Introduction to Stochastic Differential Equations

In part I of this lecture we will give an informal introduction to stochastic differential equations (SDEs), which serve as the basic tool for understanding and implementation of most important issues in interest rate modeling, and ultimately the analysis of inflation-linked products, which will be our main purpose here. Our introduction will be neither entirely formal nor rigorous, but it serves to set the notations and define the computational framework which we will need for the remainder of this course.

Stochastic Differential Equations

We begin by presenting an informal introduction to SDEs. We consider the scalar (one-dimensional) case to simplify exposition. However, extension to higher dimensions is generally straightforward from the concepts that will be discussed here.

Probability Space

Our starting point will be a probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_t, P)\). This space has the usual interpretation as a generic experiment. Here, \(\Omega\) represents the set of all possible outcomes of the random experiment. An element \(\omega \in \Omega\) represents the result of the experiment, and a subset \(A \subset \Omega\) represents an event (if \(\omega \in A\) is the result of the experiment, then we say that event \(A\) has occurred). \(\mathcal{F}\) is the \(\sigma\)–field representing the set of all events, and \((\mathcal{F}_t)_{t \geq 0}\) is a filtration (with \(\mathcal{F}_t \subseteq \mathcal{F}_u \subseteq \mathcal{F}\) for all \(t \leq u\)) representing information, that increases in time and never exceeds \(\mathcal{F}\).

We use the symbol \(\mathbb{E}\) to denote expectation, and \(\mathbb{E}[. \mid \mathcal{F}]\) to denote expectation conditional on the information contained in \(\mathcal{F}\). We may use the notations \(\mathbb{E}[. \mid \mathcal{F}_t]\) and \(\mathbb{E}_t[.]\) interchangeably.

From Deterministic to Stochastic

To help introduce stochastic differential equation it is helpful to begin with an example of deterministic differential equation, a simple population growth model. Let \(x(t) = x\) denote the population at time \(t\), and assume for simplicity a constant (proportional) population growth rate, so that the change in the population at \(t\) is given by the deterministic differential equation:

\[
dx = Kx dt, \quad x(0) = x_0
\]

where \(K\) is some constant. Now suppose that due to some inherent randomness we can no longer assume that the initial condition \(x_0\) to be deterministic constant. Then we may
assume $x_0$ to be a random variable $X_0(\omega)$ and to model the population growth by the differential equation:

$$dX_t(\omega) = KX_t(\omega)dt, \quad X_0(\omega)$$  \hspace{1cm} (EQ 2)

The solution to this equation is $X_t(\omega) = X_0(\omega)\exp(Kt)$. Note that $X_t(\omega)$ is a random variable, and in this case its randomness comes from the initial condition $X_0(\omega)$. For each experiment result $\omega$, the map $t \to X_t(\omega)$ is called the path (or sample path) of $X$ associated with $\omega$.

As a further step, suppose that even $K$ is not known for certain, but that our knowledge of $K$ is perturbed by some randomness, which we will model as the increment of a stochastic process $\{W_t(\omega), t \geq 0\}$, so that

$$dX_t(\omega) = (Kdt + dW_t(\omega))X_t(\omega), \quad X_0(\omega)$$  \hspace{1cm} (EQ 3)

Here, $dW_t(\omega)$ represents a noise process that essentially adds randomness to $K$.

**Stochastic Differential Equations**

The equation above is an example of stochastic differential equation (SDE). More generally, a SDE is written as

$$dX_t(\omega) = f_t(X_t(\omega))dt + \sigma_t(X_t(\omega))dW_t(\omega), \quad X_0(\omega)$$  \hspace{1cm} (EQ 4)

The function $f_t$, corresponding to the deterministic part of the SDE, is called the drift (the subscript $t$ indicates that it may depend on time $t$, in addition to the argument $X_t(\omega)$). The function $\sigma_t$ is called the diffusion coefficient. $dW_t(\omega)$ represents the noise term as above. Such an SDE (or more specifically, its solution) is generally referred to diffusion process. Diffusion processes generally have continuous paths for all $\omega$.

**Brownian Motion**

The process with increments $dW_t(\omega)$ representing the noise in the above SDE is the Brownian motion. It is an important process that is stationary and has independent Gaussian increments. More precisely, for any $0 < s < t < u$ and any $h > 0$:

**Independent increments:** $W_u(\omega) - W_s(\omega)$ independent of $W_s(\omega) - W_t(\omega)$

**Stationary increments:** $W_{t+h}(\omega) - W_{s+h}(\omega) \sim W_t(\omega) - W_s(\omega)$

**Gaussian increments:** $W_t(\omega) - W_s(\omega) \sim N(0, t - s)$
For example, one can show that $W_{t+h}(\omega) - W_t(\omega)$ is independent of the history of $W$ up to time $t$, and can therefore assume any value independently of $\{W_s(\omega), s < t\}$.

Even though Brownian motion can be shown to have continuous paths, it turns out that the properties above imply that they are (almost surely) nowhere differentiable. In fact, it can be shown that the paths of the Brownian motion unbounded variation (velocity), and hence $W_t(\omega) = dW_t(\omega) / dt$ does not exist.

**Stochastic Integrals**

Formally, the meaning of SDE above relies on rewriting it in the integral form:

$$X_t(\omega) = X_0(\omega) + \int_0^t f_s(X_s(\omega))ds + \int_0^t \sigma_s(X_s(\omega))dW_s(\omega)$$

(EQ 5)

The first integral can be defined as an Stieltjes integral on the paths. As for the integrals like the second term, although it cannot strictly speaking be defined as an Stieltjes integral (due to its unbounded variation), it is still possible to define such integrals as Stieltjes-like integral. The price paid is that the resulting integral will depend on the chosen points of the sub-partitions used in the limit that defines the integral.

More specifically, consider the following definition. Take the interval $[0,T]$ and consider the following partition of this intervals, depending on an integer $n$:

$$T^n_i = \min(T, \frac{i}{2^n}), i = 0,1,\ldots, \infty$$

(EQ 6)

Notice that from certain $i$ on all terms collapse to $T$, i.e., $T^n_i = T$ for all $i > 2^n T$. For each $n$ we have such a partition, and when $n$ increase the partition contains more elements, giving a better discrete approximation of the continuous interval $[0,T]$. Then define the integral as

$$\int_0^T \varphi_s(\omega)dW_s(\omega) = \lim_{n \to \infty} \sum_{i=0}^\infty \varphi_{t^n_i}(\omega) \left[ W_{T^n_i}(\omega) - W_{t^n_i}(\omega) \right]$$

(EQ 7)

where $t^n_i$ is any point in the interval $[T^n_i, T^n_{i+1})$. Now by choosing $t^n_i = T^n_i$, the beginning of the interval, we have the definition of the Ito integral, whereas by taking $t^n_i = (T^n_i + T^n_{i+1}) / 2$, the middle of the interval, we obtain a different result, the Stratonovich integral.

The Ito integral has interesting properties, for example it is a martingale, but leads to a calculus where the standard chain rule is not preserved since there is non-zero
contribution from the second order terms. On the contrary, the Stratonovich integral does preserve the ordinary chain rule and is preferred form the viewpoint of the properties of the paths.

To better understand the difference between these two definitions, consider the classical example of the stochastic integral computed both using Ito calculus and the Stratonovich calculus:

\[
\text{Ito} \Rightarrow \int_0^t W_s(\omega) dW_s(\omega) = \frac{1}{2} W_t(\omega)^2 - \frac{1}{2} dt \quad \text{(EQ 8)}
\]

\[
\text{Stratonovich} \Rightarrow \int_0^t W_s(\omega) dW_s(\omega) = \frac{1}{2} W_t(\omega)^2 \quad \text{(EQ 9)}
\]

Normally, to distinguish between these two versions, the symbol “o” is often introduced to denote the Stratonovich version as \(\int_0^t W_s(\omega) \circ dW_s(\omega)\). In differential notation, one then has

\[
\text{Ito} \Rightarrow d(W_t(\omega)^2) = 2W_t(\omega)dW_t(\omega) + dt \quad \text{(EQ 10)}
\]

\[
\text{Stratonovich} \Rightarrow d(W_t(\omega)^2) = 2W_t(\omega) \circ dW_t(\omega) \quad \text{(EQ 11)}
\]

In the Ito version the \(dt\) term originates from the second order effects which are not negligible like the ordinary calculus. The first integral is known as a **martingale** (it has the constant expected value equal to zero), which is an important probabilistic property. However, it does not satisfy formal rules of ordinary calculus, as instead does the second one (which is not a martingale). We will define the concept of martingale in more detail below.

### Martingales, Driftless SDEs and Semimartingales

Consider a process \(X\) satisfying the following **measurability** and **integrability** conditions:

**Measurability**: \(\mathcal{F}_t\) includes all information on \(X\) up to time \(t\), usually expressed in literature by saying that \((X_t, \mathcal{F}_t)\) is adapted to filteration \((\mathcal{F}_t)\).

**Integrability**: the relevant expectations exist.

A **martingale** is a process satisfying the above conditions and such that the following property holds for all \(t \leq T\):

\[
\mathbb{E}[X_T | \mathcal{F}_t] = X_t \quad \text{(EQ 12)}
\]
Hence, the expected value at a future time $T$ given the information at $t$ is equal to the value at $t$. This is among other things the picture of a *fair game*, where on average is not possible to gain or lose. It turns out that the martingale property is also suited to model *absence of arbitrage* in mathematical finance, where one requires that certain processes of the economy be martingales, so that there are no way to safe make money without taking on appropriate risk.

It is simple to see that a solution of a SDE is a martingale when $f_t(.) = 0$ for all $t$:

\[ dX_t(\omega) = \sigma_t(X_t(\omega))dW_t(\omega) \]  

(EQ 13)

Therefore, in diffusion process language, martingale means *driftless* diffusion process.

A *submartingale* is a similar process $X$ satisfying instead

\[ \mathbb{E}[X_T | \mathcal{F}_t] \geq X_t \] 

(EQ 14)

This means that the expected value of the process *grows* in time, and that averages of future values of the process given the current information always exceeds (or is equal to) the current value.

Similarly, a *supermartingale* satisfies

\[ \mathbb{E}[X_T | \mathcal{F}_t] \leq X_t \] 

(EQ 15)

and the expected value of the process decreases in time, so that the averages of future values of the process given the current information are always smaller or equal to the current value.

Finally, a process that is either a supermartingale or submartingale is usually termed a *semimartingale*.

### Quadratic Variation

The *quadratic variation* of a stochastic process $Y_t$, with continuous paths is defined as follows:

\[ \langle Y \rangle_T = \lim_{n \to \infty} \sum_{i=1}^{\infty} \left( Y_{t_i}^n(\omega) - Y_{t_{i-1}}^n(\omega) \right)^2 \] 

(EQ 16)

Intuitively this could be written as a second order integral:
\[
\langle Y \rangle_T = \int_0^T \left( dY_s(\omega) \right)^2
\]

(EQ 17)

or in its differential form

\[
d\langle Y \rangle_t = dY_t(\omega)dY_t(\omega)
\]

(EQ 18)

It is easy to check that a process \( Y \) whose paths are differentiable (almost surely) for all \( \omega \) satisfies \( \langle Y \rangle_T = 0 \). For a Brownian motion, it can be shown that

\[
\langle W \rangle_T = T \quad \text{for any } T
\]

(EQ 19)

which can also be informally written as

\[
d\langle W \rangle_t = dW_t(\omega)dW_t(\omega) = dt
\]

(EQ 20)

Again this comes from the fact that the Brownian motion moves so quickly that the second order effects are not negligible. Instead a process whose trajectories are differentiable can not move so quickly, and therefore the second order effects do not contribute.

In case the process \( Y \) is equal to the deterministic process \( dY_t = dt \) we immediately retrieve the classical result from ordinary calculus

\[
dt dt = 0
\]

(EQ 21)

**Quadratic Covariation**

We can also define the *quadratic covariation* of two stochastic process \( Y \) and \( Z \) with continuous paths follows:

\[
\langle Y, Z \rangle_T = \lim_{n \to \infty} \sum_{i=1}^n \left( Y_{T_{i-1}}(\omega) - Y_{T_i}(\omega) \right) \left( Z_{T_{i-1}}(\omega) - Z_{T_i}(\omega) \right)
\]

(EQ 22)

Intuitively this could be written as a second order integral:

\[
\langle Y, Z \rangle_T = \int_0^T (dY_s(\omega))(dZ_s(\omega))
\]

(EQ 23)

or even more intuitively in the differential form

\[
d\langle Y, Z \rangle_t = dY_t(\omega)dZ_t(\omega)
\]

(EQ 24)
It is simple to check that, denoting by $t$ the deterministic process $dY = dt$, that

$$\{W, t\}_t = 0 \quad \text{for any } T$$  \hspace{1cm} (EQ 25)

which can be informally written as

$$dW_t(\omega)dt = 0$$  \hspace{1cm} (EQ 26)

**Solution to General SDE**

Let us go back to our general SDE, and let us take time-homogenous coefficients for simplicity:

$$dX_t(\omega) = f(X_t(\omega))dt + \sigma(X_t(\omega))dW_t(\omega), \quad X_0(\omega)$$  \hspace{1cm} (EQ 27)

The standard theory tells us that for above SDE has to have a unique solution it is enough for both the $f$ and $\sigma$ coefficients satisfy the Lipschitz continuity conditions (and linear growth condition which does not automatically follow in the time-inhomogeneous case or with local Lipschitz continuity). These sufficient conditions are valid for deterministic differential equations as well and can be weakened, especially in dimension one.

**Interpretation of the Coefficients of the SDE**

For a deterministic differential equation such as

$$dx_t = f(x_t)dt$$  \hspace{1cm} (EQ 28)

with a smooth function $f$ one clearly has

$$\lim_{h \to 0} \frac{x_{t+h} - x_t}{h} \bigg|_{x_t = y} = f(y)$$  \hspace{1cm} (EQ 29)

$$\lim_{h \to 0} \frac{(x_{t+h} - x_t)^2}{h} \bigg|_{x_t = y} = 0$$  \hspace{1cm} (EQ 30)

The analogous relations for the SDE

$$dX_t(\omega) = f(X_t(\omega))dt + \sigma(X_t(\omega))dW_t(\omega)$$  \hspace{1cm} (EQ 31)
are given by

\[
\lim_{h \to 0} \frac{1}{h} \left[ X_{t+h}(\omega) - X_t(\omega) \right] \bigg|_{X_t = y} = f(y) \quad \text{(EQ 32)}
\]

\[
\lim_{h \to 0} \frac{1}{h} \left[ (X_{t+h}(\omega) - X_t(\omega))^2 \right] \bigg|_{X_t = y} = \sigma^2(y) \quad \text{(EQ 33)}
\]

The second limit is non-zero because of infinite velocity of the Brownian motion, while the first is analogous to the deterministic case.

**Ito’s Formula**

Ito’s formula provides the chain rule for differentials in the stochastic context. For a deterministic differential equation such as

\[
dx_t = f(x_t) dt \quad \text{(EQ 34)}
\]
given a smooth function \( \varphi(t, x) \), one can write the evolution of this function via the standard chain rule:

\[
d\varphi(t, x_t) = \frac{\partial \varphi}{\partial t}(t, x_t) dt + \frac{\partial \varphi}{\partial x}(t, x_t) dx_t \quad \text{(EQ 35)}
\]

We already observed that whenever the Brownian motion is involved such a fundamental rule of calculus needs to be modified. The general formulation of the chain rule for stochastic differential equations is the following.

Let \( \varphi(t, x) \) be a smooth function and let \( X_t(\omega) \) be the unique solution of the stochastic differential equation above. Then the Ito’s lemma reads

\[
d\varphi(t, X_t(\omega)) = \frac{\partial \varphi}{\partial t}(t, X_t(\omega)) dt + \frac{\partial \varphi}{\partial x}(t, X_t(\omega)) dX_t(\omega) + \frac{1}{2} \frac{\partial^2 \varphi}{\partial x^2}(t, X_t(\omega)) d\langle X \rangle_t \quad \text{(EQ 36)}
\]

comparing with the deterministic counterpart above we notice an extra term

\[
\frac{1}{2} \frac{\partial^2 \varphi}{\partial x^2}(t, X_t(\omega)) d\langle X \rangle_t = \frac{1}{2} \frac{\partial^2 \varphi}{\partial x^2}(t, X_t(\omega)) d\langle X \rangle_t
\]

This term can be computed algebraically by taking into account the rules of the quadratic variation and covariation discussed above
\[ dW_i(\omega)dW_j(\omega) = dt, \quad dW_i(\omega)dt = 0, \quad dt\,dt = 0 \]  
(EQ 38)

We thus obtain:

\[ d\varphi(t, X_i(\omega)) = \left[ \frac{\partial\varphi}{\partial t}(t, X_i(\omega)) + \frac{\partial\varphi}{\partial x}(t, X_i(\omega))f(X_i(\omega)) + \frac{1}{2} \frac{\partial^2\varphi}{\partial x^2}(t, X_i(\omega))\sigma^2(X_i(\omega)) \right] dt + \frac{\partial\varphi}{\partial x}(t, X_i(\omega))dW_i(\omega) \]  
(EQ 39)

**Stochastic Leibnitz Rule**

Also the classical *Leibnitz rule* for the differentiation of a product of functions is modified, analogously with the chain rule. The related formula can be derived as a corollary of the Ito’s formula in two dimensions and is reported below.

For a deterministic and differentiable functions \(x\) and \(y\) we have the deterministic Leibnitz rule:

\[ d(x, y) = dx_i \quad dy_i \]  
(EQ 40)

For two diffusion processes (and more generally semimartingales) \(X_i(\omega)\) and \(Y_i(\omega)\) we have instead

\[ d\left(X_i(\omega)Y_i(\omega)\right) = X_i(\omega)dY_i(\omega) + Y_i(\omega)dX_i(\omega) + dX_i(\omega)dY_i(\omega) \]  
(EQ 41)

or in more compact notation

\[ d\left(X_i, Y_i\right) = X_i \, dy_i + Y_i \, dx_i + d\left\{X, Y\right\} \]  
(EQ 42)

**Discretizing SDEs for Monte Carlo: Euler Scheme**

When one cannot solve the SDE explicitly, it is possible to simulate its trajectories through a discretization scheme. Here we briefly review the Euler Scheme.

Consider again the SDE:

\[ dX_i(\omega) = f(X_i(\omega))dt + \sigma(X_i(\omega))dW_i(\omega), \quad x_0 \]  
(EQ 43)

where for simplicity we took a deterministic initial condition. Let us integrate this equation between \(s\) and \(s + \Delta s\):

\[ X_{s+\Delta s}(\omega) = X_s(\omega) + \int_s^{s+\Delta s} f(X_i(\omega))dt + \int_s^{s+\Delta s} \sigma(X_i(\omega))dW_i(\omega) \]  
(EQ 44)
The Euler scheme consists of approximating the integral equation by

\[
\bar{X}_{s + \Delta s}(\omega) = \bar{X}_{s}(\omega) + f(\bar{X}_{s}(\omega))\Delta s + \sigma(\bar{X}_{s}(\omega))(W_{s + \Delta s}(\omega) - W_{s}(\omega))
\]  

(EQ 45)

with \( \bar{X}_{0}(\omega) = x_{0} \). If we apply this formula iteratively for a given set of \( s \)’s, say, \( s_{1} = 0, s_{2}, s_{3}, \ldots, s_{m} = T \), we obtain a discretized approximation \( \bar{X} \) of the solution \( X \) of the above SDE.

A stronger approximation can be attained with a more refined scheme, called the Milstein scheme. We will leave review of this scheme to the reader, but we will mention that when the diffusion coefficient \( \sigma(\cdot) \) is deterministic, i.e. \( \sigma(X_{t}(\omega)) = \sigma(t) \) with \( \sigma(\cdot) \) a deterministic function of time, the Euler and Milstein schemes coincide. Therefore, it is preferable to apply, when possible, the Euler scheme to the SDEs with deterministic diffusion coefficients, since this ensures the same stronger convergence of the Milstein scheme.

These discretization schemes can be useful for Monte Carlo simulations. Indeed, suppose we need to compute the expected value of a functional of \( X \) of the above SDE, say for simplicity

\[
\mathbb{E}_{0}\left[\varphi(X_{s_{1}}(\omega), \ldots, X_{s_{m}}(\omega))\right] \quad s_{0} = 0, \ s_{m} = T
\]  

(EQ 46)

this is a typical pricing problem for path-dependent payoffs in mathematical finance. Assume that the times \( s \) are close to each other.

We may decide to compute an approximation of this expectation as follows:

1. Select the number \( N \) of scenarios for the Monte Carlo simulation.
2. Set the initial value to \( X_{0}(\omega) = x_{0} \) for all scenarios \( j = 1, \ldots, N \).
3. Set \( k = 1 \).
4. Set \( s = s_{k} \) and \( \Delta s = s_{k+1} - s_{k} \) so that \( s + \Delta s = s_{k+1} \).
5. Generate \( N \) new realizations \( \Delta W^{j}, j = 1, \ldots, N \) of the standard Gaussian distribution \( \mathcal{N}(0,1) \) multiplied by \( \sqrt{\Delta s} \), thus simulating the distribution of \( W_{s + \Delta s}(\omega) - W_{s}(\omega) \).
6. Apply the above formula for each scenario \( j = 1, \ldots, N \) with the generated shocks

\[
\bar{X}_{s + \Delta s}^{j} = \bar{X}_{s}^{j} + f(\bar{X}_{s}^{j})\Delta s + \sigma(\bar{X}_{s}^{j})\Delta W^{j}
\]
7. Store $\bar{X}_{s_j + \Delta s}^j$ for all $j$.

8. If $s + \Delta s = s_m$ then stop, otherwise increase $k$ by one and start again on step 4.

9. Approximate the expected value by

$$\sum_{j=1}^{N} \frac{\varphi(\bar{X}_{s_1}^j, \ldots, \bar{X}_{s_N}^j)}{N}$$

(EQ 47)

Notice that the increments can be generated as new independent draws from a Gaussian distribution at each iteration because the Brownian motion has independent (stationary Gaussian) increments.

Here we have assumed that the $s$’s in the expectation are close enough for the scheme to be applied directly, otherwise we would need to add additional points to achieve the required degree of accuracy.

**Examples of SDEs**

We now present some relevant examples of SDEs are often encountered in this course.

**Linear SDEs with Deterministic Coefficients**

A SDE is said to be **linear** if both its drift and diffusion coefficients are first order polynomials (or affine functions) of the state variable. We here consider the particular case:

$$dX_t(\omega) = (\alpha_t + \beta_t X_t(\omega))dt + \nu_t dW_t(\omega), \quad X_0(\omega) = x_0$$

(EQ 48)

where $\alpha, \beta, \nu$ are deterministic functions of time that are regular enough to ensure existence and uniqueness of a solution.

It can be shown that a stochastic integral of a deterministic function is the same both in the Ito and Stratonovich sense. Hence, by writing the above equation in integral form we see that the same equation holds in the Stratonovich sense:

$$dX_t(\omega) = (\alpha_t + \beta_t X_t(\omega))dt + \nu_t \circ dW_t(\omega), \quad X_0(\omega) = x_0$$

(EQ 49)

so that we can solve it by ordinary calculus for linear differential equations. We obtain:
\[ X_t(\omega) = e^{\int_0^t \beta_s ds} \left[ x_0 + \int_0^t e^{-\int_u^t \beta_s du} \alpha_s ds + \int_0^t e^{-\int_u^t \beta_s du} \nu_s dW_s(\omega) \right] = \]

\[ x_0 e^{\int_0^t \beta_s ds} + \int_0^t e^{\int_u^t \beta_s du} \alpha_s ds + \int_0^t e^{\int_u^t \beta_s du} \nu_s dW_s(\omega) \]

It is simple to argue that distribution of the solution \( X_t \) is normal at each time \( t \).

Intuitively, this holds since the last stochastic integral is the limit of a sum of independent normal random variables. Indeed we have:

\[ X_t \sim N \left( x_0 e^{\int_0^t \beta_s ds} + \int_0^t e^{\int_u^t \beta_s du} \alpha_s ds, \int_0^t e^{2\int_u^t \beta_s du} \nu_s^2 ds \right) \] (EQ 51)

The major examples of models based on linear SDEs are that of \textit{Vasicek} and \textit{Hull-White} instantaneous short rate models.

**Lognormal (Linear) SDEs**

The lognormal SDE can be obtained as an exponential of a solution of a linear equation (as above) with deterministic diffusion coefficient. Indeed, let us take \( Y_t = \exp(X_t) \), where \( X_t \) evolves according to Equation 48. Since \( \ln(Y_t) = X_t \) we have that:

\[ d \ln Y_t(\omega) = (\alpha_t + \beta_t \ln Y_t(\omega)) dt + \nu_t dW_t(\omega) \] (EQ 52)

Equivalently, we write by Ito’s lemma:

\[ dY_t(\omega) = e^{X_t(\omega)} dX_t(\omega) + \frac{1}{2} e^{X_t(\omega)} dX_t(\omega)dX_t(\omega) \]

\[ = Y_t(\omega) \left[ \alpha_t + \beta_t \ln Y_t(\omega) + \frac{1}{2} \nu_t^2 \right] dt + \nu_t Y_t(\omega) dW_t(\omega) \] (EQ 53)

Hence, the process \( Y_t \) has lognormal marginal density. A major example of model based on this SDE is the \textit{Black-Karasinski} model.

**Geometric Brownian Motion**

The geometric Brownian motion is a particular case of process satisfying lognormal linear SDE. Its evolution is defined according to:

\[ dX_t(\omega) = \mu X_t(\omega) dt + \sigma X_t(\omega) dW_t(\omega) \] (EQ 54)
where $\mu, \sigma$ are positive constants. To check that $X$ is indeed lognormal process, one can compute $d \ln (X_t)$ according to Ito’s lemma and obtain:

$$X_t(\omega) = X_0 \exp \left\{ \left( \mu - \frac{1}{2} \sigma^2 \right) t + \sigma W_t(\omega) \right\}$$  \hspace{1cm} (EQ 55)

From the seminal work of Black and Scholes on, processes of this type are frequently used in option pricing theory to model general asset price dynamics. Notice that this process is a submartingale, in that clearly:

$$E[X_T | \mathcal{F}_t] = e^{\mu(T-t)} X_t \geq X_t$$  \hspace{1cm} (EQ 56)

Finally, notice that by setting $Y_t(\omega) = e^{-\mu t} X_t(\omega)$ is a martingale since we have:

$$dY_t(\omega) = \sigma Y_t(\omega) dW_t(\omega)$$  \hspace{1cm} (EQ 57)

**Square-Root Processes**

An interesting case of non-linear SDE is given by

$$dX_t(\omega) = (\alpha_t + \beta_t X_t(\omega)) dt + \sqrt{X_t(\omega)} dW_t(\omega), \quad X_0(\omega) = x_0$$  \hspace{1cm} (EQ 58)

A process with this dynamics is commonly referred to as square-root process. Major examples of models based on this dynamics are Cox, Ingersoll and Ross (CIR) instantaneous short rate model, and a particular case of constant-elasticity of variance (CEV) model for stock prices:

$$dX_t(\omega) = \mu X_t(\omega) dt + \sqrt{X_t(\omega)} dW_t(\omega)$$  \hspace{1cm} (EQ 59)

In general, square-root processes are naturally linked to non-central $\chi^2$ - square distributions. In particular, there are simplified versions of above for which the resulting process $X$ is strictly positive and analytically tractable, like the case of CIR model.

**The Feynman-Kac Theorem**

The *Feynman-Kac* theorem, under certain conditions, allows us the express the solution of a given partial differential equation (PDE) as expected value of a function of a suitable diffusion process whose drift and diffusion coefficients are defined in terms of the PDE coefficients.
**Theorem [Feynman-Kac]:** Given Lipschitz continuous $f(x)$ and $\sigma(x)$ and a smooth function $\varphi$, the solution of the PDE:

$$\frac{\partial V}{\partial t}(t, x) + f(x) \frac{\partial V}{\partial x}(t, x) + \frac{1}{2} \sigma^2(x) \frac{\partial^2 V}{\partial x^2}(t, x) = rV(t, x)$$  \hspace{1cm} (EQ 60)

with terminal boundary condition

$$V(T, x) = \varphi(x)$$  \hspace{1cm} (EQ 61)

can be expressed as the following expectation:

$$V(t, x) = e^{-r(T-t)} \mathbb{E}\{\varphi(X_T) \mid X_t = x\}$$  \hspace{1cm} (EQ 62)

where the diffusion process $X$ has dynamics, starting from $x$ at time $t$, given by

$$dX_s(\omega) = f(X_s(\omega))ds + \sigma(X_s(\omega))d\tilde{W}_s(\omega), \quad s \geq t, \quad X_t(\omega) = x$$

under the probability measure $\tilde{P}$ under which the expectation $\mathbb{E}\{\cdot\}$ is taken. The process $\tilde{W}$ is a standard Brownian motion under $\tilde{P}$.

Notice that the terminal condition determines the function $\varphi$ of the diffusion process whose expectation is relevant, whereas the PDE coefficients determine the dynamics of the diffusion process.

The importance of Feynman–Kac theorem is that it establishes a link between the PDEs of traditional analysis and diffusion processes in stochastic calculus. Solutions of PDEs can be interpreted as expectations of suitable transformations of solutions of stochastic differential equations and vice versa.

**The Girsanov Theorem**

The *Girsanov theorem* shows how a SDE changes due to changes in the underlying probability measure. It is based on the fact that SDE drift depends on the particular probability measure $P$ in our probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, and that, if we change the probability measure in a “regular” manner, the drift of the equation changes while the diffusion coefficients remain the same. The Girsanov theorem is thus useful when we want to change the coefficients of the SDE.

Indeed, suppose that we are given two measures $P^*$ and $P$ on the space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$. Two such measures are said to be *equivalent*, written $P^* \sim P$, if they share the same sets of null probability (or probability one, which is equivalent). Therefore the two measures are equivalent when they agree on which events on $\mathcal{F}$ hold almost surely. Similar
definitions apply also for measures restricted to $\mathcal{F}_t$, thus expressing equivalence of two measures up to the time $t$.

When two measures are equivalent, it is possible to express the first in terms of the second through the Radon-Nikodym derivative. Indeed, there exists a martingale $\rho_t$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ such that

$$\mathbb{P}^*(A) = \int_A \rho_t(\omega) d\mathbb{P}(\omega) \quad \text{(EQ 63)}$$

which can be written in the more concise form as

$$\frac{d\mathbb{P}^*}{d\mathbb{P}}|_{\mathcal{F}_t} = \rho_t \quad \text{(EQ 64)}$$

The process $\rho_t$ is called the Radon-Nikodym derivative of $\mathbb{P}^*$ with respect to $\mathbb{P}$ restricted to $\mathcal{F}_t$.

When in need of computing the expected value of an integrable random variable $X$, it may be useful to switch from one measure to another equivalent one. Indeed, it is possible to prove that the following equivalence holds:

$$\mathbb{E}^*[X] = \int_{\Omega} X(\omega) d\mathbb{P}^*(\omega) = \int_{\Omega} X(\omega) \frac{d\mathbb{P}^*}{d\mathbb{P}}(\omega) d\mathbb{P}(\omega) = \mathbb{E}\left[ X \frac{d\mathbb{P}^*}{d\mathbb{P}} \bigg| \mathcal{F}_t \right] \quad \text{(EQ 65)}$$

where $\mathbb{E}^*$ and $\mathbb{E}$ denote expected values with respect to the probability measures $\mathbb{P}^*$ and $\mathbb{P}$, respectively. More generally, when dealing with conditional expectations, we can prove that

$$\mathbb{E}^*[X | \mathcal{F}_t] = \frac{1}{\rho_t} \mathbb{E}\left[ X \frac{d\mathbb{P}^*}{d\mathbb{P}} \bigg| \mathcal{F}_t \right] \quad \text{(EQ 66)}$$

**Theorem [Girsanov]:** Consider again the stochastic differential equation, with Lipschitz coefficients:

$$dX_t(\omega) = f(X_t(\omega)) dt + \sigma(X_t(\omega))dW_t(\omega), \quad x_0 \quad \text{(EQ 67)}$$

under $\mathbb{P}$. Let be given a new drift $f^*(x)$ and assume $(f^*(x) - f(x))/\sigma(x)$ to be bounded. Define the measure $\mathbb{P}^*$ by

$$\mathbb{P}^*$$

(EQ 68)
\[
\frac{dP^*(\omega)}{dP}(\omega) \bigg|_{\mathcal{F}_t} = \exp \left\{ -\frac{1}{2} \int_0^t \left( \frac{f^*(X_s(\omega)) - f(X_s(\omega))}{\sigma(X_s(\omega))} \right)^2 ds + \int_0^t \frac{f^*(X_s(\omega)) - f(X_s(\omega))}{\sigma(X_s(\omega))} dW_s(\omega) \right\}
\]

Then \( P^* \) is equivalent to \( P \). Moreover, the process \( W^* \) defined by

\[
dW^*_t(\omega) = -\left[ \frac{f^*(X_t(\omega)) - f(X_t(\omega))}{\sigma(X_t(\omega))} \right] dt + dW_t(\omega)
\]

is a Brownian motion under \( P^* \), and

\[
dx_t(\omega) = f^*(X_t(\omega))dt + \sigma(X_t(\omega))dW^*_t(\omega)
\]

(EQ 69)

An example of application of Girsanov’s theorem in mathematical finance is when one moves from the real-world asset price dynamics

\[
ds_t(\omega) = \mu S_t(\omega)dt + \sigma S_t(\omega)dW_t(\omega)
\]

(EQ 71)

to the risk-neutral ones

\[
ds_t(\omega) = r S_t(\omega)dt + \sigma S_t(\omega)dW^*_t(\omega)
\]

(EQ 72)

This is accomplished by setting

\[
\frac{dP^*}{dP}(\omega) \bigg|_{\mathcal{F}_t} = \exp \left\{ -\frac{1}{2} \left( \frac{\mu - r}{\sigma} \right)^2 t - \frac{\mu - r}{\sigma} W_t(\omega) \right\}
\]

(EQ 73)