Ecole Polytechnique
Applied Mathematics Department

EA MAP571

Tutor: Frederic BONNANS

American Options: conditions for artificial limits

Brahim AIT HADDOU
Tiago SALVADOR

06 December 2011

Contents

1 Introduction

An option is a contract that gives its owner the right (but not the obligation) to buy ('*call option*') or to sell ('*put option*') an underlying asset S (typically a stock) for the *strike price* E by the date T to receive some payoffs. An '*European option*' can only be exercised at the *maturity time* T while an '*American option*' can be exercised at anytime up to T. An option, just like a stock or a bond, is a security. It is also a binding contract with strictly defined terms and properties. The basic problem here is to specify a fare price to charge for permitting these rights. A closely related question is how to hedge the risks that arises when selling these options.

The famous Black-Scholes equation is an effective model for option pricing. The standard approach to solve this equation consists in transforming it into a heat equation posed on semi-unbounded domain with a free boundary, whose exact analytical expression is not known yet. Usually finite difference or finite elements are used to discretize this heat equation and an *artificial boundary condition* (ABC) is introduced in order to reduce the computational domain. If the solution on the computational domain coincides with the exact solution on the unbounded domain (restricted to the finite domain), one refers to these boundary conditions as a *transparent boundary condition* (TBC).

In this report will study the numerical resolution of the Black-Sholes equation using artificial boundary conditions. The report is organized as follows: first we introduce the Black-Scholes equation and recall the standard transformations to a forward-in-time heat equation. In §3, we derive the analytic TBC for the heat equation. We review in §4 two approaches to discretize the analytic TBC and construct a *discrete transparent boundary condition* DTBC for the Crank-Nicolson discretization. In order to try to reduce the numerical effort in the DTBC constructed, we present in §5 a simple implementation by the sum-of-exponentials approximation. In §6 we discuss the numerical treatment of the free boundary, after we analyze in §7 the stability of the resulting numerical scheme. Finally, we illustrate in the §8 the accuracy and efficiency of the new method with some numerical examples and compare it to the known discretized of Mayfield and Han and Wu.

2 The Black-Scholes equation

In this report we consider an American call option. The treatment of an American put option is analogous and we will make some remarks to emphasize that. We represent the value of the American call option by $V = V(S,t)$ since it depends on the current market price of the underlying asset, S, and the remaining time $T-t$ until the option expires. The Black-Scholes equation that models our problem is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - D_0)S\frac{\partial V}{\partial S} - rV = 0, \quad 0 < S < S_f(t), \quad 0 \leq t < T, \tag{2.1a}$$

where σ denotes the annual volatility of the asset price, r the risk-free interest rate and T is the expiry date ($t = 0$ means 'today'). We assume that dividends are paid with a continuous yields of constant level $D_0 > 0$. The equation (2.1a) is a backward-in-time parabolic equation and is posed on a time-dependent domain.

In (2.1a), $S_f(t)$, usually called '*early exercise boundary*' or '*optimal exercise price*', denotes the (a priori unknown) free boundary that helps decide if we should or not exercise our American call option: if the value of the asset S is equal or greater than $S_f(t)$ at time t, then we should exercise, otherwise the option should be held. Hence, the free boundary $S_f(t)$ separates the holding region ($S < S_f(t)$) from the exercise region ($S \geq S_f(t)$).

At the expiry date $t = T$, we only exercise our option, if the value of the asset S is greater than the strike price E and, since we have dividends, if we have more profit from the dividends then from the interest rate, that is, if $D_0 S > rE$. Hence the final condition ('*payoff condition*') can be written as

$$V(S,T) = (S-E)^+, \quad 0 \leq S < S_f(T), \tag{2.1b}$$

where $S_f(T) = \max(E, rE/D_0)$.

We only have left to see the 'spacial' or asset-price boundary conditions at $S = 0$ and $S = S_f(t)$. At $S = 0$, the option is worthless, therefore

$$V(0,t) = 0, \quad 0 \leq t \leq T. \tag{2.1c}$$

Note that we need two conditions at the free boundary $S = S_f(t)$. One condition is necessary for the solution of (2.1a) and the other one is need for determining the position of the free boundary $S_f(t)$. Hence, we have

$$V(S_f(t),t) = (S_f(t)-E)^+, \quad \frac{\partial V}{\partial S}(S_f(t),t) = 1, \quad 0 \le t \le T. \tag{2.1d}$$

The first condition ('value matching' condition) is the continuity of the mapping $S \mapsto V(S,t)$ since $V(S,t) = (S-E)^+ = S-E$, in the exercise region $S \ge S_f(t)$. At $S = S_f(t)$ one requires additionally that $V(S,t)$ touches the payoff function tangentially ('high contract condition'), i.e., the function $S \mapsto \frac{\partial V(S,t)}{\partial S}$ should be continuous at $S = S_f(t)$. The conditions (2.1d) are jointly referred as the 'smooth-pasting conditions'.

Remark 2.1 *As we have already mentioned, the treatment of an American put option is analogous. In such case, one has to consider the Black-Scholes equation (2.1a) on the domain $S > S_f(t)$. The terminal condition at the expiry date $t = T$ then reads*

$$V(S,T) = (E-S)^+, \quad S > S_f(T), \tag{2.2a}$$

and the 'spatial' boundary conditions at $S = S_f(t)$ and $S \to \infty$ are given by

$$V(S_f(t),t) = (E-S_f(t))^+, \quad \frac{\partial V}{\partial S}(S_f(t),t) = -1, \quad 0 \le t \le T, \tag{2.2b}$$

$$\lim_{S\to\infty} V(S,T) = 0, \quad 0 \le t \le T. \tag{2.2c}$$

As a consequence of the non-arbitrage principal, we can deduce the following a priori bound for our solution:

$$V(S,t) \ge (S-E)^+, \quad S \ge 0, \quad 0 \le t \le T.$$

If $V(S,t) < (S-E)^+$ for one value of $S > E$ and $t \le T$, then we could purchase a call option for V, exercise it immediatly, buying the underlying asset for E and selling it for its value S. This way, we would have an instantaneous risk-free profit of $S - V - E > 0$, which is a violation to the non-arbitrage principle (in this reasoning we ignore transaction costs).

2.1 The transformation to the heat equation

In this section, we will see how we can transform (2.1a) into the heat equation.

First, it is convenient to apply a time reversal and transform (2.1a) into a forward-in-time equation. In order to achieve that, we make the following change of variable: $t = T - 2\tau/\sigma^2$. The new variable τ represents (up to the scaling $\sigma^2/2$) the remaining life time of the call option. We denote the new variables by:

$$\begin{aligned}
&\widetilde{V}(S,\tau) = V(S,t) = V(S, T - \frac{2\tau}{\sigma^2}), \quad \widetilde{S}_f(\tau) = S_f(T - \frac{2\tau}{\sigma^2}),\\
&\widetilde{r} = \frac{2}{\sigma^2} r, \quad \widetilde{D}_0 = \frac{2}{\sigma^2} D_0, \quad \widetilde{T} = \frac{\sigma^2}{2} T.
\end{aligned}$$

We are now going to see which equation $\widetilde{V}$ satisfies. We have that

$$\begin{aligned}
\frac{\partial \widetilde{V}}{\partial \tau} &= -\frac{2}{\sigma^2}\frac{\partial V}{\partial t}\\
&= S^2 \frac{\partial^2 V}{\partial S^2} + \left(\frac{2r}{\sigma^2} - \frac{2D_0}{\sigma^2}\right)\frac{\partial V}{\partial S} - \frac{2r}{\sigma^2} V \qquad \text{by (2.1a)}\\
&= S^2 \frac{\partial^2 V}{\partial S^2} + (\widetilde{r} - \widetilde{D}_0)\frac{\partial \widetilde{V}}{\partial S} - \widetilde{r}\widetilde{V}
\end{aligned}$$

which is valid for

$$0 \leq T - \frac{2\tau}{\sigma^2} < T \Leftrightarrow 0 < \tau \leq \frac{\sigma^2}{2} T = \widetilde{T}$$

and

$$0 < S < S_f(T - \frac{2\tau}{\sigma^2}) \Leftrightarrow 0 < S < \widetilde{S}_f(\tau).$$

Hence the resulting forward-in-time equation is

$$\frac{\partial \widetilde{V}}{\partial \tau} = S^2 \frac{\partial^2 V}{\partial S^2} + (\widetilde{r} - \widetilde{D}_0)\frac{\partial \widetilde{V}}{\partial S} - \widetilde{r}\widetilde{V}, \quad 0 < S < \widetilde{S}_f(\tau), \quad 0 < \tau \leq \frac{\sigma^2}{2} T = \widetilde{T}, \tag{2.3a}$$

with initial condition

$$\widetilde{V}(S,0) = (S - E)^+, \quad 0 \leq S < \widetilde{S}_f(0) = \max(E, \widetilde{r}E/\widetilde{D}_0) \tag{2.3b}$$

and the boundary conditions

$$\widetilde{V}(0,\tau) = 0, \quad 0 \leq \tau \leq \widetilde{T}, \tag{2.3c}$$

$$\widetilde{V}(\widetilde{S}_f(\tau),\tau) = (\widetilde{S}_f(\tau) - E)^+, \quad \frac{\partial \widetilde{V}}{\partial S}(\widetilde{S}_f(t), t) = 1, \quad 0 \leq \tau \leq \widetilde{T}. \tag{2.3d}$$

We now transform (2.3a) into the heat equation. To do so, let

$$\alpha = -\frac{1}{2}(\widetilde{r} - \widetilde{D}_0 - 1), \quad \beta = -\alpha^2 - \widetilde{r},$$

use the change of variables

$$S = Ee^x, \quad \widetilde{V}(S,\tau) = Ee^{\alpha x + \beta\tau} v(x,\tau)$$

and define the function g as

$$g(x,\tau) = e^{-\alpha x - \beta\tau}\left(e^x - 1\right)^+ .$$

To see that we in fact obtain the heat equation, we will compute $\frac{\partial v}{\partial \tau}$ and $\frac{\partial^2 v}{\partial x^2}$, but first we start by observing that we can write the change of variables as

$$x = \log\left(S/E\right), \quad v(x,\tau) = \frac{e^{-\alpha x - \beta\tau}}{E}\widetilde{V}(Ee^x,\tau).$$

Therefore

$$\begin{aligned}\frac{\partial v}{\partial \tau} &= -\beta\frac{e^{-\alpha x - \beta\tau}}{E}\widetilde{V}(Ee^x,\tau) + \frac{e^{-\alpha x - \beta\tau}}{E}\frac{\partial \widetilde{V}}{\partial \tau} \\ &= -\beta v(x,\tau) + \frac{e^{-\alpha x - \beta\tau}}{E}\frac{\partial \widetilde{V}}{\partial \tau},\end{aligned} \tag{2.4}$$

$$\begin{aligned}\frac{\partial v}{\partial x} &= -\alpha\frac{e^{-\alpha x - \beta\tau}}{E}\widetilde{V}(Ee^x,\tau) + \frac{e^{-\alpha x - \beta\tau}}{E}Ee^x\frac{\partial \widetilde{V}}{\partial S} \\ &= -\alpha v(x,\tau) + \frac{\partial \widetilde{V}}{\partial S}e^{(1-\alpha)x - \beta\tau}\end{aligned} \tag{2.5}$$

and, finally,

$$\begin{aligned}\frac{\partial^2 v}{\partial x^2} &= -\alpha\frac{\partial v}{\partial x} + Ee^x\frac{\partial^2 \widetilde{V}}{\partial S^2}e^{(1-\alpha)x - \beta y} + (1-\alpha)\frac{\partial \widetilde{V}}{\partial S}e^{(1-\alpha)x - \beta\tau} \\ &= \alpha^2 v(x,\tau) - \alpha\frac{\partial \widetilde{V}}{\partial S}e^{(1-\alpha)x - \beta\tau} + \frac{e^{-\alpha x - \beta\tau}}{E}S^2\frac{\partial^2 \widetilde{V}}{\partial S^2} + (1-\alpha)\frac{\partial \widetilde{V}}{\partial S}e^{(1-\alpha)x - \beta\tau} \quad \text{by (2.5)} \\ &= (-\beta - \widetilde{r})v(x,\tau) + \frac{e^{-\alpha x - \beta\tau}}{E}S^2\frac{\partial^2 \widetilde{V}}{\partial S^2} + (1-2\alpha)\frac{e^{-\alpha x - \beta\tau}}{E}S\frac{\partial \widetilde{V}}{\partial S},\end{aligned}$$

and by (2.3a), we can write

$$\frac{\partial^2 v}{\partial x^2} = \frac{e^{-\alpha x - \beta\tau}}{E}\frac{\partial \widetilde{V}}{\partial \tau} - \beta v(x,\tau). \tag{2.6}$$

Hence by (2.4) and (2.6), we have

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2}, \quad -\infty < x < x_f(\tau), \quad 0 < \tau \le \widetilde{T}, \tag{2.7a}$$

where $x_f(\tau) = \log\left(\widetilde{S}_f(\tau)/E\right)$.

As for the initial condition, we see that

$$\begin{aligned} v(x,0) &= \frac{\widetilde{V}(S,0)}{E}e^{-\alpha x} \\ &= \left(\frac{S}{E}e^{-\alpha x} - e^{-\alpha x}\right)^+ \\ &= \left(e^{\frac{1}{2}\left(\widetilde{r}-\widetilde{D}_0+1\right)x} - e^{\frac{1}{2}\left(\widetilde{r}-\widetilde{D}_0-1\right)x}\right)^+ \\ &= g(x,0). \end{aligned}$$

Then, the initial condition is given by

$$v(x,0) = g(x,0), \quad x < x_f(0) \tag{2.7b}$$

with $x_f(0) = \log\left(\max(1, r/D_0)\right)$.

For the boundary conditions, we compute $v(x_f(\tau),\tau)$ and $\frac{\partial \widetilde{V}}{\partial S}\left(\widetilde{S}_f(\tau),\tau\right)$:

$$\begin{aligned} v(x_f(\tau),\tau) &= \frac{\widetilde{V}(Ee^{x_f(\tau)},\tau)}{E}e^{-\alpha x_f(\tau)-\beta\tau} \\ &= \frac{\widetilde{V}(\widetilde{S}_f(\tau),\tau)}{E}e^{-\alpha x_f(\tau)-\beta\tau} \\ &= \frac{\left(\widetilde{S}_f(\tau)-E\right)^+}{E}e^{-\alpha x_f(\tau)-\beta\tau} \\ &= \left(e^{x_f(\tau)}-1\right)^+ e^{-\alpha x_f(\tau)-\beta\tau} \\ &= g(x_f(\tau),\tau), \end{aligned}$$

$$
\begin{aligned}
\frac{\partial \widetilde{V}}{\partial S}\left(\widetilde{S}_f(\tau),\tau\right) &= \frac{\partial}{\partial S}\left(E\,(S/E)^{\alpha}\,e^{\beta\tau}v\,(\log(S/E),\tau)\right)\Big|_{S=\widetilde{S}_f(\tau)} \\
&= E\frac{\alpha}{E}\left(\widetilde{S}_f(\tau)/E\right)^{\alpha-1}e^{\beta\tau}v\left(\log(\widetilde{S}_f(\tau)/E),\tau\right)+E\left(\widetilde{S}_f(\tau)/E\right)^{\alpha}e^{\beta\tau}\frac{\partial v}{\partial x}\frac{1/E}{\widetilde{S}_f(\tau)/E} \\
&= e^{(\alpha-1)x_f(\tau)+\beta\tau}\left(\alpha v(x_f(\tau),\tau)+\frac{\partial v}{\partial x}(x_f(\tau),\tau)\right).
\end{aligned}
$$

Hence, we have

$$
\lim_{x\to-\infty} v(x,\tau)=0, \quad 0\le\tau\le\widetilde{T}, \tag{2.7c}
$$

$$
v(x_f(\tau),\tau)=g(x_f(\tau),\tau), \quad 0\le\tau\le\widetilde{T}, \tag{2.7d}
$$

$$
e^{(\alpha-1)x_f(\tau)+\beta\tau}\left(\alpha v(x_f(\tau),\tau)+\frac{\partial v}{\partial x}(x_f(\tau),\tau)\right)=1, \quad 0\le\tau\le\widetilde{T}. \tag{2.7e}
$$

3 The Transparent Boundary Condition

The problem (2.7) is posed on an unbounded and time-dependent domain $\Omega(\tau)$:

$$\Omega(\tau) = \left\{(x,\tau) \in \mathbb{R}^2 \mid x < x_f(\tau), 0 \le \tau \le \tilde{T}\right\}$$

In this section we present the derivation of TBC at the artificial boundary $x = a$. For this we split the domain $\Omega(\tau)$ into the bounded time-dependent *interior domain*

$$\Omega_{int}(\tau) = \left\{(x,\tau) \in \mathbb{R}^2 \mid a < x < x_f(\tau), 0 \le \tau \le \tilde{T}\right\}$$

and the unbounded time-independent *exterior domain*

$$\Omega_{ext}(\tau) = \left\{(x,\tau) \in \mathbb{R}^2 \mid x < a, 0 \le \tau \le \tilde{T}\right\}.$$

3.1 Deriviation of TBC

We want to determine the TBC at $x = a < 0$ in such a way that the resulting initial boundary value problem coincides with the solution of the problem (2.7) restricted to Ω_{int}.

In the interior domain Ω_{int}, since $1 - e^{-x} < 0$, we have that

$$g(x,0) = e^{\alpha}(1 - e^{-x})^+ = 0 \quad \forall x \in \Omega_{ext}(0)$$

with $\alpha = \frac{1}{2}(\tilde{r} - \tilde{D_0} + 1)$.

For the derivation of the TBC at $x = a$ we consider the *interior problem*

$$\begin{aligned} \frac{\partial v}{\partial \tau} &= \frac{\partial^2 v}{\partial x^2}, \quad (x,\tau) \in \Omega_{int}(\tau) \\ v(x,0) &= g(x,0), \quad a < x < x_f(0), \\ v_x(a,\tau) &= (T_a v)(a,\tau), \quad 0 \le \tau \le \tilde{T}, \end{aligned} \tag{3.1}$$

together with the boundary conditions (2.7d) and (2.7e) at the free boundary $x = x_f(\tau)$.

We obtain the *Direchlet-to-Nueman* map T_a by solving the exterior problem:

$$\begin{aligned}
\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}, \quad (x,\tau) \in \Omega_{int}(\tau), \\
u(x,0) = 0, \quad x < a, \\
u(a,\tau) = \phi(\tau), \quad 0 \leq \tau \leq \tilde{T}, \quad \phi(0) = 0, \\
u(-\infty,\tau) = 0, \quad 0 \leq \tau \leq \tilde{T}, \\
T_a(\phi)(\tau) = u_x(a,\tau), \quad 0 \leq \tau \leq \tilde{T},
\end{aligned} \tag{3.2}$$

The problem on the exterior domain Ω_{ext} is coupled to the problem on the interior problem on the domain Ω_{int} by the assumption that v and v_x are continuous across the artificial boundary at $x = a$. One can solve (3.2) explicitly by the *Laplace-Method*, i.e., we use the Laplace transformation of u:

$$\widehat{u}(x,s) = \mathcal{L}\left\{u(x,\tau)\right\}(s) = \int_0^{+\infty} u(x,\tau)e^{-s\tau}\, d\tau$$

where we set $s = \zeta + i\xi$, $\xi \in \mathbb{R}^2$, and $\zeta > 0$ is fixed. Hence, we have transformed the exterior problem (3.2) in

$$\begin{aligned}
\widehat{u}_{xx} - s\widehat{u} = 0, \quad x < a, \\
\widehat{u}(a,s) = \widehat{\phi}(s).
\end{aligned} \tag{3.3}$$

The solution of this equation, which decays as $x \to -\infty$, is

$$\widehat{u}(x,s) = \widehat{\phi}(s)e^{\sqrt[+]{s}(x-a)} \text{ for } x < a$$

where $\sqrt[+]{\ }$ denotes the branch of the square root with nonnegative real part. Consequently, the *transformed TBC* is

$$\widehat{u}_x(a,s) = \sqrt[+]{s}\widehat{u}(a,s)$$

and we can one easily compute the TBC. Using the fact that our solutions are equal at the artificial boundary at $x = a$ and applying the inverse Laplace transformation, we can write the TBC at $x = a$ as

$$\begin{aligned} v_x(a,\tau) &= \mathcal{L}^{-1}\left\{\widehat{v_x}\right\}(\tau) \\ &= \mathcal{L}^{-1}\left\{\frac{1}{\sqrt{s}}\left(s\,\widehat{u}\,(a,s)\right)\right\}(\tau) \\ &= \mathcal{L}^{-1}\left\{\frac{1}{\sqrt{s}}\right\} * \mathcal{L}^{-1}\left\{s\,u(a,s)\right\}(\tau) \\ &= \left(\frac{1}{\sqrt{\pi\xi}} * v_\tau(a,\xi)\right)(\tau), \end{aligned}$$

where we use the notation $f * g$ for the convolution between f and g. Finally, the TBC at $x = a$ reads

$$v_x(a,\tau) = \frac{1}{\sqrt{\pi}}\int_0^\tau \frac{v_\tau(a,\xi)}{\sqrt{\tau-\xi}}\,d\xi \tag{3.4}$$

We observe that (3.4) is a memory-type function of τ, i.e., the computation of the solution at some time uses the solution at all previous times. After integration of this equation in space,the solution in Ω_{ext} can also be computed with :

$$v(x,\tau) = -\frac{x-a}{2\sqrt{\pi}}\int_0^\tau e^{-\frac{(x-a)^2}{4(\tau-\xi)}}\,\frac{v(a,\xi)}{(\tau-\xi)^{3/2}},d\xi \tag{3.5}$$

4 Discrete Transparent Boundary Conditions

In this section we shall adress the question how to adequately discretize the analytic TBC (3.4) for a chosen full discretization of (2.7) which in this example will be the *Crank-Nicolson scheme*. Considering the uniform grid points

$$\begin{cases} x_j = a + j\Delta x, & j = 0, 1, \dots \\ \tau_n = n\Delta\tau, & n = 0, 1, \dots \end{cases}$$

and the approximation $v_j^{(n)} \approx v(x_j, \tau_n)$, we can write the *Crank-Nicolson scheme* for solving the heat equation as

$$v_j^{(n+1)} - v_j^{(n)} = \rho \left(v_{j+1}^{(n+1/2)} - 2v_j^{(n)} + v_{j-1}^{(n+1/2)} \right) \tag{4.1}$$

with the abbreviation $v_j^{(n+1/2)} = \left(v_j^{(n+1)} + v_j^{(n)} \right) /2$ and *parabolic mesh ratio* $\rho = \frac{\Delta\tau}{(\Delta x)^2}$. We can show that this scheme is unconditionally stable and has second order accuracy in time and space. Furthermore, it obeys a discrete maximum principle.

In the sequel we present different strategies to incorporate the analytic TBC (3.4) into the finite difference scheme (4.1).

4.1 Discretization strategies for the TBC

In this section we want to compare three strategies to discretize the TBC (3.4). First we review two known discretization techniques from Mayfield and Han and Wu.

4.1.1 Discretized TBC of Mayfield

We first review the *ad-hoc discretization strategy of Mayfield* applied to the heat equation (2.7). We begin by observing that, since v satisfies the heat equation, the TBC (3.4) can be written as

$$v(a, \tau) = \frac{1}{\sqrt{\pi}} \int_0^\tau \frac{v_x(a, \tau)}{\sqrt{\tau - \xi}}\, d\xi,$$

or equivalently

$$v(a, \tau) = \frac{1}{\sqrt{\pi}} \int_0^\tau \frac{v_x(a, \tau - \xi)}{\sqrt{\xi}}\, d\xi.$$

The discretization of this form is

$$\int_0^{\tau_n} \frac{v_x(a, \tau_n - \xi)}{\sqrt{\xi}}\, d\xi \approx \frac{1}{\Delta x} \sum_{m=0}^{n-1} (v(a + \Delta x, \tau_n - \tau_m) - v(a, \tau_n - \tau_m)) \int_{\tau_m}^{\tau_{m+1}} \frac{d\xi}{\sqrt{\xi}}$$
$$= \frac{2}{\Delta x} \sum_{m=0}^{n-1} \left(v_1^{(n-m)} - v_0^{(n-m)} \right) (\sqrt{\tau_{m+1}} - \sqrt{\tau_m})$$
$$= \frac{2\sqrt{\Delta \tau}}{\Delta x} \sum_{m=0}^{n-1} \frac{v_1^{(n-m)} - v_0^{(n-m)}}{\sqrt{m+1} + \sqrt{m}}.$$

Hence, we have

$$v_0^{(n)} = \frac{2\sqrt{\Delta \tau}}{\sqrt{\pi}\Delta x}(v_1^{(n)} - v_0^{(n)}) + \frac{2\sqrt{\Delta \tau}}{\sqrt{\pi}\Delta x} \sum_{m=1}^{n-1} \tilde{l}^{(m)}(v_1^{(n)} - v_0^{(n)}).$$

where the convolution coefficients are given by:

$$\tilde{l}^{(m)} = \frac{1}{\sqrt{m+1} + \sqrt{m}}. \tag{4.2}$$

This approach leads to the following *discretized TBC* for the heat equation

$$v_1^{(n)} - v_0^{(n)} = \frac{\sqrt{\pi}\Delta x}{2\sqrt{\Delta \tau}} v_0^{(n)} - \sum_{m=1}^{n-1} \tilde{l}^{(m)}(v_1^{(n)} - v_0^{(n)}). \tag{4.3}$$

4.1.2 Discretized TBC of Han and Wu

Another way to discretize the analytic TBC was introduced by *Han* and *Wu*. The authors discretized the analytic TBC (3.4) in the following way:

$$\int_0^{\tau_n} \frac{v_\tau(a, \tau_n - \xi)}{\sqrt{\xi}}\, d\xi \approx \sum_{m=0}^{n-1} v_\tau(a, \xi_m) \int_{\tau_m}^{\tau_{m+1}} \frac{d\xi}{\sqrt{\xi}}$$
$$= 2 \sum_{m=0}^{n-1} v_\tau(a, \xi_m)(\sqrt{\tau_n - \tau_{m+1}} - \sqrt{\tau_n - \tau_m})$$
$$= 2\Delta \tau \sum_{m=0}^{n-1} \frac{v_\tau(a, \xi_m)}{\sqrt{\tau_n - \tau_{m+1}} + \sqrt{\tau_n - \tau_m}}$$
$$= \frac{2\Delta \tau}{\sqrt{\Delta \tau}} \sum_{m=0}^{n-1} \frac{v_\tau(a, \xi_m)}{\sqrt{\tau_n - \tau_{m+1}} + \sqrt{\tau_n - \tau_m}}$$

where $\xi_m \in (\tau_m, \tau_{m+1})$. This form leads to the equation

$$\begin{aligned}\frac{v_1^{(n)} - v_{-1}^{(n)}}{\Delta x} &\approx v_x(a, \tau_n) \\ &= \frac{2}{\sqrt{\pi}\sqrt{\Delta\tau}} \sum_{m=0}^{n-1} \frac{v_0^{(m+1)} - v_0^{(m)}}{\sqrt{n-m} + \sqrt{n-m-1}}\end{aligned}$$

which is equivalent to

$$v_1^{(n)} - v_{-1}^{(n)} = \frac{4}{\sqrt{\pi}} \frac{1}{\sqrt{\rho}} \sum_{m=1}^{n} \frac{v_0^{(m)} - v_0^{(m-1)}}{\sqrt{n-m} + \sqrt{n-m+1}} \tag{4.4}$$

By applying a purely implicit scheme to the heat equation at the artificial boundary $x_0 = a$, i.e.,

$$v_1^{(n)} - v_{-1}^{(n-1)} = \rho(v_1^{(n)} - 2v_0^n + v_{-1}^{(n)})$$

one can eliminate the fictitious value $v_{-1}^{(n)}$ in (4.4) to obtain the discretized TBC of *Han and Wu*

$$(1 + 2\rho + B)v_0^n - 2\rho v_1^{(n)} = (1 + B)v_0^{n-1} - B \sum_{m=1}^{n-1} \tilde{l}^{(n-m)}(v_0^{(m)} - v_0^{(m-1)}) \tag{4.5}$$

with the abbreviation $B = 4\frac{\sqrt{\rho}}{\sqrt{\pi}}$ and the convolution coefficients given in (4.2).

4.2 Derivation of the DTBC

In this section, we present a different approach to obtain the TBC. Instead of discretizing the analytic TBC (3.4), the strategy is to derive a *discrete TBC* of the fully discretized problem. For that, we will solve the *discrete exterior problem*, i.e., (4.1) for $j \leq 1$. To do so, we assume for the initial data that $v_j^{(0)} = 0$ for $j \leq 1$, and we apply the Z-transform to solve (4.1) explicitly for $j \leq 1$ fixed:

$$\mathcal{Z}\left\{j^{(n)}\right\} = \widehat{v}_j(z) := \sum_{n=0}^{\infty} v_j^{(n)} z^{-n}$$

Using the *Crank-Nicolson scheme* we have

$$\begin{aligned}
2(z-1)\widehat{v}_j(z) &= \sum_{n=0}^{\infty} 2\left(v_j^{(n)} z^{1-n} - v_j^{(n)} z^{-n}\right) \\
&= \sum_{n=0}^{\infty} 2\left(v_j^{(n+1)} - v_j^{(n)}\right) z^{-n} \\
&= \sum_{n=0}^{\infty} 2\rho\left(v_j^{(n+1/2)} - 2v_j^{(n)} + v_j^{(n+1/2)}\right) z^{-n} \\
&= \rho(z+1)\left(\widehat{v}_{j+1}(z) - 2\widehat{v}_j(z) + \widehat{v}_{j-1}(z)\right)
\end{aligned}$$

We obtain the transformed exterior scheme:

$$\frac{2}{\rho}\frac{z-1}{z+1}\widehat{v}_j(z) = \widehat{v}_{j+1}(z) - 2\widehat{v}_j(z) + \widehat{v}_{j-1}(z). \tag{4.6}$$

The two linearly independent solutions of the resulting *second-order difference equation* (4.6) take the form:

$$\widehat{v}_j(z) = (\alpha_{1,2})^{j+1}(z), \quad j \leq 1,$$

where $\alpha_{1,2}$ are the solutions of the quadratic equation:

$$\alpha^2 - 2\left[1 + \frac{1}{\rho}\frac{z-1}{z+1}\right]\alpha + 1 = 0. \tag{4.7}$$

Since we are seeking decreasing modes as $j \to -\infty$, we have to require $|\alpha_1| > 1$ and obtain the Z-transform discrete TBC as

$$\widehat{v}_1(z) = \alpha_1(z)\widehat{v}_0(z) \tag{4.8}$$

with

$$\alpha_1(z) = 1 + \frac{1-z}{\rho(1+z)} - \sqrt[+]{\frac{1-z}{\rho(1+z)}\left(1 + 2\frac{1-z}{\rho(1+z)}\right)} \tag{4.9}$$

It only remains to compute the inverse Z-transform $\alpha_1(z)$ in order to obtain the discrete TBC from (4.8). This can be performed explicitly, which we will do in the next section, and the discrete TBC becomes

$$v_1^{(n)} = l^{(n)} * v_0^{(n)} = \sum_{k=1}^{n} l^{(n-k)} v_0^k \tag{4.10}$$

where the convolution coefficients $l^{(n)}$ are given by

$$l^{(n)} = \{\mathcal{Z}^{-1}(\alpha_1)\}^{(n)}.$$

4.3 Calculation of $\mathcal{Z}^{-1}(\alpha_1)$

In this section we will compute the inverse Z-transform $\alpha_1(z)$. We start by rewriting α_1 as

$$\begin{aligned}\alpha_1(z) &= 1 + \frac{z-1}{\rho(1+z)} - \frac{1}{\rho(1+z)} \sqrt[+]{(1+2\rho)z^2 - 2z + (1-2\rho)} \\ &= 1 + \frac{z-1}{\rho(1+z)} - \frac{1}{\rho(1+z)} \sqrt[+]{(az^2 - 2z + b)}\end{aligned}$$

with the constants

$$a = 1 + 2\rho, b = 1 - 2\rho$$

For the inverse Z-transform we use

$$\sqrt[+]{az^2 - 2z + b} = \frac{1}{\sqrt[+]{a}} \frac{az^2 - 2z + b}{z} \frac{\frac{\sqrt[+]{a}}{\sqrt[+]{b}}}{\sqrt[+]{\frac{a}{b}z^2 - 2\frac{1}{b} + 1}}$$

with the following abbreviations

$$\begin{aligned}F(z,\mu) &= \frac{z}{\sqrt[+]{z^2 - 2\mu z + 1}} \\ \lambda &= \frac{\sqrt{a}}{\sqrt[+]{b}} = \frac{\sqrt{1+2\rho}}{\sqrt[+]{1-2\rho}} \\ \mu &= \frac{1}{\sqrt[+]{b}\sqrt{a}} = \frac{1}{\sqrt[+]{1+2\rho}\sqrt{1-2\rho}}\end{aligned} \tag{4.11}$$

Then, we obtain

$$\begin{aligned}\frac{1}{z+1} \sqrt[+]{az^2 - 2z + b} &= \frac{1}{\sqrt[+]{a}} \frac{az^2 - 2z + b}{\rho z(z+1)} F(\lambda z, \mu) \\ &= \frac{1}{\rho\sqrt{a}}(a + \frac{b}{z} - \frac{a+b+2}{z+1})F(\lambda z, \mu) \\ &= \frac{\sqrt[+]{b}}{\rho}(\lambda + \frac{\lambda^{-1}}{z} - 4\frac{\mu}{z+1})F(\lambda z, \mu)\end{aligned}$$

The inversion rules now yields

$$
\begin{aligned}
l^{(n)} &= \frac{\rho+1}{\rho}\delta_0^n + \frac{2}{\rho}(-1)^n + \sqrt[+]{1-2\rho}\left(\lambda\delta_0^n + \lambda^{-1}\delta_0^{n-1} - 4\mu\left[(-1)^n - \delta_0^n\right]\right) * \tilde{P}_n(\mu) \\
&= \frac{\rho+1}{\rho}\delta_0^n - \frac{\sqrt{1+2\rho}}{\rho}\left(\lambda\tilde{P}_n(\mu) + \lambda^{-1}\tilde{P}_{n-1}(\mu) - 2\mu\sum_{k=0}^{k=n-1}(-1)^{n-k}\tilde{P}_k(\mu)\right)
\end{aligned}
$$

where here $*$ denotes the discrete convolution and $\tilde{P}_n(\mu) = \lambda^{-n}P_n(\mu)$ denotes the *"damped"* Legendre polynomials ($\tilde{P}_0 \equiv 1, \tilde{P}_{-1} \equiv 0$). Since the assymptotical behavior $l^{(n)} \sim 4(-1)^n/\rho$ of the convolution coefficients may lead to subtractive cancellation in in (4.10) we prefer to use following *summed coefficients* in the implementation:

$$
\begin{aligned}
s^{(n)} &= l^{(n)} + l^{(n-1)}, \quad n \geq 1, \\
s^{(0)} &= l^{(0)}.
\end{aligned}
\tag{4.12}
$$

Thus, we obtain

$$
\begin{aligned}
s^n &= l^{(n)} + l^{(n-1)} \\
&= -\frac{\sqrt{1+2\rho}}{\rho}\lambda^{-n}\left(\tilde{p}_n - 2\mu\tilde{p}_{n-1} + \tilde{p}_{n-2}\right)
\end{aligned}
$$

Knowing the recurrence relation of the Legendre polynomials

$$
\mu P_{n-1} = \frac{nP_n + (n-1)P_{n-2}}{2n-1}
$$

we finally get

$$
s^{(n)} = -\frac{\sqrt{1+2\rho}}{\rho}\frac{\tilde{P}_n(\mu) - \lambda^{-2}\tilde{P}_{n-2}(\mu)}{2n-1}
$$

The TBC then reads

$$
v_1^{(n)} - s^{(0)}v_0^{(n)} = \sum_{k=1}^{n-1} s^{(n-k)}v_0^{(k)} - v_1^{(n-1)}, n \geq 1 \tag{4.13}
$$

with the convolution coeficients:

$$
\begin{aligned}
s^{(0)} &= \alpha_1(0) = 1 + \frac{1+\sqrt{1+2\rho}}{\rho}, \\
s^{(1)} &= \alpha_1{}'(0) + \alpha_1(0) = 1 - \frac{1}{\rho} - \frac{1}{\rho\sqrt{1+2\rho}}, \\
s^{(n)} &= -\frac{\sqrt{1+2\rho}}{\rho}\frac{\tilde{P}_n(\mu) - \lambda^{-2}\tilde{P}_{n-2}(\mu)}{2n-1}, \ n \geq 2
\end{aligned}
\tag{4.14}
$$

where $\tilde{P}_n(\mu) = \lambda^{-n} P_n(\mu)$ denotes the *"damped"* Legendre polynomials ($\tilde{P}_0 \equiv 1, \tilde{P}_{-1} \equiv 0$). The parameters λ, μ are given by:

$$\lambda = \frac{\sqrt{1+2\rho}}{\sqrt[+]{1-2\rho}}, \quad \mu = \frac{1}{\sqrt[+]{1-2\rho}\sqrt{1+2\rho}}$$

Using the recurrence relation of the Legendre polynomials

$$\mu \tilde{p}_{n-1} = \frac{n\lambda \tilde{p}_n + (n-1)\lambda^{-1}\tilde{p}_{n-2}}{2n-1},$$

the convolution coefficients can be computed by the recursion formula:

$$s^{(n+1)} = \frac{2n-1}{n+1}\mu\lambda^{-1}s^{(n)} - \frac{n-2}{n+1}\lambda^{-2}s^{(n-1)}, \quad n \geq 2, \tag{4.15}$$

which can be used after calculating $s^{(n)}, n = 0, 1, 2$ by the formula (4.14).

4.4 The asymptotic behaviour of $s^{(n)}$

We finish this chapter, by studying *the asymptotic behaviour* of $s^{(n)}$. For that we introduce the notation

$$\mu = cos(\theta), 0 \leq \theta \leq \pi$$

since $|\mu| \leq 1$ and recall a classical result on the asymptotic property of the Legendre polynomials:

Lemma 4.1 *(Laplace Formula)*

$$P_n(\cos(\theta)) = \frac{\sqrt{2}}{\sqrt{\pi}\cos(\theta)} \frac{\cos\left((n+\frac{1}{2})\theta - \frac{\pi}{4}\right)}{\sqrt{n}} + \mathcal{O}(n^{-3/2}). \tag{4.16}$$

From this lemma we conclude that

$$\begin{aligned} s^{(n)} &= -\frac{\sqrt{1+2\rho}}{\rho(2n-1)}\lambda^{-n}\left(P_n(\cos(\theta)) - P_{n-2}(\cos(\theta))\right) \\ &= 2\lambda^{-n}\frac{\sqrt{1+2\rho}}{\rho}\frac{\sqrt{\sin(\theta)}}{\sqrt{\pi}}\frac{\sin\left((n-\frac{1}{2})\theta - \frac{\pi}{4}\right)}{\sqrt{n}(2n-1)} + \mathcal{O}(n^{-3/2}). \end{aligned}$$

Since $|\lambda|^{-1} = |\frac{\sqrt{1-2\rho}}{\sqrt{1+2\rho}}| \leq 1$ we would roughly expect

$$|2\lambda^{-n}\frac{\sqrt{1+2\rho}}{\rho}\frac{\sqrt{\sin(\theta)}}{\sqrt{\pi}}\frac{\sin\left((n-\frac{1}{2})\theta-\frac{\pi}{4}\right)}{\sqrt{n}(2n-1)}| \leq 2\frac{\sqrt{1+2\rho}}{\rho}\frac{1}{\sqrt{\pi}}\frac{1}{\sqrt{n}(2n-1)} = \mathcal{O}(n^{-3/2})$$

Hence, we finally have

$$s^{(n)} = \mathcal{O}(n^{-3/2}). \tag{4.17}$$

This rapid decay property of $s^{(n)}$ motivates a simplified discrete TBC by restricting (4.13) to a convolution over the "recent past"(last M time levels) :

$$v_1^{(n)} - s^{(0)}v_0^{(n)} = \sum_{k=n-M}^{n-1} s^{(n-k)}v_0^{(k)} - v_1^{(n-1)}, n \geq 1 \tag{4.18}$$

We note that the stability of the resulting scheme is still not proven yet.

5 Approximation by Sums of Exponentials

The evaluation of (4.13) is computationally expensive as n grows. Therefore, in order to derive a fast numerical method to calculate, in particular the discrete convolution in (4.13), we approximate the coefficients $s^{(n)}$ by the following (sum of exponentials):

$$s^{(n)} \approx \tilde{s}^{(n)} := \begin{cases} s^{(n)}, & n = 0, 1 \\ \sum\limits_{l=1}^{L} b_l q_l^{-n}, & n = 2, 3, \dots \end{cases} \tag{5.1}$$

where $L \in \mathbb{N}$ is a fixed number. It is important to observe that the approximation properties of $\tilde{s}^{(n)}$ depend on L, and the corresponding set b_l, q_l. We now propose deterministic method for finding b_l, q_l for fixed L. It is important to pay attention to the 'split' definition of $\tilde{s}^{(n)}$ in (5.1). This is motivated by the different nature of the first two coefficients in (4.14). Their inclusion, would lead in the discrete-sum-exponential would yield less accurate approximation results.

Let us fix L and consider the formal power series:

$$f(x) := s^{(2)} + s^{(3)}x + s^{(4)}x^2 \dots, \quad |x| \leq 1.$$

If there is the $[L-1, L]$ Padé approximation

$$\tilde{f}(x) := \frac{P_{L-1}(x)}{Q_L(x)}$$

of f, then its Taylor series

$$\tilde{f}(x) = \tilde{s}^{(2)} + \tilde{s}^{(3)}x + \tilde{s}^{(4)}x^2 \dots \tag{5.2}$$

satisfies the conditions

$$\tilde{s}^{(n)} = s^{(n)}, \quad n = 2, 3, \dots, 2L+1 \tag{5.3}$$

due to the definition of the Padé approximation (see appendix A).

Theorem 5.1 *Let $Q_L(x)$ have L simple roots q_l with $|q_l| > 1$, $l = 1, \dots, L$. Then*

$$\tilde{s}^{(n)} = \sum_{l=1}^{L} b_l q_l^{-n}, \quad n = 2, 3, \dots$$

where

$$b_l = -\frac{P_{L-1}(q_l)}{Q'_L(q_l)} q_l \neq 0, \quad l = 1, \ldots, L. \tag{5.4}$$

Proof:

We start by observing that b_l is well defined since q_l is a simple root for $l = 1, \ldots, L$. By hypothesis, we can write

$$Q_L(x) = \prod_{l=1}^{L} (x - q_l).$$

Since

$$Q'_L(x) = \sum_{k=1}^{L} \prod_{l \neq k} (x - q_l)$$

we get

$$Q'_L(q_l) = \prod_{k \neq l} (q_l - q_k).$$

Hence, using (5.4), we have that

$$\begin{aligned}
\sum_{l=1}^{L} \frac{b_l q_l^{-1}}{q_l - x} &= \frac{1}{Q_L(x)} \sum_{l=1}^{L} -b_l q_l^{-1} \prod_{k \neq l} (x - q_k) \\
&= \frac{1}{Q_L(x)} \sum_{l=1}^{L} \frac{P_{L-1}(q_l)}{Q'_L(q_l)} \prod_{k \neq l} (x - q_k) \\
&= \frac{1}{Q_L(x)} \sum_{l=1}^{L} P_{L-1}(q_l) \prod_{k \neq l} \frac{x - q_k}{q_l - q_k} \\
&=: \frac{R_{L-1}(x)}{Q_L(x)}.
\end{aligned}$$

Since R_{L-1} has degree $L - 1$ and

$$R_{L-1}(q_l) = P_{L-1}(q_l), \quad l = 1, \ldots, L$$

we have $R_{L-1} \equiv P_{L-1}$, where we use once again use the fact that the roots q_l are simple roots. Therefore, we proved that

$$\frac{P_{L-1}(x)}{Q_L(x)} = \sum_{l=1}^{L} \frac{b_l q_l^{-1}}{q_l - x}.$$

Finally, using

$$\frac{1}{q_l - x} = \sum_{n=0}^{\infty} x^n q_l^{-n+1}$$

with $|x| < |q_l|$, we have

$$\begin{aligned}
\frac{P_{L-1}(x)}{Q_L(x)} &= \sum_{n=0}^{\infty} \sum_{l=1}^{L} b_l q_l^{-1} x^n q_l^{-n+1} \\
&= \sum_{n=0}^{\infty} \left(\sum_{l=1}^{L} b_l q_l^{-n} \right) x^n.
\end{aligned}$$

□

It follows that the set b_l, q_l defined in the Theorem 5.1 can be used in (5.1) at least for $n = 2, 3, \ldots, 2L + 1$. The main question now is to know if we can use these b_l, q_l also for $n > 2L + 1$, that is, how good is the approximation

$$\tilde{s}^{(n)} \approx s^{(n)}, \quad n > 2L + 1.$$

Before answer this question, we will see how these "exponential" coefficients help us compute the convolution in (4.13) in a much more efficient way. Let us define

$$C^{(n)} := \sum_{m=1}^{n-1} \tilde{s}^{(n-m)} v_0^{(m)}. \tag{5.5}$$

What happens now is that the $C^{(n)}$ coefficients can be calculated by a recurrence formulas, as the following proposition shows us, and therefore this will reduce the numerical effort significantly.

Proposition 5.2 *The value $C^{(n)}$ from (5.5) for $n \geq 2$ can be represented by*

$$C^{(n)} = \sum_{l=1}^{L} C_l^{(n)}, \tag{5.6}$$

where

$$\left.\begin{aligned} C_l^{(1)} &= 0, \\ C_l^{(n)} &= q_l^{-1} C_l^{(n-1)} + b_l q_l^{-1} v_0^{(n-1)}, \quad n = 2, 3, \ldots, \end{aligned}\right. \qquad l = 1, \ldots, L \tag{5.7}$$

Proof:

A straightforward calculation yields:

$$\begin{aligned} C^{(n)} &= \sum_{m=1}^{n-1} \tilde{s}^{(n-m)} v_0^{(m)} \\ &= \sum_{m=1}^{n-1} v_0^{(m)} \sum_{l=1}^{L} b_l q_l^{-(n-m)} \\ &= \sum_{l=1}^{L} \sum_{m=1}^{n-1} b_l q_l^{-(n-m)} v_0^{(m)} \\ &= \sum_{l=1}^{L} C_l^{(n)} \end{aligned}$$

where

$$C_l^{(n)} := \sum_{m=1}^{n-1} b_l q_l^{-(n-m)} v_0^{(m)}.$$

For $n > 2$ and for each $l = 1, \ldots, L$, we have that the following recursion formula:

$$\begin{aligned} C_l^{(n)} &= \sum_{m=1}^{n-2} b_l q_l^{-(n-m)} v_0^{(m)} + b_l q_l^{-1} v_0^{(n-1)} \\ &= q_l^{-1} C_l^{(n-1)} + + b_l q_l^{-1} v_0^{(n-1)}. \end{aligned}$$

As for $n = 2$, for each $l = 1, \ldots, L$,

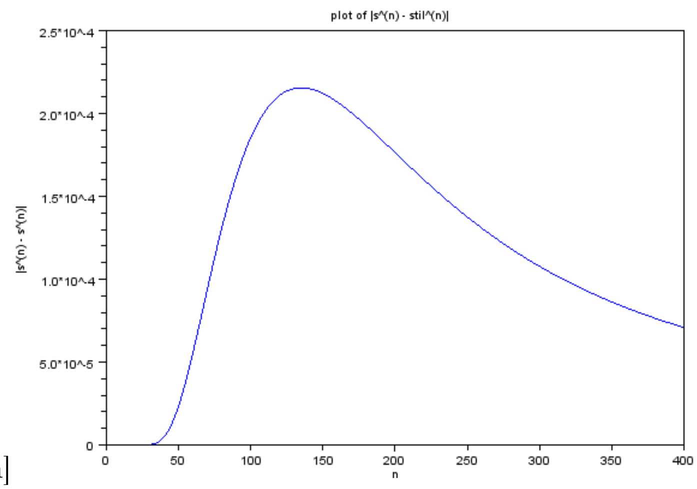
$$C_l^{(2)} = b_l q_l^{-1} v_0^{(1)}$$

hence

$$C_l^{(1)} = 0.$$

$\square$

We now summarize how one should evaluate the DTBC:



[h]

Figure 1: Absolute error $|s^{(n)} - \tilde{s}^{(n)}|$ of the approximated coefficients (5.1) for $L = 10$ and $\rho = 1$.

1. Fix L in (5.1) and calculate $s^{(n)}$, $n = 0, \dots, 2L + 1$, by formula (4.15).

2. Find the $[L - 1|L]$ Padé Approximation of (5.2) and compute $\{b_l, q_l\}$ for (5.1) in accordance with Theorem 5.1.

3. Use the recurrence formulas (5.6) and (5.7) to calculate the approximate convolutions in (4.13).

It only rest us now to see how good is the approximation

$$\tilde{s}^{(n)} \approx s^{(n)}, \quad n > 2L + 1.$$

The coefficients $\{b_l, q_l\}$ for $L = 5, 10$ and $\rho = 1$ can be seen in Table 1. For $L = 5$, we got roughly the same values as in [1] for $\{q_l\}$ except for one which is largely different. Turns out that this particular value is not in fact a root of the polynomial Q_5, since in fact this polynomial has only degree 4 (the leading coefficient is very close to 0). For this reason we don't have 5 different simple roots and therefore we can't use the $\tilde{s}^{(n)}$ as in (5.1). For $L = 10$, we get roughly the same values as in [1]. However, as we can see in Figures 1 and 2, although the absolute error is very close to 0, the relative error goes to 1, which means that in fact we can't use the $\tilde{s}^{(n)}$ to approximate $s^{(n)}$. We also tried to obtain the coefficients $\{b_l, q_l\}$ for $L = 20$ but in this case some of the q_l are complex and therefore won't yield good approximations for $s^{(n)}$, which are always real.

	q_l	b_l
$L = 5$	1.0980762	-0.0386751
	1.5	-0.2165064
	-4.0980762	-0.5386751
	3	-1.1547005
	-6.095×10^{12}	-3.575×10^{24}
$L = 10$	1.0277542	-0.0058673
	1.1183568	-0.0255837
	1.2988579	-0.0674318
	1.6386011	-0.1554760
	2.3353035	-0.3739261
	-3.2644678	-0.0591949
	-4.3955591	-0.3952131
	4.1936326	-1.2058128
	-9.6861225	-3.7500851
	17.733647	-18.210127

Table 1: Coefficients $\{q_l, b_l\}$ of the sum of exponentials 5.1 for $\rho = 1$.

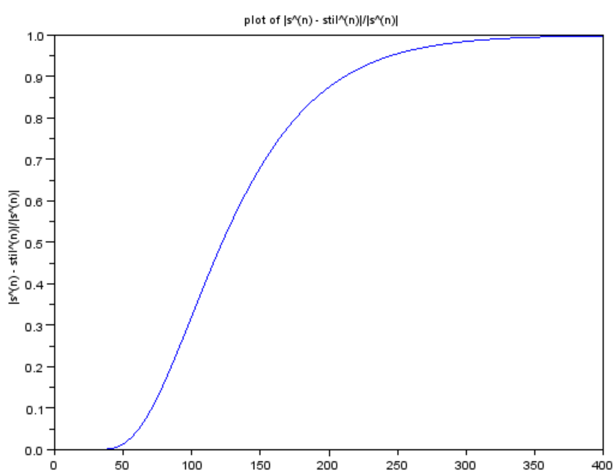


Figure 2: Relative error $\frac{|s^{(n)} - \tilde{s}^{(n)}|}{|s^{(n)}|}$ of the approximated coefficients (5.1) for $L = 10$ and $\rho = 1$.

6 Numerical treatment of the free boundary

In this section we shall describe how to treat numerically the free boundary in (2.7), but first we need an additional result that you allows to determine the position of the free boundary.

6.1 Some properties of the solution of the Black-Scholes equation:

We start this section by recalling what we did in chapter 2. The value of an American call option satisfies the Black-Scholes equation (2.1a):

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - D_0) S \frac{\partial V}{\partial S} - rV = 0, \quad 0 < S < S_f(t), \quad 0 \le t < T, \tag{6.1a}$$

where $S_f(t)$ is the *early exercise boundary*, and the final boundary conditions are given by:

$$V(S,T) = h(S), \quad 0 \le S \le S(T) = S_0 \tag{6.1b}$$

$$V(S_f(t),t) = h(S_f(t)), \quad \frac{\partial V}{\partial S} = 1, \quad 0 \le t < T, \tag{6.1c}$$

$$V(0,t) = 0, \quad 0 \le t \le T, \tag{6.1d}$$

where $h(S) = (S - E)^+$ with $E > 0$ and $S_0 = \max(E, rE/D_0)$. We introduced a change of variable for t:

$$t = T - \frac{2\tau}{\sigma^2}$$

and defined

$$\widetilde{V}(S,\tau) = V(S,t) = V(S, T - \frac{2\tau}{\sigma^2}), \quad \widetilde{S}_f(\tau) = S_f(T - \frac{2\tau}{\sigma^2}),$$
$$\widetilde{r} = \frac{2}{\sigma^2} r, \quad \widetilde{D}_0 = \frac{2}{\sigma^2} D_0, \quad \widetilde{T} = \frac{\sigma^2}{2} T.$$

Defining the operator L as

$$L\widetilde{V} := -\frac{\partial \widetilde{V}}{\partial \tau} + S^2 \frac{\partial^2 V}{\partial S^2} + (\widetilde{r} - \widetilde{D}_0) \frac{\partial \widetilde{V}}{\partial S} - \widetilde{r}\widetilde{V},$$

we can say that (6.1) is equivalent to the following problem

$$L\widetilde{V} = 0, \quad 0 < S < \widetilde{S}_f(\tau), \quad 0 < \tau < \widetilde{T} \tag{6.2a}$$

$$\widetilde{V}(S, 0) = h(S), \quad 0 \leq S \leq \widetilde{S}_f(0), \tag{6.2b}$$

$$\widetilde{V}(\widetilde{S}_f(\tau), \tau) = h(\widetilde{S}_f(\tau)), \quad \frac{\partial \widetilde{V}}{\partial S}(\widetilde{S}_f(\tau), \tau) = 1, \quad 0 \leq \tau \leq \widetilde{T}, \tag{6.2c}$$

$$\widetilde{V}(0, \tau) = 0, \quad 0 \leq \tau \leq \widetilde{T}. \tag{6.2d}$$

Furthermore, we defined

$$\alpha = -\frac{1}{2}(\widetilde{r} - \widetilde{D}_0 - 1), \quad \beta = -\alpha^2 - \widetilde{r},$$

and introduced a second the change of variables, this time in S:

$$S = Ee^x, \quad \widetilde{V}(S, \tau) = Ee^{\alpha x + \beta \tau} v(x, \tau).$$

Then the free boundary value problem (6.2) is equivalent to the following problem:

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2}, \quad -\infty \leq x \leq x_f(\tau) \tag{6.3a}$$

$$v(x, 0) = g(x, 0), \quad -\infty < x < x_f(0), \tag{6.3b}$$

$$v(x_f(\tau), \tau) = g(x_f(\tau), \tau), \quad 0 \leq \tau \leq \widetilde{T} \tag{6.3c}$$

$$\alpha v(x_f(\tau), \tau) + \frac{\partial v(x_f(\tau), \tau)}{\partial x} = e^{(1-\alpha)x_f(\tau) - \beta\tau}, \quad 0 \leq \tau \leq \widetilde{T} \tag{6.3d}$$

$$\lim_{x \to -\infty} v(x, \tau) = 0, \quad 0 \leq \tau \leq \widetilde{T} \tag{6.3e}$$

where

$$g(x, \tau) = e^{-\alpha x - \beta \tau}(e^x - 1)^+,$$
$$x_f(\tau) = \log\left(\frac{\widetilde{S}_f(\tau)}{E}\right).$$

It is known that $x_f(\tau) > 0$ for $\tau > 0$.

We know consider the problem (6.2) and define

$$W = \frac{\partial \widetilde{V}(S, \tau)}{\partial S}.$$

Then one can show that W satisfies

$$-\frac{\partial W}{\partial \tau}+S^2\frac{\partial^2 W}{\partial S^2}+\left(2-\widetilde{r}-\widetilde{D}_0\right)S\frac{\partial W}{\partial S}-\widetilde{D}_0 W=0, \quad 0<S<\widetilde{S}_f,\ 0<\tau\leq \widetilde{T}, \tag{6.4a}$$

$$W(S,0)=0, \quad 0<S<E, \tag{6.4b}$$

$$W(S,0)=\widetilde{S}_f(0), \tag{6.4c}$$

$$W(\widetilde{S}_f,\tau)=1, \quad 0<\tau<\widetilde{T} \tag{6.4d}$$

$$\lim_{S\to 0} W(S,\tau)=0, \quad 0<\tau\leq \widetilde{T}. \tag{6.4e}$$

By strong maximum principle of the parabolic equation, we have the following theorem:

Theorem 6.1 *$W(S,\tau)$ satisfies the following inequality:*

$$0<W(S,\tau)<1, \quad 0<S<\widetilde{S}_f(\tau),\ 0<\tau\leq \widetilde{T}.$$

Namely,

$$0<\frac{\partial \widetilde{V}(S,\tau)}{\partial S}<1, \quad 0<S<\widetilde{S}_f(\tau),\ 0<\tau\leq \widetilde{T}$$

and

$$0<e^{(\alpha-1)x+\beta\tau}\left(\frac{\partial v(x,\tau)}{\partial x}+\alpha v(x,\tau)\right)=<1, \quad -\infty<x<x_f(\tau),\ 0<\tau\leq \widetilde{T}.$$

For the solution $\left\{\widetilde{V}(S,\tau),\widetilde{S}_f(\tau)\right\}$ of the problem (6.2), we extend $\widetilde{V}(S,\tau)$ to the domain

$$\widetilde{S}_f(\tau)<S<+\infty, \quad 0\leq \tau\leq \widetilde{T}$$

by

$$\widetilde{V}(S,\tau)=g(S), \quad \widetilde{S}_f(\tau)<S<+\infty, \quad 0\leq\tau\leq\widetilde{T}.$$

For a given smooth boundary $S=\widehat{S}(\tau)$ and a given τ_j such that $0<\tau_j<\widetilde{T}$, satisfying

$$\widetilde{S}_f(\tau)<\widehat{S}(\tau), \quad \tau_j\leq\tau\leq\widetilde{T}$$

consider the following auxiliary problem

$$L\widehat{V} = 0, \quad 0 < S < \widehat{S}(\tau), \quad \tau_j < \tau < \widetilde{T}, \tag{6.5a}$$

$$\widehat{V}(\widehat{S}(\tau), \tau) = h(\widehat{S}(\tau)), \quad \tau_j < \tau < \widetilde{T}, \tag{6.5b}$$

$$\widehat{V}(0, \tau) = 0, \quad \tau_j \le \tau \le \widetilde{T} \tag{6.5c}$$

$$\widehat{V}(S, \tau_j) = \widetilde{V}(S, \tau_j), \quad 0 \le S \le \widehat{S}(\tau_j). \tag{6.5d}$$

Hence, problem (6.5) has a solution $\widehat{V}(S,\tau)$ on

$$\Omega_j = \left\{ (S,\tau) \in \mathbb{R}^2 : 0 < S < \widehat{S}(\tau), \tau_j \le \tau \le \widetilde{T} \right\}.$$

If we let

$$\varepsilon(S,\tau) := \widetilde{V}(S,\tau) - \widehat{V}(S,\tau)$$

we can show that

$$L\varepsilon(S,\tau) = \begin{cases} 0, & 0 < S < \widetilde{S}_f(\tau),\ \tau_j < \tau \le \widetilde{T}. \\ -D^* S + r^* E \le 0, & \widetilde{S}_f(\tau) \le S \le \widehat{S}(\tau),\ \tau_j < \tau \le \widetilde{T}. \end{cases}$$

and that $\frac{\partial \varepsilon(S,\tau)}{\partial S}$ is continuous on $S = \widetilde{S}_f(\tau),\ \tau_j < \tau \le \widetilde{T}$. From the strong maximum principle we get that

$$\varepsilon(S,\tau) > 0, \quad 0 < S < \widehat{S}(\tau),\ \tau_j < \tau \le \widetilde{T}$$

So when $\widetilde{S}_f(\tau) < S < \widehat{S}(\tau)$ we have

$$\varepsilon(S,\tau) = \widetilde{V}(S,\tau) - \widehat{V}(S,\tau) = h(S) - \widehat{V}(S,\tau) > 0$$

namely

$$\widehat{V}(S,\tau) < h(S), \quad S_f^*(\tau) < S < \widehat{V}(\tau),\ \tau_j < \tau \le \widetilde{T}. \tag{6.6}$$

Now we define $\widehat{V} = Ee^{\alpha x + \beta\tau}\widehat{v}(x,\tau)$, with $S = Ee^x$. Then the auxiliary problem (6.5) is equivalent to the problem:

$$\frac{\partial \widehat{v}}{\partial \tau} = \frac{\partial^2 \widehat{v}}{\partial x^2}, \quad -\infty < x < \widehat{x}_f(\tau),\ \tau_j < \tau < \widetilde{T} \tag{6.7a}$$

$$\widehat{v}(\widehat{x}_f(\tau), \tau) = g(\widehat{x}_f(\tau), \tau), \quad \tau_j < \tau < \widetilde{T}, \tag{6.7b}$$

$$\lim_{x\to-\infty} \widehat{v}(x, \tau) = 0, \quad \tau_j \leq \tau \leq \widetilde{T} \tag{6.7c}$$

$$\widehat{v}(x), 0) = v(x, 0), \quad -\infty < x < \widehat{x}_f(\tau),\ \tau_j \leq \tau \leq \widetilde{T} \tag{6.7d}$$

where

$$\widehat{x}_f(\tau) = \log\left(\widehat{S}(\tau)/E\right) \geq x_f(\tau), \quad \tau_j \leq \tau \leq \widetilde{T}.$$

Finally, from inequality (6.6), we obtain the following theorem:

Theorem 6.2 *For the solution $\widehat{v}(x,\tau)$ of the auxiliary problem (6.7) the following inequality holds:*

$$\widehat{v}(x,\tau) < g(x,\tau),\ x_f(\tau) < x < \widehat{x}_f,\ \tau_j < \tau < \widetilde{T}.$$

This inequality is very useful for determining the location of the free boundary in the numerical schemes.

6.2 Finite difference approximation

In this section we will see how can we deal with the free boundary. The idea is to first solve our problem in a larger domain where we know that our free boundary is contained. Then using that solution, we are to determine the position of the free boundary and hence compute the solution of our original problem.

From Theorems 6.1 and 6.2 we know that for a given τ the free boundary is the only point satisfying the partial differential equation and the condition

$$e^{(1-\alpha)x+\beta\tau}\left(\frac{\partial v(x,\tau)}{\partial x} + \alpha u(x,\tau)\right) = 1$$

and if the boundary condition $v(x,\tau) = g(x,\tau)$ is given at $x > x_f(\tau)$, then $v(x,\tau) < g(x,\tau)$ will occur on the left of the boundary. Let J be the largest number such that

$$v_J^{(n)} \geq g_J^{(n)}.$$

Hence, we know that $x_f(\tau_n) \in [x_J, x_{J+1}]$. We will then consider the approximation

$$x_f(\tau_n) \approx x_{J+1}. \tag{6.8}$$

Hence, to solve our problem we just have to know a a priori bound for the free boundary. According to [3], the free boundary $S_f(t)$ is a nondecreasing function and

$$S_f(T) \leq S_f(t) \leq S_f^*, \quad 0 \leq t \leq T,$$

with

$$S_f^* = \frac{\sqrt{-\beta} + \alpha}{\sqrt{-\beta} + \alpha - 1} E. \tag{6.9}$$

Thus, if we set $x_f^* = \log(S_f^*/E)$, then the free boundary $x_f(\tau)$ has the property

$$0 \leq x_f(\tau) \leq x_f^*, \quad 0 \leq t \leq \widetilde{T}. \tag{6.10}$$

We will then star by solving our problem in $[a, x_{max}]$ with $x_{max} = x_{jMax}$ and $x_{max} \geq x_f^*$. Finally, we see how we can efficiently solve the linear systems that arise. The *Crank-Nicolson scheme* (4.1) can also be written as

$$-\frac{\rho}{2} v_{j-1}^{(n)} + (1+\rho) v_j^{(n)} - \frac{\rho}{2} v_{j+1}^{(n)} = \widetilde{b}_j, \quad j = 1, 2, \ldots \tag{6.11}$$

with

$$\widetilde{b}_j = \frac{\rho}{2}\left(v_{j-1}^{(n-1)} + v_{j+1}^{(n-1)} + (1-\rho) v_j^{(n-1)}\right), \quad j = 1, 2, \ldots \tag{6.12}$$

Hence we will have a tridiagonal matrix which we can solve using the Thomas Algorithm. As for the second system, we can simplify things. We start by writing (6.11) as

$$(1+\rho) v_1^{(n)} - \frac{\rho}{2} v_2^{(n)} = b_1 \tag{6.13}$$

$$-\frac{\rho}{2} v_{j-1}^{(n)} + (1+\rho) v_j^{(n)} - \frac{\rho}{2} v_{j+1}^{(n)} = b_j, \quad j = 2, 3, \ldots \tag{6.14}$$

where

$$\begin{aligned} b_1 &= \frac{\rho}{2}(v_0^{(n-1)} + v_2^{(n-1)}) + (1-\rho) v_1^{(n-1)} + \frac{\rho}{2} v_0^{(n)}, \\ b_j &= \frac{\rho}{2}(v_{j-1}^{(n-1)} + v_{j+1}^{(n-1)} + (1-\rho) v_j^{(n-1)} \quad j = 2, 3, .. \end{aligned} \tag{6.15}$$

If we let $c_1 = 1 + \rho$ and $y_1 = b_1$ then we have

$$c_1 v_1^{(n)} - \frac{\rho}{2} v_2^{(n)} = y_1$$

Solving for $v_1^{(n)}$ and substituting into (6.14) we get

$$c_2 v_2^{(n)} - \frac{\rho}{2} v_2^{(n)} = y_2,$$

where

$$c_2 = 1 + \rho - \frac{\rho^2}{4c_1}, \quad y_2 = b_2 + \frac{\rho y_1}{2c_1}$$

In general, we have

$$c_j v_j^{(n)} - \frac{\rho}{2} v_{j+1}^{(n)} = y_j$$

where

$$c_j = 1 + \rho - \frac{\rho^2}{4c_{j-1}}, \quad y_j = b_j + \frac{\rho y_{j-1}}{2c_{j-1}} \tag{6.16}$$

By (6.8) we have that

$$v_{J+1}^{(n)} = g_{J+1}^{(n)}.$$

Hence,

$$\begin{aligned} v_J^{(n)} &= \frac{1}{c_J}\left(b_J + \frac{\rho}{2} g_{J+1} + \frac{\rho y_{J-1}}{2c_{J-1}}\right) \\ v_j^{(n)} &= \frac{1}{c_n}\left(y_j + \frac{\rho}{2} v_{j+1}^{(n)}\right), \quad j = J-1, J-2, ... \end{aligned} \tag{6.17}$$

Thus we have the following algorithm.

Algorithm

Choose x_{max} such that $x_{max} \geq x_f^*$. Then at each time step, do the following:

1. Set up the linear system given by (6.11) for $j = 1, \ldots, jMax$, together with $v_{jMax}^{(n)} = g_{jMax}$ and (4.3), (4.5) or (4.10).
2. Solve the linear system define in 2. using the Thomas Algorithm.
3. Determine the largest J such that $v_J^{(n)} \geq g_J^{(n)}$.
4. Use back substitution (6.17) to find all solutions in $[a, x_J]$.

7 Stability anlysis of the artificial boundary condition

Here we analyze the stability of the Crank-Nicolson scheme (4.1) along with the DTBC (4.13) or its approximated version. Since we will focus on the fact that the (approximated) DTBC does not destroy the unconditional stability of the underlying finite difference scheme, we consider the following problem on the half-space $j \geq 0$:

$$
\begin{aligned}
&v_j^{(n+1)} - v_j^{(n)} = \rho(v_j^{(n+1/2)} - 2v_j^{(n+1/2)} + v_{j-1}^{(n+1/2)}), \quad j \geq 1\\
&v_j^0 = g(x_j, 0), \quad j = 0, 1, 2, \dots \text{ with } v_0^0 = v_1^0 = 0,\\
&\widehat{v}_1(z) = \alpha_1(z)\widehat{v}_0(z).
\end{aligned}
\tag{7.1}
$$

where the transformed boundary kernel $\alpha_1(z) = \nu_1(z)$ is given by (4.9). In the sequel we want to bound the exponential growth of solutions to the numerical scheme (7.1) for a fixed mesh ratio. We will prove an estimate of the discrete solution to (7.1) in the discrete l^2-norm:

$$
\|v^{(n)}\|_2^2 := \Delta x \sum_{j=1}^{\infty} \left| v_j^{(n)} \right|^2. \tag{7.2}
$$

Theorem 7.1 ***(Growth condition)*** *The transformed boundary kernel $\alpha_1(z)$ satisfy*

$$
\Re\left\{\alpha_1(\beta e^{i\varphi})\right\} \geq 1, \quad \forall \;\; 0 \leq \varphi \leq 2\pi \tag{7.3}
$$

for some (sufficiently large) $\beta \geq 1$. Assume also that $\alpha_1(z)$ is analytic for $|z| \geq \beta$. Then the solution of (7.1) satisfies the a-priori estimate in the discrete l^2-norm:

$$
\|v^{(n+1)}\|_2 \leq \beta^n \left(\|v^{(0)}\|_2 + \sqrt{\frac{(\beta-1)\rho}{2}} \|\Delta^- v^{(0)}\|_2 \right), \quad n \in \mathbb{N}_0 \tag{7.4}
$$

where $\Delta^- u_j^{(n)} = u_j^{(n)} - u_{j-1}^{(n)}$.

For the case of the exact discrete DTBC the assumption of the Theorem 7.1 can easily be checked: tis property of $\alpha_1(z)$ can be shown for $\beta = 1$ in the following way consider the unit circle $z = e^{i\varphi}$,$0 \leq \varphi \leq 2\pi$. Then, we have

$$
y(z) = \frac{1}{\rho}\left(\frac{z-1}{z+1}\right) = \frac{1}{\rho}\left(i \tan(\varphi/2)\right), \quad 0 \leq \varphi \leq 2\pi
$$

Therefore we obtain the requested property

$$\Re\left\{\alpha_1(z)\right\} = 1 + \Re\left\{\sqrt[+]{y(z)(2+y(z))}\right\} \geq 1$$

for $z = e^{i\varphi}, \quad 0 \leq \varphi \leq 2\pi$, i.e. for the exact dicrete TBC we have the estimate

$$\|v^{(n)}\|_2 \leq \|v^{(0)}\|_2, \quad n = 0, 1, \ldots$$

7.1 Proof of the Growth Condition

In this section we present the proof, based entirely on [1] of the theorem presented in the last section. The proof is based on a discrete energy estimate for the new variable

$$u_j^{(n)} = u_j^{(n)} \beta^{-n}$$

which fulfills

$$\beta^{-n}\left(v_j^{(n+1)} \pm v_j^{(n)}\right) = u_j^{(n+1)} \pm u_j^{(n)} + (\beta - 1)\, u_j^{(n+1)},$$

and therefore satisfies

$$\begin{aligned} u_j^{(n+1)} - u_j^{(n)} &= \rho\left(u_{j+1}^{(n+1/2)} - 2u_j^{(n+1/2)} + u_{j-1}^{(n+1/2)}\right) \\ &\quad + (\beta - 1)\left[\frac{\rho}{2}\left(u_{j+1}^{(n+1)} - 2u_j^{(n+1)} + u_{j-1}^{(n+1)}\right) - u_j^{(n+1)}\right] \quad j \geq 1 \end{aligned} \tag{7.5a}$$

$$u_j^{(0)} = v_j^{(0)}, \quad j = 0, 1, 2, \ldots, \tag{7.5b}$$

$$\Delta^+ \widehat{u}_0 = (\widehat{\ell}(\beta z) - 1)\widehat{u}_0. \tag{7.5c}$$

The transformed discrete TBC (7.5c) can be written in physical space as

$$\Delta^+ u_0^{(n)} = \frac{\tilde{\ell}^{(n)}}{\beta^n} * u_0^{(n)} = \sum_{m=0}^{n}\left(\tilde{\ell}^{(n-m)}\beta^{m-n}\right) u_0^{(m)}, .$$

where $\tilde{\ell}^{(n)} := \ell^{(n)} - \delta_0^n$ and $\Delta^+ u_0^{(n)} = u_1^{(n)} - u_0^{(n)}$ denotes the usual forward difference. First we multiply (7.5a) by $u_j^{(n)}/\beta$ and then by $u_j^{(n+1)}$:

$$\begin{aligned} u_j^{(n)}\left(u_j^{(n+1)} - u_j^{(n)}\right) &= \rho u_j^{(n)}\left(u_{j+1}^{(n+1/2)} - 2u_j^{(n+1/2)} + u_{j-1}^{(n+1/2)}\right) \\ &\quad - \beta^{-1}(\beta - 1)\left[\frac{\rho}{2}\left(u_{j+1}^{(n)} - 2u_j^{(n)} + u_{j-1}^{(n)}\right) + u_j^{(n)}\right], \end{aligned} \tag{7.6a}$$

$$\begin{aligned} u_j^{(n+1)}\left(u_j^{(n+1)} - u_j^{(n)}\right) &= \rho u_j^{(n+1)}\left(u_{j+1}^{(n+1/2)} - 2u_j^{(n+1/2)} + u_{j-1}^{(n+1/2)}\right) \\ &\quad + (\beta - 1)\left[\frac{\rho}{2}\left(u_{j+1}^{(n+1)} - 2u_j^{(n+1)} + u_{j-1}^{(n+1)}\right) - u_j^{(n+1)}\right] \end{aligned} \tag{7.6b}$$

Note that we use the equation (7.5a) to modify the last term of (7.6a). Next we add (7.6a) and (7.6b), sum it up for the range $j = 1, 2, 3 \ldots$ and obtain using the summation by parts rule:

$$
\begin{aligned}
\sum_{j=1}^{\infty} \left((u_j^{(n+1)})^2 - (u_j^{(n)})^2 \right) = &-2\rho \sum_{j=1}^{\infty} \left(\Delta^- u_j^{(n+1/2)} \right)^2 \\
&- (\beta - 1) \frac{\rho}{2} \sum_{j=1}^{\infty} \left(\Delta^- u_j^{(n+1)} \right)^2 \\
&- \frac{(\beta - 1)}{\beta} \frac{\rho}{2} \sum_{j=1}^{\infty} \left(\Delta^- u_j^{(n)} \right)^2 \\
&- (\beta - 1) \sum_{j=1}^{\infty} \left(u_j^{(n+1)} \right)^2 - \frac{(\beta - 1)}{\beta} \sum_{j=1}^{\infty} \left(u_j^{(n)} \right)^2 \\
&- \frac{\rho}{2\beta} \left(u_0^{(n)} + \beta u_0^{(n+1)} \right) \Delta^+ \left(u_0^{(n)} + \beta u_0^{(n+1)} \right), \qquad (7.7)
\end{aligned}
$$

where $\Delta^- u_0^{(n)} = u_j^{(n)} - u_{j-1}^{(n)}$ denotes the backward difference. Now summing (7.7) from time level $n = 0$ to $n = N$ yields:

$$
\begin{aligned}
\beta \| u^{(N+1)} \|_2^2 = &\beta^{-1} \| u^{(0)} \|_2^2 - \frac{\beta^2 - 1}{\beta} \sum_{n=1}^{N} \| u^{(0)} \|_2^2 \\
&- 2\rho \sum_{n=0}^{N} \| \Delta^- u^{(n+1/2)} \|_2^2 - \frac{(\beta - 1)^2}{\beta} \frac{\rho}{2} \sum_{n=1}^{N} \| \Delta^- u^{(n)} \|_2^2 \qquad (7.8) \\
&+ \frac{\beta - 1}{\beta} \frac{\rho}{2} \| \Delta^- u^{(0)} \|_2^2 - (\beta - 1) \frac{\rho}{2} \| \Delta^- u^{(N+1)} \|_2^2 \\
&- \frac{\rho}{2\beta} \sum_{n=0}^{N} \left(u_0^{(n)} + \beta u_0^{(n+1)} \right) \Delta^+ \left(u_0^{(n)} + \beta u_0^{(n+1)} \right)
\end{aligned}
$$

Noting that $\beta \geq 1$ we obtain from (7.8) the following estimate:

$$
\begin{aligned}
\| u^{(N+1)} \|_2^2 \leq &\beta^{-2} \| u^{(0)} \|_2^2 + \frac{\beta - 1}{\beta^2} \frac{\rho}{2} \| \Delta^- u^{(0)} \|_2^2 \qquad (7.9) \\
&- \frac{\rho}{2\beta^2} \sum_{n=0}^{N} \left(u_0^{(n)} + \beta u_0^{(n)} \right) \Delta^+ \left(u_0^{(n+1)} + \beta u_0^{(n+1)} \right)
\end{aligned}
$$

It remains to show that the boundary-memory-term in (7.9) is of *positif type.* To this end we define (for N fixed) the two sequences,

$$g^{(n)} := \begin{cases} u_0^{(n)} + \beta u_0^{(n+1)}, & n = 0, \ldots, N, \\ 0, & n > N, \end{cases}$$

$$f^{(n)} := \frac{\tilde{\ell}^{(n)}}{\beta^n} * g^{(n)} = \sum_{m=0}^{n} \frac{\tilde{\ell}^{(n-m)}}{\beta^{n-m}} g^{(m)}, \quad n \in \mathbb{N}_0$$

i.e. $\sum_{n=0}^{N} f^{(n)} g^{(n)} \geq 0$ is to show, the Z-transform $\mathcal{Z}\left\{f^{(n)}\right\} = \widehat{f}(z)$ is analytic for $|z| > 0$. The Z-transform $\mathcal{Z}\left\{f^{(n)}\right\}$ then satisfies $\widehat{f}(z) = \left(\widehat{\ell}(\beta z) - 1\right)\widehat{g}(z)$ and is analytic for $|z| \geq 1$. Using Plancherel's Theorem for Z-transforms we have

$$\begin{aligned} \sum_{n=0}^{N} f^{(n)} g^{(n)} &= \frac{1}{2\pi} \int_0^{2\pi} \widehat{f}(e^{i\varphi}) \overline{\widehat{g}(e^{i\varphi})} d\varphi \\ &= \frac{1}{\pi} \int_0^{\pi} \mid \widehat{g}(e^{i\varphi}) \mid^2 \left(\Re \left\{ \widehat{\ell}(\beta e^{i\varphi} \right\} - 1 \right) d\varphi \end{aligned} \tag{7.10}$$

where we have used the fact that $\widehat{g}(\overline{z}) = \overline{\widehat{g}(z)}$, $\widehat{f}(\overline{z}) = \overline{\widehat{f}(z)}$ since $f_n, g_n \in \mathbb{R}$. Using (7.10) for the boundary term in (7.9) now gives:

$$\begin{aligned} \|u^{(N+1)}\|_2^2 \leq \beta^{-2} \|u^{(0)}\|_2^2 &+ \frac{\beta - 1}{\beta^2} \frac{\rho}{2} \|\Delta^- u^{(0)}\|_2^2 \\ &- \frac{\rho}{2\pi\beta^2} \int_0^{\pi} \mid \left(1 + \beta e^{i\varphi}\right) \widehat{u}_0(e^{i\varphi}) \mid^2 \left(\Re \left\{ \widehat{\ell}(\beta e^{i\varphi}) \right\} - 1 \right) d\varphi \end{aligned}$$

Our assumption on $\widehat{\ell}$ therefore implies

$$\|u^{(N+1)}\|_2^2 \leq \beta^{-1} \|u^{(0)}\|_2 + \frac{\sqrt{\beta - 1}}{\beta} \sqrt{\frac{\rho}{2}} \|\Delta^- u^{(0)}\|_2, \quad \forall N \geq 0,$$

and the result of the theorem follows.

8 Numerical examples

In this section we consider two examples of an American call option from [1]. We will compare the numerical results from using the DTBC (4.10) with the exact coefficients $s^{(n)}$ to the solution using the DTBC with approximated coefficients $\tilde{s}^{(n)}$ and the discretized TBC (4.3) or (4.5). Since the DTBC (4.10) with the exact coefficients $s^{(n)}$ yields the exact numerical solution to the discrete problem (4.1), we will take this solution as a reference solution v_{ref}.

Example 8.1

We consider an American call with an expiry of $T = 0.5$ years and a dividend yield $D_0 = 0.03$. The risk-free interest rate $r = 0.03$, the volatility is $\sigma = 40\%$ and the exercise price is $E = \$100$. We choose $\Delta t = 0.001$ and $\Delta x = 0.01$ with artificial boundary conditions at $a = -1$ which corresponds to an asset price $S = Ee^a \approx 36.79$. In Figure 3, we can see the option values $V(S, 0)$ calculated with the exact DTBC (4.10). The upper bound of the free boundary $x_f(\tau)$ was calculated by (6.9) as $x_f^* = 1.49$ and so we will use $x_{max} = 1.5$. However the largest value of $x_f(\tau)$ is much smaller; it is about 0.6. The time evolution of the free-boundary $x_f(\tau)$ obtained with the exact DTBC (4.10) is plotted in Figure 4.

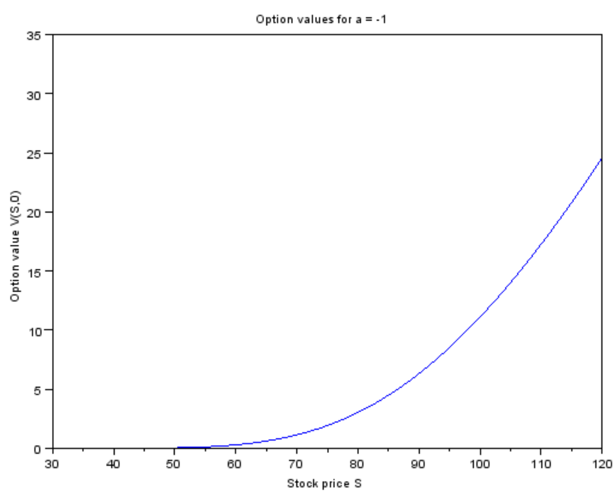


Figure 3: Option values V at time $t = 0$ (i.e. at $\tau = \widetilde{T}$) for Example 8.1.

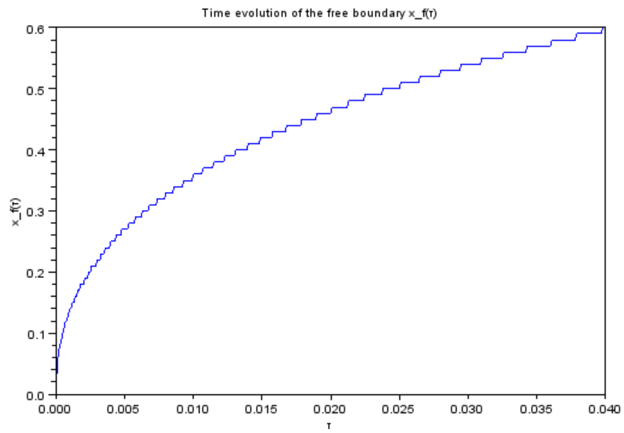


Figure 4: Time evolution of the free boundary $x_f(\tau)$ for Example 8.1.

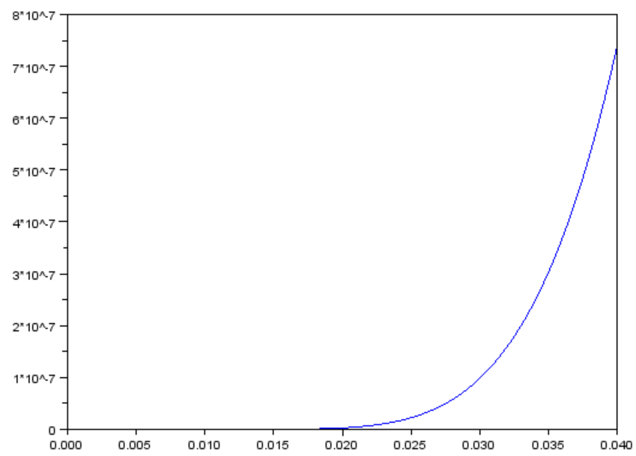


Figure 5: Error $\|v^{(n)} - v_{ref}^{(n)}\|_2^2$ obtained with the discretized TBC (4.5) for Example 8.1.

Finally, we want to compare the errors when using different artificial boundary conditions. In order to make the error more apparent, we reduce the computational domain using $a = -0.5$. We plot in Figures 5 and 6, the errors $\|v^{(}(n)) - v^{(n)}\|_2^2$ where $v^{(n)}$ is the solution obtained with the artificial TBC (4.5) and with the DTBC with approximated coefficients $\tilde{s}^{(n)}$ for $L = 10$, respectively (the norm considered is the one defined in (7.2)). We don't plot the error with TBC (4.3) since we obtained fairly the same error as with TBC (4.5). As we can see, the discretized TBC (4.5) induced a smaller error than the approximated DTBC with $L = 10$. This was somehow expected since, as we say in chapter 5, the approximated coefficients $\tilde{s}^{(n)}$ aren't a good approximation for $s^{(n)}$.

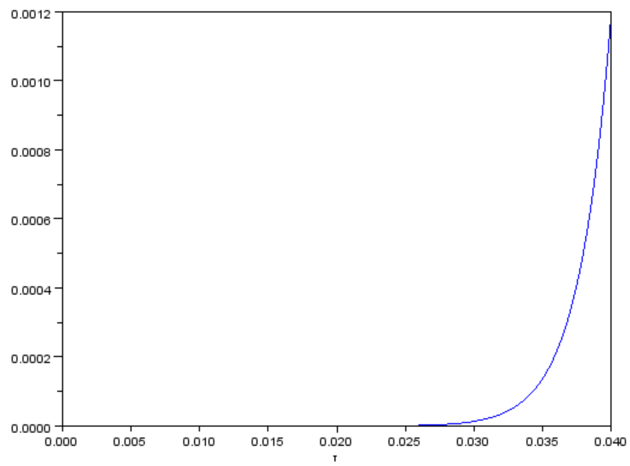


Figure 6: Error $\|v^{(n)} - v_{ref}^{(n)}\|_2^2$ obtained with the approximated DTBC (4.10) for $L = 10$ for Example 8.1.

Comparing our results with the ones from [1], as for the option values and the free boundary we obtained the same results. As for the errors, we obtained a different behavior for the error with (4.5): we have concave function, instead of a convex function.

Example 8.2

We consider an American call with an expiry of $T = 3$ years and a dividend yield $D_0 = 0.07$. The risk-free interest rate $r = 0.03$, the volatility is $\sigma = 40\%$ and the exercise price is $E = \$100$. We choose $\Delta t = 0.006$ and $\Delta x = 0.025$ with artificial boundary conditions at $a = -1$ which corresponds to an asset price $S = Ee^a \approx 36.79$. In Figure 7, we can see the option values $V(S, 0)$ calculated with the exact DTBC (4.10). The upper bound of the free boundary $x_f(\tau)$ was calculated by (6.9) as $x_f^* = 0.8722$ and so we will use $x_{max} = 0.88$. In this case we got a good estimate since the largest value of $x_f(\tau)$ is 0.675. The time evolution of the free-boundary $x_f(\tau)$ obtained with the exact DTBC (4.10) is plotted in Figure 8.

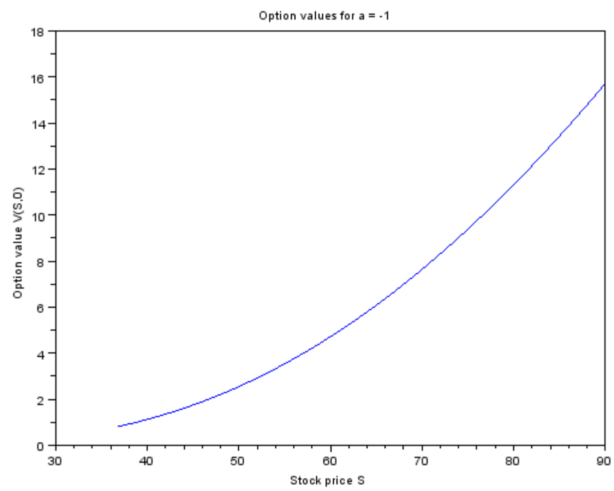


Figure 7: Option values V at time $t = 0$ (i.e. at $\tau = \widetilde{T}$) for Example 8.2.

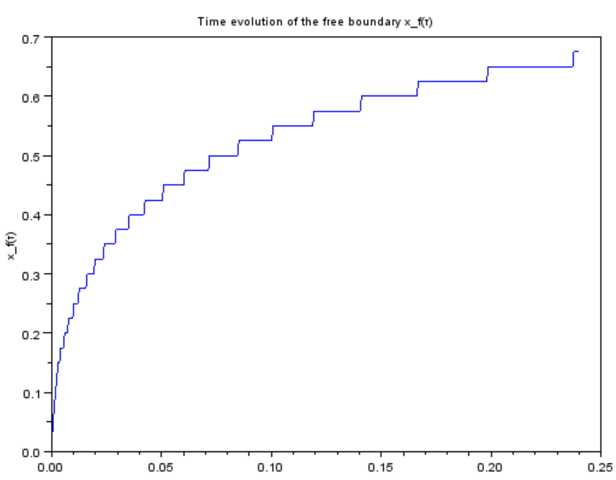


Figure 8: Time evolution of the free boundary $x_f(\tau)$ for Example 8.2.

Finally, we plot the errors in Figure 9. As in the previous example the discretized TBC (4.5) yielded more accurate results and once again (4.5) and (4.3) gave fairly the same erros. As for comparisons with [1], we remark that in this example our free boundary is smaller.

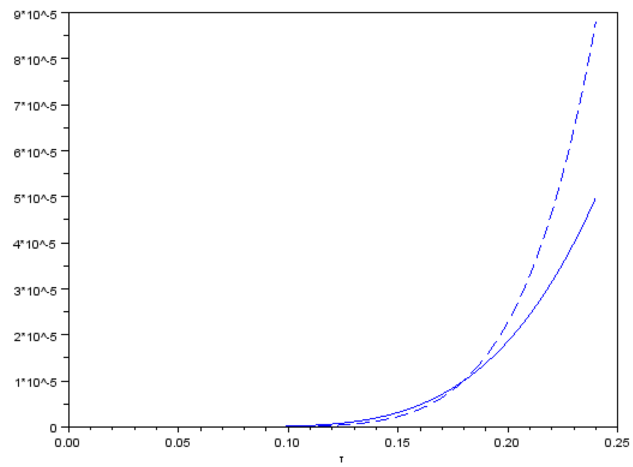


Figure 9: Errors $\|v^{(n)} - v_{ref}^{(n)}\|_2^2$ obtained with the discretized TBC (4.5) (solid line) and the approximated DTBC (4.10) for $L = 10$ (dashed line) for Example 8.2.

9 Conclusion

In this report we presented an alternative to the discretized analytic TBC by deriving an exact discrete boundary condition for the Cranc-Nicolson scheme for solving the Black-Scholes equation for the pricing of American options and which could be extended to other option as forward and future contracts.

In both numerical examples, apart from the approximated DTBC (4.10), all the considered artificial boundary conditions yielded satisfactory results. We didn't however obtained the same conclusions as in [1]for the accuracy of the approximated DTBC with the $\tilde{s}^{(n)}$ coefficients. We couldn't compute them for $L = 20$, and for $L = 10$ we got more accuracy with the TBC (4.5) and (4.3). Also, we couldn't distinguish any difference between the TBC (4.5) and (4.3) as opposed to [1].

Finally, we would like to point out a possible improvement to the treatment of the free boundary that we didn't have time to implement due to lack of time. As stated in §6, for each time step τ_n, we find an interval such that $x_f(\tau_n) \in [x_J, x_{J+1}]$ and make the naive and practical approximation $x_f(\tau_n) \approx x_{J+1}$. However, we could use the values $v_J^{(n)}$ and $v_{J+1}^{(n)}$ to derive by interpolation the point x^* such that $v(\tau_n, x^*) \approx g(x^*, \tau_n)$, which we would expect to be a more accurate approximation of the free boundary. However, since this point is not a point of our grid, the finite difference method will induce a bigger error which as to be taken into account. Another problem that we would have to overcome is that in fact there is no obvious way of determining if the new free boundary is in fact more accurate or not.

A Padé Approximation

The Padé Approximation can be seen as a generalization of the Taylor approximation. In the last one, given a function f such that

$$f(x) = \sum_{i=0}^{\infty} a_i x^i,$$

we approximate it by a polynomial P_n of order n, where

$$P_n(x) = \sum_{i=0}^{n} a_i x^i.$$

Hence,

$$f(x) - P_n(x) = \sum_{i=n+1}^{\infty} a_i x^i = \mathcal{O}\left(x^{n+1}\right).$$

Moreover, we have that

$$\begin{aligned} f(0) &= P_n(0) \\ f'(0) &= P_n'(0) \\ &\dots \\ f^{(n)}(0) &= P_n^{(n)}(0). \end{aligned}$$

Definition A.1 *Let f be an analytic function at 0 such that $f(0) \neq 0$ and n, m two positive integers. The Padé approximation of order $[p, q]$ for the function f is the ration function*

$$R(z) = \frac{P_n(z)}{Q_m(z)} = \frac{p_0 + p_1 z + \dots p_n z^n}{q_0 + q_1 z + \dots q_m z^m}$$

where P and Q are polynomials of order n and m, respectively, and such that agrees with f to the highest possible order, which amounts to

$$\begin{aligned} f(0) &= R(0) \\ f'(0) &= R'(0) \\ &\dots \\ f^{(n+m)}(0) &= R^{(n+m)}(0). \end{aligned}$$

Equivalently,

$$f(z) - R(z) = \mathcal{O}\left(z^{n+m+1}\right).$$

Remark A.2 *The hypothesis of the function being nonzero at the origin is not restrictive. If g is an analytic function at 0, there is necessarily k such that $g(t) = t^k f(z)$ where f satisfies the hypotheses of our definition.*

It is important to notice that the Padé approximation might not always exist. Consider for instance the function $f(z) = 1 + z^2$. If we search for Padé approximation of order $[1,1]$, we must solve the following system

$$\begin{cases} f(0) = R(0) \\ f'(0) = R'(0) \\ f''(0) = R''(0) \end{cases}$$

which is impossible. Therefore, there isn't a Padé approximation of order $[1,1]$ for the function f.

B Parabolic Equations

Consider the operator

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

in a domain $D \subseteq \mathbb{R}^{n+1}$. Let $(x,t) = (x_1, \ldots, x_n, t)$ be a point in $\mathbb{R}^{n+1}$. Let the boundary ∂D of D consist of the closure of a domain B_0 lying on $t = 0$, a domain B_T $(T > 0)$ and a (not necessarily connected) manifold S lying in the strip $0 < t \leq T$. For any set G, we denote by $\bar{G}$ the closure of G.

Definition B.1 *We say that L is parabolic if for every $(x,t) \in D$ and for any real vector $\xi \neq 0$,*

$$\sum_{i,j=1}^{n} a_{ij}(x,t)\xi_i\xi_j > 0.$$

We will assume that

- L is parabolic in D;
- the coefficients of L are continuous functions in D;
- $c(x,t) \leq 0$ in D.

We can now state the weak maximum principle.

Theorem B.2 *If D is bounded and $Lu = 0$ in D, then*

$$\max_{D} |u| \leq \max_{\bar{B}+S} |u|.$$

A proof of this theorem can be seen in [4].

References

[1] M. Ehrhardt and R.E. Mickens, A fast, stable and accurate numerical method for the Black-Scholes equation of American options, *Int. J. Theoret. Appl. Finance* Vol. 11, Issue 5, (2008), 471-501.

[2] H. Han and X. Wu, A fast numerical method for the BlackŨScholes equation of American options, *SIAM J. Numer. Anal.* **41** (2003) 2081Ũ2095.

[3] R. C. Merton, Theory of rational option pricing, *Bell J. Econ. Manag. Sci.* **4** (1973) 141Ũ183.

[4] A. Friedman, *Partial Differential Equations of Parabolic Type*, Robert E. Krieger Publishing, Huntington, NY, 1983.