Variational inequalities and the pricing of American options *

Patrick Jaillet Damien Lamberton Bernard Lapeyre

CERMA-ENPC La Courtine
93167 NOISY LE GRAND — FRANCE

Abstract

This paper is devoted to the derivation of some regularity properties of pricing functions for American options and to the discussion of numerical methods, based on the Bensoussan-Lions methods of variational inequalities. In particular, we provide a complete justification of the so-called Brennan-Schwartz algorithm for the valuation of American put options.

AMS subject classification (1980). 90A09, 60G40, 60J60, 65K10, 65M10.

Key words. Option pricing, variational inequalities, optimal stopping problem.

1 Introduction

In their celebrated paper [4], Black and Scholes derived explicit pricing formulas for European call and put options on stocks which do not pay dividends. The absence of closed form expressions for the pricing of American options caused an extensive literature on numerical methods, with, in general, little mathematical support.

Actually, the rigorous theory of American options is rather recent (see [2], [22], [23], [24]). In [2], [22], Bensoussan and Karatzas derive pricing formulas involving Snell envelopes. The aim of the present paper is to study

*Research supported in part by a contract from Banque INDOSUEZ

and exploit these formulas in the framework of diffusion models, including the classical Black-Scholes [4] and Garman-Kohlagen [17] models. In this setting, we can rely on the connection between optimal stopping and variational inequalities as established by Bensoussan and Lions (see [3], [14], chapter 16) in order to investigate regularity properties of pricing functions and discuss the accuracy of numerical methods. In particular, we will examine an algorithm due to Brennan and Schwartz [5].

In section 2, we set the basic assumptions of our model and show how the price of an American option can be obtained as a function of the current stock prices. Note that in order to deal with non-degenerate operators, we choose the logarithms of stock prices as state variables. Section 3.1 contains general results on variational inequalities which are for the most part included in [3]. We have tried to weaken some regularity assumptions, so that the results can be applied to put and call options. In section 3.2, we concentrate on one-dimensional models. We show, among other things, that the American put price, in the Black-Scholes model, is a continuously differentiable function of the stock price, before maturity (note that this property can also be derived from [32]).

In section 4, we localize the inequalities and review some results of [18] on the approximation of variational inequalities. Section 5 contains a detailed discussion of an algorithm which, except for the logarithmic change of variable, is the Brennan-Schwartz method of valuation of American put options (cf.[5]). It turns out that, although the formulation of the boundary problem in [5] was mathematically incorrect, this algorithm can be completely justified. However, it must be emphasized that this justification relies on special properties of put and call options and that the method fails for the pricing of other types of options (see remark 5.14 below). Therefore, we have included some remarks on alternate methods in the last section.

In order to complete this introduction, we wish to mention van Moerbeke's paper [32], on optimal stopping and free boundary problems, whose results on warrant pricing extend readily to American options. The main advantage of our approach is that the techniques of variational inequalities provide the adequate framework for the study of numerical methods. For regularity results on European options, we refer the reader to [16].

This paper is a detailed version of [20].

2 Assumptions and notations

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and $(W_t)_{t\geq 0}$ a standard brownian motion with values in $\mathbf{R}^n$. We denote by $(\mathcal{F}_t)_{t\geq 0}$ the $\mathbf{P}$-completion of the natural filtration of $(W_t)_{t\geq 0}$.

Consider a financial market with n risky assets, with prices $S_t^1, \dots, S_t^n$ at time t, and let X_t be the n-dimensional vector with components : $X_t^j = \log S_t^j$, for $j = 1, \dots, n$. We will assume that (X_t) satisfies the following stochastic differential equation :

$$dX_t = \beta(t, X_t)dt + \sigma(t, X_t).dW_t \tag{2.1}$$

on a finite interval $[0, T]$, where T is the horizon (namely the date of maturity of the option).

We impose the following conditions on β and σ, and on the interest rate :

(H1) $\beta(t, x)$ is a bounded C^1 function from $[0, T] \times \mathbf{R}^n$ into $\mathbf{R}^n$, with bounded derivatives.

(H2) $\sigma(t, x)$ is a bounded C^1 function from $[0, T] \times \mathbf{R}^n$ into the space of $n \times n$ matrices, with bounded derivatives. Also σ admits bounded continuous second partial derivatives with respect to x, $\partial^2 \sigma_{i,j}/\partial x_i \partial x_j$, satisfying a Hölder condition in x, uniformly with respect to (t, x) in $[0, T] \times \mathbf{R}^n$

(H3) The entries $a_{i,j}$ of $\sigma(t, x).\sigma^*(t, x)/2$ (where $*$ denotes transposition) satisfy the following coercivity property :

$$\exists \eta > 0\ \forall (t, x) \in [0, T] \times \mathbf{R}^n\ \forall \xi \in \mathbf{R}^n \sum_{1 \leq i,j \leq n} a_{i,j}(t, x)\xi_i \xi_j \geq \eta \sum_{i=1}^{n} \xi_i^2$$

(H4) The instantaneous interest rate $r(t)$ is a C^1 function from $[0, T]$ into $[0, \infty)$.

Remark 2.1 Note that condition **(H3)** can be interpreted in terms of completeness of the market (cf.[19], [2], [23], [24])

An American contingent claim is defined by an adapted process $(h(t))_{0\leq t\leq T}$, where $h(t)$ is the payoff of the claim if exercized at time t. The typical examples are put and call American options : for a put (resp. call) option on one unit of asset 1, with exercise price K, one has :

$h(t) = (K - S_t^1)_+$ (resp. $h(t) = (S_t^1 - K)_+$). Note that, for simplicity, we do not consider contingent claims allowing a payoff rate per unit of time as in [22]. Throughout the paper, we will assume that $h(t)$ depends only on the prices of risky assets at time t, so that : $h(t) = \psi(X_t)$, where ψ is a continuous function from $\mathbf{R}^n$ into $\mathbf{R}$. Note that in the case of a put (resp. a call) on asset 1, one has : $\psi(x_1) = (K - e^{x_1})_+$ (resp. $(e^{x_1} - K)_+$).

We will denote by $(X_s^{t,x})_{s \geq t}$ a continuous version of the flow of the stochastic differential equation (2.1). Therefore, $(s,t,x) \mapsto X_s^{t,x}(\omega)$ is continuous for almost all ω, $X_t^{t,x} = x$, and $X_.^{t,x}$ satisfies (2.1) on $[t,T]$.

Proposition 2.2 *Assume ψ is continuous and satisfies :$|\psi(x)| \leq Me^{M|x|}$ for some $M > 0$, and define a function u^* on $[0,T] \times \mathbf{R}^n$, by :*

$$u^*(t,x) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbf{E}\left(e^{-\int_t^\tau r(s)ds}\psi(X_\tau^{t,x})\right) \tag{2.2}$$

where $\mathcal{T}_{t,T}$ is the set of all stopping times with values in $[t,T]$. Then u^* is continuous and, for any solution (X_t) of (2.1) :

$$u^*(t,X_t) = \text{ess sup}_{\tau \in \mathcal{T}_{t,T}} \mathbf{E}\left(e^{-\int_t^\tau r(s)ds}\psi(X_\tau)|\mathcal{F}_t\right) \tag{2.3}$$

Note that if, under probability $\mathbf{P}$, the discounted vector price is a martingale, equation (2.3) means that $u^*(t,X_t)$ is the "fair price" of the American option defined by ψ, at time t (cf.[2], [22], [23], [24]).

Proof of Proposition 2.2 : Using the equality

$$X_s^{t,x} = x + \int_t^s \beta(v, X_v^{t,x})dv + \int_t^s \sigma(v, X_v^{t,x})dW_v \tag{2.4}$$

it is easy to prove that :

$$\mathbf{E}\left(\sup_{t \leq s \leq T}(e^{M|X_s^{t,x}|})\right) \leq Ce^{M|x|} \tag{2.5}$$

where C depends only on T, M, and on the bounds on β, σ.

Now, observe that if (t_1,x_1),(t_2,x_2) $\in [0,T] \times \mathbf{R}^n$, with $t_1 < t_2$, then :

$$\begin{aligned} u^*(t_2,x_2) - u^*(t_1,x_1) = \\ \sup_{\tau \in \mathcal{T}_{t_2,T}} \mathbf{E}\left(e^{-\int_{t_2}^\tau r(s)ds}\psi(X_\tau^{t_2,x_2})\right) - \sup_{\tau \in \mathcal{T}_{t_2,T}} \mathbf{E}\left(e^{-\int_{t_1}^\tau r(s)ds}\psi(X_\tau^{t_1,x_1})\right) \\ + \sup_{\tau \in \mathcal{T}_{t_2,T}} \mathbf{E}\left(e^{-\int_{t_1}^\tau r(s)ds}\psi(X_\tau^{t_1,x_1})\right) - \sup_{\tau \in \mathcal{T}_{t_1,T}} \mathbf{E}\left(e^{-\int_{t_1}^\tau r(s)ds}\psi(X_\tau^{t_1,x_1})\right) \end{aligned}$$

Therefore :

$$|u^*(t_2,x_2) - u^*(t_1,x_1)| \leq$$
$$\mathbf{E}\left(\sup_{t_2\leq s\leq T} |e^{-\int_{t_2}^s r(v)dv}\psi(X_s^{t_2,x_2}) - e^{-\int_{t_1}^s r(v)dv}\psi(X_s^{t_1,x_1})|\right)$$
$$+\mathbf{E}\left(\sup_{t_1\leq s\leq t_2} |e^{-\int_{t_1}^s r(v)dv}\psi(X_s^{t_1,x_1}) - e^{-\int_{t_1}^s r(v)dv}\psi(X_{t_2}^{t_1,x_1})|\right)$$

and the continuity of u^* follows from the continuity of ψ and of the flow and from (2.5) (which, applied with $2M$ instead of M, ensures uniform integrability).

Equality (2.3) is essentially the expression of the Snell envelope in terms of the "reduite" (see [12]). It can be proved directly, by arguing that the sup in (2.2) and the essential sup in (2.3) are the same when $\mathcal{T}_{t,T}$ is replaced by the set of stopping times with respect to the filtration $(\mathcal{G}_{t,s})_{s\geq t}$ of the increments $W_s - W_t$, $s \geq t$ (see [21] for details). ⋄

3 Variational inequalities

3.1 Existence, uniqueness and regularity results

Before stating the variational inequalities related to the computation of u^*, we introduce some function spaces. Let m be a nonnegative integer and let $1 \leq p \leq \infty$, $0 < \mu < \infty$. $W^{m,p,\mu}(\mathbf{R}^n)$ will denote the space of all functions u in $L^p(\mathbf{R}^n, e^{-\mu|x|}dx)$, whose weak derivatives of all orders $\leq m$ exist and belong to $L^p(\mathbf{R}^n, e^{-\mu|x|}dx)$. We will write H_μ (resp.V_μ) instead of $W^{0,2,\mu}(\mathbf{R}^n)$ (resp. $W^{1,2,\mu}(\mathbf{R}^n)$) and the inner product on H_μ (resp.V_μ) will be denoted by $(.,.)_\mu$ (resp. $((.,.))_\mu$).

We define a bilinear form on V_μ, for each $t \in [0,T]$, in the following way :

$$
\begin{aligned}
&\forall u, v \in V_\mu \\
&a^\mu(t; u, v) \;=\; \sum_{i,j} \int_{\mathbf{R}^n} a_{i,j}(t,x) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} e^{-\mu|x|} dx \\
&\qquad - \sum_{j=1}^{n} \int_{\mathbf{R}^n} \left(a_j(t,x) + \mu \sum_{i=1}^{n} a_{i,j}(t,x) \frac{x_i}{|x|} \right) \frac{\partial u}{\partial x_j} v \, e^{-\mu|x|} dx \\
&\qquad + \int_{\mathbf{R}^n} r(t) uv \, e^{-\mu|x|} dx
\end{aligned}
$$

where : $a_j(t,x) = \beta_j(t,x) - \sum_{i=1}^{n} \frac{\partial a_{i,j}}{\partial x_i}(t,x)$, for $j = 1, \ldots, n$.

Theorem 3.1 *Under hypotheses* **(H1)**-**(H4)**, *if* $\psi \in V_\mu$, *there is one and only one function* u, *defined on* $[0,T] \times \mathbf{R}^n$ *such that :*

$$
(I) \begin{cases} u \in L^2([0,T]; V_\mu), \ \frac{\partial u}{\partial t} \in L^2([0,T]; H_\mu) \\ u \geq \psi \ a.e. \ in \ [0,T] \times \mathbf{R}^n, u(T) = \psi \ and \\ \forall v \in V_\mu \ v \geq \psi \Rightarrow -(\frac{\partial u}{\partial t}, v - u)_\mu + a^\mu(t; u, v - u) \geq 0 \end{cases}
$$

For this theorem, we refer the reader to [3], chapter 3, section 2 (see [21] for a detailed exposition of the Bensoussan-Lions methods in this setting). Note that the solution of system (I) is also the solution of the system obtained by changing μ into any number $\nu > \mu$.

Theorem 3.2 *Let* $\psi \in W^{1,p,\mu}(\mathbf{R}^n)$, *with* $p > n$, *and let* $u^*(t,x) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbf{E}\left(e^{-\int_t^\tau r(s)ds} \psi(X_\tau^{t,x}) \right)$. *Then* u^* *is the solution of system (I).*

Note that if $\psi \in W^{1,p,\mu}(\mathbf{R}^n)$, $\exp(-(\mu/p)|x|)\psi(x)$ is in $W^{1,p,0}(\mathbf{R}^n)$, therefore $\exp(-(\mu/p)|x|)\psi(x)$ has a bounded continuous version by the Sobolev imbedding theorem (see e.g. [1]) and proposition 2.2 can be applied.

The proof of Theorem 3.2 when $\psi \in W^{2,p,\mu}(\mathbf{R}^n)$, with $p > n/2 + 1$ follows from the Bensoussan-Lions methods (see [3], chapterIII,n^{o}4 or [21] for details).For $\psi \in W^{1,p,\mu}(\mathbf{R}^n)$, with $p > n$, we need the following lemma :

Lemma 3.3 *Let $0 < \mu < \nu$. If $\psi \in W^{1,p,\mu}(\mathbf{R}^n)$, with $p > n$, there exists a sequence (ψ_m), with $\psi_m \in W^{2,p,\nu}(\mathbf{R}^n)$ for all m, such that (ψ_m) converges uniformly to ψ.*

Before proving this lemma, we complete the proof of Theorem 3.2. Let ψ and (ψ_m) be as in Lemma 3.3. Denote by u the solution of system (I) and u_m the solution of system (I) when ψ is replaced by ψ_m. With obvious notations, we have : $u_m^* = u_m$ and $u_m^* \to u^*$ uniformly as $m \to \infty$. On the other hand, it follows from [3] (chapter 3, paragraphs 1.8 and 2.13) that u_m converges uniformly to u as m goes to infinity. Therefore, $u = u^*$.

Proof of Lemma 3.3 : We can assume that ψ vanishes in a neighborhood of the origin. Let :

$$\phi(x) = \begin{cases} \psi\left(\dfrac{\log|x|}{\gamma}\dfrac{x}{|x|}\right) & \text{if } |x| > 1 \\ 0 & \text{if } |x| \leq 1 \end{cases}$$

where γ is a positive number to be chosen later. Note that ϕ vanishes in a neighborhood of the unit ball and that :

$$|\nabla\phi(x)| \leq \frac{C}{\gamma}\left(\frac{1+\log|x|}{|x|}\right)\left|\nabla\psi\left(\frac{\log|x|}{\gamma}\frac{x}{|x|}\right)\right|$$

where C depends only on n.

By an easy computation, it follows that if $\gamma(p-n) > \mu$, then $\int_{\mathbf{R}^n} |\nabla\phi(x)|^p dx < \infty$, which implies that ϕ is uniformly continuous on $\mathbf{R}^n$. From now on, we assume $\gamma(p-n) > \mu$.

Now, let ρ be a nonnegative, C^∞ function, with compact support on $\mathbf{R}^n$, satisfying $\int_{\mathbf{R}^n} \rho(x)dx = 1$. For any integer m, let :

$$\begin{aligned} \rho_m(x) &= m^n \rho(mx) \\ \phi_m &= \rho_m * \phi \\ \text{and } \psi_m(x) &= \phi_m\left(e^{\gamma|x|}\frac{x}{|x|}\right) \end{aligned}$$

Clearly, ψ_m vanishes in a neighborhood of the origin for m large enough and is C^∞. Moreover, (ψ_m) converges uniformly to ψ as m goes to infinity, since

ϕ is uniformly continuous. On the other hand, if α is a multiindex of length $|\alpha| \leq 2$, we have :

$$|D^\alpha \psi_m(x)| \leq Ce^{|\alpha|\gamma|x|}(D^\alpha \rho_m * \phi)\left(e^{\gamma|x|}\frac{x}{|x|}\right)$$

where C depends only on γ and n.

Therefore, for all positive ν :

$$\left(\int_{\mathbf{R}^n} e^{-\nu|x|}\,|D^\alpha \psi_m(x)|\right)^{1/p} \leq$$

$$C\int_{\mathbf{R}^n} dy|D^\alpha \rho_m(y)|\left(\int_{\mathbf{R}^n} dxe^{-|x|(\nu-|\alpha|\gamma p)}\left|\phi\left(e^{\gamma|x|}\frac{x}{|x|}-y\right)\right|^p\right)^{1/p} \quad (3.1)$$

By the substitution $z = e^{\gamma|x|}(x/|x|)$, we get :

$$\int_{\mathbf{R}^n} dxe^{-|x|(\nu-|\alpha|\gamma p)}\left|\phi\left(e^{\gamma|x|}\frac{x}{|x|}-y\right)\right|^p =$$

$$= \frac{1}{\gamma^n}\int_{|z|\geq 1} dz\frac{(\log|z|)^{n-1}}{|z|^{n-1}}|z|^{-(1/\gamma)(\nu-|\alpha|p\gamma)}|\phi(z-y)|^p$$

$$= \frac{1}{\gamma^n}\int_{|x+y|\geq 1} dx\frac{(\log|x+y|)^{n-1}}{|x+y|^n}|x+y|^{-(1/\gamma)(\nu-|\alpha|p\gamma)}|\phi(x)|^p$$

Now, assume $\nu > 2p\gamma$ and $|y| \leq 1/3$. Then :

$$\int_{|x+y|\geq 1} dx\frac{(\log|x+y|)^{n-1}}{|x+y|^{n+(1/\gamma)(\nu-|\alpha|p\gamma)}}|\phi(x)|^p$$

$$\leq \int_{|x|\geq 2/3} dx\frac{(\log(1+|x|))^{n-1}}{(|x|-1/3)^{n+(1/\gamma)(\nu-|\alpha|p\gamma)}}|\phi(x)|^p$$

Observe that :

$$\int_{|x|\geq 1} dx\frac{(\log(|x|)^{n-1}}{|x|^{n+(1/\gamma)(\nu-|\alpha|p\gamma)}}|\phi(x)|^p$$

$$= \nu^n\int_{\mathbf{R}^n} dze^{-|z|(\nu-|\alpha|p\gamma)}|\psi(z)|^p$$

Therefore, if $\nu > 2p\gamma + \mu$ and m is large enough (so that $\text{supp}\rho_m \subset \{y \in \mathbf{R}^n|\ |y| \leq 1/3\}$), the right-hand side in (3.1) is finite for $|\alpha| \leq 2$, and, consequently, ψ_m is in $W^{2,p,\nu}(\mathbf{R}^n)$, which completes the proof. $\diamond$

Remark 3.4 One can construct a function in $W^{1,2,\mu}(\mathbf{R})$ which cannot be approximated uniformly by functions in $W^{2,2,\mu}(\mathbf{R})$. Therefore, the number ν in Lemma 3.3 has to be greater than μ.

3.2 One-dimensional models

In this section, we assume that $n = 1$ and that the drift term β and the diffusion term σ do not depend on X_t, so that (2.1) becomes :

$$dX_t = \beta(t)dt + \sigma(t)dW_t \tag{3.2}$$

where β and σ are C^1 functions on $[0, T]$, and $\inf_{t\in[0,T]} \sigma^2(t) \geq \eta > 0$.

Recall that in the classical Black-Scholes and Garman-Kohlhagen models, σ, β, r are constants, with $\beta + \sigma^2/2 = r$ in the Black-Scholes case and $\beta + \sigma^2/2$ equals the differential interest rate in the Garman-Kohlhagen case. We will also assume that ψ satisfies the following assumptions :

(H5) There exists $\mu > 0$ such that $\psi(x) \exp(-\mu|x|)$ is a bounded, continuous function and admits a bounded (weak) derivative on $\mathbf{R}$.

(H6) The function $x \mapsto \psi(\log(x))$ is convex over $(0, \infty)$

Note that (**H5**) and (**H6**) are satisfied by the put and call functions :$\psi_p(x) = (K - e^x)_+$, $\psi_c = (e^x - K)_+$.

We first observe that the flow of S.D.E. (3.2) is given by :

$$X_s^{t,x} = x + \int_t^s \beta(v)dv + \int_t^s \sigma(v)dW_v \tag{3.3}$$

Therefore, assuming (**H5**) and using the results and notations of paragraph 3.1, we have :

$$u(t, x) = u^*(t, x) = \sup_{\tau\in\mathcal{T}_{t,T}} \mathbf{E}\left(e^{-\int_t^\tau r(s)ds}\psi(X_\tau^{t,x})\right)$$

And, by (3.3) :

$$u(t, x) = \sup_{\tau\in\mathcal{T}_{t,T}} \mathbf{E}\left(e^{-\int_t^\tau r(s)ds}\psi\left(x + \int_t^\tau \beta(v)dv + \int_t^\tau \sigma(v)dW_v\right)\right) \tag{3.4}$$

With this equality the following proposition is easy to prove.

Proposition 3.5 *1. If ψ is nonincreasing (resp. nondecreasing) and satisfies (H5), then $u(t, .)$ is nonincreasing (resp. nondecreasing), for all $t \in [0, T]$.*

2. If ψ satisfies (H5) and (H6), then for all $t \in [0, T]$, $u(t, \log(.))$ is a convex function on $(0, \infty)$.

We now state the main result of this section :

Theorem 3.6 *Let ψ satisfy (H5) and (H6), then u admits partial derivatives $\partial u/\partial t$, $\partial u/\partial x$, $\partial^2 u/\partial x^2$, which are locally bounded on $[0,T) \times \mathbf{R}$. Furthermore, with the notation $A(t)u = (\sigma^2(t)/2)\partial^2 u/\partial x^2 + \beta(t)\partial u/\partial x - r(t)u$, we have :*

$$\begin{cases} \partial u/\partial t + A(t)u \leq 0 \\ (\partial u/\partial t + A(t)u)(\psi - u) = 0 \end{cases}$$

almost everywhere in $[0,T) \times \mathbf{R}$

Corollary 3.7 *If ψ satisfies (H5) and (H6), then $\partial u/\partial x$ is continuous on $[0,T) \times \mathbf{R}$*

Note that, since ψ is not assumed to be continuously differentiable, we cannot expect continuity of $\partial u/\partial x$ over $[0,T] \times \mathbf{R}$. Corollary 3.7 is an immediate consequence of Theorem 3.6 and the following classical lemma (cf.[26], chapter II, lemma 3.1).

Lemma 3.8 *Let $v(t,x)$ be a function mapping $\mathbf{R}^2$ into $\mathbf{R}$. Assume v admits partial derivatives $\partial v/\partial t$, $\partial v/\partial x$, $\partial^2 v/\partial x^2$, which are uniformly bounded on $\mathbf{R}^2$. Then, $\partial v/\partial x$ satisfies a Hölder condition in t with exponent 1/2, uniformly with respect to x.*

In order to prove Theorem 3.6, we will need the following lemma :

Lemma 3.9 *If ψ satisfies (H5) then :*

$$u(t,x) = \sup_{\theta \in \mathcal{T}_{0,1}} \mathbf{E} \left(e^{-\int_0^\theta \bar{r}(t,a)da} \psi \left(x + \int_0^\theta \bar{\beta}(t,a)da + \int_0^\theta \bar{\sigma}(t,a)dW_a \right) \right)$$

with :

- $\bar{r}(t,a) = (T-t)r(t + a(T-t))$,
- $\bar{\beta}(t,a) = (T-t)\beta(t + a(T-t))$,
- $\bar{\sigma}(t,a) = \sqrt{T-t}\,\sigma(t + a(T-t))$.

Proof : From equality (3.4), derive :

$$u(t,x) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbf{E} \left(e^{-\int_0^{\tau - t} r(t+a)da} \psi \left(x + \int_0^{\tau - t} \beta(t+a)da + \int_0^{\tau - t} \sigma(t+a)dW_{t+a} \right) \right)$$

As was recalled at the end of Section 2, we can replace $\mathcal{T}_{t,T}$ by the subset of stopping times with respect to the filtration $(\mathcal{G}_{t,s})_{s\geq t}$ of the increments $(W_s - W_t)$, $s \geq t$. Therefore, since the law of $(W_{t+a} - W_t)_{a\geq 0}$ is the same as that of $(W_a)_{a\geq 0}$, we can state :

$$\begin{aligned} u(t,x) &= \\ &= \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbf{E}\left(e^{-\int_0^\tau r(t+a)da}\psi\left(x + \int_0^\tau \beta(t+a)da + \int_0^\tau \sigma(t+a)dW_a\right)\right) \end{aligned} \tag{3.5}$$

Now, observe that $\tau \in \mathcal{T}_{0,T-t}$ if and only if τ can be expressed in the form : $\tau = \theta(T-t)$, where θ is a stopping time with values in $[0,1]$, with respect to the filtration $(\mathcal{H}_s)$, where $\mathcal{H}_s$ is the σ-field generated by all random variables $W_{a(T-t)}$, $a \leq s$. Since $(W_{a(T-t)})_{a\geq 0}$ and $(\sqrt{T-t}W_a)_{a\geq 0}$ are identically distributed, we obtain the following :

$$\begin{aligned} u(t,x) &= \\ &\sup_{\theta\in\mathcal{T}_{0,1}} \mathbf{E}\Bigg(e^{-\int_0^{\theta(T-t)} r(t+a)da} \times \\ &\times\psi\left(x + \int_0^{\theta(T-t)} \beta(t+a)da + \sqrt{T-t}\int_0^\theta \sigma(t+a(T-t))dW_a\right)\Bigg) \end{aligned} \tag{3.6}$$

and Lemma 3.9 follows from (3.6) by a change of variables in the integrals with respect to da. ◇

Proof of Theorem 3.6 : 1) Choose $\mu > 0$ so that $\psi \in V_\mu$. From Theorem 3.1, we know that $u \in L^2([0,T]; V_\mu)$ and $\partial u/\partial t \in L^2([0,T]; H_\mu)$ and that for all $v \in V_\mu$ such that $v \geq \psi$:

$$-\left(\frac{\partial u}{\partial t}, v-u\right)_\mu + a^\mu(t; u, v-u) \geq 0 \tag{3.7}$$

Let : $v(x) = u(t,x) + \phi(x)e^{\mu|x|}$, where ϕ is a nonnegative C^∞ function, with compact support. From (3.7), we derive that $-(\partial u/\partial t + A(t)u)$ defines a nonnegative measure on $(0,T)\times\mathbf{R}$. On the other hand, since by proposition 3.4, $u(t, log(.))$ is convex for all $t \in [0,T]$, $\partial^2 u/\partial x^2 - \partial u/\partial x$ must be a nonnegative measure on $(0,T) \times \mathbf{R}$. Therefore, the following inequalities hold (in the sense of measures on $(0,T)\times\mathbf{R}$) :

$$\frac{\partial u}{\partial t}+(\sigma^2(t)/2)\frac{\partial^2 u}{\partial x^2} + \beta(t)\frac{\partial u}{\partial x}-r(t)u\leq 0$$
$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial x}\geq 0$$

Hence :

$$0\leq(\eta/2)\left(\frac{\partial^2 u}{\partial x^2}-\frac{\partial u}{\partial x}\right)\leq r(t)u-\frac{\partial u}{\partial t}-(\beta(t)+(\sigma^2(t)/2))\frac{\partial u}{\partial x} \qquad (3.8)$$

Therefore $\partial^2u/\partial x^2\in L^2([0,T];H_\mu)$ and in order to prove Theorem 3.6, it suffices to show that $\partial u/\partial t$ and $\partial u/\partial x$ are locally bounded on $[0,T)\times\mathbf{R}$ (derive $(\partial u/\partial t+A(t)u)(\psi-u)=0$ by setting $v=\psi$ in (3.7)).

2)In order to prove local boundedness of $\partial u/\partial x$, fix x, y, $\in\mathbf{R}^2$ and note that, for all $\tau\in\mathcal{T}_{t,T}$

$$\begin{aligned}\left|\psi\left(x+\int_t^T\beta(a)da+\int_t^T\sigma(a)dW_a\right)-\psi\left(y+\int_t^T\beta(a)da+\int_t^T\sigma(a)dW_a\right)\right|&\leq\\ \leq|x-y|\sup_{z\in[x,y]}\left|\psi'\left(z+\int_t^T\beta(a)da+\int_t^T\sigma(a)dW_a\right)\right|&\\ \leq C|x-y|\exp\left(\mu(|x|+|y|)+\mu\int_t^\tau|\beta(a)|da+\mu|\int_t^\tau\sigma(a)dW_a|\right)&\end{aligned}$$

where C is some positive constant. Here we have used (H5). Since the stochastic process $(\exp(\mu|\int_t^s\sigma(a)dW_a|))_{s\geq t}$ is a submartingale and $\tau\leq T$, we have :

$$\begin{aligned}\mathbf{E}\left(\exp\mu|\int_t^\tau\sigma(a)dW_a|\right)&\leq\mathbf{E}\left(\exp\mu|\int_t^T\sigma(a)dW_a|\right)\\&\leq 2\exp\left(\frac{\mu^2\int_0^T\sigma^2(a)da}{2}\right)\end{aligned}$$

Therefore, using (3.4), we have :

$$|u(t,x)-u(t,y)|\leq C|x-y|e^{\mu(|x|+|y|)}$$

where C is some positive constant. It then follows that u is locally Lipschitz in x, uniformly with respect to t.Therefore, $\partial u/\partial x$ is locally bounded on $[0,T]\times\mathbf{R}$.

3)It remains to prove local boundedness for $\partial u/\partial t$. For that purpose, we will use lemma 3.9. Fix t, s in $\mathbf{R}^2$ and θ in $\mathcal{T}_{0,1}$. Then :

$$\begin{aligned}
&\Big| e^{-\int_0^\theta \bar{r}(t,a)da}\psi\left(x+\int_0^\theta \bar{\beta}(t,a)da+\int_0^\theta \bar{\sigma}(t,a)dW_a\right) \\
&\quad -e^{-\int_0^\theta \bar{r}(s,a)da}\psi\left(x+\int_0^\theta \bar{\beta}(s,a)da+\int_0^\theta \bar{\sigma}(s,a)dW_a\right)\Big| \quad \leq \\
&\quad \leq \left|e^{-\int_0^\theta \bar{r}(t,a)da}-e^{-\int_0^\theta \bar{r}(s,a)da}\right|\left|\psi\left(x+\int_0^\theta \bar{\beta}(t,a)da+\int_0^\theta \bar{\sigma}(t,a)dW_a\right)\right| \\
&\qquad +e^{-\int_0^\theta \bar{r}(s,a)da}\Big|\psi\left(x+\int_0^\theta \bar{\beta}(t,a)da+\int_0^\theta \bar{\sigma}(t,a)dW_a\right) \\
&\qquad -\psi\left(x+\int_0^\theta \bar{\beta}(s,a)da+\int_0^\theta \bar{\sigma}(s,a)dW_a\right)\Big| \\
&\quad \leq C|t-s|\exp\left(\mu(|x|+\int_0^\theta |\bar{\beta}(t,a)|da+|\int_0^\theta \bar{\sigma}(t,a)dW_a|)\right) \\
&\qquad +C\left|\int_0^\theta (\bar{\beta}(t,a)-\bar{\beta}(s,a))da+\int_0^\theta (\bar{\sigma}(t,a)-\bar{\sigma}(s,a))dW_a\right| \\
&\qquad\qquad \times\exp\mu\left[2|x|+|\int_0^\theta \bar{\sigma}(t,a)dW_a|+|\int_0^\theta \bar{\sigma}(s,a)dW_a|\right]
\end{aligned}$$

where C is some positive constant. Here, we have used the boundedness and regularity of r and (H5). Now, taking expectations of both sides and using the regularity of β, we obtain :

$$\begin{aligned}
&|u(t,x)-u(s,x)| \leq \\
&\quad C|t-s|e^{2\mu|x|}\sup_{\theta\in\mathcal{T}_{0,1}}\mathbf{E}\left(\exp\mu\left(|\int_0^\theta \bar{\sigma}(t,a)dW_a|+|\int_0^\theta \bar{\sigma}(s,a)dW_a|\right)\right) \\
&\qquad +Ce^{2\mu|x|}\sup_{\theta\in\mathcal{T}_{0,1}}\mathbf{E}\Bigg(\left|\int_0^\theta (\bar{\sigma}(t,a)-\bar{\sigma}(s,a))dW_a\right|\times \\
&\qquad\qquad \times\exp\mu\left(\left|\int_0^\theta \bar{\sigma}(t,a)dW_a\right|+\left|\int_0^\theta \bar{\sigma}(s,a)dW_a\right|\right)\Bigg)
\end{aligned}$$

To estimate the second term of the righthand side, we use the Cauchy-

Schwarz inequality and the following equality :

$$\begin{aligned}\mathbf{E}\left|\int_0^\theta (\bar\sigma(t,a)-\bar\sigma(s,a))dW_a\right|^2 &= \mathbf{E}\int_0^\theta (\bar\sigma(t,a)-\bar\sigma(s,a))^2 da \\ &\le \mathbf{E}\left(\int_0^T (\bar\sigma(t,a)-\bar\sigma(s,a))^2 da\right)\end{aligned}$$

Using the definition of σ and the fact that $t \mapsto \sqrt{T-t}$ is continuously differentiable on $[0,T)$, we obtain that $\partial u/\partial t$ is locally bounded on $[0,T)\times\mathbf{R}$, which completes the proof. $\diamond$

Remark 3.10 Regularity results similar to corollary 3.6 can be obtained without using variational inequalities (see [32], [13], remark 5.7).

We conclude this section with some remarks concerning the American put price in the Black-Scholes model. Therefore, we assume σ, β, r constant and $\beta = r - \sigma^2/2$, with $r > 0$ and we set : $\psi(x) = (K - e^x)_+$.

Remark 3.11 Let $P(t,x)$ be the American put price at time t when the current stock price is x : $P(t,x) = u(t, \log(x))$. It follows from proposition 3.5 that for all $t \in [0,T]$, $P(t,x)$ is a convex, nonincreasing function of x. From equality (3.4), we derive :

$$P(t,x) = \sup_{\tau\in\mathcal{T}_{t,T}} \mathbf{E}\left(Ke^{-r(\tau-t)} - xe^{\sigma(W_\tau-W_t)-(\sigma^2/2)(\tau-t)}\right)_+$$

It follows that $\lim_{x\to 0} P(t,x) = K$ and that $P(t,x) \ge P_0(t,x)$, where

$$P_0(t,x) = \mathbf{E}\left(Ke^{-r(T-t)} - xe^{\sigma(W_T-W_t)-(\sigma^2/2)(T-t)}\right)_+$$

is the price of the European put option. Therefore, $P(t,x) > 0$ if $(t,x) \in [0,T) \times [0,\infty)$. The above properties yield that, for each $t \in [0,T)$, there exists a real number $s(t) \ge 0$ such that :

$$\begin{aligned} 0 \le x \le s(t) &\Rightarrow P(t,x) = (K-x)_+ \\ x > s(t) &\Rightarrow P(t,x) > (K-x)_+ \end{aligned}$$

In the financial literature, $s(t)$ is refered to as the “critical stock price” since the option should be exercized as soon as the stock price becomes smaller than $s(t)$. A study of the function $t \mapsto s(t)$ (for the case of a warrant or a call option rather than a put option) can be found in [32] (see also [15]). Using equality (3.5), it is easy to show that $P(t,x)$ is a non increasing function of t. Therefore, $s(t)$ must be a nondecreasing function of t.

Remark 3.12 We now observe that for a put option, there is only one stopping time τ maximizing $\mathbf{E}(Ke^{-r(\tau-t)}-xe^{\sigma(W_\tau-W_t)-(\sigma^2/2)(\tau-t)})_+$. Indeed, we know (see e.g. [11]) that the smallest optimal stopping time is given by :

$$\tau = \inf\{\theta \in [t,T] | X_\theta^{t,x} \leq \log(s(\theta))\}$$

It follows from well-known properties of brownian motion that :

$$\inf\{\theta \geq \tau | X_\theta^{t,x} < X_\tau^{t,x}\} = \tau$$

almost surely. Therefore, on the set $\{\tau < T\}$, there exists a sequence (θ_n) decreasing to τ, which satisfies : $\forall n \geq 0\ X_{\theta_n} < X_\tau$. Therefore $X_{\theta_n} < \log(s(\tau)) \leq \log(s(\theta_n))$. But, on the open set $\mathcal{O} = \{(\theta,x) | x < \log s(\theta)\}$ we have :$u(t,x) = K - e^x$. Therefore :

$$\frac{\partial u}{\partial t} + A(t)u = -\frac{\sigma^2}{2}e^x - (r - \frac{\sigma^2}{2})e^x - ru(t,x) < 0$$

Hence :

$$\forall \theta > \tau \quad \int_\tau^\theta (\frac{\partial u}{\partial t} + A(t)u)(s, X_s)ds < 0$$

since $(X_s^{t,x})_{s\geq t}$ spends some time in $\mathcal{O}$ on any interval $[\tau,\theta]$. It follows that τ coincides with the largest optimal stopping time (cf.[11] or [28]).

4 Localization and discretization of variational inequalities

This section deals with numerical computation of solutions for variational inequalities involved in the pricing of American options.

We assume that **(H1)**, **(H2)**, **(H3)**, **(H4)**, **(H5)** are satisfied and moreover we suppose that the differential operator involved in the variational inequality does not depend on time $a^\mu(t,.,.) = a^\mu(.,.)$. Under these assumptions, we have shown in previous sections that the price of an American option is the unique solution $u(t,x)$ of the following system of inequalities :

$$\begin{cases} \text{Find } u \in L^2\left([0,T]; V_\mu\right) \text{ such that :} \\ \qquad \dfrac{\partial u}{\partial t} \in L^2\left([0,T], H_\mu\right) \\ \qquad u(T) = \psi \\ \qquad u \geq \psi \quad p.p. \text{ in } [0,T] \times \mathbf{R}^n \\ \qquad \forall v \in V_\mu \text{ if } v \geq \psi \text{ then} \\ \qquad -\left(\dfrac{\partial u}{\partial t}, v-u\right)_\mu + a^\mu(u, v-u) \geq 0 \end{cases} \tag{4.9}$$

In order to compute u we proceed in two steps :

- we localize the variational inequality in an open bounded set of $\mathbf{R}^n$.
- we compute the solution to the localized problem using finite difference methods.

The second part is well known. The reader is refered to [18] for background and proofs. We will only show how to use these results in our framework.

4.1 Localization

Let $(X_s^{t,x})_{s\geq t}$ be the solution of the stochastic differential equation :

$$\begin{cases} dX_s &= \beta(X_s)ds + \sigma(X_s)dW_s \quad \text{if } s \geq t \\ X_t &= x. \end{cases}$$

We want to compute :

$$u^*(t,x) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbf{E}\left(e^{-\int_t^\tau r(s)ds}\psi\left(X_\tau^{t,x}\right)\right)$$

Let $T_K^{t,x} = \inf\{s > t,\ |X_s^{t,x}| > K\}$ and :

$$u_K^*(t,x) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbf{E}\left\{\exp\left(-\int_t^{\tau \wedge T_K^{t,x}} r(s)ds\right)\psi\left(X_{\tau \wedge T_K^{t,x}}^{t,x}\right)\right\}$$

Let $B_R = \{x \in \mathbf{R}^n,\ |x| \leq R\}$. The following proposition compares u^* with u_K^*.

Proposition 4.1 *Under assumptions* $(\mathbf{H1}), (\mathbf{H2}), (\mathbf{H3}), (\mathbf{H4}), (\mathbf{H5})$, u_K^* *converges to* u *(uniformly on compact sets) as* K *goes to infinity. For all* $R > 0$ *:*

$$\lim_{K \to \infty} \sup_{(t,x) \in [0,T] \times B_R} |u^*(t,x) - u_K^*(t,x)| = 0$$

Proof : Clearly :

$$|u^*(t,x) - u_K^*(t,x)| \leq 2\mathbf{E}\left(\sup_{[0,T]}\left|\psi\left(X_s^{t,x}\right)\right| \mathbf{1}_{\{T_K^{t,x} < T\}}\right)$$

So using assumption $(\mathbf{H5})$ we obtain :

$$|u^*(t,x) - u_K^*(t,x)| \leq 2M\sqrt{\mathbf{E}\left(e^{2M\sup_{[t,T]}|X_s^{t,x}|}\right)}\sqrt{\mathbf{P}\left(T_K^{t,x} < T\right)}$$

where M is a positive constant. Using inequality (2.5) we have :

$$\sup_{(t,x) \in [0,T] \times B_R} \mathbf{E}\left(e^{2M\sup_{[t,T]}|X_s^{t,x}|}\right) \leq C < +\infty$$

where C is a constant depending only on T, R and the bounds on β and σ. In order to complete the proof, we need to prove that :

$$\lim_{K \to +\infty} \sup_{[0,T] \times B_R} \mathbf{P}\left\{T_K^{t,x} < T\right\} = 0$$

Now since :

$$\left\{T_K^{t,x} < T\right\} = \left\{\sup_{[t,T]}\left|X_s^{t,x}\right| > K\right\}$$

By using the stochastic differential equation defining $X^{t,x}$ and bounds $\tilde{\beta}$ and $\tilde{\sigma}$ on β and σ we get :

$$\left\{\sup_{[t,T]}\left|X_s^{t,x}\right| > K\right\} \subset \bigcup_{1 \leq i \leq n}\left\{\sup_{s \in [t,T]}\left|\left(\int_t^s \sigma(X_s)dW_s\right)_i\right| > \frac{1}{n}\left(K - R - \tilde{b}T\right)\right\}.$$

Therefore :

$$\mathbf{P}\left\{\sup_{[t,T]}\left|X_s^{t,x}\right| > K\right\} \leq n\mathbf{P}\left\{\sup_{s\in[0,\tilde{\sigma}^2T]}|\beta_s| > \frac{1}{n}\left(K - R - \tilde{b}T\right)\right\}$$

and the desired result follows easily. ◇

Remark 4.2 We can obtain the same kind of results if u_K^* is replaced by :

$$\tilde{u}_K^*(t,x) = \sup_{\tau\in\mathcal{T}_{t,T}} \mathbf{E}\left\{e^{-\int_t^\tau r(s)ds}\psi\left(X_{\tau\wedge T_K^{t,x}}^{t,x}\right)\mathbf{1}_{\{\tau<T_K^{t,x}\}}\right\}$$

or by :

$$\tilde{\tilde{u}}_K^*(t,x) = \sup_{\tau\in\mathcal{T}_{t,T}} \mathbf{E}\left\{e^{-\int_t^\tau r(s)ds}\psi\left(\hat{X}_\tau^{t,x}\right)\right\}$$

where $\hat{X}$ is the solution of the stochastic differential equation reflected in point K and $-K$ in the sense of Skorohod.

We will suppose now that a number K is given such that u^* and u_K^* are close enough for practical purposes.

4.2 Discretization

Let us denote :

$$\begin{aligned}
\Omega_K &= \{x\in\mathbf{R}^n,\ |x|<K\}\\
\partial\Omega_K &= \{x\in\mathbf{R}^n,\ |x|=K\}\\
H_K &= L^2(\Omega_K)\\
V_K &= \left\{f\in H_K,\ \frac{\partial f}{\partial x}\in H_K\right\}
\end{aligned}$$

Let :

$$\begin{aligned}
a_K(u,v) &= -\int_{\Omega_K} Auvdx\\
&= \sum_{1\leq i,j\leq n}\int_{\Omega_K} a_{ij}(x)\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}dx - \int_{\Omega_K} a_j(x)\frac{\partial u}{\partial x}vdx + \int_{\Omega_K} ruvdx
\end{aligned}$$

where :

$$a_j(x) = \beta_j(x) - \sum_{1\leq i\leq n}\frac{\partial a_{ij}}{\partial x_j}(x)$$

Under assumptions (**H1**), (**H2**), (**H3**), (**H4**), (**H5**), $u_K^*(t,x)$ is the unique solution of the following variational inequality :

$$\begin{cases} u \in L^2([0,T];V_K) \\ \dfrac{\partial u}{\partial t} \in L^2([0,T];H_K) \\ u \geq \psi \text{ a.e. in } [0,T] \times \Omega_K \\ \forall v \in V_\Omega \text{ if } v \geq \psi \text{ then } -\left(\dfrac{\partial u}{\partial t}, v-u\right)_K + a_K\left(u, v-u\right) \geq 0 \\ u = \psi \text{ if } x \in \partial\Omega_K \\ u(T) = \psi \end{cases}$$

This result can be obtained by using theorems of [3] (chapter 3, section 4) and the method of paragraph 3.

Let us rewrite the previous inequality as an homogeneous inequality. Let $\tilde{\psi}$ be a smooth function ($C^2(\mathbf{R}^n, \mathbf{R})$ for instance) such that $\psi = \tilde{\psi}$ on an open neighbourhood of $\partial\Omega_K$. Such a function will be easy to find in all practical cases of option pricing. Let u' be $u - \tilde{\psi}$, then u' is solution of :

$$\begin{cases} u' \in L^2(V_K) \\ \dfrac{\partial u'}{\partial t} \in L^2([0,T];H_K) \\ u' \geq \psi - \tilde{\psi} \text{ p.p. in } [0,T] \times \Omega_K \\ \forall v \in V_\Omega \text{ if } v \geq \psi - \tilde{\psi} \text{ then } -\left(\dfrac{\partial u'}{\partial t}, v-u'\right)_K + a_K\left(u', v-u'\right) \geq 0 \\ u = 0 \text{ if } x \in \partial\Omega_K \\ u'(T) = \psi - \tilde{\psi} \end{cases} \tag{4.10}$$

Discretization of inequation 4.10 can be obtained using results of [18]. Let h be $(h_1, h_2, \cdots, h_n)$, with $h_i > 0$. h is supposed to be small. Let Ω_h be the approximation of Ω_K obtained by :

- $R_h = \{M \in \mathbf{R}^n,\ M = (m_1 h_1, \cdots, m_n h_n),\ m_i \in \mathbf{Z}\}$
- if $M \in R_h$ then let :

$$\omega_h^0(M) = \prod_{i=1}^{n}](m_i - 1/2)\, h_i, (m_i + 1/2)\, h_i[$$

- $\Omega_h = \left\{M \in R_h,\ \forall i = 1, \cdots, n \quad \omega_h^0\left(M \pm \frac{h_i}{2}\right) \subset \Omega_K\right\}$.

Then let $H_h = \text{Vect}\left(\theta_h^M\right)_{M\in\Omega_h}$, where, $\theta_h^M = \mathbf{1}_{\omega_h^0}(M)$. We approximate the partial derivative $\dfrac{\partial}{\partial x_i}$ by a finite difference operator :

$$(\delta_i\psi)(x) = \frac{1}{h_i}\left[\psi\left(x+\frac{h_i}{2}e_i\right) - \psi\left(x-\frac{h_i}{2}e_i\right)\right].$$

If $u_h, v_h \in V_h$ we define :

- $a_h\left(u_h, v_h\right)$ by

$$\sum_{1\le i,j\le n}\int_{\Omega_K} a_{ij}(x)\delta_i u_h \delta_j v_h dx - \int_{\Omega_K} a_j(x) v_h \delta_i u_h dx + \int_{\Omega_K} r u_h v_h dx,$$

- $(.,.)_h$ by $(u_h, v_h)_h = \int_{\Omega_K} u_h v_h dx$,
- A_h by $a_h(u_h, v_h) = -(A_h u_h, v_h)_h$ for all $u_h, v_h \in H_h$.

Application :
Let $n = 1$ and :

$$A = \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2} + \alpha\frac{\partial}{\partial x} - r,$$

then :

$$\left(A_h u_h\right)_j = \frac{1}{2}\frac{\sigma^2}{h^2}\left(u_h^{j+1} - 2u_h^j + u_h^{j-1}\right) + \frac{\alpha}{2h}\left(u_h^{j+1} - u_h^{j-1}\right) - r u_h^i.$$

Let $\bar{\psi}$ be $\psi - \tilde{\psi}$, $\mathcal{K}$ be $\{v \in V_K,\ u \ge \bar{\psi}\}$ and :

$$\psi_h(M) = \frac{1}{\prod_{i=1}^n h_i}\int_{\omega_0^h(M)}\psi(x)dx.$$

$f = -A\tilde{\psi}$ is approximated by $f_h = -A_h\tilde{\psi}_h$, $\mathcal{K}$ by $\mathcal{K}_h$ where :

$$\mathcal{K}_h = \left\{(u_M)_{M\in\Omega_h},\ u_M \ge \bar{\psi}_h(M)\right\},$$

and the initial value by :

$$u_{T,h}(M) = \bar{\psi}_h(M)$$

The derivative of v_h with respect to time is approximated by,

$$\bar{\delta} v_{h,k}(t) = \frac{1}{k}\left(v_{h,k}(t) - v_{h,k}(t-k)\right),$$

where k is the time step. Let N be an integer and $k = T/N$, the discretized inequality can be stated as :

$$(\mathbf{D}) \quad \left\{ \begin{array}{l} \text{find } (u_h^i)_{1 \le i \le N} \text{ such that :} \\ u_h^N(M) = \tilde{\psi}_h(M) \\ \text{if } 1 \le i < N, \; u_h^i \text{ is obtained from } u_h^{i+1} \text{ by solving :} \\ \quad \left\{ \begin{array}{l} \quad u_h^i \in K_h \\ \forall v_h \in K_h \quad \left(u_h^{i+1} - u_h^i + kA_h u_h^{i+\theta} - k f_h^i, v_h - u_h^i \right)_h \le 0 \\ u_h^{i+\theta} = u_h^{i+1} + \theta \left(u_h^i - u_h^{i+1} \right) \end{array} \right. \end{array} \right.$$

The approximating solution is now :

$$u_{h,k}(t,x) = \sum_{1 \le i \le N} u_h^i(x) \mathbf{1}_{](i-1)k, ik]}(t)$$

It can be proved using theorem 3.1 of [18] that $u_{h,k}$ converges to u^*. More precisely :

Theorem 4.3 *Assume that* $\theta \in [0,1]$ *and that* h *and* k *go to* 0*, and* if $\theta < 1$ that k/h^2 goes to 0, then *:*

- $\lim u_{h,k} = u^*$ *in* $L^2\left([0,T], L^2(\Omega_K)\right)$
- *for all* $1 \le i \le n$*,* $\lim \delta_i u_{h,k} = \dfrac{\partial u^*}{\partial x_i}$ *in* $L^2\left([0,T], L^2(\Omega_K)\right)$

Remark 4.4

- Let $u_h^i \leftarrow u_h^i + \tilde{\psi}_h$, we can restate $(\mathbf{D})$ as :

  $$\left\{ \begin{array}{l} u_h^N(M) = \psi_h(M), \\ \text{Find } u_h^i \ge \psi_h \text{ such that :} \\ \quad \forall v_h \ge \psi_h \quad \left(u_h^{i+1} - u_h^i + kA_h u_h^{i+\theta}, v_h - u_h^i \right)_h \le 0 \\ \quad u_h^{i+\theta} = u_h^{i+1} + \theta \left(u_h^i - u_h^{i+1} \right) \end{array} \right. .$$

  Clearly regularity assumptions on $\tilde{\psi}$ prove that the convergence results for $u_{h,k} + \tilde{\psi}_h$ are consequences of the ones for $u_{h,k}$.

- It seems that unconditional stablity has been proved only if $\theta = 1$, that is to say for the fully implicit scheme. Recall that in the case of equations the schemes obtained by letting $\theta = 1/2$ (Crank-Nicholson scheme) or $\theta > 1/2$ are unconditionaly stable.

- The same kind of convergence results can be obtained in the case of Neumann boundary conditions on Ω_K.

5 The Brennan-Schwartz algorithm

5.1 Variational inequalities in finite dimensional spaces

This subsection is meant to prepare the way for a discussion of the Brennan-Schwartz algorithm. The inner product of two vectors u, $v \in \mathbf{R}^n$ will be denoted by (u, v) and we will write $u \geq v$ if :

$$\forall i \in \{1, \cdots, n\} \quad u_i \geq v_i$$

and $|u|^2$ for (u, u).

Proposition 5.1 *Let A be an $n \times n$ matrix and u, θ, $\phi \in \mathbf{R}^n$. The following two systems are equivalent :*

$$(I) \qquad \begin{cases} Au \geq \theta \\ u \geq \phi \\ (Au - \theta, \phi - u) = 0 \end{cases}$$

$$(I') \qquad \begin{cases} u \geq \phi \\ \forall v \geq \phi, \quad (Au - \theta, v - u) \geq 0 \end{cases}$$

This proposition, which expresses the linear complementarity problem as a variational inequality is well known (see e.g. [25] chapter 1 for a proof).

The following result is due to Samuelson, Thrall and Weisler [30].

Theorem 5.2 *The linear complementary problem (I') has a unique solution for all vectors θ, ϕ if and only if A has positive principal minors.*

Note that this property is satisfied if A is coercive in the following sense :

$$\exists C > 0 \; \forall x \in \mathbf{R}^n \quad (Ax, x) \geq C|x|^2.$$

For a direct proof of Theorem 5.2 in the coercive case we refer the reader to [18] chapter 1, paragraph 3.1 or [25] chapter 1, paragraph 4.

The following characterization of the solution, which can be found in [10], has some connection with the notion of "reduite" in potential theory :

Proposition 5.3 *Assume the nondiagonal coefficients of a coercive matrix are non positive. Then the solution of (I) is the smallest vector u satisfying :*

$$Au \geq \theta \quad u \geq \phi$$

In the next two propositions, we assume that A is a tridiagonal matrix, which we denote in the following way :

$$(T) \qquad A = \begin{pmatrix} b_1 & c_1 & 0 & \cdots & 0 \\ a_2 & b_2 & c_2 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & a_{n-1} & b_{n-1} & c_{n-1} \\ 0 & \cdots & 0 & a_n & b_n \end{pmatrix}$$

This type of matrix appears in finite difference schemes of approximations in one dimension.

Proposition 5.4 *If A is a matrix of type (T) with $c_i = 0$ for all $i \in \{1, \cdots, n-1\}$, and $b_i > 0$ for all $i \in \{1, \cdots, n\}$, then, for all θ, ϕ in $\mathbf{R}^n$, the unique solution of system (I) can be computed by solving the following recursive relations :*

$$\begin{aligned} u_1 &= \tfrac{\theta_1}{b_1} \vee \phi_1 \\ u_j &= \tfrac{1}{b_j}(\theta_j - a_j u_{j-1}) \vee \phi_j \quad \textit{for } 2 \le j \le n \end{aligned}$$

This proposition is an immediate consequence of the following equalities :

$$\begin{aligned} (Au)_1 &= b_1 u_1 \\ (Au)_j &= b_j u_j + a_j u_{j-1} \text{ for } j \in \{2, \cdots, n\} \end{aligned}$$

Note that, here, A need not be coercive.

Now, given any coercive matrix of type (T) consider the matrix $\tilde{A}$ defined by :

$$\tilde{A} = \begin{pmatrix} \tilde{b}_1 & 0 & 0 & \cdots & 0 \\ a_2 & \tilde{b}_2 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_n & \tilde{b}_n \end{pmatrix}$$

where $\tilde{b}_n = b_n$ and $\tilde{b}_{i-1} = b_{i-1} - (c_{i-1}/\tilde{b}_i)a_i$, for all $i \in \{2, \cdots, n\}$.

Lemma 5.5 *If A is coercive $\tilde{b}_i > 0$ for all $i \in \{1, \cdots, n\}$.*

Proof : Note that, for $1 \le i \le n$, $\prod_{j=i}^{n} \tilde{b}_j$ is the determinant of the matrix derived from A by erasing its $(i-1)$ first lines and columns. This matrix, being a submatrix of A is coercive and therefore, its determinant is positive. Hence, $\tilde{b}_i > 0$ for all $i \in \{1, \cdots, n\}$. $\diamond$

If $\theta \in \mathbf{R}^n$, denote $\tilde{\theta}$ the vector with coordinates satisfying $\tilde{\theta}_n = \theta_n$ and $\tilde{\theta}_{i-1} = \theta_{i-1} - c_{i-1}\tilde{\theta}_i/\tilde{b}_i$, whenever $i \in \{1, \cdots, n\}$. Clearly the two systems $Ax = \theta$ and $\tilde{A}x = \tilde{\theta}$ are equivalent. The following proposition generalizes this property to some variational inequalities.

Proposition 5.6 *Assume A is a coercive matrix of type (T) with $c_j \leq 0$ for all $j \in \{1, \cdots, n-1\}$. If the solution u of system (I) satisfies $u_i = \phi_i$ whenever $1 \leq i \leq k$, and $u_i > \phi_i$ whenever $k < i$, for some $k \in \{1, \cdots, n\}$, then u is also a solution to system $(\tilde{I})$:*

$$(\tilde{I}) \left\{ \begin{array}{l} \tilde{A}u \geq \tilde{\theta} \\ u \geq \phi \\ (\tilde{A}u - \tilde{\theta}, \phi - u) = 0 \end{array} \right.$$

Proof : Since $c_{i-1} \leq 0$ and $\tilde{b}_i > 0$ (by lemma 5.5), we have $-c_{i-1}/\tilde{b}_i \geq 0$ for all $i \in \{2, \cdots, n\}$. It follows that, for all $i \in \{1, \cdots, n-1\}$, there exist non negative numbers $\lambda_{i,i+1}, \cdots, \lambda_{i,n}$ such that :

$$\tilde{\theta}_i = \theta_i + \lambda_{i,i+1}\theta_{i+1} + \cdots + \lambda_{i,n}\theta_n$$

and $\left(\tilde{A}u\right)_i = (Au)_i + \lambda_{i,i+1}(Au)_{i+1} + \cdots + \lambda_{i,n}(Au)_n$. Therefore if u solves system (I), we have $\left(\tilde{A}u\right)_i \geq \tilde{\theta}_i$. Moreover, under the assumptions of Proposition 5.6, equality holds whenever $i > k$, since $(Au)_j$ is then equal to θ_j for all $j \geq i$. Hence u solves $\left(\tilde{I}\right)$. $\diamond$

Remark 5.7 From Propositions 5.4 and 5.6, a very simple algorithm can be derived for solving (I). The difficulty in practice will come from the verification of the assumption on u in Proposition 5.6. This assumption is essential as the following example proves. Let $n = 3$,

$$A = \left[\begin{array}{ccc} 1 & -1 & 0 \\ -\epsilon & 1 & 0 \\ 0 & 0 & 1 \end{array} \right]$$

with $0 < \epsilon < 1$,

$$\theta = \left(\begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right) \quad \phi = \left(\begin{array}{c} 1 \\ 2 \\ 0 \end{array} \right)$$

The unique solution of system $(\tilde{I})$ is given by :

$$\tilde{u} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

and satisfies $\tilde{u}_1 = \phi_1$, $\tilde{u}_2 = \phi_2$ and $\tilde{u}_3 > \phi_3$ but does not solve system (I).

Remark 5.8 Similarly, a very simple algorithm can be derived if A is coercive of type (T), with $a_j \leq 0$ and if the solution u satisfies $u_i > \phi_i$ whenever $1 \leq i \leq k$ and $u_i = \phi_i$ whenever $i > k$. Indeed we only need to renumber the indices in decreasing order and apply the above results.

5.2 The validity of the Brennan-Schwartz algorithm

Recall that for classical models, the calculation of the price of an American option reduces to solving the following system, in which we have made the change of variable $t \to (T-t)$.

$$\begin{cases} u \geq \psi \\ \dfrac{\partial u}{\partial t} - \dfrac{\sigma^2}{2}\dfrac{\partial^2 u}{\partial x^2} - \beta\dfrac{\partial u}{\partial x} + ru \geq 0 \\ \left(\dfrac{\partial u}{\partial t} - \dfrac{\sigma^2}{2}\dfrac{\partial^2 u}{\partial x^2} - \beta\dfrac{\partial u}{\partial x} + ru\right)(\psi - u) = 0 \\ u(0) = \psi \end{cases}$$

with $\psi(x) = (K - e^x)_+$ for a put option and $\psi(x) = (e^x - K)_+$ for a call option. For simplicity we assume $\psi(x) = (K - e^x)_+$ (see remark 5.13 below about call options). The implicit discretization scheme leads to approximate the vector :

$$(u\,(i\Delta t, j\Delta x))_{0 \leq i \leq N,\; n_1 \leq j \leq n_2}$$

by the vector :

$$\left(\bar{u}^i_j\right)_{0 \leq i \leq N, n_1 \leq j \leq n_2}$$

obtained by solving the following set of equations :

$$(D) \quad \begin{cases} \bar{u}^0_j = \psi(j\Delta x) \quad \text{For } n_1 + 1 \leq j \leq n_2 + 1. \\ \text{for } 1 \leq i \leq N \text{ and } n_1 + 1 \leq j \leq n_2 + 1 : \\ a\bar{u}^i_{j-1} + b\bar{u}^i_j + c\bar{u}^i_{j+1} \geq \bar{u}^{i-1}_j \\ \bar{u}^i_j \geq \psi\,(j\Delta x) \\ \left(a\bar{u}^i_{j-1} + b\bar{u}^i_j + c\bar{u}^i_{j+1} - \bar{u}^{i-1}_j\right)\left(\bar{u}^i_j - \psi\,(j\Delta x)\right) = 0 \\ \bar{u}^i_{n_1} = \psi\,(n_1\Delta x) \qquad \bar{u}^i_{n_2} = 0 \end{cases}$$

where :

$$\begin{cases} a &= -\frac{\Delta t \sigma^2}{2(\Delta x)^2} + \beta \frac{\Delta t}{2\Delta x} \\ b &= 1 + \frac{\Delta t \sigma^2}{(\Delta x)^2} + r\Delta t \\ c &= -\frac{\Delta t \sigma^2}{2(\Delta x)^2} - \beta \frac{\Delta t}{2\Delta x} \end{cases}$$

and $N\Delta t = T$, $[n_1\Delta x, n_2\Delta x]$ beeing a sufficiently large interval about $\log(K)$.

Note that the last line of system (D) corresponds to Dirichlet boundary conditions, whereas Brennan and Schwartz use mixed conditions, Dirichlet in $n_1\Delta x$ and Neumann in $n_2\Delta x$ (see remark 5.12 below). The Brennan-Schwartz algorithm method [5], applied to system (D) goes as follows :

1. From the first line of (D) get $(\bar{u}_j^0)_{n_1 \le j \le n_2}$.

2. Derive $\left(\bar{u}_j^i\right)_{n_1 \le j \le n_2}$ from $\left(\bar{u}_j^{i-1}\right)_{n_1 \le j \le n_2}$ through the following equations :

$$\begin{aligned} \bar{u}_{n_1}^i &= \psi(n_1\Delta x) \\ \bar{u}_j^i &= \left(\frac{1}{\tilde{b}_j}\left(\tilde{u}_j^{i-1} - a\bar{u}_{j-1}^{i-1}\right)\right) \vee \psi(j\Delta x) \quad n_1 + 1 \le j \le n_2 - 1 \\ \bar{u}_{n_2}^i &= \psi(n_2\Delta x) = 0 \end{aligned}$$

in which the following notations are used :

$$\tilde{b}_{n_2-1} = b \quad \text{and} \quad \tilde{b}_{j-1} = b_{j-1} - \frac{c}{\tilde{b}_j}a$$

and :

$$\tilde{u}_{n_2-1}^{i-1} = \bar{u}_{n_2-1}^{i-1}, \quad \tilde{u}_{j-1}^{i-1} = \bar{u}_{j-1}^{i-1} - \frac{c}{\tilde{b}_j}\tilde{u}_j^{i-1}.$$

for all $j \in \{n_1 + 2, \cdots, n_2 - 1\}$

We will give a justification of this procedure, using the results of paragraph 5.1 and special properties of ψ. We first restate system (D) as a set of variational inequalities.

Let $n = n_2 - n_1 - 1$ and u^i be the vector in $\mathbf{R}^n$, with components $u_j^i = \bar{u}_{j+n_1}^i$ for $1 \le j \le n$. Given any vector $v \in \mathbf{R}^n$, denote $\dot{v}$ the vector with components $\dot{v}_1 = v_1 - a\psi(n_1\Delta x)$ and $\dot{v}_j = v_j$ whenever $2 \le j \le n$, and let A be the following $n \times n$ matrix :

$$A = \begin{pmatrix} b & c & 0 & \cdots & 0 \\ a & b & c & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & c \\ 0 & \cdots & 0 & a & b \end{pmatrix}$$

System (D) is equivalent to the following set of inequalities :

$$\begin{cases} \bar{u}^i_{n_1} = \psi\left(n_1 \Delta x\right) \qquad \bar{u}^i_{n_2} = \psi\left(n_2 \Delta x\right) \quad 0 \leq i \leq N \\ (D') \quad \begin{cases} u^0 = \phi \\ \text{And for } 1 \leq i \leq N : \\ \qquad Au^i \geq \dot{u}^{i-1} \qquad u^i \geq \phi \\ \qquad (Au^i - \dot{u}^{i-1}, u^i - \phi) = 0 \end{cases} \end{cases}$$

where ϕ is the vector with components :

$$\phi_j = \psi\left((n_1 + j)\Delta x\right), \ 1 \leq i \leq n.$$

It is now easy to see that in order to justify the Brennan-Schwartz algorithm, it suffices to prove that the assumptions of Proposition 5.6 are satisfied for each of the N variational inequalities in (D'). To this end we will assume that the following condition is satisfied :

$$\beta < \frac{\sigma^2}{\Delta x} \tag{5.1}$$

Note that 5.1 is satisfied when Δx is small enough and that it implies $a < 0$.

Lemma 5.9 *If (5.1) is satisfied, A is a coercive matrix and satisfies the folowing property :*

$$Ax \geq 0 \text{ and } x \neq 0 \quad \Longrightarrow \quad \forall i \in \{1, \cdots, n\} \ x_i \geq 0$$

Proof : Using the fact that a and c are negative, we get :

$$\begin{aligned} (Ax, x) &= \sum_{j=2}^{n} a x_{j-1} x_j + \sum_{j=1}^{n} b x_j^2 + \sum_{j=1}^{n-1} c x_j x_{j+1} \\ &\geq \frac{a}{2} \sum_{j=2}^{n} \left(x_{j-1}^2 + x_j^2\right) + b \sum_{j=1}^{n} x_j^2 + \frac{c}{2} \sum_{j=1}^{n-1} \left(x_j^2 + x_{j+1}^2\right) \\ &\geq (a + b + c) \sum_{i=1}^{n} x_j^2 = (1 + r\Delta t) \sum_{i=1}^{n} x_j^2 \end{aligned}$$

This proves coercivity. $\diamond$

Now suppose $Ax \geq 0$. Then, with the notations of paragraph 5.1, $\tilde{A}x \geq 0$. Indeed the numbers $\lambda_{i,j}$ introduced in the proof of Proposition 5.6 are positive since $a < 0$ and $c < 0$. If, in addition, $x \neq 0$, it follows that $(\tilde{A}x)_1 > 0$ and using the negativity of a again, we conclude that $x_j > 0$ for all $j \in \{1, \cdots, n\}$.

Lemma 5.10 *Let :*

$$\begin{aligned}\Gamma(\psi) &= \{j \in \mathbf{Z},\ \psi(j\Delta x) > 0 \text{ and} \\ &\qquad a\psi((j-1)\Delta x) + b\psi(j\Delta x) + c\psi((j+1)\Delta x) \geq \psi(j\Delta x)\},\end{aligned}$$

The set $\Gamma(\psi)$ has the following property :

$$j_0 \in \Gamma(\psi) \quad \Longrightarrow \quad \forall j \leq j_0 \quad j \in \Gamma(\psi).$$

Proof : Let $\bar{\psi}(x) = K - e^x$. Using $c < 0$, it is easy to prove $\Gamma(\psi) \subset \Gamma(\bar{\psi})$. On the other hand, we have :

$$\{j \in \mathbf{Z},\ \psi((j+1)\Delta x) \neq 0\} \subset \Gamma(\psi).$$

Now, for any integer j :

$$\begin{aligned} a\bar{\psi}((j-1)\Delta x) + b\bar{\psi}(j\Delta x) + c\bar{\psi}((j+1)\Delta x) - \bar{\psi}(j\Delta x) &= \\ (a+b+c-1)K - e^{j\Delta x}\left(ae^{-\Delta x} + b + ce^{\Delta x} - 1\right) \end{aligned}$$

In the latter equation, the righthand side is a monotonous function of j, whose limit as j goes to $-\infty$ is : $K(a+b+c-1) = Krh \geq 0$. It follows that :

$$j_0 \in \Gamma(\bar{\psi}) \quad \Longrightarrow \quad \forall j \leq j_0 \quad j \in \Gamma(\bar{\psi}).$$

And Lemma 5.10 follows from the inclusion relations between $\Gamma(\psi)$ and $\Gamma(\bar{\psi})$. $\diamond$

The following proposition together with the results of paragraph 5.1 completes the justification of the Brennan-Schwartz algorithm for solving system (D).

Proposition 5.11 *Let $u^0, u^1, \cdots, u^N$ be the solution of (D'). The following properties are satisfied :*

1. $\forall i \in \{1, \cdots, N\} \quad u^i \geq u^{i-1}$

2. *For each i, there exists an index k_i such that :*
   - $u^i_j = \phi_j$ *whenever* $1 \leq j \leq k_i$,
   - *and* $u^i_j > \phi_j$ *whenever* $j > k_i$.

Proof : The first property is proved by induction. Clearly $u^1 \geq u^0 = \phi$. If $u^i \geq u^{i-1}$ then $Au^{i+1} \geq \dot{u}^{i-1}$ and $u^{i+1} \geq \phi$, therefore, by Proposition 5.3 $u^{i+1} \geq u^i$.

For the second assertion we first assume that $i = 1$. Since $\phi \neq 0$, it follows from Lemma 5.9 that $u^1_j > 0$ for all $j \in \{1, \cdots, n\}$. Now we assume that there exists an interval of integers $[l_1, l_2]$ with $1 \leq l_1 \leq l_2 \leq n$ such that :

$$u^1_{l_2} = \phi_{l_2} \text{ and } u^1_j > \phi_j \quad \text{whenever } j \in [l_1, l_2]$$

By diminishing l_1 if necessary, we can assume that either $l_1 = 1$ or $u_{l_1-1} = \phi_{l_1-1}$. From $\left(Au^1\right)_{l_2} \geq \phi_{l_2}$ and $u^1_{l_2} = \phi_{l_2}$, we derive :

$$au^1_{l_2-1} + b\phi_{l_2} + cu^1_{l_2+1} \geq \phi_{l_2}$$

and, since a and c are negative :

$$a\phi_{l_2-1} + b\phi_{l_2} + c\phi_{l_2+1} \geq \phi_{l_2}.$$

Therefore $n_1 + l_2 \in \Gamma(\phi)$ and by Lemma 5.10, $[n_1 + 1, n_1 + l_2] \subset \Gamma(\phi)$, therefore $(A\phi)_j \geq \dot{\phi}_j$ for $j \in [1, l_2]$. Since for all j in $[l_1, l_2[$, $\left(Au^1\right)_j = \dot{\phi}_j$, it follows that u^1 satisfies $\left(Au^1\right)_j \leq (A\phi)_j$ for j in $[l_1, l_2[$. Therefore, the vector $v = \phi - u^1$ satisfies $(Av)_j \geq 0$ for j in $[l_1, l_2[$, $v_{l_2} = 0$ and, if $l_1 > 1$ $v_{l_1-1} = 0$. Applying Lemma 5.9 to a submatrix of A, we infer that $v_j \geq 0$ for j in $[l_1, l_2[$. Hence $u^1_j \leq \phi_j$ for $j \in [l_1, l_2[$, which leads to a contradiction. Therefore the second assertion must be satisfied for $i = 1$.

It is now easy to complete the proof by induction. Assume the second assertion is true for i and that $[l_1, l_2]$ is an interval such that $u^{i+1}_{l_2} = \phi_{l_2}$ and $u^{i+1}_j > \phi_j$ for j in $[l_1, l_2[$. Then using the first property we have $u^i_{l_2} = \phi_{l_2}$ therefore $u^i_j = \phi_j$ for $j \leq l_2$ and the same reasoning as above leads to a contradiction. $\diamond$

Remark 5.12 The above results can be extended to the solution of a discretized system involving a Neumann condition at $n_2\Delta x$: $\bar{u}^i_{n_2} = \bar{u}^i_{n_2-1}$. The matrix A is then slightly different since its last diagonal entry must be $b+c$. Note that A is coercive if the following condition is satisfied :

$$1 + r\Delta t - \beta \frac{\Delta t}{2\Delta x} \geq 0$$

Remark 5.13 In the case of a call option, a similar algorithm can be justified along the same lines, using remark 5.8 (whereas a direct application of the Brennan-Schwartz algorithm would lead to a false solution). Instead of Lemma 5.10, the property to be used in that case is the following :

$$j_0 \in \Gamma(\psi) \quad \Longrightarrow \quad \forall j \geq j_0, \quad j \in \Gamma(\psi)$$

Remark 5.14 For some functions ψ, there seems to be no justification of the Brennan-Schwartz techniques. Consider, for instance, $\psi(x) = |e^x - K|$. The application of the Brennan-Schwartz algorithm to ψ in the Garman-Kohlhagen model does not solve the finite dimensional variational inequality. The owner of the American option defined by this function would have the right to buy or sell one unit of foreign currency at price K, before the date of maturity (note that this is not equivalent to holding one American call and one American put simultaneously).

6 Pivoting methods

6.1 Generalities

In this section, we survey general methods for solving the following linear complementarity problem:

$$(I)\left\{\begin{array}{c} Au \geq \theta \\ u \geq \phi \\ (Au-\theta, \phi - u) = 0 \end{array}\right.$$

If we let $\lambda = Au - \theta$, $s = u - \phi$, and $\delta = A\phi - \theta$, problem (I) is equivalent to:

$$(II)\left\{\begin{array}{c} \lambda = \delta + As \\ \lambda \geq 0,\ s \geq 0 \\ (\lambda, s) = 0 \end{array}\right.$$

Several direct methods have been proposed to solve problems of type (II). The algorithm of Lemke (see [27]) is a general-purpose method for these problems. For the class of all matrices with positive principal minors, a principal pivoting method has been developed (see Cottle and Dantzig [8]) based on the fact that this class is invariant under principal pivoting (see Tucker [31]). These algorithms are in the worst case exponential which is not surprising in view of the NP-hardness of problem (II) (see [7]).

In our specific problem, however, the matrix A has other interesting properties besides the fact of being coercive: it belongs to the class of matrices with nonpositive off-diagonal entries, and it is also tridiagonal. We will see next how to take advantage of this fact.

6.2 Algorithm for matrices with nonpositive off-diagonal entries

As applied to problem (II), pivoting methods consist of successively exchanging the roles of dependent (initially the λ's) and independent (initially the s's) variables in order to obtain an equivalent system in which the transform of δ is nonnegative. Having done this, one would set the independent variables to zero and obtain the values of the corresponding dependent variables. Specifically, suppose that δ_i is initially negative and a_{ii} positive, an exchange of λ_i and s_i (i.e., a principal pivot on a_{ii}) would cause s_i to be dependent (or 'basic') and δ_i to be transformed to $-\delta_i/a_{ii}$.

An important fact about a matrix with nonpositive off-diagonal entries is that once a s-variable s_i becomes dependent, it never decreases as a result of future principal pivot operations and thus stays dependent. This fact can be illustrated by considering a line (say, the i^{th}) of the system (II) in which a pivot is called for:

$$\lambda_i = \delta_i + a_{i1}s_1 + \cdots + a_{ii}s_i + \cdots + a_{ij}s_j + \cdots + a_{in}s_n \tag{6.1}$$

If $\delta_i < 0$ and $a_{ii} > 0$, a pivot on a_{ii} can be interpreted as augmenting the value of s_i (while keeping the value of the other independent variables to 0) until λ_i $(= \delta_i + a_{ii}s_i)$ becomes zero, in which case λ_i becomes independent and is thus exchanged with s_i. This being done, any future pivot, say on a_{jj}, would result in an increase of s_j and, since $a_{ij} \leq 0$ in equation 6.1, in a non-decrease of s_i.

This fact has been exploited by Saigal (see [29]) who proposed the following direct method for solving (II) when A is a matrix with nonpositive off-diagonal entries:

Algorithm(Saigal)

1. Let $I = \{i : \delta_i < 0\}$
2. If I is empty, stop: $s = 0$ and $\lambda = \delta$ is a solution.
   Else choose i in I.
   If $a_{ii} \leq 0$, stop: no solution.
   Else pivot on a_{ii}.
   Rename the transform system as $\lambda = \delta + As$.
   Drop column i from matrix A and go to Step 1.

This algorithm gives a solution (if any) in at most n pivots. The overall complexity is thus $O(n^3)$.

6.3 The tridiagonal case

In the case the matrix A is also tridiagonal, one can adapt the previous algorithm (see Cottle and Sacher [9]) so that the number of nonzero entries in each rows of the system is, at any time, bounded by 4. The running time of the method is then $O(n^2)$. In [9] the practical implementation of this algorithm seemed very efficient for solving problems quite similar to the one discussed here. It is an open problem to know if this algorithm could possibly have a linear running time on average. It would also be interesting to investigate thoroughly the iterative method developed by Chandrasekaran ([6]) for these special matrices.

The authors are grateful to O. Chateau and D. Gabay for fruitful conversations about this work

References

[1] R. A. Adams, *Sobolev Spaces*, Academic press, 1975.

[2] A. Bensoussan, On the theory of option pricing, *Acta Applicandae Mathematicae*, 2, 1984, p. 139-158.

[3] A. Bensoussan, J. L. Lions, *Applications des inéquations variationelles en contrôle stochastique*, Dunod, Paris, 1978.

[4] F. Black, M. Scholes, The pricing of options and corporate liabilities, *J. Polit. Econ.*, 81, 1973, p.637-659.

[5] M. J. Brennan, E. S. Schwartz, The valuation of the American put option, *J. of finance*, 32, 1977, p. 449-462

[6] R. Chandrasekaran, A special case of the complementary pivot problem, *Opsearch*, 7, 1970, p. 263-268.

[7] S.J. Chung, A note on the complexity of LCP: the Lcp is strongly NP-complete, Technical Report 792, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, 1979.

[8] R.W. Cottle, G.B. Dantzig, Complementary pivot theory of mathematical programming, *Linear Algebra and Applications*, 1, 1968, p. 103-125.

[9] R.W. Cottle, R.S. Sacher, On the solution of large, structured linear complementary problems: the tridiagonale case, *Applied Math. Opt.* , 3, 1977, p. 321-340.

[10] R.W. Cottle, A. F. Veinott, Jr, Polyhedral sets having a least element, *Math. Programming* 3, 1972, p. 238-249.

[11] N. El Karoui, Les aspects probabilistes du contrôle stochastique, *Lecture Notes in Mathematics* 876, 72-238, Springer Verlag 1981.

[12] N. El Karoui, A.Millet, J.-P. Lepeltier, A probabilistic approach to the reduite, *Preprint*, 1988.

[13] N. El Karoui, I. Karatzas, Integration of the optimal stopping risk and a new approach to the Skorohod problem, *preprint*, 1988.

[14] A. Friedman, *Stochastic differential equations*, 2 vol., Academic Press, 1976.

[15] A. Friedman, Parabolic variational inequalities in one space dimension and smoothness of the free boundary, *J. of functional analysis* 18, 1975, 151-176.

[16] D. Gabay, Modéles stochastiques des marchés financiers : l'équation de valuation de la finance, *IFAC conference on distibuted parameter systems*, Toulouse, 1982.

[17] M. Garman, S. Kohlhagen, Foreign currency option values, *J. of international money and finance*, 2, 1983, p. 231-237.

[18] R. Glowinski, J.-L. Lions, R. Trémollières, Analyse numérique des inéquations variationnelles, Dunod, Paris, 1976

[19] J.M. Harrison, S. R. Pliska, Martingales and stochastic integrals in the theory of continuous trading, *Stochastic processes and applications* 11, 1981, p. 215-260.

[20] P. Jaillet, D. Lamberton, B. Lapeyre, Inéquations variationnelles et théorie des options, *C.R. Acad. Sci. Paris* 307, Série I, 1988, p. 961-965.

[21] P. Jaillet, D. Lamberton, B. Lapeyre, *Analyse numérique des options américaines*, Cahier du CERMA n^{o}9, 1988, p. 66-126.

[22] I. Karatzas, On the pricing of American options, *Appl. Math. Optimization*, 17, 1988, p. 37–60.

[23] I. Karatzas, Applications of stochastic calculus in financial economics, *to appear in Lecture Notes in Control and information sciences*

[24] I. Karatzas, Optimization Problems in the Theory of Continuous Trading, *SIAM J.Control and Optimization*, Vol. 27, No.6, pp. 1221–1259, 1989.

[25] D. Kinderlehrer, G. Stampacchia, *An introduction to variational inequalities and their applications*, Academic Press 1980.

[26] O. A. Ladyzenskaja, V. A. Solonnikov, N. N. Ural'ceva, *Linear and quasilinear equations of parabolic type*, Translations of mathematical monographs 23, AMS 1968.

[27] C.E. Lemke, Bimatrix equilibrium points and mathematical programming, *Management Science*, 11, 1965, p. 681-689.

[28] M. A. Maingueneau, Temps d'arrêts optimaux et théorie générale, *Séminaire de Probabilités XII*, Lect. Notes in Mathematics 649, Springer Verlag 1978, p. 457-467.

[29] R. Saigal, A note on a special linear complementary problem, *Opsearch*, 7, 1970, p. 175-183.

[30] H. Samelson, R. M. Thrall, O. Wesler, A partition theorem for Euclidean n-space, *Proc. Amer. Math. Soc.* 9, 1958, p.805-807.

[31] A.W. Tucker, Principal pivoting transforms of square matrices, *SIAM Review*, 5, 1963, p. 305.

[32] P. van Moerbeke, On optimal stopping and free boundary problems, *Arch. Rational Mech. and Analysis*, 60, 1976, p.101-148.