FEYNMAN-KAC FORMULAS FOR BLACK-SCHOLES TYPE OPERATORS

SVANTE JANSON* AND JOHAN TYSK*

Abstract. There are many references showing that a classical solution to the Black–Scholes equation is a stochastic solution. However, it is the converse of this theorem which is most relevant in applications and the converse is also more mathematically interesting. In the present article we establish such a converse. We find a Feynman–Kac type theorem showing that the stochastic representation yields a classical solution to the corresponding Black–Scholes equation with appropriate boundary conditions under very general conditions on the coefficients. We also obtain additional regularity results in the one-dimensional case.

1. Introduction

Stochastic formulas for option prices are often easy to formulate and to implement in for instance Monte Carlo algorithms. It is natural from the point of view of applications to let zero be an absorbing barrier for processes describing the risky assets. In Section 3 we discuss such a representation in the case of time- and level-dependent volatilities. However, it is often advantageous to instead solve the corresponding Black–Scholes equation and thus be able to use results from the theory of partial differential equations. In the literature one often omits specifying the boundary conditions on the lateral part of the parabolic boundary. This causes no problem when the risky assets are modelled with geometric Brownian motion, since assets in this model reach zero with probability zero. However, for many models, such as the constant elasticity of variance models, the values of the assets can vanish with positive probability and boundary conditions need to be specified. Furthermore, in numerical applications one is often helped by knowing the boundary behavior of the solution even if these conditions are mathematically redundant to specify.

In fact, there are many references showing that a classical solution to the Black–Scholes equation is a stochastic solution, compare Theorem 2.5. However, it is the converse of this theorem which is most relevant in applications as described above and the converse is also more mathematically subtle. In the present article we establish such a converse. We find a a Feynman–Kac type theorem showing that the stochastic representation yields a classical solution to the corresponding Black–Scholes equation *with appropriate boundary conditions*, compare Theorem 5.5. We also obtain additional regularity results in the one-dimensional case.

Date: July 2, 2004; revised October 28, 2004.

Key words and phrases. Parabolic equations; stochastic representation.

* Partially supported by the Swedish Research Council.

One should note that in the standard theory of parabolic equations, boundary regularity is only obtained for operators that are uniformly parabolic near the boundary. In fact, basic results from that theory fail in this more general setting of operators degenerating at the boundary. For instance, the standard regularity result saying that if the initial (or terminal condition in a financial setting) condition is continuous and if the operator has smooth coefficients the solution is smooth for any positive time (or any time strictly before expiration), fails. This is seen by considering contract functions of the form x^α and the stock price modelled by geometric Brownian motion. Here the solution will have the form $f(t)x^\alpha$, for some function f of time, see Example 6.4, and this solution is not smooth if α is not a non-negative integer. Another example is the Hopf boundary point lemma, see page 10 of [7]. This lemma says that at a boundary minimum of a solution to a parabolic equation the inner normal derivative must be strictly positive. This fails for instance in such a well-known example as the Black–Scholes formula for the call option: the inner normal derivative at the origin is zero even though this is a minimum point for the option price. In this formula the stock price is modelled by geometric Brownian motion, so the corresponding parabolic operator does degenerate at the boundary. On the other hand, a standard tool such as the maximum principle, is still available in our setting. We will indeed use this fact below.

It is somewhat surprising how little attention that has been paid to the issues desribed above given the importance of the type of equation under consideration and the general mathematical interest of existence and regularity questions also for this class of operators. However, there are of course references dealing with this type of problem. A general treatment can be found in Chapter 15 of Friedman [2] and an example closer to the present article is given by the work of Heath and Schweizer [3]. In the latter reference the processes are assumed to reach the boundary with zero probability and in the book by Friedman [2] the coefficients of the equation are assumed to be continuous up to and including the boundary. In contrast to these references, we will allow processes that reach the boundary with positive probability as well as coefficients that are not continuous at the boundary, compare Section 3.

Our result should be applicable also in biology and chemistry when modeling systems on a meso scale, compare [4], and more generally whenever one uses systems of stochastic processes of non-negative values. Stock prices should then be replaced by the number of molecules of various compounds. What is crucial for the results of this paper to hold is the absorbing property of the boundary, i.e. once a compound has vanished it does not reappear, and further that on the boundary the process is governed by the number of remaining molecules and their stochastic properties.

Finally we remark that the notation we use for stochastic processes is influenced by the various needs we have in the article but also by the standard notation in the theory of stochastic processes and financial applications, respectively. Thus, when noting that the process X depends on the time variable t we write X_t. However, when the point in time is of some special significance, such as the expiration date T of an option, we write $X(T)$, as is common in finance. Also, when noting that the process X at time t, say, is at some specified point x, we write $X_{x,t}$.

Acknowledgement. We thank a referee for careful reading and insightful criticism.

2. Classical and stochastic solutions

In this section we collect some general, largely known, facts on classical and stochastic solutions to linear parabolic partial differential equations. Let $\Omega = B \times (0,T)$ where B is a domain in $\mathbb{R}^n$. Consider the equation

$$\frac{\partial F}{\partial t} + \mathcal{L}F = 0, \tag{1}$$

where

$$\mathcal{L} = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x,t)\frac{\partial}{\partial x_i} + c(x,t). \tag{2}$$

Here (a_{ij}) is a positive definite symmetric matrix, and b_i and c as well as the entries of (a_{ij}) are continuous functions on Ω. Let $C^{2,1}(\Omega)$ be the space of functions with two continuous spatial derivatives and one continuous time-derivative on the indicated set, and let $C(\overline{\Omega})$ be the space of continuous functions on the closure of Ω.

Definition 2.1. *A function $F \in C^{2,1}(\Omega) \cap C(\overline{\Omega})$ satisfying equation* (1) *in Ω is called a classical solution.*

To introduce the corresponding stochastic solution we let X_t be an $\mathbb{R}^n$-valued stochastic process with each component X^i solving

$$X_t^i = x_0^i + \sum_{j=1}^{n}\int_{t_0}^{t} \sigma_{ij}(X_s, s)dW_s^j + \int_{t_0}^{t} b_i(X_s,s)ds, \tag{3}$$

where $W^j, j = 1,\ldots,n$, are independent Wiener processes. For these problems to correspond we have to choose the matrix (a_{ij}) in the operator of equation (2) to be $a = \sigma\sigma^*$ with the superscript $*$ denoting transpose. We will assume that $\sigma = (\sigma_{ij})$ and $b = (b_i)$ are, apart from the continuity in mentioned above, locally Lipschitz in the spatial variables in Ω to ensure existence and uniqueness of the solution, see [10, Chapter IX: (2.4), (2.10)]. In one spatial dimension, the condition on σ can be relaxed to a Hölder($\frac{1}{2}$) condition, see [10, Chapter IX:(3.5)] and [5, Section 5.5]. Further, we will make the standard assumptions

$$|\sigma(x,t)| + |b(x,t)| \le D(1+|x|) \tag{4}$$

to avoid exploding solutions, see [10, Chapter IX: (2.10)]; we will also assume $|c| \le D$ and that c is locally Hölder in the spatial coordinates for some positive exponent.

We should note that by Friedman [2, Lemma 6.1.1], a locally Lipschitz (or Hölder($\frac{1}{2}$)) σ exists if (a_{ij}) is locally Lipschitz (or Hölder($\frac{1}{2}$), respectively) and positive definite.

Let $\tau \le T$ be the first exit time of the process X from the set Ω. Then we are ready to define the stochastic solution to equation (1).

Definition 2.2. *A stochastic solution to equation* (1) *is a function F defined on $\overline{\Omega}$ satisfying, with the notation introduced above,*

$$F(x_0, t_0) = E e^{\int_{t_0}^{\tau} c(X_s,s)\,ds} F(X_\tau, \tau), \tag{5}$$

for every $(x_0, t_0) \in \Omega$.

Note that in general, we may have both $\tau = T$ and $\tau < T$ with positive probability; the latter means that X_t hits the boundary ∂B before time T. The formula (5) thus corresponds to an initial-boundary problem in classical differential equation theory.

We have the following lemma.

Lemma 2.3. *If F is a classical solution then, with the notation introduced above,*

$$e^{\int_{t_0}^{t\wedge\tau} c(X_s,s)\,ds} F(X_{t\wedge\tau}, t \wedge \tau) \tag{6}$$

is a local martingale.

Proof. Follows by applying Itô's formula. □

From this the following theorems follow.

Theorem 2.4. *If F is a classical solution and Ω is bounded then F is a stochastic solution.*

Proof. The result follows since a bounded local martingale is a martingale. □

Theorem 2.5. *Let $\Omega = \mathbb{R}_+^n \times (0, T)$ and assume that* (4) *holds. If F is a classical solution in Ω that is polynomially bounded, i.e. $|F(x, t)| \le C(1 + |x|)^m$ for some C and m, then F is a stochastic solution.*

Proof. Let τ_M be the stopping time $\inf\{t : |X_t| \ge M\}$, with $\tau_M = T$ if no such t exists. By Lemma 5.1, which holds also in the present generality with coefficients b by the same proof, $E|X_{\tau_M}|^{m+1} \le C_1$ for some constant C_1, and thus, for $M > 0$

$$P(\sup_{t\le T} |X_t| \ge M) = P(|X_{\tau_M}| \ge M) \le M^{-m-1} E|X_{\tau_M}|^{m+1} \le C_1 M^{-m-1}.$$

Consequently, $\sup_t |F(X_t, t)| \le C \sup_t (1 + |X_t|)^m$ is integrable, and thus the local martingale in (6) is a martingale. □

Before proceeding we need the following lemma on the existence of classical solutions which we formulate precisely as it is needed here. This existence is not immediate from standard Schauder theory since the most common assumption on the coefficients in this theory is Hölder continuity in space and time whereas we require only continuity in time. However, the necessary estimates are available also in our setting, see [1] and [6]. Our lemma follows directly from a related result in [8] and we therefore only sketch the argument.

Lemma 2.6. *Using the notation and assumptions of this section, consider a cylinder $\Omega_1 = B_1 \times (t_1, t_2)$ with $\overline{\Omega_1} \subset \Omega$ and let ϕ be a continuous function on the parabolic boundary $\partial_p \Omega_1$. Then there is a unique classical solution to equation (1) in Ω_1 with boundary values ϕ.*

Proof. According to [8, Theorem 16.1] there is a unique classical solution if ϕ is smooth, i.e. the restriction to $\partial_p\Omega_1$ of a smooth function on a neighborhood of $\overline{\Omega_1}$. Otherwise we approximate ϕ uniformly with such smooth functions. The sequence of solutions to (1) with these approximating smooth functions as boundary data converge uniformly to a solution of (1) by the maximum principle and interior Schauder estimates [6, Theorem 1], compare also [8, Theorem 15.6]. □

We also have the following partial converse of the previous theorems. Note that we assume the solution to be continuous, so it remains to establish this assumption to obtain a complete converse. Our main result is Theorem 5.5 which establishes this in the setting described in Section 3.

Theorem 2.7. *Assume that F is a continuous stochastic solution. Then F is a classical solution.*

Proof. Consider a cylinder $\Omega_1 = B_1 \times (t_1, t_2)$ with $\overline{\Omega_1} \subset \Omega$ and assume $(x_0, t_0) \in \Omega_1$. Let τ_1 be the hitting time of $\partial\Omega_1$. By the strong Markov property, see [5, Theorem 5.4.20],

$$\begin{aligned} E\big(e^{\int_{t_0}^{\tau} c(X_s,s)\,ds} F(X_\tau,\tau)|\mathcal{F}_{\tau_1}\big) &= e^{\int_{t_0}^{\tau_1} c(X_s,s)\,ds} E\big(e^{\int_{\tau_1}^{\tau} c(X_s,s)\,ds} F(X_\tau,\tau)|\mathcal{F}_{\tau_1}\big) \\ &= e^{\int_{t_0}^{\tau_1} c(X_s,s)\,ds} F(X_{\tau_1},\tau_1) \end{aligned} \tag{7}$$

Refering to the lemma above, we let $h \in C^{2,1}(\Omega_1)$ be the classical solution in Ω_1 with boundary values F on the parabolic boundary $\partial_p\Omega_1$. By Theorem 2.4, h is also the stochastic solution in Ω_1, so by equation (7),

$$\begin{aligned} h(x_0,t_0) &= E\big(e^{\int_{t_0}^{\tau_1} c(X_s,s)\,ds} F(X_{\tau_1},\tau_1)\big) \\ &= E\big(e^{\int_{t_0}^{\tau} c(X_s,s)\,ds} F(X_\tau,\tau)\big) = F(x_0,t_0). \end{aligned} \tag{8}$$

Thus $F = h$ in Ω_1; in particular $F \in C^{2,1}(\Omega_1)$ and $\mathcal{L}F = \mathcal{L}h = 0$ in Ω_1 and thus everywhere in Ω. □

3. Stochastic representation formulas

We consider a market consisting of a bank account with price process

$$B(t) = B(0)\exp\big\{\int_0^t r(u)\,du\big\},$$

where the interest rate r is a deterministic continuous function, and n risky assets, with the price X^i of the ith asset satisfying the stochastic differential equation

$$dX_t^i = r(t)X_t^i\,dt + \sum_{j=1}^n \sigma_{ij}(X_t,t)\,dW_t^j \tag{9}$$

before some given time horizon $T > 0$ under some risk neutral measure $\mathbb{Q}$, where X^i is absorbed at zero. Using the notation of the previous section we have $\Omega = \mathbb{R}_+^n \times (0,T)$ and we make the corresponding regularity assumptions on the coefficients. Here $\mathbb{R}_+^n = \{(x_1,\ldots,x_n) : x_i > 0, 1 \le i \le n\}$ and thus its closure $\overline{\mathbb{R}_+^n}$ is instead described by the inequalities $x_i \ge 0, 1 \le i \le n$. In this case, however, we make some additional assumptions on the coefficients since we need them to be defined also on the boundary, compare our discussion below. More precisely we will assume the following.

Hypothesis 3.1. *Each σ_{ij} is defined and continuous in the time variable on $\overline{\mathbb{R}^n_+} \times [0,T]$ and is Lipschitz continuous in the spatial variables on every compact subset of $\{(x,t) \in \overline{\mathbb{R}^n_+} \times [0,T] : x_i > 0\}$. On the set where $\{x_i = 0\}$, we have $\sigma_{ij} = 0$ for all $1 \le j \le n$. The rank of the matrix $\sigma_{ij}(x,t)$, at each point (x,t), is equal to the number of non-zero spatial coordinates.*

Note that this means that we allow discontinuities of the coefficients at the boundary. For instance, we include such cases as Brownian motion absorbed at zero ($\sigma_{11}(x,t) = 1$ for $x > 0$, but $\sigma_{11}(0,t) = 0$). Also Lipschitz continuity need not hold up to the boundary as we allow examples such as $\sigma_{11} = x^\alpha$ for $0 < \alpha < 1$. The assumption on the rank of the matrix σ_{ij} is standard in financial applications and is related to the absence of arbitrage; it is equivalent to parabolicity of the corresponding partial differential equation, see Section 4. For an illustration of the generality of models which is allowed under Hypothesis 3.1, see Example 5.6.

Now, let $g : \overline{\mathbb{R}^n_+} \to \mathbb{R}$ be continuous and of at most polynomial growth. Standard arbitrage theory yields that the price at time t of the option which at time T pays $g(X(T))$ is $F(X(t),t)$, where

$$F(x,t) = \exp\Big\{-\int_t^T r(u)\,du\Big\} E_{x,t} g(X(T)). \tag{10}$$

Here E denotes expected value with respect to the some so called risk neutral measure $\mathbb{Q}$ and the indices indicate that $X_t = x$. The dynamics of the processes X^i are those of equation (9). Let us elaborate on the interpretation of equation (10) in view of Hypothesis 3.1. If we start the process X at some interior point x at time t, it may at some time hit an $(n-1)$-dimensional face $\{x_i = 0\}$. It will then continue in this face, since X^i is absorbed at 0, by a stochastic differential equation in the remaining variables. Another coordinate may hit 0 at a later time, and X then continues in this lower dimensional face and so on until time T. The stochastic solution (10) obtained in this way can thus evidently also be otained by first finding the solution for the zero dimensional face $\{0\}$, which amounts to solving an ordinary differential equation, and then inductively solving the equation for higher-dimensional faces by Definition 2.2, using some of the previously obtained solutions as boundary values.

4. A partial differential equation formulation

The pricing function F should solve the Black–Scholes parabolic differential equation

$$\frac{\partial F}{\partial t} + \mathcal{L}F = 0, \tag{11}$$

where

$$\mathcal{L} = \frac{1}{2}\sum_{i,j=1}^n a_{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n r x_i \frac{\partial}{\partial x_i} - r, \tag{12}$$

with terminal condition

$$F(x,T) = g(x).$$

In this equation the coefficients $a_{ij} = a_{ij}(x,t)$ are the entries of the $n \times n$-matrix $\sigma\sigma^*$. Note that the assumption of full rank of σ guarantees parabolicity of the equation (11) in $\mathbb{R}^n_+$ (since the direction of the time variable is opposite to the customary one). We also see, however, that a_{kl} vanishes, if k or l agrees with i, on the set $\{x : x_i = 0\} \cap \partial\mathbb{R}^n_+$ from the assumptions on the diffusion matrix. The restriction of the operator to these faces is parabolic, due to the assumptions on the rank of the diffusion matrix. Moreover, by Friedman [2], Lemma 6.1.1, the Lipschitz conditions on σ_{ij} can be translated to the same conditions on the coefficients a_{ij} as mentioned above. We can also define a classical solution inductively in the same way as described in our discussion at the end of the previous section. This is the same as demanding that $F \in C(\mathbb{R}^n_+ \times [0,T])$ is a classical solution in the sense of Definition 2.1.

Theorem 4.1. *There is at most one solution to Black–Scholes equation, as specified above, of polynomial growth.*

Proof. We solve the equation on each face of the boundary starting with the ordinary differential equation on the Cartesian product of the origin and a time-ray. The maximum principle is valid as long as the equation is parabolic in the interior of each face. Applying the maximum principle on each face and finally on $\mathbb{R}^n_+$ gives the desired uniqueness. □

5. Continuity at the boundary

To establish continuity at the boundary of the stochastic solution we need the following lemmas. The first is well-known and is included for the sake of convenience of the reader. Lemmas 5.3 and 5.4 imply that if $(y,u) \to (x,t)$, then $X_{y,u}(T) \to X_{x,t}(T)$ in probability. This is known in the case of processes on $\mathbb{R}^n$, see e.g. [9, Theorem V.37]. However, the presence of an absorbing boundary does seem to entail additional technical difficulties under our quite general conditions on the coefficients σ_{ij}. (Presumably the convergence holds almost surely too, as for the corresponding equations in $\mathbb{R}^n$, but we will need only this weaker result.)

We change notation and let $X_{x,t}(u)$ be the solution X_u to (9), starting at time t at the point x.

In the proofs below we assume for simplicity that $r(t) = 0$; thus (9) simplifies and shows that X is a local martingale (and by Lemma 5.1 below a martingale). This can be done without loss of generality, since we in general can transform to the forward prices $\exp\{\int_t^T r(u)\,du\}X(t)$.

Lemma 5.1. *Suppose that the coefficients σ_{ij} satisfy* (4). *For every $k \geq 0$ and A there exists a constant C, depending only on n, T, k, A and the constant in* (4), *such that if $|x| \leq A$, then $E|X_{x,0}(\tau)|^k \leq C$ for every stopping time $\tau \leq T$.*

Proof. By stopping at $\inf\{t : |X_{x,0}(t)| \geq M\}$, we may assume that $X_{x,0}(t \wedge \tau)$ is bounded; the result then follows by letting $M \to \infty$. If $m \geq 0$ is an integer, let

$$f(t) = E(1 + |X_{x,0}(t \wedge \tau)|^2)^m.$$

Itô's lemma and equation (3) easily yield

$$\begin{aligned} f(t) &\leq f(0) + \\ &\quad C_1 \sum_{i,j,l} E \int_0^{t\wedge\tau} |\sigma_{ij}(X_{x,0}(s),s)\sigma_{lj}(X_{x,0}(s),s)|(1+|X_{x,0}(s)|^2)^{m-1}\,ds \\ &\leq f(0) + C_2 E \int_0^{t\wedge\tau} (1+|X_{x,0}(s)|^2)^m\,ds \\ &\leq (1+|x|^2)^m + C_2 \int_0^t f(s)\,ds, \end{aligned}$$

and the result follows by Gronwall's lemma, see e.g. [10, Appendix §1]. □

We also need the following.

Lemma 5.2. *Suppose that the coefficients σ_{ij} satisfy (4) and Hypothesis 3.1. For every ε, η and $A > 0$, there exists a $\delta > 0$, depending only on $n, T, \varepsilon, \eta, A$, the constant in (4) and the Lipschitz constants of σ_{ij} on compact sets, such that if k is an integer with $0 \leq k \leq n$ and x, y satisfy $0 \leq x^i, y^i \leq 2\delta$ for $1 \leq i \leq k$, and $|x^i - y^i| \leq \delta$ and $x^i \leq A$ for $1 \leq i \leq n$, and further $\tau \leq T$ is a stopping time such that $X^i_{x,0}(t), X^i_{y,0}(t) \geq \varepsilon$ for $0 \leq t \leq \tau$ and $k < i \leq n$, then with probability at least $1-\eta$,*

$$X^i_{x,0}(\tau), X^i_{y,0}(\tau) \leq \varepsilon/2, \tag{13}$$

for $1 \leq i \leq k$ and

$$|X^i_{x,0}(\tau) - X^i_{y,0}(\tau)| \leq \varepsilon, \tag{14}$$

for $1 \leq i \leq n$.

Proof. Since $X_{x,0}$ and $X_{y,0}$ are martingales, we have for $i \leq k$ that $EX^i_{x,0}(\tau) = x^i \leq 2\delta$ and thus $P(X^i_{x,0}(\tau) > \varepsilon/2) \leq 4\delta/\varepsilon$ and similarly for $X^i_{y,0}$. Hence, if we choose $\delta < \eta\varepsilon/16n$, inequality (13) will fail with probability less than $\eta/2$. Note that (13) implies (14) for $i \leq k$. Now let $A' = 8n(A+\delta)/\eta$. By Doob's inequality, for each i,

$$P(\sup_{t\leq\tau} X^i_{x,0}(t) > A') \leq x^i/A' \leq \eta/8n,$$

and similarly for $X^i_{y,0}(t)$, so if we let τ' be the first time before τ, if any, such that $X^i_{x,0}(t)$ or $X^i_{y,0}(t) \geq A'$ for some i, then $P(\tau' \neq \tau) \leq \eta/4$. We may thus replace τ by τ'; for convenience we write τ and assume that $X^i_{x,0}(t), X^i_{y,0}(t) \leq A'$ for $t \leq \tau$ and every i; further $X^i_{x,0}(t), X^i_{y,0}(t) \geq \varepsilon$ for $i > k$ by assumption. Our local Lipschitz assumption thus yields, for some C_1,

$$|\sigma_{ij}(X_{x,0}(t),t) - \sigma_{ij}(X_{y,0}(t),t)| \leq C_1|X_{x,0}(t) - X_{y,0}(t)| \tag{15}$$

for $i > k$. Let $Z = X_{x,0} - X_{y,0}$. Then

$$Z^i(t\wedge\tau) = x^i - y^i + \sum_j \int_0^{t\wedge\tau} \big(\sigma_{ij}(X_{x,0}(s),s) - \sigma_{ij}(X_{y,0}(s),s)\big)dW_j(s)$$

and thus, using inequality (15), for $i > k$,

$$\begin{aligned} E|Z^i(t\wedge\tau)|^2 &= (x^i-y^i)^2 + \\ &\qquad \sum_j \int_0^{t\wedge\tau} |\sigma_{ij}(X_{x,0}(s),s) - \sigma_{ij}(X_{y,0}(s),s)|^2\,ds \qquad (16) \\ &\le \delta^2 + C_2\int_0^{t\wedge\tau} |Z(s)|^2\,ds. \end{aligned}$$

For $i \le k$ we use

$$E|Z^i(t\wedge\tau)| \le E(X_{x,0}(t\wedge\tau) + X_{y,0}(t\wedge\tau)) = x^i + y^i \le 4\delta$$

which by Hölder's inequality and Lemma 5.1 yields

$$E|Z^i(t\wedge\tau)|^2 \le (E|Z^i(t\wedge\tau)|)^{2/3}(E|Z^i(t\wedge\tau)|^4)^{1/3} \le C_3\delta^{2/3}. \qquad (17)$$

Hence, if $f(t) = E|Z^i(t\wedge\tau)|^2$, we find by summing (17) for $i \le k$ and (16) for $i > k$, assuming $\delta < 1$,

$$f(t) \le C_4\delta^{2/3} + C_5\int_0^t f(s)\,ds,$$

which by Gronwall's lemma yields

$$f(t) \le C_4\delta^{2/3}e^{C_5 t}.$$

Choosing δ so small that $C_4\delta^{2/3}e^{C_5 T} = \eta\varepsilon^2/2$ we thus find $E|Z(\tau)|^2 = f(T) \le \eta\varepsilon^2/2$ and thus $P(|Z(\tau)| > \varepsilon) < \eta/2$, which proves that (14) fails with probability less than $\eta/2$. ☐

Lemma 5.3. *Suppose that the coefficients σ_{ij} satisfy* (4) *and Hypothesis 3.1. For every ε, η and $A > 0$, there exists a $\delta > 0$, depending on the same parameters as in Lemma 5.2 only, such that if $|x-y| \le \delta$, $|x| \le A$ and $0 \le t \le T$, then $P(|X_{x,t}(T) - X_{y,t}(T)| > \varepsilon) < \eta$.*

Proof. We may assume that $t = 0$. Let $A' = 2n(n+1)A/\eta$, and note that since each $X^i_{x,0}$ is a martingale with mean $x^i \le A$, we have

$$\sum_i P(X^i_{x,0}(\tau) > A') < \eta/2(n+1)$$

for every stopping time τ. We now define a sequence of small numbers $0 < \delta_0 < \delta_1 < \cdots < \delta_{n+1} = \varepsilon/2n$ by backwards recursion: for $k = n,\dots,0$ we let δ_k be the δ given by Lemma 5.2 if (ε,η,A) is replaced by $(\delta_{k+1},\eta/2(n+1),A')$. We let $\delta = \delta_0$. Next, we define a sequence of stopping times $0 = \tau_0 \le \tau_1 \le \cdots \le \tau_{n+1} = T$ by letting $\tau_0 = 0$ and, for $k = 1,2,\dots,n+1$, by letting τ_k be the first time $t \ge \tau_{k-1}$ such that $\inf_{s\le t} X^i_{x,0}(s) \wedge X^i_{y,0}(s) \le \delta_k$ for at least k indices i; if no such t exists we put $\tau_k = T$. We apply Lemma 5.2 inductively to each interval $[\tau_k,\tau_{k+1}]$, conditioning on $\mathcal{F}_{\tau_k}$, i.e. on everything that has happened up to τ_k, and making an obvious time shift. We assume that $|X^i_{x,0}(\tau_k) - X^i_{y,0}(\tau_k)| \le \delta_k$ and $X^i_{x,0}(\tau_k) \le A'$ for every i, and that $X^i_{x,0}(\tau_k), X^i_{x,0}(\tau_k) \le 2\delta_k$ for at least k indices i; we may assume that this holds for $i = 1,\dots,k$.

By the definition of τ_{k+1}, $X^i_{x,0}(t), X^i_{y,0}(t) > \delta_{k+1}$ for $i > k$ and $\tau_k \le t < \tau_{k+1}$, and thus Lemma 5.2 shows that with probability at least $1-\eta/2(n+1)$ we have $|X^i_{x,0}(\tau_{k+1}) - X^i_{y,0}(\tau_{k+1})| \le \delta_{k+1}$ for every i and $X^i_{x,0}(\tau_{k+1})$,

$X^i_{y,0}(\tau_{k+1}) \leq \delta_{k+1}/2$ for $i \leq k$; further if $\tau_{k+1} < T$, then $X^i_{x,0}(\tau_{k+1}) \wedge X^i_{y,0}(\tau_{k+1}) = \delta_{k+1}$ and thus $X^i_{x,0}(\tau_{k+1}), X^i_{y,0}(\tau_{k+1}) \leq 2\delta_{k+1}$ for at least one other i. Finally, as observed above, $X^i_{x,0}(\tau_{k+1}) \leq A'$ for all i with probability at least $1 - \eta/2(n+1)$.

Hence, given that the assumptions above hold for k, they hold for $k+1$ too, with an error probability of at most $\eta/(n+1)$. These assumptions hold for $k = 0$, and thus, with probability at least $1 - \eta$ for $k = n+1$, yielding $|X^i_{x,0}(T) - X^i_{y,0}(T)| \leq \delta_{k+1} = \varepsilon/n$ for each i and thus $|X_{x,0}(T) - X_{y,0}(T)| \leq \varepsilon$. □

To show continuity in time of the stochastic solution we need the following result.

Lemma 5.4. *Suppose that the coefficients σ_{ij} satisfy* (4) *and Hypothesis 3.1. For every ε, η, and $A > 0$ there exists a $\delta > 0$, with the same parameter dependence as above, such that if $|x| \leq A$ and $|t - u| \leq \delta$, with $0 \leq t, u \leq T$, then*

$$P(|X_{x,t}(T) - X_{x,u}(T)| > \varepsilon) < \eta.$$

Proof. We may assume that $0 \leq u \leq t \leq T$. Then, using Lemma 5.1,

$$\begin{aligned} E|X_{x,u}(t) - x|^2 &= \sum_{i,j} E \int_u^t |\sigma_{ij}(X_{x,u}(s), s)|^2 \, ds \leq C_1 E \int_u^t (1 + |X_{x,u}(s)|^2) \, ds \\ &\leq C_2(t - u) \leq C_2\delta. \end{aligned} \tag{18}$$

Lemma 5.3 shows, by conditioning on $X_{x,u}(t)$, that for some $\delta_0 > 0$, the conditional probability $P\big(|X_{x,t}(T) - X_{x,u}(T)| > \varepsilon \;\big|\; |X_{x,u}(t) - x| \leq \delta_0\big) < \eta/2$. If we choose $\delta = \eta\delta_0^2/2C_2$, then (18) yields

$$P(|X_{x,u}(t) - x| > \delta_0) < C_2\delta/\delta_0^2 = \eta/2,$$

and the result follows. □

Lemmas 5.3 and 5.4 imply that if $(y, u) \to (x, t)$ then $X_{y,u}(T) \to X_{x,t}(T)$ in probability.

Theorem 5.5. *Suppose that the coefficients σ_{ij} satisfy* (4) *and Hypothesis 3.1 and that the function $g : \overline{\mathbb{R}^n_+} \to \mathbb{R}$ is continuous and of polynomial growth. Then the stochastic solution in* (10) *is continuous and is thus a classical solution to* (11)*. Moreover, this solution is of polynomial growth.*

Proof. Lemmas 5.3 and 5.4 and the continuity of g imply that if $(y, u) \to (x, t)$ in $\Omega = \overline{\mathbb{R}^n_+} \times [0, T]$, then $g(X_{y,u}(T)) \to g(X_{x,t}(T))$ in probability. Further, $Eg(X_{y,u}(T))^2 \leq C^2 E(1 + |X_{y,u}(T)|)^{2m}$ is bounded (for $|y| \leq |x| + 1$, say) by Lemma 5.1, where we use the fact that g satisfies some bound $|g(x)| \leq C(1 + |x|)^m$. Hence the family $\{g(X_{y,u}(T))\}$ is uniformly integrable and $Eg(X_{y,u}(T)) \to Eg(X_{x,t}(T))$. This proves that the stochastic solution is continuous in Ω and thus it is a classical solution by Theorem 2.7. Since the same argument holds in any face of $\mathbb{R}^n_+$, the results extend to the solutions defined as in Section 3 by induction over faces. Estimates as in the proof of Lemma 5.1 show that the solution is of polynomial growth. □

We end this section with an example, of a type that has not been studied rigorously previously, illustrating the theorem above. For the sake of notational simplicity, X and Y will denote $\mathbb{R}$-valued processes below.

Example 5.6. We consider a market, where again for the sake of simplicity it is assumed that the interest rate $r(t) = 0$, with two risky assets X and Y modeled by

$$dX = c_{11} X dW^1 + C_{12} X^\alpha Y dW^2,$$

where $0 < \alpha < 1$ and

$$dY = c_{22}(\varepsilon + Y) dW^2,$$

where ε is some positive number and c_{ij} are arbitrary constants, with c_{11} and c_{22} non-zero. The conditions of Hypothesis 3.1 are satisfied since $\sigma_{11} = c_{11} X$ and $\sigma_{12} = c_{12} X^\alpha Y$ are locally Lipschitz continuous on the set $\{(x, y, t) \in \overline{\mathbb{R}^2_+} \times [0, T] : x > 0\}$. The coefficient $\sigma_{22} = c_{22}(\varepsilon + Y)$ is obviously Lipschitz on the set where $y > 0$. Note, however, that σ_{22} is assumed to be zero where y vanishes, corresponding to Y being absorbed, and thus this coefficient is not continuous up to the part of the boundary where $y = 0$ which indeed is allowed by Hypothesis 3.1.

Let $F(x, y, t)$ be the option price given by (10) for some contract function $g(x, y)$ of polynomial growth. According to the theorem above F is the unique (unicity by Theorem 4.1) classical solution to the corresponding Black-Scholes equation (11) with g as terminal condition and with appropriate boundary conditions. Let us discuss these boundary condition in some detail. Since the interest rate is assumed to be zero, $F(0, 0, t) = g(0, 0)$. Furthermore, on set $\{(x, y, t) \in \overline{\mathbb{R}^2_+} \times [0, T] : y = 0\}$, $F(x, 0, t)$ is the unique classical solution of polynomial growth of

$$\frac{\partial F}{\partial t} + \frac{1}{2} c_{11}^2 x_1^2 \frac{\partial^2 F}{\partial x^2} = 0$$

with boundary condition $F(0, 0, t) = g(0, 0)$ and of course terminal condition $g(x, 0)$. In this case the boundary condition $F(0, 0, t) = g(0, 0)$ is redundant to specify; in fact there exists a solution *only* with this boundary condition. Note that second Wiener process, and in particular the value of α, in the dynamics of X does not influence the equation on this set. Similarly, $F(0, y, t)$ is the unique classical solution of polynomial growth of

$$\frac{\partial F}{\partial t} + \frac{1}{2} c_{22}^2 (\varepsilon + y)^2 \frac{\partial^2 F}{\partial y^2} = 0$$

with boundary condition $F(0, 0, t) = g(0, 0)$ and terminal condition $g(0, y)$. Here the boundary condition $F(0, 0, t) = g(0, 0)$ is essential: the equation is non-singular and thus solvable with *any* continuous boundary data.

6. Regularity in the one-dimensional case

In this section we assume $n = 1$, i.e. there is only one underlying asset. In this case, we can prove sharper results.

First, in one spatial dimension, the Lipschitz condition on σ can be relaxed to a Hölder($\frac{1}{2}$) condition, see [10, §IX.3]. Furthermore, if $x \leq y$, then $X_{x,t}(u) \leq X_{y,t}(u)$ a.s. for all $u \in [t, T]$ since paths with different initial values may not cross because of pathwise uniqueness.

We can now relax the assumptions of Theorem 5.5.

Theorem 6.1. *Suppose that the coefficient σ satisfies* (4) *and Hypothesis 3.1 with the local Lipschitz condition relaxed to Hölder*(1/2) *and that the function $g : \overline{\mathbb{R}_+} \to \mathbb{R}$ is continuous and of polynomial growth. Then the stochastic solution in* (10) *is continuous and is thus a classical solution to* (11).

Proof. If $0 \le x \le y$, then

$$E|X_{x,t}(T) - X_{y,t}(T)| = E\big(X_{y,t}(T) - X_{x,t}(T)\big) = y - x$$

because $X_{x,t}$ and $X_{y,t}$ are martingales (assuming again $r = 0$) and $X_{x,t} \le X_{y,t}$. Hence a (stronger) version of Lemma 5.3 holds by Chebyshev's inequality. Thus Lemma 5.4 holds too, by the same proof, and the proof is completed as for Theorem 5.5. □

Remark. Keeping the notation of the theorem above, and using Theorem 4.1, we can phrase this result more concretely. The option price $F(x,t)$ given by the stochastic representation formula (10) is the unique classical solution, of polynomial growth, of the corresponding Black-Scholes equation (11), with terminal condition given by the contract function g and boundary condition given by

$$F(0,t) = e^{-\int_t^T r(u)du} g(0).$$

This has certainly been used for numerous special cases of contract functions and stock models, such as put options i.e. $g(x) = \max(K - x, 0)$, where K is the so-called strike price. In this case the boundary condition is evidently given by the discounted strike price. However, it has, to our knowledge, not previously been formulated or proved for general continuous contract functions and Hölder(1/2)-volatilities.

If we make additional assumptions on g we can say more about the boundary behavior of the solution. Let us consider two such results.

We have the following result on Lipschitz continuity of the stochastic solutions for contract functions $g : \overline{\mathbb{R}_+} \to \mathbb{R}$ that are uniformly Lipschitz, i.e. satisfying

$$|g(x) - g(y)| \le K|x - y|$$

for some constant K and all $x, y \ge 0$.

Proposition 6.2. *Suppose that the coefficient σ satisfies the assumptions of the previous theorem. Further, let the contract function $g : \overline{\mathbb{R}_+} \to \mathbb{R}$ be uniformly Lipschitz continuous with Lipschitz constant K. Then the option price F as given by equation* (10) *is Lipschitz continuous in the spatial variable with the same Lipschitz constant.*

Proof. Fix $t < T$. Then

$$\begin{aligned} |F(x,t) - F(y,t)| &= \exp\Big\{-\int_t^T r(u)\,du\Big\}|Eg(X_{x,t}(T)) - Eg(X_{y,t}(T))| \\ &\le K\exp\Big\{-\int_t^T r(u)\,du\Big\}E|X_{x,t}(T) - X_{y,t}(T)| \\ &= K|x - y| \end{aligned}$$

where the equality follows from the martingale property of discounted asset prices and the fact that if $x \geq y$ then $X_{x,t} \geq X_{y,t}$ and therefore

$$E|X_{x,t}(T) - X_{y,t}(T)| = E(X_{x,t}(T) - X_{y,t}(T)), \tag{19}$$

and that if instead $x \leq y$ the equality holds with a right-hand-side of the opposite sign. □

Remark. In financial terminology this means that the delta of the option is bounded by the delta of the contract function. The argument above does not carry over to several underlying assets, since equation (19) is not generally true in that case.

We have a related result for Hölder(α) continuity for $\alpha < 1$. We say that a function $g : \overline{\mathbb{R}_+} \to \mathbb{R}$ is *locally Hölder continuous of polynomial growth* if the following estimate holds in $\overline{\mathbb{R}_+}$ for some constant C, some integer m and some number $0 < \alpha < 1$,

$$|g(x) - g(y)| \leq C(1 + x^m + y^m)|x - y|^\alpha. \tag{20}$$

We have the following continuity result of the price function for contracts of this type.

Proposition 6.3. *Let the coeffient σ satisfy the same assumptions as in Theorem 6.1. Further, let the contract function $g : \overline{\mathbb{R}_+} \to \mathbb{R}$ be locally Hölder continuous of polynomial growth as defined above with exponent α. Then the option price F given by the stochastic representation formula* (10) *(or equivalently the equation* (11)*) is locally Hölder(α) continuous in $\overline{\mathbb{R}_+}$, i.e., for every A there is a C' such that if $x, y \leq A$, then*

$$|F(x,t) - F(y,t)| \leq C'|x - y|^\alpha.$$

Proof. We fix $t < T$, and assume again for convenience that $r = 0$. Then

$$|F(x,t) - F(y,t)| = |E\big(g(X_{x,t}(T)) - g(X_{y,t}(T))\big)|.$$

By (20) and the Hölder inequality, the right-hand-side of this equation is dominated by

$$C\big(E(1 + |X_{x,t}(T)|^m + |X_{y,t}(T)|^m)^q\big)^{1/q}\big(E|X_{x,t}(T) - X_{y,t}(T)|\big)^\alpha$$

where $q = \frac{1}{1-\alpha}$. Note that the first expected value is locally bounded as a function of x and y by Lemma 5.1. Finally, $E|X_{x,t}(T) - X_{y,t}(T)| = |x-y|$ as in the proof of Proposition 6.2 by monotonicity and the martingale property. □

Remark. In the proposition above we see that a bound for the modulus of continuity of the solution at the boundary is given by bound of the modulus of continuity of the contract function. This estimate is in fact sharp, as is seen by considering geometric Brownian motion, see the example below. Of course, in the non-degenerate case, with coefficients regular enough, we obtain regularity across the boundary, compare Chapter VII of [7], for $t < T$, even if the contract function is not regular at the boundary.

Example 6.4. We consider geometric Brownian motion:

$$dX = rXdt + \sigma XdW, \tag{21}$$

where r and σ are constants. We assume that $X(t) = x$. By considering $\ln X$ and using Itô's formula one finds the explicit solution

$$X_{x,t}(T) = x \exp\left\{(r - \frac{1}{2}\sigma^2)(T - t) + \sigma(W(T) - W(t))\right\}.$$

Thus, if $g(x) = x^\alpha$, the function F given by the stochastic representation formula (10) is equal to

$$F(x,t) = x^\alpha \exp\left\{(r(\alpha - 1) + \frac{1}{2}\sigma^2(\alpha^2 - \alpha))(T - t)\right\},$$

showing that the estimate of modulus of continuity in the proposition above is indeed sharp.

References

[1] Brandt, A.: Interior Schauder estimates for parabolic differential- (or difference-) equations via the maximum principle. *Israel J. Math.* **7** (1969), 254–262.

[2] Friedman, A.: *Stochastic Differential Equations and Applications, Vol. 1 and 2.* Academic Press, New York, 1976.

[3] Heath, D., Schweizer, M.: Martingales versus PDEs in finance: an equivalence result with examples. *J. Appl. Prob.* **37** (2000), 947–957.

[4] van Kampen, N. G.: *Stochastic processes in physics and chemistry.* North-Holland, Amsterdam, 1992.

[5] Karatzas, I., Shreve, S.: *Brownian Motion and Stochastic Calculus.* 2nd ed., Springer New York, 2000.

[6] Knerr, B. F.: Parabolic interior Schauder estimates by the maximum principle. *Arch. Rational Mech. Anal.* **75** (1980/81), 51–58.

[7] Lieberman, G.: *Second order parabolic differential equations.* World Scientific, Singapore, 1998.

[8] Lieberman, G.: Intermediate Schauder theory for second order parabolic equations. IV. Time irregularity and regularity. *Differential Integral Equations* **5** (1992), 1219–1236.

[9] Protter, P.: *Stochastic Integration and Differential Equations.* 2nd ed., Springer, Berlin, 2004.

[10] Revuz, D., Yor, M.: *Continuous martingales and Brownian motion.* 3rd ed., Springer, Berlin, 1999.

Department of Mathematics, Box 480, 751 06 Uppsala, Sweden