The Annals of Probability
1973, Vol. 1, No. 5, 772–787

ON THE AMOUNT OF VARIANCE NEEDED TO ESCAPE FROM A STRIP[1]

BY DAVID BLACKWELL AND DAVID FREEDMAN

University of California, Berkeley

On the intrinsic time-scale, a martingale with uniformly bounded increments escapes from a strip at the same asymptotic rate as Brownian motion.

We would like to dedicate this paper to the memory of our friend and colleague, Jimmie Savage.

1. Introduction. Let τ_b be the time at which standard Brownian motion starting from y first hits $\pm b$, where $|y| \leqq b$. As is known, (or see Section 7, below),

$$P\{\tau_b > v\} \approx \frac{4}{\pi} \cos \frac{\pi y}{2b} \exp\left(-\frac{\pi^2}{8b^2} v\right) \qquad \text{as} \quad v \to \infty\,, \tag{1}$$

where $\exp u = e^u$ and $f(v) \approx g(v)$ means $f(v)/g/(v) \to 1$. A more precise result, Fürth's formula, is known: a reference is (Feller (1968), page 359).

Our main result is that (1) nearly holds for all martingales with bounded increments, on the intrinsic time scale. To state this precisely, suppose $Y_1, Y_2 \cdots$ are random variables on a triple $(\Omega, \mathscr{F}, P)$ which satisfy

$$|Y_i| \leqq 1 \qquad \text{and} \qquad E(Y_i \mid Y_1, \cdots, Y_{i-1}) = 0\,. \tag{2}$$

Let

$$V_i = E(Y_i^2 \mid Y_1, \cdots, Y_{i-1})\,; \tag{3}$$

so V_i is a random variable, measurable on $(Y_1, \cdots, Y_{i-1})$, with $0 \leqq V_i \leqq 1$. It is the conditional variance of Y_i given the past.

(4) DEFINITION. The intrinsic time, or total amount of conditional variance W_b used by $Y_1, Y_2, \cdots$ in escaping the strip $\pm b$ is $\sum_1^{\tau_b} V_i$, where τ_b is the least n if any with $|Y_1 + \cdots + Y_n| > b$, and $\tau_b = \infty$ if there is no such n.

At first sight, τ_b is a natural measure of the time to escape, and W_b an artifice. Here is some motivation. Suppose *first* that $Y_1, Y_2, \cdots$ are independent and identically distributed, taking the values ± 1 with chance $\frac{1}{2}$ each. Then $\tau_b = W_b$, because Var $Y_i = 1$. According to (5) and (6) below, W_b is approximately exponential, with parameter $\beta = \pi^2/(8b^2)$. Suppose *second* that $Y_1, Y_2, \cdots$ are independent and identically distributed, with $|Y_i| \leqq 1$ and $E(Y_i) = 0$ and Var $Y_i > 0$.

Received August 11, 1972; revised February 22, 1973.

[1] Research sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, USAF, under Grant AFOSR-71-2100A. This manuscript is submitted for publication with the understanding that the United States government is authorized to reproduce and distribute reprints for governmental purposes.

AMS 1970 *Subject classifications.* Primary 60G45

Key words and phrases. Martingale, crossing time, hitting time.

Then $W_b = (\text{Var } Y_i)\tau_b$, because Var Y_i is constant. If the Y_i are small, for instance if $P(Y_i = 0)$ is near 1, then τ_b is enormous: but Var Y_i is small. And W_b is still approximately exponential with the same parameter β. Suppose *third* that $Y_1, Y_2, \cdots$ are independent, $|Y_i| \leqq 1$, and $E(Y_i) = 0$, and $\sum \text{Var } Y_i = \infty$. Then

$$W_b = \sum_{i=1}^{\tau_b} \text{Var } Y_i .$$

So moves with large variance contribute a large amount to W_b, and moves with small variance contribute a small amount. Again, W_b is approximately exponential with the same parameter β. *Finally*, replace the independence-mean 0 assumptions by (2). This still does not affect the approximate distribution of W_b, if you replace variance by conditional variance. The idea here is that different paths move at different speeds; you have to take this into account in W_b. Total conditional variance as a measure of intrinsic time was introduced by Lévy (1954), and used by Dubins and Savage (1965). It is also used in the invariance principle; one exposition is in Freedman (1971).

(5) THEOREM. *Suppose* $(Y_1, Y_2, \cdots)$ *satisfies* (2). *Define* W_b *by* (4). *If* $b \geqq 16$ *and* $v > 0$, *then*

$$P\{W_b > v\} \leqq A \exp(-k\beta v), \qquad \textit{where}$$

$A = (5b)^{\frac{1}{2}}$ *and* $k = 1 - 3b^{-\frac{1}{2}}$ *and* $\beta = \pi^2/(8b^2)$.

(6) THEOREM. *Suppose* $(Y_1, Y_2, \cdots)$ *satisfies* (2). *Define* τ_b *and* W_b *by* (4). *If* $b \geqq 16$ *and* $v > 0$, *then*

(a) $P\{\tau_b < \infty$ *and* $W_b \leqq v\} \leqq 1 - \theta(v)$ *where* $\theta(v) = \exp[-k\beta(v+1)]$, $k = 1 + 15b^{-\frac{2}{3}}$ *and* $\beta = \pi^2/(8b^2)$. *If also* $P\{\tau_b < \infty\} = 1$, *then*

(b) $P\{W_b > v\} \geqq \theta(v)$.

The cut-off $b \geqq 16$ is arbitrary: it does simplify calculations. However, (6) is false for $b \leqq 1$. Part (b) of (6) is immediate from (a). To see why conditions on τ_b are needed, think about Y_i which are identically 0. So $P\{W_b < v\} = 1$.

Theorems (5) and (6) hold for all processes satisfying (2), all $b \geqq 16$, and all $v > 0$. They imply

(7) If $v \to \infty$ and then $b \to \infty$,

$$\frac{b^2}{v} \log P\{W_b > v\} \to -\frac{\pi^2}{8}$$

uniformly in processes satisfying (2) with $P\{\tau_b < \infty\} = 1$. If $b \to \infty$ first, W_b/b^2 converges in distribution to the time for standard Brownian motion to escape from ± 1. This follows, for instance, from the invariance principle. But (7) seems to be quite different, for there we estimate very small tail probabilities, as $v \to \infty$.

In Section 7, we make some estimates like (1) for the random walk. If the Y_i are independent and ± 1 with equal small probability p, and 0 with large probability $(1 - 2p)$, then

(8) $$\lim_{v\to\infty} \frac{1}{v} \log P\{W_b > v\} > -\frac{\pi^2}{8b^2}.$$

Once you have two different limits, it is easy to make an example where

(9) $$\frac{1}{v} \log P\{W_b > v\} \quad \text{oscillates as} \quad v \to \infty .$$

Think of two distributions with different limits. Let $1 \leqq n_0 < n_1 < \cdots$ increase quickly. Let the Y_i be independent, having the first distribution for $n_{2k} \leqq i < n_{2k+1}$, and the second for $n_{2k+1} \leqq i < n_{2k+2}$.

The constant multipliers, like $A = (5b)^{\frac{1}{2}}$ in (5), probably are not very important. Their effect is drowned out, for large v, by small changes in the constant appearing inside the exponential. Our error terms inside the exponential, like $3b^{-\frac{1}{2}}$ in (5), are probably of the wrong order of magnitude: but they do go to 0 as b increases.

We started working on (5) to get another proof of a known result:

(10) COROLLARY. *Suppose* (2) *and* (3). *Then* $P\{\sup_n |Y_1 + \cdots + Y_n| < \infty$ *and* $\sum V_i = \infty\} = 0$.

PROOF. It is enough to show that for all $b = 16, 17, \cdots$

$$P\{|Y_1 + \cdots + Y_n| > b \text{ for no } n \text{ and } \sum V_i = \infty\} = 0 .$$

But definition (4) makes

$$\{|Y_1 + \cdots + Y_n| > b \text{ for no } n \text{ and } \sum V_i = \infty\} \subset \{W_b = \infty\} .$$

So (5) proves (10). □

In particular, $P\{\sum V_i = \infty\} = 1$ implies $P\{\tau_b < \infty\} = 1$. We remind you that

(11) FACT. $P\{\sum_1^\infty Y_i$ diverges and $\sum_1^\infty V_i < \infty\} = 0$.

This holds even for unbounded Y_i. A reference is Lemma (10) of Dubins–Freedman (1965).

We might also note:

(12) Suppose $Y_1, Y_2, \cdots$ satisfies (2). Define τ_b and W_b by (4). Suppose $P\{\tau_b < \infty\} = 1$. Then $b^2 \leqq E\{W_b\} \leqq (b + 1)^2$.

This is almost immediate, once you observe that $(Y_1 + \cdots + Y_n)^2 - (V_1 + \cdots + V_n)$ is a martingale. For a complete discussion, see Dubins (1972). It suggests there is no process which maximizes or minimizes $P\{W_b > v\}$.

Our proofs of (5) and (6) depend on a gambling technique of Dubins and Savage. The original reference is Theorem (2.12.1) of Dubins and Savage (1965). A secondary reference is Theorem (22) of Dubins and Freedman (1965). We will review the technique in Section 2. We prove some easy inequalities on trigonometric and exponential functions in Section 3. Then we can prove (5), in Section 4, and (6), in Section 5.

We discuss a related estimate in Section 6—we got it while trying for (5). Relax (2) to

$$E(Y_i \mid Y_1, \cdots, Y_{i-1}) = 0 \qquad \text{and} \qquad |Y_1 + \cdots + Y_{\tau_b}| \leqq b + 1$$

where τ_b and W_b are still defined by (4). Then

$$P\{W_b > v + 1\} \leqq \frac{(b+1)^2}{2b+1} \exp\left[-\frac{v}{(b+1)^2}\right]. \tag{13}$$

We have an example to show this bound is of the right order of magnitude.

2. Gambling. In this section, we present a special case of Theorem 2.12.1 of Dubins and Savage (1965). We will use these results to prove (5) and (6). We follow Section 2 of Dubins and Freedman (1965).

Let F be the set of fortunes, endowed with a σ-field Σ. Let u be a nonnegative Σ-measurable function on F. This u is the utility or payoff function: if you stop at $f \in F$, you win $u(f)$; if you never stop, you win nothing. For each fortune f, there is a set $\Gamma(f)$ of available gambles γ, each γ being a (countably additive) probability on Σ. If your present fortune is f, you are allowed to select any $\gamma \in \Gamma(f)$, to pick a new fortune f' at random from F according to γ, and to move to f'. A *Γ-process starting from f* on the triple $(\Omega, \mathscr{F}, P)$ is an F-valued stochastic process $\phi_0, \phi_1, \cdots$ such that: ϕ_0 is f almost surely; and the conditional distribution $D_{n+1}(\phi_0, \cdots, \phi_n)$ of ϕ_{n+1} given $\phi_0, \cdots, \phi_n$ is an element of $\Gamma(\phi_n)$ almost surely. Think of $\phi_0, \phi_1, \cdots$ as your sequence of fortunes when you start from f, and use the strategy defined by $\phi_0, \phi_1, \cdots$. A strategy is a rule for selecting gambles based on past results, and the strategy defined by $\phi_0, \phi_1, \cdots$ is to use $D_{n+1}(\phi_0, \cdots, \phi_n) \in \Gamma(\phi_n)$ when you have moved through the fortunes $\phi_0, \cdots, \phi_n$. A *stopping time* s is a random variable taking the values $0, 1, 2, \cdots, \infty$: where the set $\{s = n\}$ must be an element of the σ-field generated by $\phi_0, \cdots, \phi_n$, for all $n = 0, 1, 2, \cdots$. Informally, you decide to stop at $s = n$ by looking at $\phi_0, \cdots, \phi_n$: and you win $u(\phi_n)$. You win nothing on $\{s = \infty\}$.

The next proposition gives an upper bound on $E\{u(\phi_s)\}$, valid for all Γ-processes ϕ and stopping times s.

(14) PROPOSITION. *Let Q be a real-valued, Σ-measurable function on F. Suppose that for all $f \in F$,*

(a) $Q(f) \geqq u(f)$

(b) $Q(f) \geqq \int_F Q(f')\gamma(df')$ *for all* $\gamma \in \Gamma(f)$.

Then for each fortune f, each Γ-process $\{\phi_0, \phi_1, \cdots\}$ starting from f, and each stopping time s:

$$Q(f) \geqq E\{u(\phi_s)\},$$

when $u(\phi_s) = 0$ on $\{s = \infty\}$.

PROOF. Condition (b) makes $Q(\phi_0), Q(\phi_1), \cdots$ an expectation-decreasing martingale (that is, a submartingale). To see this clearly, remember that $D_{n+1}(\phi_0, \cdots, \phi_n) = \gamma_n(\cdot, \cdot)$ say is a regular conditional distribution for ϕ_{n+1} given $\phi_0, \cdots, \phi_n$. So $\gamma_n(\omega, \cdot)$ is a countably additive probability on (F, Σ), and $\int Q(f')\gamma_n(\cdot, df')$ is a version of $E(Q(\phi_{n+1}) | \phi_0, \cdots, \phi_n)$. But $\gamma_n(\omega, \cdot) \in \Gamma(\phi_n(\omega))$ for almost all ω, because $\phi_0, \phi_1, \cdots$ is a Γ-process. For such ω, condition (b) ensures $\int_F Q(f')\gamma_n(\omega, df') \leqq Q[\phi_n(\omega)]$. If A is in the σ-field generated by $\phi_0, \cdots, \phi_n$,

$$\begin{aligned}\int_A Q(\phi_{n+1})dP &= \int_A E(Q(\phi_{n+1}) | \phi_0, \cdots, \phi_n)dP \\ &= \int_A \int_F Q(f')\gamma_n(\cdot, df')dP \\ &\leqq \int_A Q(\phi_n)dP .\end{aligned}$$

So $Q(\phi_0), Q(\phi_1), \cdots$ is an expectation decreasing martingale. Now s is a stopping time, so

$$\int Q(\phi_{s\wedge n})dP \leqq Q(\phi_0) .$$

Define $Q(\phi_s) = 0$ on $\{s = \infty\}$. So $Q(\phi_s) \leqq \liminf_{n\to\infty} Q(\phi_{s\wedge n})$. And Fatou makes

$$\int Q(\phi_s)dP \leqq Q(\phi_0) .$$

But $u(\phi_s) \leqq Q(\phi_s)$: use (a) on $\{s < \infty\}$ and the convention that $u(\phi_s) = 0$ on $\{s = \infty\}$. So

$$E\{u(\phi_s)\} \leqq E\{Q(\phi_s)\} \leqq Q(\phi_0) = Q(f) . \tag{15}$$

Here, $Q(\phi_0) = Q(f)$ because $\phi_0, \phi_1, \cdots$ starts from f. □

To prove (5) and (6), we let F be the set of pairs (v, y) of real numbers with $-1 \leqq v < \infty$ and $|y| \leqq b + 1$. We endow F with the Borel σ-field Σ. There are two kinds of fortunes: the *stuck positions* (v, y) with $v < 0$ or $|y| > b$, and the *moving positions* (v, y) with $v \geqq 0$ and $|y| \leqq b$. If you reach a stuck position (v, y), you stay there: so $\Gamma(v, y)$ has just one element, point mass at (v, y). To define $\Gamma(v, y)$ for moving positions (v, y), consider pairs (V, Y) which satisfy

(16) Y is random variable with mean 0 and absolute value at most 1; the real number V is the variance of Y.

From a moving position (v, y), you can move to $(v - V, y + Y)$ for any (V, Y) satisfying (16). So we put $\gamma \in \Gamma(v, y)$ iff γ is the distribution of $(v - V, y + Y)$ for some (V, Y) satisfying (16).

The Γ-process we use is defined in terms of the process $Y_1, Y_2, \cdots$ of Section 1. Fix $b \geqq 16$ and $v \geqq 0$. Let s be the least n if any with

$$|Y_1 + \cdots + Y_n| > b \quad \text{or} \quad V_1 + \cdots + V_n > v ;$$

let $s = \infty$ if there is no such n. Let $f = (v, 0)$.

Let

$$\begin{aligned}\phi_n &= (v - V_1 - \cdots - V_n, Y_1 + \cdots + Y_n) &&\text{for } n \leqq s \\ &= \phi_s &&\text{for } n > s .\end{aligned}$$

Conditions (2)—(3) ensure that $\phi_0, \phi_1, \cdots$ is a Γ-process starting from $(v, 0)$. Clearly, s is a stopping time, and $s \leqq \tau_b$.

To prove (5), we introduce the utility function u

$$u(v, y) = 1 \quad \text{for} \quad v < 0$$
$$= 0 \quad \text{for} \quad v \geqq 0 .$$

We say $u(\phi_s)$ is 1 if $W_b > v$, and is 0 if $W_b \leqq v$. Indeed, suppose $W_b > v$. Then $V_1 + \cdots + V_n > v$ for some n, so $s < \infty$. If s is the least n with $|Y_1 + \cdots + Y_n| > b$, then $s = \tau_b$ and $V_1 + \cdots + V_s = W_b > v$. If s is the least n with $V_1 + \cdots + V_n > v$, still $V_1 + \cdots + V_s > v$. Either way, the first coordinate $v - V_1 - \cdots - V_s$ of ϕ_s is negative, and $u(\phi_s) = 1$. On the other hand, suppose $W_b \leqq v$. Then $V_1 + \cdots + V_s \leqq W_b \leqq v$, because $s \leqq \tau_b$, so $u(\phi_s) = 0$ when s is finite. Consequently,

$$P\{W_b > v\} = E\{u(\phi_s)\} .$$

We will produce a Q which satisfies (14a)—(14b), for which $Q(v, 0)$ is the bound in (5). So (14) proves (5).

To prove (6), we introduce the utility function

$$u(v, y) = 1 \quad \text{for} \quad v \geqq 0 \quad \text{and} \quad |y| > b$$
$$= 0 \quad \text{elsewhere} .$$

Then $u(\phi_s)$ is 1 if $\tau_b < \infty$ and $W_b \leqq v$; otherwise $u(\phi_s)$ is 0. So

$$P\{\tau_b < \infty \quad \text{and} \quad W_b \leqq v\} = E\{u(\phi_s)\} .$$

We will produce a Q which satisfies (14a)—(14b), for which $Q(v, 0) = 1 - \theta(v)$. So (14) proves (6).

Our Q's were suggested by (1). Let

$$Q^*(v, y) = q(y) \exp(-\beta v) ,$$

where

$$q(y) = \cos \frac{\pi y}{2b} \quad \text{and} \quad \beta = \frac{\pi^2}{8b^2} .$$

Then

(17) $E[Q^*(v - V, y + Y)]$ is nearly $Q^*(v, y)$ for moving positions (v, y) and (V, Y) satisfying (16).

Indeed,

$$\frac{1}{Q^*(v, y)} E[Q^*(v - V, y + Y)] = \exp(\beta V) E\left[\cos \frac{\pi Y}{2b} - \tan \frac{\pi y}{2b} \sin \frac{\pi Y}{2b}\right]$$
$$= [1 + \beta V + O(b^{-4})][1 - \beta V + O(b^{-3})]$$
$$= 1 + O(b^{-3}) .$$

We will tinker with q and β to get rigorous inequalities instead of approximate equality. Here is a more careful statement. Continue to abbreviate

$$\beta = \frac{\pi^2}{8b^2}.$$

(18) THE GAMBLING PROOF OF (5). Fix $b \geqq 16$. Let $k = 1 - 36^{-\frac{1}{2}}$. In Section 4, we will produce a function $q(y)$ for $|y| \leqq b + 1$ such that

(a) $q(y) \geqq 1$
(b) $E\{q(y + Y)\} \leqq q(y) \exp\{-k\beta V\}$ for $|y| \leqq b$ and (V, Y) satisfying (16).
(c) $q(0) = (5\mathrm{b})^{\frac{1}{2}}$.

Then $Q(v, y) = q(y) \exp\{-k\beta v\}$ satisfies (14a)—(14b), and $Q(v, 0)$ is the bound in (5) by (c), so (14) proves (5). Condition (a) gives (14a) and (b) gives (14b), as you can easily check. □

(19) THE GAMBLING PROOF OF (6). Fix $b \geqq 16$. Let $k = 1 + 15\mathrm{b}^{-\frac{3}{4}}$. In Section 5, we will produce a function $q(y)$ for $|y| \leqq b + 1$ such that

(a) $0 \leqq q(y) \leqq 1$, and $q(y) = 0$ for $|y| > b$
(b) $E\{q(y + Y)\} \geqq q(y) \exp(-k\beta V)$ for $|y| \leqq b$ and (V, Y) satisfying (16)
(c) $q(0) = 1$.

Then $Q(v, y) = 1 - q(y) \exp[-k\beta(v + 1)]$ satisfies (14a)—(14b), and $Q(v, 0) = 1 - \theta(v)$ by (c); so (14) proves (6). Condition (a) gives (14a)—remember there is a $v + 1$ in the exponential; and (b) gives (14b), as you can easily check. □

3. The lemmas.

We collect here some estimates of the trigonometric and exponential functions, for use in Sections 4 and 5.

(20) LEMMA. *Suppose* $0 < \varepsilon < 1$.

(a) $\sin(\pi/2)\varepsilon \geqq \varepsilon$.
(b) $\cos(\pi/2)(1 - \varepsilon) \geqq \varepsilon$.
(c) $\tan(\pi/2)(1 - \varepsilon) \leqq 1/\varepsilon$.

(21) LEMMA. $\tan x \leqq x + x^3 \cos^{-4}\gamma$ *for* $0 \leqq x \leqq \gamma < \pi/2$.

PROOF. Make a finite Taylor expansion of $\tan x$. □

The next lemma is the key to (18b—19b).

(22) LEMMA. *Suppose* $0 < \varepsilon < 1$ *and* $|x| \leqq (1 - \varepsilon)\pi/2$. *Suppose* X *is a random variable, with* $|X| \leqq \alpha$ *and* $E(X) = 0$. *Then*

$$1 - (1 + \theta)\tfrac{1}{2}E(X^2) \leqq \frac{1}{\cos x} E[\cos(x + X)] \leqq 1 - (1 - \theta)\tfrac{1}{2}E(X^2),$$

where

$$\theta = \frac{\alpha^2}{3} + \frac{\alpha}{3\varepsilon}.$$

PROOF. By Taylor's expansion,

(23) $$\cos(x+X) = \cos x - X\sin x - \tfrac{1}{2}X^2\cos x + R\,;$$

where $R = \frac{1}{6}X^3\sin(x+X')$, and X' is a random variable with $|X'| \leqq |X| \leqq \alpha$. We claim

(24) $$|E(R)| \leqq E(|R|) \leqq \tfrac{1}{2}\theta E(X^2)\cos x\,.$$

Taking E of (23) and using (24) proves (22). But

$$\sin(x+X') = [\sin X' + (\tan x)\cos X']\cos x\,.$$

Now $|\sin X'| \leqq |X'| \leqq \alpha$, and $|\cos X'| \leqq 1$, and $|\tan x| \leqq 1/\varepsilon$ by (20c). So $|\sin(x+X')| \leqq B = (\alpha + \varepsilon^{-1})\cos x$. Furthermore, $|X^3| \leqq \alpha X^2$, so $|R| \leqq \frac{1}{6}\alpha B X^2$, proving (24). □

(25) LEMMA. *If* $0 < x < \frac{1}{2}$, *then* $(1-x) > \exp[-(1+x)x]$.

PROOF. Make an infinite Taylor expansion of $\log(1-x)$. □

4. The upper bound. We return to (18). Fix $b \geqq 16$. Let

(26) $$q(y) = A\cos[(1-\varepsilon)(\pi y/2b)] \quad \text{with} \quad A = (5b)^{\frac{1}{2}} \quad \text{and} \quad \varepsilon = (2b)^{-\frac{1}{2}}\,.$$

We claim this q satisfies (18).

We needed large A and positive ε to get (18a). For the check, confirm that $A(\varepsilon - b^{-1}) \geqq 1$, using the condition $b \geqq 16$. Furthermore, $(1-\varepsilon)(1+b^{-1}) < 1$. So q is minimized when $y = b+1$. Using (20b),

$$\begin{aligned} q(y) &\geqq q(b+1) \\ &= A\cos[(1-\varepsilon)(1+b^{-1})(\pi/2)] \\ &\geqq A[\varepsilon - b^{-1} + \varepsilon b^{-1}] \\ &\geqq 1\,. \end{aligned}$$

Condition (18c) is obvious: that is how we came to the multiplier $(5b)^{\frac{1}{2}}$ in (5).

Condition (18b) is harder to check. To make it hold, we need $\varepsilon > 0$ for (22), the difficulty being $\tan\pi/2 = \infty$.

The smaller k is, the less informative (5) is. So we wanted k to be reasonably large. It had to be less than 1, as (9) shows; but we wanted it close. Our k drops out of (22), and putting $\varepsilon = (2b)^{-\frac{1}{2}}$ approximately maximizes it.

Here is the systematic argument. We claim

(27) $E\{q(y+Y)\} \leqq q(y)[1 - k\beta V]$ for $|y| \leqq b$ and (V, Y) satisfying (16), where q is defined by (26) and $k = 1 - 3b^{-\frac{1}{2}}$ and $\beta = \pi^2/(8b^2)$.

This proves (18b), because $1 - u \leqq \exp(-u)$. We now argue (27), by using (22) with

$$x = (1-\varepsilon)\frac{\pi y}{2b} \qquad \text{and} \qquad X = (1-\varepsilon)\frac{\pi Y}{2b}\,.$$

If $|y| \leqq b$, then $|x| \leqq (1-\varepsilon)\pi/2$. If (V, Y) satisfies (16), then $\frac{1}{2}E(X^2) = (1-\varepsilon)^2\beta V$, and $|X| \leqq \alpha = \pi/(2b)$. So (22) proves (27) for any $k \leqq k_0$, where

$$k_0 = (1-\theta)(1-\varepsilon)^2 \quad \text{and} \quad \theta = \frac{\alpha^2}{3} + \frac{\alpha}{3\varepsilon}.$$

We now estimate k_0 fairly crudely: $\alpha^2/3 = \pi^2/(12b^2) < 1/b^2$, and $\alpha/3 = \pi/(6b) < 1/b$. Regrouping,

$$\theta \leqq b^{-2} + (\varepsilon b)^{-1}. \tag{28}$$

Now

$$\begin{aligned} k_0 &= (1-\theta)(1-\varepsilon)^2 \\ &\geqq 1 - \theta - 2\varepsilon \\ &\geqq 1 - b^{-2} - (\varepsilon b)^{-1} - 2\varepsilon. \end{aligned}$$

We chose ε so $2\varepsilon = (\varepsilon b)^{-1}$, approximately maximizing k_0. In any case, putting $\varepsilon = (2b)^{-\frac{1}{2}}$ gives

$$k_0 \geqq 1 - b^{-2} - 2 \cdot 2^{\frac{1}{2}} \cdot b^{-\frac{1}{2}}.$$

Now $b \geqq 16$, so

$$b^{-2} = b^{-\frac{3}{2}} \cdot b^{-\frac{1}{2}} \leqq 64^{-1}b^{-\frac{1}{2}}$$

and

$$\begin{aligned} k_0 &\geqq 1 - (64^{-1} + 2 \cdot 2^{\frac{1}{2}})b^{-\frac{1}{2}} \\ &\geqq 1 - 3b^{-\frac{1}{2}}. \end{aligned}$$

This proves (27), so (18b), so Theorem (5).

5. The lower bound. We return to (19). The most natural q was $q(y) = \max[\cos(\pi y/2b), 0]$. As before, this fails because $\tan \pi/2 = \infty$, spoiling (22). Rescaling by $1-\varepsilon$ spoils (19a). Our solution was to replace the last bit of the cosine function by a line segment. To avoid subscripts, fix $b \geqq 16$ and set

$$\varepsilon = \tfrac{1}{2}b^{-\frac{1}{2}}, \qquad \text{so } 0 < \varepsilon \leqq \tfrac{1}{4}. \tag{29}$$

This choice of ε roughly minimizes the k in (6). Draw the tangent line λ to the cosine at $(1-\varepsilon)\pi/2$. Extend λ downward and to the right until it meets the horizontal axis. Call the point of intersection $(1+\delta)\pi/2$. So δ depends on ε, and we will compute it later. Let $\varphi = \varphi_\varepsilon$ be the cosine on $[0, (1-\varepsilon)\pi/2]$, the tangent line λ on $[(1-\varepsilon)\pi/2, (1+\delta)\pi/2]$, and 0 on $[(1+\delta)\pi/2, \infty)$. Let $\varphi(-x) = \varphi(x)$. Now rescale so $(1+\delta)\pi/2$ becomes b: let

$$q(y) = \varphi\left[(1+\delta)\frac{\pi}{2}\frac{y}{b}\right]. \tag{30}$$

We will show that q satisfies (19). Conditions (19a) and (19c) are easy to check. Condition (19b) is harder. To begin with, we have to compute δ. The slope of the tangent can be computed both as a derivative and a ratio of lengths, giving the equation

$$\sin\left[(1-\varepsilon)\frac{\pi}{2}\right] = \frac{\cos[(1-\varepsilon)\pi/2]}{(\varepsilon+\delta)\pi/2},$$

or

$$\frac{\pi}{2}\delta = \tan\frac{\pi}{2}\varepsilon - \frac{\pi}{2}\varepsilon \,. \tag{31}$$

So $\delta \sim \varepsilon^3$. We need an upper bound. Remember $\varepsilon \leqq \frac{1}{4}$ from (29). Use estimate (21):

$$\delta \leqq \frac{\pi^2}{4}\left(\cos\frac{\pi}{8}\right)^{-4}\varepsilon^3 \,.$$

But $\cos^2 \pi/8 = \frac{1}{2}(1 + \cos \pi/4) = \frac{1}{2}(1 + 2^{-\frac{1}{2}})$, so $\cos^4 \pi/8 = \frac{1}{8}(3 + 2 \cdot 2^{\frac{1}{2}})$. Estimate $\pi^2 \leqq 10$, and eliminate the surd:

$$\delta \leqq 20(3 - 2 \cdot 2^{\frac{1}{2}})\varepsilon^3 \leqq \tfrac{1}{2}b^{-\frac{3}{4}} \leqq \tfrac{1}{16} \,, \tag{32}$$

where ε is obtained from (29).

We are now ready to tackle (19b). As the main step, we claim

(33) $E\{q(y + Y)\} \geqq q(y)[1 - h\beta V]$ for $|y| \leqq b$ and (V, Y) satisfying (16), where q is defined by (30) and $h = 1 + 14b^{-\frac{3}{4}}$ and $\beta = \pi^2/(8b^2)$.

By symmetry, it is enough to prove (33) when $0 \leqq v \leqq b$. Let $c = b(1 - \varepsilon)/(1 + \delta)$. There are three cases to consider: $0 \leqq y \leqq c$, where q is a rescaled cosine; $c < y < c + 1$, where $y + Y$ can fall in both intervals $[0, c]$ and $[c, b]$; finally, $c + 1 \leqq y \leqq b$, where q is linear. These cases are handled by (34), (38), and (45) respectively.

The first case of (33). We claim

(34) $E[q(y + Y)] \geqq q(y)[1 - k_1\beta V]$ for $0 \leqq y \leqq c$ and (V, Y) satisfying (16), where $k_1 = 1 + 3b^{-\frac{3}{4}}$ and $\beta = \pi^2/(8b^2)$.

In proving (34), it is convenient to abbreviate

$$g(y) = \cos\,(1 + \delta)\frac{\pi y}{2b} \,, \tag{35}$$

so by definition (30),

$$q(y) = g(y) \qquad \text{for } |y| \leqq c \,. \tag{36}$$

By convexity,

$$q(y) > g(y) \qquad \text{for } c < |y| \leqq c + 1 \,. \tag{37}$$

To pin (37) down, you have to check that the cosine did not get positive again, past $3\pi/2$. After rescaling, you need to know

$$c + 1 < b + 1 < \frac{3b}{(1 + \delta)} \,,$$

which follows from (32) and the assumption $b \geqq 16$. So

$$\begin{aligned} E\{q(y + Y)\} &\geqq E\{g(y + Y)\} && \text{by (36)—(37)} \\ &\geqq g(y)[1 - k_1\beta V] && \text{by (22)} \\ &= q(y)[1 - k_1\beta V] && \text{by (36)} \end{aligned}$$

all provided $k_1 \geqq k_0$, where

$$k_0 = (1 + \theta)(1 + \delta)^2$$
$$\theta = \frac{\alpha^2}{3} + \frac{\alpha}{3\varepsilon}$$
$$\alpha = (1 + \delta)\frac{\pi}{2b}\,.$$

Using (29) and (32) and $b \geqq 16$, confirm

$$\begin{aligned} \alpha &\leqq 2/b \\ \theta &\leqq (\tfrac{4}{3}b^{-\frac{5}{4}} + \tfrac{4}{3})b^{-\frac{3}{4}} \leqq \tfrac{3}{2}b^{-\frac{3}{4}} \\ k_0 &= 1 + (1 + \delta)^2\theta + (2 + \delta)\delta \\ &\leqq 1 + (\tfrac{17}{16})^2\theta + \tfrac{33}{16}\delta \\ &\leqq 1 + 3b^{-\frac{3}{4}}\,. \end{aligned}$$

This completes the proof of (34).

The second case of (33). We claim

(38) $E[q(y + Y)] \geqq q(y)[1 - k_2\beta V]$ for $c < y < c + 1$ and (V, Y) satisfying (16), where $k_2 = 1 + 14b^{-\frac{3}{4}}$ and $\beta = \pi^2/(8b^2)$.

This is the most delicate step. To begin with, the function q is convex on $[-b, b]$ and linear on $[c, b]$. Consequently, if you slide the graph down the tangent line at c, so $(c, q(c))$ moves to $(y, q(y))$, the new graph is below the old. Algebraically,

$$q(y + Y) \geqq q(y) + q(c + Y) - q(c)\,. \tag{39}$$

So (34) on $y = c$ shows

$$E[q(y + Y)] \geqq q(y) - k_1\beta Vq(c)\,. \tag{40}$$

But q is linear on $[c, b]$, with slope

$$-(1 + \delta)\frac{\pi}{2b}\sin(1 - \varepsilon)\frac{\pi}{2} > -\frac{2}{b}\,,$$

using (32). Since $c < y < c + 1$,

$$q(c) \leqq q(y) + 2b^{-1}\,. \tag{41}$$

Use estimate (20b) and (35)—(36):

$$q(c) = \cos(1 - \varepsilon)\frac{\pi}{2} \geqq \varepsilon\,. \tag{42}$$

Combine (41)—(42):

$$q(y) \geqq \varepsilon - 2b^{-1}\,. \tag{43}$$

Combine (41)—(43):

$$\frac{q(c)}{q(y)} \leqq 1 + \frac{2}{b(\varepsilon - (2/b)}$$

(44)
$$= 1 + \frac{2}{b\varepsilon - 2}$$
$$\leqq 1 + 8b^{-\frac{3}{4}}$$

by (29), using $b \geqq 16$. This and (40) prove (38) for

$$k_2 \geqq (1 + 8b^{-\frac{3}{4}})k_1$$
$$= (1 + 8b^{-\frac{3}{4}})(1 + 3b^{-\frac{3}{4}}) .$$

But $(1 + 8b^{-\frac{3}{4}})(1 + 3b^{-\frac{3}{4}}) \leqq 1 + 14b^{-\frac{3}{4}}$. This completes the proof of (38).

The third case of (33). We claim

(45) $E[q(y + Y)] \geqq q(y)$ for $c + 1 \leqq y \leqq b$ and (V, Y) satisfying (16).

To argue this, let l be the linear function which agrees with q on $[c, b]$. So $q \geqq l$ on $[c, b + 1]$, and

$$E[q(y + Y)] \geqq E[l(y + Y)]$$
$$= l(y)$$
$$= q(y) .$$

THE PROOF OF (33). Combine (34), (38) and (45) to prove (33).

THE PROOF OF (19b). Let $x = h\beta V$. Using the condition $b \geqq 16$ and $V \leqq 1$, and the estimate $\pi^2 \leqq 10$, you find

$$x \leqq 4b^{-2} \leqq \tfrac{1}{2} .$$

With the help of (25),

$$(1 - x) \geqq \exp[-(1 + x)x]$$
$$\geqq \exp[-(1 + 4b^{-2})h\beta V] .$$

But $(1 + 4b^{-2})h \leqq (1 + 15b^{-\frac{3}{4}})$, so (38) proves (19b) for q defined by (30). This completes the proof of (6).

6. A more general inequality. Suppose $Y_1, Y_2, \cdots$ are random variables which satisfy

(46)
$$E(Y_i \mid Y_1, \cdots, Y_{i-1}) = 0 .$$

Do not assume $|Y_i| \leqq 1$. Define V_i by (3) and τ_b, W_b by (4). In general, there is nothing to say about $P\{W_b > v\}$. It can be made arbitrarily close to 1 by taking the Y_i to be independent and identically distributed, 0 with high probability, and $\pm N$ with the remaining probability, where N is enormous. In the other direction, if $Y_1 = \pm (b + 1)$, then $P\{W_b > v\} = 0$ for $v > (b + 1)^2$. Suppose, however, that

(47)
$$P\{|Y_1 + \cdots + Y_{\tau_b}| \leqq b + 1\} = 1 .$$

This is clearly weaker than $|Y_i| \leqq 1$. We claim:

(48) PROPOSITION. *If* $Y_1, Y_2, \cdots$ *satisfy* (46)—(47), *where* τ_b *and* W_b *are defined by* (4), *then*

$$P\{W_b > v + (b+1)^2\} \leqq A \exp(-kv),$$

where $A = (b+1)^2/(2b+1)$ *and* $k = 1/(b+1)^2$.

PROOF. Consider this gambling situation. Your fortunes are the set of pairs (v, y), with $-\infty < v < \infty$ and $|y| \leqq b+1$. If $v < 0$ or $|y| > b$, you are stuck at (v, y). On the continuation region $v \geqq 0$ and $|y| \leqq b$, you can move to $(v - V, y + Y)$, where $E(Y) = 0$, $E(Y^2) = V$, and $|y + Y| \leqq b + 1$. Your utility function u is 1 on the win region $\{(v, y): v < 0 \text{ and } |y| \leqq b\}$, and 0 elsewhere. Let

$$Q(v, y) = A[1 - (b+1)^{-2}y^2] \exp(-kv).$$

Then $Q \geqq u$, and

$$E[Q(v - V, y + Y] \leqq Q(v, y)$$

on the continuation region, as you can easily check. If $W_b > v + (b+1)^2$, then $(V_1, Y_1), (V_2, Y_2), \cdots$ move you from the starting position $(v, 0)$ to the win region. This has chance at most $Q(v, 0)$, by (14). □

Here is an example to show (48) is of the right order of magnitude. Let $\varepsilon > 0$. If $Y_1 + \cdots + Y_n = 0$, let $Y_{n+1} = \pm\varepsilon$ with equal conditional probability $\frac{1}{2}$. The conditional mean is 0, the conditional variance is ε^2. If $Y_1 + \cdots + Y_n = \varepsilon$, let Y_{n+1} be $-\varepsilon$ or $b + 1 - \varepsilon$, with conditional probabilities $1 - (b+1)^{-1}\varepsilon$ and $(b+1)^{-1}\varepsilon$ respectively. The conditional mean is 0, the conditional variance is $\varepsilon(b + 1 - \varepsilon)$. If $Y_1 + \cdots + Y_n = -\varepsilon$, reverse the sign. Starting at 0, you have chance $1 - (b+1)^{-1}\varepsilon$ of returning to 0 in two moves, and chance $(b+1)^{-1}\varepsilon$ of absorption in two moves. Either way, you pick up

$$\varepsilon^2 + \varepsilon(b + 1 - \varepsilon) = \varepsilon(b + 1)$$

units of conditional variance. So W_b is distributed like $\varepsilon(b+1)N_\varepsilon$, where N_ε is the number of tosses it takes to get the first head with a coin that lands heads with probability $(b+1)^{-1}\varepsilon$. So

(49) $$P\{W_b > v\} \to \exp\{-(b+1)^{-2}v\} \qquad \text{as } \varepsilon \to 0.$$

7. Brownian motion and random walk. Our results here are very close to known ones, so we will report the facts and only sketch our argument. We got (1) this way. Let B be standard Brownian motion starting from y, and let $b \geqq |y|$. Let τ be the least t with $|B(t)| = b$. Wald's identity (Freedman (1971) Section 1.11) shows

(50) $$E\{\exp[\lambda B(\tau) - \tfrac{1}{2}\lambda^2\tau]\} = e^{\lambda y}.$$

Replace λ by $-\lambda$ and solve the two equations:

(51) $$E[\exp(-\tfrac{1}{2}\lambda^2\tau)] = \frac{\cosh \lambda y}{\cosh \lambda b},$$

where $\cosh u = \frac{1}{2}(e^u + e^{-u})$. This holds for real λ. Some complex variable arguing allows you to put $\lambda = i(2z)^{\frac{1}{2}}$, getting

(52) $$E(\exp(-z\tau)) = \cos y(2z)^{\frac{1}{2}} / \cos b(2z)^{\frac{1}{2}} ,$$

for nonnegative real z less than the first zero of $\cos b(2z)^{\frac{1}{2}}$, namely $0 \leqq z < \pi^2/(8b^2)$. Expanding the right side and using a Tauberian theorem (Feller (1966) Section 13.5) gets (1). The square root looks nonanalytic, but cosine has only even powers. Incidentally, if you make a partialfraction expansion of the right side of (51), you get Fürth's formula (Feller (1968) page 359).

The situation for random walks is very similar. Let $Y_1, Y_2, \cdots$ be independent and identically distributed, with $P(Y_i = \pm 1) = p \leqq \frac{1}{2}$ and $P(Y_i = 0) = 1 - 2p$. So (2) holds. Now edge effects are important. Let $\hat{\tau}_b$ be the least n with $|Y_1 + \cdots + Y_n| = b$. Let $\hat{W}_b = \operatorname{Var} Y_1 \cdot \hat{\tau}_b$. So $\hat{W}_b \leqq W_b$.

Our first result shows that the ordinary coin tossing game ($p = \frac{1}{2}$) *hits* $\pm b$ an order of magnitude faster than Brownian motion. To get a process satisfying (2) which *leaves* $\pm b$ an order of magnitude faster than Brownian motion, stop the coin tossing game when it first hits $\pm b$. Then, move ε away from 0 with probability $1/(1 + \varepsilon)$; move 1 toward 0 with probability $\varepsilon/(1 + \varepsilon)$.

(53) If $p = \frac{1}{2}$, then $\lim_{v\to\infty} (1/v) P\{\hat{W}_b > v\} = \log\cos \pi/2b < -\pi^2/8b^2$.

Of course

$$\log\cos\frac{\pi}{2b} = -\frac{\pi^2}{8b^2} - O(b^{-4}) ,$$

so the constant is even closer than (5) and (6) predict.

Our second result shows that for small p, the walk uses an order of magnitude more variance than Brownian motion to hit $\pm b$; even more is needed to leave.

(54) $\lim_{p\to 0} P\{\hat{W}_b > v\} = H_b(v)$ exists, and $\lim_{v\to\infty} 1/v \log H_b(v) = -2\sin^2 \pi^2/8b^2$.

Incidentally, you can identify $H_b(v)$ as the time it takes a Poisson process with jumps ± 1, of equal probability $\frac{1}{2}$, and unit rate of jumping, to hit $\pm b$. We give an explicit formula in (63).

Let $|j| < b$, and let τ be the least n with $|S_n| = b$, where $S_n = j + Y_1 + \cdots + Y_n$. We will obtain a variant of the Lagrange formula for $P\{\tau = n\}$. The idea is to make a partial fraction expansion of the Laplace transform, computed by Wald's identity. Namely, let

(55) $$\phi(\lambda) = E[\exp(\lambda Y_1)] = 1 + 2p(\cosh\lambda - 1) ,$$

so $\phi(\lambda) \geqq 1$ for real λ. Let $\psi(\lambda) = \log\phi(\lambda)$. Then

$$E\{\exp(\lambda S_\tau - \psi(\lambda)\tau)\} = \exp j\lambda ,$$

so

(56) $$E\{\exp(-\psi(\lambda)\tau)\} = \frac{\cosh j\lambda}{\cosh b\lambda} .$$

The right side of (56), call it $F(\lambda)$, is analytic in the whole plane, except for simple poles at the zeros of $\cosh b\lambda$, namely, at

$$\lambda_m = (2m+1)\frac{\pi}{2b}i\,. \tag{57}$$

Because of the periodicity, only the roots with $m = 0, \cdots, b-1$ are interesting. We now make a partial fraction expansion of F, but in terms of ϕ. That is, we ask for a_m such that

$$F(\lambda) = \sum_{m=0}^{b-1} \frac{a_m}{\phi(\lambda) - \phi(\lambda_m)}\,. \tag{58}$$

To compute a_m, multiply across by $\phi(\lambda) - \phi(\lambda_m)$ and let $\lambda \to \lambda_m$. So

$$a_m = \cosh(j\lambda_m) \lim_{\lambda\to\lambda_m} \frac{\phi(\lambda) - \phi(\lambda_m)}{\cosh b\lambda}\,.$$

Remember that $\cosh iz = \cos z$, and use L'Hospital:

$$a_m = (-1)^m \frac{2p}{b} \cos\left(j(2m+1)\frac{\pi}{2b}\right) \sin\left((2m+1)\frac{\pi}{2b}\right). \tag{59}$$

But why does (58) hold? Well, $\cosh k\lambda$ is a polynomial of degree $|k|$ in $\cosh \lambda$, so in $\phi(\lambda)$. Consequently, (58) is right modulo a polynomial. Both sides of (58) tend to 0 as $|\lambda| \to \infty$, so the polynomial is zero.

Now expand (58) as a power series in $1/\phi(\lambda)$:

$$F(\lambda) = \sum_{m=0}^{b-1} \frac{a_m}{\phi(\lambda)} \sum_{n=0}^{\infty} \left[\frac{\phi(\lambda_m)}{\phi(\lambda)}\right]^n. \tag{60}$$

But $\exp(-\psi(\lambda)) = 1/\phi(\lambda)$, so the left side of (56) is $E\{\phi(\lambda)^{-\tau}\}$. Comparing this with (60), we get

$$P\{\tau = n\} = \sum_{m=0}^{b-1} a_m r_m{}^{n-1} \qquad \text{for } n = 1, 2, \cdots \tag{61}$$

where the a_m are defined by (59) and

$$r_m = \phi(\lambda_m) = 1 - 2p\left\{1 - \cos\left[(2m+1)\frac{\pi}{2b}\right]\right\}.$$

When $p = \frac{1}{2}$, relation (61) is the Lagrange formula. The leading terms correspond to $m = 0$ and $b - 1$, and the signs depend on the relative parity of n, j, and b. If you bother sorting it out, you get (53).

When $p < \frac{1}{2}$, it is more convenient to write

$$r_m = 1 - 4p \sin^2\left[(m + \tfrac{1}{2})\frac{\pi}{2b}\right].$$

Summing,

$$P\{\tau > n\} = \sum_{m=0}^{b-1} a_m (1 - r_m)^{-1} r_m{}^n\,. \tag{62}$$

So

$$\begin{aligned} H_b(j, v) &= \lim_{p\to 0} P\{2p\tau > v\} \\ &= \sum_{m=0}^{b-1} \alpha_m \rho_m{}^v\,, \end{aligned} \tag{63}$$

where

$$\alpha_m = a_m(1 - r_m)^{-1} = (-1)^m \frac{1}{b} \cos\left[j(2m + 1)\frac{\pi}{2b}\right] \cot\left[(m + \tfrac{1}{2})\frac{\pi}{2b}\right]$$

$$\rho_m = \lim_{p\to 0} r_m^{1/2p} = \exp\left[-2\sin^2(m + \tfrac{1}{2})\frac{\pi}{2b}\right].$$

This proves (54).

REFERENCES

DUBINS, L. E. (1972). Sharp bounds for the total variance of uniformly bounded semimartingales. *Ann. Math. Statist.* **43** 1559–1565.

DUBINS, L. E. and FREEDMAN, D. A. (1965). A sharper form of the Borel-Cantelli lemma and the strong law. *Ann. Math. Statist.* **36** 800–807.

DUBINS, L. E. and SAVAGE, L. J. (1965). *How to Gamble if You Must.* McGraw-Hill, New York.

DUBINS, L. E. and SCHWARTZ G. (1965). On continuous martingales. *Proc. Nat. Acad. Sci. U.S.A.* **53** 913–916.

FELLER, W. (1966). *An Introduction to Probability Theory and its Applications* **2**. Wiley, New York.

FELLER, W. (1968). *An Introduction to Probability Theory and its Applications.* **1** 3rd. ed. Wiley, New York.

FREEDMAN, D. A. (1971). *Brownian Motion and Diffusion.* Holden-Day, San Francisco.

LÉVY, P. (1954). *Théorie de l'addition des variables aléatoires.* Gauthier-Villars, Paris.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720