.$$

- Now make the substitutions $x = X(t)$ and $dx = dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t)$.

- For a first order (linear) approximation, we have no problem in discarding the higher order $dt^2$ term.

- Nor does the mixed term present a problem:

$$dX \, dt = (f(X(t), t) \, dt + g(X(t), t) \, dW(t)) \, dt.$$

But $dW(t) \sim \sqrt{dt} N(0, 1)$, so $dX \, dt$ is of order $dt^{3/2}$. 
For a function $F(X, t)$ of a stochastic process $X(t)$:

- Start over with a Taylor expansion

  $$dF = \frac{\partial F(x, t)}{\partial x} dx + \frac{\partial F(x, t)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F(x, t)}{\partial x^2} dx^2$$

  $$+ \frac{\partial^2 F(x, t)}{\partial x \partial t} dx \, dt + \frac{1}{2} \frac{\partial^2 F(x, t)}{\partial t^2} dt^2.$$

- Now make the substitutions $x = X(t)$ and $dx = dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t)$.

- For a first order (linear) approximation, we have no problem in discarding the higher order $dt^2$ term.

- Nor does the mixed term present a problem:  

  $$dX \, dt = (f(X(t), t) \, dt + g(X(t), t) \, dW(t)) \, dt.$$ 

  But $dW(t) \sim \sqrt{dt} N(0, 1)$, so $dX \, dt$ is of order $dt^{3/2}$.
Stochastic Chain Rule

For a function $F(X, t)$ of a stochastic process $X(t)$:

- Start over with a Taylor expansion

$$dF = \frac{\partial F(x, t)}{\partial x} dx + \frac{\partial F(x, t)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F(x, t)}{\partial x^2} dx^2$$

$$+ \frac{\partial^2 F(x, t)}{\partial x \partial t} dx \, dt + \frac{1}{2} \frac{\partial^2 F(x, t)}{\partial t^2} dt^2.$$

- Now make the substitutions $x = X(t)$ and $dx = dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t)$.

- For a first order (linear) approximation, we have no problem in discarding the higher order $dt^2$ term.

- Nor does the mixed term present a problem:

$$dX \, dt = (f(X(t), t) \, dt + g(X(t), t) \, dW(t)) \, dt.$$

But $dW(t) \sim \sqrt{dt} N(0, 1)$, so $dX \, dt$ is of order $dt^{3/2}$.
Ito’s Formula

Ito formula for \( dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t) \):

- The problem is with the second order term in \( dx^2 \) because \( dW^2 \sim \delta t \, \chi^2(1) \), which has mean \( \delta t \) and variance \( 2\delta t^2 \). So it is reasonable that the term approaches \( \delta t \) as \( \delta t \to 0 \).

- The net result is that

\[
\begin{aligned}
dF &= \frac{\partial F(X, t)}{\partial X} \, dX + \frac{\partial F(X, t)}{\partial t} \, dt + \frac{1}{2} \frac{\partial^2 F(X, t)}{\partial X^2} \, dX^2.
\end{aligned}
\]

- Substitute \( dX = f \, dt + g \, dW \), discard \( dW \, dt \) and \( dt^2 \) terms and get

\[
\begin{aligned}
dF &= \left( F_X f + F_t + \frac{1}{2} F_{XX} g^2 \right) \, dt + F_X g \, dW.
\end{aligned}
\]
Ito’s Formula

Ito’s formula for \( dX(t) = f(X(t), t) \, dt + g(X(t), t) \, dW(t) \):

- The problem is with the second order term in \( dx^2 \) because \( dW^2 \sim \delta t \chi^2(1) \), which has mean \( \delta t \) and variance \( 2\delta t^2 \). So it is reasonable that the term approaches \( \delta t \) as \( \delta t \to 0 \).

- The net result is that

\[
    dF = \frac{\partial F(X, t)}{\partial X} \, dX + \frac{\partial F(X, t)}{\partial t} \, dt + \frac{1}{2} \frac{\partial^2 F(X, t)}{\partial X^2} \, dX^2.
\]

- Substitute \( dX = f \, dt + g \, dW \), discard \( dW \, dt \) and \( dt^2 \) terms and get

\[
    dF = \left( F_X f + F_t + \frac{1}{2} F_{XX} g^2 \right) \, dt + F_X g \, dW
\]
Ito’s Formula

Ito formula for $dX(t) = f(X(t), t)\, dt + g(X(t), t)\, dW(t)$:

- The problem is with the second order term in $dx^2$ because $dW^2 \sim \delta t \chi^2(1)$, which has mean $\delta t$ and variance $2\delta t^2$. So it is reasonable that the term approaches $\delta t$ as $\delta t \to 0$.

- The net result is that
  
  $$dF = \frac{\partial F(X, t)}{\partial X} dX + \frac{\partial F(X, t)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F(X, t)}{\partial X^2} dX^2.$$

- Substitute $dX = f \, dt + g \, dW$, discard $dW \, dt$ and $dt^2$ terms and get
  
  $$dF = \left( F_X f + F_t + \frac{1}{2} F_{XX} g^2 \right) dt + F_X g \, dW$$
Applications

Example

The linear model for volatile stock price $X(t)$ with drift $\lambda$ and volatility $\mu$

$$dX(t) = \lambda X(t) \, dt + \mu X(t) \, dW(t).$$

Suppose a portfolio consists of an option (buy or sell) for a share of the stock with price $p(X, t)$, and a short position of $\Delta$ shares of it. It's value: $F = p(X, t) - \Delta X$. By the Ito formula,

$$dF = \left( (p_X - \Delta) \lambda X + p_t + \frac{1}{2} p_{XX} \mu^2 X^2 \right) dt + (p_X - \Delta) \mu X \, dW.$$
Some Elementary Inequalities

Theorem

(Markov Inequality) Let \( u(X) \) be a non-negative function of the r.v. \( X \) with finite expected value. For all positive \( a \),

\[
P(u(X) \geq a) \leq \frac{E[u(X)]}{a}.
\]

Theorem

(Chebychev Inequality) If the r.v. \( X \) has finite variance \( \sigma^2 \) and expected value \( \mu \), then for all positive \( k \)

\[
P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.
\]
Some Elementary Inequalities

Theorem

(Markov Inequality) Let \( u(X) \) be a non-negative function of the r.v. \( X \) with finite expected value. For all positive \( a \),

\[
P(u(X) \geq a) \leq \frac{E[u(X)]}{a}.
\]

Theorem

(Chebychev Inequality) If the r.v. \( X \) has finite variance \( \sigma^2 \) and expected value \( \mu \), then for all positive \( k \)

\[
P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}.
\]
Remarkably simple!

- For Markov, let event $E = \{ x \mid u(x) \geq a \}$ and $f(x)$ be p.d.f. of $X$.

$$E [u(X)] = \int_{-\infty}^{\infty} u(x) f(x) \, dx \geq \int_E a f(x) \, dx = a P(E).$$

Now divide by $a$ and we’re done!

- For Chebychev, take $u(X) = (X - \mu)^2$, $a = \sigma^2 k^2$ and obtain from Markov

$$P(|X - \mu| \geq k\sigma) \equiv P\left((X - \mu)^2 \geq \sigma^2 k^2\right) \leq \frac{\sigma^2}{\sigma^2 k^2} = \frac{1}{k^2}.$$

Done!
Remarkably simple!

- For Markov, let event $E = \{ x \mid u(x) \geq a \}$ and $f(x)$ be p.d.f. of $X$.

$$E[u(X)] = \int_{-\infty}^{\infty} u(x) f(x) \, dx \geq \int_{E} a f(x) \, dx = a P(E).$$

Now divide by $a$ and we’re done!

- For Chebychev, take $u(X) = (X - \mu)^2$, $a = \sigma^2 k^2$ and obtain from Markov

$$P(|X - \mu| \geq k\sigma) \equiv P \left( (X - \mu)^2 \geq \sigma^2 k^2 \right) \leq \frac{\sigma^2}{\sigma^2 k^2} = \frac{1}{k^2}.$$

Done!
Recall that a sequence $X_n$ of r.v.’s converges in probability to r.v. $X$ if for all $\epsilon > 0$,

$$\lim_{n \to \infty} P (|X_n - X| \geq \epsilon) = 0.$$

With this definition and the previous theorems, we can explain “it is reasonable that...”
(In derivation for $\int_0^T W(t) \, dW(t)$),

\[
\sum_{j=0}^{N-1} (dW_j)^2 = \sum_{j=0}^{N-1} (W_{j+1} - W_j)^2 \sim \delta t \chi^2(N).
\]

Thus, this sum has mean $N \delta t = T$ and variance $\delta t^2 2N = 2T \delta t$. So it is reasonable that the sum approaches $T$ as $\delta t \to 0$.

Let $X_N = \sum_{j=0}^{N-1} (dW_j)^2$ and let $k = 1/\sqrt{\delta t}$, so that

\[
k\sigma = \frac{1}{\sqrt{\delta t}} 2T \delta t = 2T \frac{\sqrt{T}}{\sqrt{N}} = \frac{2T^{3/2}}{\sqrt{N}}, \quad \text{and} \quad k^2 = \frac{1}{\delta t} = \frac{N}{T}.
\]

Hence

\[
P\left(\left|X_N - T\right| \geq \frac{2T^{3/2}}{\sqrt{N}}\right) \leq \frac{T}{N} \to 0 \quad \text{as} \quad N \to \infty.
\]

Hence $X_N$ converges to $T$ in probability.
(In derivation for $\int_0^T W(t) \, dW(t)$),

$\sum_{j=0}^{N-1} (dW_j)^2 = \sum_{j=0}^{N-1} (W_{j+1} - W_j)^2 \sim \delta t \chi^2(N)$. Thus, this sum has mean $N \delta t = T$ and variance $\delta t^2 2N = 2T \delta t$. So it is reasonable that the sum approaches $T$ as $\delta t \to 0$.

Let $X_N = \sum_{j=0}^{N-1} (dW_j)^2$ and let $k = 1/\sqrt{\delta t}$, so that

$$k \sigma = \frac{1}{\sqrt{\delta t}} 2T \delta t = 2T \frac{\sqrt{T}}{\sqrt{N}} = \frac{2T^{3/2}}{\sqrt{N}}, \quad \text{and} \quad k^2 = \frac{1}{\delta t} = \frac{N}{T}.$$

Hence

$$P\left[ |X_N - T| \geq \frac{2T^{3/2}}{\sqrt{N}} \right] \leq \frac{T}{N} \to 0 \text{ as } N \to \infty.$$

Hence $X_N$ converges to $T$ in probability.
“It is reasonable that…”

- (In derivation for $\int_0^T W(t) \, dW(t)$),
  
  $$\sum_{j=0}^{N-1} (dW_j)^2 = \sum_{j=0}^{N-1} (W_{j+1} - W_j)^2 \sim \delta t \chi^2(N).$$
  
  Thus, this sum has mean $N \delta t = T$ and variance $\delta t^2 2N = 2T \delta t$. So it is reasonable that the sum approaches $T$ as $\delta t \to 0$.

- Let $X_N = \sum_{j=0}^{N-1} (dW_j)^2$ and let $k = 1/\sqrt{\delta t}$, so that
  
  $$k\sigma = \frac{1}{\sqrt{\delta t}} 2T \delta t = 2T \frac{\sqrt{T}}{\sqrt{N}} = \frac{2T^{3/2}}{\sqrt{N}},$$
  
  and $k^2 = \frac{1}{\delta t} = \frac{N}{T}$.

- Hence

  $$P \left[ |X_N - T| \geq \frac{2T^{3/2}}{\sqrt{N}} \right] \leq \frac{T}{N} \to 0 \text{ as } N \to \infty.$$ 

- Hence $X_N$ converges to $T$ in probability.
(In derivation for $\int_0^T W(t) \, dW(t)$),

\[
\sum_{j=0}^{N-1} (dW_j)^2 = \sum_{j=0}^{N-1} (W_{j+1} - W_j)^2 \sim \delta t \chi^2(N).
\]

Thus, this sum has mean $N \delta t = T$ and variance $\delta t^2 2N = 2T \delta t$. So it is reasonable that the sum approaches $T$ as $\delta t \to 0$.

Let $X_N = \sum_{j=0}^{N-1} (dW_j)^2$ and let $k = 1/\sqrt{\delta t}$, so that

\[
k \sigma = \frac{1}{\sqrt{\delta t}} 2T \delta t = 2T \frac{\sqrt{T}}{\sqrt{N}} = \frac{2T^{3/2}}{\sqrt{N}}, \quad \text{and} \quad k^2 = \frac{1}{\delta t} = \frac{N}{T}.
\]

Hence

\[
P \left[ |X_N - T| \geq \frac{2T^{3/2}}{\sqrt{N}} \right] \leq \frac{T}{N} \to 0 \quad \text{as} \quad N \to \infty.
\]

Hence $X_N$ converges to $T$ in probability.
(In derivation of Ito’s formula) “The problem is with the second order term in $dx^2$ because $dW^2 \sim \delta t \chi^2(1)$, which has mean $\delta t$ and variance $2\delta t^2$. So it is reasonable that the term approaches $\delta t$ as $\delta t \to 0$.”

Take $\delta t = 1/N$, $X_N = dW^2/\delta t$, $k = 1/(2\sqrt{\delta t})$ and as above obtain that

$$P \left[ |X_N - 1| \geq \frac{1}{\sqrt{N}} \right] \leq \frac{4}{N} \to 0 \text{ as } N \to \infty.$$  

Hence $dW^2/\delta t$ converges to 1 in probability.
(In derivation of Itō’s formula) “The problem is with the second order term in $dx^2$ because $dW^2 \sim \delta t \chi^2 (1)$, which has mean $\delta t$ and variance $2\delta t^2$. So it is reasonable that the term approaches $\delta t$ as $\delta t \to 0$.”

Take $\delta t = 1/N$, $X_N = dW^2 / \delta t$, $k = 1 / \left(2\sqrt{\delta t}\right)$ and as above obtain that

$$P \left[ |X_N - 1| \geq \frac{1}{\sqrt{N}} \right] \leq \frac{4}{N} \to 0 \text{ as } N \to \infty.$$

Hence $dW^2 / \delta t$ converges to 1 in probability.
(In derivation of Ito’s formula) “The problem is with the second order term in $dx^2$ because $dW^2 \sim \delta t \chi^2 (1)$, which has mean $\delta t$ and variance $2\delta t^2$. So it is reasonable that the term approaches $\delta t$ as $\delta t \to 0$.”

Take $\delta t = 1/N$, $X_N = dW^2/\delta t$, $k = 1/(2\sqrt{\delta t})$ and as above obtain that

$$P \left[ |X_N - 1| \geq \frac{1}{\sqrt{N}} \right] \leq \frac{4}{N} \to 0 \text{ as } N \to \infty.$$ 

Hence $dW^2/\delta t$ converges to 1 in probability.
Suppose that an iterative method is strongly convergent at \( \tau = T \) of order \( \gamma \), so that

\[
E [\| X_n - X(\tau) \|] \leq C \Delta t^\gamma.
\]

It follows that

\[
\frac{E [\| X_n - X(\tau) \|]}{\Delta t^{\gamma/2}} \leq C \Delta t^{\gamma/2}.
\]

By Markov,

\[
P \left( |X_n - X(\tau)| \geq \Delta t^{\gamma/2} \right) \leq C \Delta t^{\gamma/2},
\]

which is a strong statement about individual paths. For example, EM has \( \gamma = 1/2 \). Compare this with weak convergence.
Suppose that an iterative method is strongly convergent at $\tau = T$ of order $\gamma$, so that

$$E[|X_n - X(\tau)|] \leq C\Delta t^\gamma.$$ 

It follows that

$$\frac{E[|X_n - X(\tau)|]}{\Delta t^{\gamma/2}} \leq C\Delta t^{\gamma/2}.$$ 

By Markov,

$$P\left(|X_n - X(\tau)| \geq \Delta t^{\gamma/2}\right) \leq C\Delta t^{\gamma/2},$$

which is a strong statement about individual paths. For example, EM has $\gamma = 1/2$. Compare this with weak convergence.
Suppose that an iterative method is strongly convergent at $\tau = T$ of order $\gamma$, so that

$$E\left[|X_n - X(\tau)|\right] \leq C\Delta t^\gamma.$$  

It follows that

$$E\left[\frac{|X_n - X(\tau)|}{\Delta t^{\gamma/2}}\right] \leq C\Delta t^{\gamma/2}.$$  

By Markov,

$$P\left(|X_n - X(\tau)| \geq \Delta t^{\gamma/2}\right) \leq C\Delta t^{\gamma/2},$$

which is a strong statement about *individual paths*. For example, EM has $\gamma = 1/2$. Compare this with weak convergence.