The Cereteli-Davis Solution to the H^1-embedding Problem and an Optimal Embedding in Brownian Motion

by

Edwin Perkins

Department of Mathematics

University of British Columbia

Vancouver, B.C.

V6T 1Y4

Canada

Summary

Necessary and sufficient conditions are found on a mean-zero probability, μ, for the existence of a stopping time, T, and a Brownian motion, B, such that B_T has law μ and B_T^* is integrable. This result, due to Burgess Davis (the classical analogue was first solved by O. D. Cereteli), leads naturally to a stopping time, T, that stochastically minimizes both $\sup_{s\leq T} B_s$ and $-\inf_{s\leq T} B_s$.

1. Introduction.

Consider a mean-zero probability on the line, μ, and a one-dimensional $\{F_t\}$-Brownian motion, B_t defined on some (Ω, F, F_t, P) and satisfying $B_0 = 0$. An $\{F_t\}$-stopping time, T, is an embedding of μ if $B(T \wedge t)$ is a uniformly integrable martingale such that $L(B_T) = \mu$ ($L(Z)$ denotes the law of the r.v. Z). T is an H^p-embedding if, in addition, $B_T^* \equiv \sup\{|B_s| : s \leq T\}$ is in L^p. The existence of an embedding is due to Skorokhod (1965). A particularly explicit one is described in Azéma-Yor (1978 a,b). If $p > 1$, Doob's strong L^p inequalties show that an H^p-embedding of μ exists iff $\int |x|^p d\mu(x) < \infty$ iff every embedding of μ is an H^p-embedding, and, if $p < 1$, Doob's weak L^1 inequality shows that every embedding is an H^p-embedding. The situation for $p = 1$ is more delicate. It is well-known that there are laws μ for which some embeddings are H^1-embeddings and some are not (eg. compare Proposition 2.1 below with Theorem 2.2 of Azéma-Yor (1978 b)). A natural question, which I learned from John Walsh, is therefore:

(1.1) Problem. Find necessary and sufficient conditions on μ for it to have an H^1-embedding.

Doob's LlogL inequality shows that $\int |x| \log^+|x| d\mu(x) < \infty$ ($\log^+ x = \max(\log x, 0)$) is sufficient for every embedding to be an H^1-embedding and by using optional stopping it is easy to see that it is also necessary if $\mathrm{supp}(\mu)$ ($\mathrm{supp}(\mu) \equiv$ support of μ) is bounded below or above. It is simple enough to obtain better sufficient conditions by doing some computations with one's favourite embedding, providing of course it is an embedding that allows one to compute such things as $P(\sup_{s \leq T} B_s \geq \lambda)$. This is done in the beginning of section 2, where we adopt the Skorohod embedding as our favorite embedding (it will be until section 3) and arrive at a sufficient condition due to Walsh (2.5).

Getting necessary conditions seems harder as there are a lot of embeddings to check. In fact, a complete solution to (1.1) already exists in the literature. As this seems to have escaped the notice of many probabilists, and as the history of the subject is complicated by the close connection between classical H^p-theory and probability, shown by Burkholder, Gundy and Silverstein (1971), a short historical account is in order.

Let ∂D be the unit circle in the complex plane. If f is an integrable, real-valued function on ∂D, let $\tilde{f}$ denote the conjugate function of f, and write $f \in \mathrm{Re}\, H^1(\partial D)$ if $\tilde{f}$ is integrable on ∂D. The space $\mathrm{Re}\, H^1(\partial D)$ can be considered as a subspace of the H^1-embedding stopping times and furthermore the spaces $\mathrm{Re}\, H^1(R^n)$, $n \geq 1$, are less closely, but still strongly, connected to these times (see Davis (1980) for the necessary definitions and motivation). The analogue of (1.1) for $\mathrm{Re}\, H^1(\partial D)$ was first answered completely by O. D. Cereteli in a series of papers (see Cereteli (1976)). He gives a condition on the distribution of an integrable function, f, on ∂D, that is necessary and sufficient for the existence of a rearrangement (of f) that belongs to $\mathrm{Re}\, H^1(\partial D)$. We learned of Ceretel̆i's work from Burgess Davis, who used probability to give a different (but of course equivalent) necessary and sufficient condition in Davis (1980). (He was unaware of Ceretel̆i's work at that time.) From a classical viewpoint the contribution of this part of Davis' paper is that the natural extension of his condition to functions on R^n was shown to characterize the distributions of functions in $\mathrm{Re}\, H^1(R^n)$, whereas Ceretelli's condition did not extend to this setting.

For the probabilistic question (1.1) we are studying in this paper, the answer follows immediately from the probabilistic arguments in Davis (1980), as is pointed out on p.218 of that work, and the characterization is the same as in the classical unit circle setting. However, Davis tells me that, "Any probabilist knowing Ceretelli's work, as well as the results of Burkholder, Gundy and Silverstein, would have been able to answer (1.1)."

In section 2 we follow Davis' proof of necessity and show his condition is equivalent to Ceretelli's condition and Walsh's sufficient condition. The main result of this section is stated as Theorem 2.7. This approach to the Ceretelli-Davis theorem has the advantage of showing that if an H^1-embedding of μ exists, then the Skorokhod embedding will be such an embedding, a result we found a little surprising.

Davis' proof of necessity leads naturally to the definition of an explicit extremal embedding of μ, much in the spirit of Azéma-Yor (1978a), that stochastically minimizes both $\sup_{s \leq T} B(s)$ and $-\inf_{s \leq T} B(s)$ over all embeddings T (Theorems 3.7, 3.8). This construction is carried out in section 3.

Throughout this work X denotes a random variable with the fixed law, μ, and if Y_t is a real-valued process,

$$T_Y(\lambda) = \begin{cases} \inf\{t > 0 : Y_t \geq \lambda\} & \text{if } \lambda \geq 0 \\ \inf\{t > 0 : Y_t \leq \lambda\} & \text{if } \lambda < 0 \end{cases} \qquad (\inf \emptyset = \infty).$$

We write $\lambda_n \uparrow\uparrow \lambda$ to denote that $\{\lambda_n\}$ is strictly increasing to λ.

2. The Ceretelí–Davis Solution to the H^1-embedding Problem

We start by obtaining sufficient conditions for the existence of an H^1-embedding. Assume first that μ is atomless so that there is an explicit description of the Skorokhod embedding.

Definition. If $\lambda \geq 0$, let

$$-\rho(\lambda) = \inf\{y: \int I(x \leq y \text{ or } x \geq \lambda)x d\mu(x) \leq 0\}.$$

It is easy to see that $\rho(\lambda) = +\infty$ iff $\mu[\lambda,\infty) = 0$ and ρ: $[0,\infty) \to [0,\infty]$ is non-decreasing and right-continuous. One can also show that

$$\mu[-\lambda,0] = \int I(0 \leq x,\ \rho(x) \leq \lambda)x\rho(x)^{-1}\,d\mu(x). \tag{2.1}$$

If μ has a smooth, strictly positive density this is an easy calculus argument and the technical problems one faces in general are uninteresting and easily overcome. Let $R \geq 0$ be independent of B and have distribution function

$$P(R \leq x) = \int I(y \leq x)(1 + y\rho(y)^{-1})d\mu(y).$$

The right side defines a probability by (2.1). If

$$T_s = \inf\{t: B_t \notin (-\rho(R),R)\}, \tag{2.2}$$

then T_s is an embedding of μ. This is essentially the embedding studied in Skorokhod (1965). Although $\int x^2 d\mu(x) < \infty$ is assumed there, it is an easy exercise to check that $B(t \wedge T_s)$ is uniformly integrable without this condition.

Notation. $M_t = \sup\limits_{s\leq t} B_s$, $m_t = -\inf\limits_{s\leq t} B_s$, $H(\mu) = \int_0^\infty \lambda^{-1}\left|\int_{-\infty}^\infty x\, I(|x| \geq \lambda)d\mu(x)\right|d\lambda$.

<u>Proposition 2.1.</u> <u>Assume</u> μ <u>is a mean-zero, atomless probability on the line and</u> T_s <u>is given by (2.2).</u>

(a) $$E(M(T_s)) = \int_0^\infty (x + \rho(x)) \log(1 + \frac{x}{\rho(x)}) d\mu(x)$$

$$E(m(T_s)) = \int_0^\infty (x + \rho(x)) x\, \rho(x)^{-1} \log(1 + \frac{\rho(x)}{x}) d\mu(x)$$

(b) $$E(M(T_s) + m(T_s)) \le 2\Big(\int_{-\infty}^\infty |x| d\mu(x) + \int_0^\infty x \Big|\log \frac{x}{\rho(x)}\Big| d\mu(x)\Big)$$

$$= 2\Big(\int_0^\infty |x| d\mu(x) + H(\mu)\Big).$$

<u>Proof</u>. (a) $$P(M(T_s) \ge \lambda) = \int_0^\infty P(M(T_s) \ge \lambda | R = x) \ (1 + \frac{x}{\rho(x)}) d\mu(x)$$

$$= \int_\lambda^\infty \frac{\rho(x)}{\lambda + \rho(x)} (1 + \frac{x}{\rho(x)}) d\mu(x)$$

$$\therefore \quad E(M(T_s)) = \int_0^\infty \int_\lambda^\infty \frac{x + \rho(x)}{\lambda + \rho(x)} d\mu(x) d\lambda$$

$$= \int_0^\infty (x + \rho(x)) \log(1 + \frac{x}{\rho(x)}) d\mu(x).$$

A similar argument gives the required expression for $E(m(T_s))$.

(b) Use the inequalities $\log(1 + y) \le y$ and $\log(1 + y) \le 1 + |\log y|$ for all $y \ge 0$, to see that

$$\int_0^\infty (x + \rho(x)) \log(1 + \frac{x}{\rho(x)}) d\mu(x) + \int_0^\infty (x + \rho(x)) \frac{x}{\rho(x)} \log(1 + \frac{\rho(x)}{x}) d\mu(x)$$

$$\le \int_0^\infty x(1 + |\log \frac{x}{\rho(x)}|) + x \, d\mu(x) + \int_0^\infty x + x(1 + |\log(\frac{\rho(x)}{x})|) d\mu(x)$$

$$= 2\Big(\int_{-\infty}^\infty |x| d\mu(x) + \int_0^\infty x \Big|\log \frac{x}{\rho(x)}\Big| d\mu(x)\Big),$$

and hence obtain the first inequality in (b).

Let $\lambda_n \uparrow\uparrow \lambda$ and take limits in

$$\int x\, I(x \le -\rho(\lambda_n) \text{ or } x \ge \lambda_n)\, d\mu(x) = 0$$

to see that

$$\int x\, I(x \le -\rho(\lambda-) \text{ or } x \ge \lambda)\, d\mu(x) = 0.$$

As the same equation holds with $\rho(\lambda)$ in place of $\rho(\lambda-)$, it must be that

(2.3) $$\mu[-\rho(\lambda), -\rho(\lambda-)] = 0 \quad \text{for each } \lambda \ge 0.$$

In particular, if $\rho^{-1}(x)$ denotes the right-continuous inverse of ρ, then $-\lambda \le -\rho(\rho^{-1}(\lambda)-)$ and so $\mu[-\rho(\rho^{-1}(\lambda)), -\lambda] = 0$. It follows that

$$\int x\, I(x \le -\lambda \text{ or } x \ge \rho^{-1}(\lambda))\, d\mu(x) = \int x\, I(x \le -\rho(\rho^{-1}(\lambda)) \text{ or } x \ge \rho^{-1}(\lambda))\, d\mu(x)$$
$$= 0,$$

and therefore

(2.4) $$\left|\int xI(x \le -\lambda \text{ or } x \ge \lambda)\, d\mu(x)\right| = \int xI(\lambda \wedge \rho^{-1}(\lambda) \le x < \lambda \vee \rho^{-1}(\lambda))\, d\mu(x).$$

An argument similar to that used to show (2.3) gives us $\mu[r,s] = 0$ whenever $r < s$ satisfy $\rho(r) = \rho(s)$. This implies that μ does not charge the "flat spots" of ρ and hence that $\rho^{-1}(\rho(x)) = x$ for μ-a.a. $x \ge 0$.

This gives us

$$\rho^{-1}(\lambda) \le x \iff \lambda \le \rho(x) \quad \text{for } \mu\text{-a.a. } x \ge 0$$

and hence, by (2.4),

$$\left|\int x\, I(x \le -\lambda \text{ or } x \ge \lambda)\, d\mu(x)\right| = \int_0^\infty x\, I(x \wedge \rho(x) < \lambda \le \rho(x) \vee x)\, d\mu(x).$$

It follows that

$$H(\mu) = \int_0^\infty \int_{x\wedge\rho(x)}^{x\vee\rho(x)} \lambda^{-1} d\lambda\, x\, d\mu(x) = \int_0^\infty x \left|\log \frac{x}{\rho(x)}\right| d\mu(x)$$

and the proof is complete. □

As an immediate corollary we see that, when μ is atomless, either of the two equivalent conditions

(2.5) $$\int_0^\infty x\left|\log \frac{x}{\rho(x)}\right| d\mu(x) < \infty$$

or

(2.6) $$H(\mu) < \infty$$

is sufficient for the existence of an H^1-embedding, namely T_s. These conditions are symmetry conditions on the tails of μ. Both conditions hold if μ is symmetric (the integrals are zero) and are equivalent to $\int|x|\log^+|x|\,d\mu(x) < \infty$ if supp(μ) is bounded above or below. (2.6) appears in Cereteli (1976), and also in Vallois (1982). (2.5) was shown by Walsh (private communication) to be necessary and sufficient for $E(B^*(T_s)) < \infty$ and led him to make the

(2.7) Conjecture. (2.5) is necessary and sufficient for the existence of an H^1-embedding of the atomless measure μ.

The necessity of (2.7) is not at all obvious, since T_s is in no way an "optimal embedding".

Our immediate task is to extend these results to the case when μ may have atoms. It will be easier to work with (2.6) than (2.5). Let α_n denote the uniform law on $[-\frac{1}{n},\frac{1}{n}]$, $\mu_n = \alpha_n * \mu$ (* denotes convolution) and $T_{s,n}$ denote the Skorokhod embedding of μ_n. Then there are random variables $(U_n,V_n) \in (-\infty,0] \times [0,\infty)$, independent of B such that

$$T_{s,n} = \inf\{t \geq 0: B_t \notin (U_n,V_n)\}.$$

By changing the underlying probability space we may assume there is a subsequence such that $(U_{n_k},V_{n_k}) \overset{a.s.}{\to} (U,V) \in [-\infty,0] \times [0,\infty]$, where $\{U_{n_k},V_{n_k} : k \in \mathbb{N}\}$ is independent of the Brownian motion B (Skorokhod (1965, Ch.1.6)). It follows that

$$T_{s,n_k} \overset{a.s.}{\to} T_s \equiv \inf\{t \geq 0: B_t \notin (U,V)\} \leq \infty.$$

Note also that

$$\begin{aligned} P(T_{s,n} \geq K) &\leq P(T_{s,n} \geq K,\ B^*_{T_{s,n}} \leq M) + P(B^*_{T_{s,n}} \geq M) \\ &\leq P(B^*_K \leq M) + M^{-1}E(|B_{T_{s,n}}|) \\ &\leq P(B^*_1 \leq M\,K^{-1/2}) + M^{-1}(\int |x|d\mu(x) + n^{-1}). \end{aligned}$$

The last line may be made arbitrarily small, uniformly in n, by first choosing M and then K large enough. Therefore $T_s < \infty$ a.s. and so $B(T_{s,n_k}) \overset{a.s.}{\to} B(T_s)$. This implies $\mathcal{L}(B(T_s)) = \mu$. Moreover we have

$$\lim_{k\to\infty} E|B(T_{s,n_k})| = \lim_{k\to\infty} \int |x| d\mu_{n_k}(x) = E(|B(T_s)|).$$

This clearly shows $B(T_s \wedge t)$ is uniformly integrable and hence T_s is an embedding of μ, which we call the Skorokhod embedding of μ (although, given this nebulous procedure, "the" may be rather strong language).

Theorem 2.2. *Let μ be a mean-zero probability on the line and let T_s be defined as above. Then*

$$E(M_{T_s} + m_{T_s}) \leq 2(\int |x| d\mu(x) + H(\mu)),$$

and in particular $H(\mu) < \infty$ implies T_s is an H^1-embedding of μ.

Proof. Recall X denotes a r.v. with law μ. Let U_n be independent of X and have law α_n, and let $X_n = X + U_n$. Then

$$\left| \int_1^\infty \lambda^{-1}(|E(X_n I(|X_n| \geq \lambda))| - |E(X\ I(|X| \geq \lambda))|)d\lambda \right|$$

$$\leq \int_1^\infty \lambda^{-1}|E(X_n I(|X_n| \geq \lambda)) - E(X\ I(|X| \geq \lambda))|d\lambda$$

$$\leq \int_1^\infty \lambda^{-1}(n^{-1}P(|X_n| \geq \lambda) + |E(X(I(|X| \geq \lambda) - I(|X_n| \geq \lambda)))|)d\lambda$$

$$\leq n^{-1}E(\log^+|X_n|) + \int_1^\infty \lambda^{-1}E(|X|I(|X-\lambda| \leq n^{-1} \text{ or } |X+\lambda| \leq n^{-1}))d\lambda$$

$$\leq n^{-1}(E(\log^+|X_n|) + 4\ E|X|) \to 0 \quad \text{as } n \to \infty.$$

Note also that

$$\lim_{n\to\infty} \int_0^1 \lambda^{-1}|E(X_n I(|X_n| \geq \lambda))|d\lambda = \int_0^1 \lambda^{-1}|E(X\ I(|X| \geq \lambda))|d\lambda$$

because the integrands converge for Lebesgue-a.a. λ and are bounded by one (recall $E(X_n) = 0$). We have shown that

$$\lim_{n\to\infty} H(\mu_n) = H(\mu). \tag{2.8}$$

Let $\{n_k\}$ be the subsequence used to construct T_s. Then

$$E(M(T_s) + m(T_s)) \leq \liminf_{k\to\infty} E(M(T_{s,n_k}) + m(T_{s,n_k})) \text{ (Fatou's lemma)}$$

$$\leq \liminf_{k\to\infty} 2(\int_{-\infty}^{\infty} |x|d\mu_{n_k}(x) + H(\mu_{n_k}))$$

(Proposition 2.1 (b))

$$= 2(\int_{-\infty}^{\infty} |x|d\mu(x) + H(\mu)) \qquad \text{(by (2.8))}.\square$$

To find necessary conditions for the existence of an H^1-embedding introduce the

<u>Notation</u>. $-\alpha = -\alpha^\mu = \inf \operatorname{supp}(\mu)$, $\beta = \beta^\mu = \sup \operatorname{supp}(\mu)$

$$-\gamma_+(\lambda) = -\gamma_+^\mu(\lambda) = \begin{cases} \sup\{y: E(X|X \leq y \text{ or } X \geq \lambda) \geq \lambda\} & \text{if } P(X \geq \lambda) > 0 \\ -\alpha & \text{otherwise} \end{cases}$$

$$= \sup\{y: \int (x-\lambda)I(x \leq y \text{ or } x \geq \lambda)d\mu(x) \geq 0\} \text{ (sup } \emptyset = -\alpha, \lambda \geq 0)$$

$$\gamma_-(\lambda) = \gamma_-^{\mu}(\lambda) = \begin{cases} \inf\{y: E(X|X \leq -\lambda \text{ or } X \geq y) \leq -\lambda\} & \text{if } P(X \leq -\lambda) > 0 \\ \beta & \text{otherwise} \end{cases}$$

$$= \inf\{y: \int(x+\lambda)I(x \leq -\lambda \text{ or } x \geq y) \leq 0\} \quad (\inf \emptyset = \beta,\ \lambda \geq 0)$$

$$\psi(y) = \begin{cases} \int x\, I(x \geq y)d\mu(x)/\mu[y,\infty) & \text{if } \mu[y,\infty) > 0 \\ y & \text{if } \mu[y,\infty) = 0 \end{cases}$$

$$\phi(\lambda) = \inf\{y: \psi(y) \geq \lambda\} \qquad (\lambda \geq 0)$$

$$\tilde{\mu}(-\infty,x] = \mu[-x,\infty).$$

Hence ϕ is the increasing left-continuous inverse of the increasing, left-continuous barycentre function ψ (see Azéma-Yor (1978a)). $\gamma_{\pm}$ are increasing left-continuous functions from $[0,\infty)$ to $[0,\infty]$. These and other properties of $\gamma_{\pm}$ will be discussed in the next section (Lemma 3.2). For now, we will need the following results, which follow easily from the definitions:

(2.9) $$\int(x-\lambda)I(x \geq \phi(\lambda))\, d\mu(x) \leq 0 \leq \int(x-\lambda)I(x > \phi(\lambda))d\mu(x)$$

(2.10) $$\int(x-\lambda)I(x \leq -\gamma_+(\lambda) \text{ or } x \geq \lambda)d\mu(x)$$
$$\leq 0 \leq \int(x-\lambda)I(x < -\gamma_+(\lambda) \text{ or } x \geq \lambda)d\mu(x)$$

(2.11) $$\int(x+\lambda)I(x \leq -\lambda \text{ or } x > \gamma_-(\lambda))d\mu(x)$$
$$\leq 0 \leq \int(x+\lambda)I(x \leq -\lambda \text{ or } x \geq \gamma_-(\lambda))d\mu(x).$$

(2.12) $$\gamma_-^{\mu} = \gamma_+^{\tilde{\mu}}$$

<u>Notation</u>. If $\lambda > 0$, let

$$p(\lambda) = \begin{cases} \int (x-\lambda)I(x > \phi(\lambda))d\mu(x)\ (\lambda - \phi(\lambda))^{-1} & \text{if } \phi(\lambda) < \lambda \\ \mu(\{\phi(\lambda)\}) & \text{if } \phi(\lambda) = \lambda \end{cases}$$

$$q_+(\lambda) = \int (x-\lambda)I(x < -\gamma_+(\lambda) \text{ or } x \geq \lambda)d\mu(x)(\gamma_+(\lambda)+\lambda)^{-1}$$

$$q_-(\lambda) = \int (x+\lambda)I(x \leq -\lambda \text{ or } x > \gamma_-(\lambda))d\mu(x)\ (\gamma_-(\lambda)+\lambda)^{-1}$$

$$\mu^*(\lambda) = \mu(\phi(\lambda),\infty) + p(\lambda)$$

$$\mu_+(\lambda) = \mu(-\infty, -\gamma_+(\lambda)) + q_+(\lambda) + \mu[\lambda,\infty)$$

$$\mu_-(\lambda) = \mu(\gamma_-(\lambda), \infty) + q_-(\lambda) + \mu(-\infty,-\lambda].$$

<u>Lemma 2.3</u>. (2.13) $0 \leq p(\lambda) \leq \mu(\{\phi(\lambda)\})$

(2.14) $0 \leq q_+(\lambda) \leq \mu(\{-\gamma_+(\lambda)\})$

(2.15) $0 \leq q_-(\lambda) \leq \mu(\{\gamma_-(\lambda)\})$

(2.16) $q_-^{\mu}(\lambda) = q_+^{\tilde{\mu}}(\lambda), \quad \mu_- = \tilde{\mu}_+$

<u>Proof</u>. (2.9) implies that

$$0 \leq \int (x - \lambda)I(x > \phi(\lambda))d\mu(x) \leq (\lambda - \phi(\lambda))\mu(\{\phi(\lambda)\}).$$

Divide the above by $\lambda - \phi(\lambda)$ to obtain (2.13). Similarly one can use (2.10) and (2.11) to prove (2.14) and (2.15). (2.16) is an easy consequence of (2.12).□

The key idea in the derivation of necessary conditions for the existence of an H^1-embedding is

<u>Lemma 2.4 (Davis' Law of the Lever)</u>. <u>Assume</u> $L(X) = \mu$, $\lambda > 0$ <u>and A is a measurable set such that</u> $\{X \geq \lambda\} \subset A$.

(a) (i) <u>If</u> $\int_A X - \lambda \, dP \geq 0$, <u>then</u>

$$P(A) \leq \mu^*(\lambda). \tag{2.17}$$

(ii) <u>If, in addition, equality holds in (2.17), then</u>

$$\{X > \phi(\lambda)\} \subset A \subset \{X \geq \phi(\lambda)\} \text{ a.s.} \tag{2.18}$$

(iii) <u>Conversely if (2.18) holds and</u> $\int_A X - \lambda \, dP = 0$ <u>then</u> $P(A) = \mu^*(\lambda)$.

(b) (i) <u>If</u> $\int_A X - \lambda \, dP \leq 0$, <u>then</u>

$$P(A) \geq \mu_+(\lambda). \tag{2.19}$$

(ii) _If, in addition, equality holds in (2.19), then_

(2.20) $\{X < -\gamma_+(\lambda) \text{ or } X \geq \lambda\} \subset A \subset \{X \leq -\gamma_+(\lambda) \text{ or } X \geq \lambda\}$ a.s.

(iii) _Conversely if (2.20) holds and_ $\int_A X - \lambda dP = 0$, _then_ $P(A) = \mu_+(\lambda)$.

Remark. (2.17) was observed in Blackwell-Dubins (1963). The idea of (b) appears in Davis (1980, p.215, 1982, p.157).

These results are intuitively obvious. Sand is distributed along a see-saw according to μ. The fulcrum is at λ and sand is initially added to the right of λ. (a) says that if we want to add the maximum amount of sand without tipping the see-saw to the left, we should add it as close as possible to the fulcrum. (b) says that if we want to add the minimum amount of sand needed to tip the see-saw to the left or at least put it in equilibrium, we should add it as far from the fulcrum as possible. Although a proof is clearly not needed, we include a derivation of (b) because of its importance in what follows.

Proof of (b). The definition of $q_+(\lambda)$ gives us

$$\int_A X - \lambda \, dP \leq 0 = \int I(X < -\gamma_+(\lambda) \text{ or } X \geq \lambda) \, (X - \lambda)dP - (\gamma_+(\lambda) + \lambda)q_+(\lambda)$$

and therefore

$$(2.21) \quad \int I(A, -\gamma_+(\lambda) \leq X < \lambda)(X - \lambda)dP \leq \int I(A^c, X < -\gamma_+(\lambda))(X - \lambda)dP - (\gamma_+(\lambda) + \lambda)q_+(\lambda).$$

This implies that

$$(2.22) \quad (-\gamma_+(\lambda) - \lambda)P(A, -\gamma_+(\lambda) \leq X < \lambda) \leq \text{LHS of } (2.21)$$
$$\leq \text{RHS of } (2.21) \leq (-\gamma_+(\lambda)-\lambda)[P(A^c, X<-\gamma_+(\lambda)) + q_+(\lambda)].$$

If $\gamma_+(\lambda) = \infty$, (2.19) is trivial. Assuming $\gamma_+(\lambda) < \infty$, we may divide the above by $-\gamma_+(\lambda) - \lambda$ and then add $P(A, X < -\gamma_+(\lambda)) + P(X \geq \lambda)$ to both sides to complete the proof of (2.19).

If $P(A) = \mu_+(\lambda)$, then reversing the final steps in the above argument, we see that the extreme left and right sides of (2.22) are equal. This means that

$$\begin{aligned} &\int I(A, -\gamma_+(\lambda) \leq X < \lambda)\ (-\gamma_+(\lambda) - \lambda)dP \\ &= \int I(A, -\gamma_+(\lambda) \leq X < \lambda)(X - \lambda)dP \\ &= \int I(A^c, X < -\gamma_+(\lambda))(X - \lambda)dP - (\gamma_+(\lambda) + \lambda)q_+(\lambda) \\ &= \int I(A^c, X < -\gamma_+(\lambda))\ (-\gamma_+(\lambda)-\lambda)\ dP - (\gamma_+(\lambda) + \lambda)q_+(\lambda). \end{aligned}$$

The last equality implies that $P(A^c, X < -\gamma_+(\lambda)) = 0$ and hence the first inclusion in (2.20) holds. The first equality shows that $P(A, -\gamma_+(\lambda) < X < \lambda) = 0$ and hence the second inclusion in (2.20) holds.

Finally note that, under the hypotheses of (iii), if

$$A_\lambda \equiv A - \{X < -\gamma_+(\lambda) \text{ or } X \geq \lambda\}$$

then $A_\lambda \subset \{X = -\gamma_+(\lambda)\}$ and so

$$0 = \int_A (X - \lambda)dP = \int I(X < -\gamma_+(\lambda) \text{ or } X \geq \lambda)dP - (\gamma_+(\lambda) + \lambda)P(A_\lambda).$$

Solve for $P(A_\lambda)$ to get $P(A_\lambda) = q_+(\lambda)$ and therefore $P(A) = \mu_+(\lambda)$.□

<u>Theorem 2.5</u>. <u>Assume</u> $\{X_t: t \geq 0\}$ <u>is a uniformly integrable (right-continuous) martingale such that</u> $L(X_\infty) = \mu$.

(a) (Blackwell-Dubins (1963)). <u>For all</u> $\lambda > 0$,

$$P(\sup_t X_t \geq \lambda) \leq \mu^*(\lambda). \tag{2.23}$$

<u>If equality holds in (2.23) then</u>

$$\{X_\infty > \phi(\lambda)\} \subset \{\sup_t X_t \geq \lambda\} \subset \{X_\infty \geq \phi(\lambda)\} \text{ a.s.} \tag{2.24}$$

<u>Conversely if</u> X_t <u>is a.s. continuous,</u> $X_0 = 0$ <u>and (2.24) holds, then equality</u>

holds in (2.23).

(b) (Davis) Assume, in addition that X_t is a.s. continuous and $X_0 = 0$.

Then for all $\lambda > 0$,

$$P(\sup_t X_t \geq \lambda) \geq \mu_+(\lambda) \tag{2.25}$$

$$P(-\inf_t X_t \geq \lambda) \geq \mu_-(\lambda). \tag{2.26}$$

Equality holds in (2.25) (respectively, (2.26)) iff

$$\{X_\infty < -\gamma_+(\lambda) \text{ or } X_\infty \geq \lambda\} \subset \{\sup_t X_t \geq \lambda\} \subset \{X_\infty \leq -\gamma_+(\lambda) \text{ or } X_\infty \geq \lambda\} \text{ a.s.} \tag{2.27}$$

(respectively,

$$\{X_\infty \leq -\lambda \text{ or } X_\infty > \gamma_-(\lambda)\} \subset \{-\inf_t X_t \geq \lambda\} \subset \{X_\infty \leq -\lambda \text{ or } X_\infty \geq \gamma_-(\lambda)\} \text{ a.s.)} \tag{2.28}$$

Remark. It is not hard to show that the right side of (2.23) equals $\bar{\mu}[\lambda,\infty)$ where $\bar{\mu}$ is the distribution of the Hardy-Littlewood maximal function associated with μ (see eg. Dubins-Gilat (1978)). Thus (2.23) really is Theorem 3(a) of Blackwell-Dubins (1963) (see also Theorem 1 of Dubins-Gilat (1978)).

Proof. (b) The optional stopping theorem shows that for $\lambda > 0$,

$$\int I(X(T_X(\lambda)) \geq \lambda)(X_\infty - \lambda)dP = \int I(X(T_X(\lambda)) \geq \lambda)\ (X(T_X(\lambda)) - \lambda)dP = 0.$$

Apply Lemma 2.4(b) with $A = \{X(T_X(\lambda)) \geq \lambda\}$ and $X = X_\infty$ to obtain (2.25), and the equivalence between (2.27) and equality holding (2.25). The rest of (b) is obtained by replacing X with -X and μ with $\tilde{\mu}$ (use (2.12) and (2.16) here).

(a) Use Lemma 2.4(a) as above. In this case the possibility of jumps as well as X_0 exceeding λ means that

$$\int I(X(T_X(\lambda)) \geq \lambda)\ (X(T_X(\lambda)) - \lambda)dP \geq 0.$$

Therefore continuity of X and $X_0 = 0$ is needed for the last statement in (a). □

By integrating out (2.25) and (2.26) we see that a necessary condition for the existence of an H^1-embedding of μ is

$$\int_0^\infty \mu_+(\lambda) + \mu_-(\lambda)d\lambda < \infty.$$

It remains to show that this is equivalent to our earlier sufficient condition, $H(\mu) < \infty$.

Lemma 2.6.

(2.29) $$H(\mu) \le 2 \int_0^\infty \mu_+(\lambda) + \mu_-(\lambda)d\lambda$$

(2.30a) $$\int_0^\infty \mu_+(\lambda)d\lambda \le \int_{-\infty}^\infty |x|d\mu(x) + H(\mu)$$

(2.30b) $$\int_0^\infty \mu_-(\lambda)d\lambda \le \int_{-\infty}^\infty |x|d\mu(x) + H(\mu)$$

Proof. Fix $\lambda > 0$ and note that

(2.31) $$\int I(x < -\gamma_+(\lambda) \text{ or } x \ge \lambda)\ (x - \lambda)\ d\mu(x) - (\lambda + \gamma_+(\lambda))q_+(\lambda) = 0.$$

case 1. $\lambda > \gamma_+(\lambda)$

$$\int_{-\infty}^\infty xI(|x| \ge \lambda)d\mu(x) = \int_{-\infty}^\infty xI(|x| \ge \lambda)d\mu(x) - (2.31)$$

$$= -\int_{-\infty}^\infty xI(-\lambda < x < -\gamma_+(\lambda))d\mu(x) + \gamma_+(\lambda)q_+(\lambda) + \lambda\mu_+(\lambda)$$

$$\Rightarrow \mu_+(\lambda) \le \lambda^{-1}\left|\int_{-\infty}^\infty x\ I(|x| \ge \lambda)d\mu(x)\right| \text{ (by (2.14))}$$

(2.32) $$\le \int_{-\infty}^\infty \frac{-x}{\lambda}\ I(-\lambda < x < -\gamma_+(\lambda))d\mu(x) + \frac{\gamma_+(\lambda)}{\lambda} q_+(\lambda) + \mu_+(\lambda) \le 2\mu_+(\lambda).$$

<u>case 2</u>. $\lambda > \gamma_-(\lambda)$.

Replace μ with $\tilde{\mu}$ in the above to get

$$\mu_-(\lambda) \leq \lambda^{-1} \left| \int_{-\infty}^{\infty} x\, I(|x| \geq \lambda) d\mu(x) \right| \leq 2\mu_-(\lambda). \tag{2.33}$$

<u>case 3</u>. $\lambda \leq \gamma_+(\lambda)$ and $\lambda \leq \gamma_-(\lambda)$

$$\int_{-\infty}^{\infty} xI(|x| \geq \lambda)\, d\mu(x) = \int_{-\infty}^{\infty} x\, I(|x| \geq \lambda) d\mu(x) - \quad (2.31)$$

$$= \int_{-\infty}^{\infty} x\, I(-\gamma_+(\lambda) \leq x \leq -\lambda) d\mu(x) + \gamma_+(\lambda) q_+(\lambda) + \lambda\, \mu_+(\lambda)$$

$$\leq \int_{-\infty}^{\infty} x\, I(-\gamma_+(\lambda) < x \leq -\lambda) d\mu(x) + \lambda\mu_+(\lambda) \quad \text{(by (2.14))}.$$

$$\leq \lambda\mu_+(\lambda).$$

By symmetry we may conclude that

$$-\lambda\mu_-(\lambda) \leq \int xI(|x| \geq \lambda) d\mu(x) \leq \lambda\, \mu_+(\lambda)$$

$$\therefore \lambda^{-1} \left| \int xI(|x| \geq \lambda) d\mu(x) \right| \leq (\mu_+(\lambda) + \mu_-(\lambda)). \tag{2.34}$$

(2.29) follows by using the upperbounds on $\lambda^{-1}\left|\int xI(|x| \geq \lambda) d\mu(x)\right|$ in (2.32), (2.33), (2.34) and then integrating out λ.

For (2.30a) note that if $\lambda \leq \gamma_+(\lambda)$ then (2.14) shows that

$$\mu_+(\lambda) \leq \mu(\{|x| \geq \lambda\}) \tag{2.35}$$

This, together with the first inequality in (2.32) gives (2.30a) upon integrating. Replace μ with $\tilde{\mu}$ to get (2.30b) from (2.30a).□

It is now an easy matter to prove the main result of this section. Recall the definitions of $H(\mu)$ and $\mu_\pm(\lambda)$ given prior to Proposition 2.1 and Lemma 2.3, respectively.

<u>Theorem 2.7</u>. Let μ be a mean-zero probability on the line and T_s be the Skorokhod embedding of μ. The following are equivalent:

(a) There exists an H^1-embedding of μ.

(b) T_s is an H^1-embedding of μ.

(c) $H(\mu) < \infty$

(d) $\int_0^\infty \mu_+(\lambda) + \mu_-(\lambda) < \infty$

<u>Proof</u>. (d) <=> (c) Lemma 2.6

(c) => (b) Theorem 2.2

(b) => (a) obvious

(a) => (d) Theorem 2.4 (b). □

<u>Remarks</u>. 1. In particular the equivalence of (2.5) and (d) in the atomless case (Prop. 2.1(b)) shows that Walsh's conjecture (2.7) is true.

2. It is a little surprising that if an H^1-embedding of μ exists then T_s must be an H^1-embedding. Clearly there must be other embeddings with this property. Vallois (1982) describes an interesting embedding, T_V, that uses local time, and shows if $\mu\{0\} = 0$, it is an H^1-embedding iff(c) holds in the above (Vallois (1982, Prop. 4.23)). The filling scheme, T_C, has been studied extensively (see eg. Rost (1971), Baxter [3]) and is known to minimize $E(\sqrt{T})$ over all embeddings of μ (see P. Chacon (1985)). Davis' inequality shows that if an H^1-embedding exists then T_C is such an embedding. Indeed, this suggested a direct method of attack on the original problem, namely find

NASC on μ for $E(B^*_{T_c}) < \infty$. Unfortunately, the filling scheme does not seem to lend itself to such explicit calculations.

3. An Optimal Embedding

How should one define an embedding of μ, T, that minimizes B^*_T, M_T or m_T? Theorem 2.5 tells us how one might hope to define such a T. To illustrate the idea let us first try to maximize M_T. According to Theorem 2.5(a), $P(M_t \geq \lambda) \leq \mu^*(\lambda)$, and, if equality holds, T must satisfy

$$\{B_T > \phi(\lambda)\} \subset \{M_T \geq \lambda\} \subset \{B_T \geq \phi(\lambda)\}.$$

The left-continuity of ψ now implies

$$\psi(B_T) > \lambda \Rightarrow B_T > \phi(\lambda) \Rightarrow M_T \geq \lambda.$$

Let $\lambda \uparrow \psi(B_T)$ to conclude that $M_T \geq \psi(B_T)$. This suggests the

<u>Definition</u>. $T_a = \inf\{t > 0: M_t \geq \psi(B_t)\}$ $(\inf \phi = \infty)$.

Only an optimist would expect T_a to be an embedding of μ. In fact it is precisely the embedding studied by Azéma and Yor (1978a,b). The point of this digression is that Theorem 2.5 provides a natural route to their stopping time. Moreover, it is now easy to show that T_a stochastically maximizes M_T over all embeddings T, as was observed in Azéma-Yor (1978b).

<u>Theorem 3.1.</u>

(3.1) $$P(M(T_a) \geq \lambda) = \mu^*(\lambda) \qquad \text{for all } \lambda \geq 0.$$

<u>If</u> $\{X_t: t \geq 0\}$ <u>is a uniformly integrable (right-continuous) martingale such that</u> $L(X_\infty) = \mu$, <u>then</u>

$$P(\sup_t X_t \geq \lambda) \leq P(M(T_a) \geq \lambda) \qquad \text{for all } \lambda \geq 0.$$

Proof. We first show that

$$\phi(B(T_a)) \leq M(T_a). \tag{3.2}$$

If (3.2) fails there must at least exist a sequence $t_n \downarrow\downarrow T_a$ such that $\phi(B(t_n)) \leq M(t_n)$. Choose $u_n \in (T_a, t_n]$ such that $B(u_n) \leq \min(B(T_a), B(t_n))$ (if $B(t_n) \leq B(T_a)$, let $u_n = t_n$). Then $\phi(B_{u_n}) \leq M(t_n)$ and $B(u_n) \leq B(T_a)$, so, letting $n \to \infty$ in the first inequality, we get (3.2) by the left-continuity of ϕ. It follows that for each $\lambda \geq 0$,

$$\{B(T_a) > \phi(\lambda)\} \subset \{M(T_a) \geq \lambda\}. \tag{3.3}$$

The definition of T_a allows one to conclude $\{B(T_a) < \phi(\lambda)\} \subset \{M(T_a) < \lambda\}$ and hence for each $\lambda \geq 0$,

$$\{M(T_a) \geq \lambda\} \subset \{B(T_a) \geq \phi(\lambda)\}. \tag{3.4}$$

(3.3) and (3.4) allow us to apply Theorem 2.5 (a) with $X = B(T_a)$ and $A = \{M(T_a) \geq \lambda\}$ and conclude that (3.1) holds. The rest of the result is then immediate from the Blackwell-Dubins theorem (Theorem 2.5(a)).□

To stochastically minimize M_T, use Theorem 2.5(b) to show that if $P(M_T \geq \lambda) = \mu_+(\lambda)$, then T must satisfy

$$\{B_T < -\gamma_+(\lambda) \text{ or } B_T \geq \lambda\} \subset \{M_T \geq \lambda\} \subset \{B_T \leq -\gamma_+(\lambda) \text{ or } B_T \geq \lambda\}. \tag{3.5}$$

Using the latter inclusion and letting $\lambda = M_T$, we see that

$$\text{if } B_T < 0, \text{ then } B_T \leq -\gamma_+(M_T). \tag{3.6}$$

To simultaneously minimize m_T, we see in the same way that T should also satisfy

$$\text{if } B_T > 0, \text{ then } B_T \geq \gamma_-(m_T). \tag{3.7}$$

(3.6) and (3.7) together suggest the

Definition. $T_d = \inf\{t > 0: B_t \notin (-\gamma_+(M_t), \gamma_-(m_t))\}$ $(\inf \emptyset = \infty)$.

There is a slight problem with this definition. If $\mu_\alpha = \alpha\mu + (1-\alpha)\delta_0$ $(0 < \alpha \leq 1)$, then it is easy to see that $\gamma_\pm^{\mu_\alpha} = \gamma_\pm^\mu$ and hence T_d would be the same for all of these laws. To handle atoms at zero we may, and shall, assume our probability space is rich enough to support a r.v., U, uniformly distributed on [0,1] and independent of B, and make the

Definition. $T_b = \begin{cases} T_d & \text{if } U > \mu(\{0\}) \\ 0 & \text{if } U \leq \mu(\{0\}) \end{cases}$.

We sometimes write T_b^μ or T_d^μ to denote the dependence on μ.

The optimality properties of T_b are fairly easy to show, once one knows that T_b is an embedding of μ. For this we need some further properties of $\gamma_\pm$.

Lemma 3.2. (a) *γ_+ and γ_- are non-decreasing, left-continuous functions from $[0,\infty)$ to $[0,\alpha]$ and $[0,\beta]$, respectively.*

(b) $\lambda < \beta \Rightarrow \gamma_+(\lambda) < \infty$, $\lambda \geq \beta \Rightarrow \gamma_+(\lambda) = \alpha$

$\lambda < \alpha \Rightarrow \gamma_-(\lambda) < \infty$, $\lambda \geq \alpha \Rightarrow \gamma_-(\lambda) = \beta$

(c) $\gamma_\pm(\lambda) > 0$ *if* $\lambda > 0$.

(d) *If* $a, b \geq 0$, $a + b > 0$, $\gamma_+(b) \leq a$, *and* $\gamma_-(a) \leq b$, *then* $\mu([-a,b]^c) = 0$.

Proof. By replacing μ with $\tilde{\mu}$, it suffices to consider γ_+.

(a) It is clear from the definition that γ_+ is non-decreasing and takes values in $[0,\alpha]$. If $\lambda_n \uparrow \lambda$ and $y > -\gamma_+(\lambda)$, then

$$0 > \int (x - \lambda)I(x \leq y \text{ or } x \geq \lambda)d\mu(x) = \lim_{n\to\infty} \int (x - \lambda_n)I(x \leq y \text{ or } x \geq \lambda_n)d\mu(x),$$

and so for large enough n we have $y \geq -\gamma_+(\lambda_n)$. This shows that $-\gamma_+(\lambda) \geq \lim_{n\to\infty} -\gamma_+(\lambda_n)$. As the opposite inequality is obvious by monotonicity, we see that γ_+ is left continuous.

(b) If $\lambda < \beta$, then

$$\lim_{y\to-\infty} \int(x - \lambda)I(x \leq y \text{ or } x \geq \lambda)d\mu(x) = \int(x - \lambda)I(x \geq \lambda)d\mu(x) > 0$$

and hence $-\gamma_+(\lambda) > -\infty$. If $\lambda \geq \beta$ and $y > -\alpha$ then

$$\int(x - \lambda)I(x \leq y \text{ or } x \geq \lambda)d\mu(x) = \int(x - \lambda)I(x \leq y)d\mu(x) < 0$$

and so $-\gamma_+(\lambda) = -\alpha$.

(c) If $\lambda > 0$, then

$$\lim_{\varepsilon\to 0+} \int(x - \lambda)I(x \leq -\varepsilon \text{ or } x \geq \lambda)d\mu(x) = \int (x - \lambda)I(x < 0 \text{ or } x \geq \lambda)d\mu(x)$$

$$\leq -\lambda\ \mu(-\infty,0) < 0.$$

Thus the integral on the left is negative for ε small enough and for such an ε, $\gamma_+(\lambda) \geq \varepsilon > 0$.

(d) $\gamma_+(b) \leq a$ and $\gamma_-(a) \leq b$, together with (2.10) and (2.11) give

$$\int I(x < -a \text{ or } x > b)(x + a)d\mu(x) \leq 0$$

$$\int I(x < -a \text{ or } x > b)(x - b)d\mu(x) \geq 0.$$

Subtracting, we get

$$(a + b)\ \mu([-a,b]^c) \leq 0,$$

and hence the result.□

Notation. Let $\sigma_\pm(\gamma) = \sigma_\pm^\mu(\gamma)$ denote the left-continuous inverse of $\gamma_\pm$, i.e., $\sigma_\pm(\gamma) = \inf\{\lambda \geq 0 : \gamma_\pm(\lambda) \geq \gamma\}$ (inf $\phi = +\infty$).

Lemma 3.3. (a) $\gamma < \alpha \Rightarrow \sigma_+(\gamma) < \infty$, $\gamma < \beta \Rightarrow \sigma_-(\gamma) < \infty$

(b) $\sigma_\pm(0+) = 0$

(c) (3.8) $\int I(x \leq -\gamma \text{ or } x \geq \sigma_+(\gamma))(x - \sigma_+(\gamma))d\mu(x) = 0 \quad (0 \cdot (-\infty) = 0)$

(3.9) $\int I(x \leq -\sigma_-(\gamma) \text{ or } x \geq \gamma)\ (x + \sigma_-(\gamma))d\mu(x) = 0 \quad (0 \cdot \infty = 0)$

(d) $\sigma_-(\sigma_+(s)) \leq s$ for all s in $[0,\alpha]$.

Proof. As usual, it suffices to consider σ_+.

(a) If $\gamma < \alpha$, then

$$\lim_{\lambda \to +\infty} \int (x - \lambda) I(x \leq -\gamma \text{ or } x \geq \lambda)d\mu(x) = -\infty.$$

Therefore $-\gamma_+(\lambda) < -\gamma$ for λ large enough, whence $\sigma_+(\gamma) < \infty$.

(b) is immediate from Lemma 3.2(b).

(c) If $\lambda < \sigma_+(\gamma)$, then $\gamma_+(\lambda) < \gamma$ and so (2.10) shows that

$$\int (x - \lambda) I(x \leq -\gamma \text{ or } x \geq \lambda)d\mu(x) \geq 0 \quad .$$

Let $\lambda \uparrow \sigma_+(\gamma) \leq \infty$, to get

$$\int (x - \sigma_+(\gamma)) I(x \leq -\gamma \text{ or } x \geq \sigma_+(\gamma))d\mu(x) \geq 0 \qquad (-\infty \cdot 0 = 0).$$

If $\sigma_+(\gamma) = \infty$, the above integrand is $(-\infty)I(x \leq -\gamma) \leq 0$ so that the integral must be zero. Assume therefore that $\sigma_+(\gamma) < \infty$ and let $\lambda > \sigma_+(\gamma)$. Then $\gamma_+(\lambda) \geq \gamma$ and so

$$\int (x - \lambda) I(x \leq -\gamma \text{ or } x \geq \lambda)d\mu(x) \leq 0 \qquad \text{(by (2.10))}.$$

Let $\lambda \downarrow \sigma_+(\gamma)$ to get

$$\int (x - \sigma_+(\gamma)) I(x \leq -\gamma \text{ or } x \geq \sigma_+(\gamma))d\mu(x) \leq 0.$$

This, together with the above converse inequality, proves (c).

(d) would follow from

$$(3.10) \qquad \gamma_+(\gamma_-(s)) \geq s \qquad \text{for } 0 \leq s \leq \alpha.$$

If $0 < s < \alpha$ and $\gamma_+(\gamma_-(s)) \leq s$, then Lemma 3.2(d) with $a = s$ and $b = \gamma_-(s)$ shows that $s \geq \alpha$, a contradiction. This proves (3.10) for $0 < s < \alpha$. It is trivial for $s = 0$ and holds for $s = \alpha$ by left-continuity. □

Notation. $G(x) = \mu(-\infty,x]$, $K(\gamma) = \int_0^\gamma (1 - G(x))dx$, $H(\gamma) = \int_{-\gamma}^0 G(x)dx$,

$$f_\pm(t) = \exp\{ \int_t^1 (s + \sigma_\pm(s))^{-1} ds\} \qquad (t > 0).$$

Proposition 3.4. (a) $c_\pm = \lim_{\gamma\to 0+} \gamma f_\pm(\gamma)$ exists and satisfies $0 < c_\pm \leq 1$.

(b) (H,K) satisfies the integral equations

$$(3.11) \qquad H(\gamma) = c_+ G(0-) f_+(\gamma)^{-1} + \int_0^\gamma K(\sigma_+(s)) df_+(s)\ f_+(\gamma)^{-1}, \quad \gamma \geq 0$$

$$(3.12) \qquad K(\gamma) = c_-(1 - G(0)) f_-(\gamma)^{-1} + \int_0^\gamma H(\sigma_-(s)) df_-(s)\ f_-(\gamma)^{-1}, \quad \gamma \geq 0.$$

Proof. As H and f_+ are constant on $\{\gamma: \sigma_+(\gamma) = \infty\}$, it suffices to consider (3.11) for $0 < \gamma$ such that $\sigma_+(\gamma) < \infty$.

Integrate (3.8) by parts to see that for γ as above,

$$(-\gamma - \sigma_+(\gamma))G(-\gamma) - \int I(x \leq -\gamma)G(x)dx - \int I(x \geq \sigma_+(\gamma))(G(x) - 1)dx = 0$$

$$\Rightarrow (\gamma + \sigma_+(\gamma))^{-1}\Big(\int_0^\infty x\ dG(x) - K(\sigma_+(\gamma))\Big)$$

$$= G(-\gamma) + (\gamma + \sigma_+(\gamma))^{-1}\Big(\int_{-\infty}^0 -x\ dG(x) - H(\gamma)\Big)$$

$$\Rightarrow - (\gamma + \sigma_+(\gamma))^{-1} K(\sigma_+(\gamma)) = G(-\gamma) - (\gamma + \sigma_+(\gamma))^{-1} H(\gamma)$$

$$f_+'(\gamma)\ K(\sigma_+(\gamma)) = \frac{d^-}{d\gamma}\ (Hf_+)(\gamma). \tag{3.13}$$

Note that

$$Hf_+(\gamma) \leq H(\gamma)\ \exp\{\int_\gamma^1 s^{-1}ds\} \to G(0-) \text{ as } \gamma \downarrow 0.$$

As $Hf_+(\gamma)$ increases as $\gamma \downarrow 0$ (by (3.13)), it follows that $L_+ \equiv \lim_{\gamma\to 0+} Hf_+(\gamma)$ exists and belongs to $(0, G(0-)]$. We can now integrate (3.13) and conclude that

$$\int_0^\gamma K(\sigma_+(s))df_+(s) = Hf_+(\gamma) - L_+$$

for $0 < \gamma$ such that $\sigma_+(\gamma) < \infty$ and hence for all $\gamma \geq 0$. To obtain (3.11), simply note that

$$\lim_{\gamma\to 0+} \gamma\ f_+(\gamma) = \lim_{\gamma\to 0+} \gamma H(\gamma)^{-1} \lim_{\gamma\to 0+}\ H(\gamma)f_+(\gamma) = G(0-)^{-1}L_+ \equiv c_+ \in (0,1].$$

The rest of (a) and (3.12) follow upon replacing μ with $\tilde{\mu}$.□

<u>Proposition 3.5</u>. <u>If</u> μ_1, μ_2 <u>are mean-zero laws such that</u> $\gamma_\pm^{\mu_1} = \gamma_\pm^{\mu_2}$ <u>and</u> $\mu_1(\{0\}) = \mu_2(\{0\})$, <u>then</u> $\mu_1 = \mu_2$.

<u>Proof</u>. Let $G_i(x) = \mu_i(-\infty,x]$, define H_i and K_i as above but with G_i in place of G, and write $\gamma_\pm$, $\sigma_\pm$, and $f_\pm$ for $\gamma_\pm^{\mu_i}$, $\sigma_\pm^{\mu_i}$, and $f_\pm^{\mu_i}$, respectively ($i = 1,2$). Note that $\alpha^{\mu_i} = \gamma_+^{\mu_i}(\infty)$ and $\beta^{\mu_i} = \gamma_-^{\mu_i}(\infty)$, so we may write α and β for α^{μ_i} and β^{μ_i}, respectively. (3.11) and (3.12) become

$$(3.14)_i \qquad H_i(\gamma) = c_+G_i(0-)f_+(\gamma)^{-1} + \int_0^\gamma K_i(\sigma_+(s))df_+(s)f_+(\gamma)^{-1},\ \gamma \geq 0$$

$$(3.15)_i \quad K_i(\gamma) = c_-(1 - G_i(0))f_-(\gamma)^{-1} + \int_0^\gamma H_i(\sigma_-(s))df_-(s)f_-(\gamma)^{-1},\ \gamma \geq 0.$$

Proposition 3.4(a), together with $f_-(u) \geq u^{-1}$ 1, shows there is a $K > 0$ such that $f_-(u) \geq K^{-1}u^{-1}$ for all $u \geq 0$. Therefore if $\varepsilon < 1 < \gamma$, then we have

$$\begin{aligned}
\int_\varepsilon^\gamma f_-(\sigma_+(s))^{-1}d(-f_+(s)) &\leq K \int_\varepsilon^\gamma \sigma_+(s)f_+(s)(s + \sigma_+(s))^{-1}I(\sigma_+(s) < \infty)ds \\
&\leq K \int_\varepsilon^1 \sigma_+(s)(s^2 + s\,\sigma_+(s))^{-1}I(\sigma_+(s) < \infty)ds \\
&+ K \int_1^\gamma \sigma_+(s)\ (s + \sigma_+(s))^{-1}\ I(\sigma_+(s) < \infty)ds \\
&\leq -K \log(f_+(\varepsilon)\varepsilon) + K \int_1^\gamma I(\sigma_+(s) < \infty)ds \\
&\to -K \log c_+ + K \int_1^\gamma I(\sigma_+(s) < \infty)ds \quad \text{as} \quad \varepsilon \to 0.
\end{aligned}$$

Therefore we may define continuous, non-decreasing functions on $[0,\infty)$ by

$$g_+(\gamma) = \int_0^\gamma f_-(\sigma_+(u))^{-1}d(-f_+(u))$$

and symmetrically,

$$g_-(\gamma) = \int_0^\gamma f_+(\sigma_-(u))^{-1}d(-f_-(u)).$$

Substitute $(3.15)_i$ into $(3.14)_i$ to get

$$f_+(\gamma)H_i(\gamma) = c_+G_i(0-) + c_-(G_i(0) - 1)g_+(\gamma) + \int_0^\gamma \int_0^{\sigma_+(s)} H_i(\sigma_-(u))f_+(\sigma_-(u))\, dg_-(u)dg_+(s).$$

Take differences and recall that $\Delta G_1(0) = \Delta G_2(0)$ to see that

$$(3.16) \qquad f_+(\gamma)(H_1(\gamma) - H_2(\gamma)) = (G_1(0) - G_2(0))(c_+ + c_-g_+(\gamma)) + \int_0^\gamma \int_0^{\sigma_+(s)} (H_1(\sigma_-(u)) - H_2(\sigma_-(u)))f_+(\sigma_-(u))dg_-(u)dg_+(s).$$

Assume $G_1(0) > G_2(0)$. Then $\sigma_-(\sigma_+(s)) \leq s$ for $s \leq \alpha$ (Lemma 3.3(d)) and (3.16) show that $H_1(\gamma) > H_2(\gamma)$ for $\gamma \leq \alpha$ and hence for all $\gamma > 0$ because $H_i(\gamma) = H_i(\gamma \wedge \alpha)$. Let $\gamma \to +\infty$ in (3.16) to see that

$$(3.17) \qquad \int_{-\infty}^{0} -x \, dG_1(x) - \int_{-\infty}^{0} -x \, dG_2(x) = H_1(\infty) - H_2(\infty) > 0.$$

Take differences in $(3.15)_i$ to get

$$(3.18) \qquad K_2(\gamma) - K_1(\gamma) = (G_1(0) - G_2(0))c_-f_-(\gamma)^{-1}$$
$$+ \int_0^{\gamma} (H_1 - H_2)(\sigma_-(s))d(-f_-)(s)f_-(\gamma)^{-1}.$$

Letting $\gamma \to \infty$, we obtain

$$(3.19) \qquad \int_0^{\infty} xdG_2(x) - \int_0^{\infty} xdG_1(x) \geq (G_1(0) - G_2(0))c_-f_-(\gamma)^{-1} \geq 0.$$

Add (3.17) and (3.19) and conclude that $\int_{-\infty}^{\infty} x \, d(G_2 - G_1)(x) > 0$, contradicting the fact that G_1 and G_2 have mean zero. Hence our original assumption was false and we may conclude that $G_1(0) = G_2(0)$. (3.16) simplifies to

$$(3.20) \; f_+(\gamma)(H_1(\gamma)-H_2(\gamma)) = \int_0^{\gamma} \int_0^{\sigma_+(s)} (H_1(\sigma_-(u))-H_2(\sigma_-(u)))f_+(\sigma_-(u))dg_-(u)dg_+(s).$$

Proposition 3.4(a) shows that

$$M(u) \equiv \sup_{0 \leq \gamma \leq u} f_+(\gamma)|H_1(\gamma) - H_2(\gamma)| < \infty, \text{ for all } u \geq 0,$$

and (3.20) implies

$$M(u) \leq \int_0^u M(\sigma_-(\sigma_+(s)))g_-(\sigma_+(s))dg_+(s)$$
$$\leq g_-(\sigma_+(u)) \int_0^u M(s)dg_+(s) \qquad \text{for } u \leq \alpha, \text{ by Lemma 3.3(d).}$$

An appropriate version of Gronwall's lemma shows that $M(u) = 0$ on $[0,\alpha) \subset \{\gamma:\sigma_+(\gamma) < \infty\}$ (Lemma 3.3(a)). As $H_i(\gamma) = H_i(\gamma \wedge \alpha)$, we have proved $H_1 = H_2$ and hence $K_1 = K_2$ by (3.18) and the fact that $G_1(0) = G_2(0)$. Differentiate to see that $G_1 = G_2$.□

Lemma 3.6. $B(T_b) = \begin{cases} -\gamma_+(M(T_b)) = -m(T_b), & \underline{\text{if}}\ B(T_b) \le 0 \\ \gamma_-(m(T_b)) = M(T_b), & \underline{\text{if}}\ B(T_b) \ge 0 \end{cases}$ a.s.

Proof. If $T_b = 0$, the result is obvious. By symmetry it suffices to consider the case when $B(T_b) \le 0$ and $T_b > 0$. By definition there are $t_n \downarrow T_b$ such that $B(t_n) \le -\gamma_+(M(t_n))$ and $M(t_n) = M(T_b)$ a.s. (the latter because $M(T_b) > 0 \ge B(T_b)$ a.s.). Therefore $B(t_n) \le -\gamma_+(M(T_b))$ and we can let $n \to \infty$ to see that $B(T_b) \le -\gamma_+(M(T_b))$ a.s. If $u_n \uparrow\uparrow T_b$, then for a.a. ω and large enough n we have

$$B(u_n) > -\gamma_+(M(u_n)) = -\gamma_+(M(T_b)).$$

Let $n \to \infty$ in the above to obtain $B(T_b) \ge -\gamma_+(M(T_b))$. This proves $-\gamma_+(M(T_b) = B(T_b)$. If $0 < t < T_b$, then

$$B(t) \ge -\gamma_+(M_t) \ge -\gamma_+(M(T_b)) = B(T_b),$$

and therefore $B(T_b) = -m(T_b)$.□

Notation. $\delta = \delta^\mu = \sup\{x \ge 0: \mu[0,x) = 0\}$.

$-\varepsilon = -\varepsilon^\mu = \inf\{x \le 0: \mu(x,0] = 0\}$

Theorem 3.7. T_b *is an embedding of* μ.

Proof. case 1. $-\infty < \alpha \leq -\varepsilon < 0 < \delta \leq \beta < \infty$.

In this case $-\alpha \leq -\gamma_+(\lambda) \leq -\varepsilon$ and $\delta \leq \gamma_-(\lambda) \leq \beta$ for $\lambda > 0$ and hence $B(t \wedge T_b)$ is uniformly bounded, $0 < T_b < \infty$, and $B(T_b) \neq 0$ a.s. Let ν denote the law of $B(T_b)$ and continue to write $\gamma_\pm$ for $\gamma_\pm^\mu$. We will use Proposition 3.5 to show $\nu = \mu$. The previous lemma shows

$$(3.21)\quad \{B(T_b) \geq \lambda \text{ or } B(T_b) < -\gamma_+(\lambda)\}$$
$$\subset \{M(T_b) \geq \lambda\} \subset \{B(T_b) \geq \lambda \text{ or } B(T_b) \leq -\gamma_+(\lambda)\} \qquad \text{for all } \lambda \geq 0,$$

which in turn implies

$$\int I(B(T_b) \geq \lambda \text{ or } B(T_b) < -\gamma_+(\lambda))\, B(T_b)dP \geq \int I(M(T_b) \geq \lambda)\, B(T_b)dP$$
$$= \lambda\, P(M(T_b) \geq \lambda) \text{ (optional stopping)}$$
$$\geq \lambda\, P(B(T_b) \geq \lambda \text{ or } B(T_b) < -\gamma_+(\lambda)).$$

It follows immediately that $\gamma_+^\nu(\lambda) \leq \gamma_+(\lambda)$ for all $\lambda \geq 0$. If $\lambda < \beta^\nu$ and $\lambda' \in (\lambda, \beta^\nu)$, then

$$\int I(B(T_b) \geq \lambda' \text{ or } B(T_b) \leq -\gamma_+(\lambda))\, B(T_b)dP$$
$$\leq \int I(M(T_b) \geq \lambda)\, B(T_b)dP - \int I(\lambda \leq B(T_b) < \lambda')\, B(T_b)dP \text{ (by (3.21))}$$
$$\leq \lambda\, P(M(T_b) \geq \lambda) - \lambda\, P(\lambda \leq B(T_b) < \lambda')$$
$$< \lambda'\, P(B(T_b) \geq \lambda' \quad \text{or} \quad B(T_b) \leq -\gamma_+(\lambda)),$$

the last by (3.21) and the fact that $\lambda' < \beta^\nu$. This shows that for λ, λ' as above, $\gamma_+^\nu(\lambda') \geq \gamma_+(\lambda)$. First let $\lambda' \downarrow\downarrow \lambda$ and then take limits from below (using the left continuity of γ_+^ν, γ_+) to see that $\gamma_+^\nu(\lambda) \geq \gamma_+(\lambda)$ for $\lambda \leq \beta^\nu$. We have therefore shown

$$\gamma_+^\nu(\lambda) = \gamma_+(\lambda) \qquad \text{for } 0 \leq \lambda \leq \beta^\nu \leq \beta$$

(the last inequality is clear because $\gamma_- \leq \beta$), and symmetrically,

$$\gamma_-^{\nu}(\lambda) = \gamma_-(\lambda) \quad \text{for} \quad 0 \le \lambda \le \alpha^{\nu} \le \alpha.$$

In particular, $\gamma_+(\beta^{\nu}) = \gamma_+^{\nu}(\beta^{\nu}) = \alpha^{\nu}$ and $\gamma_-(\alpha^{\nu}) = \gamma_-^{\nu}(\alpha^{\nu}) = \beta^{\nu}$, results that allow us to apply Lemma 3.2(d) and conclude that $\mu([-\alpha^{\nu}, \beta^{\nu}]^c) = 0$. This means $\alpha^{\nu} = \alpha$, $\beta^{\nu} = \beta$ and hence $\gamma_{\pm}^{\nu}(\lambda) = \gamma_{\pm}(\lambda)$ for all $\lambda \ge 0$. As $\mu(\{0\}) = \nu(\{0\}) = 0$, Proposition 3.5 implies $\nu = \mu$.

<u>case 2</u>. $-\infty < \alpha, \beta < \infty$, $\mu(\{0\}) = 0$.

Choose $\varepsilon_n \downarrow\downarrow 0$ and let $K_n = \mu[-\varepsilon_n, \varepsilon_n]$,

$$m_n = \int I(-\varepsilon_n \le x \le \varepsilon_n) x \, d\mu(x)/K_n \qquad (0/0 = 0).$$

Pick r_n in $[0,1]$ such that $m_n = r_n(-\varepsilon_n) + (1 - r_n)\varepsilon_n$ and let

$$\mu_n(A) = \mu([-\varepsilon_n, \varepsilon_n]^c \cap A) + K_n r_n \delta_{-\varepsilon_n}(A) + K_n(1 - r_n)\delta_{\varepsilon_n}(A).$$

μ_n is a mean-zero probability satistying the conditions of case 1.

Therefore, if we write T_n for $T_b^{\mu_n}$ and $\gamma_{\pm}^n$ for $\gamma_{\pm}^{\mu_n}$, then $\mathcal{L}(B(T_n)) = \mu_n$. If $\lambda \ge \varepsilon_n$ and $\gamma_+(\lambda) \ge \varepsilon_n$ then

$$\int I(x \le -\varepsilon_n \text{ or } x \ge \lambda) \; (x - \lambda) d\mu_n(x) < 0$$

and so

$$\begin{aligned} -\gamma_+^n(\lambda) &= \sup\{y < -\varepsilon_n : \int I(x \le y \text{ or } x \ge \lambda) \; (x - \lambda) d\mu_n(x) \ge 0\} \\ &= \sup\{y < -\varepsilon_n : \int I(x \le y \text{ or } x \ge \lambda) \; (x - \lambda) \, d\mu(x) \ge 0\} \\ &= -\gamma_+(\lambda) \qquad\qquad (\because \; \gamma_+(\lambda) \ge \varepsilon_n). \end{aligned}$$

By symmetry we have

$$\gamma_{\pm}^n(\lambda) = \gamma_{\pm}(\lambda) \quad \text{if } \lambda \ge \varepsilon_n \quad \text{and} \quad \gamma_{\pm}(\lambda) \ge \varepsilon_n. \tag{3.22}$$

Choose $q_n \downarrow\downarrow 0$ such that $\mu[-q_n, q_n] \le 2^{-n}$ and then $p_n \downarrow\downarrow 0$ such that

$$P(\max(T_B(p_n), T_B(-p_n)) > T_{|B|}(q_n)) \le 2^{-n}. \tag{3.23}$$

As $\gamma_\pm(\lambda) > 0$ if $\lambda > 0$, (3.22) shows that we may choose $\{\varepsilon_n\}$ so that

$$\gamma_\pm^n(\lambda) = \gamma_\pm(\lambda) \quad \text{for } \lambda \geq p_n \quad \text{and} \quad \mu = \mu_n \quad \text{on } [-q_n, q_n]^c.$$

This shows that

$$T_n' \equiv \inf\{t \geq \max(T_B(p_n), T_B(-p_n)): B_t \notin (-\gamma_+^n(M_t), \gamma_-^n(m_t))\}$$

$$\text{(3.24)} \qquad = \inf\{t \geq \max(T_B(p_n), T_B(-p_n)): B_t \notin (-\gamma_+(M_t), \gamma_-(m_t))\},$$

and therefore

$$\begin{aligned} P(T_n \neq T_n') &\leq P(T_n \leq \max(T_B(p_n), T_B(-p_n))) \\ &\leq P(T_n \leq T_{|B|}(q_n)) + 2^{-n} \qquad \text{(by (3.23))} \\ &\leq \mu_n[-q_n, q_n] + 2^{-n} \\ &= \mu[-q_n, q_n] + 2^{-n} \leq 2^{-n+1}. \end{aligned}$$

The Borel-Cantelli lemma implies

$$\text{(3.25)} \qquad T_n = T_n' \quad \text{for large enough } n \text{ a.s.}$$

(3.24) shows that $T_n' \downarrow T_\infty' \geq T_b$ a.s. Let $t \in (0, T_\infty')$ and choose n large enough so that $\max(T_B(p_n), T_B(-p_n)) \leq t$. We must have $B_t \in (-\gamma_+(M_t), \gamma_-(m_t))$ because $t < T_n'$. This shows that $T_\infty' \leq T_b$ and hence (3.25) shows that $T_n \to T_\infty' = T_b$ a.s. Therefore $B(T_n) \to B(T_b)$ a.s. This shows that $\mathcal{L}(B(T_b)) = \mu$ because $\mathcal{L}(B(T_n)) = \mu_n \overset{w}{\to} \mu$. T_b is an embedding of μ because $B(T_b \wedge t)$ is bounded.

<u>case 3</u>. $\mu(\{0\}) = 0$.

Let $-\alpha_n \downarrow\downarrow -\alpha$ ($\alpha_n > 0$) and define

$$\beta_n = \inf\{\lambda \geq 0: \int_{-\infty}^{0} (-x) \wedge \alpha_n \, d\mu(x) = \int_0^\infty x \wedge \lambda \, d\mu(x)\}$$

Then $0 < \beta_n \uparrow \beta$, $\beta_n < \beta$ and

$$\mu_n(A) \equiv \mu(-\infty, -\alpha_n]\delta_{\alpha_n}(A) + \mu(A \cap (-\alpha_n, \beta_n)) + \mu[\beta_n, \infty)\delta_{\beta_n}(A).$$

is a mean-zero probability with compact support. Therefore if we write T_n for $T_b^{\mu_n}$ and $\gamma_\pm^n$ for $\gamma_\pm^{\mu_n}$, then $L(B(T_n)) = \mu_n$. An argument similar to that given in case 2 shows that

(3.26) $$\gamma_+^n(\lambda) = \gamma_+(\lambda) \quad \text{if} \quad \lambda \leq \beta_n \text{ and } \gamma_+(\lambda) \leq \alpha_n$$

(3.27) $$\gamma_-^n(\lambda) = \gamma_-(\lambda) \quad \text{if} \quad \lambda \leq \alpha_n \quad \text{and} \quad \gamma_-(\lambda) \leq \beta_n.$$

Note that since μ and μ_n are mean-zero laws that agree on $(-\alpha_n, \beta_n)$, one has

$$\begin{aligned}\int I(x \leq -\alpha_n \text{ or } x \geq \beta_n)\, x d\mu(x) &= \int I(x \leq -\alpha_n \text{ or } x \geq \beta_n) x d\mu_n(x)\\ &= -\alpha_n \mu(-\infty, -\alpha_n] + \beta_n \mu[\beta_n, \infty)\\ &< \beta_n(\mu(-\infty, \alpha_n] + \mu[\beta_n, \infty)).\end{aligned}$$

This shows that $\gamma_+(\beta_n) \geq \alpha_n$ and hence $\gamma_+(\lambda) \geq \alpha_n \geq \gamma_+^n(\lambda)$ for $\lambda \geq \beta_n$ or $\gamma_+(\lambda) \geq \alpha_n$. Combine this with (3.26) to conclude that $\gamma_+(\lambda) \geq \gamma_+^n(\lambda)$ for all $\lambda \geq 0$. As μ may be replaced by μ_{n+1} in the definition of μ_n, this in fact shows $\gamma_\pm^n \leq \gamma_\pm^{n+1} \leq \gamma_\pm$ and therefore $T_n \uparrow T_\infty \leq T_b$, and $\gamma_\pm^n \uparrow \gamma_\pm^\infty \leq \gamma_\pm$. Fix $\lambda \geq 0$ and choose $\gamma > \gamma_+^\infty(\lambda)$ such that $\mu(\{-\gamma\}) = \mu_n(\{-\gamma\}) = 0$ for each n. $\mu_n \overset{w}{\to} \mu$ and

(3.28) $$\int |x| d\mu_n \to \int |x| d\mu \qquad \text{as} \qquad n \to \infty.$$

It follows that

$$\begin{aligned}\int I(x \leq -\gamma \text{ or } x \geq \lambda)(x - \lambda) d\mu(x) &= \lim_{n\to\infty} \int I(x \leq -\gamma \text{ or } x \geq \lambda)(x - \lambda) d\mu_n(x)\\ &\geq 0 \quad (\gamma > \gamma_+^\infty(\lambda) \geq \gamma_+^n(\lambda)).\end{aligned}$$

Therefore $\gamma \geq \gamma_+(\lambda)$ and letting $\gamma \downarrow \gamma_+^\infty(\lambda)$ we see that $\gamma_+^\infty(\lambda) \geq \gamma_+(\lambda)$. By symmetry we have shown that $\lim_{n\to\infty} \gamma_\pm^n(\lambda) = \gamma_\pm(\lambda)$ for all $\lambda \geq 0$.

We now show $T_b < \infty$ a.s. If α or β is finite this is obvious because $T_b \leq \min(T_B(-\alpha), T_B(\beta))$. Assume therefore $\alpha = \beta = \infty$. (3.26) and (3.27) show $T_b = T_n$ if $M(T_n) \leq \beta_n \wedge \sigma_+(\alpha_n) \equiv a_n$ and $m(T_n) \leq \alpha_n \wedge \sigma_-(\beta_n) \equiv b_n$. These latter conditions are implied by

$$-b_n \vee (-\gamma_+(a_n)) < B(T_n) < a_n \wedge \gamma_-(b_n)$$

(Lemma 3.6). Therefore

$$\begin{aligned} P(T_b = T_n) &\geq P(-b_n \vee (-\gamma_+(a_n)) < B(T_n) < a_n \wedge \gamma_-(b_n)) \\ &= \mu((-\gamma_+(a_n) \vee (-b_n), \ \gamma_-(b_n) \wedge a_n) \\ &\to 1 \quad \text{as} \quad n \to \infty \end{aligned}$$

because $\sigma_{\pm}(\infty) = \gamma_{\pm}(\infty) = \infty$ if $\alpha = \beta = \infty$. The fact that $T_n < \infty$ for all n a.s. now shows that $T_b < \infty$ a.s., and hence $T_\infty < \infty$ a.s. also.

If $B(T_n) = \gamma_-^n(m_{T_n})$ for infinitely many n then, taking limits, we see that $B(T_\infty) > 0$ a.s. (recall $P(T_n = 0) = 0$ by case 2). Therefore if $B(T_\infty) \leq 0$, Lemma 3.6 shows

$$B(T_\infty) = \lim_{n\to\infty} - \gamma_+^n(M(T_n)) = -\gamma_+(M(T_\infty)).$$

The last equality holds because $\gamma_+^n \uparrow \gamma_+$ and the limit is left-continuous (Lemma 3.2). This result, together with a similar conclusion if $B(T_\infty) \geq 0$, shows that $T_\infty \geq T_b$. Therefore $T_n \uparrow T_b$ and so $\mathcal{L}(B(T_b))$ is the weak limit of $\mathcal{L}(B(T_n)) = \mu_n$, namely μ. (3.28) implies that $\{B(T_n): n \in \mathbb{N}\}$, and hence $\{B(T_b \wedge t): t \geq 0\}$, is uniformly integrable.

<u>case 4</u>. General μ.

Assuming without loss of generality that $\mu(\{0\}) < 1$, let $\nu(A) = \mu(A|\mathbb{R} - \{0\})$. Then $\gamma_{\pm}^{\nu} = \gamma_{\pm}^{\mu}$ and therefore $T_b^{\nu} = T_d^{\mu}$ is an embedding of ν

by the previous case. This implies

$$\mathcal{L}(B(T_b)) = \mu(\{0\})\delta_0 + (1 - \mu(\{0\}))\mathcal{L}(B(T_d^{\mu}))$$
$$= \mu(\{0\})\delta_0 + (1 - \mu(\{0\}))\nu = \mu.$$

The fact that $T_b^{\mu} \leq T_b^{\nu}$ shows that $\{B(t \wedge T_b^{\mu}): t \geq 0\}$ is uniformly integrable. □

THEOREM 3.8. *Let* T *be any embedding of* μ.

(a) *For all* $\lambda > 0$,

(i) $P(M(T_b) \geq \lambda) = \mu_+(\lambda) \leq P(M(T) \geq \lambda)$

(ii) $P(m(T_b) \geq \lambda) = \mu_-(\lambda) \leq P(m(T) \geq \lambda)$

(iii) $P(B^*(T_b) \geq \lambda) = \max\{\mu_+(\lambda), \mu_-(\lambda), \mu(|x| \geq \lambda)\} \leq P(B^*(T) \geq \lambda)$.

(b) *If* $E(M(T) + m(T)) = E(M(T_b) + m(T_b))$, *then* $T = T_d^{\mu}$ *on* $\{T > 0\}$ a.s. and $P(T = 0) = \mu(\{0\})$. *In particular, if* $\mu(\{0\}) = 0$, *then* $T = T_b^{\mu}$ a.s.

Proof. (a) (3.21) shows that we may use Theorem 2.5(b) to conclude that $P(M(T_b) = \lambda) = \mu_+(\lambda)$. By symmetry one gets $P(m(T_b) \geq \lambda) = \mu_-(\lambda)$. The inequalities in (i) and (ii) are immediate from Theorem 2.5(b).

Lemma 3.6 implies that for $\lambda > 0$,

$$\begin{aligned} P(M(T_b) \geq \lambda,\ m(T_b) < \lambda) &= P(M(T_b) \geq \lambda,\ m(T_b) < \lambda,\ B(T_b) > 0) \\ &\quad + P(M(T_b) \geq \lambda,\ m(T_b) < \lambda,\ B(T_b) < 0) \\ &\leq P(\lambda \leq B(T_b) \leq \gamma_-(\lambda)) + P(-\lambda < B(T_b) \leq -\gamma_+(\lambda)). \end{aligned} \tag{3.29}$$

Replace B with $-B$ and μ with $\tilde{\mu}$ to see that

$$\begin{aligned} P(m(T_b) \geq \lambda,\ M(T_b) < \lambda) &\leq P(-\gamma_+(\lambda) \leq B(T_b) \leq -\lambda) \\ &\quad + P(\gamma_-(\lambda) \leq B(T_b) < \lambda). \end{aligned} \tag{3.30}$$

To prove (iii) we consider four cases.

case 1. $\gamma_-(\lambda) < \lambda \le \gamma_+(\lambda)$.

$$P(B^*(T_b) \ge \lambda) = P(m(T_b) \ge \lambda) + P(M(T_b) \ge \lambda,\ m(T_b) < \lambda)$$
$$= \mu_-(\lambda),$$

by (3.29) and (ii).

case 2. $\gamma_+(\lambda) < \lambda \le \gamma_-(\lambda)$.

Use (3.30) as above to see that $P(B^*(T_b) \ge \lambda) = \mu_+(\lambda)$.

case 3. $\lambda \le \gamma_+(\lambda)$ and $\lambda \le \gamma_-(\lambda)$.

Lemma 3.6 shows that

$$P(B^*(T_b) \ge \lambda) \le P(B(T_b) \ge \min(\lambda, \gamma_-(\lambda)) + P(B(T_b) \le \max(-\lambda, -\gamma_+(\lambda)))$$
$$= P(|B(T_b)| \ge \lambda) = \mu(|x| \ge \lambda).$$

case 4. $\lambda > \gamma_+(\lambda)$ and $\lambda > \gamma_-(\lambda)$.

Choose $\lambda' < \lambda$ such that $\lambda' > \gamma_\pm(\lambda')$. Lemma 3.2(d), with $a = b = \lambda'$, shows that $\mu([-\lambda',\lambda']^c) = 0$ and therefore $\gamma_\pm(t) \le \lambda'$ for all $t \ge 0$. This in turn implies $P(B^*(T_b) \ge \lambda) \le P(B^*(T_b) > \lambda') = 0$.

(iii) follows easily from the above, and (i) and (ii).

(b) If $E(M(T) + m(T)) = E(M(T_b) + m(T_b))$, then (a) shows that

$$P(M(T) \ge \lambda) = P(M(T_b) \ge \lambda) = \mu_+(\lambda)$$
$$P(m(T) \ge \lambda) = P(m(T_b) \ge \lambda) = \mu_-(\lambda)$$

for all $\lambda > 0$. Theorem 2.5(b) gives

$$\{B_T \ge \lambda \text{ or } B_T < -\gamma_+(\lambda)\} \subset \{M_T \ge \lambda\} \subset \{B_T \ge \lambda \text{ or } B_T \le -\gamma_+(\lambda)\} \quad \text{for all rational } \lambda \ge 0 \quad \text{a.s.}$$

Approximating M_T from below by rationals in the latter inclusion, we obtain

(3.31) $$B_T = M_T \quad \text{or} \quad B_T \le -\gamma_+(M_T) \quad \text{a.s.}$$

Symmetrically we have

(3.32) $$B_T = -m_T \quad \text{or} \quad B_T \ge \gamma_-(m_T) \quad \text{a.s.}$$

(3.31) and (3.32), together with Lemma 3.2(c), show that

(3.33) $$\text{if } T > 0, \text{ then } T \ge T_d \quad \text{a.s.}$$

(3.34) $$\{T = 0\} = \{B_T = 0\} \quad \text{a.s.,}$$

whence $P(T = 0) = \mu(\{0\})$. Assuming, without loss of generality that $\mu(\{0\}) < 1$, let $Q(A) = P(A|T > 0) = P(A|B_T \neq 0)$ and $\nu(C) = \mu(C|\{0\}^c)$. Then B_t is a Q-Brownian motion and T (on$(\Omega,\mathcal{F},Q)$) is an embedding of ν. (B,T_d) is independent of $\{T = 0\}$ because T_d is measurable function of B. Therefore

$$Q(B(T_d) \in A) = P(B(T_d) \in A) = \nu(A).$$

(recall from case 4 of Theorem 2.7 that $T_d^\mu = T_b^\nu$). Hence T and T_d are both embeddings of ν (on$(\Omega,\mathcal{F},Q)$) and $T \ge T_d$ Q-a.s. by (3.33). This implies $T = T_d$ Q-a.s. (see Chacon-Ghoussoub (1979, p.27)) and therefore $T = T_d$ a.s. on $\{T > 0\}$. If $\mu(\{0\}) = 0$, then $P(T = 0) = P(T_b = 0) = 0$ and one has $T = T_b$ a.s. □

Remarks. 1. (b) shows that T_b is the essentially unique embedding that minimizes $E(M_T + m_T)$ over all embeddings. The corresponding uniqueness theorem for $E(B_T^*)$ is false. Indeed, if μ is symmetric, then the Skorokhod time T_s, the filling scheme T_c, and T_b all satisfy $L(B_T^*) = L(|B_T|)$.

2. It is now of some interest to compute T_b in some specific cases. If μ assigns probability 1/4 to each of the points ±2, ±1 then $\gamma_\pm^\mu(\lambda) = 1 + I(\lambda > 1/3)$ (for $\lambda > 0$). If $\tau(A) = \inf\{t \ge 0: B_t \in A\}$, then

$$T_b = \begin{cases} \tau(1) & \text{if} \quad m(\tau(1)) \le 1/3 \\ \tau(-1) & \text{if} \quad M(\tau(-1) \le 1/3 \\ \tau(\{\pm 2\}) & \text{otherwise} \end{cases} .$$

If μ is the uniform distribution on $[-1,1]$, then

$$\gamma^{\mu}_{\pm}(\lambda) = \begin{cases} 2\sqrt{\lambda} - \lambda, & \lambda \le 1 \\ 1 \quad , & \lambda > 1 \end{cases} ,$$

and so

$$T_b = \inf\{t > 0\colon B_t \ge 2\sqrt{m_t} - m_t \quad \text{or} \quad B_t \le -2\sqrt{M_t} + M_t\}.$$

These results are most impressive if you start with the definition of T_b and ask an unsuspecting friend for the law of $B(T_b)$.

3. The existence of a unique, and fairly explicit, extremal embedding should be compared to section 4 of Davis (1980), where a similar question is considered for rearrangements of an integrable function f on the unit circle. Here the problem is to find a rearrangement of f of minimal H^p-norm for $0 < p < 2$ and maximal H^p-norm for $2 < p < \infty$. In this setting the extremal problem is harder to solve because one must work with a restricted class of continuous martingales. Indeed there need not be an extremal rearrangement (in the above sense) in general, and even if one exists, it may be rather difficult to describe explicitly. Note also the extremality properties of T_b are stronger than those of the extremal rearrangement obtained by Davis. This is essentially caused by the restriction $\int \tilde{f}^2 dm = \int f^2 dm$ where $\tilde{f}$ is the conjugate function of f.

4. If ν is a second mean-zero probability on $\mathbb{R}$, write $\nu \prec \mu$ if there is a Brownian motion B and a stopping time T such that $\mathcal{L}(B_0) = \nu$, $\mathcal{L}(B_T) = \mu$ and $B_{t\wedge T}$ is uniformly integrable. Such a T is an embedding from ν to μ, and is

called an H^1-embedding from ν to μ if, in addition, $E(B_T^*) < \infty$. It is easy to see that

(3.35) $\qquad \nu \prec \mu \iff \mu = \nu * \eta$ for some probability η.

If $h_\nu(t) = \int e^{itx} d\nu(x)$ and

(3.36) $\qquad \{t: h_\nu(t) \neq 0\}$ is dense in $\mathbb{R}$,

then the law, η, appearing in (3.35) is unique because $h_\eta(t) = h_\mu(t)/h_\nu(t)$ on a dense set of t.

Let $\tilde{B}$ be an $\{\mathcal{F}_t\}$-Brownian motion starting at zero, B_0 an $\mathcal{F}_0$-measurable r.v. with law ν (a mean-zero law), and $B_t = B_0 + \tilde{B}_t$. Assume (3.35) and let $T_b = \tilde{T}_b^\eta$ denote the embedding of η in $\tilde{B}$ considered in Theorem 3.7 (η is some fixed law obtained from (3.35)). We may, and shall, assume $(\tilde{B}, T_b)$ is independent of B_0. Then

$$\mathcal{L}(B(T_b)) = \mathcal{L}(B_0 + \tilde{B}(\tilde{T}_b^\eta)) = \nu * \eta = \mu,$$

and hence T_b is an embedding from ν to μ.

Let T be any embedding from ν to μ and let $\mathcal{L}(B_T - B_0) = \eta$. If $T_b = \tilde{T}_b^\eta$, as above, then

$$\begin{aligned} P(B_0 < \lambda \leq M_T) &= \int I(x < \lambda)\, P(\tilde{M}_T \geq \lambda - x)\, d\nu(x) \qquad (\tilde{M}_T = \inf_{s \leq T} \tilde{B}_s) \\ &\geq \int I(x < \lambda)\, P(\tilde{M}_{T_b} \geq \lambda - x)\, d\nu(x) \qquad \text{(Theorem 3.8)} \\ &= P(B_0 < \lambda \leq M_{T_b}), \end{aligned}$$

and therefore

(3.37) $\qquad P(M_T \geq \lambda) \geq P(M_{T_b} \geq \lambda) \quad$ for all $\lambda > 0$.

Similarly we have

(3.38) $\qquad P(m_T \geq \lambda) \geq P(m_{T_b} \geq \lambda) \quad$ for all $\lambda > 0$.

Assume (3.36). Then (3.37) and (3.38) holds for any embedding from ν to μ. If $E(M_T + m_T) = E(M_{T_b} + m_{T_b})$ then $E(\tilde{M}_T + \tilde{m}_T) = E(\tilde{M}_{T_b} + \tilde{m}_{T_b})$. The uniqueness of η in (3.35) shows that T is an embedding of η in $\tilde{B}$. Therefore Theorem 3.8(b) shows that $T = \tilde{T}_d^\eta$ on $\{T > 0\}$ a.s. (the ~ indicates the underlying Brownian motion is $\tilde{B}$) and in particular $T = T_b$ a.s. if $\eta(\{0\}) = 0$.

Finally the above remarks (especially (3.37), (3.38)) together with Theorems 3.8 and Lemma 2.6 prove

Theorem 3.9. *Let ν, μ be a mean-zero probabilities on $\mathbb{R}$. There is an H^1-embedding from ν to μ iff there is a probability η such that $\mu = \nu * \eta$ and $H(\eta) < \infty$.*

Acknowledgement. I wish to thank Burgess Davis and John Walsh for explaining their results to me and also Paul-Andrè Meyer and Michel Emery for an enjoyable visit to the University of Strasbourg, during which most of this work was done.

References

1. J. Azéma, M. Yor (1978a). Une solution simple au probléme de Skorokhod. Séminaire de Probabilités XIII, p. 90-115. Lect. Notes in Math. 721, Springer-Verlag, Berlin, Heidelberg, New York.

2. J. Azéma, M. Yor (1978b). Le probléme de Skorokhod: compléments à l'exposé précedent. Séminaire de Probabilités XIII, p.626-633. Lect. Notes in Math. 721, Springer-Verlag, Berlin, Heidelberg, New York.

3. J. R. Baxter. Balayage in least time, unpublished manuscript.

4. D. Blackwell, L. E. Dubins (1963). A converse to the dominated convergence theorem, Illinois J. Math. 7, 508-514.

5. D. L. Burkholder, R. F. Gundy and M. L. Silverstein (1971). A maximal function characterization of the class H^p, Trans Amer. Math. Soc. 157, 137-153.

6. O. D. Cereteli (1976). A metric characterization of the set of functions whose conjugate functions are integrable, Bull. of the Academy of Sciences of the Georgian S.S.R. 81, 281-283 (in Russian).

7. P. Chacon (1985). Ph.D. thesis, U. of Washington.

8. R. Chacon, N. Ghoussoub (1979). Embeddings in Brownian motion, Annales de l'Institut Henri Poincaré, Section B 15, 287-292.

9. B. Davis (1980). Hardy spaces and rearrangements, Trans. Amer. Math. Soc. 261, 211-233.

10 B. Davis (1982). On the integrability of the ergodic maximal function, Studia Mathematica 73, 153-167.

11. L. E. Dubins, D. Gilat (1978). On the distribution of maxima of martingales, Proc. of Amer. Math. Soc. 68, 337-338.

12. H. Rost (1971). The stopping distributions of a Markov process, Inventiones Math. 14, 1-16.

13. A. Skorokhod (1965). Studies in the theory of random processes, Addison-Wesley, Reading.

14. P. Vallois (1982). Le probléme de Skorokhod sur R, une approche avec le temps local. Séminaire de Probabilités XVII, p.227-239. Lect. Notes in Math. 986, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo.