Partial Differential Equations in Mathematical Finance *

M. R. Grasselli
Dept. of Mathematics and Statistics
McMaster University
Hamilton,ON, L8S 4K1

November 19, 2003

1 The Feynman-Kac formula

Suppose we want to solve the Cauchy problem associated with the heat equation in n dimensions:

$$\begin{cases} u_t - \frac{1}{2}\Delta u = 0 & \text{on} \quad (0,\infty)\times \mathbb{R}^n \\ u(0,x) = f(x) & \text{on} \quad \{t=0\}\times \mathbb{R}^n. \end{cases} \tag{1}$$

Then for bounded initial data $f \in C_b(\mathbb{R}^n)$, the bounded solutions of (1) are known to be given by

$$u(t,x) = \int_{\mathbb{R}^n} p(t,x,y)f(y)dy, \tag{2}$$

where

$$p(t,x,y) = \frac{1}{\sqrt{2\pi t}^n} e^{-\frac{|x-y|^2}{2t}} \tag{3}$$

is the n–dimensional heat kernel.

A probabilistic interpretation of (2) goes as follows. Define the sample space

$$\Omega = \{\omega : [0,\infty) \to \mathbb{R}^n, \omega \text{ continuous}\}. \tag{4}$$

*Draft notes for a talk for the thematic year on PDE at The Fields Institute

Physicists call this the "path space", because it can be thought as the path followed by a particle whose position at time $t \in [0, \infty)$ depends on the result of a random experiment with outcome ω and is given by the function

$$X_t(\omega) := \omega(t).$$

We now use $X_t : \Omega \to \mathbb{R}^n$ to endow Ω with a measurable space structure by defining for each $t \in [0, \infty)$ the σ–algebras

$$\mathcal{F}_t := \sigma\text{–algebra generated by } \{X_s^{-1}(A) : \text{ for all } 0 \le s \le t, A \in \mathcal{B}(\mathbb{R}^n)\}$$

and

$$\mathcal{F} := \sigma\text{–algebra generated by } \{X_s^{-1}(A) : \text{ for all } 0 \le s < \infty, A \in \mathcal{B}(\mathbb{R}^n)\},$$

where $\mathcal{B}(\mathbb{R}^n)$ denotes the Borel σ–algebra of $\mathbb{R}^n$. These have the properties that $\mathcal{F}_t \subset \mathcal{F}_s$ and $\mathcal{F}_t \subset \mathcal{F}$ whenever $t \le s < \infty$.

On each measurable space $(\Omega, \mathcal{F}_t)$ we can define the probability

$$P_x\{X_t^{-1}(A)\} := \int_A p(t, x, y)dy, \quad A \in \mathcal{B}(\mathbb{R}^n), \tag{5}$$

which can then be extended to the entire $\mathcal{F}$ (via the Kolmogorov extension theorem) due to the semigroup property of the heat kernel:

$$p(t, x, y) = \int_{\mathbb{R}^n} p(t - s, x, z)p(s, z, y)dz.$$

That is, for each $x \in \mathbb{R}^n$, we have that $(\Omega, \mathcal{F}_t, \mathcal{F}, P_x)$ is a filtered probability space on which the stochastic process X_t satisfies the properties

1. $P_x\{X_0 = x\} = 1$,
2. $P_x\{X_t \in A\} = E_x[P_{X_s}\{X_{t-s} \in A\}], \forall A \in \mathcal{B}(\mathbb{R}^n)$,
3. $E_x[X_t - x] = 0$ and $E_x[(X_t - x)^2] = t$.

These are the defining properties of a standard Brownian motion starting at the point $x \in \mathbb{R}^n$.

The last ingredient for the desired probabilistic interpretation of (2) is to observe that, for a given measurable function $f : \mathbb{R}^n \to \mathbb{R}$, the composition

$f \circ X_t : \Omega \to \mathbb{R}$ is a random variable on Ω whose expectation under the measure P_x is

$$E_x[f(X_t)] = \int_\Omega f(X_t(\omega))dP_x(\omega) = \int_{\mathbb{R}^n} p(t,x,y)f(y)dy,$$

from which we can rewrite (2) as our first example of the Feynman–Kac formula:

$$u(t,x) = E_x[f(X_t)]. \tag{6}$$

To proceed further it is convenient to focus on the scalar problem. Let us turn our ingredients around and start with an arbitrary filtered probability space $(\Omega, \mathcal{F}_t, \mathcal{F}, P)$ on which we define the standard Brownian motion W_t as the unique (up to indistinguishability) process starting at $W_0 = 0$ with independent increments $(W_t - W_s)$ distributed according to a Gaussian law with variance $(t-s)$. Then the process X_t used in the representation (6) of the solution to (1) is the solution of the *stochastic differential equation*

$$\begin{cases} dX_t = dW_t \\ X_0 = x. \end{cases} \tag{7}$$

As a rather trivial generalization, we might consider the scalar heat equation with a diffusion constant σ

$$\begin{cases} u_t - \frac{1}{2}\sigma^2 u_{xx} = 0 & \text{on} \quad (0,\infty) \times \mathbb{R} \\ u(0,x) = f(x) & \text{on} \quad \{t=0\} \times \mathbb{R}. \end{cases} \tag{8}$$

Then exactly the same construction leads to the representation of its solution in the form of (6) but with X_t being the solution to the SDE

$$\begin{cases} dX_t = \sigma dW_t \\ X_0 = x \end{cases} \tag{9}$$

The Feynman–Kac formula provides a probabilistic representation of solutions of PDEs whose generators are associated with general SDEs and, given its far reaching consequences, is surprisingly easy to prove once the technicalities involved in the definition of stochastic integrals are properly overcome. We state it as a theorem and outline its proof making reference to the necessary technical steps while avoiding any lengthy explanation about them. To conform with applications to mathematical finance, we henceforth consider a backward parabolic problem for the time interval $[0,T]$.

Theorem 1.1 (Feynman–Kac) *Let $(\Omega, \mathcal{F}_t, \mathcal{F}, P)$ be a filtered probability space and W_t be a standard Brownian motion. Assume that u is a solution to the Cauchy problem*

$$\left\{ \begin{array}{rl} u_t(t,x) + \mu(t,x)u_x(t,x) + \frac{1}{2}\sigma^2(t,x)u_{xx}(t,x) = 0 & \text{on} \quad [0,T) \times I\!R \\ u(T,x) = f(x) & \text{on} \quad \{t = T\} \times I\!R, \end{array} \right. \tag{10}$$

where the $\mu : [0,T] \times I\!R \to I\!R$ and $\sigma : [0,T] \times I\!R \to I\!R$ are measurable functions satisfying

$$|\mu(t,x)| + |\sigma(t,x)| \leq C(1 + |x|); \qquad t \in [0,T], \quad x \in I\!R$$

for some constant C and

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq D|x - y|; \qquad t \in [0,T], \quad x, y \in I\!R$$

for some constant D. Let X be the unique solution to the SDE

$$\left\{ \begin{array}{l} dX_s = \mu(t, X_s)ds + \sigma(s, X_s)dW_s \\ X_t = x \end{array} \right. \tag{11}$$

and assume further that $\int_0^t E[\sigma^2(s,X_s)u_x^2(s,X_s)]ds < \infty$ for all $t \in [0,T]$. Then

$$u(t,x) = E[f(X_T)|\mathcal{F}_t]. \tag{12}$$

Proof: Consider the process $Y_s = u(s, X_s)$. It follows from Ito's formula that

$$\begin{aligned} dY_s \;=\; & \left[u_t(s,X_s) + \mu(s,X_s)u_x(s,X_s) + \frac{1}{2}\sigma^2(s,X_s)u_{xx}(s,X_s) \right] ds \\ & +\sigma(s,X_s)u_x(s,X_s)dW_s, \end{aligned}$$

which when integrated on the interval (t,T) gives

$$\begin{aligned} u(T,X_T) \;=\; & u(t,X_t) + \int_t^T \sigma(s,X_s)u_x(s,X_s)dW_s \\ & + \int_t^T \left[u_t(s,X_s) + \mu(s,X_s)u_x(s,X_s) + \frac{1}{2}\sigma^2(s,X_s)u_{xx}(s,X_s) \right] ds. \end{aligned}$$

Now notice that the second integrand above vanishes, since u is a solution to the PDE. Moreover, because of the integrability condition we imposed on $\sigma(t, X_t)u_x(t, X_t)$, its stochastic integral with respect to the Brownian motion is a martingale. To conclude the proof we take expectations with respect to the law of the process X_s satisfying (11) evaluated at $s = t$, obtaining

$$E_{t,x}[u(T, X_T)] = E_{t,x}[u(t, X_t)] = u(t, x).$$

2 The Black–Scholes equation

Consider a financial market consisting of a risky asset with price S_t and a risk-free bank account B_t whose dynamics on the filtered probability space $(\Omega, \mathcal{F}_t, \mathcal{F}, P)$ is governed by

$$\begin{aligned} dB_t &= rB_t dt, \qquad\qquad\qquad B_0 = 1, &(13)\\ dS_t &= \mu(t, S_t)S_t dt + \sigma(t, S_t)S_t dW_t, \quad S_0 = s_0, &(14)\end{aligned}$$

where the risk-free interest rate r is assumed to be constant.

Let us introduce in this market a contingent claim of the form $\Phi(S_T)$, that is, a financial instrument whose pay-off depends on the terminal value of the risk asset. The celebrated result by Black, Scholes and Merton is that the only price process of the form $\pi_t = F(t, S_t)$, for some smooth function $F : [0, T] \times \mathbb{R}_+ \to \mathbb{R}$, which is consistent with the absence of arbitrage in the extended market (B_t, S_t, π_t) is (P–almost surely) the unique solution of the following boundary value problem on $[0, T] \times \mathbb{R}_+$:

$$\left\{ \begin{array}{r} F_t(t, s) + rsF_s(t, s) + \frac{1}{2}s^2\sigma^2(t, s)F_{ss}(t, s) - rF(t, s) = 0 \\ F(T, s) = \Phi(s). \end{array} \right. \tag{15}$$

Now observe that the boundary value problem above is closely related to the Cauchy problem of theorem 1.1, the only significant difference being the term $rF(t, s)$. A straightforward modification of the argument presented in the previous section shows that, under technical conditions on the function $\sigma(t, s)$, the solution of (15) admits the following Feynman-Kac representation:

$$F(t, s) = e^{-r(T-t)}E^Q_{t,s}[\Phi(S(T))], \tag{16}$$

where S_t is the solution of

$$\left\{ \begin{array}{l} dS_u = rS_u du + S_u\sigma(u, S_u)dW^Q_u \\ S_t = s, \end{array} \right. \tag{17}$$

for a Brownian motion W_t^Q on the filtered probability space $(\Omega', \mathcal{F}'_t, \mathcal{F}', Q)$. We have used the same letter S to denote both the original stock price process under the initial measure P and the process appearing in the Feynman-Kac formula, which is primarily just a technical tool. The important step now is to realize that we are free to choose $\Omega' = \Omega$, $\mathcal{F}'_t = \mathcal{F}_t$ and $\mathcal{F}' = \mathcal{F}$, as long as W^Q is a Brownian motion under Q. We then view the two different dynamics for S as realizations of the same stochastic process under two equivalent measures. The advantage of such interpretation is that it provides an *ansatz* for an explicit connection between the Brownian motions W_t and W_t^Q, namely

$$dW_t^Q = dW_t + \lambda dt, \tag{18}$$

where $\lambda = \frac{\mu - r}{\sigma}$ is called the *market price of risk*. Moreover, we can now use Girsanov's theorem (provided that λ_t satisfies the so called Novikov condition) to explicitly obtain the measure Q in terms of the measure P by putting

$$\frac{dQ}{dP} = \rho_T, \tag{19}$$

where ρ_T is the final value of the exponential P-martingale

$$\rho_t = \exp\left(-\int_0^t \lambda_s dW_s - \frac{1}{2}\int_0^t \lambda_s^2 ds \right). \tag{20}$$

That is, Q turns out to be equivalent to P, with density given by the stochastic exponential of $-\lambda_t$. Finally, it is easy to show that the discounted price process $B_t^{-1}S_t$ is a Q–martingale, from which Q is called the *equivalent martingale measure* for the market (B_t, S_t).

After the work of Harrison, Pliska and others, the main focus of derivative pricing has shifted from the Black-Scholes equation to the paradigm of "pricing by expectation", as expressed by (16), with the concept of equivalent martingale measures playing a central role. For instance, what is now called the First Fundamental Theorem of Asset Pricing asserts that the existence of a (local) equivalent martingale measure is equivalent to a version of absence of arbitrage (namely, "no free lunch with vanishing risk" – NFLVR). In the same vein, the Second Fundamental Theorem of Asset Pricing tells us that a market is complete, in the sense that every claim on the underlying assets can be uniquely replicated a combination of the assets themselves, if and only if the martingale measure is unique.

3 Stochastic Control

Consider an n–dimensional Brownian motion W_t on a filtered probability space $(\Omega, \mathcal{F}_t, \mathcal{F}, P)$ and a system whose state is given by an m–dimensional Ito process

$$\begin{cases} dZ_t^h = \mu(t, Z_t^h, h_t)dt + \sigma(t, Z_t^h, h_t)dW_t \\ Z_0^h = x_0, \end{cases} \tag{21}$$

where $\mu : \mathbb{R} \times \mathbb{R}^m \times \mathcal{B} \to \mathbb{R}$, $\sigma : \mathbb{R} \times \mathbb{R}^m \times \mathcal{B} \to \mathbb{R}^{m \times n}$ and $h_t \in \mathcal{B} \subset \mathbb{R}^k$.

The state process Z_t is deemed to be "controlled" by the parameter h_t taking values on a Borel set $\mathcal{B}$ of $\mathbb{R}^k$. A stochastic control problem consists of solving

$$u(x_0) = \sup_{h \in \mathcal{B}} E[\mathcal{U}(Z_T^h)], \tag{22}$$

where $\mathcal{U} : \mathbb{R}^n \to \mathbb{R}$ is a continuous function. That is, one tries to find the optimal control parameter $\hat{h}_t \in \mathcal{B}$ that we stir the system through (28) in order to produce the maximum expected value for the "utility" of the terminal state Z_T.

In the most general case, h_t needs only to be a random variable adapted to $\mathcal{F}_t$. As a special case, we restrict ourselves to Markov controls, that is, we assume that

$$h_t = h(t, Z_t). \tag{23}$$

Under these circumstances, the technique used to solve the optimization problem (22) is to embed it into the larger class of problems

$$u(t, x) = \sup_{h \in \mathcal{B}} E_{t,x}[\mathcal{U}(Z_T^h)], \tag{24}$$

where $E_{t,x}[\cdot]$ denotes expectation under the probability law of the solution Z_s^h to the SDE

$$\begin{cases} dZ_s^h = \mu(s, Z_s^h, h(s, Z_s^h))ds + \sigma(s, Z_s^h, h(s, Z_s^h))dW_s \\ Z_t^h = x, \end{cases} \tag{25}$$

evaluated at $s = t$, where we assume that μ and σ are regular enough so that Z_s^h exists as a well defined stochastic integral. We then have the following result from the theory of dynamic programming.

Theorem 3.1 (Hamilton–Jacobi–Bellman) *Suppose that $u(t, x)$ in (31) is bounded and $C^{1,2}$ on $[0, T] \times \mathbb{R}^m$ and assume that an optimal control $\hat{h}$ exists. Then*

1. The function $u(t,x)$ satisfies the Hamilton–Jacobi–Bellman equation

$$\begin{cases} u_t(t,x) + \sup_{h\in\mathcal{B}} \mathcal{L}^h u(t,x) = 0 \\ u(T,x) = \mathcal{U}(x), \end{cases} \tag{26}$$

where $\mathcal{L}^h$ is the generator of (25), that is,

$$\mathcal{L}^h(f) = \sum_{i=1}^m \mu_i^h \frac{\partial f}{\partial x_i} + \sum_{i,j=1}^m \frac{1}{2}(\sigma\sigma^\dagger)_{ij}^h \frac{\partial^2 f}{\partial x_i \partial x_j}. \tag{27}$$

2. For each $(t,x) \in [0,T] \times I\!R$, the supremum above is attained by $\hat{h}(t,x)$.

The above theorem has a convenient converse, namely that if $u(t,x)$ is a sufficiently integrable solution to the HJB equation and if the supremum of $\mathcal{L}^h$ is attained at every (t,x) by a function $\hat{h}(t,x)$, then $u(t,x)$ is the value function for the associated optimization problem with optimal Markov control given by $\hat{h}(t,x)$. We end this section with an even more comforting observation: under further technical conditions (which are easy to check for the financial markets considered in the literature), one can always obtain as good a performance with a Markov control as with an arbitrary $\mathcal{F}_t$–adapted control, that is, the optimal solution obtained from the HJB equation is as general as it can be expected.

4 Optimal Hedging in Incomplete Markets

As an application of stochastic control, we consider a financial market whose state is given by

$$\begin{cases} dZ_t^h = \mu(t,Z_t)dt + \sigma(t,Z_t)dW_t \\ Z_0^h = z_0, \end{cases} \tag{28}$$

where $\mu : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$, $\sigma : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^{m\times n}$. Let us say that the state decomposes as $Z_t = (S_t^1, \ldots, S_t^d, Y_t^1, , Y_t^{m-d})$, where the first d components are the prices of traded assets and the last $(m-d)$ factors are non-traded variables, such as market volatility, employment rates, inflation, etc. The control parameters come into the problem once we introduce the portfolio process $H_t = (H_t^1, \ldots, H_t^d)$ corresponding to an investor's asset allocations. We now let

$$X_t = x_0 + \int_0^t H_u dS_u \tag{29}$$

denoted the wealth process for a self–financing portfolio. Assume further that at the terminal time T the agent is faced with a liability in the form of a bounded $\mathcal{F}_T$–random variable B. The optimal hedging problem consists of solving

$$u(x_0) = \sup_{H\in\mathcal{A}} E\left[U\left(X_T - B\right)\right], \tag{30}$$

where $U : \mathbb{R} \to \mathbb{R}$ is an increasing, strictly convex, differentiable utility function and H ranges through a convex set of admissible portfolios $\mathcal{A}$.

The technique of dynamic programming for Markov controls can be applied, provided we also assume that the liability B is Markovian in the state variables, that is $B = \Phi(S_T, Y_T)$. One then considers the optimization problems

$$u(t,x,s,y) = \sup_{h\in\mathcal{B}} E_{t,x,s,y}[U(X_T^h - \Phi(S_T, Y_T))], \tag{31}$$

where $E_{t,x,s,y}[\cdot]$ now refers to expectations under the probability law of the process (X_t^h, S_t, Y_t), with $H_t = h(t, X_t^h, S_t, Y_t)$. Accordingly, the value function is now the solution of the modified HJB equation

$$\begin{cases} u_t(t,x,s,y) + \sup_{h\in\mathcal{B}} \mathcal{L}^h u(t,x,s,y) = 0 \\ u(T,x,s,y) = U(x - \Phi(s,y)), \end{cases} \tag{32}$$

where $\mathcal{L}^h$ is now the generator of the multidimensional process (X_t^h, S_t, Y_t) and the supremum above is attained at the optimal hedging portfolio $\hat{h}$.

The optimal hedging portfolio can be used to define a price for the liability B as follows. Denote the optimal wealth process for problem (30) by $\widehat{X}_t$. For each $t \in [0,T]$, define the *certainty equivalent process* B_t for the liability B by the relation

$$U(\widehat{X}_t - B_t) = u(t,x,s,y). \tag{33}$$

For the particular case $B = 0$, which corresponds to an optimal investment problem also known as Merton's problem, let us denote the certainty equivalent process by B_t^M. Then the *indifference price* for the liability B is defined to be the process

$$\pi_t := B_t - B_t^M. \tag{34}$$

5 Pricing with stochastic volatility

As a concrete example of the pricing concept introduced in the previous section, consider the market

$$\begin{aligned} dS_t &= S_t[\mu(t,Y_t)dt + Y_t dW_t^1] \\ dY_t &= a(t,Y_t)dt + b(t,Y_t)[\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2] \end{aligned} \tag{35}$$

for deterministic functions μ, a, b and $|\rho| \leq 1$, where W_t^1 and W_t^2 are uncorrelated one–dimensional Brownian motions.

For an exponential utility of the form $U(x) = -e^{-x}$, the indifference price $\pi = \pi(t,s,y)$ of a liability $B = \Phi(S_T, Y_T)$ satisfies

$$\pi_t + \mathcal{L}(\pi) + \frac{1}{2}(1-\rho^2)b^2(\pi_y)^2 = 0 \tag{36}$$

$$\pi(T,s,y) = \Phi(s,y) \tag{37}$$

where for any function $f : [0,T] \times D \to \mathbb{R}$,

$$\mathcal{L}(f) = \frac{1}{2}\left(y^2 s^2 f_{ss} + 2\rho bys f_{sy} + b^2 f_{yy} + \left[a - \frac{\rho b \mu}{y} + (1-\rho^2) b^2 B_y^0 \right] f_y \right).$$