

A GENERAL ‘BANG–BANG’ PRINCIPLE FOR PREDICTING THE MAXIMUM OF A RANDOM WALK

PIETER ALLAART,* *University of North Texas*

Abstract

Let $(B_t)_{0 \leq t \leq T}$ be either a Bernoulli random walk or a Brownian motion with drift, and let $M_t := \max\{B_s : 0 \leq s \leq t\}$, $0 \leq t \leq T$. In this paper we solve the general optimal prediction problem $\sup_{0 \leq \tau \leq T} E[f(M_T - B_\tau)]$, where the supremum is over all stopping times τ adapted to the natural filtration of (B_t) and f is a nonincreasing convex function. The optimal stopping time τ^* is shown to be of ‘bang–bang’ type: $\tau^* \equiv 0$ if the drift of the underlying process (B_t) is negative and $\tau^* \equiv T$ if the drift is positive. This result generalizes recent findings of Toit and Peskir (2009) and Yam, Yung and Zhou (2009), and provides additional mathematical justification for the dictum in finance that one should sell bad stocks immediately, but keep good stocks as long as possible.

Keywords: Bernoulli random walk; Brownian motion; optimal prediction; ultimate maximum; stopping time; convex function

2010 Mathematics Subject Classification: Primary 60G40; 60G50; 60J65
Secondary 60G25

1. Introduction

Suppose that you buy a share of stock at time $t = 0$, which you must sell on or before some finite time $T > 0$. Your goal is to sell the stock at such a time τ so as to maximize the expectation of some function g of the ratio P_τ/H_T , where P_t denotes the price of the stock at time t and $H_T := \max\{P_t : 0 \leq t \leq T\}$, the highest price over the time interval $[0, T]$. What is the optimal selling time?

Under the model of geometric Brownian motion and with g the identity function, this problem was formulated and partially solved by Shiryaev *et al.* [5], and was soon afterwards solved completely by Du Toit and Peskir [1]. The solution is as simple as it is natural: if the logarithm of the stock price trends upward, the stock should be held until the time horizon, and if it trends downward, the stock should be sold immediately. Thus, in the words of Du Toit and Peskir, the optimal strategy is of ‘bang–bang’ type. Shortly after their paper appeared, an analogous result was obtained by Yam *et al.* [6] for exponentiated random walks in discrete time.

The purpose of the present paper is to generalize these results to a larger class of functions g . It will be shown that the ‘bang–bang’ principle holds, under both the geometric Brownian motion and exponentiated random walk models, whenever g is increasing and convex. Rather than directly considering price ratios, however, it is more convenient to study the logarithms of the price process. For instance, if $\{P_t\}$ is a geometric Brownian motion then we can write

Received 10 February 2010; revision received 28 June 2010.

* Postal address: Department of Mathematics, University of North Texas, 1155 Union Circle #311430, Denton, TX 76203-5017, USA. Email address: allaart@unt.edu

Supported in part by the Japanese GCOE Program G08: ‘Fostering Top Leaders in Mathematics—Broadening the Core and Exploring New Ground’.

$P_t = \exp(\sigma B_t^\lambda)$, where B_t^λ is Brownian motion with a suitable drift $\lambda \in \mathbb{R}$, and $\sigma > 0$. Setting $M_t^\lambda := \max\{B_s^\lambda : 0 \leq s \leq t\}$ and $f(x) = g(e^{-\sigma x})$, the problem described in the first paragraph can then be put in the form

$$\sup_{0 \leq \tau \leq T} E[f(M_T^\lambda - B_\tau^\lambda)], \tag{1.1}$$

the supremum being over all stopping times $0 \leq \tau \leq T$ adapted to the process $\{B_t^\lambda\}$. It is not difficult to verify that if g is nondecreasing and convex, then f is nonincreasing and convex. A similar transformation can of course be made for random walks in discrete time. Thus, for the remainder of this paper we will study (1.1) and its discrete-time analog (2.1) below for a nonincreasing convex f , keeping in mind that via the above transformation the problem has a practical financial interpretation.

For the specific function $f(x) = e^{-\sigma x}$, where $\sigma > 0$, our results reduce to those of [1] and [6]. The proofs involve only a minimum of technicalities, and bring to the foreground the essential feature hidden within the arguments in the aforementioned papers, namely convexity of the function f . In addition, we present simple conditions on f in order for the optimal stopping rules to be unique.

To end this introduction, we point out that problems of the type (1.1) were first formulated (albeit as a penalty-minimization problem) in the paper by Graversen *et al.* [2], where f was the (nonconvex) function $f(x) = -x^2$. Their results were extended by Pedersen [4] to $f(x) = -x^q$ for arbitrary $q > 0$. Note that, for $0 < q < 1$, Pedersen’s result is a special case of Theorem 2.3 below.

2. Main results

This section is devoted to a precise formulation of the problem and statements of the main results. First, let $\{S_n\}_{n=0,1,\dots}$ be a Bernoulli random walk with parameter $p \in (0, 1)$. That is, $S_0 \equiv 0$, and, for $n \geq 1$, $S_n = X_1 + \dots + X_n$, where X_1, X_2, \dots are independent, identically distributed random variables with $P(X_1 = 1) = p$, and $P(X_1 = -1) = q := 1 - p$. Let a finite time horizon $N \in \mathbb{N}$ be given, let $f: \{0, 1, \dots, N\} \rightarrow \mathbb{R}$ be nonincreasing, and consider the optimal stopping problem

$$\sup_{0 \leq \tau \leq N} E[f(M_N - S_\tau)], \tag{2.1}$$

where $M_N := \max\{S_0, S_1, \dots, S_N\}$ and the supremum is over the set of all stopping times $\tau \leq N$ adapted to the natural filtration $\{\mathcal{F}_k\}_{0 \leq k \leq N}$ of the process $\{S_k\}_{0 \leq k \leq N}$.

As a concrete example, taking $f(0) = 1$ and $f(k) = 0$ for $k \geq 1$ turns the expectation in (2.1) into the probability $P(S_\tau = M_N)$, so that (2.1) becomes a ‘best-choice’ or ‘secretary’ problem for the random walk, where the goal is to maximize the probability of stopping at the ultimate maximum of the walk; see [3], where this problem is solved in a somewhat more general setting for the case $p = \frac{1}{2}$. Yam *et al.* [6] recently solved the problem for arbitrary p , and showed the (unique) optimal rule to be $\tau \equiv 0$ when $p < \frac{1}{2}$ and $\tau \equiv N$ when $p > \frac{1}{2}$. When $p = \frac{1}{2}$, it is optimal to stop at time 0, or at time N , or at any time at which the walk is at its running maximum. (For a continuous-time analog of this problem, where the objective is to stop a Brownian motion within a distance $\varepsilon > 0$ from its ultimate maximum, see [4].)

Curiously, as shown by Yam *et al.* [6] in the second half of their paper, the same rule is optimal when $f(k) = d^k$ for a constant $0 < d < 1$. This leads one to believe that there must be some general principle at work. A brief look at the graphs reveals that in both examples, f is in fact convex. The first aim of this paper is to show that the optimal rule is of the above simple

form for any nonincreasing convex objective function f , thereby generalizing the results of [6]. Recall that a function $f: \{0, 1, \dots, N\} \rightarrow \mathbb{R}$ is *convex* if $f(k - 1) - 2f(k) + f(k + 1) \geq 0$ for all k with $0 < k < N$, and is *strictly convex* if the inequality is strict for all such k .

Theorem 2.1. *Let $f: \{0, 1, \dots, N\} \rightarrow \mathbb{R}$ be nonincreasing and convex, and consider the optimal stopping problem (2.1).*

- (i) *If $p \leq \frac{1}{2}$, the rule $\tau \equiv 0$ is optimal.*
- (ii) *If $p \geq \frac{1}{2}$, the rule $\tau \equiv N$ is optimal.*
- (iii) *If $p = \frac{1}{2}$, any stopping time τ satisfying $P(S_\tau = M_\tau \text{ or } \tau = N) = 1$ is optimal.*

Thus, in the words of Du Toit and Peskir [1], the optimal strategy τ^* is of ‘bang–bang’ type: $\tau^* \equiv 0$ if $p < \frac{1}{2}$ and $\tau^* \equiv N$ if $p > \frac{1}{2}$.

Convexity of f is essential, as the following example shows.

Example 2.1. Let $f(0) = f(1) = 1$ and $f(k) = 0$ for $k \geq 2$. Thus, there are two possible outcomes, ‘winning’ and ‘losing’, and we win if we stop with one of the two highest values of the walk. Let $N = 2$. It is easy to see that the rule $\tau \equiv 1$ gives a winning probability of 1. On the other hand, the winning probability for the rule $\tau \equiv 0$ is $1 - p^2$ and, for the rule $\tau \equiv 2$, it is $1 - q^2$.

We might ask when the optimal rules in Theorem 2.1 are unique. The next theorem gives simple sufficient conditions to this effect.

Theorem 2.2. *Let f be as in Theorem 2.1.*

- (i) *If $p < \frac{1}{2}$ and f is nonconstant, then $\tau \equiv 0$ is the only optimal rule.*
- (ii) *If $p > \frac{1}{2}$ and f is strictly decreasing, then $\tau \equiv N$ is the only optimal rule.*
- (iii) *If $p = \frac{1}{2}$ and f is strictly convex, then the only optimal rules are those that satisfy $P(S_\tau = M_\tau \text{ or } \tau = N) = 1$.*

It is left to the interested reader to verify that the above conditions cannot be substantially weakened.

Next, let $B := (B_t)_{t \geq 0}$ be a standard Brownian motion, and let λ be a real parameter. Then the process $(B_t^\lambda)_{t \geq 0}$ defined by $B_t^\lambda := B_t + \lambda t$ is a Brownian motion with drift λ . Let $M_t^\lambda := \max\{B_s^\lambda : 0 \leq s \leq t\}$. We seek a stopping time τ (with respect to the natural filtration $(\mathcal{F}_t^B)_{t \geq 0}$ of B) that will attain the maximum in (1.1). Since Brownian motion is the scaling limit of the Bernoulli random walk, we might expect the result to be the same as in Theorem 2.1. This is indeed the case, except that the conditions for uniqueness of the optimal rules are weaker.

Theorem 2.3. *Let $f: [0, \infty) \rightarrow \mathbb{R}$ be nonconstant, nonincreasing, and convex, and consider the optimal stopping problem (1.1).*

- (i) *If $\lambda < 0$, the rule $\tau \equiv 0$ is the unique optimal rule.*
- (ii) *If $\lambda > 0$, the rule $\tau \equiv T$ is the unique optimal rule.*
- (iii) *If $\lambda = 0$, any stopping time τ satisfying*

$$P(B_\tau^\lambda = M_\tau^\lambda \text{ or } \tau = T) = 1 \tag{2.2}$$

is optimal. (In particular, the rules $\tau \equiv 0$ and $\tau \equiv T$ are optimal.) If, furthermore, f is not linear, then all optimal rules satisfy (2.2).

Theorem 2.3 was proved in [1] for the case $f(x) = e^{-\sigma x}$, after the problem was first discussed in [5]. Note that if f is constant, or if f is linear and $\lambda = 0$, then any stopping time is optimal in view of the optional sampling theorem. Thus, the uniqueness conditions in Theorem 2.3 are the best possible. That they are weaker than in the discrete case is due essentially to the fact that the increments of Brownian motion can be arbitrarily large, whereas the increments of the Bernoulli random walk are bounded.

Finally, we note that by putting $\tilde{f} := -f$, problems (1.1) and (2.1) may be formulated equivalently as penalty-minimization problems. For instance, (1.1) can be represented alternatively in the form

$$\inf_{0 \leq \tau \leq T} E[\tilde{f}(M_T^\lambda - B_\tau^\lambda)], \tag{2.3}$$

where $\tilde{f}: [0, \infty) \rightarrow \mathbb{R}$ is nondecreasing and concave. Thus, the above results apply to a variety of natural penalty functions, including $\tilde{f}(x) = x^q$, where $0 < q \leq 1$, $\tilde{f}(x) = \log(1 + x)$, etc.

For nonconcave \tilde{f} , the solution of (2.3) is generally of a more intricate form, and is usually found by applying the principle of ‘smooth fit’ and solving an appropriate free boundary problem. See [2] for $\tilde{f}(x) = x^2$, [4] for $\tilde{f}(x) = x^q$ with $q > 1$, or Section 3 of [1] for $\tilde{f}(x) = e^{\sigma x}$, where $\sigma > 0$.

Theorems 2.1 and 2.2 are proved in Section 3, and Theorem 2.3 is proved in Section 4. Many of the ideas of the proofs are adapted from [1] and [6], and some details, in as far as they can be found in these papers, are therefore omitted here. The novel contributions of the present paper are the explicit use of the convexity of f (see Lemmas 3.1 and 4.1 below) and the investigation of uniqueness of the optimal stopping times, which requires some finesse in the case of general f .

3. The maximum of the Bernoulli random walk

This section is devoted to the proofs of Theorems 2.1 and 2.2. It will be useful to consider an infinite family of random walks, defined on the same probability space. The following construction is standard. Let U_1, U_2, \dots be independent random variables, uniformly distributed on the interval $[0, 1]$. For $k \in \mathbb{N}$ and $p \in (0, 1)$, define

$$X_k^p := \begin{cases} 1 & \text{if } U_k \leq p, \\ -1 & \text{if } U_k > p. \end{cases}$$

Set $S_0^p \equiv 0$ and $S_k^p := X_1^p + \dots + X_k^p$ for $k \geq 1$. Then, for each $p \in (0, 1)$, $\{S_k^p\}_k$ is a Bernoulli random walk with parameter p . Furthermore, if $p \geq p'$ then $X_k^p \geq X_k^{p'}$ for all k .

Let $M_k^p := \max\{S_0^p, \dots, S_k^p\}$ and $Z_k^p := M_k^p - S_k^p$. Observe that, for each p , the process $\{Z_k^p\}_k$ is Markovian. Moreover, it is easy to see that

$$p \geq p' \implies Z_k^p \leq Z_k^{p'} \quad \text{for all } k. \tag{3.1}$$

Finally, and most importantly, the Bernoulli random walk satisfies the well-known reflection property

$$(M_n^p - S_n^p, S_n^p) \stackrel{D}{=} (M_n^q, -S_n^q) \tag{3.2}$$

for each fixed $n \in \mathbb{N}$. (The easiest way to see this is to observe that the time-reversed process $\tilde{S}_k := S_{n-k}^p - S_n^p$, $k = 0, 1, \dots, n$, is a Bernoulli random walk with parameter q , starting at

0 and ending at $-S_n^p$, with maximum value $M_n^p - S_n^p$.) In particular (reversing the roles of p and q),

$$M_n^p \stackrel{D}{=} M_n^q - S_n^q = Z_n^q. \tag{3.3}$$

It is almost amusing to see how many times this identity must be used in order to prove Theorem 2.1.

The following lemma holds the key to the proof of Theorem 2.1.

Lemma 3.1. *Let $f: \{0, 1, \dots, N\} \rightarrow \mathbb{R}$ be nonincreasing and convex.*

(i) *If $p \geq \frac{1}{2}$ then*

$$E[f(i \vee M_n^p - S_n^p)] \geq E[f(i \vee (M_n^p - S_n^p))] \tag{3.4}$$

for all $0 \leq n \leq N$ and $0 \leq i \leq N - n$.

(ii) *If $p > \frac{1}{2}$ and f is strictly decreasing, then strict inequality holds in (3.4) for all $0 < n \leq N$ and $0 < i \leq N - n$.*

(iii) *If $p \geq \frac{1}{2}$ and f is strictly convex, then strict inequality holds in (3.4) for all $0 < n \leq N$ and $0 < i \leq N - n$.*

Proof. (i) Let $p \geq \frac{1}{2}$. We begin by writing

$$\begin{aligned} & E[f(i \vee M_n^p - S_n^p) - f(i \vee (M_n^p - S_n^p))] \\ &= \sum_{l \in \mathbb{Z}} \sum_{k \geq l} [f(i \vee k - l) - f(i \vee (k - l))] P(M_n^p = k, S_n^p = l) \\ &= \left(\sum_{l > 0} \sum_{k \geq l} + \sum_{l < 0} \sum_{k \geq 0} \right) [f(i \vee k - l) - f(i \vee (k - l))] P(M_n^p = k, S_n^p = l) \\ &=: \Sigma^+ + \Sigma^-. \end{aligned}$$

(Note that the terms with $l = 0$ vanish.) By (3.2) and the change of variables $k' = k - l$, $l' = -l$, the second summation becomes

$$\begin{aligned} \Sigma^- &= \sum_{l < 0} \sum_{k \geq 0} [f(i \vee k - l) - f(i \vee (k - l))] P(M_n^q = k - l, S_n^q = -l) \\ &= \sum_{l' > 0} \sum_{k' \geq l'} [f(i \vee (k' - l') + l') - f(i \vee k')] P(M_n^q = k', S_n^q = l'). \end{aligned}$$

The key to further progress is that, for $l > 0$,

$$P(M_n^p = k, S_n^p = l) \geq P(M_n^q = k, S_n^q = l).$$

(This follows easily by considering the probability of a single path ending at l with maximum k .) Since f is nonincreasing and $i \vee k - l \leq i \vee (k - l)$, we have

$$f(i \vee k - l) - f(i \vee (k - l)) \geq 0,$$

and, therefore,

$$\Sigma^+ \geq \sum_{l > 0} \sum_{k \geq l} [f(i \vee k - l) - f(i \vee (k - l))] P(M_n^q = k, S_n^q = l). \tag{3.5}$$

Combining these results, we conclude that

$$\Sigma^+ + \Sigma^- \geq \sum_{l>0} \sum_{k \geq l} \psi(i, k, l) P(M_n^q = k, S_n^q = l), \tag{3.6}$$

where

$$\begin{aligned} \psi(i, k, l) &:= [f(i \vee k - l) - f(i \vee (k - l))] + [f(i \vee (k - l) + l) - f(i \vee k)] \\ &= [f(i \vee k - l) - f(i \vee k)] - [f(i \vee (k - l)) - f(i \vee (k - l) + l)]. \end{aligned}$$

Since $i \vee k - l \leq i \vee (k - l)$ and f is convex, it is easy to see that $\psi(i, k, l) \geq 0$. This yields (3.4).

(ii) Suppose that $p > \frac{1}{2}$ and f is strictly decreasing. Let $n > 0$ and $i > 0$, and set $k = l = n$. Then

$$f(i \vee k - l) - f(i \vee (k - l)) = f((i - n)^+) - f(i) > 0.$$

Since $P(M_n^p = S_n^p = n) > P(M_n^q = S_n^q = n)$, strict inequality holds in (3.5), and, hence, in (3.4).

(iii) Finally, suppose that $p \geq \frac{1}{2}$ and f is strictly convex. Let $n > 0$ and $i > 0$. Since $i \vee n - n = (i - n)^+ < i = i \vee (n - n)$, the strict convexity of f implies that $\psi(i, n, n) > 0$. This, together with (3.6) and the obvious fact that $P(M_n^q = S_n^q = n) > 0$, gives strict inequality in (3.4).

Corollary 3.1. *Let f be as in Lemma 3.1. If $p \geq \frac{1}{2}$ then*

$$E[f(i \vee M_n^p - S_n^p)] \geq E[f(i \vee M_n^q)] \tag{3.7}$$

for all $0 \leq n \leq N$ and $0 \leq i \leq N - n$. Moreover, if $p > \frac{1}{2}$ and f is strictly decreasing, then strict inequality holds in (3.7) for all $0 < n \leq N$ and $0 \leq i \leq N - n$.

Proof. Let $p \geq \frac{1}{2}$. Note that in view of (3.2), inequality (3.4) can be stated alternatively as

$$E[f(i \vee M_n^p - S_n^p)] \geq E[f(i \vee M_n^q)]. \tag{3.8}$$

Since $M_n^q \leq M_n^p$ and f is nonincreasing, we furthermore have

$$E[f(i \vee M_n^q)] \geq E[f(i \vee M_n^p)]. \tag{3.9}$$

This, together with (3.8), gives (3.7). If $p > \frac{1}{2}$ and f is strictly decreasing, then strict inequality holds in (3.9) for $i = 0$, since $P(M_n^q < M_n^p) > 0$ for $n > 0$. Furthermore, Lemma 3.1(ii) gives strict inequality in (3.7) for all $i > 0$.

Proof of Theorem 2.1. Define the σ -algebras $\mathcal{F}_k := \sigma(\{U_1, \dots, U_k\})$, $k = 0, 1, \dots, N$. We prove that even among stopping rules that can use complete information about the U_k s, the rules given in the statement of the theorem are optimal. Recall that, for a stopping time τ adapted to $\{\mathcal{F}_k\}$, the sigma algebra \mathcal{F}_τ is defined by the rule $A \in \mathcal{F}_\tau$ if and only if to $A \cap \{\tau \leq k\} \in \mathcal{F}_k$ for all k .

(i) Consider first the case $p \leq \frac{1}{2}$. The argument below is adapted from [6]. Let τ be a stopping time relative to $\{\mathcal{F}_k\}$. By conditioning on \mathcal{F}_τ we can write

$$E[f(M_N^p - S_\tau^p)] = E[G(N - \tau, Z_\tau^p)],$$

where

$$G(k, i) := E[f(i \vee M_k^p)]. \tag{3.10}$$

Using (3.3) and the stationary and independent increments of the random walk, we similarly obtain

$$E[f(M_N^p)] = E[f(Z_N^q)] = E[E[f(Z_N^q) | \mathcal{F}_\tau]] = E[D(N - \tau, Z_\tau^q)],$$

where

$$D(k, i) := E[f(i \vee M_k^q - S_k^q)].$$

(See [6, pp. 654, 660] for the details of these calculations in the case $f(k) = e^{-\delta k}$.) Since f is nonincreasing, $G(k, i)$ is nonincreasing in i for fixed k , which, by (3.1), implies that $G(N - \tau, Z_\tau^p) \leq G(N - \tau, Z_\tau^q)$. But, by (3.8), with the roles of p and q reversed,

$$D(k, i) \geq G(k, i)$$

for all k and all i . It follows that

$$\begin{aligned} E[f(M_N^p - S_\tau^p)] &= E[G(N - \tau, Z_\tau^p)] \leq E[G(N - \tau, Z_\tau^q)] \\ &\leq E[D(N - \tau, Z_\tau^q)] = E[f(M_N^p)] \end{aligned} \tag{3.11}$$

for any stopping time τ . Thus, the rule $\tau \equiv 0$ is optimal.

(ii) Assume next that $p \geq \frac{1}{2}$. Define $G(k, i)$ by (3.10), and let

$$\tilde{D}(k, i) := E[f(i \vee M_k^p - S_k^p)].$$

By Corollary 3.1, $\tilde{D}(k, i) \geq G(k, i)$, and, hence, for any stopping time τ ,

$$\begin{aligned} E[f(M_N^p - S_\tau^p)] &= E[G(N - \tau, Z_\tau^p)] \leq E[\tilde{D}(N - \tau, Z_\tau^p)] \\ &= E[f(Z_N^p)] = E[f(M_N^p - S_N^p)]. \end{aligned} \tag{3.12}$$

Therefore, the rule $\tau \equiv N$ is optimal.

(iii) Finally, consider the case $p = \frac{1}{2}$. Observe that $G(0, i) = D(0, i) = f(i)$ for all i , and $G(k, 0) = E[f(M_k^p)] = E[f(Z_k^q)] = D(k, 0)$ for all k . Thus, for any stopping time τ with $S_\tau = M_\tau$ or $\tau = N$ almost surely,

$$G(N - \tau, Z_\tau^p) = D(N - \tau, Z_\tau^p) = D(N - \tau, Z_\tau^q)$$

(since $p = q$), and, hence, for any such τ ,

$$E[f(M_N^p - S_\tau^p)] = E[f(M_N^p)] = \sup_{\tau'} E[f(M_N^p - S_{\tau'})], \tag{3.13}$$

where the last equality follows by part (i).

Proof of Theorem 2.2. (i) Let $p < \frac{1}{2}$, and suppose that f is not constant. Since f is nonincreasing and convex, this implies that $f(0) > f(i)$ for all $i > 0$. It follows that $G(k, 0) > G(k, i)$ for all $i > 0$ and all k , since, obviously, $f(M_k^p) \geq f(i \vee M_k^p)$, and

$$P(f(M_k^p) > f(i \vee M_k^p)) \geq P(M_k^p = 0) > 0.$$

Now consider a stopping time τ with $\tau > 0$. Then

$$\begin{aligned} \mathbb{P}(Z_\tau^q = 0, Z_\tau^p > 0) &\geq \mathbb{P}(Z_k^q = 0 \text{ and } Z_k^p > 0 \text{ for } k = 1, \dots, N) \\ &\geq \mathbb{P}(X_k^q = 1 \text{ and } X_k^p = -1 \text{ for } k = 1, \dots, N) \\ &= (q - p)^N \\ &> 0. \end{aligned}$$

(Note that this holds for *any* random time τ , not just for stopping times.) It therefore follows that $\mathbb{E}[G(N - \tau, Z_\tau^q)] > \mathbb{E}[G(N - \tau, Z_\tau^p)]$, which is the strict inequality in (3.11).

(ii) Suppose next that $p > \frac{1}{2}$ and f is strictly decreasing. Then strict inequality holds in Corollary 3.1 for $n > 0$ and all i . But this yields strict inequality in (3.12) for any stopping time τ with $\mathbb{P}(\tau < N) > 0$.

(iii) Finally, assume that $p = \frac{1}{2}$, and let f be strictly convex. If $N = 1$, the only stopping times are $\tau \equiv 0$ and $\tau \equiv 1$, which both satisfy the condition in Theorem 2.1(iii). So assume that $N \geq 2$. By Lemma 3.1(iii), strict inequality holds in (3.4) for all $i > 0$. Thus, if τ is a stopping time with the property that $\mathbb{P}(M_\tau^p - S_\tau^p > 0 \text{ and } \tau < N) > 0$ then

$$\mathbb{E}[D(N - \tau, Z_\tau^p)] > \mathbb{E}[G(N - \tau, Z_\tau^p)],$$

and so the first equality in (3.13) is replaced with ' $<$ '.

4. The maximum of Brownian motion

The key to the proof of Theorem 2.3 is the following analog of Lemma 3.1. It makes use of the well-known fact, analogous to (3.2), that

$$(M_t^\lambda - B_t^\lambda, B_t^\lambda) \stackrel{D}{=} (M_t^{-\lambda}, -B_t^{-\lambda}) \tag{4.1}$$

for every fixed $t \geq 0$.

Lemma 4.1. *Let $f : [0, \infty) \rightarrow \mathbb{R}$ be nonincreasing and convex.*

(i) *If $\lambda \geq 0$ then*

$$\mathbb{E}[f(x \vee M_t^\lambda - B_t^\lambda)] \geq \mathbb{E}[f(x \vee (M_t^\lambda - B_t^\lambda))] \tag{4.2}$$

for all $t \geq 0$ and all $x \geq 0$.

(ii) *If $\lambda > 0$ and f is not constant, then strict inequality holds in (4.2) for all $t > 0$ and all $x > 0$.*

(iii) *If $\lambda = 0$ and f is not linear, then strict inequality holds in (4.2) for all $t > 0$ and all $x > 0$.*

Proof. (i) The inequality is trivial when $t = 0$, so assume that $t > 0$. Let $h(s, b; \lambda)$ be the joint density function of $(M_t^\lambda, B_t^\lambda)$. Note that in view of (4.1), or by (4.4) below,

$$h(s, b; \lambda) = h(s - b, -b; -\lambda). \tag{4.3}$$

As in the proof of Lemma 3.1, we begin by writing

$$\begin{aligned} & E[f(x \vee M_t^\lambda - B_t^\lambda) - f(x \vee (M_t^\lambda - B_t^\lambda))] \\ &= \int_{b \in \mathbb{R}} \int_{s > b} [f(x \vee s - b) - f(x \vee (s - b))]h(s, b; \lambda) \, ds \, db \\ &= \left(\int_{b > 0} \int_{s > b} + \int_{b < 0} \int_{s > 0} \right) [f(x \vee s - b) - f(x \vee (s - b))]h(s, b; \lambda) \, ds \, db \\ &=: I^+ + I^-. \end{aligned}$$

Using (4.3) and the change of variables $z = s - b$, $b' = -b$, we can write I^- as

$$\begin{aligned} I^- &= \int_{b' > 0} \int_{z > b'} [f(x \vee (z - b') + b') - f(x \vee z)]h(z, b'; -\lambda) \, dz \, db' \\ &= \int_{b > 0} \int_{s > b} [f(x \vee (s - b) + b) - f(x \vee s)]h(s, b; -\lambda) \, ds \, db, \end{aligned}$$

where the last equality follows simply by renaming the variables. Recall (see, e.g. Equation (3.2) of [1]) that, for fixed t , $h(s, b; \lambda)$ is given by the formula

$$h(s, b; \lambda) = \sqrt{\frac{2}{\pi}} \frac{2s - b}{t^{3/2}} e^{-(2s-b)^2/2t} e^{\lambda(b-\lambda t/2)} \tag{4.4}$$

for all $s \geq 0$ and $b \leq s$. It follows that, for all $b > 0$ and $s \geq b$,

$$h(s, b; \lambda) \geq h(s, b; -\lambda),$$

with strict inequality if $\lambda > 0$. (Note that there does not seem to be a direct probabilistic argument for this last inequality; instead, we must rely on the specific form of the density formula (4.4).) Since f is nonincreasing and $x \vee s - b \leq x \vee (s - b)$ for $b > 0$, we have

$$f(x \vee s - b) - f(x \vee (s - b)) \geq 0 \quad \text{for } b > 0.$$

Thus,

$$I^+ \geq \int_{b > 0} \int_{s > b} [f(x \vee s - b) - f(x \vee (s - b))]h(s, b; -\lambda) \, ds \, db. \tag{4.5}$$

Putting these results together, we conclude that

$$I^+ + I^- \geq \int_{b > 0} \int_{s > b} \psi(x, s, b)h(s, b; -\lambda) \, ds \, db, \tag{4.6}$$

where

$$\psi(x, s, b) := f(x \vee s - b) - f(x \vee (s - b)) + f(x \vee (s - b) + b) - f(x \vee s).$$

As in the proof of Lemma 3.1, the convexity of f implies that $\psi(x, s, b) \geq 0$. Thus, the proof of (4.2) is complete.

(ii) Suppose now that $\lambda > 0$ and f is not constant. Fix $x > 0$. Since f is nonincreasing and convex, we can choose $\delta > 0$ so small that $2\delta < x$, and $f(2\delta) > f(x)$. But then, on the small square $x - \delta < b < x < s < x + \delta$, we have

$$f(x \vee s - b) - f(x \vee (s - b)) = f(s - b) - f(x) \geq f(2\delta) - f(x) > 0.$$

Since $h(s, b; \lambda) > h(s, b; -\lambda)$ on this small square, strict inequality results in (4.5), and, hence, in (4.2).

(iii) Finally, suppose that $\lambda = 0$ and f is not linear. Then there exists a point $x_0 > 0$ such that, for all $x > x_0$ and all $u > 0$, $f(0) - f(u) > f(x) - f(x + u)$. Choose $n \in \mathbb{N}$ such that $nx > x_0$. Then, for $s = b = nx$, $\psi(x, s, b) = f(0) - f(x) + f((n + 1)x) - f(nx) > 0$. By continuity of ψ , it follows that $\psi > 0$ on a small square of positive $h(s, b; -\lambda)$ -density. Putting this back in (4.6) gives strict inequality in (4.2).

Remark 4.1. Inequality (4.2) was derived by Du Toit and Peskir [1] for the special case $f(x) = e^{-\sigma x}$. In fact, for this choice of f , (4.2) holds even if $-\frac{1}{2} \leq \lambda < 0$; for this and other related inequalities for the exponential case, see Remark 4 of [1].

Corollary 4.1. *Let f be as in Lemma 4.1. If $\lambda \geq 0$ then*

$$E[f(x \vee M_t^\lambda - B_t^\lambda)] \geq E[f(x \vee M_t^\lambda)] \tag{4.7}$$

for all $t \geq 0$ and all $x \geq 0$. If $\lambda > 0$ and f is not constant, then strict inequality holds in (4.7) for all $t > 0$ and all $x \geq 0$.

Proof. Let $\lambda \geq 0$. In view of (4.1), inequality (4.2) is equivalent to

$$E[f(x \vee M_t^\lambda - B_t^\lambda)] \geq E[f(x \vee M_t^{-\lambda})]. \tag{4.8}$$

(Note that (4.8) generalizes the key inequality (4.28) of [1].) Since $M_t^{-\lambda} \leq M_t^\lambda$ and f is nonincreasing, we have

$$E[f(x \vee M_t^{-\lambda})] \geq E[f(x \vee M_t^\lambda)]. \tag{4.9}$$

This, together with (4.8), gives (4.7).

Now suppose that $\lambda > 0$ and f is not constant. By Lemma 4.1(ii), it suffices to verify strict inequality for $x = 0$. Since f is nonincreasing and convex, f is strictly decreasing on $[0, x_0]$ for some $x_0 > 0$. Clearly, $P(M_t^{-\lambda} < M_t^\lambda < x_0) > 0$ for $t > 0$. As a result, strict inequality holds in (4.9) for $x = 0$.

Proof of Theorem 2.3. Optimality. Let

$$Z_t^\lambda := M_t^\lambda - B_t^\lambda, \quad t \geq 0,$$

and note that, for fixed t , Z_t^λ is pointwise nonincreasing in λ .

(i) Assume first that $\lambda \leq 0$. Define the functions

$$G(t, x) := E[f(x \vee M_t^\lambda)], \quad D(t, x) := E[f(x \vee M_t^{-\lambda} - B_t^{-\lambda})].$$

Let $\tau \leq T$ be any stopping time adapted to the filtration (\mathcal{F}_t^B) . As in the proof of Theorem 2.1, we have

$$E[f(M_T^\lambda - B_T^\lambda)] = E[G(T - \tau, Z_\tau^\lambda)].$$

Using (4.1), the stationary and independent increments of Brownian motion, and the strong Markov property of the process (Z_t) , we obtain

$$E[f(M_T^\lambda)] = E[f(Z_T^{-\lambda})] = E[D(T - \tau, Z_\tau^{-\lambda})].$$

(For the details of these calculations, see [1, pp. 987, 1004].) Since f is nonincreasing, $G(t, x)$ is nonincreasing in x for fixed t . It follows that $G(T - \tau, Z_\tau^\lambda) \leq G(T - \tau, Z_\tau^{-\lambda})$.

Furthermore, (4.8), with λ replaced by $-\lambda$, gives $D(t, x) \geq G(t, x)$ for all t and all x . As a result,

$$E[f(M_T^\lambda - B_\tau^\lambda)] = E[G(T - \tau, Z_\tau^\lambda)] \leq E[G(T - \tau, Z_\tau^{-\lambda})] \tag{4.10}$$

$$\leq E[D(T - \tau, Z_\tau^{-\lambda})] = E[f(M_T^\lambda)]. \tag{4.11}$$

Since this holds for any stopping time τ , it follows that the rule $\tau \equiv 0$ is optimal.

(ii) Next consider the case $\lambda \geq 0$. Let

$$\tilde{D}(t, x) := E[f(x \vee M_t^\lambda - B_t^\lambda)].$$

Then Corollary 4.1 implies that $\tilde{D}(t, x) \geq G(t, x)$, and, hence,

$$E[f(M_T^\lambda - B_\tau^\lambda)] = E[G(T - \tau, Z_\tau^\lambda)] \leq E[\tilde{D}(T - \tau, Z_\tau^\lambda)] \tag{4.12}$$

$$= E[f(Z_\tau^\lambda)] = E[f(M_T^\lambda - B_T^\lambda)]$$

for any stopping time τ . Thus, the rule $\tau \equiv T$ is optimal.

(iii) Finally, suppose that $\lambda = 0$. Then $G(0, x) = D(0, x)$ for all x and $G(t, 0) = D(t, 0)$ for all t . Thus, for any stopping time τ satisfying (2.2),

$$G(T - \tau, Z_\tau^\lambda) = D(T - \tau, Z_\tau^\lambda) = D(T - \tau, Z_\tau^{-\lambda}),$$

so that (see (4.10) and (4.11))

$$E[f(M_T^\lambda - B_\tau^\lambda)] = E[f(M_T^\lambda)]. \tag{4.13}$$

By part (i) of the theorem, this implies that τ is optimal.

Uniqueness. We next verify the uniqueness claims in Theorem 2.3.

(i) Assume first that $\lambda < 0$. While Lemma 4.1 provides strict inequality in (4.11) for the majority of stopping times, it does not do so for stopping times τ of the form (2.2). Therefore, we establish strict inequality in (4.10) instead. First, since f is nonconstant, nonincreasing, and convex, there exists a point $x_0 > 0$ such that f is strictly decreasing on $[0, x_0]$. It is easy to see that the same is then true for $G(t, \cdot)$ for any fixed t , including $t = 0$. Let $\tau \leq T$ be a stopping time with $P(\tau > 0) > 0$. We show first that

$$P(0 < Z_\tau < x_0) > 0, \tag{4.14}$$

where we write Z_t for Z_t^λ . Choose $t_0 > 0$ so that $P(\tau > t_0) > 0$, and let

$$\tau_0 := \min \left\{ t_0, \tau \left(\frac{x_0}{2} \right) \right\},$$

where $\tau(x) := \inf\{t > 0 : Z_t \geq x\}$ for $x > 0$. Then τ_0 is a stopping time adapted to (\mathcal{F}_t^B) , and so $\{\tau > \tau_0\} \in \mathcal{F}_{\tau_0}^B$. Moreover, $P(\tau > \tau_0) \geq P(\tau > t_0) > 0$ and $P(Z_{\tau_0} > 0) = P(Z_{t_0} > 0) = 1$. Thus, the set $\{\tau > \tau_0, Z_{\tau_0} > 0\}$ lies in $\mathcal{F}_{\tau_0}^B$ and has positive probability. On this set,

$$P(0 < Z_\tau < x_0 \mid \mathcal{F}_{\tau_0}^B) \geq P(0 < Z_t < x_0 \text{ for } \tau_0 \leq t \leq T \mid \mathcal{F}_{\tau_0}^B) > 0,$$

by the strong Markov property of (Z_t) and the fact that (Z_t) behaves like Brownian motion with drift as long as it does not hit 0. But then

$$P(0 < Z_\tau < x_0) = E[P(0 < Z_\tau < x_0 \mid \mathcal{F}_{\tau_0}^B)] > 0,$$

proving (4.14).

Next, a moment of reflection shows that $Z_t^\lambda = Z_t^{-\lambda}$ if and only if $Z_t^\lambda = 0$. Thus, by (4.14),

$$P(Z_\tau^{-\lambda} < Z_\tau^\lambda < x_0) = P(0 < Z_\tau^\lambda < x_0) > 0.$$

Along with the fact that $G(t, \cdot)$ is strictly decreasing on $[0, x_0]$ for all $t \geq 0$, this yields strict inequality in (4.10).

(ii) Consider next the case $\lambda > 0$. Then strict inequality holds in Corollary 4.1 for $t > 0$ and all x . But this yields strict inequality in (4.12) for any stopping time τ with $P(\tau < T) > 0$.

(iii) Finally, assume that $\lambda = 0$, and f is not linear. By Lemma 4.1(iii), strict inequality holds in (4.2) for all $t > 0$ and all $x > 0$. Thus, for any stopping rule τ such that $P(M_\tau^\lambda - B_\tau^\lambda > 0$ and $\tau < T) > 0$,

$$E[D(T - \tau, Z_\tau^\lambda)] > E[G(T - \tau, Z_\tau^\lambda)],$$

and so the equality in (4.13) is replaced with ' $<$ '.

Acknowledgements

This work was prepared while the author was on sabbatical in Kyoto, Japan. The author wishes to thank the Kyoto University Mathematics Department and the Research Institute for Mathematical Sciences (RIMS) for their warm hospitality during 2009.

References

- [1] DU TOIT, J. AND PESKIR, G. (2009). Selling a stock at the ultimate maximum. *Ann. Appl. Prob.* **19**, 983–1014.
- [2] GRAVERSEN, S. E., PESKIR, G. AND SHIRYAEV, A. N. (2000). Stopping Brownian motion without anticipation as close as possible to its ultimate maximum. *Theory Prob. Appl.* **45**, 41–50.
- [3] HLYNKA, M. AND SHEAHAN, J. N. (1988). The secretary problem for a random walk. *Stoch. Process. Appl.* **28**, 317–325.
- [4] PEDERSEN, J. L. (2003). Optimal prediction of the ultimate maximum of Brownian motion. *Stoch. Stoch. Reports* **75**, 205–219.
- [5] SHIRYAEV, A. N., XU, Z. AND ZHOU, X. Y. (2008). Thou shalt buy and hold. *Quant. Finance* **8**, 765–776.
- [6] YAM, S. C. P., YUNG, S. P. AND ZHOU, W. (2009). Two rationales behind the 'buy-and-hold or sell-at-once' strategy. *J. Appl. Prob.* **46**, 651–668.