J. Appl. Prob. **37**, 947–957 (2000)
Printed in Israel

MARTINGALES VERSUS PDEs IN FINANCE: AN EQUIVALENCE RESULT WITH EXAMPLES

DAVID HEATH,* *University of Technology, Sydney*
MARTIN SCHWEIZER,** *Technische Universität Berlin*

Abstract

We provide a set of verifiable sufficient conditions for proving in a number of practical examples the equivalence of the martingale and the PDE approaches to the valuation of derivatives. The key idea is to use a combination of analytic and probabilistic assumptions that covers typical models in finance falling outside the range of standard results from the literature. Applications include Heston's stochastic volatility model and the Black–Karasinski term structure model.

Keywords: Option valuation; martingale approach; partial differential equations; finance; Feynman–Kac formula

AMS 2000 Subject Classification: Primary 91B28; 60H30; 62P05

Introduction

Valuing derivative products is one of the most common problems in mathematical finance. Apart from approximations by discrete-time models, there are two main methods to obtain a valuation formula for a given derivative: via martingales and via partial differential equations (PDEs). In the more general *martingale approach*, one first specifies a stochastic process for the underlying primary asset and possibly other factors. Then one chooses an equivalent probability measure turning the discounted underlying asset into a (possibly local) martingale and computes the derivative's value as the conditional expectation of its discounted payoff under this risk-neutral measure. If the model has a Markovian structure, this value turns out to be some function u, say, of the state variables. In the *PDE approach*, one describes the state variables by a stochastic differential equation (SDE) and then derives for the function u based on the underlying martingale valuation a PDE involving the coefficients of the given SDE. One can also have state variables that follow processes with jumps; in that case, there will be additional integral terms. We remark in passing that both approaches can be used for complete as well as incomplete markets, even though there is some arbitrariness about the valuation rule in the latter case. For the martingale approach, this is reflected in the choice of martingale measure; in the PDE approach, one has the equivalent freedom of fixing the market price of risk for the nontraded variables.

From the preceding description as well as from economic intuition, it seems obvious that the two approaches should be equivalent. However, it turns out to be surprisingly tricky to give a rigorous proof for this. One possibility is to show that the valuation function derived from the martingale approach is sufficiently smooth for an application of Itô's formula; then it

Received 5 May 1998; revision received 17 April 2000.
* Postal address: University of Technology, Sydney, PO Box 123, Broadway, NSW 2007, Australia.
** Postal address: Technische Universität Berlin, Fachbereich Mathematik, MA 7–4, Str. des 17. Juni 136, D–10623 Berlin, Germany. Email address: mschweiz@math.tu-berlin.de

is easy to argue that it must satisfy the corresponding valuation PDE, at least on the support of the underlying diffusion. A second possibility is to prove existence of a nice solution to the valuation PDE; this is then readily seen to coincide with the valuation function from the martingale approach. The difficulty is that standard Feynman–Kac type results for both of these arguments only hold under restrictive analytic conditions on the SDE coefficients — assumptions not satisfied in many models used in practice. Quite remarkably, this problem is often glossed over or not mentioned in the literature.

In this paper, we provide a set of sufficient conditions for proving the equivalence of the martingale and the PDE approaches in a number of applied examples. We cannot give a general solution to this problem; our main contribution is to specify a mixture of analytic and probabilistic assumptions strong enough to allow us proving results, but still weak enough to be satisfied in some typical examples from finance. In Section 1, we formulate and prove for this a Feynman–Kac type result in a general multidimensional setting. Mathematically, our key assumptions are that the underlying process does not exit from a given domain almost surely and that its coefficient functions are sufficiently smooth on the *interior* of that domain; this allows us to handle degeneracies on the boundary that appear in many finance models. Neither the result nor the argument for its proof are surprising; the usefulness of our contribution is that the required conditions are often easy to verify in practice. Section 2 illustrates this in a range of examples including Heston's stochastic volatility model and the Black–Karasinski term structure model.

1. The theoretical result

In this section, we present our main theoretical result in a rather general form. Examples are deferred to the next section. Let $T \in (0, \infty)$ be a fixed time horizon and D a domain in $\mathbb{R}^d$, i.e., an open connected subset of $\mathbb{R}^d$. We consider the stochastic differential equation (SDE)

$$\mathrm{d}X_s^{t,x} = b(s, X_s^{t,x})\,\mathrm{d}s + \sum_{j=1}^{m} \Sigma_j(s, X_s^{t,x})\,\mathrm{d}W_s^j, \qquad X_t^{t,x} = x \in D \tag{1.1}$$

for continuous functions $b : [0, T] \times D \to \mathbb{R}^d$ and $\Sigma_j : [0, T] \times D \to \mathbb{R}^d$, $j = 1, \dots, m$, with an m-dimensional Brownian motion $W = (W^1 \, \dots \, W^m)^\top$. We write b and each Σ_j as a $(1 \times d)$ column vector and define the $(d \times m)$ matrix-valued function Σ by $\Sigma^{ij} := (\Sigma_j)^i$. For given measurable functions $h : D \to [0, \infty)$, $g : [0, T] \times D \to (-\infty, 0]$ and $c : [0, T] \times D \to \mathbb{R}$, we then define $u : [0, T] \times D \to [0, \infty]$ by

$$u(t,x) := \mathrm{E}\bigg[h(X_T^{t,x}) \exp\bigg(\int_t^T c(s, X_s^{t,x})\,\mathrm{d}s\bigg) - \int_t^T g(s, X_s^{t,x}) \exp\bigg(\int_t^s c(u, X_u^{t,x})\,\mathrm{d}u\bigg)\,\mathrm{d}s\bigg]; \tag{1.2}$$

this is well-defined in $[0, \infty]$ if $X^{t,x}$ does not explode or leave D before time T. We define the operator $\mathcal{L}$ on sufficiently smooth functions $f : [0, T] \times D \to \mathbb{R}$ by

$$(\mathcal{L}f)(t,x) := \sum_{i=1}^{d} b^i(t,x) \frac{\partial f}{\partial x^i}(t,x) + \frac{1}{2} \sum_{i,k=1}^{d} a^{ik}(t,x) \frac{\partial^2 f}{\partial x^i \partial x^k}(t,x) + c(t,x) f(t,x)$$

with

$$a^{ik}(t,x) := \sum_{j=1}^{m} \Sigma^{ij}(t,x) \Sigma^{kj}(t,x) = (\Sigma(t,x)\Sigma^\top(t,x))^{ik}.$$

Our goal is then to give sufficient conditions on X, D, b, Σ, h, c, g to ensure that the function u satisfies the partial differential equation (PDE)

$$\frac{\partial u}{\partial t} + \mathcal{L}u = g \quad \text{on } (0, T) \times D \tag{1.3}$$

with boundary condition

$$u(T, x) = h(x) \quad \text{for } x \in D. \tag{1.4}$$

This question is of course by no means new, and there are well-known sets of sufficient conditions for (1.3) and (1.4); see for instance Chapter 6 of Friedman (1975), Section 2.9 of Krylov (1980) or Appendix E of Duffie (1992). But many finance applications do not satisfy the very restrictive assumptions imposed by these standard results. For instance, b and Σ may be unbounded or grow faster than linearly or have unbounded derivatives etc. Our contribution here is to provide a mixture of analytic and probabilistic assumptions that still allow us to derive (1.3) and (1.4) while being weak enough to be satisfied in a number of applications.

The following result is the main theoretical contribution of this paper; we shall comment below on its assumptions and on possible ways of verifying them.

Theorem 1. *Suppose that the following conditions hold:*

(A1) *The coefficients b and Σ_j, $j = 1, \ldots, m$, are on $[0, T] \times D$ locally Lipschitz-continuous in x, uniformly in t, i.e., for each compact subset F of D, there is a constant $K_F < \infty$ such that*

$$|G(t, x) - G(t, y)| \leq K_F |x - y|$$

for all $t \in [0, T]$, $x, y \in F$ and $G \in \{b, \Sigma_1, \ldots, \Sigma_m\}$.

(A2) *For all $(t, x) \in [0, T) \times D$, the solution $X^{t,x}$ of (1.1) neither explodes nor leaves D before T, i.e., $\mathrm{P}[\sup_{t \leq s \leq T} |X_s^{t,x}| < \infty] = 1$ and $\mathrm{P}[X_s^{t,x} \in D$ for all $s \in [t, T]] = 1$.*

(A3) *There exists a sequence $(D_n)_{n \in \mathbb{N}}$ of bounded domains contained in D such that $\bigcup_{n=1}^{\infty} D_n = D$ and such that for each n, the PDE*

$$\frac{\partial w}{\partial t} + \mathcal{L}w = g \quad \text{on } (0, T) \times D_n$$

with boundary condition $w(t, x) = u(t, x)$ on $(0, T) \times \partial D_n \cup \{T\} \times D_n$ has a classical solution $w_n(t, x)$.

Then u satisfies the PDE (1.3) with boundary condition (1.4). In particular, u is in $C^{1,2}$ and there exists a unique classical solution to the PDE (1.3) and (1.4).

Proof. By Theorem II.5.2 of Kunita (1984), (A1) implies that (1.1) has a unique solution $X^{t,x}$ up to a possibly finite random explosion time. By (A2), this explosion time must be greater than T P-almost surely, so that $X^{t,x}$ is well-defined on $[t, T]$. Since h and $-g$ are nonnegative, the expectation in (1.2) is then well-defined in $[0, \infty]$ and (A3) implicitly contains the assumption that $u(t, x) < \infty$ on $(0, T) \times \partial D_n \cup \{T\} \times D_n$ for all n.

For fixed $(t, x) \in (0, T) \times D$, (A3) allows us to find $n \in \mathbb{N}$ such that $x \in D_n$. If we denote by $\tau_n := \{s \geq t \mid X_s^{t,x} \notin D_n\} \wedge T$ the first exit time of $X^{t,x}$ from D_n before T, then continuity of $X^{t,x}$ implies that $(\tau_n, X^{t,x}_{\tau_n}) \in (0,T) \times \partial D_n \cup \{T\} \times D_n$, so that $u(\tau_n, X^{t,x}_{\tau_n}) < \infty$. By Theorem 6.5.2 of Friedman (1975), we then have

$$w_n(t,x) = \mathrm{E}\Bigg[u(\tau_n, X^{t,x}_{\tau_n}) \exp\left(\int_t^{\tau_n} c(s, X^{t,x}_s)\,\mathrm{d}s\right) - \int_t^{\tau_n} g(s, X^{t,x}_s) \exp\left(\int_t^s c(u, X^{t,x}_u)\,\mathrm{d}u\right)\mathrm{d}s\Bigg]; \tag{1.5}$$

this can easily be verified by applying Itô's formula to w_n and using the PDE and boundary condition for w_n and the boundedness of D_n (to conclude that the appearing stochastic integral of W is a martingale). On the other hand, (A1) and (A2) imply that $X^{t,x}$ is a strong Markov process; see Ikeda and Watanabe (1989), Theorem IV.2.3 and the remark following Theorem IV.6.1. More precisely, these results are stated for b and Σ independent of t, but defined on all of $\mathbb{R}^d$. However, (A2) allows us to replace $\mathbb{R}^d$ by D throughout, and with the help of Chapter 6 of Stroock and Varadhan (1979), all the results easily generalize to the case where b and Σ depend on t. Hence, the strong Markov property yields

$$\begin{aligned}\mathrm{E}\Bigg[h(X^{t,x}_T)\exp\left(\int_t^T c(s, X^{t,x}_s)\,\mathrm{d}s\right) - \int_t^T g(s, X^{t,x}_s)\exp\left(\int_t^s c(u, X^{t,x}_u)\,\mathrm{d}u\right)\mathrm{d}s \,\Bigg|\, \mathcal{F}_{\tau_n}\Bigg] \\ = u(\tau_n, X^{t,x}_{\tau_n})\exp\left(\int_t^{\tau_n} c(s, X^{t,x}_s)\,\mathrm{d}s\right) - \int_t^{\tau_n} g(s, X^{t,x}_s)\exp\left(\int_t^s c(u, X^{t,x}_u)\,\mathrm{d}u\right)\mathrm{d}s\end{aligned}$$

and, therefore,

$$u(t,x) = w_n(t,x)$$

by (1.2) and (1.5). Thus, u and w_n coincide on $(0,T) \times D_n$ for all n and this implies, by (A3), that u satisfies (1.3) on $(0,T) \times D$. The boundary condition (1.4) is evident from (1.2) and (1.1), and uniqueness follows from the probabilistic representation (1.2).

Remark. While the examples in Section 2 clearly illustrate the scope of Theorem 1, one general comment seems appropriate here. In quite a number of cases, one can explicitly solve the PDE (1.3) and (1.4) by either analytic or probabilistic methods if the SDE coefficients do not depend on time. One major achievement of our result is that even in the time-dependent case, we still obtain smoothness of u and an existence result for the PDE. To the best of our knowledge, general results of this type under our weak assumptions have not been available so far.

At first sight, some of our conditions may look deterringly abstract and hard to verify. The most harmless one is probably (A1); it is, for instance, satisfied if b and Σ are differentiable in x on the *open* set $(0,T) \times D$ with derivatives that are continuous on $[0,T] \times D$. This is easy to check and notably also covers situations where some derivatives become infinite on the boundary $(0,T) \times \partial D$. Condition (A2) has to be verified individually in each case and involves a more careful study of the process X under consideration. To make (A3) more palatable, we first note that by Theorem 6.3.6 and the remark before Theorem 6.5.2 of Friedman (1975), (A3) is implied by the combination of:

(A3′) there exists a sequence $(D_n)_{n\in\mathbb{N}}$ of bounded domains with $\overline{D_n} \subseteq D$ such that $\bigcup_{n=1}^\infty D_n = D$, each D_n has a C^2-boundary and for each n,

(A3a′) the functions b and $a = \Sigma\Sigma^\top$ are uniformly Lipschitz-continuous on $[0,T] \times \overline{D_n}$, (A3b′) $a(t,x)$ is uniformly elliptic on $\mathbb{R}^d$ for $(t,x) \in [0,T) \times D_n$, i.e., there is $\delta_n > 0$ such that $y^\top a(t,x) y \geq \delta_n |y|^2$ for all $y \in \mathbb{R}^d$,

(A3c′) c is uniformly Hölder-continuous on $[0,T] \times \overline{D_n}$,

(A3d′) g is uniformly Hölder-continuous on $[0,T] \times \overline{D_n}$, and

(A3e′) u is finite and continuous on $[0,T] \times \partial D_n \cup \{T\} \times \overline{D_n}$.

Examples in the next section will show how one can often readily verify (A3′). The crucial difference from the standard results in the literature is that the restrictive uniform assumptions on a, b, c, g are not imposed globally for $x \in D$, but only locally on the bounded domains D_n. This allows us to handle certain degeneracies on the boundary of D.

Condition (A3e′) requires continuity of u, which at first sight seems difficult to verify. But since we only need this in the interior of D and since (A1) and (A2) ensure that X is well-behaved there, we can easily give a simple sufficient condition for (A3e′). Its boundedness assumptions on h, g and c are stronger than required; they could be replaced by uniform integrability of the family

$$\left\{h(X_T^{r,y}) \exp\left(\int_r^T c(s, X_s^{r,y})\,\mathrm{d}s\right) - \int_r^T g(s, X_s^{r,y}) \exp\left(\int_r^s c(u, X_u^{r,y})\,\mathrm{d}u\right) \mathrm{d}s\right\},$$

where (r,y) runs through a neighbourhood of (t,x), to obtain the same conclusion.

Lemma 2. *Assume that (A1) and (A2) hold. If h, g and c are continuous, h and g are bounded and c is bounded from above, then u is continuous on* $[0,T] \times D$.

Proof. In view of the definition (1.2) of u and the boundedness assumptions on h, g and c, the assertion will follow from the dominated convergence theorem once we show that

$$(t,x) \mapsto h(X_T^{t,x}) \exp\left(\int_t^T c(s, X_s^{t,x})\,\mathrm{d}s\right) - \int_t^T g(s, X_s^{t,x}) \exp\left(\int_t^s c(u, X_u^{t,x})\,\mathrm{d}u\right) \mathrm{d}s \tag{1.6}$$

is P-almost surely continuous. By Theorem II.5.2 of Kunita (1984), (A1) and (A2) imply that $X^{t,x}$ has a version such that the mapping

$$(t,x,s) \mapsto X_s^{t,x} \quad \text{is P-almost surely continuous.}$$

Hence $(t,x) \mapsto h(X_T^{t,x})$ is P-almost surely continuous and $(t,x,s) \mapsto c(s, X_s^{t,x})$, $(t,x,s) \mapsto g(s, X_s^{t,x})$ are P-almost surely uniformly continuous and bounded on compact subsets of $[0,T] \times D \times [t,T]$. Because this readily implies that $(t,x) \mapsto \int_t^T c(s, X_s^{t,x})\,\mathrm{d}s$ and $(t,x) \mapsto \int_t^T g(s, X_s^{t,x})\,\mathrm{d}s$ are P-almost surely continuous, so is the mapping in (1.6) and this completes the proof.

The seemingly unpleasant ellipticity condition (A3b′) can also be verified quite simply. Note that $a(t,x)$ cannot be uniformly elliptic unless $\det a(t,x) \neq 0$, so that the conditions of Lemma 3 are in a sense almost optimal.

Lemma 3. *Assume that Σ is continuous in (t,x) and fix any bounded domain $D' \subseteq D$. If $\det a(t,x) \neq 0$ for all $(t,x) \in [0,T] \times \overline{D'}$, then $a(t,x)$ is uniformly elliptic on $\mathbb{R}^d$ for $(t,x) \in [0,T) \times D'$.*

Proof. Let $\alpha(t,x)$ be the square root of the symmetric nonnegative definite matrix $a(t,x)$ so that $a(t,x) = \alpha(t,x)\alpha^\top(t,x)$. Because $\det a(t,x) \neq 0$, $\alpha^\top(t,x)$ is invertible for any $(t,x) \in [0,T] \times \overline{D'}$, and so we obtain, for $y \neq 0$,

$$\frac{y^\top a(t,x)y}{|y|^2} = \frac{|\alpha^\top(t,x)y|^2}{|(\alpha^\top(t,x))^{-1}\alpha^\top(t,x)y|^2} \geq \frac{1}{\|(\alpha^\top(t,x))^{-1}\|^2},$$

where $\|A\| = \sup_{|z|\leq 1} |Az|$ is the operator norm. Each component of the matrix $\Sigma(t,x)$ is, by assumption, continuous in (t,x); the same then clearly holds for $a(t,x) = \Sigma(t,x)\Sigma^\top(t,x)$ and, by Lemma 6.1.1 of Friedman (1975), for $\alpha(t,x)$ as well. Since $\det\alpha^\top(t,x) \neq 0$, each component of the inverse matrix $(\alpha^\top(t,x))^{-1}$ is then also continuous in (t,x) and

$$B := \sup_{(t,x)\in[0,T]\times\overline{D'}} \|(\alpha^\top(t,x))^{-1}\| = \sup\{|(\alpha^\top(t,x))^{-1}z| \mid (t,x) \in [0,T] \times \overline{D'}, |z| \leq 1\}$$

is therefore the supremum of a continuous function over a compact set, hence finite. Clearly, $\delta := 1/B^2$ is then sufficient for the uniform ellipticity condition.

Combining Theorem 1 with Lemmas 2 and 3 gives us a fairly handy set of assumptions under which we rigorously obtain the equivalence of the martingale and the PDE approaches. The next section shows how to exploit this in specific examples.

2. Examples

This section illustrates the usefulness of Theorem 1 by several examples. We have chosen these to cover a range of problems and models relevant for practical applications. The general format of our presentation will always be the same: we briefly outline the model, explain how it fits into our framework and comment on the contribution made by our result.

2.1. The Heston stochastic volatility model

Our first example uses a model with two assets B and S. The bank account, B, is given by $B_t = \mathrm{e}^{rt}$, where r is the instantaneous riskless interest rate. The stock, S, satisfies the SDE

$$\frac{\mathrm{d}S_u}{S_u} = \mu_u\,\mathrm{d}u + \sqrt{v_u}\,\mathrm{d}\bar{W}_u, \qquad S_0 > 0 \tag{2.1}$$

where the squared volatility v is itself stochastic and given, as in Heston (1993), by

$$\mathrm{d}v_u = \kappa(\vartheta - v_u)\,\mathrm{d}u + \sigma\sqrt{v_u}\,\mathrm{d}\bar{W}'_u, \qquad v_0 > 0 \tag{2.2}$$

for nonnegative constants $\kappa, \vartheta, \sigma$. The processes S and v are defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathrm{Q})$ and $\bar{W}, \bar{W}'$ are Q-Brownian motions with instantaneous correlation ϱ, that is, $\mathrm{d}\langle\bar{W}, \bar{W}'\rangle_u = \varrho\,\mathrm{d}u$ for some constant $\varrho \in (-1,1)$. The goal is to value a European put option on S; we consider puts rather than calls in order to have a bounded payoff function.

To make Theorem 1 applicable, we first have to recast this problem as the computation of a suitable expectation. Let P be an equivalent local martingale measure for S/B, i.e., a probability $\mathrm{P} \approx \mathrm{Q}$ such that S/B is a local P-martingale. Like Heston (1993), we moreover choose P such that the market price of volatility risk is $(\lambda/\sigma)\sqrt{v}$, for some constant λ. Note that, in contrast to Heston (1993), here we use the standard definition of market price of risk as in Hull (1997) or Ingersoll (1987). This means that under P, the pair (S, v) is given by

$$\begin{aligned} \mathrm{d}S_u &= rS_u\,\mathrm{d}u + \sqrt{v_u}\,S_u\,\mathrm{d}W_u, \\ \mathrm{d}v_u &= (\kappa(\vartheta - v_u) - \lambda v_u)\,\mathrm{d}u + \sigma\sqrt{v_u}\,\mathrm{d}W'_u, \end{aligned} \tag{2.3}$$

where W, W' are now P-Brownian motions with instantaneous correlation ϱ. The P-price at time t of a European put on S with maturity T and strike K is then

$$\mathrm{E}[\mathrm{e}^{-r(T-t)}(K - S_T)^+ \mid \mathcal{F}_t] = u(t, S_t, v_t)$$

by the Markov property of (S, v) under P.

Consider now the two-dimensional process X with coordinates $X^1 := S$ and $X^2 := v$. To construct W and W' as in (2.3), choose independent P-Brownian motions W^1, W^2 and set $W' := W^2$ and $W := \varrho W^2 + \sqrt{1-\varrho^2}\, W^1$. Then X satisfies the SDE (1.1) with coefficients

$$b(t,x) = \begin{bmatrix} rx^1 \\ \kappa(\vartheta - x^2) - \lambda x^2 \end{bmatrix}, \qquad \Sigma(t,x) = \begin{bmatrix} \sqrt{1-\varrho^2}x^1\sqrt{x^2} & \varrho x^1 \sqrt{x^2} \\ 0 & \sigma\sqrt{x^2} \end{bmatrix}$$

for $x \in D := (0,\infty)^2$. Since b and Σ do not depend on t and are obviously C^1 in x on D, it is clear that (A1) is satisfied. If we assume that $\kappa\vartheta \geq \frac{1}{2}\sigma^2 > 0$, then Feller's test for explosion shows, as in Example IV.8.2 of Ikeda and Watanabe (1989), that v with probability 1 neither hits 0 nor explodes to $+\infty$, and so the same is true for $S = S_0 \exp(\int \sqrt{v_s}\, \mathrm{d}W_s + \int (r - \frac{1}{2}v_s)\, \mathrm{d}s)$. Thus (A2) is also satisfied. To compute the option price as $u(t, S_t, v_t)$ with u given by (1.2), we finally choose $g(t,x) \equiv 0$, $c(t,x) \equiv -r$ and $h(x) := (K - x^1)^+$. In order to verify (A3) via (A3′), we take as domains D_n the squares $(1/n, n)^2$ with smoothed corners so that they satisfy (A3′). Because b and Σ are C^1 in x, (A3a′) is obvious; so is (A3d′), and (A3e′) follows by Lemma 2. Finally, an elementary calculation gives $\det a(t,x) = \sigma^2 |x^1|^2 |x^2|^2 (1-\varrho^2) > 0$ on $[0,T] \times D$, and so Lemma 3 implies that (A3b′) is also satisfied. By applying Theorem 1, using the more suggestive variables (s, v) instead of (x^1, x^2) and writing subscripts for partial derivatives, we conclude that the put pricing function $u(t,x) = u(t,s,v)$ satisfies the PDE

$$u_t + rsu_s + (\kappa(\vartheta - v) - \lambda v)u_v + \tfrac{1}{2}s^2 v u_{ss} + \varrho\sigma s v u_{sv} + \tfrac{1}{2}\sigma^2 v u_{vv} - ru = 0 \tag{2.4}$$

on $(0,T) \times D$ with boundary condition

$$u(T,s,v) = (K-s)^+ \quad \text{for } (s,v) \in (0,\infty)^2. \tag{2.5}$$

This PDE already appears in Heston (1993) and is handled there by analytically finding the Fourier transform of the price and inverting this numerically. One may thus wonder what we have gained from our approach, and we can give at least two answers. First of all, we have rigorously proved that in this example, the PDE approach (solving (2.4) and (2.5) for u) and the martingale approach (computing u via (1.2)) lead to the same option pricing function. This is, of course, intuitively clear and could in the present simple situation also be deduced by other probabilistic methods; see for example Leblanc (1996). More importantly, though, our result immediately generalizes to situations where r, κ, ϑ, λ, σ are not constants, but sufficiently smooth (C^1 in t on $[0,T]$, say) deterministic *functions* of time. In that case, our method proves the existence and uniqueness of a classical solution to the PDE (2.4) and (2.5) and even gives via (1.2) a probabilistic recipe for its computation. To the best of our knowledge, these results are new and look very helpful for practical applications.

A third benefit from our approach, which actually provided the original motivation for this study, appears if we look at option prices under different martingale measures. Suppose, for instance, that our initial model under Q is as in (2.1) and (2.2) with specific drift $\mu_u = \gamma v_u$ for some constant γ and with correlation $\varrho = 0$. If we use, instead of Heston's martingale measure P, the variance-optimal martingale measure $\widetilde{\mathrm{P}}$, we have to replace (2.3) by

$$\begin{aligned} \mathrm{d}S_u &= rS_u\,\mathrm{d}u + \sqrt{v_u}\,S_u\,\mathrm{d}\widetilde{W}_u, \\ \mathrm{d}v_u &= (\kappa(\vartheta - v_u) - \sigma^2\gamma^2\alpha(T-u)v_u)\,\mathrm{d}u + \sigma\sqrt{v_u}\,\mathrm{d}\widetilde{W}'_u, \end{aligned} \tag{2.6}$$

where $\widetilde{W}$, $\widetilde{W}'$ are independent $\widetilde{\mathrm{P}}$-Brownian motions and the function $\alpha(\tau)$ is given by

$$\alpha(\tau) = \frac{2(\mathrm{e}^{\Gamma\tau} - 1)}{(\Gamma + \kappa)(\mathrm{e}^{\Gamma\tau} - 1) + 2\Gamma}$$

with $\Gamma = \sqrt{2\gamma^2\sigma^2 + \kappa^2} > 0$. For a more detailed derivation of (2.6), we refer to Laurent and Pham (1999) or Heath *et al.* (1999). The $\widetilde{\mathrm{P}}$-price of a European put is now

$$\widetilde{\mathrm{E}}[\mathrm{e}^{-r(T-t)}(K - S_T)^+ \mid \mathcal{F}_t] = \widetilde{u}(t, S_t, v_t)$$

and we want to deduce from Theorem 1 that $\widetilde{u}$ is sufficiently smooth and satisfies a PDE determined by the coefficients of the SDE (2.6). Since α is at least C^1 on $[0, T]$, it is immediately clear that (A1) and (A3) are again satisfied. Moreover, (A2) holds because we know already that v does not leave $(0, \infty)$ up to time T with P-probability 1, and $\widetilde{\mathrm{P}}$ and P are equivalent on $\mathcal{F}_T$. Thus Theorem 1 is applicable under $\widetilde{\mathrm{P}}$ instead of P and gives the desired conclusion about $\widetilde{u}$.

Remark. As the preceding considerations show, there are no unique option prices in the Heston model. This is, of course, clear, since the model is incomplete: there are two sources, $\bar{W}$ and $\bar{W}'$, of uncertainty, but only one risky asset, S, available for trade. Hence, option prices are only determined once a particular martingale measure has been chosen (or, equivalently, a market price of risk). In particular, each (reasonable) martingale measure also gives rise to an associated PDE; this means that many different PDEs can be found in this example.

2.2. The Black–Karasinski term structure model

For our second example, we consider the pricing of a zero coupon bond in the term structure model of Black and Karasinski (1991). They describe the short rate r by the SDE

$$\mathrm{d}(\log r_u) = \varphi(u)(\log\mu(u) - \log r_u)\,\mathrm{d}u + \sigma(u)\,\mathrm{d}W_u, \qquad r_0 > 0 \tag{2.7}$$

with deterministic functions φ, μ, σ. The price at time t of a zero coupon bond with maturity T is then

$$\mathrm{E}\left[\exp\left(-\int_t^T r_s\,\mathrm{d}s\right)\,\middle|\,\mathcal{F}_t\right] = u(t, r_t)$$

by the Markov property of r, and we want to obtain a PDE for the function u.

By Itô's formula, the one-dimensional process $X_t := r_t$ satisfies the SDE (1.1) with

$$b(t,x) = x(\varphi(t)\log\mu(t) + \tfrac{1}{2}\sigma^2(t) - \varphi(t)x), \qquad \Sigma(t,x) = \sigma(t)x.$$

If we choose $D := (0, \infty)$ and assume that φ, $\log\mu$ and σ are all C^1 in t on $[0, T]$, (A1) is clearly satisfied. It is straightforward to verify that the solution of (2.7) is explicitly given by

$$\log r_t = \frac{1}{G(t)}\left(\log r_0 + \int_0^t G(s)\varphi(s)\log\mu(s)\,\mathrm{d}s + \int_0^t G(s)\sigma(s)\,\mathrm{d}W_s\right)$$

with $G(t) := \exp(\int_0^t \varphi(s)\,\mathrm{d}s)$, so that (A2) is also satisfied. With $D_n := (1/n, n)$, (A3′)–(A3b′) are easily seen to be fulfilled. Since we want to compute

$$u(t,x) := \mathrm{E}\left[\exp\left(-\int_t^T X_s^{t,x}\,\mathrm{d}s\right)\right],$$

we choose $g(t,x) \equiv 0$, $h(x) \equiv 1$ and $c(t,x) := -x$. Then c is Lipschitz-continuous, hence uniformly Hölder-continuous, so that (A3d′) holds, and since c is continuous and nonpositive on $[0,T] \times D$, (A3e′) follows from Lemma 2. Thus, Theorem 1 implies that $u(t,x)$ satisfies the PDE

$$u_t + b(t,x)u_x + \tfrac{1}{2}\Sigma^2(t,x)u_{xx} - xu = 0 \quad \text{on } (0,T) \times D$$

with boundary condition $u(T,x) = 1$, for $x \in (0,\infty)$. As in the previous example, our contribution here is to prove existence and uniqueness of a classical solution to the above PDE and to show rigorously that the PDE and martingale approaches lead to the same result.

2.3. The CEV model

In our third and final example, we study the pricing of a European put in the constant elasticity of variance (CEV) model introduced by Cox (1996); see also Hull (1997). There are again two assets; the bank account B is given by $B_t = \mathrm{e}^{rt}$ and the stock X follows the SDE

$$\mathrm{d}X_u = rX_u\,\mathrm{d}u + \sigma X_u^\alpha\,\mathrm{d}W_u, \qquad X_0 > 0 \tag{2.8}$$

with constants $\sigma > 0$ and $\alpha > 0$. (We omit the case $\alpha = 0$, since it yields an Ornstein–Uhlenbeck process for X, and thus leads to negative stock prices.) To value a European put with strike K and maturity T, we have to compute

$$u(t,x) := \mathrm{E}[\mathrm{e}^{-r(T-t)}(K - X_T^{t,x})^+],$$

which is just (1.2) with $g \equiv 0$, $c \equiv -r$ and $h(x) := (K-x)^+$. By now, routine arguments show that (A1) and (A3′)–(A3d′) are satisfied with $D := (0,\infty)$ and (A3e′) will follow by Lemma 2 if (A2) is satisfied. Once we verify (A2), Theorem 1 will therefore imply that u solves the PDE

$$u_t + rxu_x + \tfrac{1}{2}\sigma^2 x^{2\alpha}u_{xx} - ru = 0 \quad \text{on } (0,T) \times D \tag{2.9}$$

with boundary condition

$$u(T,x) = (K-x)^+ \quad \text{for } x \in (0,\infty). \tag{2.10}$$

To deal with (A2), we use Feller's test for explosion as explained in Ikeda and Watanabe (1989). Since X satisfies (1.1) with

$$b(t,x) = rx, \qquad \Sigma(t,x) = \sigma x^\alpha,$$

we have to examine the function

$$\kappa(x) := \int_1^x \exp\left(-\int_1^y \frac{2b(t,z)}{\Sigma^2(t,z)}\,\mathrm{d}z\right)\int_1^y \exp\left(\int_1^s \frac{2b(t,z)}{\Sigma^2(t,z)}\,\mathrm{d}z\right)\frac{1}{\Sigma^2(t,s)}\,\mathrm{d}s\,\mathrm{d}y$$

for finiteness at the boundaries $x = 0$ and $x = +\infty$ of D. The case $\alpha = 1$ is explicitly solvable and yields for X the familiar model of geometric Brownian motion which satisfies (A2). If $\alpha \neq 1$, then

$$\int_1^s \frac{2b(t,z)}{\Sigma^2(t,z)}\,\mathrm{d}z = \frac{2r}{\sigma^2}\int_1^s z^{1-2\alpha}\,\mathrm{d}z = \frac{r}{\sigma^2(1-\alpha)}(s^{2(1-\alpha)} - 1) =: G(s)$$

and therefore

$$\kappa(x) = \frac{1}{\sigma^2}\int_1^x \mathrm{e}^{-G(y)}\int_1^y \mathrm{e}^{G(s)}s^{-2\alpha}\,\mathrm{d}s\,\mathrm{d}y = \frac{1}{2r}\int_1^x \mathrm{e}^{-G(y)}\int_1^y \mathrm{e}^{G(s)}G'(s)\frac{1}{s}\,\mathrm{d}s\,\mathrm{d}y \quad (2.11)$$

for $x \geq 1$. Since $s \leq y$ and all integrands are nonnegative, we obtain

$$\kappa(x) \geq \frac{1}{2r}\int_{x_0}^x \mathrm{e}^{-G(y)}\frac{1}{y}(\mathrm{e}^{G(y)} - 1)\,\mathrm{d}y$$

for $x \geq x_0 \geq 1$. Now, if $\alpha > 1$, then $G(y)$ is bounded on $[1, \infty)$ and increases from 0 to some finite positive value; for x_0 large enough, we thus obtain $\mathrm{e}^{G(y)} - 1 \geq$ const. > 0 and $\mathrm{e}^{-G(y)} \geq$ const. > 0 for $y \geq x_0$. If $\alpha < 1$, then $G(y)$ increases to $+\infty$ as $y \to \infty$; hence $\mathrm{e}^{G(y)} - 1 \geq \frac{1}{2}\mathrm{e}^{G(y)}$ and $\mathrm{e}^{-G(y)}(\mathrm{e}^{G(y)} - 1) \geq$ const. > 0 for $y \geq x_0$. This implies that

$$\kappa(x) \geq \text{const.}\int_{x_0}^x \frac{1}{y}\,\mathrm{d}y,$$

so that, clearly,

$$\lim_{x \nearrow +\infty} \kappa(x) = +\infty \quad \text{for all } \alpha > 0,\ \alpha \neq 1. \quad (2.12)$$

For $x \leq 1$ and $\alpha > 1$, we rewrite (2.11) as

$$\kappa(x) = \frac{1}{2r}\int_x^1 \mathrm{e}^{-G(y)}\int_y^1 \mathrm{e}^{G(s)}G'(s)\frac{1}{s}\,\mathrm{d}s\,\mathrm{d}y \geq \frac{1}{2r}\int_x^{x_0} \mathrm{e}^{-G(y)}(1 - \mathrm{e}^{G(y)})\,\mathrm{d}y$$

for $x \leq x_0 \leq 1$, since $s \leq 1$. Now, $\alpha > 1$ yields $\lim_{y \searrow 0} G(y) = -\infty$ and therefore $1 - \mathrm{e}^{G(y)} \geq$ const. > 0 for $0 < y \leq x_0$ and $x_0 > 0$ small enough. Moreover,

$$\mathrm{e}^{-G(y)} = \text{const.}\exp(y^{-2(\alpha-1)}) \geq \frac{\text{const.}}{k!}y^{-2k(\alpha-1)}$$

for any $k \in \mathbb{N}$, and so we get

$$\kappa(x) \geq \text{const.}\int_x^{x_0} y^{-2k(\alpha-1)}\,\mathrm{d}y \quad \text{for } x \leq x_0 \text{ and } x_0 > 0 \text{ small.}$$

Since $\alpha > 1$, we can choose k large enough to get $2k(\alpha - 1) > 1$ and this leads to

$$\lim_{x \searrow 0} \kappa(x) = +\infty \quad \text{for } \alpha > 1. \quad (2.13)$$

If $\alpha < 1$, then G is bounded on $[0, 1]$. Hence (2.11) gives

$$\kappa(x) \leq \text{const.}\int_x^1 \int_y^1 s^{-2\alpha}\,\mathrm{d}s\,\mathrm{d}y$$

for $x < 1$, and elementary computations then yield

$$\lim_{x \searrow 0} \kappa(x) < +\infty \quad \text{for } 0 < \alpha < 1. \quad (2.14)$$

From (2.12)–(2.14) and Theorem VI.3.2 of Ikeda and Watanabe (1989), we conclude that (A2) is satisfied for all $\alpha \geq 1$, whereas X leaves D in finite time with positive probability for $\alpha < 1$.

For $\alpha \geq 1$, (A2) is therefore satisfied; Theorem 1 is then applicable and provides existence and uniqueness of a solution to (2.9) and (2.10). For $\alpha < 1$, (A2) is not satisfied and our approach does not give any new results.

We round off this subsection by giving some comments on the model itself. From the SDE (2.8) for X, it is obvious that the discounted price process X/B is a local martingale under P and so the model is certainly arbitrage-free if X is well-defined. The only possible problem is that (2.8) might fail to have a solution. But, for $\alpha \geq 1$, the coefficients in (2.8) are locally Lipschitz-continuous, which guarantees the existence of a (possibly exploding) solution, and we have just seen that the explosion time is actually $+\infty$. For $\alpha < 1$, the results of Delbaen and Shirakawa (1995) tell us that (2.8) has a unique weak solution; this is even a strong solution if one stops the process X when it hits 0.

The preceding analysis or the explicit result in Delbaen and Shirakawa (1995) shows that for $\alpha < 1$, stock prices will hit 0 with positive probability in finite time. It is surprising that this fact is not usually mentioned in the CEV literature, because it may give the model some rather unpleasant properties. If one does not stop X when it hits 0, then X/B is only a local martingale, but not a martingale. (The same thing happens for $\alpha > 1$.) But this implies for instance that put-call parity is not satisfied or that the stock is not priced by taking conditional expectations after discounting. By going short in the stock, one can even produce the final payoff X_T at a price lower than X_0. This does not contradict the absence of arbitrage, because the strategy of shorting the stock is not admissible if X/B is only a local martingale. These rather subtle issues are not usually mentioned in the literature on the CEV model and it may be useful to point out explicitly the existence of such pitfalls.

Acknowledgement

We thank an unknown referee for a number of helpful comments.

References

Black, F. and Karasinski, P. (1991). Bond and option pricing when short rates are lognormal. *Financial Analysts J.* **Jul./Aug. 1991,** 52–59.

Cox, J. C. (1996). The constant elasticity of variance option pricing model. *J. Portfolio Management*, **spec. iss. Dec. 1996,** 15–17.

Delbaen, F. and Shirakawa, H. (1995). Option pricing for constant elasticity of variance model. Preprint 96-03, Tokyo Institute of Technology.

Duffie, D. (1992). *Dynamic Asset Pricing Theory*. Princeton University Press.

Friedman, A. (1975). *Stochastic Differential Equations and Applications*, Vol. 1. Academic Press, New York.

Heath, D., Platen, E. and Schweizer, M. (1999). A comparison of two quadratic approaches to hedging in incomplete markets. Preprint, Technical University of Berlin.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Stud.* **6,** 327–343.

Hull, J. C. (1997). *Options, Futures, and other Derivatives*, 3rd edn. Prentice-Hall, New York.

Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*. North-Holland, Amsterdam.

Ingersoll, J. E. Jr. (1987). *Theory of Financial Decision Making*. Rowman & Littlefield, Totowa, New Jersey.

Krylov, N. V. (1980). *Controlled Diffusion Processes*. Springer, New York.

Kunita, H. (1984). Stochastic differential equations and stochastic flows of diffeomorphisms. In *École d'Été de Probabilités de Saint-Flour XII–1982* (Lecture Notes in Math. **1097**), ed. P. H. Hennequin. Springer, Berlin, pp. 143–303.

Laurent, J. P. and Pham, H. (1999). Dynamic programming and mean-variance hedging. *Finance Stoch.* **3,** 83–110.

Leblanc, B. (1996). Une approche unifiée pour une forme exacte du prix d'une option dans différents modèles à volatilités stochastiques. *Stoch. Stoch. Rep.* **57,** 1–35.

Stroock, D. W. and Varadhan, S. R. S. (1979). *Multidimensional Diffusion Processes*. Springer, New York.