ITÔ CALCULUS AND DERIVATIVE PRICING WITH RISK-NEUTRAL MEASURE

MAX CYTRYNBAUM

ABSTRACT. This paper will develop some of the fundamental results in the theory of Stochastic Differential Equations (SDE). After a brief review of stochastic processes and the Itô Calculus, we set our sights on some of the more advanced machinery needed to work with SDE. Chief among our results are the Feynman-Kac Theorem, which establishes a link between stochastic methods and the classic PDE approach, and the Girsanov theorem, which allows us to change the drift of an Itô diffusion by switching to an equivalent martingale measure. These results are also valuable on a practical level, and we will consider some of their applications to derivative pricing calculations in mathematical finance.

CONTENTS

1. INTRODUCTION AND MOTIVATION

In many applications, the evolution of a system with some fixed initial state is subject to random perturbations from the environment. For example, a stock price may rise in the long term yet be subject to seemingly random short-term fluctuations. Similarly, a small object in a liquid suspension may experience random impulses from collisions with surrounding water particles. This leads us to the consideration of a differential equation of the following form:

$$\frac{dX_t}{dt} = b(t, X_t) + \sigma(t, X_t) \cdot W_t \tag{1.1}$$

where W_t represents "white noise". The way to formulate this precisely is through Itô calculus and Brownian motion. We will briefly review some of the basic results from Itô calculus. We assume the reader has a working knowledge of measure theoretic probability and is comfortable with the properties of conditional expectation.

2. ITÔ CALCULUS AND BROWNIAN MOTION

We will closely follow the exposition in [4]. Let $(\Omega, \mathcal{F}, P)$ be a probability space.

Definition 2.1 (Filtration)**.** We say that $\{\mathcal{F}_t\}_{t\geq 0}$ is a *filtration* on the space $(\Omega, \mathcal{F}, P)$ if

(1) $\mathcal{F}_t$ is sub σ-algebra of $\mathcal{F}$ for each t

(2) $\mathcal{F}_t \subset \mathcal{F}_s$ for all $t \leq s$.

A measurable function $f : [0, \infty) \times \Omega \to \mathbb{R}^n$ is said to be *adapted to the filtration* $\{\mathcal{F}_t\}_{t \geq 0}$ if $\omega \to f(t, \omega)$ is an $\mathcal{F}_t$ measurable function for each t.

The sub σ-algebra $\mathcal{F}_t$ represents the amount of information available to us at time t. Intuitively, this says that the value of the $\mathcal{F}_t$-adapted random variable $f_t(\omega)$ can be determined by only using information in the σ-algebras up to $\mathcal{F}_t$. This inability to see into the future is important in mathematical finance, where $f(t, \omega)$ may represent some investment strategy and $\mathcal{F}_t$ may be the price information available up to time t. Thus, we cannot use future price information to determine our strategy.

Definition 2.2 (Brownian Motion)**.** We call a stochastic process $B(t, \omega) = B_t(\omega)$ a *version of Brownian motion* starting at $x \in \mathbb{R}$ when

1. $B_0 = x$ for all ω
2. For all times $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$, $B(t_2) - B(t_1), B(t_2) - B(t_3), \cdots, B(t_n) - B(t_{n-1})$ are independent random variables
3. $B(t + h) - B(t)$ is normally distributed with mean 0 and variance h for all $t, h \in [0, \infty]$
4. $B(t)$ is P-a.s. t-continuous

Similarly, we may define $B_t \in \mathbb{R}_n$ to be a Brownian motion if $B(t) = (B_1(t), \ldots, B_n(t))$ where each $B_i(t)$ is a version of 1-dimensional brownian motion and all the $B_i(t)$ are independent. For a construction of Brownian motion, we refer the reader to [5].

Definition 2.3 (Martingale)**.** A stochastic process $\{M_t\}$ is said to be a *martingale* with respect to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ if

1. $\{M_t\}$ is adapted to $\{\mathcal{F}_t\}_{t \geq 0}$
2. $E[|M_t|] < \infty$ for all t
3. For all $0 \leq s < t$ and for a.s. ω
$$E[M_t | \mathcal{F}_s] = M_s$$

A martingale neither increases nor decreases on average and can thus be thought of as a model of a fair game. From the above it is also clear that for any t we have $E[M_t] = E[M_0]$. In what follows, we let $\{\mathcal{F}_t\}_{t \geq 0}$ be the filtration generated by $\{B_t\}$ (i.e. $\mathcal{F}_t$ is the σ-algebra generated by $\omega \to B(t, \omega)$ for each t.) Then using independence of increments and properties of conditional expectation, we calculate:

$$E[B_t | \mathcal{F}_s] = E[B_t - B_s | \mathcal{F}_s] + E[B_s | \mathcal{F}_s] = E[B(t) - B(s)] + B(s) = B(s)$$

so that Brownian motion is a martingale with respect to its own filtration. For more information on martingales in discrete-time, see [7]. We now attempt to find a precise interpretation of the SDE (1.1). To do this, we rewrite (1.1) in the more suggestive form:

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \tag{2.1}$$

where we have interpreted white noise in terms of brownian motion as $W_t dt = dB_t$, which leads to the tentative integral equation:

$$X_t = \int b(t, X_t)dt + \int \sigma(t, X_t)dB_t \tag{2.2}$$

Thinking in terms of Riemann-Stieltjes integration, it seems natural to interpret an integral of the form $\int f(t, \omega)dB_t(\omega)$ as a limit of sums of the form:

$$\sum_j f(t_j, \omega)\Delta B_{t_j}(\omega)$$

Intuitively, the increments $f(t_j,\omega)\Delta B_{t_j}(\omega)$ reflect the random change in X_t due to white noise. We now make this notion precise.

Definition 2.4 (Itô Integrable)**.** We say that a function $f(t,\omega)$ is *Itô Integrable on* $[S,T]$ if the following conditions are satisfied:

(1) $f(t,\omega)$ is $\mathcal{B}\otimes\mathcal{F}$ measurable
(2) $f(t,\cdot)$ is $\mathcal{F}_t$-measurable for each t
(3) $E\left[\int_S^T f^2(t,\omega)dt\right] < \infty$

In this case we write $f \in \mathcal{V}[S,T]$, following the notation in [4]. We say that a function $\phi(t,\omega)$ is *elementary* if $\phi \in \mathcal{V}[S,T]$ and $\phi(t,\omega) = \sum_j \phi(t_j,\omega)\mathcal{X}_{[t_j,t_{j+1}]}$ where $\{t_j\}$ is a partion of the interval $[S,T]$. We define:

$$\int_S^T \phi\, dB_t := \sum_j \phi(t_j)\Delta B_j$$

where $\Delta B_j := B(t_{j+1}) - B(t_j)$. This yields the following

Lemma 2.5 (Itô Isometry)**.** *If ϕ is elementary and $\phi \in \mathcal{V}[S,T]$, then we have*

$$E\left[\left(\int_S^T f(t,\omega)dB_t\right)^2\right] = E\left[\int_S^T f^2(t,\omega)dt\right]$$

Proof.

$$E\left[\left(\int_S^T f(t,\omega)dB_t\right)^2\right] = E\left[\left(\sum_j f(t_j,\omega)\Delta B_j\right)^2\right]$$
$$= E\left[\sum_{i,j} f(t_i,\omega)f(t_j,\omega)\Delta B_j\Delta B_i\right]$$

Now since the the f_{t_j} are $\mathcal{F}_{t_j}$ adapted and by the independence of increments for Brownian motion, we get that ΔB_j is independent from $\Delta B_i f_{t_j} f_{t_i}$ when $i < j$. Since $E(\Delta B_j) = 0$, $E(f_{t_j} f_{t_i}\Delta B_j\Delta B_i) = 0$ for $i < j$. Then the above simplifies to:

$$E\left[\sum_j f_{t_j}^2(\Delta B_j)^2\right] = \sum_j \Delta t_j E[f_{t_j}^2] = E\left[\sum_j f_{t_j}^2\Delta t_j\right] = E\left[\int_S^T f^2(t,\omega)dt\right]$$

The first identity uses the fact the ΔB_j and f_{t_j} are independent and the result from definition (2.2) that $E(\Delta B_j^2) = \Delta t_j$. □

We can now use this result to extend the Itô integral to all of $\mathcal{V}[S,T]$. Specifically, given $f \in \mathcal{V}[S,T]$ let $\{\phi_n\}$ be a sequence of elementary functions such that:

$$E\left[\int_S^T (f(t,\omega) - \phi_n(t,\omega))^2 dt\right] \to 0 \qquad as \quad n\to\infty \tag{2.3}$$

As the reader can check, Lemma 2.5 and equation (2.3) imply that $\{\phi_n\}$ is a Cauchy sequence in $L^2(P)$, where P is the law of standard Brownian motion. Since $L^2(P)$ is complete, we can define:

$$\int_S^T f(t,\omega)dB_t(\omega) = \lim_{n\to\infty}\int_S^T \phi_n(t,\omega)dB_t(\omega) \tag{2.4}$$

where the limit is taken in $L^2(P)$. Standard analysis arguments can be used to show that this limit is independent of the chosen sequence $\{\phi_n\}$. For a closer look at the existence of such a sequence $\{\phi_n\}$, see [4]. We now provide a brief review of some of the properties of $\int f dB_t$:

Proposition 2.6 (Properties of the Itô Integral). *Let $[S,T] \subset \mathbb{R}^+$ and let $f \in \mathcal{V}[S,T]$.*

(1) $f \to \int_S^T f dB_t$ *is a linear operator on* $\mathcal{V}[S,T]$
(2) $E\left[\int_S^T f dB_t\right] = 0$
(3) $\int_S^T f dB_t$ *is* $\mathcal{F}_T$*-measurable*
(4) $E\left[\left(\int_S^T f dB_t\right)^2\right] = E\left[\int_S^T f^2 dt\right]$

These all follow by proving the statement for elementary functions and taking limits using equation (2.4). With a little more work we can show

Theorem 2.7 (Martingale Property). *Let $f \in \mathcal{V}[0,T]$, $0 \le t \le T$.*

(1) *There exists a t-continuous version of* $\int_0^t f dB_s$
(2) *Put* $M_t(\omega) = \int_0^t f dB_s(\omega)$. *Then* M_t *is a* $\{\mathcal{F}_t\}_{t\ge 0}$ *martingale.*

Theorem 2.7(2) follows from the fact that B_t is a martingale with respect to its own filtration. We refer the reader to page 32 of [4] for the proof. The above extends naturally to n dimensions. Let $\mathcal{V}^{n,m}[S,T]$ be the space of $n \times m$ matrices $v(t,\omega)$ such that each $v_{ij}(t,\omega) \in \mathcal{V}[S,T]$. Let $B_t \in \mathbb{R}^m$. Then we define

$$\int_S^T v dB_t = \int_S^T \begin{pmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{nm} \end{pmatrix} dB_t$$

to be the $n \times 1$ vector with ith component

$$\sum_{k=1}^{m} \int_S^T v_{ik}(t,\omega) dB_k(t,\omega)$$

Definition 2.8 (Itô Process). Let $v \in \mathcal{V}[0,T]$ and let $u(t,\omega)$ be a measurable stochastic process adapted to the filtration $\{\mathcal{F}_t\}_{t\ge 0}$ and such that

$$P\left[\int_0^t |u(s,\omega)| ds < \infty \quad \text{for all} \quad t \ge 0\right] = 1$$

Then we say that $X(t,\omega)$ is *Itô Process* if X_t has the form

$$X_t = X_0 + \int_0^t u(s,\omega) ds + \int_0^t v(s,\omega) dB_s$$

This extends naturally to the n-dimensional case by considering expressions of the form

$$X(t) = X(0) + \int_0^t u\, ds \int_0^t v\, dB_s$$

where u is a $1 \times n$ matrix and v is an $n \times m$ matrix with all components satisfying the conditions of Definition 2.8. This leads us to the following fundamental theorem, which shows that Itô processes are invariant under sufficiently smooth maps. This can be considered the stochastic version of the chain rule.

Theorem 2.9 (Itô Lemma). *Let*

$$X(t) = X(0) + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dB_s$$

be an n-dimensional Itô process. Let $f : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^m$, where f is a C^2 map. Define a new stochastic process $Y(t) = f(t, X_t)$. Then Y(t) is an Itô process as above, and the ith coordinate is given by

$$dY_i(t) = \frac{\partial f_i}{\partial t}(t, X_t)dt + \sum_k \frac{\partial f_i}{\partial x_k}(t, X_t)dX_k + \sum_{k,j} \frac{\partial^2 f_i}{\partial x_k \partial x_j}(t, X_t)(dX_k)(dX_j)$$

where $(dX_i)(dX_j)$ is calculated according to the rule $(dt)^2 = 0$, $(dt)(dB_k) = (dB_k)(dt) = 0$ for all i, and $(dB_k)(dB_j) = (dB_j)(dB_k) = \delta_{jk}dt$.

We omit the proof (see page 46 of [4]) in order to start solving SDE as quickly as possible, merely noting that the second order term above, notably absent in the classic chain rule, comes from the fact that Brownian motion has non-zero quadratic variance.

Proposition 2.10 (Integration by Parts). *Let X_t and Y_t be 1-dimensional stochastic processes. Then*

$$d(X_tY_t) = X_tdY_t + Y_tdX_t + dX_t \cdot dY_t \tag{2.5}$$

so that

$$\int_0^t X_s dY_s = X_tY_t - X_0Y_0 - \int_0^t Y_s dX_s - \int_0^t (dX_s)(dY_s)$$

Proof. Let $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x_1, x_2) = x_1x_2$. Then g is smooth so applying the Itô lemma,

$$\begin{aligned} d(g(X_t, Y_t)) &= \frac{\partial g}{\partial x_1}(X_t, Y_t) + \frac{\partial g}{\partial x_2}(X_t, Y_t) + \frac{1}{2}\left(2 \cdot \frac{\partial^2 g}{\partial x_1 \partial x_2}(X_t, Y_t)\right) \\ &= Y_tdX_t + X_tdY_t + (dX_t)(dY_t) \end{aligned}$$

Then the integral equation above follows immediately from equation (2.5). □

3. Stochastic Differential Equations

In this section, we illustrate the power of Itô's Lemma by explicitly solving an elementary SDE. We will then prove an important existence and uniqueness result for well-behaved SDE. Let us consider a model for population growth in a stochastic environment.

Example 3.1 (Geometric Brownian Motion). Put $\theta_t = \alpha + \beta \cdot W_t$ where W_t is white noise and define

$$dX_t = \theta_t X_t dt$$

This seems a reasonable model for a growing population subject to random environmental shocks. Using the Brownian motion interpretation, we can rewrite this as

$$dX_t = \alpha X_t dt + \beta X_t dB_t \tag{3.1}$$

Expecting some sort of exponential behavior, we apply the Itô Lemma to the C^2 function $f(x) = \ln(x)$. Using Theorem 2.9, we get

$$\begin{aligned} d(f(X_t)) &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2 \\ &= \frac{1}{X_t}dX_t - \frac{1}{2}\frac{1}{X_t^2}(\beta^2 X_t^2 dt) \\ &= \frac{1}{X_t}dX_t - \frac{\beta^2}{2}dt \end{aligned}$$

Combining this with equation 3.1, we get that

$$d(\ln(X_t)) = \left(\alpha - \frac{\beta^2}{2}\right) dt + \beta dB_t$$

so that

$$\ln\left(\frac{X_t}{X_0}\right) = \left(\alpha - \frac{\beta^2}{2}\right) t + \beta B_t$$

or

$$X_t = X_0 \exp\left(\left(\alpha - \frac{\beta^2}{2}\right) t + \beta B_t\right)$$

From the law of iterated logarithm for Brownian motion (see [5]), we see that if $\alpha > \frac{\beta^2}{2}$, $X_t(\omega) \to \infty$ as $t \to \infty$ for a.s. ω, while for lower values of α the stochastic term takes over. Given empirically determined coefficients α and β, when do we expect the population to reach a certain size N? We will return to this question later.

Using the Itô Lemma, a wide range of classic ODE methods such as matrix exponentiation and integrating factors can be applied to suitable SDE. See the exercises in Chapter 5 of [4] for a wealth of examples. Now we turn to the interesting question of existence and uniqueness of solutions to SDE.

Lemma 3.2 (Gronwall Inequality)**.** *Let $v(t)$ be a non-negative, Borel function such that for some constants C, A, we have*

$$v(t) \leq C + A \int_0^t v(s) ds$$

then we have

$$v(t) \leq C \exp(At)$$

Theorem 3.3 (Existence and Uniqueness of SDE)**.** *Let $T \geq 0$ and $b(\cdot,\cdot) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma(\cdot,\cdot) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ be measurable functions such that we have*

$$|b(t,x)| + |\sigma(t,x)| \leq C(1 + |x|) \tag{3.2}$$

for some constant C and for all $t \in [0,T]$, $x \in \mathbb{R}^n$ (where $|\sigma| = \sum |\sigma_{ij}|$). Assume also that

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq D|x - y| \tag{3.3}$$

for some constant D and all $t \in [0,T]$, $x, y \in \mathbb{R}^n$. Let Z be a random variable with finite second moment such that Z is independent of $\mathcal{F}_\infty^{(m)} := \sigma(\mathcal{F}_t : t \geq 0)$, where $\mathcal{F}_t$ is generated by B_t. Then the SDE

$$dX_t = b(t, X_t) + \sigma(t, X_t) \qquad X_0 = Z(\omega)$$

has a unique t-continuous solution that is adapted to $\mathcal{F}_t^Z := \sigma(\mathcal{F}_t, \sigma(Z))$ and

$$E\left[\int_0^T |X_t|^2 dt\right] < \infty \tag{3.4}$$

Proof. Uniqueness: Suppose there exists another solution $\tilde{X}_t = \tilde{Z} + \int_0^t b(s,\tilde{X}_s)ds + \int_0^t \sigma(s,\tilde{X}_s)ds$. put $f(t,\omega) = b(t,X_t) - b(t,\tilde{X}_t)$ and put $g(t,\omega) = \sigma(t,X_t) - \sigma(t,\tilde{X}_t)$. Then we calculate

$$\begin{aligned}
E(|X_t - \tilde{X}_t|^2) &= E\left[\left(Z - \tilde{Z} + \int_0^t f ds + \int_0^t g dB_s\right)^2\right] \\
&\leq 3E\left[(Z - \tilde{Z})^2\right] + 3E\left[\left(\int_0^t f ds\right)^2\right] + 3E\left[\left(\int_0^t g dB_s\right)^2\right] \\
&\leq 3E\left[(Z - \tilde{Z})^2\right] + 3tE\left[\int_0^t f^2 ds\right] + 3E\left[\int_0^t g^2 ds\right]
\end{aligned}$$

The above squares should be interpreted in the matrix sense. The first inequality uses $(x+y+z)^2 \leq 3x^2+3y^2+3z^2$ and the 2nd inequality uses the Itô isometry. Then equation (3.3) implies that the above is

$$\begin{aligned}
&\leq 3E\left[(Z-\tilde{Z})^2\right] + 3tD^2E\left[\int_0^t |X_s-\tilde{X}_s|^2 ds\right] + 3D^2E\left[\int_0^t |X_s-\tilde{X}_s|^2 ds\right] \\
&= 3E\left[(Z-\tilde{Z})^2\right] + 3D^2(t+1)\int_0^t E\left[|X_s-\tilde{X}_s|^2\right] ds
\end{aligned}$$

Condition (3.4) allows us to use Fubini in the second line. Put $B = 3D^2(T+1)$, $A = 3E[|Z-\tilde{Z}|^2]$. Then with $v(t) = E[|X_t - \tilde{X}_t|^2]$, we have that

$$v(t) \leq A + B\int_0^t v(s)ds$$

Then Gronwall implies that

$$v(t) \leq A\exp(Bt)$$

Set $Z = \tilde{Z}$, then we have that $v(t) = 0$ so that for a.s. ω, $|X_t - \tilde{X}_t| = 0$. We get that

$$P[|X_t - \tilde{X}_t| = 0 \quad \text{for all} \quad t \in \mathbb{Q}\cap[0,T]] = 1$$

By the assumed t-continuity of solutions, we find that $t \to |X_t - \tilde{X}_t|$ is continuous, so that we actually have $X_t = \tilde{X}_t$ for all $t \in [0,T]$ for a.s. ω, which completes the proof of uniqueness. The proof of existence uses Picard iterations and is very similar to the proof for ODE. Because of its length, it is ommitted. See [4, p.70]. □

The solution given by the previous theorem is called a *strong* solution because it is constructed with respect to a specific version of Brownian motion B_t. By a *weak* solution, we mean *any* pair of processes $(\widehat{X_t}, \widehat{B_t})$ that satisfy the equation

$$\widehat{X_t} = b(t,\widehat{X_t})dt + \sigma(t,\widehat{X_t})d\widehat{B_t} \tag{3.5}$$

An SDE has a *strongly unique* solution if any two solutions are equal for a.s. (t,ω). An SDE is said to have a *weakly unique* solution if any two solutions have the same law. This distinction is subtle but will be necessary for our proof of the Girsanov theorem. What we require is the following:

Lemma 3.4. *Let b and σ be as in the previous theorem. Then any solution to equation (3.5) is weakly unique.*

4. Itô Diffusions

An Itô diffusion is a model of a well-behaved stochastic process that shares some fundamental characteristics with nice processes like n-dimensional Brownian motion. Specifically,

Definition 4.1 (Itô Diffusion)**.** We say a stochastic process $X_t \in \mathbb{R}^n$ is a an *Itô diffusion* if it satisfies a stochastic differential equation

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$

where $b : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are functions satisfying the conditions of Theorem 3.3. Since we have no time argument, this just means that

$$|b(x)-b(y)| + |\sigma(x)-\sigma(y)| \leq D|x-y| \qquad \text{for all} \quad x,y \in \mathbb{R}^n$$

Theorem 3.3 shows that the above is well-defined. The next concept is fundamental to the study of both martingales and stochastic processes.

Definition 4.2 (stopping time)**.** Let $\{\mathcal{M}_t\}_{t\geq 0}$ be a filtration on the probability space $(\Omega, \mathcal{M}, P)$ and let $\tau : \Omega \to \mathbb{R}$. We say that τ is a *stopping time w.r.t the filtration* $\{\mathcal{M}_t\}_{t\geq 0}$ if for each t we have

$$\{\omega : \tau(\omega) \leq t\} \in \mathcal{M}_t$$

Sometimes we will write $\{\tau \leq t\}$ for $\{\omega : \tau(\omega) \leq t\}$. The intuition behind the above definition is that one can know whether or not a stopping time has passed at time t without looking into the future (only using the information in $\mathcal{M}_t$). Thus if X_t represents the value of a portfolio, and $\{\mathcal{M}_t\}_{t\geq 0}$ is the filtration generated by X_t, then $\sup\{t : X_t > 20\}$ would not be a stopping time while $\inf\{t : X_t > 20\}$ would be. The reader can show that any deterministic time t is trivially a stopping time w.r.t. any filtration.

Proposition 4.3. *Let τ_1 and τ_2 both be stopping times w.r.t. $\{\mathcal{M}_t\}$. Then the following random variables are also stopping times w.r.t. $\{\mathcal{M}_t\}$*

(1) $\tau_1 \wedge \tau_2$
(2) $\tau_1 \vee \tau_2$

Here, $(\tau_1 \wedge \tau_2)(\omega) = \tau_1(\omega) \wedge \tau_2(\omega) := \min(\tau_1(\omega), \tau_2(\omega))$. The maximum in (2) is interpreted likewise. Moreover, if τ_n is a sequence of stopping times such that $\lim_{n\to\infty} \tau_n = \tau$ and such that, for each ω, $\tau_n(\omega) \leq \tau(\omega)$ for large n, then τ is a stopping time w.r.t the same filtration.

Proof. For (1) and (2) Note that $\{\tau_1 \wedge \tau_2 \leq t\} = \{\tau_1 \leq t\} \cup \{\tau_2 \leq t\} \in \mathcal{M}_t$ and that $\{\tau_1 \vee \tau_2 \leq t\} = \{\tau_1 \leq t\} \cap \{\tau_2 \leq t\} \in \mathcal{M}_t$. For the last statement, the reader can check that in this case we have

$$\{\tau \leq t\} = \bigcup_{n=1}^{\infty} \bigcap_{m \geq n} \{\tau_m \leq t\} \in \mathcal{M}_t$$

□

Definition 4.4. Let $\{\mathcal{M}_t\}$ be a filtration on Ω and let $\mathcal{M}_\infty$ be the smallest σ-algebra containing all the $\mathcal{M}_t$. By $\mathcal{M}_\tau$ we mean the σ-algebra of all sets in $M \in \mathcal{M}_\infty$ such that

$$M \cap \{\tau \leq t\} \in \mathcal{M}_t$$

Intuitively, $\mathcal{M}_\tau$ is the σ-algebra of all events that occured before the stopping time τ. In the following, $X_t^{s,x}$ will mean a diffusion started at a point $x \in \mathbb{R}^n$ and started at time s. $E^{s,x}$ will denote expectation w.r.t. the law afforded by $X_t^{s,x}$. One of the properties that Itô diffusions share with Brownian motion is their "forgetfulness". Formally, $E^x[X(t+h)|\mathcal{F}_t] = E^{X(t)}[X_h]$. In fact, more is true:

Theorem 4.5 (Strong Markov Property)**.** *Let $X_t^x \in \mathbb{R}^n$ be an Itô diffusion such that $X_0 = x$ and let τ be a stopping time w.r.t. the filtration $\{\mathcal{F}_t^{(m)}\}_{t\geq 0}$. Then given any bounded Borel function $f : \mathbb{R}^n \to \mathbb{R}^n$,*

$$E^x[f(X_{\tau(\omega)+h})|\mathcal{F}_\tau] = E^{X_\tau(\omega)}[f(X_h)]$$

The proof involves a lot of technical bookkeeping, so we omit it (see page 117 of [4]). Let $\mathcal{M}_\infty = \sigma(\mathcal{M}_t : t \geq 0)$, where $\mathcal{M}_t$ is the σ-algebra generated by X_t. If we define the operator θ_t so that $\theta_t(X_s) = X_{s+t}$. It can be shown that the strong Markov property extends to $\mathcal{M}_\infty$ so that for f a bounded, $\mathcal{M}_\infty$ measurable function:

$$E^x[\theta_\tau f|\mathcal{F}_\tau^{(m)}] = E^{X_\tau(\omega)}[f] \tag{4.1}$$

The next lemma is fundamental:

Lemma 4.6. *Let Y_t be an Itô process in $\mathbb{R}^n$ such that*

$$Y_t = x + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dB_s$$

Let τ be a stopping time w.r.t. the filtration $\{\mathcal{F}_t^{(m)}\}_{t\geq 0}$. Let $f \in C_0^2(\mathbb{R}^n)$ and let u and v above be bounded on the set of (s,ω) such that $Y(t,\omega) \in supp(f)$. Also, suppose $E^x[\tau] < \infty$. Then we get

$$E^x(f(Y_\tau)) = f(x) + E^x\left[\int_0^\tau \left(\sum_i \frac{\partial f}{\partial x_i}(Y_s)u_i(s,\omega) + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}(Y_s)(vv^T)_{ij}(s,\omega)\right) ds\right] \quad (4.2)$$

Proof. Note that $dY_i(t) = u_i dt + \sum_k v_{ik} dB_k$. Applying the Itô lemma to the C^2 function f, we find that

$$\begin{aligned} d(f(Y)) &= f(Y_0) + \sum_i \frac{\partial f}{\partial x_i} dY_i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}(dY_i)(dY_j) \\ &= f(x) + \sum_i \frac{\partial f}{\partial x_i} u_i dt + \sum_{i,k}\frac{\partial f}{\partial x_i} v_{ik} dB_k + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}\sum_k v_{ik}v_{jk}dt \end{aligned}$$

where we have used that $(dY_i)(dY_j) = \sum_k v_{ik}v_{jk}dt$ from the Itô lemma. Note that this can be rewritten as $(vv^T)_{ij}$. Writing this differential equation in integral form, replacing t with the stopping time τ and taking an expectation, we get that

$$\begin{aligned} E^x[f(Y_\tau)] = f(x) &+ E\left[\int_0^\tau \frac{\partial f}{\partial x_i}(Y_s)u_i(s,\omega) + \sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}(Y_s)(vv^T)_{ij}(s,\omega)ds\right] \\ &+ E\left[\int_0^\tau \sum_{i,k}\frac{\partial f}{\partial x_i}(Y_s)v_{ik}(s,\omega)dB_k\right] \end{aligned}$$

Therefore it will suffice to prove that the last expectation is 0. Let g be a bounded borel function on $\mathbb{R}^n$ s.t. $|g| < K$. then apparently $g \in \mathcal{V}[0,T]$ for all T, so with $k > 0$ we calculate

$$E^x\left[\int_0^{\tau\wedge k} g(Y_s)dB_s\right] = E^x\left[\int_0^k \mathcal{X}_{\{s\leq\tau\}} g(Y_s)dB_s\right] = 0 \quad (4.3)$$

because if τ is a stopping time then the product of g and the indicator $\mathcal{X}_{\{s\leq\tau\}}$ is in $\mathcal{V}[0,T]$, so we can apply Proposition 2.6(2). Now note that

$$\begin{aligned} E^x\left[\left(\int_0^\tau g(Y_s)dB_s - \int_0^{\tau\wedge k} g(Y_s)dB_s\right)^2\right] &= E^x\left[\int_{\tau\wedge k}^\tau g(Y_s)^2 ds\right] \\ &\leq K^2 E^x[\tau - \tau\wedge k] \end{aligned}$$

and since $E^x[\tau] < \infty$ we get that this tends to 0 as $k \to \infty$. Note that the equality above uses the Itô isometry. Then we have that

$$E^x\left[\int_0^{\tau\wedge k} g(Y_s)dB_s\right] \to E^x\left[\int_0^\tau g(Y_s)dB_s\right]$$

in $L^2(P^x)$ as $k \to \infty$. Then equation (4.3) implies that the limit is identically 0, which completes the proof of the lemma. □

In many applications, we will wish to associate to a second-order partial differential operator to each Itô diffusion. This operator encodes a wealth of information about the process X_t and will allow us to forge a direct link between stochastic theory and PDE through the Feynman-Kac formula.

Definition 4.7 (Infinitesimal Generator)**.** Let X_t be an Itô diffusion. We define the *infinitesimal generator*:

$$Af(x) = \lim_{t \downarrow 0} \frac{E^x(f(X_t)) - f(x)}{t}$$

We let D_A denote the set of functions f for which the above limit exists at all $x \in \mathbb{R}^n$.

Theorem 4.8. *Let $X_t = b(X_t)dt + \sigma(X_t)dB_t$ be an Itô diffusion in $\mathbb{R}^n$. If $f \in C_0^2(\mathbb{R}^n)$, then $f \in D_A$ and for all $x \in \mathbb{R}^n$ we have that*

$$Af(x) = \sum_i b_i(x)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\sigma\sigma^T)_{ij}(x)$$

Proof. Apply Lemma 4.6 to the Itô diffusion X_t with $\tau = t$. Using bounded convergence to pass the limit through the expectation, the theorem is a simple consequence of the fundamental theorem of calculus. □

Thinking of the operator A as a derivative of sorts, the following important theorem can be seen as a generalization of the fundamental theorem of calculus:

Theorem 4.9 (Dynkin's Formula)**.** *Let $f \in C_0^2(\mathbb{R}^n)$ and let X_t be an Itô diffusion as above. Assume also that τ is a stopping time w.r.t. the filtration $\{\mathcal{F}_t^{(m)}\}_{t \geq 0}$. Then we have*

$$E^x[f(X_\tau)] = f(x) + E^x\left[\int_0^\tau Af(X_s)ds\right]$$

Proof. This follows immediately from Theorem 4.8 and the above lemma. □

Example 4.10 (Population Growth and Hitting Times)**.** In Section 3, we considered a stochastic model of population growth and produced the equation

$$X_t = x\exp\left(\left(\alpha - \frac{\beta^2}{2}\right)t + \beta B_t\right) \tag{4.4}$$

If $\alpha > \frac{\beta^2}{2}$, $X_t \to \infty$ for a.s. ω. If $\alpha < \frac{\beta^2}{2}$, $X_t \to 0$ for a.s. ω. Given coefficients α and β, we ask

(i) for $\alpha > \frac{\beta^2}{2}$ when do we expect the population to reach a certain size R?
(ii) for $\alpha < \frac{\beta^2}{2}$, do we expect the population to ever reach size R?

For fixed $\gamma \in \mathbb{R}$ we can use Theorem 4.8 to compute X_t's infinitesimal generator for $f(x) = x^\gamma$. By letting f go to 0 in a smooth way, we may assume that $x^\gamma \in C_0^2(\mathbb{R})$. Then since $dX_t = \alpha X_t dt + \beta X_t dB_t$, we get

$$Af(x) = \beta x f'(x) + \frac{1}{2}f''(x)\alpha^2 x^2 = \gamma x^\gamma\left(\beta + \frac{\alpha^2(\gamma - 1)}{2}\right) \tag{4.5}$$

If we put $\gamma_1 = 1 - \frac{2\alpha}{\beta^2}$ and $f(x) = x^{\gamma_1}$ then from equation (4.5) we see that $Af(x) = 0$. We define $\tau_n = \inf\{t > 0 : X_t \notin [\frac{1}{n}, R]$. This is known as a *hitting time*, and it can be shown that the hitting time for any Borel set is a stopping time w.r.t. $\{\mathcal{F}_t\}$. Let $g(x) = \ln(x)$ for $x < \ln(R)$ and let g be 0 outside some compact set containing $[0, \ln(R)]$ (i.e. let g go to 0 in some smooth way). Then Theorem 4.8 implies that $Ag(x) = \alpha - \frac{1}{2}\beta^2$. Applying the Dynkin formula to g with the stopping time $\tau_n \wedge k$ we get

$$E^x[g(X_{\tau_n \wedge k})] = \ln(x) + E^x\left[\int_0^{\tau_n \wedge k} Ag(X_s)ds\right] = \ln(x) + (\alpha - \frac{\beta^2}{2})E^x[\tau_n \wedge k]$$

Since g is bounded, we can apply bounded convergence on both sides and let $k \to \infty$. Since $\tau_n \wedge k \to \tau_n$ pointwise for a.s. ω, the above implies that $E^x[\tau_n] < \infty$. This allows us to apply Dynkin to the function $f(x) = x^{\gamma_1}$ with the stopping time τ_n. Since $Af = 0$,

$$E^x[f(X_{\tau_n})] = f(x) + 0 = x^{1-\frac{2\alpha}{\beta^2}} \quad \text{for all n} \tag{4.6}$$

If we let $(1 - p_n)$ be the probability that X_t exits the bottom of the interval $[\frac{1}{n}, R]$ before it exits at R, (4.6) becomes

$$(1 - p_n)\left(\frac{1}{n}\right)^{\gamma_1} + p_n(R)^{\gamma_1} = x^{\gamma_1}$$

so that

$$p_n = \left(\frac{x}{R}\right)^{\gamma_1} - (1 - p_n)\left(\frac{1}{R \cdot n}\right)^{\gamma_1}$$

Since $\alpha < \frac{1}{2}\beta^2$ implies that $\gamma_1 > 0$, we see that $\lim_n p_n = \left(\frac{x}{R}\right)^{\gamma_1}$. Let E_n be the event that $X_{\tau_n} = R$. Then $E_n \subset E_{n+1}$

$$A := \{\omega : X_t(\omega) = R \quad \text{for some } t > 0\} = \bigcup_n E_n$$

So from basic probability,

$$P^x(A) = P^x\left(\bigcup_n E_n\right) = \lim_n p_n = \left(\frac{x}{R}\right)^{\gamma_1}$$

where x is the initial population $x = X_0$, which answers question (i). Now suppose that $\alpha > \frac{\beta^2}{2}$. For (ii), if we apply the Dynkin formula with $f(x) = \ln(x)$ as above, similar arguments can be used to show that $E^x[\tau_R]$, the expected time for a population with initial size X_t to reach population R, is given by

$$E^x[\tau_R] = \frac{\ln \frac{R}{x}}{\alpha - \frac{1}{2}\beta^2}$$

We have given the above proof in great detail to illustrate some important techniques, such as truncation of stopping times and limiting procedures with bounded convergence. For more information on this model, see *Exercise 7.9* of [4]. Using the A operator, we will now prove an interesting connection between classic PDE theory and stochastic processes.

5. Feynman-Kac and PDE

Theorem 5.1 (Feynman-Kac)**.** *Let $X_t \in \mathbb{R}^n$ be an Itô diffusion with generator A. Let $f \in C_0^2(\mathbb{R}^n)$ and $q \in C(\mathbb{R}^n)$ such that q is lower bounded. Set*

$$v(t, x) = E^x\left[\exp\left(-\int_0^t q(X_s)ds\right) f(X_t)\right]$$

Then we have that

(i)
$$\frac{\partial v}{\partial t} = Av - qv \tag{5.1}$$
$$v(0, x) = f(x) \tag{5.2}$$

(ii) *If there is another function $g(t, x) \in C^{1,2}(\mathbb{R}\times\mathbb{R}^n)$ such that g is bounded on $K\times\mathbb{R}^n$ for every compact $K \subset \mathbb{R}$, then if g satisfies (5.1) and (5.2) above we must have $g(t, x) = v(t, x)$.*

Proof. We let Z_t denote the stochastic process $\exp\left(-\int_0^t q(X_s)ds\right)$. We set $N_t = -\int_0^t q(X_s)ds$ and $h(x) = \exp(x)$. Then using the Itô lemma, we see that

$$d(h(N_t)) = h(N_t)dN_t + \frac{1}{2}h(N_t)(dN_t)^2 = -h(N_t)q(X_t)dt$$

so that we have

$$dZ_t = -Z_t q(X_t)dt \tag{5.3}$$

Note that $d(f(X_t))$ can be calculated using Lemma 4.6. Using Proposition 2.10, we see that if $Y_t := f(X_t)$,

$$d(Z_tY_t) = Z_tdY_t + Y_tdZ_t + (dZ_t)(dY_t) = Z_tdY_t + Y_tdZ_t$$

because $dZ_t = -Z_tq(X_t)dt$ implies that $(dZ_t)(dY_t) = 0$ by the Itô lemma. Then Z_tY_t is an Itô process. By the above assumptions, Z_tY_t is bounded for all ω, so using Fubini's theorem, equation (4.2) immediately implies that $E^x(f(X_t)Z_t)$ is differentiable. Then we directly calculate Av using the limit definition:

$$\begin{aligned}\frac{1}{s}E^x[v(t,X_s) - v(t,x)] &= \frac{1}{s}E^x[E^{X_s}[Z_tf(X_t)]] - E^x[f(X_t)Z_t] \\ &= \frac{1}{s}E^x[E^x[f(X_{t+s})\exp\left(-\int_0^t q(X_{s+r})dr\right)|\mathcal{F}_s] - E^x[f(X_t)Z_t|\mathcal{F}_s]]\end{aligned}$$

The second equality follows from the Markov property and properties of conditional expectation. We can rewrite this as

$$\begin{aligned}&= \frac{1}{s}E^x\left[E^x\left[f(X_{t+s})Z_{t+s}\exp\left(\int_0^s q(X_r)dr\right)\right] - E^x[f(X_t)Z_t]|\mathcal{F}_s\right] \\ &= \frac{1}{s}E^x[Z_{t+s}f(X_{t+s}) - Z_tf(X_t)] + \frac{1}{s}E^x\left[f(X_{t+s})Z_{t+s}\left(\exp\left(\int_0^s q(X_r)dr\right) - 1\right)\right]\end{aligned}$$

A calculation similar to the one that produced (5.3) shows that we can rewrite

$$\exp\left(\int_0^s q(X_r)dr\right) = \int_0^s Z_rq(X_r)dr + 1$$

Since the integrand is continuous for a.s. ω, it follows that $\left(\exp\left(\int_0^s q(X_r)dr\right) - 1\right)$ is differentiable, and its derivative at $t = 0$ is $Z_0q(X_0) = q(x)$. It follows from the lower boundedness of q that for each t, $Z_tf(X_t)$ is bounded for a.s. ω. Then since we already showed that $E^x[f(X_t)Z_t]$ is differentiable with respect to t, we can apply bounded convergence to show that

$$\frac{1}{s}E^x[Z_{t+s}f(X_{t+s}) - Z_tf(X_t)] \to \frac{\partial}{\partial t}E^x[f(X_t)Z_t] \quad \text{as} \quad s \to 0$$

and

$$\frac{1}{s}E^x\left[f(X_{t+s})Z_{t+s}\left(\exp\left(\int_0^s q(X_r)dr\right) - 1\right)\right] \to q(x)v(t,x) \quad \text{as} \quad s \to 0$$

Then we have shown that $Av = \frac{\partial v}{\partial t} - qv$. For (ii), we refer the reader to p.142 of [4] □

In the special case that $q = 0$ and with $v(t,x)$ as above, we find that $\frac{\partial v}{\partial t} = Av$, which is known as the *Kolmogorov Backward Equation*. To illustrate the theorem's power, we will consider two examples of deterministic PDE solved with stochastic methods.

Example 5.2 (Cauchy Problem)**.** Let $\phi \in C_0^2(\mathbb{R}^n)$. We seek a bounded solution g of the initial value problem

$$\frac{\partial g(t,x)}{\partial t} = \frac{1}{2}\Delta_x g(t,x) \tag{5.4}$$

$$g(0,x) = \phi(x) \tag{5.5}$$

Let $X_t = B_t \in \mathbb{R}^n$. Clearly $dB_t = I_n dB_t$, so using theorem (3.8) we get that $Af = \frac{1}{2}\Delta f$ for any $f \in C_0^2(\mathbb{R}^n)$. Let $v(t,x) = E^x[\phi(B_t)]$. Then since $\phi \in C_0^2(\mathbb{R}^n)$, we can apply Fubini to show that

$$v(t,x) = E^x[\phi(B_t)] = \int_{\mathbb{R}^n} \phi(y) \frac{1}{(2\pi t)^{\frac{n}{2}}} \exp\left(\frac{|x-y|^2}{2t}\right) dy$$

Then since ϕ is bounded, it is clear that for each t, we have $x \to v(t,x) \in C_0^2(\mathbb{R}^n)$. Therefore, for each t we have that $Av_t(x) = \frac{1}{2}\Delta v_t(x)$. Applying the Kolmogorov backward equation, we conclude that $v(t,x)$ is a solution to (5.4) and (5.5) and is bounded because ϕ is.

Example 5.3. In this example, we will give an explicit formula for the solution $u(t,x)$ to the initial value problem

$$\frac{\partial u}{\partial t} = \lambda u + \frac{1}{2}\Delta u$$
$$u(0,x) = f(x)$$

where $x \in \mathbb{R}^n$, λ is a constant and $f \in C_0^2(\mathbb{R}^n)$ is given. Set

$$v(t,x) = E^x\left[\exp\left(\int_0^t -\lambda ds\right) f(B_t)\right]$$

i.e. set $q(x) = -\lambda$ in the Feynman-Kac theorem. Then if we can prove that $v(t,x) \in C_0^2(\mathbb{R}^n)$, Feynman-Kac immediately implies that $v(t,x)$ is a solution to the above initial value problem (we showed above that for Brownian motion $Ag = \frac{1}{2}\Delta g$ for $g \in C_0^2(\mathbb{R})$.) Using the density of B_t, we compute the expectation:

$$v(t,x) = E^x[\exp(-\lambda t) f(B_t)] = \int_{\mathbb{R}^n} \frac{\exp(-\lambda t)}{(2\pi t)^{\frac{n}{2}}} f(y) \exp\left(\frac{|x-y|^2}{2t}\right) dy \tag{5.6}$$

whence it is clear that $x \to v(t,x) \in C_0^2(\mathbb{R}^n)$ for each t. Then $v(t,x)$ is indeed a solution to the above differential equation and is given explicitly by equation (5.6).

The next result tells us that if we change the drift coefficient of an Itô diffusion, the law of the new process is absolutely continuous w.r.t. the original process. The Radon-Nikodym derivative provided by the theorem is important for doing computations with risk-neutral measure in mathematical finance.

6. The Girsanov Theorem

First we prove a general result from probability theory:

Theorem 6.1 (Bayes' Rule)**.** *Let $(\Omega, \mathcal{N})$ be a measurable space equipped with a measure μ. Let $f \in L^1(\mu)$ and put $d\nu = f d\mu$. Let $\mathcal{H}$ be a sub σ-algebra of $\mathcal{N}$. Then if X is an r.v. such that*

$$\int_\Omega |X(\omega)| f(\omega) d\mu(\omega) < \infty$$

we get that

$$E_\nu[X|\mathcal{H}] \cdot E_\mu[f|\mathcal{H}] = E_\mu[fX|\mathcal{H}] \tag{6.1}$$

Proof. Using the elementary properties of conditional expectation, it will suffice to show that the left and right hand side of (6.1) have the same expectation over any set $H \in \mathcal{H}$:

$$\int_H E_\mu[Xf|\mathcal{H}]d\mu = \int_H Xf d\mu = \int_H X d\nu = \int_H E_\nu[X|\mathcal{H}]d\nu \tag{6.2}$$

where the second equality uses the $L^1(\mu)$ condition above. Note that

$$\int_H E_\nu[X|\mathcal{H}]d\nu = E_\mu[E_\nu[X|\mathcal{H}]f\mathcal{X}_H] = E_\mu[E_\mu[E_\nu[X|\mathcal{H}]f\mathcal{X}_H|\mathcal{H}]] \tag{6.3}$$

Since $E_\nu[Xf]$ and $\mathcal{X}_H$ are $\mathcal{H}$-measurable, we can pull them out of the inner expectation, giving

$$E_\mu[\mathcal{X}_H E_\nu[X|\mathcal{H}] \cdot E_\mu[f|\mathcal{H}]] = \int_H E_\nu[X|\mathcal{H}] \cdot E_\mu[f|\mathcal{H}]d\mu$$

Then we have shown that

$$\int_H E_\mu[Xf|\mathcal{H}]d\mu = \int_H E_\nu[X|\mathcal{H}] \cdot E_\mu[f|\mathcal{H}]d\mu$$

for all $H \in \mathcal{H}$, so (6.1) follows. □

We will need the following

Theorem 6.2 (Lévy Characterization of Brownian Motion)**.** *Let $X(t) \in \mathbb{R}^n$ be a stochastic process on the probability space $(\Omega, \mathcal{H}, P)$. Then the following are equvalent*

(i) *$X(t)$ is a Brownian motion w.r.t. the probability measure P (i.e. $P \circ X_t^{-1}$ is the law of Brownian motion on $\mathbb{R}^n$).*

(ii) *(a) $X(t)$ is a martingale w.r.t. its own filtration under the measure P (i.e. $E_P[M_t|\mathcal{M}_s] = M_s$)*
(b) $X_i(t)X_j(t) - \delta_{ij}t$ is a martingale w.r.t. its own filtration under the measure P

We refer the reader to Peres (2010) for more information. We may now state the main theorem of this section.

Theorem 6.3 (Girsanov Theorem)**.** *Let $X(t) \in \mathbb{R}^n$ be an Itô process such that $X_0 = 0$ and for fixed T with $0 \le t \le T \le \infty$*

$$dX_t = a(t, \omega)dt + dB_t$$

where $B_t \in \mathbb{R}^n$ is standard Brownian motion w.r.t. the measure P. Define

$$M_t = \exp\left(-\int_0^t a(s,\omega)dB_s - \frac{1}{2}\int_0^t a^2(s,\omega)ds\right)$$

and suppose that M_t is a martingale w.r.t. the filtration $\{\mathcal{F}_t^{(m)}\}$ generated by B_t under the measure P. Define the measure $dQ = M_T dP$. Then Q is a probability measure on $\mathcal{F}_T^{(m)}$ and X_t is n-dimensional Brownian motion w.r.t. Q.

Proof. It is easy to show that Q is a probability measure on $\mathcal{F}_T^{(m)}$:

$$Q(\Omega) = E_P[M_T] = E_P[M_0] = 1$$

We note that on $\mathcal{F}_t^{(m)}$, we actually have $dQ = M_t dP$. More precisely, let f be a bounded $\mathcal{F}_t^{(m)}$-measurable function. Then we get

$$E_Q[f] = E_P[fM_T] = E_P[E_P[fM_T|\mathcal{F}_t^{(m)}]] = E_P[fE_P[M_T|\mathcal{F}_t^{(m)}]] = E_P[fM_t]$$

In context of the above definition of M_t, let $dN_t = -adB_t - \frac{1}{2}a^2dt$. Applying the Itô lemma to the function $g(x) = \exp(x)$, we find that

$$\begin{aligned} d(M_t) = d(\exp(N_t)) &= \exp(N_t)dN_t + \frac{1}{2}\exp(N_t)(dN_t)^2 \\ &= M_t\left(\sum_i -a_i dB_i(t) - \frac{1}{2}\sum_i a_i^2 dt\right) + \frac{1}{2}M_t\sum_i a_i^2 dt \\ &= -M_t\left(\sum_i a_i dB_i(t)\right) \end{aligned}$$

Put $Z_t = M_t X_t$. From (2.10), we have that

$$\begin{aligned} dZ_i(t) &= M_t dX_i(t) + X_i(t)dM_t + (dM_t)(dX_i(t)) \\ &= M_t(a_i dt + dB_i(t)) + X_i(t)M_t\left(\sum_j -a_j dB_j(t)\right) + B_i(t)M_t\left(\sum_j -a_j dB_j(t)\right) \\ &= M_t\left(dB_i(t) - X_i(t)\left(\sum_j a_j dB_j(t)\right)\right) \end{aligned} \tag{6.4}$$

In matrix notation, (6.4) can be written as $M_t V_t dB_t$, where $V_t \in \mathbb{R}^n$ is defined by

$$V_j(t) = (\delta_{ij} - X_j(t)a_j(t))$$

By Theorem 2.7, we have shown that $M_t X_i(t)$ is a martingale w.r.t. $\{\mathcal{F}_t^{(m)}\}$ under the measure P. Using Bayes' Theorem, we have that for $s < t \le T$,

$$E_Q[X_i(t)|\mathcal{F}_s^{(m)}] = \frac{E_P[X_i(t)M_t|\mathcal{F}_s^{(m)}]}{E_P[M_t|\mathcal{F}_s^{(m)}]} = \frac{X_i(s)M_s}{M_s} = X_i(s)$$

So that $X_i(t)$ is a martingale w.r.t. $\{\mathcal{F}_t^{(m)}\}$ under the measure Q. The proof that $X_iX_j - \delta_{ij}t$ is also a martingale is similar. By the Lévy characterization of Brownian motion, this completes the proof. □

Exponential martingales such as M_t above are actually a commonly used tool in stochastic calculus. The following theorem is useful

Theorem 6.4 (Novikov Condition)**.** *Let M_t be defined as in the previous theorem. A sufficient condition for M_t to be a martingale w.r.t. $\{\mathcal{F}_t^{(m)}\}$ under the measure P is that*

$$E\left[\exp\left(\int_0^t a^2(s,\omega)ds\right)\right] < \infty$$

For the proof, see, for instance, Karatzas and Shreve (1991). Any continuous, deterministic function trivially satisfies the Novikov condition.

Theorem 6.5 (Girsanov II)**.** *Let $X_t \in \mathbb{R}^n$ be an Itô process*

$$dX_t = \beta(t,\omega)dt + \theta(t,\omega)dB_t$$

where $B_t \in \mathbb{R}^m$. Now suppose that we can find $\alpha(t,\omega) \in \mathcal{V}^n[0,T]$ and $u(t,\omega) \in \mathcal{V}^{n\times m}[0,T]$ such that

$$\theta(t,\omega)u(t,\omega) = \beta(t,\omega) - \alpha(t,\omega)$$

Define M_t as in the previous theorem with $u(s,\omega)$ in place of $a(s,\omega)$ and assume that this is a martingale w.r.t. $\{\mathcal{F}_t^{(m)}\}$ under the measure P. Define $dQ = M_T dP$ on $\mathcal{F}_T^{(m)}$. Then the process

$$\widehat{B_t} := \int_0^t u(s,\omega)ds + B_t \tag{6.4}$$

is standard Brownian motion w.r.t. Q, and we can write

$$dX(t) = \alpha(t,\omega)dt + \theta(t,\omega)d\widehat{B_t} \tag{6.5}$$

Proof. In view of Theorem 6.3, it suffices to prove that (6.5) holds. This is clear in view of (6.4):

$$\begin{aligned} dX(t) &= \beta(t,\omega)dt + \theta(t,\omega)dB_t \\ &= \beta(t,\omega)dt + \theta(t,\omega)(d\widehat{B_t} - u(t,\omega)dt) \\ &= \beta(t,\omega)dt + \theta(t,\omega)d\widehat{B_t} - (\alpha(t,\omega) - \beta(t,\omega))dt \\ &= \alpha(t,\omega)dt + \theta(t,\omega)d\widehat{B_t} \end{aligned}$$

□

The final version applies specifically to Itô diffusions and contains a helpful uniqueness statement:

Theorem 6.6 (Girsanov III)**.** *Let $X_t \in \mathbb{R}^n$ be an Itô diffusion and define Y_t such that*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$
$$Y_t = (b(X_t) + \gamma(t,\omega))dt + \sigma(X_t)dB_t$$

where $B_t \in \mathbb{R}^n$ and $\gamma \in \mathcal{V}^n[0,T]$. Suppose we can find $u(t,\omega) \in \mathcal{V}^m[0,T]$ such that $\sigma(Y_t)u(t,\omega) = \gamma(t,\omega)$. Define M_t, Q, and $\widehat{B_t}$ as in the last theorem, where M_t is a martingale w.r.t. $\{\mathcal{F}_T^{(m)}\}$ under the measure P. Then we get that

$$dY_t = b(Y_t)dt + \sigma(Y_t)d\widehat{B_t}$$

Also, with $x \in \mathbb{R}^n$ a starting point for both diffusions, the Q^x law of Y_t^x is equal to the P^x law of X_t^x.

Proof. This follows by applying Girsanov II with appropriate choices of α and β. The statement about equivalence of laws is an immediate consequence of weak uniqueness of solutions to SDE Lemma 3.4, which from Theorem 3.3 we can clearly apply to the above diffusions. □

The theorem essentially states that we can change the drift coefficient of an Itô diffusion and it will keep the same law up to a change of measure.

Remark 6.7. With Q, P, and M_T as above, we clearly have that $Q \ll P$. Suppose that $A \in \mathcal{F}_T^{(m)}$ and note that if $Q(A) = 0$ then $E_P[M_T \mathcal{X}_A] = 0$ so that $M_T \mathcal{X}_A = 0$ for P-a.s. ω. Since $M_T > 0$ for P-a.s. ω, we have that $\mathcal{X}_A = 0$ for P-a.s. ω. Then immediately $P(A) = 0$. This shows that $P \ll Q$ so that $P \sim Q$. Because of this, Q is sometimes called an *equivalent martingale measure*.

Example 6.8. Let

$$dX(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} dt + \begin{bmatrix} 1 & 3 \\ -1 & -2 \end{bmatrix} dB_t$$

If we set

$$\begin{bmatrix} 1 & 3 \\ -1 & -2 \end{bmatrix} u(t,\omega) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

we get that $u = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$. Since u trivially satisfies the Novikov condition, and

$$d\widehat{B_t} := \begin{bmatrix} -3 \\ -1 \end{bmatrix} + dB_t$$

is standard Brownian motion w.r.t. the measure $dQ = M_T dP$, where M_T is the exponential martingale defined as in the previous theorems. Girsanov also implies that we can rewrite $X(t)$ as

$$X(t) = \begin{bmatrix} 1 & 3 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} d\widehat{B_1}(t) \\ d\widehat{B_2}(t) \end{bmatrix}$$

7. Risk-Neutral Measure and Black-Scholes

In this section, we will apply some of the previous results to asset pricing theory and mathematical finance. We will see that stochastic processes provide a natural framework for the analysis of derivative securities. Our discussion is brief and informal. For a comprehensive introduction, see [2]. Let $(\Omega, \mathcal{N}, P)$ be a probability space. In this context, we will model risky assets as random variables on Ω. Consider the *European call*: at $t = 0$ an investor purchases the right to buy a certain security at time $T > 0$ for a specified price K. Since the buyer is not obligated to exercise this right if the price at time T if $X_T(\omega) < K$, the payoff at time T is given by:

$$\max(X_T(\omega) - K, 0) \tag{7.1}$$

We let $X(t, \omega)$ be an Itô process representing the value of the security at time t. How should we price such a contingent claim with payoff at time T given by (7.1)? One such procedure, known as *pricing by arbitrage*, takes as given a collection of primitive securities (bonds, currencies, etc...) with known price processes that can be used to price the claim. More precisely:

Definition 7.1 (Replicable Claims)**.** A contingent claim C_T with payoff at maturity given by $C_T(\omega)$ is said to be *replicable* if

(i) There exists a portfolio of primitive securities with price process $S(t, \omega)$ such that $C_T(\omega) = S(T, \omega)$ a.s.
(ii) The portfolio is *self-financing* i.e. there are no net cash infusions into the portfolio between $t = 0$ and maturity.

A simple economic argument can be used to show that in the absence of arbitrage the contingent claim's price is uniquely determined. In this situation, the claim is said to be *priced by arbitrage*. A market in which all contingent claims can be priced by arbitrage is said to be *complete*. Using this approach to price derivatives has the drawback of being computationally inefficient as a different portfolio must be constructed to price each contingent claim. Luckily, we have another method:

Definition 7.2 (Risk-Neutral Measure)**.** A measure Q on the space $(\Omega, \mathcal{N}, P)$ is said to be *Risk-Neutral* if

(i) $Q \sim P$ i.e. $Q \ll P$ and $P \ll Q$
(ii) Any price process $X(t, \omega)$ is a martingale w.r.t. its own filtration under the measure Q.

To explain the second condition, let us suppose that the market has some risk free security such as a bond. By way of normalization, we discount every price process by the bond's price process. If we wish to create a "risk-neutral" measure and avoid arbitrage opportunities, the expected change in value of every asset should be the same. Because the bond's discounted expected value change is 0, this must hold for all other assets, which is reflected in condition (ii).

Definition 7.3 (Risk-Neutral Price)**.** Let $X(t, \omega)$ be the price process of an asset in a complete, arbitrage free market. Let Q be a risk-neutral measure for the associated probability space. Given a contingent claim C_T on the security X_t we define the *risk-neutral price* $F(C_T)$ by

$$F(C_T) = E_Q\left[\frac{X_T}{V(T)}\right]$$

where $V(t)$ is the price process of a fixed risk-free asset.

Note that the existence of such a measure Q is a consequence of a result known as the *Fundamental Theorem of Asset Pricing*. This brings us to a crucial result, the proof of which can be found in any text on mathematical finance:

Theorem 7.4. *Let C_T be a replicable claim in a complete, arbitrage free market. Then the arbitrage free price (using a replicating portfolio of primitive securities) and the risk-neutral price $F(C_T)$ are the same.*

We will see that the Girsanov theorem gives us the tools to compute a claim's risk-neutral price.

Example 7.5 (Black-Scholes Formula)**.** Consider the Black-Scholes model where a certain security's price process X_t obeys geometric Brownian motion:

$$dX_t = \alpha X_t dt + \beta X_t dB_t$$

where α and β are constants. We think of αX_t as the mean rate of change in price and βX_t as the asset's uncertainty. Let $V(t)$ be the price process of a bond, which we assume is a risk-free security.

$$dV(t) = rV(t)dt$$

Then the discounted price process for X_t is $Z_t = \frac{X_t}{V_t}$. We can apply the Itô lemma to the function $g(x_1, x_2) = \frac{x_1}{x_2}$ to conclude that

$$\begin{aligned} dZ_t &= \frac{dX_t}{V_t} - \frac{dV_t}{V_t^2} X_t \\ &= \alpha Z_t dt + \beta Z_t dB_t - r Z_t dt \\ &= (\alpha - r) Z_t dt + \beta Z_t dB_t \end{aligned}$$

Let us use the Girsanov Theorem to get rid of the drift coefficient of the above diffusion. We need to find a $u(t, \omega)$ such that

$$\beta Z_t u(t, \omega) = (\alpha - r) Z_t$$

and clearly $u = \frac{\alpha - r}{\beta}$ will work. This is a constant, so it trivially satisfies the Novikov condition whence by the Girsanov Theorem we get that

$$\widehat{B_t} = \left(\frac{\alpha - r}{\beta}\right) t + B_t$$

is standard Brownian motion w.r.t. the measure Q, where $dQ = M_T dP$ and M_t is as in the Girsanov Theorem. Then the discounted price process Z_t has the representation $dZ_t = \beta Z_t d\widehat{B_t}$. Using our formula for the solution to the geometric Brownian motion SDE, this is just

$$Z_t = Z_0 \exp\left(-\frac{1}{2}\beta^2 + \beta \widehat{B_t}\right) \tag{7.2}$$

From Theorem 2.7, it is clear that Z_t is a martingale w.r.t. its own filtration under the measure Q. From Remark 6.7, we also have that $Q \sim P$. *Then we have shown that Q is a risk-neutral measure according to Definition 7.2.* By Theorem 7.4, we now know how to price replicable claims on the security X_t. We once again consider the European claim described above. According to the theorem, the expectation to be calculated is

$$\frac{1}{V_T} E_Q[\max(X_T - K, 0)] = E_Q[\max(Z_T - \exp(-rT)K, 0]$$

by evaluating the ODE for the bond price. From equation (7.2) and using the fact that $\widehat{B_t}$ is standard Brownian motion w.r.t. the measure Q, the above expectation is given by

$$\int_{\mathbb{R}} \max\left(Z_0 \exp\left(-\frac{1}{2}\beta^2 T + \beta y\right) - \exp(-rT)K, 0\right) \frac{1}{\sqrt{2\pi T}} \exp\left(\frac{-y^2}{2T}\right) dy \tag{7.3}$$

A trivial computation shows that

$$Z_0 \exp\left(-\frac{1}{2}\beta^2 T + \beta y\right) - \exp(-rT)K \geq 0 \quad \text{iff} \quad y \geq a$$

where a is defined by

$$a = \frac{1}{\beta}\left(-\ln\left(\frac{Z_0}{K}\right) + \left(\frac{1}{2}\beta^2 - r\right)T\right)$$

Thus, we may rewrite eqn. (7.3) as

$$\int_a^\infty \frac{Z_0}{\sqrt{2\pi T}} \exp\left(-\frac{1}{2}\beta^2 T + \beta y - \frac{y^2}{2T}\right) dy - \exp(-rT) \int_a^\infty \frac{K}{\sqrt{2\pi T}} \exp\left(\frac{y^2}{2T}\right) dy \qquad (7.4)$$

where we have merged the exponentials. Notice that $-\frac{1}{2T}(y-\beta T)^2 = -\frac{1}{2T}\beta^2 T+\beta y-\frac{y^2}{2T}$. Removing the constants from the integrals, the above are just integrals of the density functions for two normal random variables with distribution $N(\beta T, T)$ and $N(0, T)$ respectively. We can thus rewrite equation (7.4) in terms of the CDF of the standard normal distribution as

$$X_0\Phi(-w_1) - K\exp(-rT)\Phi(-w_2) \qquad (7.5)$$

where $w_1 = \frac{a-\beta T}{\sqrt{T}}$ and $w_2 = \frac{a}{\sqrt{T}}$, and we have used the fact that $Z_0 = X_0$, the initial price of the security. Then Theorem 7.4 implies that the arbitrage-free price of the European call described above is given by

$$F(C_T) = X_0\Phi(-w_1) - K\exp(-rT)\Phi(-w_2)$$

Acknowledgments

I would like to thank Peter May for organizing this REU, which has been both productive and enjoyable. I also want to thank my mentors Marcelo Alvisio, Andrew Lawrie, and Casey Rodriguez for their enthusiasm to meet with me every week and for the mathematical insights they shared with me during the course of the REU.

References

[1] Kai Lai Chung, *Lectures from Markov Processes to Brownian Motion*, Springer, 1982
[2] Darrell Duffie *Dynamic Asset Pricing Theory* Princeton University Press 2001
[3] Gerald B. Folland, *Real Analysis: Modern Techniques and Their Applications*, Wiley-Interscience, Second Edition, 1999
[4] Bernt Øksendal, *Stochastic Differential Equations*, Wiley, Sixth Edition, 2007
[5] Peter Mörters and Yuval Peres, *Brownian Motion*, Cambridge University Press, 2010
[6] Rangarajan Sundaram "Equivalent Martingale Measures and Risk-Neutral Pricing: An Expository Note" *Journal of Derivatives* 1997
[7] David Williams, *Probability with Martingales*, Cambridge University Press, 1991