American Options: Theory and Numerical Analysis

Andrew Carverhill*
and
Nick Webber**

*Lecturer
Dept of Finance and Economics
University of Science and Technology
Hong Kong

**Lecturer
Dept of Accounting and Finance
Warwick Business School
University of Warwick

July 1988

**Chapter 6b in *Options: Recent Advances in Theory and Practice*,
Ed S D Hodges, Manchester University Press, 1990, pp80-94**

Financial Options Research Centre
Warwick Business School
University of Warwick
Coventry
CV4 7AL
Phone: 0203 523606

FORC Preprint: 88/2

6b *Andrew Carverhill and Nick Webber*

American options: theory and numerical analysis

I Introduction

The aim at this paper is to develop a formula for the value of an American option, and to give a simple and efficient procedure for determining this value numerically. Many of the ideas that we will use have been presented before (see references); we see the value of this paper largely as presenting these ideas in a coherent and rigorous way. We will indicate the points of contact with these papers, and our work can serve as an introduction to and review of them.

We will deal mostly with the American put: the option to sell a stock for a price, c, at any time before T. The case of the call is similar and in some ways simpler than that of the put, because optimally it will not be exercised between dividend dates. We will denote the value of this put option at time t by $\phi_t(x)$, if the price of the stock at time t is x. Our standing assumption about the stock is that it pays no dividend, and its price s_t follows a geometric Brownian motion with constant drift, μ, and variance parameter, σ^2, so that in the notation of the Itô calculus, we have

$$ds_t = s_t(\mu \, dt + \sigma \, dB_t), \tag{1}$$

where B_t is a standard Brownian motion. We indicate in Section VI how to relax these assumptions. Also, we assume throughout that the riskless interest rate is a constant, $r > 0$.

Our point of departure in this paper is the work by Black and Scholes (1973) on European options (see also Smith, 1976). We will make the assumptions about market behaviour which are usual to this work, namely that continuous trading is possible with no transactions costs, that there is no penalty for selling short, and that there are no taxes.

Our basic technical device is to approximate the American option by restricting the exercise opportunities to a finite set of times $\pi \equiv \{0 < t_1 < t_2 < \ldots < t_q \equiv T\}$. This is done in Section III. The value $\phi_t^\pi(x)$ of this approximation can be determined by regarding it as a succession of European options (a 'compound' option), and it is intuitively clear that $\phi_t^\pi(x)$ is a good approximation to $\phi_t(x)$ itself if mesh(π) (that is, $\max\{t_i - t_{i-1}: i = 1, \ldots, q\}$) is small.

In Sections IV and V we present procedures for the numerical evaluation of $\phi_t(x)$. The 'analytic formula' of Section IV actually values the approximation $\phi_t^\pi(x)$ for a given partition π of the interval $[0, T]$, and was first presented by Geske (1979) (see also Geske, 1977; Geske and Johnson, 1979). This formula involves multinormal distribution functions up to the dimension q, that is, the number of partition points, but it is quite accurate for $q = 3$ or 4. The 'dynamic' procedure of Section V is to solve the Black–Scholes equation by a finite difference method in reverse time, as though the option were European, but to adjust the solution at each step to allow for the possibility of early exercise. This procedure is intuitively reasonable, and is implemented, for instance, in Geske and Shastri (1985). From our work it is clear that this gives an accurate evaluation.

II Remarks on European options

The material of this section is well known: our purpose here is to prepare for subsequent sections. Denote by $\phi_t^{\mathrm{E}}(x)$ the value of a put option if the only exercise opportunity is T (that is, if the option is European). Then we have the celebrated Black and Scholes (1973) equation:

$$\frac{\partial \phi_t^{\mathrm{E}}(x)}{\partial t} = r\phi_t^{\mathrm{E}}(x) - rx\frac{\partial \phi_t^{\mathrm{E}}(x)}{\partial x} - \frac{1}{2}x^2\sigma^2\frac{\partial^2 \phi_t^{\mathrm{E}}(x)}{\partial x^2} \tag{2}$$

with initial (final!) condition

$$\phi_T^{\mathrm{E}}(x) = (c - x)^+ \tag{3}$$

where $(c - x)^+ = \max\{(c - x), 0\}$ (see also Smith, 1976). Equation (2) is a diffusion equation in reverse time, and so its solution is determined in a stable fashion by its final condition (see Oksendal, 1980).

The following result (Proposition 1) shows how to construct a continuously hedged portfolio from the stock and the European option.

Proposition 1. Consider a continuously adjusted portfolio comprising at time t quantities Q_t^{o} of the European put option, and Q_t^{s} of the stock. Then the value of this portfolio is given by

$$v_t = Q_t^{\mathrm{o}}\phi_t^{\mathrm{E}}(s_t) + Q_t^{\mathrm{s}}s_t. \tag{4}$$

Suppose we have

$$Q_t^{\mathrm{o}} = 1, \qquad Q_t^{\mathrm{s}} = -\frac{\partial \phi_t^{\mathrm{E}}(x)}{\partial x}. \tag{5}$$

Then the value of the portfolio grows risklessly at the risk-free rate, r, that is, we have

$$dv_t = rv_t\,dt. \tag{6}$$

Proof. Formula (4) is clear. Now, suppose the portfolio is adjusted according to equations (5) at times $\{0 \equiv t_0 < t_1 < \ldots < t_q \equiv T\}$ ($\equiv \pi$, say). Then the change in value over the interval $[t_i, t_{i+1}]$ is given by

$$\Delta^i v_t = \Delta^i \phi_t(s_t) - \frac{\partial \phi^{\mathrm{E}}_{t_i}}{\partial x}(s_{t_i})\,\Delta^i s_t.$$

where $\Delta^i v_t$ is shorthand for $v_{t_{i+1}} - v_{t_i}$, etc. Summing over $i = 1, \ldots, q-1$ and letting mesh$(\pi) \to 0$, this gives

$$\int_0^T dv_t = \int_0^T d(\phi_t(s_t)) - \int_0^T \frac{\partial \phi^{\mathrm{E}}_t}{\partial x}(s_t)\,ds_t,$$

where these integrals are stochastic Itô integrals. (This follows from the definition of the Itô integral.) Applying the Itô formula to the first term in the RHS of this equation, we obtain

$$\int_0^T dv_t = \int_0^T \frac{\partial \phi^{\mathrm{E}}_t}{\partial t}(s_t)\,dt + \int_0^T \frac{\partial \phi^{\mathrm{E}}_t}{\partial x}(s_t)\,ds_t$$
$$+ \frac{1}{2}\int_0^T \sigma^2 s_t^2 \frac{\partial^2 \phi^{\mathrm{E}}_t}{\partial x^2}(s_t)\,dt - \int_0^T \frac{\partial \phi^{\mathrm{E}}_t}{\partial x}(s_t)\,ds_t$$
$$= r\int_0^T [Q^{\mathrm{o}}_t \phi^{\mathrm{E}}_t(s_t) + Q^{\mathrm{s}}_t(s_t)\,s_t]\,dt$$

(using equation (2))

$$= r\int_0^T v_t\,dt.$$

Differentiating yields equation (6) as required. □

Proposition 2.

(i)

(1) The value $\phi^{\mathrm{E}}_t(x)$ of the European put (at time t, if the stock price then is x) satisfies equation (2) with final condition (3).

(2) Also, we have

$$\phi^{\mathrm{E}}_t(x) = \exp(-r(T-t))\,E[(c - \tilde{s}_T)^+ | \tilde{s}_t = x] \tag{7}$$

where $\tilde{s}_t$ satisfies

$$d\tilde{s}_t = \tilde{s}_t(r\,dt + \sigma\,dB_t). \tag{8}$$

(Replace μ by r to obtain equation (8) for $\tilde{s}_t$ from equation (1) for s_t.)

(ii) If at time t the market price of the option differs from $\phi_t^{\mathrm{E}}(x)$, then an arbitrage profit can be gained at time T.

Proof. If the market price of the option is below (above) $\phi_t(x)$, then one can buy it (sell it short), then construct the hedged portfolio of Proposition 1, and then realise the original value of the portfolio (with the theoretical option value, and with interest) at time T, thereby making an arbitrage profit. This proves part (ii). To see that $\phi_t(x)$ must satisfy equation (2), note that we used equation (2) in the proof of Proposition 1.

Equation (7) is the standard expression of the solution to equation (2), given by the stochastic calculus (see Oksendal, 1980). To prove it, consider the stochastic process Q_u given by $\exp(-r(u-t))\,\phi_u^{\mathrm{E}}(\tilde{s}_u)$ for fixed t and with time u increasing from t. If we calculate the differential of this using the Itô formula, we see that it is a martingale, and so

$$E[Q_u|\tilde{s}_t = x] = E[Q_t|\tilde{s}_t = x].$$

Putting $u = T$ in this last expression yields equation (7). □

The equation (2), together with equation (3), can be solved numerically by standard techniques for diffusion equations (for example finite difference methods; see Geske and Shastri, 1985), applied in reverse time. We refer to this as the 'dynamic' approach to the valuation, and it is extended to American options in Section V. In fact equation (2), taken together with equation (3), also has an 'analytic' (closed-form) solution, which is as follows:

$$\phi_t^{\mathrm{E}}(x) = c\exp(-r(T-t))\,N(-d_2) - xN(-d_1) \tag{9}$$

where $N(\cdot)$ is the cumulative normal distribution function and

$$d_1 = \frac{\ln(x/c) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\ln(x/c) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

(See Black and Scholes, 1973; Smith, 1986. Note that call and put are related via parity: $\mathrm{Put}_t - \mathrm{Call}_t = c\exp(-r(T-t)) - x$.)

In section IV we give an extension of this formula to deal with American options. (This extension is due to Geske, 1979; Geske and Johnson, 1979; Selby and Hodges, 1987). We refer to this as the 'analytic' approach to the solution.

To help with Section IV we make the following remarks about formula (7). The distribution of $\tilde{s}_T$ given $\tilde{s}_t = x$ is log-normal; denote its density function by $F_x(y)$. Then from equation (8) we have

$$\phi_t^{\mathrm{E}}(x) = \exp(-r(T-t)) \int_{y=0}^{c} F_x(y)(c-y)\,\mathrm{d}y$$

$$= \exp(-r(T-t))\, c \int_{y=0}^{c} F_x(y)\,\mathrm{d}y - \exp(-r(T-t))\, c \int_{y=0}^{c} yF_x(y)\,\mathrm{d}y$$

$$= \exp(-r(T-t))\, cP[\tilde{s}_T \leqslant c | \tilde{s}_t = x]$$

$$- \exp(-r(T-t))\, E[\tilde{s}_T \chi_{\{\tilde{s}_T \leqslant c\}} | \tilde{s}_t = x] \qquad (10)$$

and the terms in this last expression correspond to those in equation (9). (Here $\chi_{\{\tilde{s}_T \leqslant c\}}$ is the characteristic function of the 'event' $\{\tilde{s}_T \leqslant c\}$, which is 1 if $\tilde{s}_T \leqslant c$ and 0 otherwise.)

Remark on the Girsanov transformation

This is explained in detail in Oksendal (1980, pp. 115–18); it relates two stochastic processes which differ in their drifts but not in their noise. In this paper we use it to relate the processes s_t and $\tilde{s}_t$, which are given by equations (1) and (8). The standard Brownian motion, B_t, is governed by the underlying probability measure P (the Wiener measure) on the space of continuous paths, and equation (1) can be regarded as telling us how s_t is governed by P. The Girsanov idea is to transform P to a new measure $\tilde{P}$, such that the process s_t governed by $\tilde{P}$ is the same as the process $\tilde{s}_t$ governed by P. Thus we can rewrite equations (7) and (10) as

$$\phi_t^{\mathrm{E}}(x) = \exp(-r(T-t))\, \tilde{E}[(c-s_T)^+ | s_{T-t} = x]$$

$$= \exp(-r(T-t))\, c\tilde{P}[s_T \leqslant c | s_t = x]$$

$$- \exp(-r(T-t))\, \tilde{E}[s_T \chi_{\{s_T \leqslant c\}} | s_t = x], \qquad (11)$$

where $\tilde{E}$ is the expectation associated with $\tilde{P}$. The advantage of equation (11) is that it refers to the original stock process, s_t, and not to the artifical process, $\tilde{s}_t$.

In fact P and $\tilde{P}$ are related via

$$\mathrm{d}\tilde{B}_t = \mathrm{d}B_t + \left(\frac{r-u}{\sigma}\right) \mathrm{d}t, \qquad (12)$$

where $\tilde{B}_t$ is the Brownian motion governed by $\tilde{P}$. Note that if we denote s_t driven by $\tilde{P}$ as $\bar{s}_t$, then we have

$$\mathrm{d}\bar{s}_t = \bar{s}_t(\mu\,\mathrm{d}t + \sigma\,\mathrm{d}\tilde{B}_t),$$

and formally substituting for $\mathrm{d}\tilde{B}_t$ in this equation, using equation (12), yields equation (8) for $\bar{s}_t$, from which we see that $\bar{s}_t = \tilde{s}_t$. The direct relationship between P and $\tilde{P}$ is given by the formula for M_t in Oksendal

(1980, p. 118). This M_t is the 'density' (Radon Nikodym derivative) of $\tilde{P}$ with respect to P.

III American options via a discretisation of the exercise opportunity set

The problem of determining a value for the American option is intimately related to the problem of determining an optimal time to exercise it. This time can be random (that is, it can depend on the behaviour of the stock price), but to be a feasible exercise strategy (which does not rely on clairvoyance), it must be a stopping time, that is, at any given time the question of whether the option has already been exercised must be independent of the future behaviour of the stock price (see Oksendal, 1980). In fact since in our model the stock price is Markovian, at any given time t, the optimal decision about whether to exercise or wait (given that the option has not already been exercised), must be independent of the past as well as the future, and must depend only on whether the present stock price, s_t, is greater than some critical price $\bar{s}_t$ (exercise the put option if $s_t < \bar{s}_t$).

Denote by $\phi_t^\pi(x)$ the value of the American option if we restrict the exercise opportunities to the finite set $\pi = \{0 < t_1 < t_2 < \ldots < t_q \equiv T\}$. Under this restriction it is easy to find that value and optimal exercise time.

Theorem 3.

(i) Starting from time t, the optimal exercise time τ^π and value $\phi_t^\pi(x)$ are determined by either of the following equivalent criteria:

(1) The final condition (at time $t_q \equiv T$) is

$$\phi_{t_q}^\pi(x) = (c - x)_-^+.$$

Take the critical price $\bar{s}_{t_q}$ to be just c.

Solve equation (2) with this final condition to determine $\phi_t^\pi(x)$ for $t \in (t_{q-1}, t_q)$. To obtain $\phi_{t_{q-1}}^\pi(x)$ continue this solution back to $t = t_{q-1}$ (and denote it by ϕ, say) and put $\phi_{t_{q-1}}^\pi(x) = \max\{\phi(x), (c - x)^+\}$. Define the critical price $\bar{s}_{t_{q-1}}$ to be the solution to $\phi(x) = (c - x)^+$ and exercise at time t_{q-1} if $s_{t_{q-1}} < \bar{s}_{t_{q-1}}$ (which is equivalent to $\phi(s_{t_{q-1}}) < (c - s_{t_{q-1}})^+$.)

Now solve equation (2) on the interval $[t_{q-2}, t_{q-1}]$ with final condition $\phi_{t_{q-1}}^\pi(x)$ and replace the solution at time t_{q-2} (say, $\phi(x)$ again) by $\max\{\phi(x), (c - x)^+\}$. Also define the critical price, $\bar{s}_{t_{q-2}}$, to be the solution to $\phi(x) = (c - x)^+$ and exercise at time t_{q-2} if $s_{t_{q-2}} < \bar{s}_{t_{q-2}}$.

Continue inductively backwards, the last step being over the interval $[t, t_{k+1}]$, where k is such that $t \in [t_k, t_{k+1}]$.

Starting from t, the optimal exercise time τ^π is the first time t_j for which $s_{t_j} < \bar{s}_{t_j}$.

(2) $$\phi_t^\pi(x) = \sup_{\tau \in S_t^\pi} E[\exp(-r(\tau - t))(c - \tilde{s}_\tau)^+ | \tilde{s}_t = x],$$

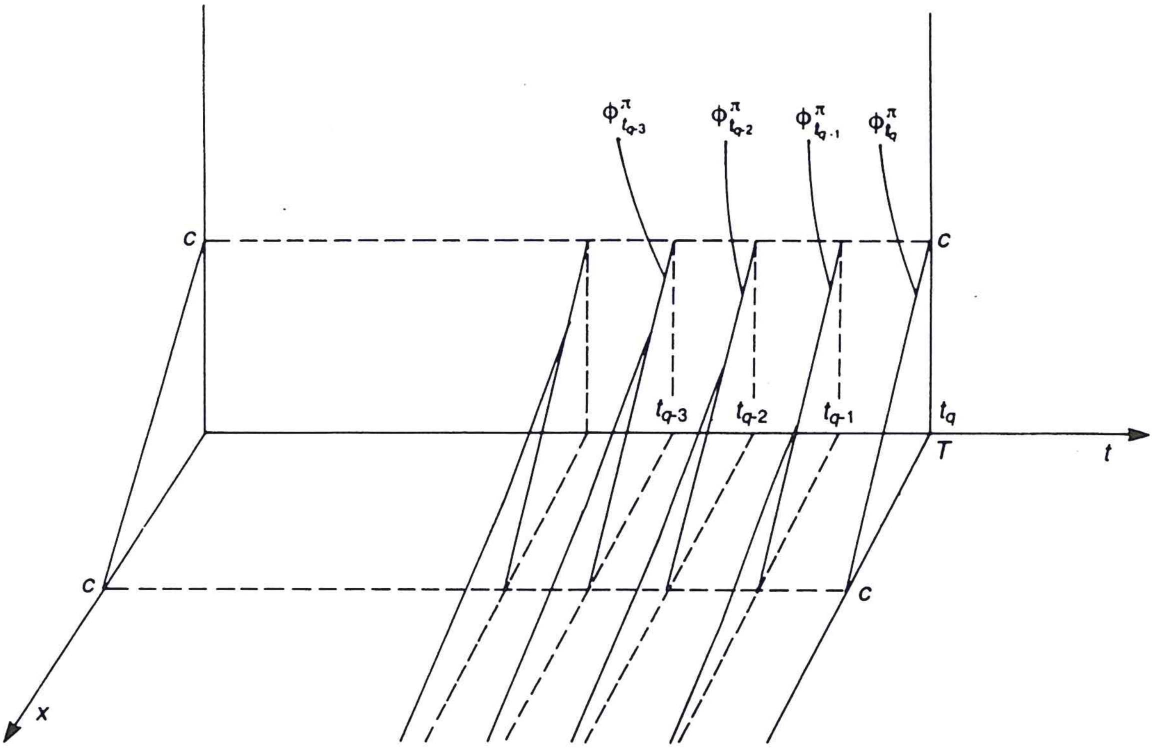

Figure 6b.1

where S_t^π is the collection of stopping times τ which take values in π and for which $t \leqslant \tau \leqslant T$. The optimum τ^π is where the sup occurs.
(ii) The continuous hedge represented by equation (5) but with ϕ_t^{E} replaced by ϕ_t^π will risklessly yield the riskless rate of return at time τ^π. Also any deviation of the market price from ϕ_t^π will allow an arbitrage profit at time τ^π.

Proof.

(i)

(1) The Black–Scholes equation (2) is very generally applicable: it will hold so long as there is no exercise opportunity. In particular it will hold between the partition points. Replacing the solution $\phi(x)$ at each partition point by $\max\{\phi(x), (c - x)^+\}$ accounts for the possibility of early exercise.
(2) For this we use Part (1) and the general expression for the solution $\psi_t(x)$ to equation (2) for $t \leqslant t_k$, given final condition, say, $\psi(x)$ at t_k:

$$\psi_t(x) = \exp(-r(\tau - t))\, E[\psi(\tilde{s}_{t_k}) | \tilde{s}_t = x]. \qquad (13)$$

(Cf. equation (7)). Thus for $t \in [t_{q-1}, t_q]$ we have

$$\phi_t^\pi(x) = \exp(-r(t_q - t))\, E[(c - \tilde{s}_{t_q})^+ | \tilde{s}_t = x] \qquad (14)$$

$$\text{or } (c - x)^+ \text{ if } t = t_{q=1} \text{ and } x \leqslant \bar{s}_{t_{q-1}} \qquad (15)$$

(i.e. if $t = t_{q=1}$ and we exercise immediately). For $t \in [t_{q-2}, t_{q-1}]$ we have

$$\begin{aligned} \phi_t^\pi(x) &= \exp(-r(t_{q-1} - t))\, E[\phi^\pi_{t_{q-1}}(\tilde{s}_{t_{q-1}}) | \tilde{s}_t = x] \\ &= \exp(-r(t_{q-1} - t))\, E[\phi^\pi_{t_{q-1}}(\tilde{s}_{t_{q-1}})(\chi_t^{t_{q-1}} + \tilde{\chi}_t^{t_{q-1}}) | \tilde{s}_t = x] \qquad (16) \end{aligned}$$

where $\chi_t^{t_{q-1}}$ is the characteristic function of the event 'exercise at time t_{q-1} having started at time t', and $\tilde{\chi}_t^{t_{q-1}}$ is the complement, that is, $\tilde{\chi}_t^{t_{q-1}} + \chi_t^{t_{q-1}} = 1$. In equation (16), substitute for $\phi^{\pi}_{t_{q-1}}(\tilde{s}_{t_{q-1}})$ using equation (14) when $\tilde{\chi}_t^{t_{q-1}}$ is activated and using equation (15) when $\chi_t^{t_{q-1}}$ is activated, to obtain

$$\phi_t^{\pi}(x) = \exp(-r(t_{q-1} - t))\, E[(c - \tilde{s}_{t_{q-1}})^+ \chi_t^{t_{q-1}} | \tilde{s}_t = x]$$
$$+ \exp(-r(t_{q-1} - t)) \exp(-r(t_q - t_{q-1}))\, E[(c - \tilde{s}_{t_q})^+ \tilde{\chi}_t^{t_{q-1}} | \tilde{s}_t = x]. \qquad (17)$$

Also we have $\phi_t^{\pi}(x) = (c - x)^+$ rather than equations (16) and (17) if $t = t_{q-2}$ and $x < \bar{s}_{t_{q-2}}$ (that is, we exercise immediately).

The result follows again by continuing inductively backwards.

(ii) For this part proceed as in the proof of Propositions 1 and 2. □

In Theorem 4 we value the full American put option. The result is very similar to Theorem 3 for the discretised option (as one would expect since the restriction of the exercise opportunities to π should matter very little if π is a fine discretisation); however, Theorem 4 is more difficult technically than Theorem 2.

Theorem 4.

(i) Starting from time t, the optimal exercise time, τ^*, and value, $\phi_t(x)$, for the full American option are determined by either of the following equivalent criteria:

(1) The stopping time, τ^*, is the first time, r, after t at which the stock price, s_r, finds itself below the critical price, $\bar{s}_r$; the critical price function, $t \to \bar{s}_t$, and value function, $\phi_t(x)$, are the solution to the free boundary problem:

$$\frac{\partial \phi_t(x)}{\partial t} = r\phi_t(x) - rx\frac{\partial \phi_t(x)}{\partial x} - \frac{1}{2}x^2\sigma^2\frac{\partial^2 \phi_t(x)}{\partial x^2} \qquad (2')$$

for $x > \bar{s}_t$,

$$\frac{\partial \phi_t(x)}{\partial t} = r\phi_t(x) - rx\frac{\partial \phi_t(x)}{\partial x} - \frac{1}{2}x^2\sigma^2\frac{\partial^2 \phi_t(x)}{\partial x^2}$$

for $x < \bar{s}_t$,

$$\phi_t(x) = (c - x)^+ \text{ and } \frac{\partial \phi_t(x)}{\partial x} = \frac{\partial}{\partial x}(c - x)^+ \equiv -1$$

for $x = \bar{s}_t$;

for all $t \in [0, T]$ we have

$$\phi_t(x) \to 0 \text{ as } x \to \infty.$$

(2) $$\phi_t(x) = \sup_{\tau \in S_t} E[\exp(-r(\tau - t))(c - \tilde{s}_\tau)^+ | \tilde{s}_t = x],$$

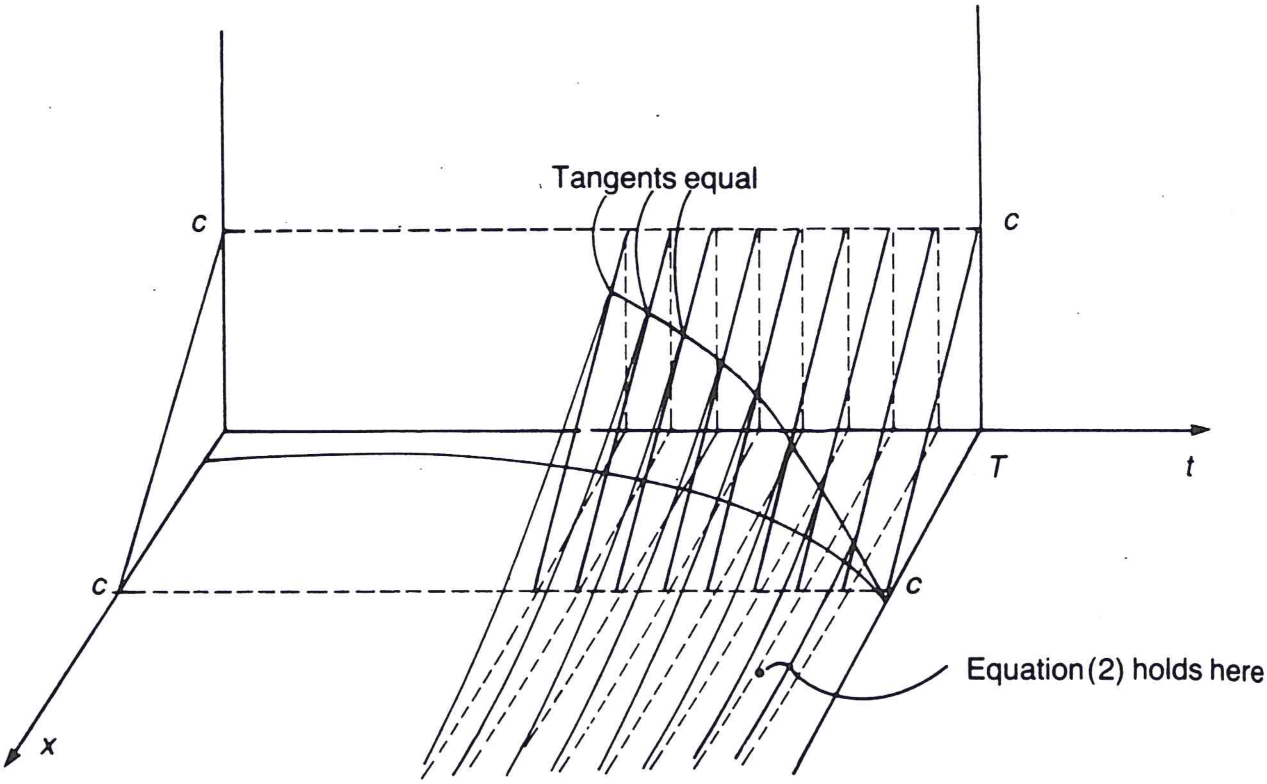

Figure 6b.2

where S_t is the collection of stopping times τ such that $t \leqslant \tau \leqslant T$. The optimum τ^* is where the sup occurs.
(ii) The continuous hedge represented by equation (5) with ϕ_t^{E} replaced by ϕ_t will risklessly yield the riskless rate of return at time τ^*. Any deviation of the market price from ϕ_t will allow an arbitrage profit at time τ^*.

Proof.
(i) The existence of the solution τ^* and $\phi_t(x)$, and the equivalence of Parts (1) and (2), follow from general optimal stopping theory. Note that the criterion of Part (1) in Theorem 3 is the general characterisation of the superharmonic majorant function, which is discussed in Oksendal (1980). The function is harmonic in the continuation region and strictly superharmonic in the stopping region (see Jacka, 1988).

We will prove Part (2). Any sensible exercise (stopping) time, τ, must be of the form 'first time r after t at which stock price s_r is less than the critical price $F(r)$', where F is some function $[0, T] \to \mathbb{R}^{\geqslant 0}$. This follows from the first paragraph of this section. Denote by $\phi_t^\tau(x)$ the value of the option, having decided on τ and F. Then above the function F (that is, at (x, t) for which $x > F(t)$), equation (2′) must hold for $\phi_t^\tau(x)$, because the essential condition for equation (2′) is that we do not exercise. It follows from equation (2′) that for fixed t and given $\tilde{s}_t$ and as u increases from t, the process $\exp(-r(u - t))\, \phi_u^\tau(\tilde{s}_u)$ is a martingale so long as $(\tilde{s}_u, u)$ is above F (that is, $\tilde{s}_u > F(u)$). (To see this, calculate the differential $\mathrm{d}\exp(-r(u - t))\, \phi_u^\tau(\tilde{s}_u)$ using the Itô formula. Also Cf. Proposition 2.)

Thus, if ρ is a stopping time such that $t \leqslant \rho \leqslant \tau$ (that is, ρ stops above F), then we have

$$\phi_t^{\tau}(x) = E[\exp(-r(\rho - t))\,\phi_{\rho}^{\tau}(\tilde{s}_{\rho})|\tilde{s}_t = x]. \tag{18}$$

(Equation (18) is a more refined version of equation (7).) Applying the Girsanov transformation to equation (17), replacing ρ by τ, and noting that $\phi_{\tau}^{\tau}(s_{\tau}) = (c - s_{\tau})^+$ (because we exercise at τ), we obtain

$$\phi_t^{\tau}(x) = \tilde{E}[\exp(-r(\tau - t))(c - s_{\tau})^+|s_t = x]. \tag{19}$$

The result follows by applying the reverse Girsanov transformation to (19).
(ii) As in Theorem 3(ii). □

Our final result in this section shows how $\phi_t(x)$ is approximated by $\phi_t^{\pi}(x)$:

Theorem 5. Put $\text{mesh}(\pi) \equiv \max\{(t_k - t_{k-1}) : k = 1, \ldots, q\} = \varepsilon$. Then we have, for all x and $t \in [0, T]$, that

$$0 \leqslant \phi_t(x) - \phi_t^{\pi}(x) \leqslant c(1 - \exp(-r\varepsilon)),$$

that is, convergence is uniformly order ε as $\varepsilon \to 0$.

Proof.
(1) Take τ^* to be the optimal exercise time and denote by τ_{π}^* the first time in the set π, after τ^*. Note that τ_{π}^* is still a stopping time, and $0 \leqslant \tau_{\pi}^* - \tau^* \leqslant \varepsilon$.

We will show that

$$E[\exp(-r(\tau^* - t))(c - \tilde{s}_{\tau^*})^+|\tilde{s}_t = x] - E[\exp(-r(\tau_{\pi}^* - t))(c - \tilde{s}_{\tau_{\pi}^*})^+|\tilde{s}_t = x] \leqslant c(1 - \exp(-r\varepsilon)), \tag{20}$$

and our result will follow from this because

$$\phi_t(x) \geqslant \phi_t^{\pi}(x) \equiv \sup_{\tau \in S_t^{\pi}} E[\exp(-r(\tau - t))(c - \tilde{s}_{\tau})^+|\tilde{s}_t = x]$$
$$\geqslant E[\exp(-r(\tau_{\pi}^* - t))(c - \tilde{s}_{\tau_{\pi}^*})^+|\tilde{s}_t = x].$$

(2) The idea of the proof is to note that waiting until time τ_{π}^* beyond τ^* to exercise the option (that is, a time $\tau_{\pi}^* - \tau^*$ past the optimum) will reduce its value by at most $c(1 - \exp(-r\varepsilon)$. For this note that this extra wait is equivalent to taking out a European option when time τ^* arrives, to mature at time $\tau_{\pi}^* - \tau^*$ ($\leqslant \varepsilon$). This wait must be disadvantageous because τ^* is the optimum, and the disadvantage is equal to

$$\phi^0(y) - \phi^{\eta}(y),$$

where $s_{\tau} = y$ and $\eta = \tau_{\pi}^* - \tau^*$, and $\phi^{\eta}(y)$ (or $\phi^0(y)$) is the current value of a European option which matures at time, η in the future (or immediately).

But we must have $y < c$ if exercise is optimal for $s_{\tau} = y$, and so

$$\phi^0(y) - \phi^\eta(y) \equiv (c - y)^+ - \exp(-r\eta)\, E[(c - \tilde{s}_\eta)^+ | \tilde{s}_0 = y]$$
$$\leqslant (c - y) - \exp(-r\eta)\, E[(c - \tilde{s}_\eta) | \tilde{s}_0 = y]$$

(including negative values makes minus the expectation bigger)

$$= (c - y) - \exp(-r\eta)\, c + E[\exp(-r\eta)\, \tilde{s}_\eta | \tilde{s}_0 = y].$$

But this last term is just y, because $\exp(-r\eta)\,\tilde{s}_\eta$ is a martingale as η increases. (To see this, calculate the differential $\mathrm{d}[\exp(-r\eta)\,\tilde{s}_\eta]$.) Thus

$$\phi^0(y) - \phi^\eta(y) \leqslant c(1 - \exp(-r\eta)) \tag{21}$$

as required.
(3) Here we formally deduce equation (20) using equation (21):
Now,

$$E[\exp(-r(\tau_\pi^* - t)(c - \tilde{s}_{\tau_\pi^*})^+ | \tilde{s}_t = x] = E[\exp(-r(\tau^* - \mathrm{t}))$$
$$E[\exp(-r(\tau_\pi^* - \tau^*))(c - \tilde{s}_{\tau_\pi^*})^+ | (\tilde{s}_\tau, \tau) \text{ and } \tilde{s}_t = x] | \tilde{s}_t = x]$$

(conditioning on (s_τ, τ)). Therefore, the LHS of equation (20) is equal to

$$E[\exp(-r(\tau^* - t)\{\phi^0(\tilde{s}_{\tau^*}) - \phi^{(\tau_\pi^* - \tau^*)}(\tilde{s}_{\tau^*})\} | \tilde{s}_t = x] \leqslant c(1 - \exp(-r\eta))$$

by equation (21) as required. □

Notes
(i) Theorem 5 crucially uses the fact that $\phi_t(x)$ corresponds to an optimal stopping time.
If τ is not optimal and τ_π is the first time after τ in the set π then we only have

$$|E[\exp(-r(\tau - t))(\mathrm{c} - \tilde{s}_\tau)^+ | \tilde{s}_t = x]$$
$$- E[\exp(-r(\tau_\pi - t))(c - \tilde{s}_{\tau_\pi})^+ | \tilde{s}_t = x]| = 0(\mathrm{mesh}(\pi)^{\frac{1}{2}}).$$

(ii) Using the ideas of Theorem 5, one can deduce that ϕ^π is a Cauchy sequence as $\mathrm{mesh}(\pi) \to 0$, and hence Theorem 4, Criterion (2).

IV The analytic approach for American options

This approach gives an 'analytic' formula, which is an extension of equation (9) for European options, and is actually a formula for the approximation $\phi_t^\pi(x)$ rather than $\phi_t(x)$ itself. This approximation is surprisingly accurate in view of Theorem 5 above: if the annual interest rate is 10% (so that $r = \ln(11/10)$) then the approximation for a nine-month option with three quarterly exercise opportunities has error at most $c(1 - \exp(-r/4)) \simeq 0.024c$, that is, about $2\frac{1}{2}$ per cent of the exercise price. The approach was developed by Geske (1979); see also Geske (1977); Geske and Johnson (1979); Selby and Hodges (1987).

To help our presentation, we will assume that the present time, t, is 0. The approach is to go through the procedure of Theorem 3 Part (1) calculating $\phi^{\pi}_{t_k}$ in turn for $k = q, q-1, \ldots,$ by using an explicit formula (which is an extension of equation (9)) for the solution to equation (2) over each interval $[t_k, t_{k+1}]$. We present the procedure in detail:

First, at time $t_q \equiv T$ we have

$$\phi^{\pi}_{t_q}(x) = (c - x)^+, \bar{s}_{t_q} = c.$$

Now, the solution, say, $\tilde{\phi}^{\pi}_{t_{q-1}}$ at time t_{q-1} to equation (2), with final condition $\phi^{\pi}_{t_q}$, is given by formula (9) for $\phi^{E}_{t_{q-1}}$. Having this, we can put $\pi^{\pi}_{t_{q-1}}(x) = \max\{\tilde{\phi}^{\pi}_{t_{q-1}}(x), (c - x)^+\}$, and take the critical price $\bar{s}_{t_{q-1}}$ to be the solution to the equation $\tilde{\phi}^{\pi}_{t_{q-1}}(x) = (c - x)$. (Note that this equation must be solved numerically.)

To obtain $\phi^{\pi}_{t_{q-2}}$ we solve equation (2) with final condition $\phi^{\pi}_{t_{q-1}}$, now using a more complicated version of formula (9), which we give in its general form below as formula (23), and then we replace this solution (say, $\tilde{\phi}^{\pi}_{t_{q-2}}(x)$) by $\max\{\tilde{\phi}^{\pi}_{t_{q-2}}(x), (c - x)\}$. To obtain the critical price, $\bar{s}_{t_{q-2}}$, we solve the equation $\tilde{\phi}^{\pi}_{t_{q-2}}(x) = (c - x)$.

Now continue inductively backwards. The formula for the solution $\tilde{\phi}^{\pi}_{t_k}$ to equation (2) over the time interval $[t_k, t_{k+1}]$, with initial condition $\phi^{\pi}_{t_{k+1}}$, is as follows. Suppose we have critical prices $\bar{s}_{t_{k+1}}, \ldots, \bar{s}_{t_q}$.

Put $$\Omega_{kj} = \text{Event}\{s_{t_{k+1}} > \bar{s}_{t_{k+1}}, \ldots, s_{t_{j-1}} > \bar{s}_{t_{j-1}}, s_{t_j} \leq \bar{s}_{t_j}\}$$

(that is, 'exercise at time t_j having started at time t_k'). If the event Ω_{kj} occurs then we profit by $(c - s_{t_j})$ at time t_j, and so

$$\begin{aligned}\tilde{\phi}^{\pi}_{t_k}(x) &= \sum_{j=k+1}^{q} \tilde{E}[\exp(-r(t_j - t_k))(c - s_{t_j})\chi_{\Omega_{kj}} | s_{t_k} = x] \\ &= c \sum_{j=k+1}^{q} \exp(-r(t_j - t_k))\, \tilde{P}[\Omega_{kj} | s_{t_k} = x] \\ &\quad - \sum_{j=k+1}^{q} \exp(-r(t_j - t_k))\, \tilde{E}[s_{t_j}\chi_{\Omega_{kj}} | s_{t_k} = x].\end{aligned}$$

(We have used the Girsanov transformation in equation (22); we refer to s_t (rather than $\tilde{s}_t$), and to the transformed measure $\tilde{P}$, and $\tilde{E}$. Compare equation (22) to equations (10) and (11).) s_t, driven by $\tilde{P}$, is a geometric Brownian motion. This fact can be used to rewrite equation (22) in terms of multinormal distributions (cf. equation (7)). We obtain

$$\tilde{\phi}^{\pi}_{t_k}(x) = c \sum_{j=k+1}^{q} \exp(-r(t_j - t_k))\, N_{j-k}(\mathbf{d}_{kj2}, \mathbf{R}_{kj}) - \sum_{j=k+1}^{q} N_{j-k}(\mathbf{d}_{kj1}, \mathbf{R}_{kj}) \tag{23}$$

where $N_{j-k}(\cdot, \mathbf{R})$ is the multinormal cumulative distribution with dimension $j - k$ and covariance matrix $\mathbf{R}$, and

$$\mathbf{d}_{kj1} = (d_{k,k+1}, \ldots, d_{k,j-1}, -\mathrm{d}_{k,j}),$$
$$\mathbf{d}_{kj2} = \mathbf{d}_{kj1} - \sigma(\sqrt{k_{k+1} - t_k}, \ldots, \sqrt{t_{j-1} - t_k}, -\sqrt{t_j - t_k})$$

where

$$d_{k,l} = \frac{\ln(x/\bar{s}_{t_l}) + (r + \frac{1}{2}\sigma^2)(t_l - t_k)}{\sigma\sqrt{t - t_k}},$$

and

$$\mathbf{R}_{kl} = \begin{bmatrix} +1 & , +r_{k,k+1}, & \ldots, & +r_{k,j-1}, & -r_{kj} \\ +r_{k,k+1}, & +1 & & & \cdot \\ \vdots & & \ddots & & \vdots \\ +r_{k,j-1} & & & & \cdot \\ -r_{kj} & \cdots & \cdots & \cdots & +1 \end{bmatrix}$$

where

$$r_{pq} = \sqrt{\frac{t_p - t_k}{t_q - t_k}}.$$

Notes on the analytic approach

(i) The efficient evaluation of equation (23) is studied by Selby and Hodges (1987). Note that we might have to calculate cumulative multinormal functions of dimension 3 or 4 in this approach, which is time-consuming for the computer, but not prohibitively so.

(ii) The analytic approach can be adapted to deal with stocks paying dividends, but this requires even more complicated formulae (see Whaley, 1981). To adapt the approach for more general price processes seems very difficult.

(iii) Geske and Shastri (1980) conclude that it might be almost optimal to exercise just after a dividend is paid. Given this, the analytic approach is very accurate if we take π to be the set of dividend dates.

V The dynamic approach for American options

Our approach here is to apply a standard finite difference method for diffusion equations to equation (2) (developing the solution backwards from time T), and simply to replace the calculated solution, say, $\tilde{\phi}_t(x)$ at each time by $\max\{\tilde{\phi}_t(x), (c - x)\}$ to account for the possibility of early exercise. This approach seems reasonable and is widely implemented (see for example, Geske and Shastri, 1985), but it has not quite been made clear that it will succeed. Geske and Shastri (1985) discuss various finite difference methods for equation (2), and conclude that the explicit version applied to a logarithmic transformation of equation (2) is efficient. (Note that such a transformation causes equation (2) to have constant coefficients.)

If the (logarithmically transformed) x-axis is split up into intervals of length h for the finite difference method, and the time axis to intervals of length k, then Geske and Shastri (1985) show that for stability of the explicit finite difference method we require $k = O(h^2)$ and that the error in the finite difference solution is order $O(k)$ and order $O(h^2)$. The device of replacing this solution, say, $\tilde{\phi}_t(x)$ at each step by $\max\{\tilde{\phi}_t(x), (c - x)\}$ means that we are actually calculating the discretised option value ϕ_t^π with $\text{mesh}(\pi) = k$; thus, since $|\phi_t^\pi - \phi_t| = O(k)$ (Theorem 5), we see that with this device the error in the explicit finite difference method is still $O(k)$.

Note that the dynamic approach is easily adapted to dealing with stocks paying dividends, and to more general price processes. Thus, it is more flexible than the analytic approach.

Our device can be used similarly with the binomial evaluation (see also Geske and Shastri, 1985), which is in theory almost the same as an explicit finite difference method.

Note

We would like to thank Stewart Hodges, Michael Selby and Saul Jacka for their help in clarifying many of the ideas presented here.

References

Bensoussan, A. (1984), On the Theory of Option Pricing, *Acta Applicandae Mathematicae*, 2, 139–58.

Black, F. and Scholes, M. (1973), The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, 81, 637–59.

Brennan, M. J. and Schwartz, E. S. (1977), The Valuation of American Put Options, *Journal of Finance*, 32, 449–62.

Geske, R. (1977), The Valuation of Corporate Liabilities as Compound Options, *Journal of Financial and Qualitative Analysis*, 12, 541–52.

Geske, R. (1979), The Valuation of Compound Options, *Journal of Financial Economics*, 7, 63–81.

Geske, R. and Johnson, H. E. (1979), The American Put Valued Analytically, *Journal of Finance*, 39, 1511–24.

Geske, R. and Shastri, K. (1980), *The Early Exercise of American Puts*, UCLA Working Paper 13–80 (Graduate School of Management).

Geske, R. and Shastri, K. (1985), Valuation by Approximation: A Comparison of Option Valuation Techniques, *Journal of Financial and Qualitative Analysis*, 20, 1, 45–72.

Jacka, S. (1988), *Optimal Stopping and the American Put*. Statistics Dept. Research Series 149, Warwick University.

Karatzas, I. (1988) On the Pricing of American Options, *Applied Mathematics Optimisation*, 17, 37–60.

Oksendal, B. (1980), *Stochastic Differential Equations*, Universitext Series, Springer Verlag, Berlin/Heidelberg/New York/Tokyo.

Parkinson, M. (1977), Option Pricing: The American Put, *Journal of Business*, 50, 21–36.

Selby, M. and Hodges, S. (1987), On the Evaluation of Compound Options, *Management Science*, 33, 3, 347–55.

Smith, C. (1976), Option Pricing: A Review, *Journal of Financial Economics*, 3, 3–51.

Whaley, R. E. (1981), On the Evaluation of American Call Options on Stocks with Known Dividends, *Journal of Financial Economics*, 9, 207–11.