DIMACS Technical Report 2000-30
October 2000

A Model for Stock Price Fluctuations Based on Information

by

Larry Shepp[1]
Dept. of Statistics
Rutgers University
New Brunswick, New Jersey 08903

[1]Supported in part by NSF STC 91-19999

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, NEC Research Institute and Telcordia Technologies (formerly Bellcore).

DIMACS was founded as an NSF Science and Technology Center, and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

This paper is dedicated to the memory of Aaron Wyner, the well known information theorist. It gives a new model for stock price fluctuations based on a concept of "information". In contrast, the usual Black-Scholes-Merton-Samuelson model is based on the *explicit* assumption that information plays *no* role in stock prices. The new model is based on the non-uniformity of information in the market and the time delay until new information becomes generally known.

The new model is expected to give more accurate predictions of future prices and more accurate formulas for hedge option valuations. The new valuations have been calculated for the various standard options inside the new model in a recent PhD thesis by Xin Guo. Because the concept of information is the driving one in the new model it seemed appropriate to discuss this in the present volume even though "information" is used in a somewhat different sense here than in communication theory.

In communication theory, information is *intended* to be communicated. One is concerned with designing a means for the successful transfer of messages from a source to the receiver and the value of the information is in its successful transmission. In the stock market, the reverse is usually true; information is hoarded and is of value to the owner *only* until it becomes known to others. Some messages are avidly communicated in the market but are really meant to *disinform* the receivers. This may of course also be true occasionally in communication theory which usually avoids concerning itself with the *meaning* of the message to be transmitted. Despite these differences in the role of information in the two situations, there are similarities in the mathematics used to study them. Certainly probabilistical models play a prominent role in both theories.

In Shannon's communication theory quantitative value is assigned to a channel (its rate or its capacity). Is there anything similar in the market? We are able to assign a quantitative value to knowing an item of information *which is unknown to others* say that a company is in a strong (or weak) position at a given point in time, via pricing of options on the company's stock. This paper is intended as a first step towards the goal of making the use of information into a quantitative tool.

Problems with explicit solutions are of value in obtaining insights. At the end of the paper we compare several problems of mathematical interest in order to better understand which optimal stopping problems have explicit solutions.

Section 1. Modelling stock price fluctuations in the real world.

The price X_t of a stock at time t is not well-defined since it is only after a transaction has taken place that the price becomes clear. Before this point there are two prices: bid and ask. Actual transactions occur in discrete points in time and at prices which are discrete multiples of a unit. It is nevertheless convenient to model $X_t, t \geq 0$ as a stochastic process continuous in both time and space variables.

Bachelier's ([B],1900) model takes the additive form

$$X_t = x + \mu t + \sigma W_t, \qquad t \geq 0,$$

in modern terminology where x is the price at time $t = 0$, μ represents *drift*, σ *volatility*, and W_t is a standard Brownian motion process. Samuelson ([S1],1965) used instead the exponential form

$$X_t = xe^{\sigma W_t + (\mu - \frac{\sigma^2}{2})t}, \qquad t \geq 0$$

which has the advantage that X_t remains nonnegative at all times. The two models appear more similar in their stochastic differential equation form. The differential in the Bachelier model is then

$$dX_t = \mu dt + \sigma dW_t, \qquad t \geq 0$$

while the Samuelson model has

$$dX_t = X_t(\mu dt + \sigma dW_t), \qquad t \geq 0$$

in which the change in X_t is proportional to the size of X_t, which is perhaps more sensible on dimensional considerations. Ito's calculus where $(dW_t)^2 = dt$ is used to obtain the differential form. Samuelson's model is now widely used to study stock prices. Each model seems crude in that each depends on only the two parameters μ and σ but it is perhaps reasonable to assume that this is about the level of knowledge possessed by a typical investor and the price is really determined by the typical investor, and so the model may not be so bad. Indeed it has stood the test of time.

The model is used to estimate the future price of the stock and in particular to price options on the stock which are sold. An example of an option is the perpetual American option in which the buyer (who expects the price to rise) pays a certain amount of money V (to be determined) and in turn receives the right to obtain at any time t of his choice a share of the stock at an agreed price k. His profit is then $(X_t - k)^+$. What is V, the "fair price" the broker should ask for this option from the buyer?

One way to determine V is to solve a certain optimal stopping problem, as Paul Samuelson ([S1],1965) posed to Henry McKean, to find:

$$V(x) = V(x; r, \mu, \sigma) = \sup_\tau E_x (X_\tau - k)^+ e^{-r\tau}$$

considering r to be (an estimated) rate at which the value of money declines with time and the supremum taken over all *stopping times*, τ. At the time of stopping the buyer *exercises* his option without of course knowing what the future values of X (or W) will be. McKean's elegant explicit solution to the determination of $V(x)$ is repeated here to illustrate the methods of the subject. We give it in Appendix 1 so as not to interrupt the main flow of the discussion.

It is reasonable to consider $V(x)$ above as the fair price of the option because from the viewpoint of the intelligent buyer, the best he can do would be to choose an optimal time to exercise his option. Thus $V(x)$ plus a profit is what the buyer should willing to pay for the service and also what the broker should be willing to accept as his charge. It should be emphasized that the buyer and the broker should and probably do have different values in mind for μ and σ, especially if one or the other has, or thinks he has, some "inside" knowledge or "information" about the company issuing the stock.

McKean ([S1], Appendix,1965) solved the problem and determined $V(x)$, but not long afterwards, Paul Samuelson rejected his own model for pricing options, replacing it by another method in which the price is determined not by solving an optimal stopping problem but instead on the basis of an *explicit* assumption that *information* plays no role. The idea is that any non-uniform distribution of information is quickly removed by the market itself, and the price of a stock already incorporates all the information because it is determined by the market. It is further assumed that there is no arbitrage and this means that there must be a self-financing portfolio of stocks and bonds with an equivalent performance to the option. In an efficient market it is argued that if it were possible to take a position in stocks and bonds such that no matter what happened a profit would be made with no risk then this would have already happened, some “arbitrageur” would have already extracted the profit, and therefore it is safe to assume that no such “free lunch” can exist. This automatically makes μ the “riskless rate of interest”, and ensures that X_t is a martingale itself. Any bets made in the market are fair and there is no value to information because it has already been exploited. You may not know the facts but the market knows them and so all bets are martingales.

Mathematicians seem to have largely agreed with the new point of view of Samuelson ([S2],1973) which was the brain-child of Robert Merton([S2],1973,Appendix). Mathematicians were pleased that martingales were elevated to center stage in Merton's model and this may have caused them to forego the skepticism that mathematicians often bring to modelling discussions. The new Merton model was widely accepted and option pricing was widely based on the Merton picture.

How does one incorporate the role of information into a model of stock fluctuations in which fair prices might instead be influenced by information disparity?

Section 2. A new model based on *information* .

We modify the Black-Scholes model to incorporate the existence of increased or non-uniform information at certain times by using a hidden Markov model for the fluctuations of the price X_t of a stock in the form

$$dX_t = X_t(\mu_{\epsilon(t)}dt + \sigma_{\epsilon(t)}dW_t)$$

where $\epsilon(t)$ is a new stochastic process which represents the *state of information* in the investor community. Thus $\epsilon(t) = 0$ at those times t at which everyone has the same information. But the process ϵ may take other values than zero. We assume that $\epsilon = \epsilon(t)$ is a Markov process which moves among a few (2 or 3) states. $\epsilon(t) = 1$ means that there is a group of people with inside information at time t and μ_1 may be larger or smaller than μ_0 depending on the nature of inside information, therefore this state may divide into two two extra states where they believe the company will prosper or will decline. Human nature dictates that $\sigma_1 > \sigma_0$ because the inside group will cause more activity in the stock. For each state, i, there is a *known* drift parameter μ_i and a known volatility parameter σ_i estimated via modelling considerations and on past observational data. If we assume that the σ's are distinct then it is no loss of generality to assume that $\epsilon(t)$ is actually observable, because the local quadratic variation of X_t in any small interval to the left of t will give $\sigma_{\epsilon(t)}$ exactly and if the values of σ_i are distinct, we may assume that σ and hence $\epsilon(t)$ is observable. One could argue for a special state corresponding to *disinformation*: some inside groups may actually be misled and so one might want to include a state which would indicate that there is a group of investors who *erroneously believe* that the company's fortunes are going to change for the positive, and another state for the negative. Note that μ_i cannot always be taken as the "riskless rate of interest" as in the Black-Scholes theory because the new model does mot to assume any such riskless rate. Instead however μ_i, σ_i and the transition rates to leave each information state have to be estimated on the basis of past data on the particular stock, or on other considerations of modeling.

The fair price of the American, Russian and European options has been obtained for this presumably more accurate *incomplete* market model, in ([Guo],1999). The method used there is quite similar to that used by McKean described at the end of my paper except the details are much more complicated and the optimal τ is not nearly as simple as a threshold rule as in ([S1], Appendix,1965). For the case of the Russian option, described below, the pricing for the complete market model was given in ([Sh1],1993). We hope that the new prices show greater accuracy in retrospective statistical studies but this remains to be investigated.

To further expand on the comparison of the role of information here and in communication theory, one might well assign the *quantitative value* of knowing that say $\epsilon = 1$ as the difference in value between the calculated of a hedge option with $\epsilon = 1$ and the same option with $\epsilon = 0$, given by Guo's formula, $V(x, 1) - V(x, 0)$. Another quantitative sense of quantifying the value of information would be based on the difference $EV_{new}(x, \epsilon) - V_{old}(x)$ in the values of an option calculated on the simple Black-Scholes-Merton-Samuelson model and that obtained in ([Guo]). Either of these values assigned to information would be an intrinsic definition of information which might play a role in finance (of course only crudely) similar to the role in communication theory played by Shannon's concepts of information and channel capacity. The retrospective study might demonstrate the possibity of extracting profit from knowing the new value.

An interesting alternative theory of option pricing based exclusively on entropy considerations has been given by Gulko ([Gulko]).

Let's look in particular at one such hedge option to see how the V's are calculated.

Section 3. The Russian option.

The Russian option is an American type option. It is a contract that allows the owner to choose an exercise date, represented by the stopping time τ and then pays the owner either s or the maximum stock price achived up to the exercise date, whichever is larger, discounted by e^{-rt} where r is chosen by the agreement between the buyer and the broker. In other words, the Russian option allows you to *reduce your regret* if you miss the peak point of the stock price in the sense that it allows you to choose the highest *past* price of the path, with a reasonable discount by e^{-rt}. To our knowledge, this very reasonable hedge option has never actually been offerred by any firm. There is both a long-term buy ([Sh1]) and a short sell ([Sh2]) version, priced under the old model.

The owner of the option will naturally seek an exercise strategy that will maximize the expected value of this future reward. Mathematically, it is the following question, Let $X = \{X_t, t \geq 0\}$ be the price process for a stock, with $X_0 = x > 0$. Given a constant $s \geq x$, let $S_t = \max\{s, \sup_{0\leq u\leq t} X_u\}$, then what is

$$V(x, s, r, \mu, \sigma) = \sup_{\tau} E_x e^{-r\tau} S_\tau ?$$

Note that it is necessary to consider the case where $x < s$ even though we may start with $x = s$ because after an instant, we have $X_t < S_t$. The exponential or memoryless discounting eliminates t and $\{(x, s)\}$ is the phase space for this control problem. Inside the new model, the phase space must also include the state ϵ and is $\{(x, s, \epsilon)\}$. It is a well-posed mathematical problem, and the solution for the *incomplete* market model has been found ([Guo]).

Earlier, Shiryaev and I obtained the fair value of the Russian option based on the usual Black-Scholes model. The value we obtained was simple,

Assume $r > \max(0, \mu)$ and let $\gamma = \gamma_1$, $\gamma = \gamma_2$, $\gamma_1 < 0 < 1 < \gamma_2$ be the two roots of the quadratic equation

$$\tfrac{1}{2}\sigma^2\gamma^2 + \gamma(\mu - \tfrac{\sigma^2}{2}) = r, \quad \gamma_{1,2} = \frac{\frac{\sigma^2}{2} - \mu \mp \sqrt{(\frac{\sigma^2}{2} - \mu)^2 + 2\sigma^2 r}}{\sigma^2}$$

and set

$$\alpha = \left(\frac{1-\frac{1}{\gamma_1}}{1-\frac{1}{\gamma_2}}\right)^{\frac{1}{\gamma_2-\gamma_1}}.$$

Then

$$\hat{V}(x,s) = \begin{cases} \frac{s}{\gamma_2-\gamma_1}\left(\gamma_2\left(\frac{\alpha x}{s}\right)^{\gamma_1} - \gamma_1\left(\frac{\alpha x}{s}\right)^{\gamma_2}\right), & \frac{s}{\alpha} \le x \le s \\ s, & 0 < x \le \frac{s}{\alpha} \end{cases}$$

is the answer to the above problem, i.e., $V(x,s) \equiv \hat{V}(x,s)$.

The basic idea in the proof in ([Sh1]) is the same as McKean's proof in Appendix 1 and is based on martingales. Namely one shows that

$$Y_t = e^{-rt}\hat{V}(X_t, S_t)$$

is a supermartingale, and so $EY_\tau \le EY_0$ for every τ. One first shows that $\hat{V}(x,s) \ge s$ for every x, s and then this gives that $\hat{V}$ is an upper bound on V since for every τ we have:

$$E_{x,s}e^{-r\tau}S_\tau = EY_\tau \le EY_0 = \hat{V}(x,s).$$

The difficulty is only in guessing the right function to use for $\hat{V}$. In the same way, ([Guo]) invents the appropriate $\hat{V}$ so that the same argument works for

$$Y_t = e^{-rt}\hat{V}(X_t, S_t, \epsilon(t))$$

and proves that it is a supermartingale (by proving that Y is expectation decreasing, $E[dY_t|\mathcal{F}_t] \le 0$), which gives the bound $V \le \hat{V}$.

She proves that Y_t is uniformly integrable, and then the usual proof shows that the bound $\hat{V}$ is achievable. Note that martingales are still involved and probabilists need not fear obsolescence!

But how does one guess $\hat{V}$? The method used is identical to the method used to solve linear programming problems and *cannot be codified* but relies on intuition and insight into the form of the solution. One can merely give some general clues on how to guess. One clue to aid in guessing involves the so-called "principle of smooth fit". That optimal stochastic control and linear programming are closely related is not generally appreciated, but see ([Tak]). One seeks $\hat{V}$ to satisfy the o.d.e. in the continuation region and the stopping reward in the stop region:

$$r\hat{V}(x,s) = \mu x \hat{V}_x(x,s) + \tfrac{1}{2}\sigma^2 x^2 \hat{V}_{xx}, \qquad f(s) < x < s$$

while $\hat{V}(x,s) = s$ for $x \le f(s)$. We need $\hat{V}_s(x,s) = 0$ along $x = s$ to make Y_t into a supermartingale so we impose this, and finally, we impose *just enough* smoothness along $x = f(s)$ to uniquely determine $\hat{V}$. As shown in ([Sh1]) this makes f a linear function.

Later it was argued in ([Sh3]) that there is an easier method of solution based on the fact that $\frac{S_t}{X_t}$ is an embedded Markov process (which is true by Levy's theorem). I now believe that this argument works *only* because the answer, $\hat{V}$, happens to involve $\frac{s}{x}$. I try to clarify this point in Appendix 2, where some new variants of the Russian option are treated by giving examples where one can solve the problem but the embedded Markov process $\frac{S_t}{X_t}$ plays no role. On the other hand, ([Guo]) makes good use of $\frac{S_t}{X_t}$, which is again a Markov process even with the hidden Markov process $\epsilon(t)$ in place, in order to simplify her proof of optimality and also to do numerical calculations as a check. A broad class of problems similar to those in the appendix have been treated in ([Peskir]).

Section 4. A remark on completing the market.

There may be a way out for the non-arbitrage school as was kindly pointed out by Darrell Duffie in personal communication. Duffie proposed that a tool, called a "ticket", could or would be devised to *complete* the market. Suppose that at each time t, suppose there is a market for a security that pays a dollar at the next time $\tau(t) = \inf\{u > t : \epsilon(u) \neq \epsilon(t)\}$ that the Markov chain ϵ changes state. Thereafter this contract becomes worthless (has no future dividends), and a new contract is issued that pays at the next change of state. It is easy to see the market is complete with this new ticket, which actually functions to eliminate the possible difference of information. Whether or not Duffie's clever idea resurrects the non-arbitrage assumption depends on whether one can find someone to market such a ticket. This seems somewhat doubtful to me, although the final verdict remains to be obtained.

Section 5. A remark on Nash theory of valuation of multi-person games.

Von Neumann theory is useful for assigning a value to a two person zero sum game and to pricing stock options via stochastic optimal control theory as indicated above. Merton's method assigns a value to an option but it is based on a dubious model of reality. An analogy can be drawn to the problem of assigning a value to a game which is not two-person zero sum based on an axiomatic approach called equilibrium theory.

How does one assign a value to a non-zero sum or multiperson game? There is no linear programming method, and some games are very difficult to understand. In such cases an axiomatic approach has been suggested by John Nash, and the value assigned is the Nash equilibrium, if such exists. As with the value assigned to options by the axioms of completeness and non-arbitrage, relying on axiomatics can lead to results that do not seem real if the axioms do not apply so well.

Consider the well-known multi-person game where a rich man gives a million dollars to 3 "friends" if they can agree by majority rule on how to divide it. This game is very unstable and although one can devise axioms which suggests that the value to each of the 3 friends ought to be $\frac{1}{3}$, it seems that the decision rule will strongly depend on the psychological stability of the 3 friends who may well turn out to be enemies.

The moral of both stories is that axiomatic pricing is only as good as the reality of the underlying axioms, both for Nash theory in the above case and for pricing options by assuming the market is free of arbitrage.

Appendix 1. McKean's solution to Samuelson's problem.

Samuelson asked McKean to find:

$$V(x) = V(x; r, \mu, \sigma) = \sup_{\tau} E_x (X_\tau - k)^+ e^{-r\tau}$$

McKean guessed the answer and used martingale theory to prove it. It's clear that one would never exercise when $X_t \leq k$ and so one guesses that there is some $a > k$ such that the optimal $\tau = \tau_a$ the first time the stock price gets to the level a. If this is so then the value denoted by $\hat{V}(x)$ while we are just guessing must satisfy $\hat{V}(x) = x - k$ for $x \geq a$. Also for $x < a$ we ought to play for some small time $dt = h$ and then the average reward would be $\hat{V}(x) = Ee^{-rh}\hat{V}(x + dx)$. Since $dx = x(\mu dt + \sigma dw)$, we see that $\hat{V}$ ought to satisfy:

$$\begin{aligned} r\hat{V}(x) &= \hat{V}'(x)x\mu + \hat{V}''(x)x^2\sigma^2, \qquad && x < a \\ \hat{V}(x) &= x - k, && x > a \end{aligned}$$

but this is really a guess at the *form* of the optimal strategy. We see that

$$\hat{V}(x) = Ax^{\gamma_+} + Bx^{\gamma_-}, \qquad 0 < x < a$$

where γ_+, γ_- are the two roots of the indicial equation,

$$0 = -r + \gamma\mu + \tfrac{\sigma^2}{2}\gamma(\gamma - 1).$$

It's easy to see that we may take $\gamma_+ > 0 > \gamma_-$. The negative root, γ_- is discarded because we need $\hat{V}(x)$ bounded near $x = 0$, and so $B = 0$.

The smooth fit principle now completes the guessing procedure and determines both A and a. We need $\hat{V}$ to be in C_1 at $x = a$ and this gives the two equations for a and A:

$$Aa^{\gamma_+} = a - k, \qquad A\gamma_+ a^{(\gamma_+ - 1)} = 1$$

As long as $r < \mu$ there is a solution, $a = \frac{k}{1 - \frac{1}{\gamma_+}}$ with $a > k$, and then $A = \frac{a^{(\gamma_+ - 1)}}{\gamma_+}$.

Now the proof is easy that $V(x) \equiv \hat{V}(x)$. Simply verify that the process $Y_t = \hat{V}(X_t)e^{-rt}$ is a supermartingale, that is Y_t is expectation decreasing. The supermartingale inequality now states that for any stopping time τ, $EY_\tau \leq EY_0$. Since it's easy to check that $\hat{V}(x) \geq x - k$ for any $0 < x < \infty$ it follows that for any τ, $E_x(X_\tau - k)^+ e^{-rt} \leq EY_\tau \leq EY_0 = \hat{V}(x)$. This says that $V(x) \le \hat{V}(x)$ for all x. If we use the optimal rule then equality holds throughout by the fact that Y_t is a uniformly integrable local martingale and so also $V(x) \le EY_\tau = \hat{V}(x)$. qed

Remark. It seems clear that the finite horizon ($\tau \le T$) version of Samuelson's problem is not solvable in closed form. It is the memorylessness of the exponential discounting that makes it solvable.

Appendix 2. A comparison of problems related to the Russian option.

Problem 1. Find for $w = W(0), t > 0$

$$V(w,t) = \sup_\tau E_{w,t} \frac{W_\tau}{\tau}$$

where the supremum is taken over the set of all stopping times τ and W_t is the ordinary Brownian motion process. The solution is given in ([Sh],1969) and takes the form τ is the first time t that $W_t \ge \alpha\sqrt{t}$ (which may be $\tau = 0$). Here $\alpha = .83992\ldots$ is seen from the smooth fit principle to be the root of $\alpha = (1-\alpha^2)\int_0^\infty e^{\alpha u - \frac{u^2}{2}} du$.

Problem 2. Find for $w \le m, t > 0$

$$V(w,m,t) = \sup_\tau E_{w,m,t} \frac{M_\tau}{\tau}$$

where W_t is the ordinary Brownian motion as in problem 1, and

$$M_t = max(s, \sup_{0 \le u \le t} W_u).$$

As in the finite horizon case discussed in the Remark in Appendix 1, Problem 2 is surely impossible to solve in terms of elementary functions because V depends on 3 parameters and the optimal stopping rule is of the form: stop the first time t that $W_t > f(M_t, t)$ but f is given by the solution of a p.d.e.

Problem 3. Find

$$V(w) = \sup_\tau E_w W_\tau e^{-r\tau}$$

where X is as in problem 1. The exponential discounting makes the problem "memoryless", t can be ignored, and the problem is easy. The optimal τ is the first t such that X_t exceeds a threshold a. Since the optimal $\hat{V}$ satisfies $\frac{1}{2}\hat{V}''(w) = r\hat{V}(w)$ or $\hat{V}(w) = w$, this prompts the guess that $V(w) = Ae^{\sqrt{2r}w}$ for some A and for $w < a$ and $\hat{V}(w) \equiv w, w \geq a$, where A and a are to be found by the smooth fit principle. It's easy to verify that the same argument as in Appendix 1, works if we take $a = \frac{1}{\sqrt{2r}}$ and $A = ae^{-\sqrt{2r}}$.

Problem 4. Find

$$V(w, m) = \sup_\tau E_{w,m} M_\tau e^{-r\tau}$$

where W and M are as in problem 2. This is a problem of "intermediate" difficulty as discussed below and one seeks a solution of the form: τ is the first t for which $W_t \leq f(M_t)$ where f is a function to be determined. The same technique we used above says that in the region $\{w \leq f(m)\}$ we must have $\hat{V}(w, m) = m$ since we stop there, and in the region $\{f(m) < w < m\}$ we must have $\frac{1}{2}\hat{V}_{ww} = r\hat{V}$. This means that we guess the form

$$\hat{V}(w, m) = m \cosh(w - f(m)), \qquad f(m) \leq w \leq m; \qquad \hat{V}(w, m) = m, w \leq f(m).$$

The smooth fit principle says that we need $\hat{V}$ to be smooth along $w = f(m)$, where the degree of smoothness is just enough to make $\hat{V}$ determined. It is also necessary to impose the condition

$$\hat{V}_m(w, m) = 0 \text{ along } w = m$$

as in the Russian option to make Y_t a martingale along $w = m$. This guesswork all easily leads to the o.d.e. for f,

$$amf'(m) \tanh(m - f(m)a) \equiv 1$$

where $a = \sqrt{2r}$. The transformation $f(m) = \frac{g(am)}{a}$ reduces the problem to the case $a = 1$, but the solution still appears difficult. The boundary condition is that $m - f(m) \to 0$ as $m \to \infty$.

The problem becomes nicer in terms of the function $\hat{V}(m) = \hat{V}(m, m)$. This satisfies $\hat{V}(m) = m \cosh(m - f(m))$, where, since we have $a = 1$, f satisfies the o.d.e.: $mf'(m) \tanh(m - f(m)) \equiv 1$. Now a simple verification shows that $\hat{V}(m) = m \cosh(m - f(m))$ satisfies the simpler o.d.e

$$\hat{V}'(m) = \sqrt{\hat{V}^2(m) - m^2}$$

and the boundary condition $\hat{V}(m) - m \to 0$ as $m \to \infty$. This is a very simple o.d.e., but apparently not simple enough to be solved.

The problem is rather more difficult than that of ([Sh1]) even though $M_t - W_t$ is a Markov process which does not seem to help here. The same observation was *claimed* in ([Sh3]) to give a simpler solution to that of ([Sh1]) for the Russian option. I want to argue here that this trick was successful in ([Sh3]) *only* because the answer had a simple form. I think Problem 4 demonstrates this point since the same trick works here but does not help at all. Note further that for the problem of the additive Bachelier type Russian option the problem is simple again:

Problem 5. Find

$$V(w,m) = \sup_{\tau} E_{w,m}(M_\tau - r\tau)$$

where W and M are as in problem 4. The optimal τ is the first t for which $W_t \leq M_t - \theta$, as shown in ([Sh1]). Now the embedded Markov process technique works but it's only because the answer has the right form, that is $\hat{V}(w,m)$ is a function of $m - w$. But we see in Problem 4 that if one uses exponential discounting instead of additive discounting then the fact that $M_t - W_t$ is Markovian no longer helps because $\hat{V}$ does not depend on the difference of m and w.

A free boundary problem of similar type was completely solved in ([BSW]) but the present o.d.e., $V'(m) = \sqrt{V^2(m) - m^2}$, as simple as it looks, seems to be more difficult than the o.d.e. in ([BSW]), and does not seem to be explicitly expressible in terms of standard functions. It appears that the series approximation at $m = \infty$,

$$V(m) = m + \tfrac{1}{2m} + \tfrac{a_3}{m^3} + \tfrac{a_4}{m^4} + \ldots$$

is only asymptotic and not a power series and is apparently divergent.

Finally, the more general class of similar problems replacing in Problem 4 M_t by $\phi(M_t)$ has special cases which *can be solved explicitly in closed form* (this is why I used the term "intermediate" difficulty in the above discussion):

Problem 6. Find

$$V(w,m) = \sup_{\tau} E_{w,m}\phi(M_\tau)e^{-r\tau}$$

where W and M are as in problem 4, and ϕ is a function to be specified later. The solution for an arbitrary ϕ is guessed by the same method as that of problem 4 as the solution to:

$$\tfrac{1}{2}\hat{V}_{ww} = rV \text{ in } \{f(m) < w < m\}$$

and $\hat{V}(w,m) = \phi(m)$ in $\{w \leq f(m)\}$ since we stop there.

The smooth fit principle says that we need $\hat{V}$ to be smooth along $w = f(m)$, and we again need $\hat{V}_(w,m) = 0$ along $w = m$. This again gives the o.d.e. for f, $a\phi(m)f'(m)\tanh(m - f(m)a) \equiv \phi'(m)$, where $a = \sqrt{2r}$. The transformation $f(m) = \frac{g(am)}{a}$ again reduces the problem to the case $a = 1$. The boundary condition is again that $m - f(m) \to 0$ as $m \to \infty$.

The problem again becomes nicer in terms of the function $\hat{V}(m) = \hat{V}(m,m)$. This satisfies $\hat{V}(m) = \phi(m)\cosh(m - f(m))$, and, since we have $a = 1$, $\hat{V}$ satisfies the simple o.d.e.

$$\hat{V}'(m) = \sqrt{\hat{V}^2(m) - \phi(m)^2}$$

and the boundary condition $\hat{V}(m) - m \to 0$ as $m \to \infty$.

This is a very simple o.d.e., but is made even simpler if we regard $\hat{V}$ as given and this as an *algebraic equation* for $\phi(m)$. Of course we need to specify a $\hat{V}$ with all the nice conditions to make the martingale argument work, but this is not hard to do. Indeed, we see that once $\hat{V}(m)$ is specified, then

$$\phi(m) = \sqrt{\hat{V}^2(m) - \hat{V}'(m)^2},$$

and

$$f(m) = m - \cosh^{-1}\left(\tfrac{\hat{V}(m)}{\phi(m)}\right).$$

Now we set

$$\hat{V}(w,m) = \phi(m)\cosh(w - f(m))$$

and then

$$\hat{V}'(w,m) = \phi(m)\sinh(w - f(m))$$

It is easy to see, parallelling the proof in Appendix 1, that the process

$$Y_t = e^{-rt}\hat{V}(W_t, M_t)$$

is a supermartingale and since $\hat{V}(w, m) \geq \phi(m)$ because $\cosh \geq 1$, it follows easily that

$$V(w, m) \leq \hat{V}(w, m) \text{ for all } w, m.$$

To prove that the last inequality is an equality we need the boundary condition on V or equivalently of f,

$$\begin{gathered} V(m) - m \to 0 \text{ as } m \to \infty \\ m - f(m) \to 0 \text{ as } m \to \infty. \end{gathered}$$

More precisely, what is needed is that the equality $V(w, m) = \hat{V}(w, m)$ holds and to prove this it is sufficient that Y_t stopped at $t = \tau$ where τ is the first t for which $W_t \leq f(M_t)$ is a uniformly integrable martingale. Then equality holds in the chain

$$EY_\tau = E_{w,m} e^{-r\tau}\phi(M_\tau) = EY_0 = \hat{V}(w, m)$$

and so $V(w, m) \geq \hat{V}(w, m)$. Thus $V \equiv \hat{V}$.

$$E_{w,m} \sup\nolimits_{t \leq \tau} e^{-rt}\hat{V}(W_t, M_t) < \infty$$

Since $\hat{V}(w, m) \leq \phi(m)\cosh(m - f(m))$ because cosh increases in $w > 0$, we need enough conditions on $\phi(m)$, $f(m)$, $V(m)$ to ensure this.

An example where this all can be carried out and all the variables found explicitly is:

$$\begin{aligned} V(m) &= \sqrt{m^2 + m + \tfrac{1}{2}}, && m \geq 0 \\ f(m) &= m - \log(1 + \tfrac{1}{m}), && m \geq 0 \\ \phi(m) &= \tfrac{m(m+1)}{\sqrt{m(m+1) + \frac{1}{2}}}, && m \geq 0. \\ V(w, m) &= \phi(m)\cosh(w - f(m)), && f(m) \leq x \leq s \end{aligned}$$

and $V(w,m) = \phi(m)$ for $w \leq f(m)$. Note $f(m) \to -\infty$ as $m \to 0$. Note that $\phi(m)$ is like m in that it vanishes at $m = 0$ and is asymptotic to m as $m \to \infty$.

If we define $\hat{V}(w,m) = V(w,m)$ given by the explicit formule above, these definitions ensure that the process $Y_t = e^{-rt}\hat{V}(W_t, M_t), t \geq 0$ is a supermartingale and that $Y_t, t \leq \tau$ is a uniformly integrable martingale if τ is the hitting time when first $W_t = f(M_t)$. Thus $\hat{V}(w,m)$ is the solution to Problem 6 for this particular choice of ϕ, i.e., $V = \hat{V}$.

No imbedded Markov process was used or needed to prove this result. It is not true that the process $Z_t = W_t - f(M_t), t \geq 0$ is an embedded Markov process. I think this example shows clearly that the fact that the embedded Markov process $\frac{S_t}{X_t}$ helped in solving the problem for the Russian option in both the new or the old model for the exponential Brownian motion has more to do with coincidence than a general rule. It's nice but not really necessary to solve the Russian option valuation problem in the way given in ([Sh3]).

References

[B] **Bachelier, L.** *Theorie de la speculation*, Ann. Ecole Norm. Sup., 17 (1900), pp. 21-86. (Reprinted in: *The Random Character of Stock Market Prices*, ed. by P.H.Coothner, MIT Press Cambridge, Mass. 1967, pp. 17-78).

[BSW] **Benes, V.E., Shepp, L.A., Witsenhausen, H.S.** *Some Solvable Stochastic Control Problems* Stochastics, 4 (1980), pp. 39-83.

[S1] **Samuelson, P.A.** *Rational Theory of Warrant Pricing*, Ind. Mngt. Rev., 6, (1965), pp.13-31, Appendix by **McKean, H.P.**,pp. 32-39.

[S2] **Samuelson, P.A.** *Mathematics of Speculative Price*, SIAM Rev., 15, (1973), pp.1-34, Appendix by **Merton, R.C.**, pp. 34-42.

[Gulko] **Gulko, Les** *The Entropy Theory of Bond Option Pricing*. ms.

[Guo] **Guo, X.** *Inside Information and Stock Fluctuations, PhD Thesis*, Rutgers U., 1999

[Peskir] **Peskir, Goran** *Optimal Stopping of the Maximum Process: The Maximality Principle* Ann. Prob., 26, (1998), pp. 1614-1640.

[Sh] **Shepp, L.A.** *Explicit Solutions to Some Problems of Optimal Stopping* Ann. Math. Statist, 40, (1969), pp. 993-1010.

[Sh1] **Shepp, L.A., Shiryaev, A.N.** *The Russian Option: Reduced Regret*, Annals of Applied Probability, 3, (1993), pp. 631-640.

[Sh2] **Shepp, L.A., Shiryaev, A.N.** *A Dual Russian Option for Selling Short* Probability Theory and Mathematical Statistics, (1996) Ibragimov, I.A. and Zaitsev, A.M., Gordon and Breach pp. 209-218.

[Sh3] **Shepp, L.A., Shiryaev, A.N.** *A New Look at the Russian Option* Theory Probability and Applications, 39 (1994), pp. 103-119.

[Tak] **Taksar, M.I.** *Infinite-dimensional linear programming approach to singular stochastic control* SIAM J. Control Optim., 35 (1997), pp. 604-625.