New features on finite-difference methods for option pricing

S. P. Oliveira and P. Pulino

This paper adds new features to the implementation of finite-difference and finite-element methods for option models based on the Black-Scholes equation. Stability analysis for explicit methods without transformations and non-uniform meshes for the asset price are focused. The American option pricing model was formulated as a variational inequality whose approximation leads in general to a non-symmetric linear complementarity problem. Numerical experiments reveal significant improvements on European and American "vanilla" options.

1 Introduction

Finite-difference methods for option pricing were extensively studied in the past two decades (Brennan and Schwarz (1978), Courtadon (1982), Geske and Shastri (1985), Hull and White (1990)). Starting from the approximation of the Black-Scholes formula (Black and Scholes (1973)) by solving the corresponding boundary-value problem, several modifications of the standard method were designed to solve general problems. Wilmott *et al.* (1993) and Hull (1997) present a complete description of the methodology.

This paper recovers some issues on the finite-difference approach and proposes effective ways to increase the accuracy of such methods as well. We present the effect of the proposed modifications on European call options.

We consider the standard problem of American put options, using a formulation that does not involve the free-boundary (Wilmott *et al.* (1993)). The same formulation can be rewriten as a variational inequality, what allows the use of finite-element methods.

The remainder of the paper is organized as follows. Sections 2 and 3 review the Black-Scholes model and the standard finite-difference method. Section 4 proposes an accurate measure of stability for explicit methods. In Section 5, we discuss the discretization of the asset price. This variable is particularly affected by logarithmic transformations. The modified methods are compared with the standard ones and with the Black-Scholes formula in Section 6. Section 7 considers the pricing of American options and introduces a finite-element method based on a variational formulation of the problem (Wilmott *et al.* (1993)). The numerical results for American options are compared with others found in the literature in Section 8. Section 9 summarizes the main results.

2 The Model

The work of Black and Scholes (1973) and Merton (1973) are a landmark in the general option pricing theory. Based on the hypothesis of continuous, frictionless trading; constant interest rate; and the principle of non-arbitrage, they had derived the following formula to the price $c(S,t)$ of European call options based on lognormal-based assets without dividend payments:

$$c(S,t) = SN(d_1) - Ee^{-r(T-t)}N(d_2) \quad , \tag{2.1}$$

$$d1 = \frac{\log(S/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}} \tag{2.2}$$

$$d2 = \frac{\log(S/E) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}} \tag{2.3}$$

where E is the exercise price, r is the interest rate, σ is the volatility, T is the expiry and $N(\cdot)$ is the standard normal distribution. The variables S and t are the asset price and time, respectively. This formula was derived from the Black-Scholes differential equation

$$\frac{\partial c}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 c}{\partial S^2} + rS\frac{\partial c}{\partial S} - rc = 0 \tag{2.4}$$

with suitable initial and boundary conditions. It was shown later that (2.1) can be derived from other approaches (Duffie (1988)).

The Black-Scholes equation is able to price several other derivatives based on the lognormal process. (Geske and Shastri (1985), Wilmott *et al.* (1993)). The formula to the price $p(x,t)$ of a put option was also derived from (2.4), and both formulas are consistent with the *put-call parity:*

$$p(S,t) = c(S,t) - S + Ee^{-r(T-t)} \tag{2.5}$$

Under the same hypothesis, the price $C(S,t)$ of an American call must be the same as it European counterpart (Merton (1973)), while the price $P(S,t)$ of an American put option is the solution of the following problem (Wilmott *et al.* (1993)):

$$(\frac{\partial P}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} - rS\frac{\partial P}{\partial S} + rP) \geq 0 \quad (2.6)$$

$$(P - (E-S)^+) \geq 0 \quad (2.7)$$

$$(\frac{\partial P}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} - rS\frac{\partial P}{\partial S} + rP)(P - (E-S)^+) = 0 \quad (2.8)$$

$$u(S,0) = (E-S)^+ \quad (2.9)$$

$$u(0,t) = E \quad (2.10)$$

$$\lim_{S\longrightarrow\infty} u(S,t) = 0 \quad (2.11)$$

Here $(E-S)^+$ denotes $\max\{E-S,0\}$. This is a free-boundary problem whose free boundary represents the *optimal exercise price* at a time t. The free boundary does not appear on the formulation above (what simplifies the analysis) but can be calculated *a posteriori*. In Appendix A we prove that this boundary is unique.

3 Finite-Difference Method

Let us consider the European call option problem with a change in the time variable (Wilmott *et al.* (1993)):

$$\frac{\partial c}{\partial t} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 c}{\partial S^2} + rS\frac{\partial c}{\partial S} - rc \ , \quad S > 0 \ , \quad t > 0 \quad (3.1)$$

$$c(S,0) = (S-E)^+ \quad (3.2)$$

$$c(S,t) \leq S \quad (3.3)$$

Let $\Omega_h = \{0 = S_0 < S_1 < \dots S_N = L\} \times \{0 = t_0 < t_1 < \dots t_M = T\}$ be an approximation of $\Omega = [0,L] \times\ [0,T]$, where $S_i = ih$ and $t_n = nk$. We denote the value of a function f at the point $(S_i,t_n) \in \Omega_h$ by f_i^n and the value of a function g at S_i by g_i (Geske and Shastri (1985), Hull (1997)).

The finite-difference methods approximate the solution $c : \Omega \to I\!R$ of the problem (3.1)-(3.2) at the set Ω_h by calculating values U_i^n such that

$$U_i^n \approx c(S_i,t_n) = c_i^n \ \ \forall\ (S_i,t_n) \in \Omega_h \quad (3.4)$$

These methods are based on local approximations of the partial derivatives at $(S_i, t_n) \in \Omega_h$, generating a set of difference equations whose solutions are the numerical solutions U_i^n.The simplest finite-difference method is the *explicit* method with centered differences at the variable S (Brennan and Schwarz (1978), Hull (1997), Wilmott *et al.* (1993))

$$\frac{U_i^{n+1} - U_i^n}{k} = \frac{1}{2}\sigma^2 S_i^2 \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{h^2} + rS_i \frac{U_{i+1}^n - U_{i-1}^n}{2h} - rU_i^n \qquad (3.5)$$

Reordering (3.5) and observing that $S_i = ih$, we have

$$U_i^{n+1} = c_{-1}U_{i-1}^n + c_0 U_i^n + c_1 U_{i+1}^n \ , \qquad (3.6)$$

where the coefficients $c_j = c_j(i)$ are given by

$$c_{-1} = \frac{1}{2}k\left(\sigma^2 i^2 - ri\right) \qquad c_o = 1 - k\left(\sigma^2 i^2 + r\right)$$

$$c_1 = \frac{1}{2}k\left(\sigma^2 i^2 + ri\right) \qquad (3.7)$$

The initial and boundary conditions are approximated in the following way:

$$U_i^0 = (ih - E)^+ \ , \ \ 0 \le i \le N \qquad (3.8)$$

$$U_0^n = \phi_n \text{ and } U_L^n = \psi_n \ , \ \ 0 \le n \le M \ , \qquad (3.9)$$

where ϕ_n and ψ_n are approximations to $c(0, t_n)$ and $c(L, t_n)$, respectively. We calculate the solution of (3.6) at each time level n, $1 \le n \le M$, using the (previous) solution at the time level $n - 1$.

Another standard scheme is given by the *Crank-Nicolson* method with parameter $\alpha = 1/2$:

$$b_{-1}U_{i-1}^{n+1} + b_0 U_i^{n+1} + b_1 U_{i+1}^{n+1} = c_{-1}U_{i-1}^n + c_0 U_i^n + c_1 U_{i+1}^n \ , \qquad (3.10)$$

where the coefficients b_j and c_j are given by:

$$c_{-1} = \frac{1}{4}k\left(\sigma^2 i^2 - ri\right) \qquad c_o = 1 - \frac{1}{2}k\left(\sigma^2 i^2 + r\right)$$

$$c_1 = \frac{1}{4}k\left(\sigma^2 i^2 + ri\right) \qquad (3.11)$$

$$b_{-1} = -c_{-1} \ \text{ and } \ b_o = 2 - c_o \ \text{ and } \ b_1 = -c_1$$

The solution at the interior points is calculated by solving a tridiagonal system derived from (3.10). The Crank-Nicolson scheme usually converges when h and k tend to zero and, e.g., $h = k$, while the convergence in the explicit method requires $k \ll h$; however, the computational cost on (3.10) is higher than in (3.6), as we need to solve linear systems at each time step. It is worth to have a precise measure of the relation between h and k on the explicit method, as discussed in the next section.

Let us return to the boundary conditions. The problem (3.1)-(3.2) doesn't need boundary conditions to be analytically solved. Moreover, there is few practical interest to find the option price when the asset price has extreme values; however, finite-difference methods do need boundary conditions.

The boundary condition at $S = 0$ can be easily derived (Black and Scholes (1973), Merton (1973))

$$U_0^n = c(0, t_n) = 0 \tag{3.12}$$

The value of S can grow without limits. We have that $S \to \infty$ implies $T \to \infty$, that is, it only makes sense to find the limit of $c(S,t)$ when $S \to \infty$ in perpetual options (Oliveira (1998)). Let us denote the perpetual European call and put options by $c_\infty(S,t)$ and $p_\infty(S,t)$. We have (Merton (1973))

$$c_\infty(S,t) = S \text{ and } p_\infty(S,t) = 0 \tag{3.13}$$

However, $c(S,t)$ approaches $c_\infty(S,t)$ when $S \gg 0$, that is,

$$\lim_{S \longrightarrow \infty} c(S,t) = \lim_{S \longrightarrow \infty} c_\infty(S,t) = S \tag{3.14}$$

Observe that (3.14) is consistent with (2.1). The next step is to approximate the "boundary condition" (3.14) by the following:

$$U_N^n = c_\infty(L, t_n) = S_N \quad , \tag{3.15}$$

with L sufficiently large. The choice of L must take into account the accuracy and the computational cost, and usually values within $[2E, 3E]$ produce good results. From the parity formula (2.5) and (3.13) we have the following alternative to (3.15) :

$$U_N^n = c_\infty(L, t_n) = S_N - Ee^{-r(T-t_n)} \tag{3.16}$$

In general, (3.16) is more accurate than (3.15).

4 Stability Analysis

It is well known that the explicit scheme (3.6) is conditionally stable, but we haven't seen any accurate estimate, as the non-constant coefficients of the Black-Scholes equation complicate the discrete Fourier transform approach and the analysis of eigenvalues. The standard approach (Courtadon (1982), Geske and Shastri (1985)) is to transform the Black-Scholes equation in another one with constant coefficients.

Based on the observations in Brennan and Schwarz (1978) and Hull and White (1990) we require the coefficients (3.7) of the explicit scheme to be positive and obtain a stability estimate for (3.6). The numerical experiments below indicate that the relation between the increments $h = \Delta S$ and $k = \Delta t$ derived from that estimate are close to be a necessary and sufficient stability condition. See Thomas (1995) for a background in stability analysis.

A explicit finite-difference scheme on the form

$$U(x, t+k) = \sum_{j=n_1}^{n_2} c_j U(x+jh, t) \tag{4.1}$$

is called positive if the components c_i satisfy the following conditions:

$$\exists h_1 > 0 \; ; \; c_j \geq 0 \quad , \quad -n_1 \leq j \leq n_2 \;\; \forall \, (x,t) \text{ and } h \quad , \quad 0 \leq h \leq h_1 \tag{4.2}$$

Brennan and Schwarz (1978) interpret the coefficients c_j as probabilities of the prices to move up ($j > 0$), down ($j < 0$) or stay the same ($j = 0$). The condition (4.2) within this interpretation simply says that those probabilities cannot be negative.

Let $U^{n+1} = [U_1, U_2, \ldots U_N]$ be a vector representing the approximate solution U at a fixed time level t_{n+1} ($U_1 = U(x_i, t_{n+1})$). We state a definition of stability following Forsythe and Wasow (1978) :

Definition 4.1 : The scheme (4.1) is stable if for any fixed n there exists $M > 0$ independent of h and k such that

$$\|U^{n+1}\| \leq M \quad \forall \, h, k > 0 \tag{4.3}$$

In general, M depends on t and on $\|U^0\|$. The theorem 1.41 on Forsythe and Wasow (1978) states that the scheme (4.1) is stable if it is positive, assuming that the initial and boundary conditions are uniformly bounded.

Back to our original problem, we should then verify when the coefficients c_{-1}, c_0 and c_1 in (3.7) are positive. Observe that c_1 is always positive. Imposing $c_0 > 0$ leads us to the following inequality:

$$\lambda \leq \frac{1 - rk}{\sigma^2 S_i^2} \quad , \quad \lambda = \frac{k}{h^2} \tag{4.4}$$

A sufficient condition for (4.4) is given by

$$\lambda \leq \frac{1 - rk}{\sigma^2 L^2} \tag{4.5}$$

As in general $r,\ k \ll 1$, we approximate (4.5) by the following inequality:

$$\lambda \leq \frac{1}{\sigma^2 L^2} \tag{4.6}$$

Analogously, the condition $c_{-1} \geq 0$ implies

$$\frac{\sigma^2 S_i}{h} \geq r \tag{4.7}$$

We can replace (4.7) by a stronger inequality:

$$\frac{\sigma^2}{r} \geq 1 \tag{4.8}$$

The condition (4.8) is too severe, as it imposes restrictions on equation parameters. As the instability due to the violation of (4.8) is numerically observed only when $r \gg \sigma$, we dismiss that condition. For the cases where that instability is observed, the explicit method diverges and we need specific numerical schemes to handle it (Zvan *et al.* (1998)).

The inequality (4.6) is thus our stability estimate. As $h * N = L$ and $k * M = T$, (4.6) provides an estimate for the minimum number of time steps given the number of price steps:

$$M \geq (\sigma N)^2 T \tag{4.9}$$

We analyze it numerically by calculating the spectral radius of the matrix A associated to the explicit scheme (Geske and Shastri (1985)), that is,

$$U^{n+1} = AU^n \tag{4.10}$$

The following plots describe the sensitivity of the spectral radius $\rho(A)$ with respect to the equation parameters. We use $T = 1$, $L = 2 * E$, $N = 50$ and fix λ to be 50% lower than λ^*, equal to λ^* and 25% higher than λ^*, where λ^* is given by (4.6). The value of h follows from N and k is calculated from λ. The starting values of the parameters E, σ and r are $E = 10$, $\sigma = 0.02$ and $r = 0.05$, respectively.

We observed stability ($\rho(A) \leq 1$) for $\lambda \leq \lambda^*$, given that r is not too large, unlike the cases when $\lambda > \lambda^*$. The instability due to the violation of (4.8) occurs when $r > 2$, outside the range of practical importance.

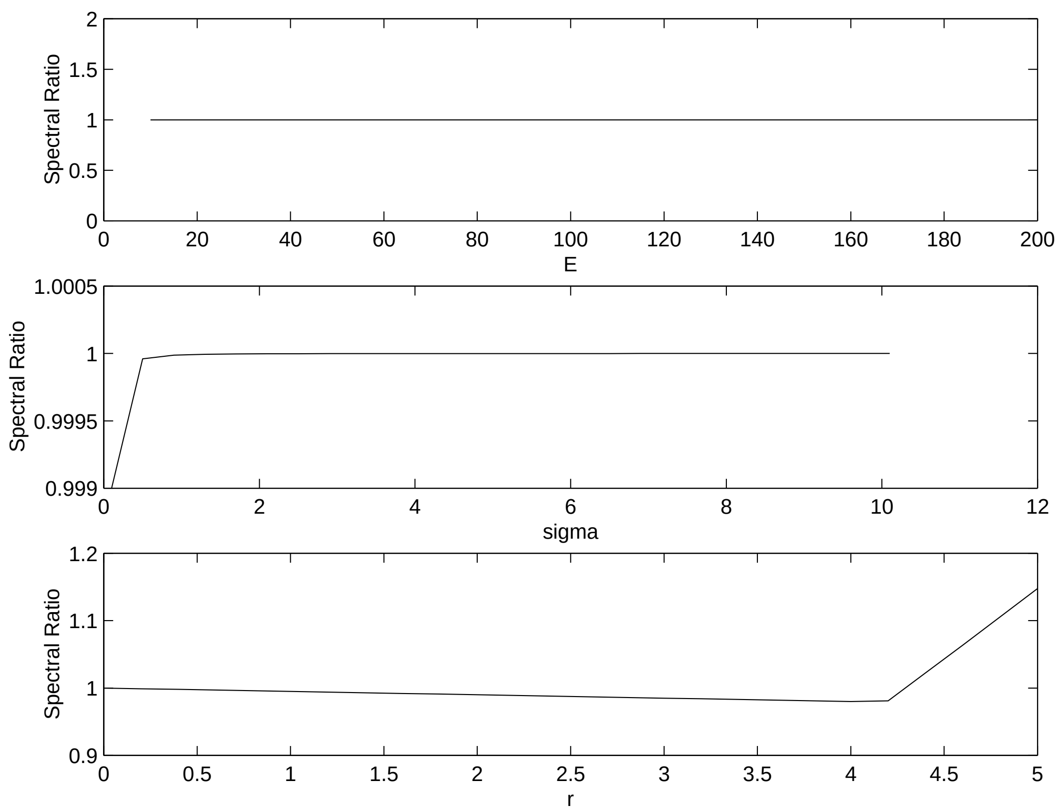

Figure 1: Sensitivity of $\rho(A)$ when λ is 50% below the estimate (4.6).

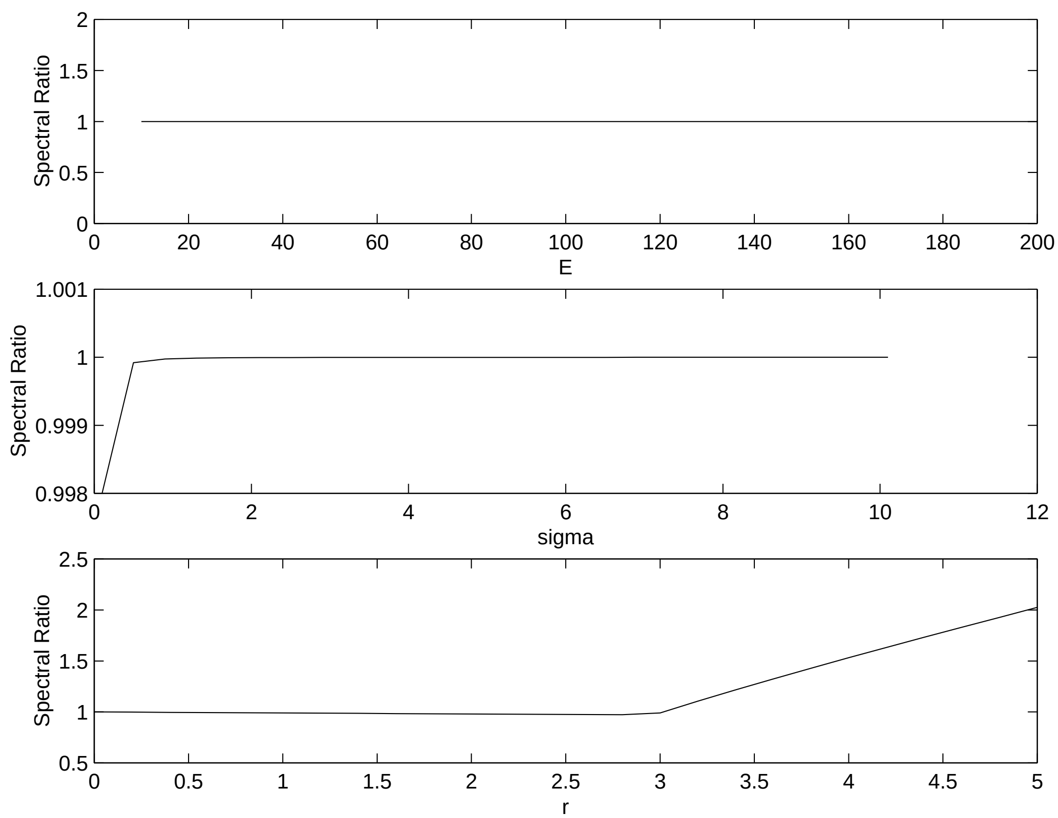

Figure 2: Sensitivity of $\rho(A)$ when λ is equal to the estimate (4.6).

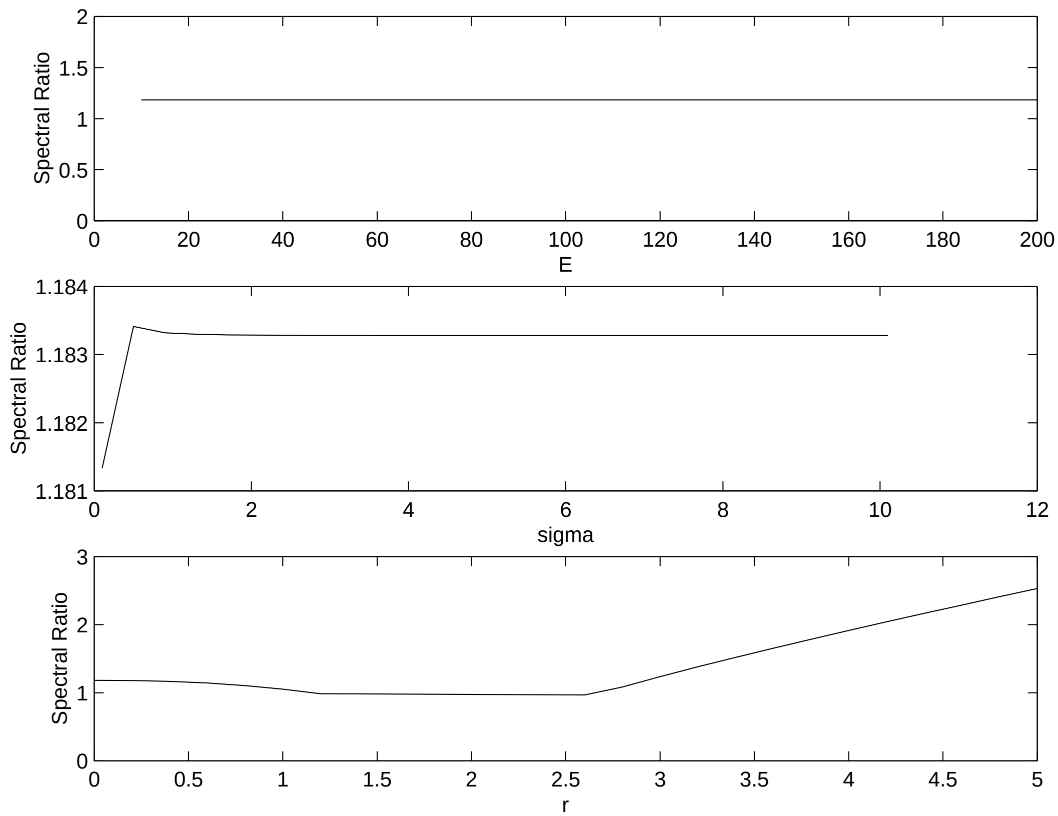

Figure 3: Sensitivity of $\rho(A)$ when λ is 25% above the estimate (4.6).

5 Non-Uniform Grids

Let us consider a more general approximation of $\Omega = [0, L] \times \ [0, T]$:

$$\Omega_h = \{0 = S_0 < S_1 < \dots S_N = L\} \times \{0 = t_0 < t_1 < \dots t_M = T\} \qquad (5.1)$$

where $t_n = nk$, $0 \leq n < M$ as before. We define $h_i = S_{i+1} - S_i$, $0 \leq i < N$. For the explicit method we have, as in (3.6) ,

$$U_i^{n+1} = c_{-1} U_{i-1}^n + c_0 U_i^n + c_1 U_{i+1}^n \ , \qquad (5.2)$$

Where the coefficients c_j now depend on S_i and are given by

$$\begin{aligned} c_{-1} &= \frac{kS_i}{h_{i-1} - h_i} \left(\frac{\sigma^2 S_i^2}{h_{i-1}} - ri \right) \\ c_0 &= 1 - k \left(\frac{\sigma^2 S_i^2}{h_i h_{i-1}} - r \right) \\ c_1 &= \frac{kS_i}{h_{i-1} - h_i} \left(\frac{\sigma^2 S_i^2}{h_i} + ri \right) \end{aligned} \qquad (5.3)$$

For the Crank-Nicolson method,

$$\bar{b}_{-1} U_{i-1}^{n+1} + \bar{b}_0 U_i^{n+1} + \bar{b}_1 U_{i+1}^{n+1} = \bar{c}_{-1} U_{i-1}^n + \bar{c}_0 U_i^n + \bar{c}_1 U_{i+1}^n \ , \quad 0 \leq i \leq N \quad (5.4)$$

where the coefficients $\bar{b}_j$, $\bar{c}_j$ can be written with respect to c_j on (5.3):

$$\bar{b}_{-1} = -\frac{1}{2} c_{-1} \quad , \quad \bar{b}_o = 1 + \frac{1}{2}(1 - c_o) \quad , \quad \bar{b}_1 = -\frac{1}{2} c_1 \qquad (5.5)$$

$$\bar{c}_{-1} = -\bar{b}_{-1} \quad , \quad \bar{b}_o = 2 - \bar{b}_o \quad , \quad \bar{b}_1 = \bar{b}_1 \qquad (5.6)$$

Non-uniform grids are employed to refine some regions of the domain where the approximate solution must be more accurate. In our case, that region is the neighborhood of $S = E$.

There are two reasons for refine the grid around the exercise price E. First, the initial condition (the *payoff*) is not differentiable at $S = E$ for both call and put options, what can jeopardize the solution there. Moreover, the most important option prices stand around the exercise price (Hull (1997)).

We also get grids with non-uniform patterns with logarithmic transformations on the asset variable (Geske and Shastri (1985), 49). The most popular transformation is $x = \ln(S/E)$ (Brennan and Schwarz (1978)), that transforms the Black-Scholes equation in a equation with constant coefficients and takes advantage from the fact that the standard deviation of $\ln(S)$ does not depend on S and t in a short period of time (Hull and White (1990)). Another approach (Geske and Shastri (1985), Wilmott *et al.* (1993)) transforms the Black-Scholes equation in the diffusion equation, a symmetric problem.

However, these transformations extend the asset price domain to $]-\infty, +\infty[$. Moreover, if we apply an uniform grid over $x = \ln(S/E)$, the grid points of the original variable S are strongly concentrated around $S = 0$. This can be easily seen if we plot the resulting solution as follows:

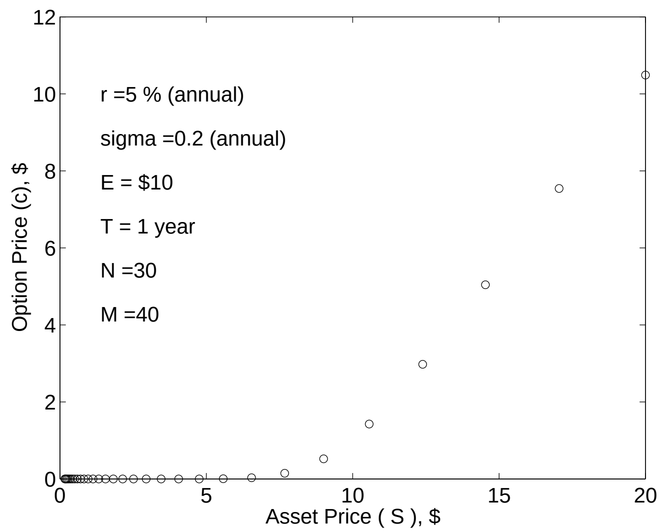

Figure 4: Transformed finite-difference solution without interpolation.

In this case, we would need to reduce the value of h (what implies a higher computational cost) to improve the solution accuracy at $S \approx E$. Further, we get an unnecessary refinement at $S \approx 0$.

Thus, applying the explicit method straight to the Black-Scholes equation leads us to stability problems, while simplifying the equation with logarithmic transformations may be inaccurate around $S = E$. A natural question is that if there is any transformation that at the same time simplifies the equation and gives a grid refined around $S = E$ in the original variable when we use a uniform mesh over the transformed one.

Focusing on that idea, we could come up with the following transformation:

$$S = E \pm E(e^y - 1) \tag{5.7}$$

Observe that (5.7) is not injective. We can circumvent that by splitting the asset price domain in $[0, E]$ and $[E, L]$. Following Wilmott *et al.* (1993), we also transform the time variable. The transformations and corresponding equations are as follow:

1. Sub-domain $[0, E]$
Transformations:

$$c = Ev \quad , \quad S = E - E(e^y - 1) = E(2 - e^y) \quad , \quad \tau = \frac{\sigma^2}{2}(T - t) \tag{5.8}$$

Transformed Equation:

$$\frac{\partial v}{\partial \tau} = g(y)^2 \frac{\partial^2 v}{\partial y^2} + g(y)\left(K - g(y)\right)\frac{\partial v}{\partial y} - Kv \tag{5.9}$$

where $g(y) = 1 - 2e^{-y}$ and $K = 2r/\sigma^2$.
Initial condition and boundary condition at $y_N \approx 0$:

$$V_i^0 = 0 \ \text{ and } \ V_N^n = 0 \tag{5.10}$$

2. Sub-domain $[E, L]$
Transformations:

$$c = Ew \quad , \quad S = E + E(e^x - 1) = Ee^x \quad , \quad \tau = \frac{\sigma^2}{2}(T - t) \tag{5.11}$$

Transformed Equation:

$$\frac{\partial w}{\partial \tau} = \frac{\partial^2 w}{\partial x^2} + (K - 1)\frac{\partial w}{\partial x} - Kw \tag{5.12}$$

Initial condition and boundary condition at $x_N = L$:

$$W_i^0 = e^x - 1 \ \text{ and } \ W_N^n = e^x \tag{5.13}$$

In principle we have no conditions at V_0^n and W_0^n, which correspond to the interior point $S = E$; however, from the symmetry on (5.7) we have that the distance from y_1 to $y_o = E$ is the same as the one from x_1 to $x_o = E$. It suggests that the points y_1 e x_1 would be shared in the following sense: x_1 becomes an artificial point on the y grid ($x_1 = y_{-1}$); then, we impose the following boundary condition over y_{-1} :

$$V^n_{-1} = W^n_1 \tag{5.14}$$

Analogously,

$$W^n_{-1} = V^n_1 \tag{5.15}$$

Despite we coudn't simplify the coefficients, the explicit finite-difference implementation of that approach, that we will refer as "E-concentrative", gave us relatively accurate solutions. But the resulting scheme is unstable and may not be generalized to other problems, in order that this approach have only theoretical interest.

However, we can still take advantage of (5.7) without modifying the original equation. We use (5.7) to map a uniform grid onto a non-uniform grid on $[0, L]$, which will be naturally refined around $S = E$, and then use the schemes (5.2)-(5.3) or (5.4)-(5.6). That new approach provided the best results for European Options.

6 Computational Results - I

We summarize the numerical experiments of the following finite-difference schemes:

E1: Standard explicit method

E2: Explicit method / transformation to the diffusion equation

E3: Explicit method / logarithmic transformation $x = \ln(S/E)$

E4: "E-Concentrative" method (see Section 5)

E5: Explicit method with a non-uniform grid defined by (5.7)

C1: Standard Crank-Nicolson method

C2: Crank-Nicolson method / transformation to the diffusion equation

C3: Crank-Nicolson with a non-uniform grid defined by (5.7)

We used a Pentium Pro with 64MB RAM. The algorithms were written in Fortran 77 and timed with Fortran's intrinsic function *Second*. The tests were performed on European call options. We denote by **BL** the solutions given by the Black-Scholes formula. We set L as twice the exercise price E and the other parameters following Geske and Shastri (1985):

S = $40,00

E = $35,00, $40,00, $45,00

$r = \ln(1 + r_F)$, $r_F = 5.00$ percent (annual)

$\sigma = 0.3$ (annual)

T = 1, 4, 7 months

The following tables show the approximate option values and absolute errors when $N = 30$ and $M = 30$ for the explicit methods and $N = 30$ and $M = 10$ for the Crank-Nicolson schemes. These values keep the average absolute errors below 5%

The non-uniform mesh improved the accuracy in both explicit and Crank-Nicolson schemes without losing stability. Both explicit and Crank Nicolson schemes had spent less than 0.01 seconds to evaluate each option price.

Table 1: European call option values.

T	1/12			4/12			7/12		
E	35	40	45	35	40	45	35	40	45
BL	5.2191	1.4614	0.1622	6.2513	3.0729	1.2549	7.1711	4.1860	2.2352
E1	5.2203	1.3494	0.1990	6.2422	3.0335	1.2703	7.1675	4.1641	2.2504
E2	5.2334	1.3894	0.2393	6.2286	3.0504	1.2759	7.1523	4.1792	2.2501
E3	5.2335	1.3895	0.2394	6.2293	3.0513	1.2766	7.1543	4.1815	2.2522
E4	5.2205	1.4544	0.1812	6.2464	3.1008	1.2859	7.1700	4.2251	2.2751
E5	5.2201	1.4103	0.1737	6.2468	3.0548	1.2549	7.1710	4.1782	2.2436
C1	5.2220	1.3416	0.2008	6.2380	3.0211	1.2662	7.1582	4.1478	2.2411
C2	5.2347	1.3854	0.2407	6.2257	3.0395	1.2732	7.1450	4.1662	2.2430
C3	5.2218	1.4036	0.1758	6.2425	3.0428	1.2563	7.1617	4.1622	2.2341

Table 2: Corresponding percent errors in absolute value.

T	1/12			4/12			7/12		
E	35	40	45	35	40	45	35	40	45
E1	0.1166	11.2037	3.6757	0.9088	3.9110	1.5427	0.3506	2.1868	1.5216
E2	1.4325	7.2058	7.7092	2.2705	2.2550	2.1007	1.8771	0.6771	1.4893
E3	1.4363	7.1962	7.7130	2.1980	2.1623	2.1745	1.6784	0.4502	1.6935
E4	0.1396	0.7040	1.8982	0.4889	2.7939	3.1025	0.1061	3.9113	3.9916
E5	0.0946	5.1143	1.1438	0.4546	1.8071	0.5835	0.0006	0.7799	0.8321
C1	0.2828	11.9859	3.8528	1.3334	5.1838	1.1250	1.2840	3.8166	0.5868
C2	1.5580	7.5995	7.8433	2.5627	3.3447	1.8276	2.6055	1.9743	0.7789
C3	0.2635	5.7788	1.3587	0.8842	3.0089	0.1405	0.9318	2.3818	0.1143

The graph of the absolute error versus the asset price reveals that the error is strongly concentrated *at-the-money*.

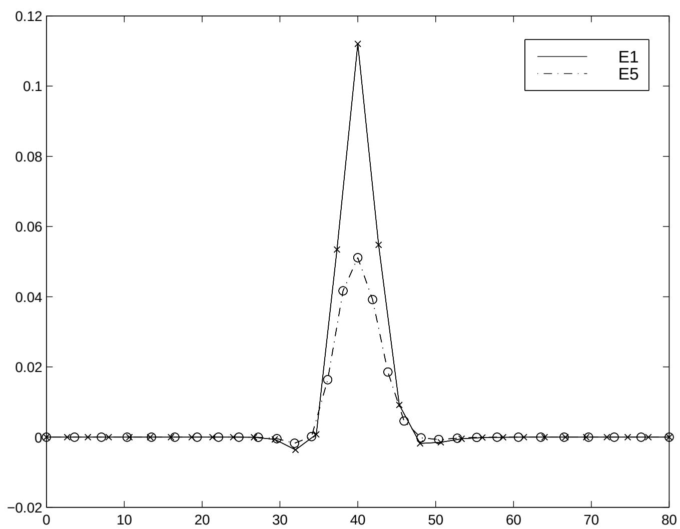

Figure 5: Absolute error (methods **E1** and **E5**) when $E = \$40$ and $T = 1/12$.

Observe that the non-uniform mesh used in **E5** has more points around $E = \$45$ than the uniform one. For both methods, the maximum error occurs at the exercise price.

The following plots illustrate the speed of convergence for some of the methods above. We compare **E1**, **E2**, **E5** and the corresponding Crank-Nicolson schemes. We calculate the option price at $S = \$40$ for $E = \$40$,

$T = 1/12$ for different values of N and M, finding the CPU time and the relative error with respect to the Black-Scholes formula at each evaluation. The results were plotted in logarithmic scale.

We chose M and N in order to satisfy stability conditions ((4.6) or Geske and Shastri (1985), eq.[23]) in the explicit case. We chose $N = 8M$ for **C2** and $N = 4M$ for **C1** and **C3**.

We used linear regression in **E2** and **C2** because the oscillatory behavior observed didn't allow interpolating the points.

All methods have similar convergence patterns. We can compare them by following one of the horizontal grid lines, that represent a fixed amount of time $\bar{T}$. The first graph intercepted by that line from the left to the right corresponds to the method that (approximately) provides the smallest relative error in $\bar{T}$ seconds.

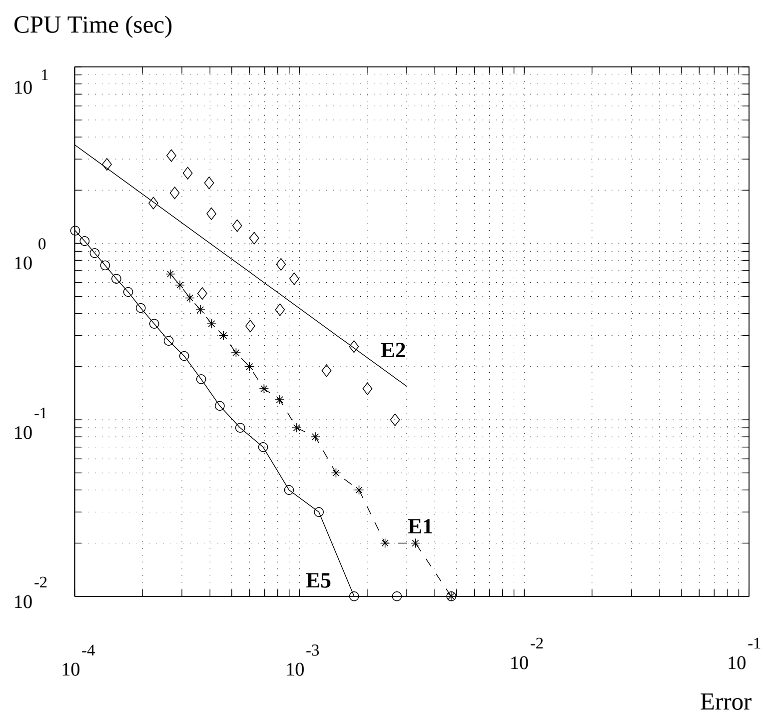

Figure 6: CPU time × relative error for explicit methods.

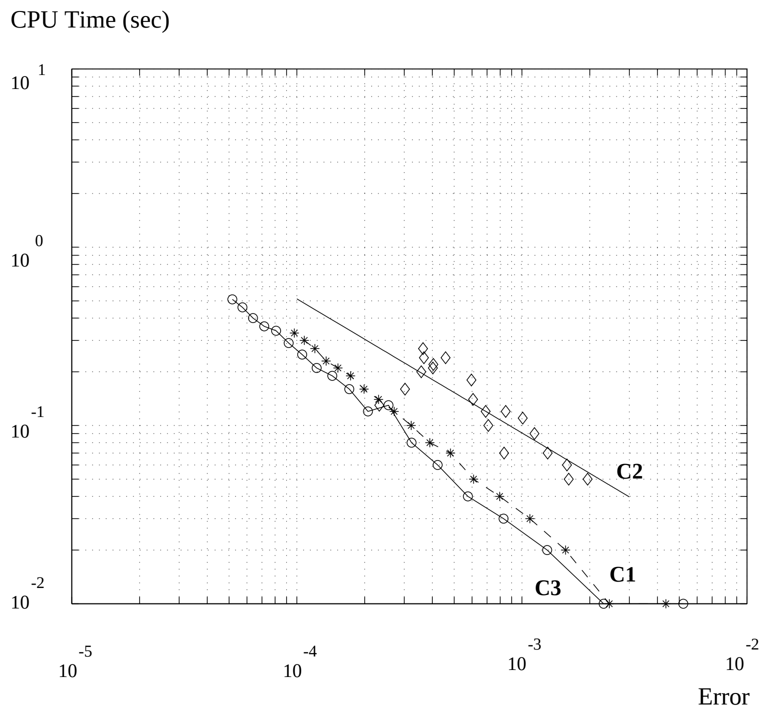

Figure 7: CPU time × relative error for the Crank-Nicolson methods.

7 Pricing American Options

The motivation to analyze numerical methods for the Black-Scholes equation is to extend the same methods to pricing problems without closed formulas. We focus on the American option pricing problem (2.6) - (2.11).

7.1 Finite-Difference Approach

The explicit method leads us to the following scheme (Wilmott *et al.* (1993)):

$$U_i^{n+1} \geq Y_i = c_{-1}U_{i-1}^n + c_o U_i^n + c_1 U_{i+1}^n \tag{7.1}$$

$$U_i^{n+1} \geq g_i \tag{7.2}$$

$$(U_i^{n+1} - Y_i)(U_i^{n+1} - g_i) = 0 \tag{7.3}$$

$$(1 \leq n \leq M-1 \ \ , \ \ 2 \leq i \leq N-1)$$

$$U_i^0 = g_i \quad , 1 \leq i \leq N \tag{7.4}$$

$$U_1^{n+1} = E \ \text{ and } \ U_N^{n+1} = 0 \quad , 1 \leq n \leq M \tag{7.5}$$

The coefficients c_j are given by (3.7) (or (5.3) in the non-uniform case). The solution of (7.1) - (7.5) is also the solution of (7.4) - (7.5) and the following equations :

$$Y_i = c_{-1} U_{i-1}^n + c_o U_i^n + c_1 U_{i+1}^n \tag{7.6}$$

$$U_i^{n+1} = \max\{Y_i, g_i\} \tag{7.7}$$

Analogously for the Crank Nicolson method,

$$b_{-1} U_{i-1}^{n+1} + b_0 U_i^{n+1} + b_1 U_{i+1}^{n+1} \geq Z_i^n = c_{-1} U_{i-1}^n + c_0 U_i^n + c_1 U_{i+1}^n \tag{7.8}$$

$$U_i^{n+1} \geq g_i \tag{7.9}$$

$$(b_{-1} U_{i-1}^{n+1} + b_0 U_i^{n+1} + b_1 U_{i+1}^{n+1} - Z_i^n)(U_i^{n+1} - g_i) = 0 \tag{7.10}$$

$$U_i^{n+1} 0 = g(i) \tag{7.11}$$

$$U_1^{n+1} = E \ \text{ and } \ U_N^{n+1} = 0 \tag{7.12}$$

$$(1 \leq n \leq M-1 \ \ , \ \ 2 \leq i \leq N-1)$$

We have that (7.8)-(7.12) can be written (Wilmott *et al.* (1993)) as the following *linear complementarity problem* (LCP):

$$\begin{cases} C u^{n+1} \geq b^n \\ u^{n+1} \geq g \\ (C u^{n+1} - b^n)^t (u^{n+1} - g) = 0 \\ u^0 = g \end{cases} \tag{7.13}$$

LCPs are usually solved by iterative methods (Oden and Kikushi (1980), Pissarra (1997), Wilmott *et al.* (1993)), and most of them are limited to symmetric matrices. This makes the transformation to the diffusion equation proposed by Wilmott *et al.* (1993) attractive. Wilmott *et al.* (1993) transform the problem and solve the resulting LCP by the Projected SOR method.

We can achieve some improvement by a suitable choice of the boundary conditions. If we would know the free boundary $S_f(t)$ at each time t, we could replace the boundary condition at $S = 0$ by the following:

$$P(S_f(t), t) = (E - S_f(t))^+ \ \ , \ \ 0 < t < T \tag{7.14}$$

As $S_f(t)$ is also an unknown, we must replace $S_f(t)$ in (7.14) by an uniform lower bound S^*. We used an accurate lower bound provided by Broadie and Detemple (1996) that only requires to find the zero of a function of one variable. As $S_f(t)$ is nondecreasing with respect to t (Myneni (1992), 13), we just need to estimate $S_f(0)$ (what corresponds to $S_f(T)$ when we apply the transformation $\tau = T - t$); therefore, we can get a good approximation for (7.14) at a reasonable cost. This choice also alleviates the excessive refinement near $S = 0$ due to the logarithmic transformation. The following plot shows the effects of the (estimate) boundary condition (7.14).

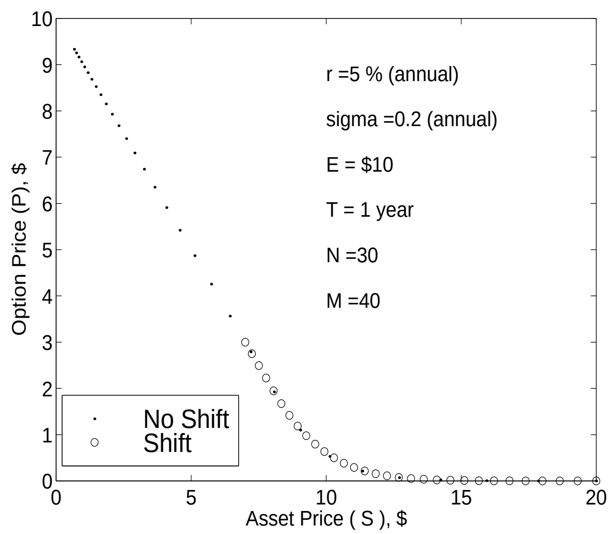

Figure 8: Transformed finite-difference solutions for American put options

7.2 Finite-Element Approach

The finite-element method approach of American option pricing is based on the formulation of the problem (2.6) - (2.11) as a variational inequality (Baiocchi and Capelo (1984), Wilmott *et al.* (1993)). In this section we consider a more general setting in order to have consistency with the standard literature.

The abstract formulation of the (elliptic) variational inequality problem is as follows: let $(V, \langle ., . \rangle)$ be a real Hilbert space with norm defined by $\|u\|^2 = \langle u, u \rangle$ and let $K \subset V$ be a closed and convex subset of V. Let $a(., .)$ be a bilinear continuous functional and $f \in V$. We consider the problem

(VI) Find $u \in K$ such that $a(u, v-u) \geq \langle f, v-u \rangle \ \ \forall\, v \in K$

The Lions-Stampacchia theorem (Lions and Stampacchia (1967)) states that the problem (VI) have a unique solution, if the bilinear form $a(.,.)$ satisfies some special properties. One particular case is the following:

$$a(v,v) = 0 \Rightarrow v \equiv 0 \tag{7.15}$$

In most of the applications, the convex set K has the following form:

$$K = \{v \in V; v(x) \geq \psi(x) \ \ \forall\, x \in \Omega\} \tag{7.16}$$

There are several different approaches to solve the problem (VI) (Glowinski *et al.* (1976), Oden and Kikushi (1980)). Here we consider the *Galerkin method*, which determines the best approximation u_h of u over a finite-dimensional subspace of V.

Let $V_h = \text{span}\{\varphi_1, \ldots, \varphi_n\}$ be a subspace of V. We approximate the solution u of (VI) and the function ψ in (7.16) as

$$u_h = \sum_{j=1}^{n} u_j \varphi_j \quad , \quad \psi_h = \sum_{j=1}^{n} \psi_j \varphi_j \tag{7.17}$$

Thus,

$$K_h = \{v = \sum_{j=1}^{n} v_j \varphi_j \in V \ ; \ v_j \geq \psi_j \ \ , \ \ 1 \leq j \leq n\} \tag{7.18}$$

The approximate solution u_h is the solution of the following problem:

(VI$_h$) Find $u_h \in K_h$ such that $a(u_h, v_h - u_h) \geq (f, v_h - u_h) \ \ \forall\, v \in K_h$

In Appendix B we verify that it is sufficient to have

$$\begin{cases} Ax \geq b \\ x \geq c \\ (Ax-b)^t (x-c) = 0 \end{cases} , \tag{7.19}$$

another LCP, where

$$\begin{array}{llll} x &= (u_1, \ldots, u_n) & c &= (\psi_1, \ldots, \psi_n) \\ & & b &= (\langle f, \varphi_1 \rangle, \ldots, \langle f, \varphi_n \rangle) \\ A = [a_{ij}], & & a_{ij} &= a(\varphi_j, \varphi_i) \ \ , \ \ 1 \leq i,j \leq n \end{array} \tag{7.20}$$

Wilmott *et al.* (1993) presented a similar result for the case when $a(\cdot,\cdot)$ is symmetric. As far as we know, this is the first time that the Galerkin method for variational inequalities is extended to non-symmetric bilinear forms.

The formulation (VI) can be extended to time-dependent problems in the following way:

$$(\mathrm{VI}_t) \text{ Find } u(t) : [0,T] \to K \text{ such that}$$

$$\begin{cases} \langle u_t, v-u\rangle + a(u, v-u) \geq \langle f, v-u\rangle \ \ \forall\, v \in K \quad \forall\, t \in [0,T] \\ \langle u(0), v\rangle = \langle u_0, v\rangle \ \ \forall\, v \in K \end{cases}$$

Consider a partition $\tau : 0 = t_1 < t_2 < \ldots t_M = T$ of $[0,T]$. Approximating u in τ by U^n, where

$$u_t \approx \frac{U^{n+1} - U^n}{\Delta t} \quad , \quad U^n = U(t_n) \in K \tag{7.21}$$

we have the following discretization of (VI_t):

$$(\mathrm{VI}_\tau) \text{ Find } U^0, \ldots, U^M \in K \text{ such that}$$

$$\begin{cases} \langle \frac{U^{n+1}-U^n}{\Delta t}, v - U^{n+1}\rangle + a(U^{n+1}, v - U^{n+1}) \geq \langle f, v - U^n + 1\rangle \quad , \ n \geq 1 \\ \\ \langle U^0, v\rangle = \langle u_0, v\rangle \qquad\qquad\qquad\qquad\qquad\qquad \forall\, v \in K \end{cases}$$

Defining

$$a_\tau(u,v) = \langle u, v\rangle + \Delta t a(u,v) \quad , \quad f_n = U^n + \Delta t f \quad , \tag{7.22}$$

we have that the solution U^{n+1} of (VI_τ) for fixed n, given that U^n is known, satisfies

$$a_\tau(U^*, v - U^*) \geq \langle f_n, v - U^*\rangle \ \ \forall\, v \in K \tag{7.23}$$

The variational inequality (7.23) has the same form as (VI) and can be solved by the Galerkin method presented before.

We can also apply the Crank-Nicolson approach to solve (VI_t). Defining $\delta U^n = U^{n+1} - U^n$, the inequality for the time step $n+1$ is given by

$$\frac{1}{\Delta t}\langle \delta U^n, v - U^{n+1}\rangle + \frac{1}{2} a(\delta U^n, v - U^{n+1}) \geq \langle f, v - U^{n+1}\rangle \ \ \forall\, v \in K \tag{7.24}$$

See Johnson (1987) for further time-space discretization procedures. Convergence estimates can be proved for implicit schemes (Johnson (1976)).

Let us use the structure above in the American option pricing problem. We start defining V as a space of smooth functions in $[0, L]$ (so that we can define derivatives), the inner product $\langle \cdot, \cdot \rangle$ as

$$\langle u, v \rangle = \int_0^L u(x)v(x)dx \tag{7.25}$$

and the convex set K as

$$K = \{v \in V; v(x) \geq (E - S)^+ \ \ \forall\, x \in [0, 1]\} \tag{7.26}$$

The bilinear form is derived by multiplying the left-hand side of (2.6) by a function $v \in V$ and integrating over $[0, L]$

$$a(u, v) = \frac{1}{2}\sigma^2 \langle S^2 \frac{\partial u}{\partial S}, \frac{\partial v}{\partial S} \rangle - (r - \sigma^2)\langle S \frac{\partial u}{\partial S}, v \rangle + r\langle u, v \rangle \tag{7.27}$$

The problem (VI_t) with K given (7.26) and $a(\cdot, \cdot)$ given by (7.27) is equivalent to (2.6)-(2.11) (Baiocchi and Capelo (1984), Oliveira (1998)). Observe that $a(\cdot, \cdot)$ satisfies (7.15).

Remark: Some important details about the space V and boundary conditions (Oliveira (1998), Wilmott *et al.* (1993)) were skipped for simplicity.

Consider the partition $\pi : 0 = S_0 < S_1 < \ldots S_N = L$ of $[0, L]$. We have used the standard piecewise linear finite-element functions, that is, we consider in the Galerkin method $n = N + 1$ and for $0 \leq j \leq N$,

$$\phi_j(S) = \begin{cases} \frac{S - S_{j-1}}{S_j - S_{j-1}} & , \ S \in [S_{j-1}, S_j] \quad (j > 0) \\ \\ \frac{S_{j+1} - S}{S_{j+1} - S_j} & , \ S \in [S_j, S_{j+1}] \quad (j < N) \\ \\ 0 & \text{elsewhere} \end{cases} \tag{7.28}$$

See Johnson (1987), Wilmott *et al.* (1993) for further details.

7.3 Iterative Methods for LCPs

Both finite-difference and finite-element approaches to solve (2.6)-(2.11) resulted in a LCP. We have used the classical Projected SOR method (Wilmott *et al.* (1993)) and a modification of the conjugate gradient squared algorithm for LCPs (Pissarra (1997)).

Given a $n \times n$ real matrix $A = [a_{ij}]$ and a column vector $b = (b_j) \in I\!R^n$, we state the LCP as follows:

$$\text{Find } x = (x_j) \in I\!R^n \text{ such that}$$

$$\begin{cases} Ax \geq b \\ x \geq \mathbf{0} \\ (Ax - b)^t x = 0 \end{cases}, \tag{7.29}$$

where $\mathbf{0}$ is the zero vector in $I\!R^n$. This formulation is the same as (7.19) up to a change of variables. Iterative methods generate a sequence $x^{(1)}$, $x^{(2)}$, $\ldots$ of approximate solutions to (LCP). We denote by $x^{(0)}$ the initial guess for x and define $r^{(k)} = b - Ax$

The Projected SOR method can only be applied to (2.6)-(2.11) after the transformation to the diffusion equation, as it requires A to be symmetric.

Oden and Kikushi (1980) present a relationship between the projected SOR method and the fixed-point method. As the fixed-point algorithm (Oden and Kikushi (1980)) does not require symmetry, the projected SOR method for non-symmetric LCPs might converge for suitable values of ω.

Algorithm: Projected SOR Method
input A, b, n, $0 < \omega < 2$, $x^{(0)}$, $\epsilon > 0$, $MAX > 0$

$k = 0$
while $(|(r^{(k)})^t x^{(k)}| > \epsilon$ **and** $k < MAX)$ **do**
  for $i = 1$ **to** n **do**

$$y = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} x_j^{(k)} a_{ij} - \sum_{j=i+1}^{n} x_j^{(k)} a_{ij} \right)$$

$$x_i^{(k+1)} = \max\{x_i^{(k)} + \omega(y - x_i^{(k)}), 0\}$$

  end
  $k = k + 1$
end

The second method, referred in Pissarra (1997) as *projected conjugate gradient squared* (CGSP), allows us to solve the non-symmetric LCPs arising from the untransformed problem. Its algorithm is as follows:

Algorithm: Method CGSP
input A, b, n, $x^{(0)}$, $\epsilon > 0$, $MAX > 0$

$k = 0$
while $(|(r^{(k)})^t x^{(k)}| > \epsilon$ and $k < MAX)$ **do**
    $P =$ **IdentityMatrix**(n)
    for $i = 1$ **to** n **do**
        if $(x_i^{(k)} < 0)$
            $P_{ii} = 0$
            $x_i^{(k)} = 0$
        end if
    end
    $R^{(k)} = Ax^{(k)} - b$
    for $i = 1$ **to** n **do**
        if $(P_{ii} = 0)$ **and** $(R_i^{(k)} > 0)$
            $P_{ii} = 1$
        end if
        $x^{(k+1)} =$**CGS**$(x^{(k)}, PAP, Pb, n, \epsilon)$
        $k = k + 1$
    end
end

Function CGS$(x^{(0)}, A, b, n, \epsilon)$
$g^{(0)} = b - Ax^{(0)}$
$d^{(-1)} = q^{(0)} = \mathbf{0}$
$\rho_{-1} = 1$
$k = 0$
while $(\|g^{(k)}\| > \epsilon)$ **do**
    $\rho_k = (g^{(0)})^t g^{(k)}$
    $\beta_k = \rho_k / \rho_{k-1}$
    $u^{(k)} = g^{(k)} + \beta_k q^{(k)}$
    $d^{(k)} = u^{(k)} + \beta_k \left(q^{(k)} + \beta_k d^{(k-1)}\right)$
    $v^{(k)} = Ad^{(k)}$
    $\sigma_k = (g^{(0)})^t v^{(k)}$
    $\alpha_k = \rho_k / \sigma_k$
    $q_{k+1} = u^{(k)} - \alpha_k v^{(k)}$
    $g^{(k+1)} = g^{(k)} - \alpha_k A \left(u^{(k)} + q^{(k+1)}\right)$

$x^{(k+1)} = x^{(k)} + \alpha_k \left(u^{(k)} + q^{(k+1)} \right)$

$k = k + 1$

end

return$(x^{(k)})$

8 Computational Results - II

The numerical experiments for American put options include the methods described above and some results from Huang *et al.* (1996) and Sullivan (1997).

BN: Binomial method with $n = 10000$ (Huang *et al.* (1996))

X1: Standard explicit method

X2: Explicit method with a non-uniform grid defined by (5.7)

S1: Crank-Nicolson method / transformation to the diffusion equation

S2: Crank-Nicolson method / transformation to the diffusion equation and shifted boundary conditions

F1: Standard Crank-Nicolson Method / Finite-Element Method

F2: Standard Crank-Nicolson, non-uniform grid / Finite-Element Method

HU: Recursive Integration Method (Huang *et al.* (1996))

QD: Extrapolated Quadrature Method (Sullivan (1997))

The parameters are the same as in Section 6. Schemes **S1** and **S2** use the projected SOR algorithm, while **F1** and **F2** use the projected conjugate gradient squared one. The method **S2** shifts the boundary condition at $S = 0$ to $S = 0.9S^*$, where S^* is a lower bound for $S_f(T)$ (Broadie and Detemple (1996)).

The following tables show the option values and absolute errors with respect to **BN**. We set $N = 100$ and $M = 950$ for the explicit methods, $N = 200$ and $M = 200$ for **F1**, **F2** and $N = 200$ and $M = 100$ for **S1**, **S2**. The entries for the methods **HU** and **SU** are the same as in Sullivan (1997).

Table 3: American put option values.

T	1/12			4/12			7/12		
E	35	40	45	35	40	45	35	40	45
BN	0.2466	1.7681	5.2868	1.3460	3.3874	6.5099	2.1594	4.3526	7.3830
X1	0.2461	1.7613	5.2887	1.3542	3.3845	6.5109	2.1545	4.3509	7.3843
X2	0.2471	1.7647	5.2879	1.3462	3.3859	6.5105	2.1552	4.3519	7.3839
S1	0.2487	1.7683	5.2919	1.3459	3.3871	6.5121	2.1544	4.3524	7.3847
S2	0.2466	1.7681	5.2868	1.3460	3.3874	6.5099	2.1594	4.3526	7.3830
F1	0.2468	1.7669	5.2865	1.3450	3.3848	6.5078	2.1527	4.3491	7.3798
F2	0.2468	1.7670	5.2868	1.3448	3.3849	6.5079	2.1526	4.3491	7.3798
HU	0.2467	1.7694	5.2853	1.3468	3.3970	6.5128	2.1603	4.3699	7.3865
QD	0.2467	1.7685	5.2868	1.3461	3.3876	6.5097	2.1594	4.3527	7.3828

Table 4: Corresponding percent errors in absolute value.

T	1/12			4/12			7/12		
E	35	40	45	35	40	45	35	40	45
X1	0.0467	0.6774	0.1851	0.0772	0.2915	0.1033	0.4939	0.1684	0.1329
X2	0.0495	0.3446	0.1108	0.0223	0.1456	0.0628	0.4179	0.0706	0.0903
S1	0.2070	0.0198	0.5099	0.0099	0.0263	0.2208	0.4956	0.0220	0.1695
S2	0.0301	0.0082	0.0084	0.0039	0.0048	0.0089	0.4502	0.0056	0.6861
F1	0.0249	0.0119	0.0223	0.0472	0.0121	0.0259	0.4160	0.0179	0.0325
F2	0.0179	0.0171	0.0414	0.0310	0.0137	0.0325	0.4329	0.0161	0.0343
HU	0.04	0.07	0.03	0.06	0.28	0.04	0.25	0.28	0.05
QD	0.04	0.02	0.00	0.01	0.01	0.00	0.00	0.00	0.00

The explicit methods had the best computational time (around 0.15 seconds) but were less accurate. The methods based on LCPs spent in mean 3.0 seconds but had an absolute error as low as the methods presented in Huang *et al.* (1996) and Sullivan (1997).

The shifted boundary condition was a remarkable improvement over the regular boundary condition for the transformation to the diffusion equation.

The non-uniform mesh on the method **F2** didn't performed well. One reason is that more refinement is needed around the free-boundary point, in order that $S = E$ is no longer the right place to refine. This also explains the good performance of **S2** relative to **S1** and all other schemes.

A probable reason for the bad performance of most methods at $E = \$35$ and $T = 7/12$ is the distance of the point $S = \$40$ to the exercise price and specially from the free boundary (that is around \$22). The option price at

S is relatively closer to the boundary condition at the right ($P = 0$), what may cause an under-pricing.

The condition (4.9) gives $M \geq 934$ for $N = 100$ and $T = 1/12$, and the choice $M = 950$ could preserve stability. It suggests that the estimate (4.6) is still adequate in explicit methods for American options.

The following plot compares the speed of convergence for some of the methods as in Section 6. We compare **X1**, **S1**, **S2** and **F1**. Now the relative error is calculated with respect to the method **BN**.

We chose M and N satisfying (4.6) for **X1**, $N = M$ for **F1** and $N = 2M$ for **S1** and **S2**.

The transformed methods (**S1** and **S2**) kept the same tendency observed on European option, while the others had a cyclic behavior. The slopes are steeper than in the methods for the European options, what represents a loss of convergence order. The explicit method seems to be a good alternative when its parameters satisfy the stability condition (4.6).

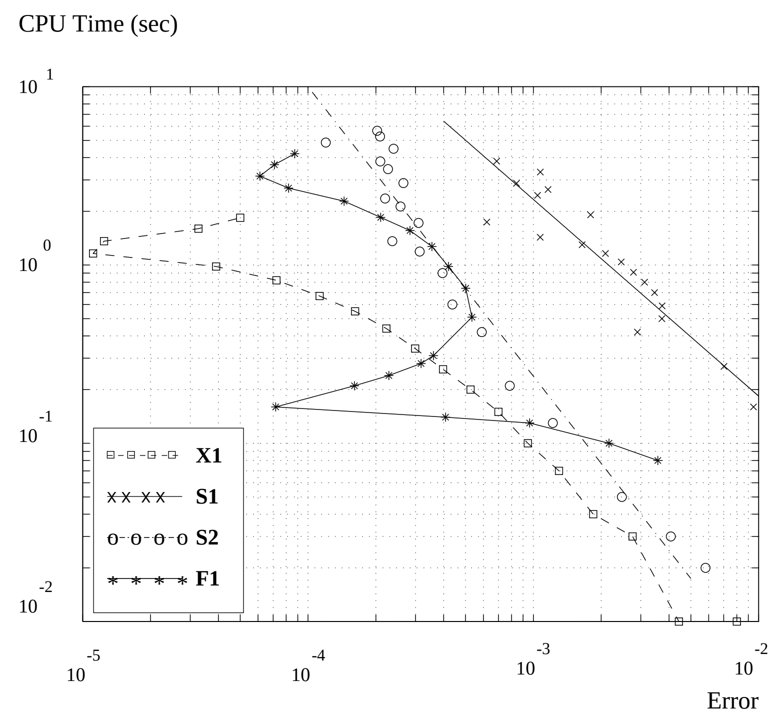

Figure 9: CPU time × relative error for American options.

9 Conclusions

Finite-differences can be relatively accurate, given that aspects like stability and choice of a suitable mesh are taken into account. We have proposed attractive alternatives for both issues based on the standard option pricing theory and several numerical experiments.

The main achievements were the stability condition (4.6) for explicit methods, which was accurate for both European and American options, the non-uniform mesh generated by (5.7), consistent with both financial and mathematical objectives, the extension of the Finite- Element technique suggested in Wilmott *et al.* (1993) for non-symmetric problems and the shifted boundary conditions for American option pricing.

The extension of the ideas above to more complex derivatives seems to be natural. Finite-difference features were exported to finite-elements methods and could be used in other numerical methods as well.

APPENDICES

Appendix A. Uniqueness of the Optimal Exercise Price

This appendix presents an explicit proof that the optimal exercise price of an American put option $P(S,t)$ is unique for each $t \geq 0$. The arguments are based on Merton (1973) and Rudin (1976).

We start by observing that the following results from Merton (1973) will be valid if $P(S,t)$ satisfies the Black-Scholes hypothesis:

$$P(S,t) \geq C(S,t) - S + Ee^{-r(T-t)} \tag{A.1}$$

$$P(S,t) \leq C(S,t) - S + E \tag{A.2}$$

$$P(\lambda S_1 + (1-\lambda)S_2, t) \leq \lambda P(S_1,t) + (1-\lambda)P(S_2,t) \tag{A.3}$$

$$(\ \forall\, S,\ S_1,\ S_2 \geq 0,\ \ \forall\, t \geq 0,\ \ \forall\, \lambda \in [0,1]\ )$$

The inequality (A.3) means that $P(S,t)$ is convex with respect to S.

Lemma A.1 : $P(0,t) = E$ and $P(S,t) > (E-S)^+ \ \forall\, S \geq E,\ t > 0$

Proof : From the Black-Scholes formula (2.1) and the fact that $C(S,t) = c(S,t)$ in absence of dividends,

$$C(S,t) > S - Ee^{-r(T-t)} \ \ \forall\, S > 0 \text{ and } t > 0 \tag{A.4}$$

$$C(0,t) = 0 \tag{A.5}$$

From (A.1) and (A.4),

$$P(S,t) > 0 = (E-S)^+ \ \ \forall\, S \geq E \text{ and } t > 0 \tag{A.6}$$

From (A.2) and (A.5),

$$P(0,t) \leq E \tag{A.7}$$

By (2.7), $P(0,t) \geq E$; thus, $P(0,t) = E$.

Theorem A.1 : If $\bar{S} \leq E$ satisfies $P(\bar{S},t) > (E-\bar{S})^+$, then $P(S,t) > (E-S)^+$ for any S such that $\bar{S} < S \leq E$.

Proof : Suppose that there is $S' \in (\bar{S}, E]$ such that

$$P(S',t) = (E-S')^+ = (E-S') \tag{A.8}$$

As $S' > \bar{S} > 0$, $\bar{S} = \lambda S' = \lambda S' + (1-\lambda)0$ for some $\lambda \in (0,1)$; therefore,

$$\begin{aligned} P(\lambda S' + (1-\lambda)0, t) \ &= P(\bar{S},t) \\ &> (E-\bar{S})^+ \\ &= E - \lambda S' + (1-\lambda)0 \\ &> \lambda(E-S') + (1-\lambda)0 \\ &= P(S',t) + (1-\lambda)0 \end{aligned} \tag{A.9}$$

From lemma A.1,

$$P(\lambda S' + (1-\lambda)0, t) > \lambda P(S',t) + (1-\lambda)P(0,t) \ \ , \tag{A.10}$$

what contradicts (A.3); thus, $P(S,t) > (E-S)^+ \ \ \forall\, S \in (\bar{S}, E]$.

Theorem A.2 : There exists a unique point $S_f = S_F(t)$ such that

$$\begin{cases} P(S,t) = (E-S)^+ \ , & S \leq S_f(t) \\ P(S,t) > (E-S)^+ \ , & S > S_f(t) \end{cases} \tag{A.11}$$

Proof : From lemma A.1, $P(S,t) = (E-S)^+$ for $S = 0$; thus, the set

$$I = \{S \geq 0\ ;\ P(S,t) = (E-S)^+\} \tag{A.12}$$

is nonempty. As P is continuous, the set I is closed and the set

$$I^c = \{S \geq 0\ ;\ P(S,t) > (E-S)^+\} \ \ , \tag{A.13}$$

the complement of I, is open. By lemma A.1, $P(S,t) > (E-S)^+$ for $S \geq E$; thus, I^c is also nonempty. As I^c is bounded below (0 is a lower bound), it has an *infimum*, i.e., there is a unique $S_f \in I$ (as I^c is open) such that

$$\begin{cases} S_f < S & \forall\, S \in I^c \\ S_f \geq S & \forall\, S \in I \end{cases} \tag{A.14}$$

Therefore, $P(S,t) = (E-S)^+$ for $S \leq S_f$ and $P(S,t) > (E-S)^+$ for $S_f < S \leq E$. From theorem A.1, $P(S,t) > (E-S)^+$ also for $S_f < S \leq E$.

Appendix B. Galerkin Method

Here we present the details of the derivation of (7.19) from the Galerkin formulation (VI_h) of the variational inequality (VI).

Let v_h be an arbitrary element in V_h given by

$$v_h = \sum_{j=1}^{n} v_j \varphi_j \tag{B.1}$$

Replacing (B.1), (7.17) in (VI_h):

$$a\left(\sum_{j=1}^{n} u_j\varphi_j, \sum_{i=1}^{n}(v_i-u_i)\varphi_i\right) \geq \left\langle f, \sum_{i=1}^{n}(v_i-u_i)\varphi_i\right\rangle$$
$$\sum_{i=1}^{n}(v_i-u_i)\sum_{j=1}^{n} u_j a(\varphi_j,\varphi_i) \geq \sum_{i=1}^{n}(v_i-u_i)\langle f,\varphi_i\rangle$$

Using the notations in (7.20), we have

$$(y-x)^t A x \geq b^t(y-x) \;\; \forall\, y \geq c \;\; , \;\; v = (v_i \ldots, v_n) \tag{B.2}$$

Let us rewrite (B.2) in order to get sufficient conditions over x:

$$u^t A^t((y-c)-(x-c)) \geq b^t((y-c)-(x-c))$$
$$u^t A^t(y-c) - b^t(y-c) \geq u^t A^t(x-c) - b^t(x-c)$$
$$(Ax-b)^t(y-c) \geq (Ax-b)^t(x-c) \;\; \forall\, y \geq c \tag{B.3}$$

As $(y-c) \geq 0$, we have from (B.3) the following sufficient conditions:

$$\begin{cases} Ax \geq b \\ x \geq c \\ (Ax-b)^t(x-c) = 0 \end{cases} \tag{B.4}$$

Acknowledgments
This work was supported in part by CAPES/MEC, Brazil, and the Organization of American States. The numerical experiments were performed at the Center for Computational Mathematics of the University of Colorado at Denver, United States.

References

Baiocchi, C. and Capelo, A. *Variational and Quasivariational Inequalities.* John Wiley & Sons, 1984.

Black, F. and Scholes, M. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, **81** (1973), 637-659.

Brennan, M. J. and Schwarz, E.S. Finite Difference Methods Arising in the Princing of Contingent Claims: a Synthesis. *Journal of Financial and Quantitative Analysis*, **13** (1978), 461-474.

Broadie, M. and Detemple, J. American Option Valuation: new bounds, Approximations and a Comparison of Existing Methods. *The Review of Financial Studies*, **9** (1996), 1211-1250.

Courtadon, G. A More Accurate Finite Difference Approximation for the Valuation of Options. *Journal of Financial and Quantitative Analysis*, **17** (1982), 697-705.

Duffie, D. *Security Markets: Stochastic Models.* Academic Press, 1988.

Forsythe, G. E. and Wasow, W.R. *Finite Difference Methods for Partial Differential Equations.* John Wiley & Sons, 1978

Geske, R. and Shastri, K. Valuation by Approximation : A Comparison of Alternative Option Valuation Techniques. *Journal of Financial and Quantitative Analysis*, **20** (1985), 45-71.

Glowinski, R. ,Lions, J. L. and Trémolières. *Analyse Numérique de Inéquations Variationnelles, Tôme 1.* Dunod, 1976.

Huang, J., Subrahmanyam, M. G. and Yu, G.G. Pricing and Hedging American Options: A Recursive Integration Method. *The Review of Financial Studies*, **9** (1996), 277-300.

Hull, J. *Options, Futures and Other Derivatives*. 3d Ed. Prentice Hall, 1997.

Hull, J. and White, A. Valuing Derivative Securities Using the Explicit Finite Difference Method. *Journal of Financial and Quantitative Analysis*, **25** (1990), 87-100.

Johnson, C. A Convergence Estimate for an Approximation of a Parabolic Variational Inequality. *SIAM Journal of Numerical Analysis*, **13** (1976), 599-606.

Johnson, C. *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, 1987.

Lions, J. L. and Stampacchia, G. Variational Inequalities. *Commununications on Pure and Applied Mathematics*, **20** (1967), 493-519.

Merton, R. C. Theory of Rational Option Pricing. *Bell Journal of Economic and Management Science*, **4** (1973) , 141-83.

Myneni, R. The Pricing of the American Option. *The Annals of Applied Probability*, **2** (1992), 1-23.

Oden, J. T. and Kikushi, K. Theory of Variational Inequalities with Applications to Problems of Flow through Porous Media. *International Journal of Engineering Science*, **18** (1980) , 1173-1284

Pissarra, C. A. Problemas de Complementaridade Linear: Aspectos Teóricos e Computacionais. Master's Thesis. State University of Campinas, Brazil, 1997.

Oliveira, S. P. Métodos Numéricos para Precificação de Opções. Master's Thesis. State University of Campinas, Brazil, 1998

Sullivan, M. A. Valuing American Options Using Gaussian Quadrature. Third International Conference on Computing in Economics and Finance. Stanford, CA, 1997.

Thomas, J. W. *Numerical Partial Differential Equations - Finite Difference Methods*. Springer, 1995

Rudin, W. *Principles of Mathematical Analysis*. 2d Ed. McGraw-Hill, 1976

Wilmott, P. , Howison, S. and Dewynne, J. *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press, 1993

Zvan, R., Forsyth, P. A. and Vetzal, K. Robust Numerical Methods for PDE models of Asian Options. *The Journal of Computational Finance*, **2** (1998), 39-78.