

STOCHASTIC PROCESSES AND THE FEYNMAN-KAC THEOREM

FORREST FLESHER

1. INTRODUCTION

In this paper, we'll prove the Feynman-Kac theorem, a result relating probability theory and analysis. It has numerous applications in fields such as physics, statistics, finance, chemistry, and others, and provides an interesting case study in the connections between solutions to elliptic and parabolic differential equations and stochastic processes. In particular, it allows one to represent the solution to a partial differential equation as an expectation of a stochastic functional. In a way, this means some deterministic partial differential equations can be seen as a type of “mean” of a stochastic differential equation. Before proving the theorem however, we'll need some results from the theory of stochastic processes, namely the construction of the stochastic integral (the Itô integral), and Itô's formula.

2. DEFINITIONS FROM PROBABILITY THEORY

In probability theory, randomness is modeled by a *probability space* (Ω, \mathcal{F}, P) , a σ -finite measure space with $P(\Omega) = 1$. Here Ω is called the *sample space*, \mathcal{F} is called the *event space*, and P is called the *probability measure*. A *random variable* is a measurable function $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, where $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ -algebra on \mathbb{R}^d . It assigns to each sample an element of Euclidean space, and to each event a Borel-subset of Euclidean space. For a random variable X on a space (Ω, \mathcal{F}, P) , the *expected value* of X is defined as

$$E(X) = \int_{\Omega} X(\omega) dP(\omega).$$

Definition 1 (Stochastic Process). A *stochastic process* is a collection of random variables $\{X_t : t \in I\}$ (typically t is thought of as time). A *sample path* is a map $t \mapsto X_t(\omega)$, where $\omega \in \Omega$ is fixed. A *filtration* is a nondecreasing family $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F} : 0 \leq s < t < \infty$, of sub σ -fields of \mathcal{F} . A stochastic process X_t is *\mathcal{F} -measurable* if the function

$$\begin{aligned} ([0, \infty) \times \Omega, \mathcal{B}([0, \infty) \otimes \mathcal{F}) &\rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \\ (t, \omega) &\mapsto X_t(\omega) \end{aligned}$$

is measurable. A process is *adapted* to a filtration $\{\mathcal{F}_t\}$ if for each t , X_t is an \mathcal{F}_t -measurable random variable.

Throughout this paper, we assume that all filtrations \mathcal{F}_t are right-continuous and contain all P -negligible events in \mathcal{F} . This is for technical purposes, but intuitively this means that for infinitesimal steps forward in time, no additional information is gained - the stream of information into the system is continuous.

Definition 2 (Martingale). A *martingale* is an adapted process X such that for all $0 \leq s < t < \infty$, $E(X_t | \mathcal{F}_s) = X_s$ (both the left and right side here are random variables). That is, the expected value given information up to time s is the process at time s . A right-continuous martingale X is *square-integrable* if $E(X_t^2) < \infty$ for all $t \geq 0$. If $X_0 = 0$, we say $X \in \mathcal{M}_2$. The *norm* of a square-integrable martingale is

$$\begin{aligned} \|X\|_t &= \sqrt{E(X_t^2)}, \\ \|X\| &= \sum_{n=1}^{\infty} \frac{\|X\|_n \wedge 1}{2^n}. \end{aligned}$$

The metric space of all square-integrable martingales X with $X_0 = 0$ induced by the norm $\|\cdot\|$ is denoted \mathcal{M}_2 , and the subspace of continuous square-integrable martingales is denoted \mathcal{M}_2^c . Note that the norm $\|X\|_t$ is an ordinary L^2 -norm on $(\Omega, \mathcal{F}_t, P)$.

Lemma 3. *For a right-continuous martingale $\{X_t, \mathcal{F}_t\}$, an interval $[0, T] \subset \mathbb{R}^+$, and $0 < \lambda \in \mathbb{R}$, we have*

$$P\left(\sup_{0 \leq t \leq T} X_t \geq \lambda\right) \leq \frac{E(\max(X_T))}{\lambda}.$$

Proof. See [5], theorem 1.3.8. □

Proposition 4. *The space \mathcal{M}_2 defined above is a complete metric space, and the subset \mathcal{M}_2^c of continuous martingales is closed.*

Proof. Take a Cauchy sequence $\{X^{(n)}\}$ in \mathcal{M}_2 . Observe that for fixed t , by definition of the \mathcal{M}_2 norm $\|\cdot\|$, the sequence $\{X_t^{(n)}\}$ is Cauchy also for $\|\cdot\|_t$. Since $\|X\|_t$ is an ordinary L^2 -norm, the Cauchy sequence $\{X_t^{(n)}\}$ has a limit X_t in $L^2(\Omega, \mathcal{F}_t, P)$.

Now, take $A \in \mathcal{F}_s$. By Cauchy-Schwarz, we have $E(\mathbb{I}_A \cdot (X_s^{(n)} - X_s)) \rightarrow 0$ and $E(\mathbb{I}_A(X_t^{(n)} - X_t)) \rightarrow 0$. Since $X^{(n)}$ is a martingale, $E(\mathbb{I}_A X_t^{(n)}) = E(\mathbb{I}_A X_s^{(n)})$ for each n , and thus $E(\mathbb{I}_A X_t) = E(\mathbb{I}_A X_s)$. So X is a martingale, and choosing a right-continuous representative, $X \in \mathcal{M}_2$. This shows that \mathcal{M} is complete.

Now, let $\{X^{(n)}\}$ be a sequence in \mathcal{M}_2^c . It has a limit X in \mathcal{M}_2 by the above. For $\epsilon > 0$, lemma 3 implies

$$P\left(\sup_{0 \leq t \leq T} |X_t^{(n)} - X_t| \geq \epsilon\right) \leq \frac{1}{\epsilon^2} \|X^{(n)} - X\|_T^2,$$

which converges to 0 as $n \rightarrow \infty$. Then for some subsequence (n_k) ,

$$P\left(\sup_{0 \leq t \leq T} |X_t^{(n_k)} - X_t| \geq \frac{1}{k}\right) \leq \frac{1}{2^k}.$$

Thus

$$\sum_{k=1}^{\infty} P\left(\sup_{0 \leq t \leq T} |X_t^{(n_k)} - X_t| \geq \frac{1}{k}\right) < \infty,$$

and by the Borel-Cantelli lemma there exists an N_k such that

$$P\left(\sup_{0 \leq t \leq T} |X_t^{n_k} - X_t| \geq \frac{1}{k}\right) = 0$$

for all $n_k \geq N_k$. That is, $X_t^{n_k}$ converges uniformly almost surely to X_t on $[0, T]$. Since the $X_t^{(n_k)}$ are continuous, then so is X , which proves that \mathcal{M}_2^c is closed. □

Definition 5 (Stopping Time). For a measurable space (Ω, \mathcal{F}) with a filtration \mathcal{F}_t , a *stopping time* T is an \mathcal{F} -measurable random variable taking values in $[0, \infty]$ such that for all $t \in \mathbb{R}^+$, the event $\{T \leq t\}$ belongs to \mathcal{F}_t . Intuitively, a time is a stopping time if given information up to time t , we can determine whether the stopping has occurred or not. For a process X and a stopping time T , X_T is defined as $X_T(\omega) = X_{T(\omega)}(\omega)$.

Definition 6 (Quadratic Variation). For a square integrable martingale X , the *quadratic variation* of X is the unique martingale $\langle X \rangle$ such that $X^2 - \langle X \rangle$ is a martingale. Alternatively let $0 = t_0 \leq t_1 \leq \dots \leq t_m = t$ be a partition of $[0, t]$. Then the quadratic variation is the limit of

$$\sum_{k=1}^m (X_{t_k} - X_{t_{k-1}})^2$$

as $\max_{1 \leq k \leq m} |t_k - t_{k-1}| \rightarrow 0$.

Definition 7 (Covariation). For square integrable processes X and Y , the *covariation* of X and Y is the process

$$\langle X, Y \rangle = \frac{1}{2} (\langle X + Y \rangle - \langle X \rangle - \langle Y \rangle).$$

Alternatively, it is the limit of

$$\sum_{k=1}^m (X_{t_k} - X_{t_{k-1}})(Y_{t_k} - Y_{t_{k-1}})$$

as $\max_{1 \leq k \leq m} |t_k - t_{k-1}| \rightarrow 0$.

Example 8. Consider the following scenario: we start at 0 and at each of 5 time steps, either add one or subtract one. In this case, the sample space Ω is the space of paths starting at the origin and moving up, down, or staying the same with equal probability for 5 steps; the event space \mathcal{F} is the set of subsets of these paths; the probability measure P takes each event $E \in \mathcal{F}$ and assigns it a number between 0 and 1; a filtration \mathcal{F}_t might be the set of $E \in \mathcal{F}$ such that all paths in E pass through the same point at time t , along with complements and unions; an example of a martingale is a function $X_t : \Omega \rightarrow \mathbb{R}$ which takes a path to its value at time t .

While the above example is very useful for intuition about the definitions and interpretations of the concepts of probability theory, we will mostly be dealing with continuous time processes. Our next example is that of Brownian motion, one of the most important continuous time stochastic processes.

Example 9 (Brownian Motion). A one dimensional *Brownian motion* (also called a *Wiener process*) is a continuous, adapted process $\{W_t, \mathcal{F}_t, t \in \mathbb{R}^+\}$ which satisfies the following properties:

- (1) $W_0 = 0$ almost surely;
- (2) W has independent increments: for all $0 \leq s < t$, the increment $W_t - W_s$ is independent of \mathcal{F}_s ;
- (3) W has normal increments: for all $0 \leq s < t$, the increment $W_t - W_s$ is normally distributed with mean 0 and variance $t - s$.

After reading the above definition, the reader may have ask: does a Brownian motion exist? The answer is yes, but the construction is non-trivial, and we do not prove it here. For the construction of Brownian motion, see [4], chapter 2.

A Brownian motion W is square integrable and satisfies $\langle W \rangle_t$ for $t \geq 0$. Such a process can often be thought of a “scaling limit” of a random walk: for $t \in [0, 1]$ and i.i.d. random variables ξ_1, ξ_2, \dots with mean 0 and variance 1, define

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{1 \leq k \leq \lfloor nt \rfloor} \xi_k.$$

Then in the limit $n \rightarrow \infty$, W_n approaches a Brownian motion, a result known as Donsker’s theorem. For a proof, see [5], theorem 2.4.20.

3. CONSTRUCTION OF A STOCHASTIC INTEGRAL

Our goal is to prove results about partial differential equations, their solutions, and stochastic processes. In order to do this, we will need a calculus of stochastic processes. We begin by constructing a stochastic integral: the Itô integral. To understand why such a construction is necessary, recall the conditions of the Lebesgue-Stieltjes integral. We can integrate $\int_a^b f(x)dg(x)$ when $f : [a, b] \rightarrow \mathbb{R}$ is a bounded Borel measurable function, and $g : [a, b] \rightarrow \mathbb{R}$ is a right continuous function of *bounded variation*. Unfortunately, martingales are almost never of bounded variation (only constant processes are of bounded variation), so defining an integral $I_T(X) = \int_0^T X_t(\omega)dM_t(\omega)$ pathwise in the Lebesgue-Stieltjes sense will not work.

However, our hope is that because continuous, square integrable martingales have finite *quadratic variation*, then we will be able to construct a suitable integral. Since the quadratic variation is finite, our starting

point is the integral $\int_0^T X_t(\omega) d\langle M \rangle_t(\omega)$, where $\langle M \rangle_t$ is the quadratic variation process defined above. We define a measure μ_M on the combined time-sample space $([0, \infty) \times \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F})$ by

$$\mu_M(A) := \int_0^\infty \chi_A((t, \omega)) d\langle M \rangle_t(\omega).$$

We then define an equivalence class of measurable \mathcal{F}_t -adapted processes by almost-everywhere equivalence in the measure μ_M . From now on we will refer to equivalence classes by their representatives.

Let \mathcal{L} be the set of measurable, \mathcal{F}_t -adapted processes such that

$$[X]_T^2 := E \left(\int_0^\infty X_t^2 d\langle M \rangle_t \right) < \infty$$

for all $T > 0$. So $[X]_T$ is an ordinary L^2 norm on the space $[0, T] \times \Omega$ under the measure μ_M . We also define more generally

$$[X] := \sum_{n=1}^\infty \frac{\min(1, [X]_n)}{2^n},$$

which gives a metric on \mathcal{L} by $[X - Y]$.

Thus far, the only assumptions we have made on M are square integrability and continuity. We now additionally make the assumption that $t \mapsto \langle M \rangle_t(\omega)$ is *absolutely continuous* (w.r.t. the Lebesgue measure) for almost all ω . While one can define an integral without this assumption, this will allow us to integrate the widest class of integrands, namely all of \mathcal{L} as defined above - without this assumption, we can integrate only the narrower class of *progressively* measurable processes inside \mathcal{L} . Since we'll be primarily concerned with the case of M being Brownian motion, which has $\langle B \rangle_t = t$, this is enough for our purposes.

The construction of the stochastic integral proceeds along similar lines as that of the Lebesgue-Stieltjes integral. We begin by integrating *simple* processes, prove that such processes are dense, and then define the integral of a general process as a limit of simple processes.

Definition 10 (Simple Process). A process H is called *simple* if it is of the form

$$H_t(\omega) = \xi_0(\omega)\chi_{\{0\}}(t) + \sum_{i=0}^\infty \xi_i(\omega)\chi_{(t_i, t_{i+1}]}(t),$$

where $\{t_n\}_{n=0}^\infty$ is a strictly increasing sequence with $t_0 = 0$, $t_n \rightarrow \infty$, and $\{\xi_n\}_{n=0}^\infty$ is a sequence of random variables such that there exists a constant C with $\sup_{n \geq 0} |\xi_n(\omega)| \leq C$. The class of simple process is denoted \mathcal{L}_0 .

Proposition 11. *For each continuous, square integrable martingale M with $t \mapsto \langle M \rangle_t(\omega)$ absolutely continuous for almost all ω , the space \mathcal{L}_0 of simple processes is dense in \mathcal{L} with respect to the metric $[\cdot]$, defined above.*

Proof. See [5], proposition 3.2.8. □

We now proceed with the definition of the stochastic integral for simple processes, and then extend to arbitrary \mathcal{L} processes.

For $H \in \mathcal{L}_0$ of the form

$$H_t(\omega) = \xi_0(\omega)\chi_{\{0\}}(t) + \sum_{i=1}^\infty \xi_i(\omega)\chi_{(t_i, t_{i+1}]}(t),$$

the stochastic integral is defined as the process

$$I_t(X) = \sum_{i=0}^\infty \xi_i(M_{t \wedge t_{i+1}} - M_{t \wedge t_i}).$$

Since $\lim_{n \rightarrow \infty} t_n = \infty$, and since t is a finite real number, this is a finite sum. Here we've used the notation \wedge , which we use throughout to mean $a \wedge b = \min(a, b)$.

Proposition 12. *If G, H are simple processes, then the following hold:*

- (1) $I_0(G) = 0$ almost surely in P ;
- (2) $E(I_t(G) \mid \mathcal{F}_s) = I_s(G)$ almost surely in P , i.e. $I_t(G)$ is a martingale;
- (3) $E(I_t(G))^2 = E \int_0^t X_r^2 d\langle M \rangle_r$;
- (4) $\|I(G)\| = [G]$.
- (5)

$$E((I_t(G) - I_s(G))^2 \mid \mathcal{F}_s) = E \left(\int_s^t X_r^2 d\langle M \rangle_r \mid \mathcal{F}_s \right),$$

almost surely in P ;

- (6) $I(aG + bH) = aI(G) + bI(H)$.

Proof. See [5], section 3.2.B. □

Given $X \in \mathcal{L}$, proposition 11 gives a sequence of simple processes $H^{(n)}$ with $[H^{(n)} - X] \rightarrow 0$ as $n \rightarrow \infty$. Then $[H^{(n)} - H^{(m)}] \rightarrow 0$ so $(H^{(n)})_n$ is Cauchy, and by above property (4), so is $(I(H^{(n)}))_n$. Since the space of continuous martingales is complete and closed, there is some process $I(X)$ such that $\|I(H^{(n)}) - I(X)\| \rightarrow 0$ as $n \rightarrow \infty$. Furthermore, if $H^{(n)}$ and $G^{(n)}$ satisfy $[H^{(n)} - X] \rightarrow 0$ and $[G^{(n)} - X] \rightarrow 0$, then so does the sequence

$$F^{(n)} := \begin{cases} H^{(n)} & n \text{ odd} \\ G^{(n)} & n \text{ even.} \end{cases}$$

Then the limit of $I(F^{(n)})$ is equal to that of both $I(G^{(n)})$ and $I(H^{(n)})$. So $I(X)$ does not depend on the choice of the sequence of simple functions converging to it, and is therefore well defined. This allows us to define the integral of X .

Definition 13. For $X \in \mathcal{L}$, the *stochastic integral*, or *Itô integral*, of X with respect to M is the unique process $I(X)$ such that $\|I(X^{(n)}) - I(X)\| \rightarrow 0$ as $n \rightarrow \infty$ for all sequences of simple functions $X^{(n)}$ with $[X^{(n)} - X] \rightarrow 0$. The integral is written

$$I_t(X) = \int_0^t X_s dM_s.$$

Proposition 14. The Itô integral satisfies the properties in proposition 12.

Definition 15 (Continuous local martingale). A continuous adapted process M is said to be a *continuous local martingale* if there exists a non decreasing sequence of stopping times (T_n) with $P(T_n \rightarrow \infty) = 1$ such that $\{M_{t \wedge T_n}; \mathcal{F}_t\}$ is a martingale for each n .

For the local case, we will be able to integrate processes satisfying a weaker condition than square integrability, namely the property that

$$P \left(\int_0^T X_t^2 d\langle M \rangle_t < \infty \right) = 1$$

for all $T \in [0, \infty)$. Recall before that we required this integral to have finite *expectation*, while now we require only that it is finite with probability 1.

Proposition 16. A bounded local martingale is a martingale.

Proof. Let T_n be the sequence of stopping times for a bounded local martingale M . Then $M_{t \wedge T_n}$ is a martingale for each n . Since M is bounded, we apply the dominated convergence theorem to get $M_s = E(M_t \mid \mathcal{F}_s)$. □

For a continuous local martingale, there is a nondecreasing sequence of stopping times $(T_n)_n$ satisfying the conditions in the definition. There also exists a sequence of stopping times $(S_n)_n$ defined by

$$S_n(\omega) = \min \left(n, \inf \left\{ t : \int_0^t X_s^2(\omega) d\langle M \rangle_s(\omega) \geq n \right\} \right).$$

This sequence is nondecreasing and satisfies $S_n \rightarrow \infty$ almost surely in P . Let $R_n = T_n \wedge S_n$, and define $M^{(n)}$ as the process M stopped at R_n , namely $M_t^{(n)} = M_{t \wedge R_n}$, and let $X^{(n)}$ be the process $X_t^{(n)} = X_t \cdot \mathbb{I}_{t \leq T_n}$. Then $M^{(n)}$ is a continuous, square integrable *ordinary* martingale, and $X^{(n)}$ is in $\mathcal{L}(M^{(n)})$. We can then take the stochastic integral $I_{t, M^{(n)}}(X^{(n)})$ as above, and we define the integral $I_t(X)$ as $I_{t, M^{(n)}}(X^{(n)})$ for $0 \leq t \leq R_n$.

4. ITÔ'S LEMMA

Definition 17. A *continuous semimartingale* is an adapted process of the form

$$X_t = X_0 + M_t + A_t,$$

where M is a local continuous martingale, and A_t is a finite variation process.

We now state Itô's lemma, the “fundamental theorem” of stochastic calculus.

Theorem 18 (Itô's Lemma). *Let X^1, \dots, X^d be d continuous semimartingales, and let $f : \mathbb{R}^+ \times \mathbb{R}^d$ be a $C^{1,2}$ function (C^1 in the first argument and C^2 in the second). Then, almost surely in P ,*

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \frac{\partial f}{\partial t}(s, X_s) ds + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(s, X_s^1, \dots, X_s^d) dX_s^i \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(s, X_s^1, \dots, X_s^d) d\langle X^i, X^j \rangle_s. \end{aligned}$$

The one-dimensional version of this, namely that for $f : \mathbb{R} \rightarrow \mathbb{R}$ of class $C^{1,2}$, and an adapted process $X_t = X_0 + M_t + A_t$, then

$$f(t, X_t) = f(0, X_0) + \int_0^t \frac{\partial f}{\partial t}(s, X_s) ds + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d\langle M \rangle_s. \quad (1)$$

This is often written in the more convenient differential form as

$$df(t, X_t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial X_t} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial X_t^2} d\langle M \rangle_t.$$

The last term is what distinguishes stochastic calculus from ordinary calculus. The term arises due to the nonzero quadratic variation - one cannot “ignore” as many higher order terms in the Taylor expansion as one often does in ordinary calculus.

Rather than prove Itô's lemma (the proof is long and somewhat technical), we will give a heuristic proof sketch of the lemma for the important special case of Itô processes, which are processes satisfying the stochastic differential equation

$$dX_t = \mu_t dt + \sigma_t dW_t,$$

where W_t is a Brownian motion, and μ and σ are \mathcal{F}_t -adapted processes. In this case, the lemma states that

$$df(X_t, t) = \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial X_t} + \frac{1}{2} \sigma_t^2 \frac{\partial^2 f}{\partial X_t^2} \right) dt + \left(\sigma_t \frac{\partial f}{\partial X_t} \right) dW_t.$$

Itô processes are continuous semimartingales, and are useful in various applications, such as option pricing in finance or drift diffusion problems in physics.

The proof sketch we give here contains a trick is often useful in stochastic calculus: the multiplication rules

$$\begin{aligned} (dt)^2 &= 0 \\ (dW_t)(dt) &= 0 \\ (dW_t)^2 &= dt. \end{aligned}$$

Proof. Expand the Taylor series for $f(X_t, t)$ this function to second order:

$$\begin{aligned} df(X_t, t) &= f(X_{t+dt}, t+dt) - f(X_t, t) \\ &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial X_t} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} (dt)^2 + \frac{1}{2} \frac{\partial^2 f}{\partial X_t^2} (dX_t)^2 + \frac{1}{2} \frac{\partial^2 f}{\partial t \partial X_t} (dX_t)(dt). \end{aligned}$$

Using the stochastic calculus multiplication rules mentioned above and expanding X_t ,

$$\begin{aligned} (dX_t)^2 &= \mu_t(dt)^2 + 2\mu_t\sigma_t(dt)(dW_t) + \sigma_t^2(dW_t)^2 \\ &= \sigma_t^2 dt, \\ (dX_t)(dt) &= \mu_t(dt)^2 + \sigma_t(dW_t)(dt) \\ &= 0 \end{aligned}$$

and so

$$\begin{aligned} df(X_t, t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial X_t} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial X_t^2} \sigma_t^2 dt \\ &= \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial X_t} + \frac{1}{2} \sigma_t^2 \frac{\partial^2 f}{\partial X_t^2} \right) dt + \left(\sigma_t \frac{\partial f}{\partial X_t} \right) dW_t, \end{aligned}$$

as desired. \square

Corollary 19 (Itô's product rule). *Suppose X_t and Y_t are two continuous semimartingales, with*

$$\begin{aligned} X_t &= X_0 + M_t + A_t \\ Y_t &= Y_0 + N_t + B_t. \end{aligned}$$

Then

$$dX_t dY_t = X_t dY_t + Y_t dX_t + d\langle M, N \rangle_t.$$

This is also known as Itô's integration by parts formula.

5. THE FEYNMAN-KAC FORMULA

We'll need the following two short lemmas in the proof of the Feynman-Kac theorem.

Lemma 20. *For $x \in (0, \infty)$, then*

$$\int_x^\infty e^{-u^2/2} du \leq \frac{1}{x} e^{-x^2/2} \quad (2)$$

For $a \in \mathbb{R}$, this implies

$$P^x(B_t \geq a) = \sqrt{\frac{t}{2\pi}} \frac{1}{a-x} e^{-(a-x)^2/2t},$$

where P^x represents the probability measure for a Brownian motion centered (started) at x .

Proof. To prove (2), note that $t > x$ in the region of integration, so

$$\int_x^\infty e^{-u^2/2} du \leq \int_x^\infty \frac{u}{x} e^{-u^2/2} du = \frac{1}{x} e^{-x^2/2}. \quad (3)$$

The second part of the lemma follows from substitution and the fact that the distribution of B_t at time t is $\mathcal{N}(x, t)$. \square

Lemma 21. *Let B_t be a Brownian motion and $b \in \mathbb{R}$, and let T_b be the first time that B_t hits b . Then $P^0(T_b < t) = 2P^0(B_t < b)$.*

Proof. If B_t has hit b by time t , then it lies on a path which passes through b . The paths which pass through b before time t are the paths such that $B_t > b$, and the paths such that $B_t < b$ which are *reflections* over the axis b . That is, for each path with $B_t > b$, there are exactly two paths lying on paths which hit b before time t . (include picture). Thus, the probability that the event $T_b < t$ occurs is twice the probability that a path lies above b at time t , or $P^0(B_t > b)$. \square

The version of the Feynman-Kac theorem we state here is a result about *representations* of solutions to PDEs. That is, we do not prove here the *existence* of solutions, rather we suppose that solutions exist and then show they have a certain form. The representation in the theorem does imply that solutions are unique.

Theorem 22 (Feynman-Kac). *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$, $k : \mathbb{R}^d \rightarrow \mathbb{R}^+$, and $g : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ be continuous functions. Let $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ be a function which is $C^{1,2}$ on $[0, T] \times \mathbb{R}^d$, such that*

$$-\frac{\partial v}{\partial t} + kv = \frac{1}{2}\Delta v + g; \quad \text{on } [0, T] \times \mathbb{R}^d, \quad (4)$$

$$v(T, x) = f(x); \quad x \in \mathbb{R}^d. \quad (5)$$

Furthermore, suppose that

$$\max_{0 \leq t \leq T} |v(t, x)| + \max_{0 \leq t \leq T} |g(t, x)| \leq Ke^{a|x|^2} \quad (6)$$

for all $x \in \mathbb{R}^d$ and some $K > 0$ and $0 < a < \frac{1}{2dT}$. Then we can write v as

$$v(t, x) = E^x \left(f(W_{T-t}) e^{-\int_0^{T-t} k(W_s) ds} + \int_0^{T-t} g(t + \theta, W_\theta) e^{-\int_0^\theta k(W_s) ds} d\theta \right).$$

To a reader who has studied physics, this might seem almost familiar: set $g = 0$, and recall that expectation values are integrals over state space, which here is a space of *paths*. Viewing this as a sort of “integral over paths” might also allow the reader to guess why the name Feynman appears in the name of this theorem. Indeed, the theorem was proved by Feynman and Kac in attempt to rigorously justify Feynman’s quantum mechanical path integrals. Why does quantum mechanics show up here? Notice that if we replace our time t by the imaginary time $-it$. For more on the physics of this theorem, the interested reader can see [2, 6].

The Feynman-Kac formula also has numerous applications in quantitative finance, for example in deducing the existence of a risk-neutral measure from the Black-Scholes equation [1]. For more on stochastic calculus methods in quantitative finance see [1, 3, 5].

The proof involves a number of calculations, but generally proceeds in three steps: first we apply Itô’s lemma to a suitable process, then integrate up to a stopping time, and finally take expectations.

Proof. Suppose that a function v satisfies (4). By Itô’s lemma, fixing a value t and introducing a variable θ , we have

$$\begin{aligned} dv(t + \theta, W_\theta) &= \left(\frac{\partial v}{\partial \theta} + \frac{1}{2}\Delta v \right) d\theta + \sum_{i=1}^d \frac{\partial v}{\partial x_i} dW_\theta^{(i)} \\ &= (k(W_\theta)v(t + \theta, W_\theta) - g(t + \theta, W_\theta))d\theta + \sum_{i=1}^d \frac{\partial v}{\partial x_i}(t + \theta, W_\theta) dW_\theta^{(i)}. \end{aligned}$$

Now consider the process $Y(\theta) = v(t + \theta, W_\theta) e^{-\int_0^\theta k(W_s) ds}$. By Itô’s product rule, we have (since the exponential part of the process is deterministic)

$$\begin{aligned} dY &= dv(t + \theta, W_\theta) e^{-\int_0^\theta k(W_s) ds} + v(t + \theta, W_\theta) \frac{\partial}{\partial \theta} \left(e^{-\int_0^\theta k(W_s) ds} \right) \\ &= \left(k(W_s)v(t + \theta, W_\theta)d\theta - g(t + \theta, W_\theta)d\theta + \sum_{i=1}^d \frac{\partial}{\partial x_i} v(t + \theta, W_\theta) dW_\theta^{(i)} \right) e^{-\int_0^\theta k(W_s) ds} + \\ &\quad + v(t + \theta, W_\theta) \cdot (-k(W_\theta)d\theta) e^{-\int_0^\theta k(W_s) ds} \\ &= \left(-g(t + \theta, W_\theta)d\theta + \sum_{i=1}^d \frac{\partial}{\partial x_i} v(t + \theta, W_\theta) dW_\theta^{(i)} \right) e^{-\int_0^\theta k(W_s) ds}. \end{aligned}$$

Next, define the stopping time $S_n = \inf\{t \geq 0 : \|W_t\| \geq n\sqrt{d}\}$, and let $0 < r < T - t$. We'll eventually pass to the limit $n \rightarrow \infty$ and $r \rightarrow T - t$. Integrating dY on $[0, r \wedge S_n]$ gives

$$\begin{aligned} Y(r \wedge S_n) - Y(0) &= \int_0^{r \wedge S_n} \left(-g(t + \theta, W_\theta) d\theta + \sum_{i=1}^n \frac{\partial}{\partial x_i} v(t + \theta, W_\theta) dW_\theta^{(i)} \right) e^{-\int_0^\theta k(W_s) ds} \\ &= - \int_0^{r \wedge S_n} g(t + \theta, W_\theta) e^{-\int_0^\theta k(W_s) ds} d\theta \\ &\quad + \int_0^{r \wedge S_n} \left(\sum_{i=1}^d \frac{\partial}{\partial x_i} v(t + \theta, W_\theta) \right) e^{-\int_0^\theta k(W_s) ds} dW_\theta \end{aligned}$$

Taking expectation values, the stochastic integral is zero - the expectation of Brownian motion given fixed information is constant. Thus

$$E^x(Y(r \wedge S_n) - Y(0)) = - \int_0^{r \wedge S_n} g(t + \theta, W_\theta) e^{-\int_0^\theta k(W_s) ds} d\theta. \quad (7)$$

We can also explicitly evaluate $Y(r \wedge S_n)$ and $Y(0)$, which are

$$\begin{aligned} Y(r \wedge S_n) &= v(t + r \wedge S_n, W_{r \wedge S_n}) e^{-\int_0^{r \wedge S_n} k(W_s) ds}, \\ Y(0) &= v(t, W_0), \\ E^x(Y(0)) &= v(t, x), \end{aligned} \quad (8)$$

so we can write the expectation of $Y(r \wedge S_n)$ as

$$\begin{aligned} E^x(Y(r \wedge S_n)) &= E^x \left(\mathbb{I}_{r < S_n} v(t + r, W_r) e^{-\int_0^r k(W_s) ds} \right) + \\ &\quad + E^x \left(\mathbb{I}_{S_n < r} v(t + S_n, W_{S_n}) e^{-\int_0^{S_n} k(W_s) ds} \right). \end{aligned} \quad (9)$$

Combining (7), (8) (8), we have

$$v(t, x) = E^x \left(\int_0^{r \wedge S_n} g(t + \theta, W_\theta) e^{-\int_0^\theta k(W_s) ds} d\theta \right) \quad (10)$$

$$+ E^x \left(\mathbb{I}_{r < S_n} v(t + r, W_r) e^{-\int_0^r k(W_s) ds} \right) \quad (11)$$

$$+ E^x \left(\mathbb{I}_{S_n < r} v(t + S_n, W_{S_n}) e^{-\int_0^{S_n} k(W_s) ds} \right). \quad (12)$$

We now pass to the limit $n \rightarrow \infty$ and $r \rightarrow T - t$. Since $k \geq 0$, the first term (10) is bounded in absolute value by

$$\int_0^{T-t} |g(t + \theta, W_\theta)| d\theta,$$

which by our assumption (6) along with dominated convergence becomes

$$E^x \left(\int_0^{T-t} g(t + \theta, W_\theta) e^{-\int_0^\theta k(W_s) ds} d\theta \right).$$

Similarly, the second term (11) converges to

$$E^x \left(v(T, W_{T-t}) e^{-\int_0^{T-t} k(W_s) ds} \right) = E^x \left(f(W_{T-t}) e^{-\int_0^{T-t} k(W_s) ds} \right).$$

To finish the proof, it suffices to show that the third term (12) is zero. Since $k \geq 0$ and $r < T - t$, the term is less than $E^x(|v(t + S_n, W_{S_n})| \cdot \mathbb{I}_{S_n \leq T-t})$. Using our assumption (6) and the definition of S_n ,

$|v(t + S_n, W_{S_n})| \leq K e^{adn^2}$. Applying this bound and using the definitions of expectation value and Brownian motion, we have

$$\begin{aligned}
 E^x(|v(t + S_n, W_{S_n})| \cdot \mathbb{I}_{S_n \leq T-t}) &\leq K e^{adn^2} P^x(S_n \leq T) \\
 &= K e^{adn^2} P^x\left(\max_{0 \leq t \leq T} \|W_t\| \geq n\sqrt{d}\right) \\
 &= K e^{adn^2} P^x\left(\max_{0 \leq t \leq T} \left(|W_t^{(i)}|^2 + \dots + |W_t^{(n)}|^2\right)^{1/2} \geq n\sqrt{d}\right) \\
 &= K e^{adn^2} P^x\left(\max_{0 \leq t \leq T} \sum_{i=1}^d |W_t^{(i)}|^2 \geq n^2 d\right) \\
 &\leq K e^{adn^2} \sum_{i=1}^d P^x\left(\max_{0 \leq t \leq T} |W_t^{(i)}| \geq n\right)
 \end{aligned}$$

By lemma 21,

$$K e^{adn^2} \sum_{i=1}^d P^x\left(\max_{0 \leq t \leq T} |W_t^{(i)}| \geq n\right) \leq 2K e^{an^2d} \sum_{i=1}^d \left(P^x(W_t^{(i)} \geq n) + P^x(W_T^{(i)} \leq -n)\right).$$

But by lemma 20,

$$e^{an^2d} P^x(W_T \geq n) \leq e^{an^2d} \sqrt{\frac{T}{2\pi}} \frac{1}{n-x} e^{-(n-x)^2/2T}.$$

In the limit $n \rightarrow \infty$, this converges to 0. Thus the third term (12) converges to 0 in the limit $n \rightarrow \infty$.

Combining these results, we have

$$v(t, x) = E^x \left(f(W_{T-t}) e^{-\int_0^{T-t} k(W_s) ds} + \int_0^{T-t} g(t + \theta, W_\theta) e^{-\int_0^\theta k(W_s) ds} d\theta \right),$$

which finishes the proof. □

REFERENCES

- [1] Tomas Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press, 1998.
- [2] Brian C. Hall. *Quantum Theory for Mathematicians*. Springer, New York, NY, 2013.
- [3] Mark S. Joshi. *The Concepts and Practice of Mathematical Finance*. Cambridge University Press, 2003.
- [4] Jean Francois Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*. Springer, Cham, 2016.
- [5] Steven Shreve and Ioannis Karatzas. *Brownian Motion and Stochastic Calculus*. Springer, New York, NY, 1998.
- [6] Barry Simon. *Functional Integration and Quantum Physics*. AMS Chelsea Publishing, 2005.