The Annals of Applied Probability
2011, Vol. 21, No. 1, 332–350
DOI: 10.1214/10-AAP698

BOUNDARY CONDITIONS FOR THE SINGLE-FACTOR TERM STRUCTURE EQUATION

BY ERIK EKSTRÖM[1,2] AND JOHAN TYSK[1]

Uppsala University

We study the term structure equation for single-factor models that predict nonnegative short rates. In particular, we show that the price of a bond or a bond option is the unique classical solution to a parabolic differential equation with a certain boundary behavior for vanishing values of the short rate. If the boundary is attainable then this boundary behavior serves as a boundary condition and guarantees uniqueness of solutions. On the other hand, if the boundary is nonattainable then the boundary behavior is not needed to guarantee uniqueness but it is nevertheless very useful, for instance, from a numerical perspective.

1. Introduction. When calculating prices of different interest rate derivatives, such as bonds and bond options, stochastic methods seem to be more commonly used than PDE methods; compare for instance [5]. This is in contrast to the case of stock option pricing where PDE methods are used extensively, in particular for low dimensional problems. We believe that one possible explanation for this phenomenon is that the correspondence between the risk neutral valuation approach and the pricing equation (henceforth referred to as the *term structure equation*) with appropriate boundary conditions is not fully developed.

For the Black–Scholes equation, the boundary condition is of Dirichlet type which corresponds to the underlying asset being absorbed if reaching zero; compare [13]. In contrast, for most interest rate models this is not the case since the short rate typically would not stay zero if the value zero is reached. Consequently, it is not clear what boundary conditions should be specified for the term structure equation. In fact, the recent monograph [6] draws attention to this issue in a section entitled "The thorny issue of boundary conditions." Moreover, in [12], two different solutions to the term structure equation are presented in the case of the CIR-model. They are both bounded and with the same terminal condition, but naturally they exhibit different boundary behavior for vanishing interest rates. The authors of that paper take the view that these solutions represent alternative possible prices. We on the other hand regard only the solution given by the stochastic

Received November 2008; revised December 2009.

[1]Supported by the Swedish Research Council (VR).

[2]Supported by the Swedish National Graduate School of Mathematics and Computing (FMB).

AMS 2000 subject classifications. Primary 91B28; secondary 35A05, 35K65, 60J60.

Key words and phrases. The term structure equation, degenerate parabolic equations, stochastic representation.

representation as the price, and the purpose of the present paper is to identify the boundary condition that this stochastic representation satisfies.

We consider the classical case of a single-factor model predicting nonnegative values of the short rate. More precisely, the rate $X(t)$ is modeled directly under the pricing measure as

$$dX(t) = \beta(X(t), t)\,dt + \sigma(X(t), t)\,dW,$$

where W is a Brownian motion and $\sigma(0, t) = 0$ and $\beta(0, t) \geq 0$. As indicated above, the option price u corresponding to a payoff function g is given using risk neutral valuation by

$$u(x, t) = E_{x,t}\big[e^{-\int_t^T X(s)\,ds} g(X(T))\big].$$

Note that if the payoff $g \equiv 1$, then bond prices are obtained. Also note that the set-up covers the case of bond options. The corresponding term structure equation is

$$u_t(x, t) + \tfrac{1}{2}\sigma^2(x, t)u_{xx}(x, t) + \beta(x, t)u_x(x, t) = xu(x, t)$$

with terminal condition $u(x, T) = g(x)$. If the price u is twice continuously differentiable up to and including the boundary $x = 0$, then plugging in $x = 0$ in the equation would give the boundary behavior

$$u_t(0, t) + \beta(0, t)u_x(0, t) = 0. \tag{1}$$

Even though there is extensive literature on equations with degenerating coefficients, compare the classical reference [14], the C^1-regularity of u at the boundary is not available in the generality that is needed here. In fact, one of the solutions in the example in [12] referred to above is bounded and continuous, but fails to be C^1 up to the boundary. (This solution, however, is not the one given by stochastic representation.)

In the present paper, sufficient regularity of the option price u for (1) to hold is established using the Girsanov theorem and scaling arguments; see Sections 3 and 4. Let us emphasize that (1) is the correct boundary behavior of the option price regardless if the boundary is hit with positive probability or not. If the boundary can be reached with positive probability, then this boundary behavior serves as a boundary condition for the term structure equation and guarantees uniqueness. On the other hand, if the boundary is reached with probability zero, equation (1) is not needed to identify the solution given by the stochastic representation, and the term "boundary condition" is perhaps misleading. However, it is still valid and certainly useful, for instance, from a numerical perspective. Indeed, the results of the present paper have already been implemented numerically in [7]. For simplicity of the exposition, we will refer to (1) as the boundary condition regardless if the boundary can be reached or not.

In Section 2 the assumptions on the model and our main result, Theorem 2.3, are presented. In Sections 3, 4 and 5, we establish regularity properties of the value function, and use them to prove Theorem 2.3. Finally, in Section 6 we also provide the link between the stochastic problem and the term structure equation for models defined on the whole real line.

2. Assumptions and the main result. When studying the term structure equation on the positive real axis, we consider models which are specified so that the short rate automatically stays nonnegative, that is, there is no need to impose any boundary behavior of the underlying diffusion process. Throughout Sections 2–5 we work under the following hypothesis:

HYPOTHESIS 2.1. *The drift* $\beta \in C([0,\infty) \times [0,T])$ *is continuously differentiable in* x *with bounded derivative, and* $\beta(0,t) \geq 0$ *for all* t. *The volatility* $\sigma \in C([0,\infty) \times [0,T])$ *is such that* $\alpha(x,t) := \frac{1}{2}\sigma^2(x,t)$ *is continuously differentiable in* x *with a Hölder continuous derivative, and* $\sigma(x,t) = 0$ *if and only if* $x = 0$. *The functions* β, σ *and* α_x *are all of, at most, linear growth*:

$$|\beta(x,t)| + |\sigma(x,t)| + |\alpha_x(x,t)| \leq C(1+x) \tag{2}$$

for all x *and* t. *The payoff function* $g : [0,\infty) \to [0,\infty)$ *is continuously differentiable with both* g *and* g' *bounded.*

Let W be a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$. Since α is continuously differentiable, σ is locally Hölder $(1/2)$ in x. It follows that there exists a unique strong solution $X(t)$ to

$$dX(t) = \beta(X(t),t)\,dt + \sigma(X(t),t)\,dW \tag{3}$$

for any initial point $x \geq 0$; compare Section IX.3 in [15]. Moreover, it follows from monotonicity results for stochastic differential equations with respect to the drift coefficient (e.g., Theorem IX.3.7 in [15]), that X remains nonnegative at all times. Indeed, if β is replaced with $\beta \wedge 0$, then the corresponding solution to (3) is absorbed at zero, so X is nonnegative if $\beta(0,t) \geq 0$. The option price $u : [0,\infty) \times [0,T] \to [0,\infty)$ corresponding to a payoff function $g : [0,\infty) \to [0,\infty)$ is given by

$$u(x,t) = E_{x,t}\big[e^{-\int_t^T X(s)\,ds} g(X(T))\big], \tag{4}$$

where the indices indicate that $X(t) = x$. As described in the Introduction, the corresponding term structure equation is given by

$$u_t(x,t) + \tfrac{1}{2}\sigma^2(x,t)u_{xx}(x,t) + \beta(x,t)u_x(x,t) = xu(x,t) \tag{5}$$

for $(x,t) \in (0,\infty) \times [0,T)$, with terminal condition

$$u(x,T) = g(x). \tag{6}$$

Moreover, by formally inserting $x = 0$ in the equation we get the boundary condition

$$u_t(0,t) + \beta(0,t)u_x(0,t) = 0 \tag{7}$$

for all $t \in [0, T)$, since $\sigma(0,t) = 0$ by assumption. One of the main efforts in this paper is to show that the option price u is continuously differentiable up to the boundary $x = 0$, and that it indeed satisfies the boundary condition (7) in the classical sense.

DEFINITION 2.2. A *classical solution* to the term structure equation is a function $v \in C([0,\infty) \times [0,T]) \cap C^1([0,\infty) \times [0,T)) \cap C^{2,1}((0,\infty) \times [0,T))$ which satisfies (5), (6) and (7).

Our main result in this article is the following:

THEOREM 2.3. *In addition to Hypothesis 2.1, also assume that Assumption 3.1 below holds. The option price u as given by (4) is then the unique bounded classical solution to the term structure equation.*

EXAMPLE. Classical short rate models such as the Cox–Ingersoll–Ross model

$$dX(t) = \big(a - bX(t)\big)\,dt + \sigma\sqrt{X(t)}\,dW, \tag{8}$$

and the Dothan model

$$dX(t) = aX(t)\,dt + \sigma X(t)\,dW, \tag{9}$$

have boundary conditions at $x = 0$ that are immediate to write down. These conditions are

$$u_t + au_x = 0 \quad \text{and} \quad u_t = 0,$$

respectively. We note that the boundary condition $u_t = 0$ for the Dothan model means that u is constant along the boundary, that is, $u(0,t) = g(0)$. This is the same type of boundary condition that appears for options on stocks in [13], which can be explained by the fact that the Dothan model is a geometric Brownian motion. Theorem 2.3 also covers the Hull–White model

$$dX(t) = \big(a(t) - b(t)X(t)\big)\,dt + \sigma(t)\sqrt{X(t)}\,dW \tag{10}$$

(which is a time-dependent generalization of the Cox–Ingersoll–Ross model), and models of, for example, the form

$$dX(t) = \big(b - aX(t)\big)\,dt + \sigma X^{\gamma}(t)\,dW, \qquad \gamma \in (1/2, 1], \tag{11}$$

which also would be natural to consider for bond pricing.

REMARK. It seems that many of the classical models for the short rate are proposed for their analytical tractability. In particular, if the drift β and the diffusion coefficient σ^2 are affine, then the model admits an affine term structure. It is easy to check that known explicit formulas for bond prices and bond options satisfy the boundary condition (7). In particular, for models admitting an affine term structure, it is a consequence of the associated Riccati equations (see [4], equation 22.25) that these boundary conditions are fulfilled.

REMARK. The assumption that g is continuously differentiable is satisfied for bonds, but not in general for bond options. However, using the Markov property, Theorem 2.3 readily extends to bounded Lipschitz payoffs provided one can show that the corresponding option price $x \mapsto u(x, T - \varepsilon)$ is continuously differentiable on $(0, \infty)$ for any $\varepsilon > 0$. The regularizing effect of parabolic equations guarantees continuous differentiability on $(0, \infty)$, so the main difficulty is to show that $x \mapsto u_x(x, T - \varepsilon)$ is continuous also at 0. If the model is convexity preserving, this is easily done in certain cases including, for example, call options written on bond prices. (Note that the corresponding payoff function g is bounded since bond prices are bounded.) For details on which short rate models are convexity preserving, see [8]. To our knowledge, all models used in practice belong to this class.

One should note that the differentiability of the option price up to the boundary $x = 0$ is not valid without some Lipschitz bound of g at 0. To see this, consider the contract function $g(x) = e^{-2\sqrt{x}}$. Then, with $\beta(x, t) = \frac{1}{2}x$ and $\sigma(x, t) = \sqrt{2}x$, it is straightforward using the Itô formula to show that the process

$$Y(s) = e^{-\int_t^s X(r)\,dr} g(X(s))$$

is a martingale. Consequently, the option price u is given by $u(x, t) = g(x)$ for all t, which fails to be a classical solution to the term structure equation since it is not differentiable at $x = 0$. One might argue, though, that the boundary condition $u_t(0, t) + \beta(0, t)u_x(0, t) = 0$ is satisfied in a weak sense.

The proof of Theorem 2.3 is carried out in several steps.

PROOF OF UNIQUENESS. Let v^1 and v^2 be two bounded classical solutions to the term structure equation, and define

$$v(x, t) = v^1(x, T - t) - v^2(x, T - t).$$

Then $v(x, t)$ is a bounded solution to

$$\begin{cases} v_t = \frac{1}{2}\sigma^2 v_{xx} + \beta v_x - xv, \\ v(x, 0) = 0, \\ v_t(0, t) = \beta(0, t)v_x(0, t). \end{cases} \tag{12}$$

Now consider the function

$$h(x, t) = (1 + x)e^{Mt},$$

where M is a positive constant. For M large enough, depending on $\beta(0,t)$ and the growth rate of β, h is a super-solution to (12) which tends to infinity at spatial infinity. Thus, according to the maximum principle, the function v is bounded above by εh and below by $-\varepsilon h$ for any $\varepsilon > 0$. It follows that $v \equiv 0$, which demonstrates uniqueness of bounded classical solutions to the term structure equation. □

PROOF OF CONTINUITY. To show that u is continuous, denote by $X^{x,t}$ the solution to (3) with initial condition $X^{x,t}(t) = x$. Let (x,t) and (y,r) be two points in $[0,\infty) \times [0,T]$. Then, if $r \leq t$, we have

$$\begin{aligned}
|u(y,r) - u(x,t)| &\leq E\big[e^{-\int_r^T X^{y,r}(s)\,ds}|g(X^{y,r}(T)) - g(X^{x,t}(T))|\big] \\
&\quad + E\big[g(X^{x,t}(T))\big|e^{-\int_r^T X^{y,r}(s)\,ds} - e^{-\int_t^T X^{x,t}(s)\,ds}\big|\big] \\
&\leq E[|g(X^{y,r}(T)) - g(X^{x,t}(T))|] \\
&\quad + C\int_t^T E[|X^{y,r}(s) - X^{x,t}(s)|]\,ds \\
&\quad + C\int_r^t E[X^{y,r}(s)]\,ds
\end{aligned} \tag{13}$$

for some constant C, where we have used that g is bounded. A similar expression can be derived if $r > t$. It follows from Remark 1 in Section 8, Chapter 2 in [11] that $X^{y,r}(t) \to x$ in L^2 as $(y,r) \to (x,t)$. Therefore, from Theorem 2.1 in [2] we have

$$E\Big[\sup_{t\leq s\leq T}\big(X^{y,r}(s) - X^{x,t}(s)\big)^2\Big] \to 0$$

as $(y,r) \to (x,t)$. (Theorem 2.1 in [2] also holds in the case of random starting points.) Since g is assumed continuous and bounded, all three terms on the right-hand side of (13) tend to 0 as $(y,r) \to (x,t)$. Thus u is continuous on $[0,\infty) \times [0,T]$. □

PROOF THAT $u \in C^{2,1}((0,\infty) \times [0,T))$ AND SATISFIES (5). For a given point $(x,t) \in (0,\infty) \times [0,T)$, let

$$R = (x_1, x_2) \times [t_1, t_2) \subseteq (0,\infty) \times [0,T)$$

be a rectangle which contains (x,t), where $x_1 > 0$. Since u is continuous, it follows from standard parabolic theory, see [9], that there exists a unique solution $U \in C^{2,1}(R)$ to the boundary value problem

$$\begin{cases} U_t + \frac{1}{2}\sigma^2 U_{xx} + \beta U_x - xU = 0, & \text{in } R, \\ U = u, & \text{on } \partial_p R, \end{cases}$$

where $\partial_p R = ([x_1,x_2] \times \{t_2\}) \cup (\{x_1,x_2\} \times [t_1,t_2])$ is the parabolic boundary of R. From Itô's formula, the process

$$Z(s) = e^{-\int_t^s X^{x,t}(r)\,dr}U(X^{x,t}(s), s)$$

is a martingale on the time interval $[t, \tau_R]$, where

$$\tau_R = \inf\{s \geq t : X^{x,t}(s) \notin R\}$$

is the first exit time from the rectangle R. Therefore,

$$U(x,t) = E\big[e^{-\int_t^{\tau_R} X^{x,t}(r)\,dr} u(X^{x,t}(\tau_R), \tau_R)\big] = u(x,t),$$

where the second equality follows from the strong Markov property. Consequently, $u \in C^{2,1}((0,\infty) \times [0,T))$. Since $u \equiv U$ on R, we also see that u satisfies (5). □

It remains to show that u is continuously differentiable up to the spatial boundary $x = 0$, and that it satisfies the boundary condition (7). This is done in Sections 3–5.

3. Continuity of the first spatial derivative. In this section we investigate regularity of the spatial derivative u_x at the boundary $x = 0$. To do this we study the stochastic representation of the terminal value problem obtained by formally differentiating the term structure equation. We show that this stochastic representation indeed is the derivative of u and that it is continuous.

Recall that α_x is assumed to be continuous on $[0,\infty) \times [0,T]$, where $\alpha(x,t) = \frac{1}{2}\sigma^2(x,t)$. Let the process Y be modeled by the stochastic differential equation

$$dY(t) = (\alpha_x + \beta)(Y(t),t)\,dt + \sigma(Y(t),t)\,dW. \tag{14}$$

Rather than specifying precise conditions under which (14) has a unique solution, we simply assume what we need.

ASSUMPTION 3.1. *The coefficients σ and β are such that*, *path-wise*, *uniqueness holds for equation* (14).

REMARK. Note that Assumption 3.1 holds for example if α is twice continuously differentiable in space, since then the drift $\alpha_x + \beta$ is locally Lipschitz continuous. Moreover, if σ and β are time-independent, then it follows from [1, 3] and Section IX.3 in [15] that Assumption 3.1 automatically holds. Thus the Cox–Ingersoll–Ross model (8), the Dothan model (9), the Hull–White model (10) and the model (11) all satisfy Assumption 3.1.

Also note that since $\alpha(0,t) = 0$, we have $\alpha_x(0,t) \geq 0$. Thus Y remains nonnegative since it has the same volatility as X but a larger drift at 0.

Next, define the function v by

$$\begin{aligned} v(x,t) = {} & E\bigg[g'(Y(T)) \exp\bigg\{\int_t^T \beta_x(Y(s),s) - Y(s)\,ds\bigg\}\bigg] \\ & - E\bigg[\int_t^T \exp\bigg\{\int_t^s \beta_x(Y(r),r) - Y(r)\,dr\bigg\} u(Y(s),s)\,ds\bigg], \end{aligned} \tag{15}$$

where Y is the solution to (14) with initial condition $Y(t) = x$.

If the term structure equation (5) is formally differentiated with respect to x, then the derivative u_x satisfies

$$(u_x)_t + \alpha(u_x)_{xx} + (\alpha_x + \beta)(u_x)_x + (\beta_x - x)u_x - u = 0$$

with terminal condition $u_x(x, T) = g'(x)$. The function v defined in (15) is the corresponding stochastic representation. In Theorem 3.4 below we show that v indeed equals the spatial derivative of u.

PROPOSITION 3.2. *The function* $v(x, t)$ *is continuous on* $[0, \infty) \times [0, T]$.

PROOF. The result follows along the same lines as the continuity of u above. Indeed, let (x_n, t_n) converge to (x, t), where $t_n \le t$, and let Y and Y^n be defined by

$$\begin{cases} dY(s) = (\alpha_x + \beta)(Y(s), s)\,ds + \sigma(Y(s), s)\,dW, \\ Y(t) = x \end{cases}$$

and

$$\begin{cases} dY^n(s) = (\alpha_x + \beta)(Y^n(s), s)\,ds + \sigma(Y^n(s), s)\,dW, \\ Y^n(t_n) = x_n, \end{cases}$$

respectively. Also define

$$I(s) := \exp\left\{\int_t^s \beta_x(Y(u), u) - Y(u)\,du\right\}$$

and

$$I^n(s) := \exp\left\{\int_{t_n}^s \beta_x(Y^n(u), u) - Y^n(u)\,du\right\}.$$

Then

$$\begin{aligned} |v(x_n, t_n) - v(x, t)| \le{}& E[|I^n(T)g'(Y^n(T)) - I(T)g'(Y(T))|] \\ &+ \int_t^T E[|I^n(s)u(Y^n(s), s) - I(s)u(Y(s), s)|]\,ds \\ &+ \int_{t_n}^t E[I^n(s)u(Y^n(s), s)]\,ds. \end{aligned}$$

The first term and the integrand in the second term are similar to the type of terms treated when proving the continuity of u. Moreover, the integrand of the third term is bounded. Thus it follows from bounded convergence that v is continuous. □

We also need a continuity result in the volatility parameter. To formulate it, let $\{\sigma^n(x, t)\}_{n=1}^{\infty}$ be a sequence of functions satisfying Hypothesis 2.1 uniformly in n, that is, with the same constant C in the bound (2). Moreover, assume that $\sigma^n(x, t)$ converges to $\sigma(x, t)$ and α_x^n converges to α_x uniformly on compacts as $n \to \infty$, where $\alpha^n = \frac{1}{2}(\sigma^n)^2$. Let u^n and v^n be defined as u and v but using the volatility function σ^n instead of σ. More explicitly,

$$u^n(x,t) = E\big[e^{-\int_t^T X^n(s)\,ds} g(X^n(T))\big]$$

and

$$\begin{aligned} v^n(x,t) &= E\bigg[g'(Y^n(T))\exp\bigg\{\int_t^T \beta_x(Y^n(s),s) - Y^n(s)\,ds\bigg\}\bigg] \\ &\quad - E\bigg[\int_t^T \exp\bigg\{\int_t^s \beta_x(Y^n(r),r) - Y^n(r)\,dr\bigg\} u^n(Y^n(s),s)\,ds\bigg], \end{aligned} \tag{16}$$

where X^n and Y^n satisfy

$$\begin{cases} dX^n(s) = \beta(X^n(s),s)\,ds + \sigma^n(X^n(s),s)\,dW(s), \\ X^n(t) = x \end{cases}$$

and

$$\begin{cases} dY^n(s) = (\alpha_x^n + \beta)(Y^n(s),s)\,ds + \sigma^n(Y^n(s),s)\,dW(s), \\ Y^n(t) = x, \end{cases}$$

respectively.

PROPOSITION 3.3. *The functions u and v are continuous in the volatility parameter. More precisely, $u^n(x,t) \to u(x,t)$ and $v^n(x,t) \to v(x,t)$ as $n \to \infty$ for any fixed point $(x,t) \in [0,\infty) \times [0,T]$.*

PROOF. It follows from Theorem 2.5 in [2] that

$$\lim_{n\to\infty} E\Big[\sup_{s\in[t,T]} \big(X(s) - X^n(s)\big)^2\Big] = 0.$$

Therefore,

$$\begin{aligned} |u^n(x,t) - u(x,t)| &\le E\big[\big|e^{-\int_t^T X^n(s)\,ds} - e^{-\int_t^T X(s)\,ds}\big| g(X^n(T))\big] \\ &\qquad + E\big[e^{-\int_t^T X(s)\,ds} |g(X^n(T)) - g(X(T))|\big] \\ &\le C\int_t^T E[|X(s) - X^n(s)|]\,ds + E[|g(X^n(T)) - g(X(T))|] \\ &\to 0 \end{aligned}$$

as $n \to \infty$. Thus u is continuous in the volatility function.

The continuity of v in the volatility function is similar. Indeed, let

$$I(s) := \exp\bigg\{\int_t^s \beta_x(Y(r),r) - Y(r)\,dr\bigg\}$$

and

$$I^n(s) := \exp\left\{\int_t^s \beta_x(Y^n(r), r) - Y^n(r)\,dr\right\}.$$

Then

$$\begin{aligned}|v^n(x,t) - v(x,t)| \le E[|I^n(T)g'(Y^n(T)) - I(T)g'(Y(T))|]\\ + \int_t^T E[|I^n(s)u^n(Y^n(s),s) - I(s)u^n(Y(s),s)|]\,ds\\ + \int_t^T E[I(s)|u^n(Y(s),s) - u(Y(s),s)|]\,ds.\end{aligned}$$

The first term and the integrand of the third term are similar to the terms appearing in the first part of the proof. The integrand of the second term can be dealt with using the fact that each u^n is Lipschitz continuous in x, uniformly in n since the Lipschitz property is inherited by the value function. Thus all terms tend to zero as $n \to \infty$, so v is continuous in the volatility. □

THEOREM 3.4. *We have* $u_x(x,t) = v(x,t)$ *on* $[0,\infty) \times [0,T]$. *Consequently,* u_x *is continuous on* $[0,\infty) \times [0,T]$.

PROOF. It suffices to prove $u_x(x,0) = v(x,0)$. We first assume that σ is continuously differentiable in x with a bounded derivative. It then follows from Section 5.5 in [10] or Section 8 in [11] that the derivative

$$\xi(t) := \frac{\partial X(t)}{\partial x}$$

of $X(t) = X^{x,0}(t)$ with respect to the initial point x exists and is continuous, and it satisfies

$$\begin{cases} d\xi(t) = \xi(t)\beta_x(X(t),t)\,dt + \xi(t)\sigma_x(X(t),t)\,dW(t),\\ \xi(0) = 1.\end{cases}$$

Moreover,

$$\begin{aligned}u_x(x,0) &= E\left[g'(X(T))\xi(T)\exp\left\{-\int_0^T X(s)\,ds\right\}\right]\\ &\quad - E\left[g(X(T))\exp\left\{-\int_0^T X(s)\,ds\right\}\int_0^T \xi(s)\,ds\right] \qquad (17)\\ &=: I_1 - I_2.\end{aligned}$$

We claim that $I_i = J_i$, $i = 1, 2$, where

$$J_1 = E\left[g'(Y(T))\exp\left\{\int_0^T \beta_x(Y(s),s) - Y(s)\,ds\right\}\right]$$

and

$$J_2 = E\left[\int_0^T \exp\left\{\int_0^s \beta_x(Y(r), r) - Y(r)\,dr\right\} u(Y(s), s)\,ds\right],$$

compare (15) above. Here Y is defined as in (14) with initial condition $Y(0) = x$.

To show that $I_1 = J_1$, define a new measure Q on $\mathcal{F}_T$ by $dQ = M(T)\,dP$, where the process M is defined by

$$M(t) = \xi(t) \exp\left\{-\int_0^t \beta_x(Y(s))\,ds\right\}. \tag{18}$$

By Itô's formula,

$$dM(t) = M(t)\sigma_x(X(t))\,dW(t),$$

so M is a martingale since σ_x is bounded. In particular, $E[M(T)] = 1$, so Q is a probability measure. From Girsanov's theorem it follows that

$$\tilde{W}(t) = W(t) - \int_0^t \sigma_x(X(s))\,ds$$

is a Q-Brownian motion, and

$$dX = (\sigma\sigma_x + \beta)(X(t), t)\,dt + \sigma(X(t), t)\,d\tilde{W}.$$

Here $\sigma\sigma_x = \alpha_x$, so by weak uniqueness, the Q-law of X is the same as the law of Y under P. Consequently,

$$\begin{aligned} I_1 &= E\left[g'(X(T))\xi^x(T) \exp\left\{-\int_0^T X(s)\,ds\right\}\right] \\ &= E^Q\left[g'(X(T)) \exp\left\{\int_0^T \beta_x(X(s), s) - X(s)\,ds\right\}\right] = J_1. \end{aligned}$$

To prove $I_2 = J_2$, note that

$$\begin{aligned} I_2 &= E\left[g(X(T)) \exp\left\{-\int_0^T X(s)\,ds\right\} \int_0^T \xi(s)\,ds\right] \\ &= \int_0^T E\left[\exp\left\{-\int_0^s X(r)\,dr\right\}\xi(s)\right. \\ &\qquad\qquad \left.\times E\left[g(X(T)) \exp\left\{-\int_s^T X(r)\,dr\right\}\Big|\mathcal{F}_s\right]\right] ds \\ &= \int_0^T E\left[\exp\left\{-\int_0^s X(r)\,dr\right\}\xi(s) u(X_s, s)\right] ds \end{aligned}$$

by the Markov property. Define a new measure $Q = Q_s$ on $\mathcal{F}_s$ by

$$dQ = M(s)\,dP,$$

where M is defined as in (18). Girsanov's theorem yields

$$\begin{aligned} &E\left[\exp\left\{-\int_0^s X(r)\,dr\right\}\xi(s)u(X_s,s)\right] \\ &\quad = E^Q\left[\exp\left\{\int_0^s \beta_x(X(r),r) - X(r)\,dr\right\}u(X_s,s)\right] \\ &\quad = E\left[\exp\left\{\int_0^s \beta_x(Y(r),r) - Y(r)\,dr\right\}u(Y_s,s)\right]. \end{aligned}$$

Consequently, $I_2 = J_2$, which finishes the proof in the case of continuously differentiable σ.

The general case follows by approximation. Let σ^n, u^n and v^n be as described before Proposition 3.3, with each σ^n being continuously differentiable in x with bounded derivative. From above, we then know that $v^n(x,t) = u^n_x(x,t)$ at all points. Moreover, by Proposition 3.3, $v^n(x,t) \to v(x,t)$ point-wise as $n \to \infty$.

On the other hand, since u^n converges to u point-wise and is uniformly bounded, it follows from standard parabolic theory that also u^n_x converges to u_x point-wise for all points (x,t) with $x > 0$. Consequently, $v = u_x$ on $(0,\infty) \times [0,T]$. Since v is continuous on $[0,\infty) \times [0,T]$ by Proposition 3.2, it is easy to check that $u_x(0,t)$ exists and that we have $v = u_x$ everywhere on $[0,\infty) \times [0,T]$. The continuity of u_x thus follows. □

4. An estimate of the second spatial derivative. Since the function v defined in (15) is continuous, it follows that [by a similar argument as in the proof that u satisfies (5)] it indeed solves the differentiated equation

$$v_t = \alpha v_{xx} + (\alpha_x + \beta)v_x + (\beta_x - x)v - u$$

on $(0,\infty) \times [0,T)$. In this section we use interior estimates to show that $\alpha v_x \to 0$ as $x \to 0$. Since $v = u_x$ by Theorem 3.4, this shows that the term αu_{xx} in (5) approaches zero close to the boundary.

PROPOSITION 4.1. *The function* $v = u_x$ *satisfies*

$$\lim_{(x,t)\to(0,t_0)} \alpha(x,t)v_x(x,t) = 0$$

for any t_0. *Consequently*, $\lim_{(x,t)\to(0,t_0)} \alpha(x,t)u_{xx}(x,t) = 0$.

PROOF. Let $\{(x_n,t_n)\}_{n=1}^{\infty} \subseteq (0,\infty) \times [0,T)$ be a sequence of points converging to $(0,t_0)$, where $t_0 \in [0,T)$. Define new coordinates (y,s) by letting $y = kx$ and $s = k(t-t_0)$, where k is specified more precisely below. Then the function w defined by

$$w(y,s) = v(x,t)$$

satisfies

$$w_s = \tilde{\alpha} w_{yy} + \tilde{\beta} w_y + \gamma w + h, \tag{19}$$

where

$$\tilde{\alpha}(y,s) = \alpha\left(\frac{y}{k}, t_0 + \frac{s}{k}\right)k,$$

$$\tilde{\beta}(y,s) = (\alpha_x + \beta)\left(\frac{y}{k}, t_0 + \frac{s}{k}\right),$$

$$\gamma(y,s) = \frac{1}{k}\beta_x\left(\frac{y}{k}, t_0 + \frac{s}{k}\right) - \frac{y}{k^2}$$

and

$$h(y,s) = -\frac{1}{k}u\left(\frac{y}{k}, t_0 + \frac{s}{k}\right).$$

Now consider a region $\mathcal{R} = \mathcal{R}^n$ which contains the point (x_n, t_n), and such that

$$1 \leq \alpha(x,t)k \leq 2 \tag{20}$$

in $\mathcal{R}$. Since $\alpha_x(x,t)$ is continuous up to the boundary, the region $\mathcal{R}$ in (y,s)-coordinates does not collapse as $n \to \infty$, but it can rather be chosen to consist of a rectangle of fixed size; the location of the rectangle is not necessarily fixed though. In this rectangle, the coefficients of the equation (19) satisfy

$$1 \leq \tilde{\alpha}(y,s) \leq 2,$$

$$|\tilde{\beta}(y,s)| \leq C,$$

$$|\gamma(y,s)| \leq C$$

and

$$|h(y,s)| \leq C/k$$

for some constant C which is independent of n. Since $w(y,s) = v(x,t)$ we have that w converges to the constant $v(0,t_0) = u_x(0,t_0)$ uniformly on $\mathcal{R}$ as $n \to \infty$. By interior Schauder estimates, w_y tends to 0 as $n \to \infty$. Since

$$\alpha(x,t)v_x(x,t) = \tilde{\alpha}(y,s)w_y(y,s),$$

and since $\tilde{\alpha}(y,s)$ is bounded on $\mathcal{R}$, the conclusion follows. □

5. The time derivative at the boundary. It follows from Proposition 4.1 and (5) that

$$\lim_{(x,t)\to(0,t_0)} u_t(x,t) + \beta(0,t_0)u_x(0,t_0) = 0 \tag{21}$$

for any $t_0 \in [0,T)$. In this section we show that the boundary condition (7) also holds *at* the boundary, that is, not merely in the limit.

PROPOSITION 5.1. *The function $u_t(x,t)+\beta(x,t)u_x(x,t)$ defines a continuous function on $[0,\infty)\times[0,T)$. Moreover, it vanishes for $x=0$.*

REMARK. Note that Proposition 5.1 finishes the proof of Theorem 2.3.

PROOF OF PROPOSITION 5.1. In view of (21) above, it suffices to show that u_t exists at the boundary and that it equals $-\beta u_x$. To do this, fix a point on the boundary with coordinates $(0,t_0)$. For notational simplicity we assume that $t_0=0$. The time (left) derivative u_t at the boundary is defined by

$$u_t(0,0)=\lim_{k\to\infty}k\left(u(0,0)-u\left(0,-\frac{1}{k}\right)\right), \tag{22}$$

provided the limit exists. To determine $u_t(0,0)$, we let X^k be defined by

$$\begin{cases} dX^k=\beta(X^k(t),t)\,dt+\sigma(X^k(t),t)\,dW,\\ X^k(-1/k)=0.\end{cases}$$

However, instead of considering the process X with different starting times, we perform a change of variables so that the starting time is independent of k. We thus introduce the process $Y^k(s)$ by

$$Y^k(s)=kX^k\left(\frac{s}{k}\right).$$

With respect to the time variable s, the dynamics of Y^k has the form

$$\begin{cases} dY^k(s)=\beta\left(\frac{1}{k}Y^k(s),\frac{s}{k}\right)ds+\sqrt{k\sigma^2\left(\frac{1}{k}Y^k(s),\frac{s}{k}\right)}\,dW^k,\\ Y^k(-1)=0,\end{cases} \tag{23}$$

where $W^k(s)$ denotes some Brownian motion. By the Markov property,

$$\begin{aligned} u(0,-1/k)&=E\big[e^{-\int_{-1/k}^0 X^k(s)\,ds}u(X^k(0),0)\big]\\ &=E\left[e^{-\int_{-1}^0(1/k^2)Y^k(s)\,ds}u\left(\frac{1}{k}Y^k(0),0\right)\right].\end{aligned}$$

Hence,

$$\begin{aligned} u_t(0,0)&=\lim_{k\to\infty}kE_{0,-1}\left[u(0,0)-e^{-\int_{-1}^0(1/k^2)Y^k(s)\,ds}u\left(\frac{1}{k}Y^k(0),0\right)\right]\\ &=\lim_{k\to\infty}E_{0,-1}\left[k\left(u(0,0)-u\left(\frac{1}{k}Y^k(0),0\right)\right)\right],\end{aligned}$$

where the second equality follows using the inequality $e^{-x}-1\ge -x$ since

$$E_{0,-1}\left[ku\left(\frac{1}{k}Y^k(0),0\right)\big|e^{-\int_{-1}^0(1/k^2)Y^k(s)\,ds}-1\big|\right]\le C\frac{1}{k}E_{0,-1}\left[\int_{-1}^0Y^k(s)\,ds\right]\to 0$$

as $k \to \infty$. Now, define the process Y by

$$\begin{cases} dY = \beta(0,0)\,ds + \sqrt{2\alpha_x(0,0)Y}\,dW, \\ Y(-1) = 0, \end{cases}$$

and redefine Y^k as in (23) above but using the same Brownian motion W (this does not change the law of Y^k). Since

$$\beta^k(y,s) := \beta\left(\frac{y}{k}, \frac{s}{k}\right) \to \beta(0,0)$$

and

$$\sigma^k(y,s) := \sqrt{k\sigma^2\left(\frac{y}{k}, \frac{s}{k}\right)} \to \sqrt{2\alpha_x(0,0)y}$$

uniformly on compacts as $k \to \infty$ (here we used the assumption that α is continuously differentiable in space), it follows from [2] that $Y^k(0) \to Y(0)$ in L^2 as $k \to \infty$. From Theorem 3.4 above we know that u is differentiable in x, so $k(u(0,0) - u(\frac{y}{k}, 0))$ converges to $-u_x(0,0)y$. By dominated convergence, we have

$$kE\left[u(0,0) - u\left(\frac{1}{k}Y(0), 0\right)\right] \to -u_x(0,0)E[Y(0)] = -\beta(0,0)u_x(0,0)$$

as $k \to \infty$. Moreover, the Lipschitz property of u yields that

$$kE\left[u\left(\frac{1}{k}Y(0), 0\right) - u\left(\frac{1}{k}Y^k(0), 0\right)\right] \le CE[|Y(0) - Y^k(0)|] \to 0$$

as $k \to \infty$. It follows that

$$u_t(0,0) + u_x(0,0)\beta(0,0) = 0.$$

As $t_0 = 0$ was chosen only for notational convenience, we have that

$$u_t(0,t) + \beta(0,t)u_x(0,t) = 0$$

for any t. To be precise, we have shown the result above only for the left t-derivative. However, this left t-derivative is continuous by the equation above, so it follows from a simple calculus lemma that in fact u is differentiable in time, thus finishing our proof. □

6. Models allowing negative interest rates. For models in which the short rate can fall below zero with positive probability, the connection between the option price, given by a risk-neutral expected value, and the term structure equation is more straightforward than for models with nonnegative rates. Nevertheless, we have not been able to find a precise reference for this case, so for completeness we provide such a result in this section.

The assumptions needed on the drift and volatility are presented below, where $x^+ = \max(x, 0)$. These assumptions now replace those of Hypothesis 2.1. To our knowledge, all models used in practice that allow negative interest rates satisfy these requirements. For example, the Vasicek model, in which

$$dX(t) = \big(a - bX(t)\big)\,dt + \sigma\,dW,$$

where a, b and σ are positive constants, is covered.

HYPOTHESIS 6.1. *The drift $\beta \in C(\mathbb{R} \times [0, T])$ and the volatility $\sigma \in C(\mathbb{R} \times [0, T])$ are both Lipschitz continuous in x. Moreover,*

$$0 < |\sigma(x, t)| \le C(1 + x^+)$$

and

$$|\beta(x, t)| \le C(1 + |x|) \tag{24}$$

for some positive constant C.

REMARK. The assumption that σ is strictly positive is not a strong assumption. Indeed, let us for simplicity consider a time-homogeneous model $dX(t) = \beta(X(t))\,dt + \sigma(X(t))\,dW(t)$ with $\sigma(a) = 0$ for some $a \in \mathbb{R}$. If $X(0) \ge a$ and $\beta(a) \ge 0$, then the process X cannot take values smaller than a, so we are essentially in the situation handled in Sections 2–5 (but with the point 0 replaced with a). If $b(a) < 0$, then X can take values below a, but if this happens then the process will stay below a forever, and we are then again back in the previous situation.

The bound on the volatility for negative rates guarantees that bond prices are finite. For models in which σ grows faster than $\sqrt{|x|}$ for negative rates, bond prices can be infinite; compare Theorem 4.1 in [16].

For a given continuous payoff function $g : \mathbb{R} \to [0, \infty)$, define the corresponding option price $u : \mathbb{R} \times [0, T]$ by

$$u(x, t) = E_{x,t}\left[\exp\left\{-\int_t^T X(s)\,ds\right\} g(X(T))\right],$$

where

$$\begin{cases} dX(s) = \beta(X(s), s)\,ds + \sigma(X(s), s)\,dW(s), \\ X(t) = x. \end{cases}$$

We require that the payoff function is bounded for positive interest rates and of, at most, exponential growth for negative rates, that is,

$$0 \le g(x) \le K \max\{1, e^{-Kx}\} \tag{25}$$

for some positive constant K. The corresponding term structure equation is given by

$$u_t(x, t) + \tfrac{1}{2}\sigma^2(x, t)u_{xx}(x, t) + \beta(x, t)u_x(x, t) = xu(x, t)$$

on $\mathbb{R} \times [0, T)$, with terminal condition $u(x, T) = g(x)$. By a *classical solution* to this equation we mean a solution which is continuous up to the boundary $t = T$ and with all derivatives appearing in the equation being continuous functions on the set $t < T$. The following is our main result about the term structure equation on the whole real line.

THEOREM 6.2. *Assume Hypothesis 6.1 and the bound (25). Then the option price $u(x,t)$ satisfies*

$$u(x,t) \le K' \max\{1, e^{-K'x}\} \tag{26}$$

for some constant K'. Moreover, $u(x,t)$ is the unique classical solution to the term structure equation satisfying this growth assumption for some constant K'.

REMARK. Note that the bound (25) is natural for models on the whole real line. In fact, even if g was bounded, the option price u would be of exponential growth for negative rates. Also note that, for example, call options on a bond are covered by Theorem 6.2. Indeed, the payoff of a bond call option with maturity T_1 is given by $g(x) = (u(x, T_1) - K)^+$, where u is the price of a bond maturing at $T_2 > T_1$. Since u satisfies (26) by Theorem 6.2, the payoff g satisfies (25). Bond put options are trivially covered since they have bounded payoff functions.

PROOF OF THEOREM 6.2. The bound (26) follows from Corollary 3.3 in [8]. To prove uniqueness of solutions, assume that v is a solution to the term structure equation with boundary value $g = 0$ such that $|v| \le K' \max\{1, e^{-K'x}\}$ for some constant K'. Let

$$h(x,t) = e^{M(T-t)}\bigl(e^{-f(t)x} + x\bigr)$$

for some large constant M. Here

$$f(t) = \frac{e^{C(T-t)} - 1}{C} + K e^{C(T-t)},$$

where C is the constant appearing in (24) and $K > K'$. Then the set

$$\{(x,t) \in \mathbb{R} \times [0,T] : \varepsilon h(x,t) < v(x,t)\}$$

is bounded, and

$$h_t + \tfrac{1}{2}\sigma^2 h_{xx} + \beta h_x - xh < 0$$

at all points provided M is chosen large enough. Standard methods used to prove the maximum principle yield that $v = 0$ at all points. Thus we have uniqueness of solutions to the term structure equation in the class of functions satisfying (26).

To show that u is a classical solution to the term structure equation, we carry out an approximation argument. We consider the term structure equation but with the discount factor x replaced by bounded functions that agree with x inside some large compact sets. We also replace the payoff function g with functions of, at most, polynomial growth that agree with g on large compact sets. The corresponding equation is in the standard class and its stochastic solution is known to be continuous and hence is a classical solution. Now let the functions approximating the discount factor grow up to x for large positive x and then decrease down to x for large negative x. By the monotone convergence theorem, the corresponding stochastic solutions converge to the stochastic solution above denoted u. Interior Schauder estimates yield interior regularity of the limiting solution and continuity at the boundary is established using the maximum principle. □

REFERENCES

[1] ATHREYA, S. R., BARLOW, M. T., BASS, R. F. and PERKINS, E. A. (2002). Degenerate stochastic differential equations and super-Markov chains. *Probab. Theory Related Fields* **123** 484–520. MR1921011

[2] BAHLALI, K., MEZERDI, B. and OUKNINE, Y. (1998). Pathwise uniqueness and approximation of solutions of stochastic differential equations. In *Séminaire de Probabilités*, *XXXII*. *Lecture Notes in Math*. **1686** 166–187. Springer, Berlin. MR1655150

[3] BASS, R. F. and PERKINS, E. A. (2003). Degenerate stochastic differential equations with Hölder continuous coefficients and super-Markov chains. *Trans*. *Amer. Math*. *Soc*. **355** 373–405 (electronic). MR1928092

[4] BJÖRK, T. (2004). *Arbitrage Theory in Continuous Time*, 2nd ed. Oxford Univ. Press.

[5] BRIGO, D. and MERCURIO, F. (2001). *Interest Rate Models—Theory and Practice*. Springer, Berlin. MR1846525

[6] DUFFY, D. J. (2006). *Finite Difference Methods in Financial Engineering*. Wiley, Chichester. MR2286409

[7] EKSTRÖM, E., LÖTSTEDT, P. and TYSK, J. (2009). Boundary values and finite difference methods for the single factor term structure equation. *Appl*. *Math*. *Finance* **16** 253–259. MR2589641

[8] EKSTRÖM, E. and TYSK, J. (2008). Convexity theory for the term structure equation. *Finance Stoch*. **12** 117–147. MR2362153

[9] FRIEDMAN, A. (1964). *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, NJ. MR0181836

[10] FRIEDMAN, A. (1975). *Stochastic Differential Equations and Applications* **1**. *Probability and Mathematical Statistics* **28**. Academic Press, New York. MR0494490

[11] GĪHMAN, Ĭ. Ī. and SKOROHOD, A. V. (1972). *Stochastic Differential Equations*. Springer, New York. MR0346904

[12] HESTON, S., LOEWENSTEIN, M. and WILLARD, G. (2007). Options and bubbles. *Rev. Financ*. *Stud*. **20** 359–390.

[13] JANSON, S. and TYSK, J. (2006). Feynman–Kac formulas for Black–Scholes-type operators. *Bull*. *London Math*. *Soc*. **38** 269–282. MR2214479

[14] OLEĬNIK, O. A. and RADKEVIČ, E. V. (1973). *Second Order Equations with Nonnegative Characteristic Form*. Plenum Press, New York. MR0457908

[15] REVUZ, D. and YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Springer, Berlin. MR1725357

[16] YONG, J. (2006). Remarks on some short rate term structure models. *J. Ind. Manag. Optim.* **2** 119–134. MR2208398

DEPARTMENT OF MATHEMATICS
UPPSALA UNIVERSITY
BOX 480, SE-75106 UPPSALA
SWEDEN
E-MAIL: ekstrom@math.uu.se
johan.tysk@math.uu.se