A general "bang-bang" principle for predicting the maximum of a random walk[*]

Pieter C. Allaart [†]

October 3, 2009

Abstract

Let $(B_t)_{0 \leq t \leq T}$ be either a Bernoulli random walk or a Brownian motion with drift, and let $M_t := \max\{B_s : 0 \leq s \leq t\}$, $0 \leq t \leq T$. This paper solves the general optimal prediction problem

$$\sup_{0 \leq \tau \leq T} \mathrm{E}[f(M_T - B_\tau)],$$

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 $\tau^* \equiv T$ is the drift is positive. This result generalizes recent findings by S. Yam, S. Yung and W. Zhou [*J. Appl. Probab.* **46** (2009), 651–668] and J. Du Toit and G. Peskir [*Ann. Appl. Probab.* **19** (2009), 983–1014], and provides additional mathematical justification for the dictum in finance that one should sell bad stocks immediately, but keep good ones as long as possible.

AMS 2000 subject classification. Primary 60G40, 60G50, 60J65; secondary 60G25.

Key words and phrases: Bernoulli random walk, Brownian motion, optimal prediction, ultimate maximum, stopping time, convex function.

1 Introduction and main results

A number of recent papers (e.g. [1, 4, 5]) have discussed the problem of stopping a random walk, or a Brownian motion, "as close as possible" to its ultimate maximum. An important motivation in these papers was the financial problem of selling a stock at a price "close" to the highest price over a given finite time

---

[†]Address correspondence to P. C. Allaart, Department of Mathematics, University of North Texas, 1155 Union Circle #311430, Denton, TX 76203-5017, USA; E-mail: allaart@unt.edu

interval, when the stock price follows a discrete binomial model (in [5]) or a geometric Brownian motion (in [1, 4]). In these three papers, "closeness" was measured by the ratio of the stopped price to the ultimate maximum price, and the striking result was that the optimal strategy is of "bang-bang" type, meaning that it is either optimal to stop at time zero, or to stop at the time horizon, depending on the quality of the stock. These results, as pointed out by the papers' authors, reinforce the widely held financial view that one should sell bad stocks quickly, but keep good ones as long as possible.

The purpose of the present paper is to provide an important generalization of the results in [1, 5]. Rather than considering price ratios, we take as the basic process either a "flat" Bernoulli random walk or a Brownian motion with drift, and measure closeness by a general nonincreasing, convex function f of the positive distance from the stopped value of the process to its eventual maximum. For this more general problem we obtain the same result, namely that it is either optimal to stop at time zero or at the time horizon, depending on the drift of the underlying process. For the specific function $f(x) = e^{-\sigma x}$, where $\sigma > 0$, our results reduce to those of [1] and [5]. 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.

The remainder of this section is devoted to a precise formulation of the problem and statements of the main results, which are nontechnical in nature. 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 $n \geq 1$, $S_n = X_1 + \cdots + X_n$, where $X_1, X_2, \dots$ are independent, identically distributed random variables with $\mathrm{P}(X_1 = 1) = p$, and $\mathrm{P}(X_1 = -1) = q := 1 - p$. Let a finite time horizon $N \in \mathbb{N}$ be given. Let $f : \{0, 1, \dots, N\} \to \mathbb{R}$ be nonincreasing, and consider the optimal stopping problem

$$\sup_{0 \leq \tau \leq N} \mathrm{E}[f(M_N - S_\tau)], \qquad (1.1)$$

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 (1.1) into the probability $\mathrm{P}(S_\tau = M_N)$, so that (1.1) 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 [2], where this problem is solved in a somewhat more general setting for the case $p = 1/2$. Yam et al. [5] solved the problem for arbitrary p, and showed the (unique) optimal rule to be $\tau \equiv 0$ when $p < 1/2$, and $\tau \equiv n$ when $p > 1/2$. When $p = 1/2$, 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. (A similar problem, where the objective is to stop a Brownian motion within a distance $\varepsilon > 0$ from its ultimate maximum, was considered in [3].)

Yam et al. [5] also treated the case $f(k) = d^k$, where $0 < d < 1$ is a constant. They showed that for this quite different objective function, the optimal rule is

nonetheless exactly the same as for the problem of maximizing the probability of stopping at the maximum.

This leads one to believe that there must be some general principle at work. Notice that in each of the above examples, f is in fact convex. The first aim of this note 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 [5]. Recall that a function $f : \{0, 1, \ldots, N\} \to \mathbb{R}$ is *convex* if $f(k) - 2f(k+1) + f(k+2) \geq 0$ for all k with $0 \leq k \leq N-2$, and is *strictly convex* if the inequality is strict for all such k.

Theorem 1.1. *Let $f : \{0, 1, \ldots, N\} \to \mathbb{R}$ be nonincreasing and convex, and consider the optimal stopping problem (1.1).*

(i) If $p \leq 1/2$, the rule $\tau \equiv 0$ is optimal.

(ii) If $p \geq 1/2$, the rule $\tau \equiv N$ is optimal.

(iii) If $p = 1/2$, any stopping time τ satisfying $S_\tau = M_\tau$ or $\tau = N$ almost surely 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 $p < 1/2$, and $\tau^* \equiv N$ if $p > 1/2$.

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

Example 1.1. 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, regardless of p, 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 that for the rule $\tau \equiv 2$ is $1 - q^2$.

One might ask when the optimal rules in Theorem 1.1 are unique. The next theorem gives simple sufficient conditions to this effect.

Theorem 1.2. *Let f be as in Theorem 1.1.*

(i) If $p < 1/2$ and f is nonconstant, then the rule $\tau \equiv 0$ is the unique optimal rule.

(ii) If $p > 1/2$ and f is strictly decreasing, then the rule $\tau \equiv N$ is the unique optimal rule.

(iii) If $p = 1/2$ and f is strictly convex, then the only optimal rules are those that satisfy $S_\tau = M_\tau$ or $\tau = N$ almost surely.

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

Next, let $B := (B_t)_{t \geq 0}$ be a standard Brownian motion, and λ 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 $M_t^\lambda := \max\{B_s^\lambda : 0 \leq s \leq t\}$. Once again we are interested in finding 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

$$\sup_{0 \leq \tau \leq T} \mathrm{E}[f(M_T^\lambda - B_\tau^\lambda)], \tag{1.2}$$

3

where $f : [0, \infty) \to \mathbb{R}$ is a given (reward) function, and $T > 0$ a fixed time horizon. Note that in particular, the choice $f(x) = e^{-\sigma x}$ for a constant $\sigma > 0$ yields the problem studied in [4] and in Section 4 of [1].

Since Brownian motion is the scaling limit of Bernoulli random walk, one might expect the result to be the same as in Theorem 1.1. This is indeed the case, except that the conditions for uniqueness of the optimal rules are weaker.

Theorem 1.3. *Let $f : [0, \infty) \to \mathbb{R}$ be non-constant, nonincreasing and convex, and consider the optimal stopping problem* (1.2).
(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

$$\mathrm{P}(B_\tau^\lambda = M_\tau^\lambda \ or \ \tau = T) = 1 \tag{1.3}$$

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 (1.3).

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 1.3 are the best possible. Essentially, the conditions for uniqueness of the optimal rules are weaker than in the discrete case because the increments of Brownian motion can be arbitrarily large, whereas the increments of Bernoulli random walk are bounded.

Finally, we note that by putting $\tilde{f} := -f$, problems (1.1) and (1.2) may be formulated equivalently as penalty-minimization problems. For instance, (1.2) can be represented alternatively in the form

$$\inf_{0 \le \tau \le T} \mathrm{E}[\tilde{f}(M_T^\lambda - B_\tau^\lambda)], \tag{1.4}$$

where $\tilde{f} : [0, \infty) \to \mathbb{R}$ is nondecreasing and concave. Thus, the above results apply to a variety of natural penalty functions, including $\tilde{f}(x) = x^\alpha$ where $0 < \alpha < 1$, $\tilde{f}(x) = \log(1 + x)$, etc. However, without concavity of $\tilde{f}$ the optimal rules are generally of a more intricate form: see, for instance, the solution in [1, Section 3] of (1.4) for the function $\tilde{f}(x) = e^{\sigma x}$, where $\sigma > 0$.

Theorems 1.1 and 1.2 are proved in Section 2, and Theorem 1.3 is proved in Section 3. Many of the ideas of the proofs are adapted from [1] and [5], and some details, in as far as they can be found in these papers, are therefore omitted here. The novel contributions of the present article are the explicit use of the convexity of f (see Lemmas 2.1 and 3.1 below), and the investigation of uniqueness of the optimal stopping times, which requires some finesse in the case of general f.

2 The maximum of Bernoulli random walk

This section is devoted to the proofs of Theorems 1.1 and 1.2. It will be useful to consider an infinite family of random walks, defined on the same probability

space. The following construction is standard. Let $U_1, U_2 \ldots$ be independent random variables, uniformly distributed on the interval $[0, 1]$. For $k \in \mathbb{N}$ and $p \in (0, 1)$, define

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

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

Let $M_k^p := \max\{S_0^p, \ldots, S_k^p\}$, and $Z_k^p := M_k^p - S_k^p$, for $k = 0, 1 \ldots$. Observe that for each p, the process $\{Z_k^p\}_k$ is Markovian. Moreover, it is easy to see that

$$p \geq p' \Rightarrow Z_k^p \leq Z_k^{p'} \text{ for all } k. \tag{2.1}$$

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

$$(M_n^p - S_n^p, S_n^p) \stackrel{d}{=} (M_n^q, -S_n^q) \tag{2.2}$$

for all $n \in \mathbb{N}$. (The easiest way to see this is to observe that the time-reversed process $\tilde{S}_k := S_{n-k}^p - S_n^p$, $k = 0, 1, \ldots, n$ 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),

$$M_n^p \stackrel{d}{=} M_n^q - S_n^q = Z_n^q. \tag{2.3}$$

It is almost amusing to see how many times this identity must be used in order to prove Theorem 1.1.

The following lemma holds the key to the proof of Theorem 1.1.

Lemma 2.1. *Let $f : \{0, 1, \ldots, N\} \to \mathbb{R}$ be nonincreasing and convex.*
  (i) If $p \geq 1/2$, then

$$\mathrm{E}[f(i \vee M_n^p - S_n^p)] \geq \mathrm{E}\left[f\left(i \vee (M_n^p - S_n^p)\right)\right] \tag{2.4}$$

for all $n \leq N$ and all $i \geq 0$.
  (ii) If $p > 1/2$ and f is strictly decreasing, then strict inequality holds in (2.4) for all $0 < n \leq N$ and all $i > 0$.
  (iii) If $p \geq 1/2$ and f is strictly convex, then strict inequality holds in (2.4) for all $0 < n \leq N$ and all $i > 0$.

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

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

5

(Note that the terms with $l = 0$ vanish.) By (2.2) and the change of variable $k' = k - l$, $l' = -l$, the second summation becomes

$$\Sigma^- = \sum_{l<0}\sum_{k\geq 0} \left[f(i \vee k - l) - f\big(i \vee (k-l)\big)\right] \mathrm{P}(M_n^q = k-l, S_n^q = -l)$$

$$= \sum_{l'>0}\sum_{k'\geq l'} \left[f\big(i \vee (k'-l') + l'\big) - f(i \vee k')\right] \mathrm{P}(M_n^q = k', S_n^q = l').$$

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

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

(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

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

and therefore,

$$\Sigma^+ \geq \sum_{l>0}\sum_{k\geq l} \left[f(i \vee k - l) - f\big(i \vee (k-l)\big)\right] \mathrm{P}(M_n^q = k, S_n^q = l). \tag{2.5}$$

Combining these results, we conclude that

$$\Sigma^+ + \Sigma^- \geq \sum_{l>0}\sum_{k\geq l} \psi(i,k,l)\,\mathrm{P}(M_n^q = k, S_n^q = l), \tag{2.6}$$

where

$$\psi(i,k,l) := \left[f(i \vee k - l) - f\big(i \vee (k-l)\big)\right] + \left[f\big(i \vee (k-l) + l\big) - f(i \vee k)\right]$$

$$= [f(i \vee k - l) - f(i \vee k)] - \left[f\big(i \vee (k-l)\big) - f\big(i \vee (k-l) + l\big)\right].$$

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 (2.4).

(ii) Suppose $p > 1/2$ and f is strictly decreasing. Let $n > 0$ and $i > 0$, and put $k = l = n$. Then

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

Since $\mathrm{P}(M_n^p = S_n^p = n) > \mathrm{P}(M_n^q = S_n^q = n)$, strict inequality holds in (2.5), and hence in (2.4).

(iii) Finally, suppose $p \geq 1/2$ 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 (2.6) and the obvious fact that $\mathrm{P}(M_n^q = S_n^q = n) > 0$, gives strict inequality in (2.4). $\qquad\square$

Corollary 2.1. *Let f be as in Lemma 2.1. If $p \geq 1/2$, then*

$$\mathrm{E}[f(i \vee M_n^p - S_n^p)] \geq \mathrm{E}[f(i \vee M_n^p)] \tag{2.7}$$

for all $n \leq N$ and all $i \geq 0$. Moreover, if $p > 1/2$ and f is strictly decreasing, then strict inequality holds in (2.7) for all $0 < n \leq N$ and all $i \geq 0$.

6

Proof. Let $p \geq 1/2$. Note that in view of (2.2), the inequality (2.4) can be stated alternatively as

$$\mathrm{E}[f(i \vee M_n^p - S_n^p)] \geq \mathrm{E}[f(i \vee M_n^q)]. \tag{2.8}$$

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

$$\mathrm{E}[f(i \vee M_n^q)] \geq \mathrm{E}[f(i \vee M_n^p)]. \tag{2.9}$$

This, together with (2.8), gives (2.7).

Now suppose $p > 1/2$ and f is strictly decreasing. By Lemma 2.1(ii), it suffices to verify strict inequality for $i = 0$. But for this value of i, (2.9) holds with strict inequality, since $\mathrm{P}(M_n^q < M_n^p) > 0$ for $n > 0$. $\qquad \square$

Proof of Theorem 1.1. Define the σ-algebras $\mathcal{F}_k := \sigma(\{U_1, \ldots, U_k\})$, for $k = 0, 1, \ldots, N$. We prove the slightly stronger statement 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 $\mathcal{F}_\tau$ is defined by the rule

$$A \in \mathcal{F}_\tau \Leftrightarrow A \cap \{\tau \leq k\} \in \mathcal{F}_k \text{ for all } k.$$

(i) Consider first the case $p \leq 1/2$. The argument below is adapted from [5]. Let τ be a stopping time adapted to $\{\mathcal{F}_k\}$. By conditioning on $\mathcal{F}_\tau$, we can write

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

where

$$G(k, i) := \mathrm{E}[f(i \vee M_k^p)]. \tag{2.10}$$

Using (2.3) and the stationary and independent increments of the random walk, we obtain similarly

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

where

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

(See [5], p. 654 and p. 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 (2.1) implies that $G(N - \tau, Z_\tau^p) \leq G(N - \tau, Z_\tau^q)$. But by (2.8) with the roles of p and q reversed,

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

for all k and all i. It follows that

$$\mathrm{E}[f(M_N^p - S_\tau^p)] = \mathrm{E}[G(N - \tau, Z_\tau^p)] \leq \mathrm{E}[G(N - \tau, Z_\tau^q)] \tag{2.11}$$
$$\leq \mathrm{E}[D(N - \tau, Z_\tau^q)] = \mathrm{E}[f(M_N^p)],$$

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

(ii) Assume next that $p \geq 1/2$. Define $G(k, i)$ by (2.10), and let

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

By Corollary 2.1, $\tilde{D}(k, i) \geq G(k, i)$, and hence, for any stopping time τ,

$$\mathrm{E}[f(M_N^p - S_\tau^p)] = \mathrm{E}[G(N - \tau, Z_\tau^p)] \leq \mathrm{E}[\tilde{D}(N - \tau, Z_\tau^p)] \qquad (2.12)$$
$$= \mathrm{E}[f(Z_N^p)] = \mathrm{E}[f(M_N^p - S_N^p)].$$

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

(iii) Consider finally the case $p = 1/2$. Observe that $G(0, i) = D(0, i) = f(i)$ for all i, and $G(k, 0) = \mathrm{E}[f(M_k^p)] = \mathrm{E}[f(Z_k^q)] = D(k, 0)$ 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,

$$\mathrm{E}[f(M_N^p - S_\tau^p)] = \mathrm{E}[f(M_N^p)] = \sup_{\tau'} \mathrm{E}[f(M_N^p - S_{\tau'})], \qquad (2.13)$$

where the last equality follows by part (i). $\qquad \square$

Proof of Theorem 1.2. (i) Let $p < 1/2$, and suppose 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

$$\mathrm{P}\left[f(M_k^p) > f(i \vee M_k^p)\right] \geq \mathrm{P}(M_k^p = 0) > 0.$$

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

$$\mathrm{P}(Z_\tau^q = 0, Z_\tau^p > 0) \geq \mathrm{P}(Z_k^q = 0 \text{ and } Z_k^p > 0 \text{ for } k = 1, \ldots, N)$$
$$\geq \mathrm{P}(X_k^q = 1 \text{ and } X_k^p = -1 \text{ for } k = 1, \ldots, N)$$
$$= (q - p)^N > 0.$$

(Note that this holds for *any* random time τ, not just for stopping times.) It therefore follows that $\mathrm{E}[G(N - \tau, Z_\tau^q)] > \mathrm{E}[G(N - \tau, Z_\tau^p)]$, which is strict inequality in (2.11).

(ii) Suppose next that $p > 1/2$ and f is strictly decreasing. Then strict inequality holds in Corollary 2.1 for $n > 0$ and all i. But this yields strict inequality in (2.12) for any stopping time τ with $\mathrm{P}(\tau < N) > 0$.

(iii) Finally, assume $p = 1/2$, 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 1.1(iii). So assume $N \geq 2$. By Lemma 2.1(iii), strict inequality holds in (2.4) for all $i > 0$. Thus, if τ is a stopping time with the property that $\mathrm{P}(M_\tau^p - S_\tau^p > 0 \text{ and } \tau < N) > 0$, then

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

and so the first equality in (2.13) is replaced with "<". $\qquad \square$

3 The maximum of Brownian motion

The key to the proof of Theorem 1.3 is the following analog of Lemma 2.1. It makes use of the well-known fact, analogous to (2.2), that

$$(M_t^\lambda - B_t^\lambda, B_t^\lambda) \stackrel{d}{=} (M_t^{-\lambda}, -B_t^{-\lambda}). \tag{3.1}$$

Lemma 3.1. *Let $f : [0, \infty) \to \mathbb{R}$ be nonincreasing and convex.*
  (i) If $\lambda \geq 0$, then

$$\mathrm{E}\left[f(x \vee M_t^\lambda - B_t^\lambda)\right] \geq \mathrm{E}\left[f\big(x \vee (M_t^\lambda - B_t^\lambda)\big)\right] \tag{3.2}$$

for all $t \geq 0$ and all $x \geq 0$.
  (ii) If $\lambda > 0$ and f is not constant, then strict inequality holds in (3.2) for all $t > 0$ and all $x > 0$.
  (iii) If $\lambda = 0$ and f is not linear, then strict inequality holds in (3.2) for all $t > 0$ and all $x > 0$.

Proof. (i) The inequality is trivial when $t = 0$, so assume $t > 0$. Let $h(s, b; \lambda)$ be the joint density function of $(M_t^\lambda, B_t^\lambda)$. Note that in view of (3.1), or by (3.4) below,

$$h(s, b; \lambda) = h(s - b, -b; -\lambda). \tag{3.3}$$

As in the proof of Lemma 2.1, we begin by writing

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

Using (3.3) and the change of variable $z = s - b$, $b' = -b$, we can write I^- as

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

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

$$h(s, b; \lambda) = \sqrt{\frac{2}{\pi}} \frac{2s - b}{t^{3/2}} e^{-(2s-b)^2/2t} e^{\lambda(b - \lambda t/2)} \tag{3.4}$$

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 (3.4).) Since f is nonincreasing and $x \vee s - b \leq x \vee (s - b)$ for $b > 0$, we have

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

Thus,

$$I^+ \geq \int_{b>0} \int_{s>b} \big[f(x \vee s - b) - f\big(x \vee (s - b)\big)\big] h(s, b; -\lambda) \, ds \, db. \qquad (3.5)$$

Putting these results together, we conclude that

$$I^+ + I^- \geq \int_{b>0} \int_{s>b} \psi(x, s, b) h(s, b; -\lambda) \, ds \, db, \qquad (3.6)$$

where

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

As in the proof of Lemma 2.1, the convexity of f implies $\psi(x, s, b) \geq 0$. Thus, the proof of (3.2) 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

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

Since $h(s, b; \lambda) > h(s, b; -\lambda)$ on this small square, strict inequality results in (3.5), and hence in (3.2).

(iii) Suppose finally 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 $s = b = nx$, $\psi(x, s, b) = f(0) - f(x) + f\big((n+1)x\big) - f(nx) > 0$. By continuity of ψ, it follows that $\psi > 0$ on a small square of positive $h(s, b; -\lambda)$-density. Putting this back in (3.6) gives strict inequality in (3.2). $\qquad \square$

Corollary 3.1. *Let f be as in Lemma 3.1. If $\lambda \geq 0$, then*

$$\mathrm{E}[f(x \vee M_t^\lambda - B_t^\lambda)] \geq \mathrm{E}[f(x \vee M_t^\lambda)] \qquad (3.7)$$

for all $t \geq 0$ and all $x \geq 0$. Moreover, if $\lambda > 0$ and f is not constant, then strict inequality holds in (3.7) for all $t > 0$ and all $x \geq 0$.

Proof. Let $\lambda \geq 0$. In view of (3.1), the inequality (3.2) is equivalent to

$$\mathrm{E}[f(x \vee M_t^\lambda - B_t^\lambda)] \geq \mathrm{E}[f(x \vee M_t^{-\lambda})]. \qquad (3.8)$$

(Note that (3.8) generalizes the key inequality (4.28) in [1].) Since $M_t^{-\lambda} \leq M_t^\lambda$ and f is nonincreasing, we have

$$\mathrm{E}[f(x \vee M_t^{-\lambda})] \geq \mathrm{E}[f(x \vee M_t^\lambda)]. \qquad (3.9)$$

This, together with (3.8), gives (3.7).

Now suppose that $\lambda > 0$ and f is not constant. By Lemma 3.1(ii), it suffices to verify strict inequality for $x = 0$. Since f is nonincreasing and convex, there exists $x_0 > 0$ such that f is strictly decreasing on $[0, x_0]$. Clearly, $\mathrm{P}(M_t^{-\lambda} < M_t^\lambda < x_0) > 0$ for $t > 0$. As a result, strict inequality holds in (3.9) for $x = 0$. $\qquad\square$

Proof of Theorem 1.3. a) *Optimality.* We first prove that the rules given in the statement of the theorem are optimal. Let

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

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

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

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

Let $\tau \leq T$ be any stopping time adapted to the filtration $(\mathcal{F}_t^B)$. As in the proof of Theorem 1.1, we have

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

Using (3.1), the stationary and independent increments of Brownian motion and the strong Markov property of the process (Z_t), we obtain

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

(For the details of these calculations, see [1], p. 987 and p. 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, (3.8) with λ replaced by $-\lambda$ gives $D(t, x) \geq G(t, x)$, for all t and all x. As a result,

$$\mathrm{E}[f(M_T^\lambda - B_\tau^\lambda)] = \mathrm{E}[G(T - \tau, Z_\tau^\lambda)] \leq \mathrm{E}[G(T - \tau, Z_\tau^{-\lambda})] \qquad (3.10)$$
$$\leq \mathrm{E}[D(T - \tau, Z_\tau^{-\lambda})] = \mathrm{E}[f(M_T^\lambda)]. \qquad (3.11)$$

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

(ii) Consider next the case $\lambda \geq 0$. Let

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

Then Corollary 3.1 implies that $\tilde{D}(t, x) \geq G(t, x)$, and hence,

$$\mathrm{E}[f(M_T^\lambda - B_\tau^\lambda)] = \mathrm{E}[G(T - \tau, Z_\tau^\lambda)] \leq \mathrm{E}[\tilde{D}(T - \tau, Z_\tau^\lambda)] \qquad (3.12)$$
$$= \mathrm{E}[f(Z_T^\lambda)] = \mathrm{E}[f(M_T^\lambda - B_T^\lambda)]$$

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

(iii) Suppose finally 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 (1.3),

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

so that (see (3.10) and (3.11))

$$\mathrm{E}[f(M_T^\lambda - B_\tau^\lambda)] = \mathrm{E}[f(M_T^\lambda)]. \tag{3.13}$$

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

b) *Uniqueness.* We next verify the uniqueness claims in Theorem 1.3.

(i) Assume first that $\lambda < 0$. While Lemma 3.1 provides strict inequality in (3.11) for the majority of stopping times, it does not do so for stopping times τ of the form (1.3). Therefore, we establish strict inequality in (3.10) instead. First, since f is non-constant, 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 $\mathrm{P}(\tau > 0) > 0$. We show first that

$$\mathrm{P}(0 < Z_\tau < x_0) > 0, \tag{3.14}$$

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

$$\tau_0 := \min\{t_0, \tau(x_0/2)\},$$

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, $\mathrm{P}(\tau > \tau_0) \geq \mathrm{P}(\tau > t_0) > 0$, and $\mathrm{P}(Z_{\tau_0} > 0) = \mathrm{P}(Z_{t_0} > 0) = 1$. Thus, the set $\{\tau > \tau_0, Z_{\tau_0} > 0\}$ lies in $\mathcal{F}_{\tau_0}^B$ and has positive probability. On this set,

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

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

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

proving (3.14).

Next, a moment of reflection shows that $Z_t^\lambda = Z_t^{-\lambda}$ if and only if $Z_t^\lambda = 0$. Thus, by (3.14),

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

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 (3.10).

(ii) Consider next the case $\lambda > 0$. Then strict inequality holds in Corollary 3.1 for $t > 0$ and all x. But this yields strict inequality in (3.12) above for any stopping time τ with $\mathrm{P}(\tau < T) > 0$.

12

(iii) Assume finally that $\lambda = 0$, and f is not linear. By Lemma 3.1(iii), strict inequality holds in (3.2) for all $t > 0$ and all $x > 0$. Thus, for any stopping rule τ such that $P(M_\tau^\lambda - B_\tau^\lambda > 0 \text{ and } \tau < T) > 0$,

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

and so the equality in (3.13) is replaced with "$<$". $\qquad\square$

Acknowledgements

References

[1] DU TOIT, J. and PESKIR, G. (2009). Selling a stock at the ultimate maximum. *Ann. Appl. Probab.* **19** 983–1014.

[2] HLYNKA, M. and SHEAHAN, J. N. (1988). The secretary problem for a random walk. *Stoch. Proc. Appl.* **28** 317–325.

[3] PEDERSEN, J. L. (2003). Optimal prediction of the ultimate maximum of Brownian motion. *Stoch. Stoch. Rep.* **75** 205–219.

[4] SHIRYAEV, A. N., XU, Z. and ZHOU, X. Y. (2008). Thou shalt buy and hold. *Quant. Finance* **8** 765–776.

[5] YAM, S. C. P., YUNG, S. P. and ZHOU, W. (2009). Two rationales behind 'buy-and-hold or sell-at-once'. *J. Appl. Probab.* **46** 651–668.