The time to a given drawdown in Brownian Motion

Isaac Meilijson*

School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, 69978 Tel-Aviv, Israel
e-mail: meilijson@math.tau.ac.il

Summary. This article deals with optimal stopping of Brownian Motion when the sampling cost is linear in time and the reward upon stopping is a non-decreasing function of the cumulative maximum. This can be viewed as pricing and management of a type of look-back American put option. The case of linear reward function was studied by Dubins & Schwarz [10].

Our treatment of the problem involves a stopped Brownian Motion formula by Taylor (see Taylor [18] and Williams [19]), first exit times by Brownian Motion from open intervals, processes with dichotomous transitions and the Azéma–Yor [2] stopping time.

Introduction

Let $\{W(t) \mid t \geqslant 0,\, W(0) = 0\}$ be Standard Brownian Motion (SBM) and let $\{B(t) \mid B(t) = \mu t + \sigma W(t),\, t \geqslant 0\}$ be Brownian Motion (BM) with *drift* μ and *diffusion parameter* σ, where $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$. For $d > 0$, define the stopping time

$$\tau_d = \min\Big\{ t \;\Big|\; \max_{0 \leqslant s \leqslant t} B(s) \geqslant B(t) + d \Big\} \tag{1}$$

to be the first time to achieve a *drawdown* of size d. That is, τ_d is the first time that BM has gone down by d from its record high value so far. As motivated by Taylor [18], an investor that owns a share whose value at time t is $V_t = V_0 \exp(B(t))$, may consider selling it at time τ_d (for some $d > 0$) because it has lost for the first time some fixed fraction $1 - \exp(-d)$ of its previously held highest value $V_0 \exp(M_d)$ (where $M_d = \max_{0 \leqslant s \leqslant \tau_d} B(s) = B(\tau_d) + d$), a possible indication of change of drift. The investor would also want to know what should be the "typical" drawdown of the share, that is, its stationary

* Work conducted in part while visiting the University of Lund, Sweden and Delft University of Technology, The Netherlands. The hospitality of the departments of Mathematical Statistics at Lund and Technical Mathematics and Informatics at Delft are warmly acknowledged.

J. Azéma et al. (Eds.): LNM 1832, pp. 94–108, 2003.

distribution. However, while τ_d and M_d have been unambiguously defined, this is not the case with the notion of drawdown, that can be equally reasonably interpreted as each of two types of process that achieve simultaneously their record heights, the (reflected Brownian Motion) *gap* process G (so termed by Dubins & Schwarz [10]) and the *extent* (so termed by Goldhirsch & Noskovicz [12]) or *downfall* (so termed by Douady, Shiryaev & Yor [7]) process X. These processes are defined via the last maximum time process L and the last maximum process M:

$$M(t) = \max_{0 \leqslant s \leqslant t} B(s) \tag{2}$$

$$L(t) = \max\{s \mid 0 \leqslant s \leqslant t,\, B(s) = M(t)\} \tag{3}$$

$$G(t) = M(t) - B(t) \tag{4}$$

$$X(t) = M(t) - \min_{L(t) \leqslant s \leqslant t} B(s) \tag{5}$$

Taylor [18] presents a closed form formula for the joint moment generating function of τ_d and $B(\tau_d)$. (Hyperbolic sine is denoted $\sinh(y) = (\exp(y) - \exp(-y))/2$. Other hyperbolic trigonometric functions such as cosh and coth are defined accordingly).

$$\mathbb{E}[\exp(\alpha B(\tau_d) - \beta\tau_d)] = \frac{\delta \exp\bigl(-(\alpha + \mu/\sigma^2)d\bigr)}{\delta\cosh(\delta d) - (\alpha + \mu/\sigma^2)\sinh(\delta d)}\,, \tag{6}$$

where $\delta = \sqrt{(\mu/\sigma^2)^2 + 2\beta/\sigma^2}$. This formula holds for $\beta > 0$ and $\alpha < \delta\coth(\delta d) - \mu/\sigma^2$ (a positive upper limit). The formula holds for $\beta = 0$ as well if $\mu \neq 0$, thus identifying the moment generating function of $B(\tau_d)$ for this case. To see the difficulty with negative β, exponential Martingale methods show that for $\mu = 0$ the k'th moment of τ_d is linearly related to the $2k$'th central moment of the *exponential* variable M_d (Williams [19]). Hence, τ_d has finite moments of all orders but infinite moment generating function on $(0, \infty)$. That is why Taylor's formula identifies its Laplace transform.

As straightforward corollaries, if $\mu \neq 0$,

$$\mathbb{E}[\tau_d] = \frac{\sigma^2}{2\mu^2}\left(\exp\Bigl(\frac{2\mu}{\sigma^2}d\Bigr) - 1 - \frac{2\mu}{\sigma^2}d\right), \tag{7}$$

$$\mathbb{E}[M_d] = d + \mu\,\mathbb{E}[\tau_d] = \frac{\sigma^2}{2\mu}\left(\exp\Bigl(\frac{2\mu}{\sigma^2}d\Bigr) - 1\right). \tag{8}$$

For $\mu = 0$ and $\beta > 0$, it is easy to see that

$$\mathbb{E}[\tau_d] = \frac{d^2}{\sigma^2};\ \operatorname{Var}(\tau_d) = \frac{2}{3}\frac{d^4}{\sigma^4};\ \mathbb{E}[M_d] = d;\ \frac{1}{\mathbb{E}[\exp(-\beta\tau_d)]} = \cosh\Bigl(\frac{d}{\sigma}\sqrt{2\beta}\Bigr). \tag{9}$$

We can further state,

Proposition 1. *For $\mu > 0$, the Markovian gap process G has exponential stationary distribution with mean $\sigma^2/(2\mu)$ and the stochastically bigger non-Markovian downfall process X has stationary distribution with expectation $(\pi^2/6) \times (\sigma^2/(2\mu))$ and cumulative distribution function*

$$F_X(d) = \frac{\mathbb{E}[\tau_d]}{\mathbb{E}[\tau_d] + d/\mu} = 1 - \frac{2\mu d/\sigma^2}{\exp(2\mu d/\sigma^2) - 1}\,. \tag{10}$$

It should be clear that the gap and downfall processes are ergodic if $\mu > 0$ and null-recurrent if $\mu = 0$. In the latter case, the downfall process, whose sample paths increase continuously and drop down discontinuously to 0, is related to a "remarkable" Martingale, introduced by Azéma [1] and so termed by Azéma & Yor [3]. Further discussion can be found in Protter [16].

The record high value M_d is exponentially distributed because as long as first hitting times of positive heights occur before achieving a drawdown of d, these times are renewal times: knowing that $M_d > x$ is the same as knowing that B has not achieved a drawdown of d by the time it first reached height x. But then it starts anew the quest for a drawdown!

Let us place this nice fact in the broader context of Skorokhod embeddings in SBM. The problem as posed and solved by Skorokhod in [17] is the following (not stated here in its fullest generality): Given a distribution F of a random variable Y with mean zero and finite variance, find a stopping time τ in SBM W, with finite mean, for which $W(\tau)$ is distributed F. The Chacon–Walsh [5] family of solutions is easiest to describe: Express Y as the limit of a Martingale $Y_n = \mathbb{E}[Y \mid \mathcal{F}_n]$ with dichotomous transitions (that is, the conditional distribution of Y_{n+1} on $\mathcal{F}_n$ is a.s. two-valued), and then progressively embed this Martingale in W by a sequence of first exit times from open intervals. Dubins [8] was the first to build such a scheme, letting $\mathcal{F}_1$ decide whether $Y \geqslant \mathbb{E}[Y]$ or $Y < \mathbb{E}[Y]$ by a first exit time of W starting at $\mathbb{E}[Y]$ from the open interval $(\mathbb{E}[Y \mid Y < \mathbb{E}[Y]], \mathbb{E}[Y \mid Y \geqslant \mathbb{E}[Y]])$. It then proceeds recursively. E.g., if the first step ended at $\mathbb{E}[Y \mid Y \geqslant \mathbb{E}[Y]]$ then the second step ends when W, re-starting at $\mathbb{E}[Y \mid Y \geqslant \mathbb{E}[Y]]$, first exits the open interval $(\mathbb{E}[Y \mid \mathbb{E}[Y] \leqslant Y < \mathbb{E}[Y \mid Y \geqslant \mathbb{E}[Y]]], \mathbb{E}[Y \mid Y \geqslant \mathbb{E}[Y \mid Y \geqslant \mathbb{E}[Y]]])$.

One of the analytically most elegant solutions to Skorokhod's problem is the Azéma–Yor stopping time T_{A-Y} (see Azéma & Yor [2] and Meilijson [15]), defined in terms of the upper barycenter function of F, $H_F(x) = \mathbb{E}[Y \mid Y \geqslant x] = \int_x^\infty y\,\mathrm{d}F(y)/(1 - F(x-))$ as

$$T_{A-Y} = \min\Big\{t \;\Big|\; \max_{0 \leqslant s \leqslant t} W(s) \geqslant H_F\big(W(t)\big)\Big\}. \tag{11}$$

This stopping time relates directly to all facets of our subject matter: if F is the exponential distribution with mean d, shifted down by d so as to have mean zero, then $H_F(x) = x + d$ and $T_{A-Y} = \tau_d$. Since $W^2(t) - t$ is a mean-zero Martingale, $\mathbb{E}[\tau_d] = \mathbb{E}[W^2(\tau_d)] = \mathrm{Var}(W(\tau_d)) = d^2$, proving the first statement of (9) up to an obvious change of scale. As mentioned above, the exponentiality of the embedded distribution holds for general μ.

Secondly, among all uniformly integrable Martingales with a given final or limiting distribution, SBM stopped at the Azéma–Yor stopping time to embed that distribution is extremal, in the sense that it stochastically maximizes the essential maximum of the Martingale (see Dubins & Gilat [9] and Azéma & Yor [2]). That is, if T_{A-Y} embeds F then $M_{T_{A-Y}}$ is stochastically bigger than the maximum of any such Martingale. Hence, if the payoff upon stopping is a non-decreasing function of M and the sampling cost is a function of the stopping time, then optimal stopping is always achieved by a T_{A-Y}, because the replacement of any other stopping time by the T_{A-Y} that embeds the same distribution will preserve the distribution of the cost while increasing stochastically that of the payoff.

The connection of the Azéma–Yor stopping time to the Chacon–Walsh family becomes apparent (see Meilijson [15]) if the random variable Y has finite support $\{x_1 < \cdots < x_k\}$. In this case, let $\mathcal{F}_n$ be the σ-field generated by $\min(Y, x_{n+1})$, that is, let the atoms of Y be incorporated one at a time, in their natural order: the first stage decides whether $Y = x_1$ (by stopping there) or otherwise (by temporarily stopping at $\mathbb{E}[Y \,|\, Y > x_1]$), etc. This is precisely the Azéma–Yor stopping rule: stop as soon as a value of Y is reached after having visited the conditional expectation of Y from this value and up.

Sequences of first exit times from open intervals will play a major role throughout this article, starting from Theorem 2. Analytical results for first exit times from open intervals are summarized in Lemma 2.

Turning now to the main subject of this paper, Dubins & Schwarz [10] considered the following optimal stopping problem for $\mu = 0$: letting $c > \mu$, find a stopping time that maximizes $\mathbb{E}[M(\tau) - c\tau]$. They proved that the best τ_d is optimal. We state their result, extended to general μ, as Theorem 1. For $\mu = 0$ it is a special case of the following Theorem 3, where the payoff upon stopping is a general non-decreasing function ϕ of the record highest value of Brownian Motion so far. Its statement and proof relies heavily on the Azéma–Yor stopping times, that provide, as described above, the structure of optimal solutions.

Theorem 1. *Let* $c > \mu$ *and* $d = \dfrac{\sigma^2}{2\mu} \log \dfrac{c}{c-\mu}$ *if* $\mu > 0$ *or its limit* $d = \dfrac{\sigma^2}{2c}$ *(as* $\mu \downarrow 0$*) if* $\mu = 0$*. Then*

$$\sup_\tau \mathbb{E}[M(\tau) - c\tau] = \mathbb{E}[M(\tau_d) - c\tau_d] = \frac{\sigma^2}{2\mu}\Big(\frac{c}{\mu} \log \frac{c}{c-\mu} - 1\Big) \tag{12}$$

where the last expression is to be interpreted as its limit $\dfrac{\sigma^2}{4c}$ *if* $\mu = 0$*.*

These theorems analyze what can be roughly seen as pricing and managing an insurance option against a drop in the value of a stock, whose premium consists in the payment of an interest exceeding the drift of the held stock. The previous theorem deals with the stock itself, the next ones with a general monotone derivative. No attempt is made here at analyzing arbitrage pricing, only optimal expected stopped value of this look-back time-unconstrained American put option.

Theorem 2. *Let ϕ be a right-continuous, non-decreasing piece-wise constant function with values $0 = \phi_0 < \phi_1 < \phi_2 < \cdots < \phi_N$ such that $\phi(x) = \phi_i$ for $a_i \leqslant x < a_{i+1}$, where $0 = a_0 < a_1 < \cdots < a_N < a_{N+1} = \infty$. Let $c > 0$. Consider the optimization problem: find a stopping time τ^* on BM B that maximizes $\mathbb{E}[\phi(M(\tau)) - c\tau]$. Define $E_N = \phi_N$ and for $n = N-1, N-2, \ldots, 0$, using the notation N and d to be introduced in Corollary 1, let*

$$E_n = \phi_n + N(\mu, a_{n+1} - a_n, E_{n+1} - \phi_n, 0) \tag{13}$$
$$x_n = a_n - d(\mu, a_{n+1} - a_n, E_{n+1} - \phi_n) \tag{14}$$

Then $\mathbb{E}[\phi(M(\tau^*)) - c\tau^*] = E_0$ and τ^* may be defined as follows. Letting B start at $a_0 = 0$, wait for the first exit time of B from the interval (x_0, a_1). If exit occurs at the top, wait for the first exit time of B from (x_1, a_2), etc. If ever exit occurs at the bottom, stop. Remark: $x_i = a_i$ is to be interpreted as instantaneous exit at the bottom.

For $\mu = 0$ and $\sigma = 1$, the explicit representation of E_n and x_n is

$$\begin{cases} E_n = \phi_n + \left(\left(\sqrt{E_{n+1} - \phi_n} - \sqrt{c}\,(a_{n+1} - a_n)\right)^+\right)^2, \\ x_n = a_n - \sqrt{(E_n - \phi_n)/c}\,. \end{cases} \tag{15}$$

More generally, if the process W starts at x and its record high value so far is $y \geqslant x$ (thus, the payoff under immediate stopping is $\phi(y) = \phi_n$ for some n), then it is best to stop immediately if and only if $x \leqslant x_n$. The optimal expected payoff at this initial state is

$$E(x, y) = \phi_n + \left(\left(\sqrt{E_{n+1} - \phi_n} - \sqrt{c}(a_{n+1} - x)\right)^+\right)^2. \tag{16}$$

A corresponding expression may be easily derived under $\mu \neq 0$.

Theorem 3. *Let $c > 0$. Let ϕ be a right-continuous, non-negative, non-decreasing function on $[0, \infty)$, such that $\phi(W(t)) - c\,t$ is a.s. negative on (t_0, ∞) for some (random) t_0, and its supremum*

$$S_c = \sup_t \big(\phi\big(W(t)\big) - c\,t\big) \tag{17}$$

is integrable (e.g., $\limsup_{x\to\infty} \phi(x)/x^{2-\varepsilon} < \infty$ for some $\varepsilon > 0$). Consider the problem of finding a stopping time τ^* on SBM W that maximizes $\mathbb{E}[\phi(M(\tau)) - c\tau]$. More generally, let $H : [0, \infty) \to \mathbb{R}$ be defined as

$$H(x) = \sup_\tau \mathbb{E}\big[\phi\big(x + M(\tau)\big) - c\tau\big]. \tag{18}$$

Then

(i) The absolutely continuous function H is the minimal solution of the differential equation

$$H(x) - \frac{1}{4c}\big(H'(x)\big)^2 = \phi(x). \tag{19}$$

If ϕ is constant on $[x_0, \infty)$, H is the unique solution of (19) that equals ϕ on $[x_0, \infty)$.
The generalization of (16), that is, the optimal expected payoff when W is at x but has been in the past at a record high value $y \geqslant x$, is

$$E(x,y) = \phi(y) + \left(\left(\sqrt{H(y)-\phi(y)} - \sqrt{c}(y-x)\right)^+\right)^2. \tag{20}$$

(ii) *The following Azéma–Yor-type stopping time is optimal for the original problem (with $x = 0$): Stop as soon (t) as the gap $M(t) - W(t)$ reaches the value $H'(M(t))/2c$.*

(iii) *For every distribution F with mean zero and finite variance there is a non-decreasing function ϕ for which the Azéma–Yor stopping time to embed that distribution in SBM, is the stopping time described in* (ii)*. This function ϕ is displayed in (19), taking $H'(x) = 2c(x - H_F^{-1}(x))$, where H_F^{-1} is the right-continuous inverse of the Hardy-Littlewood upper barycenter function H_F (see (11)). Even if ϕ does not make S_c integrable, the Azéma–Yor stopping time is optimal for this ϕ in the weaker sense where the supremum in (18) is taken over the* integrable *stopping times only.*

Now let $\limsup_{x\to\infty} \phi(x)/x^2 > c$*. Then*

$$\sup_\tau \mathbb{E}\big[\phi\big(M(\tau)\big) - c\tau\big] = \infty \tag{21}$$

even if the supremum is taken over integrable stopping times only.

As a first example, consider $\phi(x) \equiv x$. Then the Dubins & Schwarz solution (see Theorem 1) τ_d with $d = 1/(2c)$, has $H(x) = x + 1/(4c)$, that satisfies (19) and embeds in SBM a shifted exponential distribution.

As a second example, consider piecewise constant functions ϕ as in Theorem 2. Then the piecewise quadratic solution H of (19) has values $\{E_n\}$ at the points of increase $\{a_n\}$. The optimal stopping time embeds in SBM a distribution supported by a finite set. This explicit solution provides a reasonable practical way of approximating optimal solutions for general ϕ, by discretizing ϕ. This discretization is a key to proving Theorem 3.

In particular, if $\phi(x) = V * \mathbf{1}_{[b,\infty)}(x)$, the optimal stopping time is a first exit time from an interval, i.e., it embeds in SBM a dichotomous distribution whose upper atom is b (see Corollary 1 in the next section).

As an informal third example, let us try to get formally a quadratic H, (w.l.g.) $H(x) = x^2/2 + ax + a^2/(2c)$. Just as the Dubins & Schwarz case of linear H corresponds to Azéma–Yor embedding of (exponentially tailed) exponential distributions, with constant mean residual function, quadratic H would correspond to Azéma–Yor embedding of (regularly varying tailed) Pareto distributions $1 - F(x) = x^{-\alpha}$, with linear mean residual function. For some values of c (or α), this distribution has infinite variance, so the stopping time has infinite mean. Technically, equation (19) yields $\phi(x) = (1 - 1/c)(x^2/2 + ax)$, non-negative and increasing for $c > 1$ and $a > 0$, conveniently satisfying $\phi(0) = 0$, improved by would-be optimal stopping to $H(0) = a^2/(2c)$, but inconveniently failing to satisfy the integrability assumption of Theorem 3. However, by the Law of the Iterated Logarithm, if ϕ is quadratic, the supremal payoff is a.s. infinite. The Azéma–Yor solutions identify the supremum in (18) over all *integrable* stopping times, and this restricted supremum does not coincide with the more general one without some dominance assumption such as integrability of S_c.

The last section is devoted to further discussion on this issue.

1 Exponential Martingales and first exit times from open intervals

The analysis to be performed relies on the following well known facts, most of which originate with Itô & McKean [13] and are taken from Borodin & Salminen [4]. A method of proof can be taken from Itô & McKean [13] or from Karlin & Taylor [14]. Lemma 1 is a direct consequence of the formula for the moment generating function of a Gaussian random variable, via the fact that BM has independent increments. Lemma 2 contains explicit formulas for the Brownian Gambler's Ruin Problem.

Lemma 1. *For every $\lambda \in \mathbb{R}$, $\exp(\lambda(B(t) - \mu t) - (\sigma^2/2)\lambda^2 t)$ is a mean-1 Martingale and its derivatives with respect to λ are mean-0 Martingales.*

Lemma 2. *Assume that $\mu > 0$. Let $a < b$ and consider the first exit time $H_{a,b}$ from the open interval (a, b) by BM $x+B$, where the starting point x is tacitly assumed to be in the interval (a, b). Let $\mathbf{1}_b$ be the indicator function of the event that $x + B$ exits the interval at the upper endpoint b, and let $\mathbb{E}[Y;\, A]$ be understood to be the expectation of the product of Y with the indicator function of the event A, that is, $\mathbb{E}[Y;\, A] = \mathbb{E}[Y \mid A]\,\mathbb{P}(A)$. Then*

$$\mathbb{E}_x[\mathbf{1}_b] = \exp\Big(\frac{\mu}{\sigma^2}(b - x)\Big)\frac{\sinh((x - a)\frac{|\mu|}{\sigma^2})}{\sinh((b - a)\frac{|\mu|}{\sigma^2})} \tag{22}$$

$$1 - \mathbb{E}_x[\mathbf{1}_b] = \exp\Big(\frac{\mu}{\sigma^2}(a - x)\Big)\frac{\sinh((b - x)\frac{|\mu|}{\sigma^2})}{\sinh((b - a)\frac{|\mu|}{\sigma^2})} \tag{23}$$

$$\mathbb{E}_x[\exp(\theta H_{a,b});\, \mathbf{1}_b = 1] = \exp\Big(\frac{\mu}{\sigma^2}(b - x)\Big)\frac{\sinh((x - a)\frac{|\mu|}{\sigma^2}\sqrt{1 - 2\theta\sigma^2/\mu^2})}{\sinh((b - a)\frac{|\mu|}{\sigma^2}\sqrt{1 - 2\theta\sigma^2/\mu^2})} \tag{24}$$

$$\mathbb{E}_x[\exp(\theta H_{a,b});\, \mathbf{1}_b = 0] = \exp\Big(\frac{\mu}{\sigma^2}(a-x)\Big)\frac{\sinh((b-x)\frac{|\mu|}{\sigma^2}\sqrt{1-2\theta\sigma^2/\mu^2})}{\sinh((b-a)\frac{|\mu|}{\sigma^2}\sqrt{1-2\theta\sigma^2/\mu^2})} \quad (25)$$

$$\mathbb{E}_x[H_{a,b}] = \frac{1}{\mu}\big((b-a)\mathbb{E}_x[\mathbf{1}_b] - (x-a)\big) \quad (26)$$

The moment generating function $\mathbb{E}_x[\exp(\theta H_{a,b})]$ *is obtained by adding (24) and (25). The conditional moment generating functions* $\mathbb{E}_x[\exp(\theta H_{a,b}) \,|\, \mathbf{1}_b]$ *are, thus, invariant under a change of sign of* μ *(!):*

$$\mathbb{E}_x[\exp(\theta H_{a,b}) \,|\, \mathbf{1}_b = 1] = \frac{\sinh((x-a)\frac{|\mu|}{\sigma^2}\sqrt{1-2\theta\sigma^2/\mu^2})}{\sinh((b-a)\frac{|\mu|}{\sigma^2}\sqrt{1-2\theta\sigma^2/\mu^2})}\frac{\sinh((b-a)\frac{|\mu|}{\sigma^2})}{\sinh((x-a)\frac{|\mu|}{\sigma^2})} \quad (27)$$

$$\mathbb{E}_x[\exp(\theta H_{a,b}) \,|\, \mathbf{1}_b = 0] = \frac{\sinh((b-x)\frac{|\mu|}{\sigma^2}\sqrt{1-2\theta\sigma^2/\mu^2})}{\sinh((b-a)\frac{|\mu|}{\sigma^2}\sqrt{1-2\theta\sigma^2/\mu^2})}\frac{\sinh((b-a)\frac{|\mu|}{\sigma^2})}{\sinh((b-x)\frac{|\mu|}{\sigma^2})} \quad (28)$$

In particular, for $x = (a+b)/2 = 0$*, expressions simplify via* $\sinh(2t) = 2\sinh(t)\cosh(t)$ *into*

$$\mathbb{E}_0[\exp(\theta H_{-d,d}) \,|\, \mathbf{1}_b] = \frac{\cosh\big((|\mu|/\sigma^2)d\big)}{\cosh\big((|\mu|/\sigma^2)d\sqrt{1-2\,\theta\sigma^2/\mu^2}\big)}\,. \quad (29)$$

The limiting values of the moment generating functions (27) and (28) and the corresponding expected time as the initial point x *tends to an endpoint of the interval are*

$$\begin{aligned}\mathbb{E}_{a+}[\exp(\theta H_{a,b}) \,|\, \mathbf{1}_b = 1] &= \mathbb{E}_{b-}[\exp(\theta H_{a,b}) \,|\, \mathbf{1}_b = 0] \\ &= \frac{\sinh\big((b-a)(|\mu|/\sigma^2)\big)\sqrt{1-2\theta\sigma^2/\mu^2}}{\sinh\big((b-a)(|\mu|/\sigma^2)\sqrt{1-2\theta\sigma^2/\mu^2}\big)} \quad (30)\end{aligned}$$

$$\begin{aligned}\mathbb{E}_{a+}[H_{a,b} \,|\, \mathbf{1}_b = 1] &= \mathbb{E}_{b-}[H_{a,b} \,|\, \mathbf{1}_b = 0] \\ &= \frac{b-a}{|\mu|}\coth\bigg((b-a)\frac{|\mu|}{\sigma^2}\bigg) - \frac{\sigma^2}{\mu^2} \quad (31)\end{aligned}$$

The limiting behavior of (29) as $d \downarrow 0$ *is free of* μ *and depends quadratically on* d*, as should be expected from the non-differentiability and quadratic variation regularity of Brownian Motion paths:*

$$\lim_{d\downarrow 0}\frac{\mathbb{E}_0[\exp(\theta H_{-d,d}) \,|\, \mathbf{1}_b] - 1}{d^2} = \frac{\theta^2}{\sigma^2}. \quad (32)$$

Expression (31) is asymptotic to $(b-a)/|\mu| - \sigma^2/\mu^2$ *as* $b-a\to\infty$ *and to* $(|\mu|/\sigma^2)(b-a)^2$ *as* $b-a\to 0$*. The limiting value of (23) as* $b\to\infty$ *and* x*,* a *and* $d = x-a$ *are kept fixed, is*

$$1 - \lim_{b\uparrow\infty}\mathbb{E}_x[\mathbf{1}_b] = \mathbb{P}\Big\{\min_{0\leqslant t<\infty} B(t) < -d\Big\} = \exp\Big(-\frac{2\mu}{\sigma^2}d\Big), \quad (33)$$

showing that the maximum $M(\infty)$ *of BM with* negative *drift* μ *is exponentially distributed with mean* $\mathbb{E}[M(\infty)] = \sigma^2/(2|\mu|)$*.*

2 Proofs of the main results

Proof of Proposition 1. To prove the exponentiality of the stationary distribution of the gap process G (well known in Queueing Theory circles as the geometric stationary distribution of queue size in the M/M/1 queue), realize that the answer is exactly provided by the last statement of Lemma 2: At arbitrary but large t, the gap exceeds d iff the process with time reversed (i.e., BM with drift $-\mu$) ever reaches height d over its value at time t.

The stationary distribution of the downfall process X is obtained as follows: The stationary probability $F_X(d)$ should be equal to the fraction of time the process spends below d. Since the sample paths increase continuously and jump down to 0, this fraction is the ratio of the expected time $\mathbb{E}[\tau_d]$ to the expected cycle time $\mathbb{E}[\tau_d] + d/\mu$, where d/μ is the expected time it takes B to reach again height M_d and start all renewally over. This proves (10). The expectation of this distribution is

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}\{X > t\}\,\mathrm{d}t$$
$$= \int_0^\infty \frac{(2\mu/\sigma^2)t}{\exp\big((2\mu/\sigma^2)t\big) - 1}\,\mathrm{d}t = \frac{\sigma^2}{2\mu}\int_0^\infty \frac{t}{\exp(t) - 1}\,\mathrm{d}t$$

that via the change of variable $x = 1 - \exp(-t)$ becomes

$$\mathbb{E}[X] = -\frac{\sigma^2}{2\mu}\int_0^1 \frac{\log(1-x)}{x}\,\mathrm{d}x$$
$$= \frac{\sigma^2}{2\mu}\int_0^1 \Big(1 + \frac{x}{2} + \frac{x^2}{3} + \frac{x^3}{4} + \cdots\Big)\,\mathrm{d}x = \frac{\sigma^2}{2\mu}\sum_{n=1}^\infty \frac{1}{n^2} = \frac{\sigma^2}{2\mu}\frac{\pi^2}{6},$$

proving the statement. As a source for a proof that $\zeta(2) = \pi^2/6$, see for instance Edwards & Penney [11], page 582. □

As a corollary of Lemma 2, we prove the simplest case of the optimal stopping problem, the case of step-function ϕ.

Corollary 1. *Let $c > 0$ be the cost per unit time until stopping BM B with drift μ and diffusion parameter σ, starting at $x < b$, and let $V > 0$ be the reward if this process reaches b before stopping. Then it is optimal to stop at the first exit time $H_{a,b}$ from the open interval (a, b), where $a = b - d$ and $d = d(\mu, b, V)$ is $\sigma\sqrt{V/c}$ when $\mu = 0$, and the unique solution of $c\,\mathbb{E}[\tilde{\tau}_d] = V$ when $\mu \neq 0$, with $\tilde{\tau}_d$ (see (7)) being the time it takes $-B$ to develop a drawdown d. In particular, if $x \leqslant a$, stop at time zero. The optimal expected net reward is*

$$N(\mu, b, V, x) = c\bigg(\frac{x-a}{\mu} - \frac{\sigma^2}{\mu^2}\exp\Big(-\frac{\mu d}{\sigma^2}\Big)\sinh\Big(\frac{\mu}{\sigma^2}(x-a)\Big)\bigg)^+, \quad \mu \neq 0 \quad (34)$$

$$N(0, b, V, x) = \bigg(\Big(\sqrt{V} - \frac{b-x}{\sigma}\sqrt{c}\Big)^2\bigg)^+ \quad (35)$$

Sketch of a proof. This problem has two features that make it easily solvable (see [6]): incremental future payoffs are uniformly bounded from above and unstopped payoffs go to $-\infty$ a.s. As such, if the payoff function of a strategy (as a function of the initial state) is uniformly bounded and *excessive* then the strategy is optimal. Excessivity means that proceeding for a short time against the strategy and then following the strategy, is never better than following the strategy from the beginning. This shows that a first exit time from an interval is optimal: while inside the interval, it is better not to stop—by the construction of the interval, that follows—; while outside the interval by more than some small ε, if the strategy is followed (that is, stop!) only upon a first change by ε, the positive cost of the elapsed time is a pure loss because there is no chance of getting the reward. Hence, it is always preferable to heed to the strategy from the beginning. It remains to find the best choice of a. The expected net reward is $V\,\mathbb{E}_x[\mathbf{1}_b] - c\,\mathbb{E}_x[H_{a,b}]$ (see (22) and (26)), and its partial derivative with respect to a is

$$\frac{c}{\mu}(1-\mathbb{E}_x[\mathbf{1}_b])\left(\left(b-a-V\,\frac{\mu}{c}\right)\frac{\partial}{\partial a}\log(1-\mathbb{E}_x[\mathbf{1}_b])-1\right).$$

The logarithm of expression (23) is readily seen to have a derivative with respect to a that is independent of x. Hence, *the best* a is independent of x as well, an obvious property that the solution must satisfy. Existence and uniqueness of the solution is due to the claimed representation $c\,\mathbb{E}[\tilde{\tau}_d] = V$ (that is easily obtained by differentiating (23) with respect to a), since $\mathbb{E}[\tilde{\tau}_d]$ continuously strictly increases from 0 to ∞. The expected net reward $V\,\mathbb{E}_x[\mathbf{1}_b] - c\,\mathbb{E}_x[H_{a,b}]$ obviously goes to zero as $x \downarrow a$ because each summand does. It is positive because, by (26),

$$\frac{\mathbb{E}_x[H_{a,b}]}{\mathbb{E}_x[\mathbf{1}_b]} = \frac{(b-a)}{\mu} - \frac{x-a}{\mu\,\mathbb{E}_x[\mathbf{1}_b]},$$

and by (22), this ratio is an increasing function of x.

Proof of Theorem 1. The first observation is that the objective function is bounded. In fact, it is uniformly bounded in μ and τ when $c-\mu>0$ is kept constant: $M(\tau) - c\tau \leqslant M(\tau) - cL(\tau) \leqslant M^*(\infty)$, where M^* is the M process for the BM $B(t) - c\,t$, that has drift $\mu - c < 0$. The last statement of Lemma 2 proves that $\mathbb{E}[M^*(\infty)] = \sigma^2/(2(c-\mu)) < \infty$.

The second observation is that the value $V = \sup_\tau \mathbb{E}[M(\tau) - c\tau]$ is strictly positive. This is so because $\tau = \tau_d$ yields expected payoff (see (8)) $d - (c - \mu)\,\mathbb{E}[\tau_d]$, that is strictly positive for small d because $\mathbb{E}[\tau_d]$ is asymptotically quadratic in d as $d \downarrow 0$.

Now, this expected payoff function of d is strictly concave, increases at zero and has value zero at zero, and goes to $-\infty$ as $d \to \infty$. Hence, it has a unique maximum, located at the value of d as claimed in the statement. To see that τ_d is optimal, let d' be such that $cd'/\mu = V$. If at any moment the gap exceeds d', it will cost an expected amount *more than* V to wait for zero gap, and then re-start the process to obtain the optimal payoff V. Thus, the expected payoff is strictly higher by stopping now. On the other hand (see the remarks on excessivity in the sketch of a proof of Corollary 1), if the gap is less than d', there is a good rationale for not stopping: Wait for zero gap and then do whatever it takes to achieve V. Hence, $\tau_{d'}$ must be optimal. The equality $V = cd'/\mu$, with $d' = d$, yields (12). □

Proof of Theorem 2. This proof of Theorem 2 consists of the usual Dynamic Programming backwards-induction steps, each step in this case being a problem as in Corollary 1:

If for some reason stopping hasn't occurred until B reached a_{N-1}, then the decision maker (that already paid for elapsed time and collected the reward ϕ_{N-1}) is faced precisely with a Corollary 1-type problem, starting at $x = 0$ and aiming for $b = a_N - a_{N-1}$ in order to collect $V = \phi_N - \phi_{N-1}$. Corollary 1 identifies uniquely a lower bound a (denoted by $x_{N-1} - a_{N-1}$ in the notations of the statement of Theorem 2) such that sampling B is performed if and only if $0 = x > a$.

If $x \leqslant a$, then the problem has the same solution as if $\phi_N = \phi_{N-1}$ (i.e., ϕ_N is "removed"), to be inductively conceptualized.

If $x > a$, then Corollary 1 proves the statement of Theorem 2 as far as behavior beyond reaching a_{N-1} is concerned, and provides an equivalent problem with ϕ_N removed, ϕ_{N-1} replaced by E_{N-1} and all ϕ_i (with $i \leqslant N-2$), left unchanged. Apply inductively. □

Proof of Theorem 3. To see (21), consider such a non-negative ϕ and let $b_1 < b_2 < b_3 \cdots \to \infty$ be such that for some $\varepsilon > 0$, $\phi(b_i) > (c+\varepsilon)b_i^2$ for each $i = 1, 2, \ldots$ For each such i, let $\phi_i = (c+\varepsilon)b_i^2\, \mathbf{1}_{[b_i,\infty)}$. Since ϕ majorizes each ϕ_i, the optimal expected net reward for ϕ exceeds that of each ϕ_i. But corollary 1 shows that the latter—achieved by an integrable stopping time—is $((\sqrt{(c+\varepsilon)b_i^2} - b_i\sqrt{c})^+)^2 = b_i^2(\sqrt{c+\varepsilon} - \sqrt{c})^2$, that goes to ∞ with i.

Back to the dominated case, the first point of the proof is to show that the optimum may be identified as a limit of what can be achieved for bounded ϕ. Similarly to the first argument in the proof of Theorem 1, this is so because (see (17)), for $a > 0$,

$$\begin{aligned}\big(\phi\big(M(\tau)\big) - c\tau\big)\, \mathbf{1}_{\{M(\tau)>a\}} &\leqslant \big(\phi\big(M(\tau)\big) - c\tau\big)\, \mathbf{1}_{\{M(\tau)>a,\phi(M(\tau))-c\tau>0\}} \\ &\leqslant \big(\phi\big(M(\tau)\big) - cL(\tau)\big)\, \mathbf{1}_{\{M(\tau)>a,\phi(M(\tau))-c\tau>0\}} \\ &\leqslant S_c\, \mathbf{1}_{\{M(\tau)>a,\phi(M(\tau))-c\tau>0\}} \leqslant S_c\, \mathbf{1}_{Q_a}, \qquad (36)\end{aligned}$$

where Q_a is the event "$W(t)$ reaches a before $\phi(W(t)) - ct$ becomes negative forever". Since S_c is integrable by assumption and $\mathbb{P}(Q_a) \to 0$ as $a \to \infty$, it follows that $\limsup_{a\to\infty} \sup_\tau \mathbb{E}[(\phi(M(\tau)) - c\tau)\, \mathbf{1}_{\{M(\tau)>a\}}] \leqslant 0$.

To complete this stage of the proof by giving some explicit essence to the integrability condition, the statement in parentheses will be proved, namely, if $\phi(x) = x^{2-\varepsilon}$ for some $\varepsilon \in (0,2)$ then $\mathbb{E}[S_c] < \infty$ for all $c > 0$. In the sequel, Z stands for a standard normal random variable, whose density is f, its cumulative distribution function is Φ and its survival function is $\Phi^* = 1-\Phi$. We state for ease of reference the well known normal tail inequalities that hold for $x > 0$

$$\frac{f(x)}{x+1/x} < \Phi^*(x) < \frac{f(x)}{x}\,, \tag{37}$$

that follow from

$$0 < \mathbb{E}\big[(Z-x)^+\big] = f(x) - x\,\Phi^*(x) \tag{38}$$

$$0 < \mathbb{E}\Big[\big((Z-x)^+\big)^2\Big] = (x^2+1)\,\Phi^*(x) - xf(x). \tag{39}$$

Let $\{t_i\}$ be for now an arbitrary positive sequence, increasing to ∞. Evaluate

$$\begin{aligned}
\mathbb{E}\Big[\sup_t\{W_t^{2-\varepsilon} - c\,t\}\Big] &= \int_0^\infty \mathbb{P}\Big\{\sup_t\{W_t^{2-\varepsilon} - c\,t\} > a\Big\}\,\mathrm{d}a \\
&= \int_0^\infty \mathbb{P}\Big\{\exists t \in [0.\infty) \ni W_t > (a+c\,t)^{1/(2-\varepsilon)}\Big\}\,\mathrm{d}a \\
&\leqslant \sum_{i=1}^\infty \int_0^\infty \mathbb{P}\Big\{\exists t \in [t_i-1, t_i) \ni W_t > (a+c\,t_{i-1})^{1/(2-\varepsilon)}\Big\}\,\mathrm{d}a \\
&\leqslant \sum_{i=1}^\infty \int_0^\infty \mathbb{P}\Big\{\exists t \in [0, t_i) \ni W_t > (a+c\,t_{i-1})^{1/(2-\varepsilon)}\Big\}\,\mathrm{d}a \\
&= 2\sum_{i=1}^\infty \int_0^\infty \mathbb{P}\Big\{W_{t_i} > (a+c\,t_{i-1})^{1/(2-\varepsilon)}\Big\}\,\mathrm{d}a \\
&= \sum_{i=1}^\infty \int_0^\infty \mathbb{P}\Bigg\{\bigg|\frac{W_{t_i}}{\sqrt{t_i}}\bigg| > \frac{(a+c\,t_{i-1})^{1/(2-\varepsilon)}}{\sqrt{t_i}}\Bigg\}\,\mathrm{d}a \\
&= \sum_{i=1}^\infty \int_0^\infty \mathbb{P}\Big\{t_i^{(2-\varepsilon)/2}|Z|^{2-\varepsilon} - c\,t_{i-1} > a\Big\}\,\mathrm{d}a \\
&= \sum_{i=1}^\infty t_i^{(2-\varepsilon)/2}\,\mathbb{E}\Bigg[\bigg(|Z|^{2-\varepsilon} - \frac{c\,t_{i-1}}{t_i^{(2-\varepsilon)/2}}\bigg)^+\Bigg].
\end{aligned} \tag{40}$$

Let $x_i = c\,t_{i-1}/t_i^{(2-\varepsilon)/2}$ and use (38)–(39) and the following norm inequality

$$\left(\frac{\int_{x_i}^\infty z^{2-\varepsilon} f(z)\,\mathrm{d}z}{\int_{x_i}^\infty f(z)\,\mathrm{d}z}\right)^{1/(2-\varepsilon)} < \left(\frac{\int_{x_i}^\infty z^2 f(z)\,\mathrm{d}z}{\int_{x_i}^\infty f(z)\,\mathrm{d}z}\right)^{1/2} = \left(1 + \frac{x_i f(x_i)}{\Phi^*(x_i)}\right)^2 \tag{41}$$

to proceed with the evaluation (40), skipping straightforward details,

$$\mathbb{E}\Big[\sup_t\{W_t^{2-\varepsilon} - c\,t\}\Big] \leqslant \sum_{i=1}^{\infty} t_i^{(2-\varepsilon)/2}\big(x_i^2 - x_i + 2\big)\,\Phi^*(x_i). \tag{42}$$

Now a sequence $\{t_i\}$ must be displayed for which the RHS of (42) is finite. As is easy to see using the RHS of (37) to bound $\Phi^*(x_i)$, any exponential sequence $t_i = \gamma^i$ with $\gamma > 1$ will do it.

For discrete ϕ, consider the function $H(x) = E(x,x)$ (see (16)). It is continuous and piece-wise quadratic; furthermore, each such section is either increasing and strictly quadratic (with second derivative $2c$) or constant (with constant value equal to some ϕ_i. Such a section necessarily begins at a_i and ends strictly to the left of a_{i+1} if $i < N$). The first derivative $H'(x)$ at a breakpoint x between a constant and a quadratic section exists and equals zero, while at a breakpoint between two quadratic sections (necessarily some a_i), the derivative from the left is $\sqrt{E_i - \phi_{i-1}}$ while the derivative from the right is $\sqrt{E_i - \phi_i}$. This function, clearly meeting the definition (18), satisfies and is determined by (19).

It is important to note that if the range of ϕ (with finitely many values) is restricted to some interval $[\underline{\phi}, \overline{\phi}]$, then the range of H' is restricted to some interval of the form $[0, f(\overline{\phi} - \underline{\phi})]$.

Let a more general non-decreasing ϕ be restricted to such an interval. If ϕ is discretized on some grid, H is bounded between the two functions H_L and H_U corresponding to the (L)ower and (U)pper discretizations. Since these differ by at most the grid size, there is clearly convergence as the grid size tends to zero. Now the uniform boundedness of H' for discrete ϕ comes into play: By weak compactness, there is a function H' obtained as a limit of such functions corresponding to discrete cases, whose integral is H, such that (19) is satisfied. It is uniquely determined by (19) as well, by weak continuity.

Statement (*ii*) of Theorem 3 is a straightforward concise rephrasing of the stopping time. Once it is properly understood, the definition of H' in statement (*iii*) becomes apparent, and the only point still needing proof is the monotonicity of the function ϕ produced by (19) from H, that is expressible as

$$\phi(x) = \left(2\int_0^x \big(t - H_F^{-1}(t)\big)\,\mathrm{d}t - \big(x - H_F^{-1}(x)\big)^2\right)c. \tag{43}$$

It is easy to check that $2\int_0^x v(t)\,\mathrm{d}t - v^2(x)$ is non-decreasing for v non-negative such that $v(y) - v(x) \leqslant y - x$ for $x < y$. To prove these properties for $v(x) = x - H_F^{-1}(x)$, observe that $H_F(x) = \mathbb{E}[X \mid X \geqslant x]$ is non-decreasing and at least x. $\square$

3 Discussion and end of the proof of Theorem 3

The gap between sub-quadratic and quadratic-reaching ϕ is left partially open: we saw in the proof of Theorem 3 that for a quadratic ϕ with second derivative bigger than $2c$, the optimal payoff is infinite because the optimal payoffs of its bounded approximants are unbounded.

The question is whether this is the only way in which optimal payoffs can be infinite. More or less in other words, the question is whether for any mean-zero distribution with finite variance, the corresponding Azéma–Yor stopping time is optimal for the function ϕ defined by (19). The answer is *negative*: for a quadratic ϕ with second derivative less than $2c$, the bounded approximants converge to the Azéma–Yor stopping time that embeds in SBM the corresponding finite-variance Pareto distribution, and (19) is satisfied. However, this is not optimal! The Law of the Iterated Logarithm implies that for *any* quadratic ϕ the supremal expected payoff is infinite. In fact, there is even arbitrage (if we are somewhat sloppy on definitions): for any $K > 0$ there is a stopping time that guarantees a net payoff K, deterministically so, just as in SBM itself (wait until it visits K).

At best, then, what can be hoped for is that for every distribution with mean zero and finite variance, its Azéma–Yor embedding stopping time is optimal for the function ϕ defined by (19), and the supremum in (18) is finite, as long as this supremum is taken in the class of integrable stopping times. This holds clearly true, since for any stopping time τ with finite mean, monotone convergence implies that $\tau \wedge T_a$ becomes at least as good as τ as $a \to \infty$, where T_a is the first time $\phi(M) = a$ whenever well defined. But for this problem, with bounded ϕ, the optimum is provided as built in Theorems 2 and 3, that are improved by the Azéma–Yor stopping time we started with.

Acknowledgement. The author thanks Lester Dubins, David Gilat, Boris Tsirelson and Marc Yor for helpful comments and information.

References

1. Azéma, J. Sur les fermés aléatoires. *Sem. Prob. Strasb.* XIX, *Springer LN in Math.* 1123 (1985).
2. Azéma, J. and M. Yor. a. Une solution simple au problème de Skorokhod. b. Le problème de Skorokhod: compléments. *Sem. Prob. Strasb.* XIII, *Springer LN in Math.* 721 (1968).
3. Azéma, J. and M. Yor. Etude d'une martingale remarquable. *Sem. Prob. Strasb.* XXIII, *Springer LN in Math.* 1372 (1989).
4. Borodin, A. N. and P. Salminen. *Handbook of Brownian Motion - Facts and Formulae.* Probability and Its Applications, Birkhäuser Verlag: Basel (1996).
5. Chacon, R. V. and J. B. Walsh. One dimensional potential embedding. *Sem. Prob. Strasb.* X, *Springer LN in Math.* 511 (1976).
6. Chow, Y. S., H. Robbins and D. Siegmund. *Great expectations: the theory of optimal stopping.* Houghton–Mifflin: Boston (1971).
7. Douady, R., A. N. Shiryaev and M. Yor. On probability characteristics of "downfalls" in a standard Brownian Motion. *Theory Probab. Appl.* **44**, No. 1, pp. 29–38 (2000).

8. Dubins, L. E. On a theorem of Skorokhod. *Ann. Math. Statist.* **39**, pp. 2094-2097 (1968).
9. Dubins, L. E. and D. Gilat. On the distribution of maxima of Martingales. *Proc. Amer. Math. Soc.*, **68**, No. 3, pp. 337–338 (1978).
10. Dubins, L. E. and G. Schwarz. A sharp inequality for sub-martingales and stopping-times. *Société Mathématique de France, Astérisque* **157/8**, pp. 129-145 (1988).
11. Edwards, C. H. and D. E. Penney. *Calculus and Analytic Geometry.* Prentice Hall: New Jersey (1982).
12. Goldhirsch, I. and S. H. Noskovicz. The first passage time distribution in random random walk. *Physics Review*, **A42**, pp. 2047-2064 (1990).
13. Itô, K. and H. P. McKean. Diffusion processes and their sample paths. Springer-Verlag: Berlin and New York (1965).
14. Karlin, S. and H. M. Taylor. *A First Course in Stochastic Processes*, 2nd ed. Academic Press: New York (1975).
15. Meilijson, I. On the Azéma–Yor stopping time. *Sem. Prob. Strasb.* XVII, *Springer LN in Math.* 986, pp. 225-226 (1981/82).
16. Protter, P. *Stochastic integration and differential equations: a new approach.* Springer-Verlag: Berlin and Heidelberg (1990).
17. Skorokhod, A. *Studies in the theory of random processes.* Addison Wesley: Reading (1965).
18. Taylor, H. M. A stopped Brownian Motion formula. *Ann. Prob.* **3**, pp. 234–246 (1975).
19. Williams, D. On a stopped Brownian motion formula of H. M. Taylor. *Sem. Prob. Strasb.* X, *Springer LN in Math.* 511, pp. 235–239 (1976).