James W. Pitman

1-Dimensional Brownian Motion and the 3-Dimensional Bessel Process

Preprint
August
1974

6

Institute of Mathematical Statistics
University of Copenhagen

J. W. Pitman

ONE-DIMENSIONAL BROWNIAN MOTION
AND THE THREE-DIMENSIONAL BESSEL PROCESS

Preprint 1974 No. 6

INSTITUTE OF MATHEMATICAL STATISTICS

UNIVERSITY OF COPENHAGEN

July 1974

This paper was prepared while the author was supported by the Danish Science Research Council.

1. Introduction.

A Brownian motion process (BM) is a diffusion on $(-\infty,\infty)$ with infinitesimal generator

$$A = \frac{1}{2}\frac{d^2}{dx^2} \quad ,$$

while a three-dimensional Bessel process (BES(3)) is a continuous process identical in law to the radial part of three-dimensional Brownian motion, i.e. a diffusion on $[0,\infty)$ with infinitesimal generator

$$B_3 = \frac{1}{2}\frac{d^2}{dx^2} + \frac{1}{x}\frac{d}{dx}$$

which is the radial part of the three-dimensional Laplace operator. For general background on Brownian motion and diffusion see Breiman [2] , Freedman [8] or Ito-McKean [11] and for information about Bessel processes see Itô-McKean [11] and Williams [18]. The intimacy of the relationship between the two processes BM and BES(3) was brought out in Williams' paper [18], and it was Williams' striking path decomposition theorems for BES(3) which were the inspiration for the main result of the present paper, Theorem 1.3 below. This result was announced in Pitman [15].

To provide the reader with some background we mention first some of Williams' results. We shall make use of the following notation: For a continuous real-valued process $W = \{W(t),\ t \geqq 0\}$ and a real number c, τ_c^W and σ_c^W denote the first and last times that W is at c,

$$\tau_c^W = \inf\{t: t > 0,\ W(t) = c\},\ \inf \emptyset = \infty\ ,$$

$$\sigma_c^W = \sup\{t: t > 0,\ W(t) = c\},\ \sup \emptyset = 0;$$

the starting position of a BM or BES(3) process is indicated by a superscript: thus BM^c stands for Brownian motion started at c. As a start Williams pointed out that the processes BM and BES(3) are dual in the sense that BES(3) is 'upward conditioned BM' while BM is 'downward conditioned BES(3)', a statement made precise by the following proposition:

Proposition 1.1. (Williams [18]) Let $0 < b, c < \infty$, and suppose that X is a BM^b process, Y a $BES^b(3)$ process.

(i) For $b < c$ the conditional distribution of

$$\{X(t), 0 \leqq t \leqq \tau_c^X\} \text{ given } \{\tau_c^X < \tau_0^X\}$$

is identical to the distribution of

$$\{Y(t), 0 \leqq t \leqq \tau_c^Y\} .$$

(ii) For $c < b$ the conditional distribution of

$$\{Y(t), 0 \leqq t \leqq \tau_c^Y\} \text{ given } \{\tau_c^Y < \infty\}$$

is identical to the distribution of

$$\{X(t), 0 \leqq t \leqq \tau_c^X\} .$$

Part (i) of Proposition 1.1 goes back to Doob [3] . See also McKean [14], Knight [10] , and Jacobsen [9] . Williams also showed that BM and BES(3) are dual in another sense which is like the duality by reversal of increments for random walks (see Feller [6]).

Proposition 1.2. (Williams [18]) Let X be a BM^0 process, Y a BES^0 process, and let $0 < c < \infty$. Then the two processes

$$\{c - X(\tau_c^X - t), \quad 0 \leqq t \leqq \tau_c^X\}$$

and

$$\{Y(t), \quad 0 \leqq t \leqq \sigma_c^Y\}$$

are identical in distribution.

Williams used the above results to establish his path decomposition theorems for BM and BES(3), and he also showed how all these results could be extended by use of the method of random time substitution to apply to a large class of pairs of dual diffusions on the line. However the connection between BM and BES(3) which is brought out in Theorem 1·3 below is special in that it is unique to the BM/BES(3) pair with no analogue for other dual pairs of diffusions; even so this result provides an immediate explanation of William's path decomposition theorems for BM and BES(3) which greatly simplifies the original proofs in [18].

For a continuous real valued process $W = \{W(t),\ t \geqq 0\}$ let $M^W = \{M^W(t),\ 0 \leqq t < \infty\}$ be the associated past maximum process defined by

$$M^W(t) = \sup_{0 \leqq s \leqq t} W(s),$$

and let $F^W = \{F^W(t),\ 0 \leqq t < \infty\}$ be the associated future minimum process defined by

$$F^W(t) = \inf_{t \leqq s < \infty} W(s).$$

For real valued processes $U = \{U(t),\ 0 \leqq t < \infty\}$ and $V = \{V(t),\ 0 \leqq t < \infty\}$, and real numbers r and s we denote the process $\{rU(t) + sV(t), 0 \leq t < \infty\}$ by $rU + sV$.

<u>Theorem 1.3</u>. Suppose that X is a BM^0 process, Y a $BES^0(3)$ process. Then

(i) $2M^X - X$ is a $BES^0(3)$ process.

(ii) $2F^Y - Y$ is a BM^0 process.

(iii) The distributions of M^X and F^Y are identical.

Geometrically, the path of the $2M^X - X$ process is obtained from the X path by reflecting this path at each time point in the level of its previous maximum, while the path of the $2F^Y - Y$ process is obtained from the Y path by reflecting this path at each time point in the level of its future minimum. It should be observed that if we set $Y = 2M^X - X$ then as is obvious from a diagram

$$F^Y = M^X \quad \text{a.s.,} \tag{1.4}$$

the exceptional set being that on which $\overline{\lim}_{t \to \infty} M^X(t) < \infty$.

This implies that

$$X = 2M^X - Y = 2F^Y - Y \quad \text{a.s.,}$$

which makes it clear that both (ii) and (iii) are immediate consequences of (i).

Considering the BM^0 process X, it is interesting to compare the assertion (i) above with the well known result of Lévy [13] that M^X-X is reflecting Brownian motion (i.e. a one-dimensional Bessel process). The time homogenous Markov property of $2M^X - X$ is rather more subtle than that of $M^X - X$ since (1.4) makes it clear that unlike $M^X - X$ the process $2M^X - X$ is not Markov with respect to the increasing sequence of σ-fields generated by the Brownian motion X, and thus in contrast to the situation for $M^X - X$ the time homegeneous Markov property of the process $2M^X - X$ cannot be derived from that of the bivariate Markov process (M,X) by a simple transformation of state space (see Dynkin [4]§X.6).

Considering now the $BES^0(3)$ process Y, if we let $X' = 2F^Y - Y$ be the associated BM^0 process then from the fact that $F^Y = M^{X'}$ we see that

$$Y - F^Y = X' - M^{X'},$$

so that from the abovementioned result of Lévy the process $Y - F^Y$ must be a Brownian motion reflected at zero. Thus the excursions of the $BES^0(3)$ process Y above its future minimum process F^Y are Brownian, indeed identical to the excursions of the associated Brownian motion X' below its past maximum process $M^{X'} = F^Y$.

It is well known that the maximum process M^X of a BM^0 process X can be neatly described as follows: the process M^X has inverse process $A^X = \{A^X(r), r \geqq 0\}$ defined by

$$\begin{aligned} A^X(r) &= \inf\{t : t \geqq 0, M^X(t) = r\} \\ &= \inf\{t : t \geqq 0, X(t) = r\}, \qquad r \geqq 0, \end{aligned}$$

which is a process with stationary independent increments, namely the one sided stable process with exponent $\frac{1}{2}$ and rate $\sqrt{2}$ (see Ito - Mckean [11]§ 1.7). Thus the result (iii) of Theorem 1.3 can be translated at once into the following striking fact concerning Brownian motion in three dimensions: a three-dimensional Brownian motion Z started at the origin wanders steadily out to infinity in the sense that Z leaves spheres centred on the origin for the last time at a rate which does not depend on the radius of the spheres; letting $L^Z(r)$ be the last time that Z is inside the sphere S(r) of radius r,

$$L^Z(r) = \sup\{t : Z(t) \in S(r)\}, \qquad r \geqq 0,$$

the process

$$L^Z = \{L^Z(r),\ r \geqq 0\}$$

is a one sided stable process with exponent $\frac{1}{2}$ and rate $\sqrt{2}$.

Theorem 1.3 will be proved in the following way: first a discrete analogue of the theorem will be established which connects two random walks that are associated in a natural way with BM and BES(3), and then the theorem for the diffusions will be obtained by a weak convergence argument. It will be seen that this method can also be used to give illuminating proofs to both Propositions 1.1 and 1.2, thereby providing a unified approach to all these different connections between BM and BES(3).

Section 2 is denoted to the development of some notation and basic results concerning the weak convergence of the approximating random walks. Theorem 1.3 is proved in Section 3, and then in Section 4 it is explained how Theorem 1.3 can be used to establish path decomposition theorems.

2. The approximating random walks.

Let C denote the space of continuous functions

$$w = w(t), \quad t \geqq 0,$$

from $[0,\infty)$ to $\mathbb{R} = (-\infty,\infty)$. Let $\mathcal{R}$ denote the Borel σ-field of $\mathbb{R}$, and let $\mathcal{B}$ be the smallest σ-field on C with respect to which the co-ordinate mappings $X(t)$ from C to $\mathbb{R}$,

$$X(t)\colon \quad w \to w(t), \qquad (t \geqq 0)$$

are $\mathcal{B}|\mathcal{R}$ measurable for each $t \geqq 0$.

Suppose that $Y = \{Y(t),\ t \geqq 0\}$ is an $(\mathbb{R},\mathcal{R})$ valued stochastic process with continuous paths defined on a probability triple $(\Omega,\mathcal{F},P)$. Such a process induces an $\mathcal{F}|\mathcal{B}$ measurable mapping from Ω to C which we shall also denote by Y:

$$Y\colon \ \omega \to Y(\cdot,\omega)$$

where $Y(\cdot,\omega)$ is the sample function corresponding to $\omega \in \Omega$. By the P-distribution of Y we mean the probability PY induced by Y on $(C,\mathcal{B})$:

$$PY(A) = P(Y \in A), \quad A \in \mathcal{B}, \tag{2.1}$$

where the notation PY is preferred to the more conventional PY^{-1}. Thus

$$\{X(t),\ t \geqq 0;\ PY\} \sim \{Y(t),\ t \geqq 0;\ P\},$$

where '$\sim$' means 'is identical in distribution to', or 'has the same finite dimensional distributions as'.

For $x \in \mathbb{R}$ we define a probability P^x on $(C, \mathcal{B})$ as the distribution of a BM^x process, that is to say P^x is Wiener measure on $(C, \mathcal{B})$ corresponding to starting position x. Also, for $x \geqq 0$ we define a probability Q^x on $(C, \mathcal{B})$ as the distribution of a $BES^x(3)$ process.

Let C^* be the set of functions w in C for which $\lim_{t \to \infty} w(t) = \infty$. Of course $C^* \in \mathcal{B}$ and $P^x(C^*) = Q^x(C^*) = 1$. Let $\mathcal{B}^*$ denote the restriction of $\mathcal{B}$ to C^*.

Let $Z = Z_1$ denote the set of integers, and for $\delta > 0$ let Z_δ denote the grid of multiples of δ, $Z_\delta = \{j\delta, j \in Z\}$. For $m \in \mathbb{N} = \{0,1, \dots\}$ we define a function $\rho_\delta^{(m)} : C^* \to [0,\infty)$ as follows:

$$\rho_\delta^{(0)} w = \inf \{t : t \geqq 0, \quad w(t) \in Z_\delta\}$$

$$\rho_\delta^{(m+1)} w = \inf \{t : t \geqq \rho_\delta^{(m)} w, \ w(t) \in Z_\delta, \ w(t) \neq w(\rho_\delta^{(m)} w)\}.$$

Thus the times $0 \leqq \rho_\delta^{(0)} w < \rho_\delta^{(1)} w < \dots$ are the successive times at which the path w reaches a fresh point on the grid Z_δ. Now for each $m \in \mathbb{N}$ define $V_\delta^{(m)} : C^* \to Z_\delta$ by

$$V_\delta^{(m)} w = w(\rho_\delta^{(m)} w),$$

so that $V_\delta^{(0)} w, V_\delta^{(1)} w, \dots$ is the sequence of points on the grid Z_δ which is visited by the path w, with $V_\delta^{(m+1)} \neq V_\delta^{(m)}$.

It is obvious from the strong Markov property that under either P^x or Q^x the sequence $(V_\delta^{(m)}, m \in \mathbb{N})$ is a Markov chain with stationary transition probabilities. Under P^x the Markov chain is a simple symmetric random walk on Z_δ with transition probabilities from $i\delta$ to $j\delta$ given by p_{ij}, where $(p_{ij}) = P$ is the transition matrix on Z defined by

$$p_{ij} = \tfrac{1}{2} \quad \text{if } j = i+1 \text{ or } i-1$$
$$= 0 \quad \text{otherwise,}$$

while under Q^x the Markov chain $(V_\delta^{(m)}, m \in \mathbb{N})$ has transition probabilities from $i\delta$ to $j\delta$ given by q_{ij}, where $(q_{ij}) = Q$ is the transition matrix on the non-negative integers $\mathbb{N}$ defined by

$$q_{0j} = 1 \quad \text{if } j = 1$$
$$= 0 \quad \text{otherwise,}$$

and for $i > 1$

$$q_{ij} = \left(\frac{1}{i}\right)\left(\tfrac{1}{2}\right)(j) \quad \text{if } j = i+1 \text{ or } i-1,$$
$$= 0 \quad \text{otherwise.}$$

This identification of Q follows at once from the fact that BES(3) has scale function $s(y) = -1/y$ (see Breiman[2], Freedman [8] or Itô-McKean [11] for the meaning of this statement and how to derive it from the description of the infinitesimal generator of BES(3) given in the Introduction). Note that the lack of dependence of the transition probabilities on the grid size δ

is to be expected from the fact that the Brownian scale transformation $S_\delta : C \to C$ which takes the path w to the path wS_δ defined by

$$(wS_\delta)(t) = \delta w(t/\delta^2) \tag{2.2}$$

is such that it leaves both P^0 and Q^0 invariant : indeed for $x \geqq 0, \delta > 0$,

$$P^x S_\delta = P^{x\delta}, \qquad Q^x S_\delta = Q^{x\delta}\ , \tag{2.3}$$

the result for Q^x being a consequence of the standard result for P^x and the description of BES(3) as the radial part of three dimensional Brownian motion. Note that the mapping notation wU with the mapping U acting on the right of w will always be used for mappings U from C into C. This is consistent with the notation (2.1) being used in (2.3) for the transformed probabilities.

For $R = P$ or Q we shall henceforth refer to a Markov chain with transitions R as an R random walk. For facts concerning the P random walk see Feller [5], and for the Q random walk see Section 1.12 of Freedman [7]. In view of the embedding described above a great many interesting properties of P and Q random walks are inherited from their parent processes BM and BES(3). For instance both BM and the P random walk are null-recurrent, while both BES(3) and the Q random walk are transient.

But what concerns us here is the fact that the Q random walk is related to the P random walk in exactly the same ways as BES(3) is related to BM. For instance Propositions 1.1 and 1.2 of the Introduction immediately imply analogous connections between the two random walks. The key observation of the present paper is that it is usually an extremely simple matter to establish such connections between the P and Q random walks directly, and that it is then possible to deduce the coresponding results connecting BM and BES(3) by using the fact that the diffusions can be described as limits of their embedded random walks as the grid size δ goes to zero.

To make this precise we now embed the discrete time process $(V_\delta^{(m)}, m \in \mathbb{N})$ back into continuous time by defining a map $V_\delta : C^* \to C$ as follows: set

$$(wV_\delta)(m\delta^2) = V_\delta^{(m)}w, \quad w \in C^*,\ m \in \mathbb{N},$$

and define wV_δ in between multiples of δ^2 by linear interpolation: For $w \in C^*, m \in \mathbb{N}$, $m\delta^2 \leqq t \leqq (m+1)\delta^2$,

$$(wV_\delta)(t) = (t/\delta^2 - m)V_\delta^{(m)}w + (m + 1 - t/\delta^2)\, V_\delta^{(m+1)}w. \tag{2.4}$$

The map V_δ is of course $\mathcal{C}^*/\mathcal{C}$ measurable. For $w \in C^*$ the path wV_δ is a continnous broken line which moves through the same sequence of points on the grid Z_δ as w does, but at deterministic times $0, \delta^2, 2\delta^2, \ldots$ instead of random times $\rho_\delta^{(0)}w, \rho_\delta^{(1)}w, \rho_\delta^{(2)}w, \ldots$. Notice that for $x \in Z_\delta$ the quantity $m\delta^2$ is just the P^x- expectation of $\rho_\delta^{(m)}$.

Now for $x \geqq 0$ define P_δ^x and Q_δ^x as the P^x and Q^x distributions of V_δ:

$$P_\delta^x = P^x V_\delta, \quad Q_\delta^x = Q^x V_\delta.$$

Thus P_δ^x (Q_δ^x) is the distribution on $(C,\mathcal{B})$ of a linearly interpolated Markov chain which starts distributed on the two points of Z_δ closest to x [right at x if $x \in Z_\delta$], and then moves on the grid Z_δ at time intervals of δ^2 with transition probabilities determined by P(Q). Notice that if we denote by TU the composition of two mappings T and U from C to C then the mappings V_δ fit in with the Brownian scale transformations S_δ of (2.2) according to the formula

$$V_\delta = S_{1/\delta} V_1 S_\delta ,$$

and thus (2.3) implies that

$$P_\delta^{x\delta} = P_1^x S_\delta , \qquad Q_\delta^{x\delta} = Q_1^x S_\delta , \tag{2.5}$$

since for instance

$$P_\delta^{x\delta} = P^{x\delta} V_\delta = P^{x\delta} S_{1/\delta} V_1 S_\delta = P^x V_1 S_\delta = P_1^x S_\delta .$$

Let us now give C the topology of uniform convergence on compact sets. The space C is then a polish space, i.e. C can be metrised as a complete separable metric space, and the σ- field $\mathcal{B}$ generated by the co-ordinate maps is identical to the Borel σ - field of C. Let μ and $(\mu_n, n \in \mathbb{N})$ be probabilities on $(C,\mathcal{B})$. We say that the sequence (μ_n) <u>converges weakly to</u> μ, writing

$$\mu_n \Rightarrow \mu ,$$

if as $n \to \infty$

$$\int f d\mu_n \to \int f d\mu$$

for every bounded continuous function f from C to $\mathbb{R}$. For background on weak convergence of probability measures the reader is referred to Billingsley [1]. See Whitt [16] for a discussion with special reference to the space C being considered here.

The diffusions BM and BES(3) can now be described as weak limits of their embedded random walks as the grid size δ converges to zero through the sequence $(1/\sqrt{n},\ n \in \mathbb{N})$. There is also a strong version of the result now stated which gives a.s. convergence of paths, but the statement of this result is deferred to the end of the section since for the applications we have in mind here it is the following result which is of primary importance:

Theorem 2.6. Fix $x \geq 0$. As $n \to \infty$ both

$$P^x_{1/\sqrt{n}} \Rightarrow P^x \qquad (2.6)(i)$$

and

$$Q^x_{1/\sqrt{n}} \Rightarrow Q^x. \qquad (2.6)(ii)$$

Proof. The first assertion is just a particular case of Donsker's theorem (see Billingsley [1], Whitt [16]), while the second is a special case of a theorem on weak convergence of Markov chains to diffusions which was proved by Lamperti in [12].

For the reader who is familiar with the convergence (2.6)(i), strong intuitive grounds for believing (2.6)(ii) should be provided by the fact that for $x \in Z_\delta$ and a function $f : \mathbb{R} \to \mathbb{R}$ the Q_δ^x - expectation of

$$[f(X(\delta^2)) - f(X(0))]/\delta^2$$

is just

$$\tfrac{1}{2}[f(x+\delta) + f(x-\delta) - 2f(x)]/\delta^2 + \tfrac{1}{x}[f(x+\delta)-f(x-\delta)]/2\delta \ ,$$

which for twice differentiable f converges $\delta \to 0$ to

$$\mathcal{B}_3 f = \tfrac{1}{2}\frac{d^2 f}{dx^2} + \frac{1}{x}\frac{df}{dx} \ ,$$

where $\mathcal{B}_3$ is the infinitesimal generator of BES(3). This is exactly analogous to what happens in the more familiar situation corresponding to (2.6)(i). This does not of course provide proof of (2.6)(ii), but an alternative proof of the result along quite different lines will be mentioned at the end of the section.

Our applications of Theorem 2.6 rely on the fact that if μ is a probability on $(C,\mathcal{C})$ and $\Phi : C \to C$ is a $\mathcal{C}/\mathcal{C}$ measurable mapping with set of points of discontinuity which has μ measure zero, then for a sequence (μ_n) of probabilities on $(C,\mathcal{C})$,

$$\mu_n \Rightarrow \mu \quad \text{implies} \quad \mu_n\Phi \Rightarrow \mu\Phi \ , \qquad (2.7)$$

(see Billingsley [1], Theorem 5.5).

Let us now illustrate the use of Theorem 2.6 for establishing connections between BM and BES(3) by giving a proof of part (i) of Proposition 1.1. In the present terminology this result states that for $0 < b < c < \infty$

$$\{X(t),\ 0 \leqq t \leqq \tau_c;\ \ P^b|(\tau_c < \tau_0)\} \sim \{X(t), 0 \leqq t \leqq \tau_c;\ Q^b\}, \tag{2.8}$$

where for $x \in \mathbb{R}$ the first passage function $\tau_x : C \to [0,\infty]$ is defined by

$$\tau_x w = \inf\{t : t > 0,\ w(t) = x\}, \qquad w \in C.$$

The discrete analogue of (2.8) is the following: for $\delta > 0$

$$\{X(t),\ 0 \leqq t \leqq \tau_c;\ P^b_\delta\ |(\tau_c < \tau_0)\} \sim \{X(t),\ 0 \leqq t \leqq \tau_c;\ Q^b_\delta\}\ . \tag{2.9}$$

Now it is easy to see using (2.5) that to establish (2.9) it suffices to consider the case when $\delta = 1$ with b and c both integers. But under either $P^b_1\ |(\tau_c < \tau_0)$ or Q^b_1 the only paths for $\{X(t),\ 0 \leqq t \leqq \tau_c\}$ which have positive probability are broken lines started at b which more either up one or down one at each unit of time through integer values $b, j_1, j_2, \ldots\ j_n$, where $0 < j_k < c$ for $1 \leqq k \leqq n-1$ and $j_n = c$; the P^b_1 probability of this path given $(\tau_c < \tau_0)$ is just

$$2^{-n}[P^b_1(\tau_c < \tau_0)]^{-1} = 2^{-n}[bc^{-1}]^{-1} = b^{-1}2^{-n}c,$$

while the Q^b_1 probability of this path is equally

$$(b^{-1}2^{-1}j_1)(j_1^{-1}2^{-1}j_2)\ \ldots\ (j_{n-1}^{-1}2^{-1}c) = b^{-1}2^{-n}c,$$

which establishes (2.9). Finally, standard arguments show that the mapping $\Phi : C \to C$ defined by

$$(w\Phi)(t) = 1\{\tau_c < \tau_0\} w(t \wedge \tau_c)$$

is $\mathcal{C}/\mathcal{C}$ measurable with discontinuity set which has both P^x and Q^x probability zero (cf. Freedman [7], §1.7)) and (2.8) now follows quickly from Theorem (2.6) using (2.7).

The reader may like to check that the discrete result (2.9) works even for $b = 0$,

$$\{X(t),\ 0 \leqq t \leqq \tau_c;\ P_\delta^0 \mid \tau_c < \tau_0\} \sim \{X(t), 0 \leqq t \leqq \tau_c; Q_\delta^0\}, \tag{2.10}$$

though (2.8) makes no sense for $b = 0$ since the event $\{\tau_c < \tau_0\}$ has P^0 probability zero. Even so, $BES^0(3)$ prior to τ_c behaves just as if it were BM^0 'conditioned to hit c before returning to 0': witness the appearance of $BES^0(3)$ as a component in the decomposition of the Brownian path made by Williams [17]:

$$\{X(t),\ \rho_c \leqq t \leqq \tau_c;\ P^0\} \sim \{X(t),\ 0 \leqq t \leqq \tau_c;\ Q^0\} \tag{2.11}$$

where ρ_c is the last time at zero before τ_c, $\rho_c w = \sup\{t : 0 \leqq t \leqq \tau_c, w(t)=0\}$, $w \in C$.

This identity too can be deduced from the weak convergence (2.6) after observing that (2.10) implies the random walk analogue of (2.11) with P^0 and Q^0 replaced by P_δ^0 and Q_δ^0 (cf. Freedman [7] §4.5). It is interesting to note that the discrete result (2.10) also implies that we must have the following further identity pointed out to me by David Williams:

$$\{X(t),\ 0 \leqq t \leqq \tau_c;\ Q^0\} \sim \{c - X(\tau_c - t),\ 0 \leqq t \leqq \tau_c;\ Q^0\}, \qquad (2.12)$$

a fact which makes (2.11) consistent with Proposition 1.2.

Part (ii) of Proposition 1.1 and Proposition 1.2 can be established in exactly the same kind of way: once these statements are transformed into assertions concerning P^x and Q^x it is found that the identical assertions for P^x_δ and Q^x_δ can easily be established directly, and the theorems for P^x and Q^x then follow by a weak convergence argument just as above. The argument for the discrete analogue of part (ii) of Proposition 1.1 is almost identical to that discussed above for part (i), while the corresponding discrete result for
neatly Proposition 1.2 can be neatly proved using the path transformation considered in the next section. However the details are left to the reader.

To conclude this section we mention the strong version of the weak convergence theorem considered above. Recall that convergence in C means uniform convergence on compact sets.

<u>Theorem 2.13</u>. For the mapping $V_\delta : C^* \to C$ defined by (2.4), let

$$A = \{w : w \in C^*,\ \lim_{n\to\infty} wV_{1/\sqrt{n}} = w\}.$$

Then for all $x \geqq 0$

$$P^x(A) = Q^x(A) = 1.$$

Proof The result for P^x is well known (see Itô- Mckean [11], §1.10), and the assertion for Q^x follows quickly from the result for P^x put together with (2.8) for $x > 0$ and with either (2.11) or Proposition 1.2 for $x = 0$.

Since by definition $P^x_\delta = P^x V_\delta$, $Q^x_\delta = Q^x V_\delta$, the assertions of Theorem 2.6 can be deduced at once from those of Theorem 2.12 . Thus if one is prepared to take (2.8) and either (2.11) or Proposition 1.2 for granted from the above argument provides an alternative proof of Theorem 2.6.

3. Proof of Theorem 1.3.

We shall first establish the discrete analogue of part (i) of Theorem 1.3. Let $X = (X_n, n \in \mathbb{N})$ be a simple random walk on the integers Z defined on a probability triple $(\Omega, \mathcal{F}, \Pr)$. Let $M = (M_n, n \in \mathbb{N})$ be its past maximum process,

$$M_n = \max_{0 \leq k \leq n} X_k , \qquad n \in \mathbb{N}$$

and let $Y = (Y_n, n \in \mathbb{N})$ be the process defined by

$$Y_n = 2M_n - X_n, \qquad n \in \mathbb{N} .$$

The value Y_n is the reflection of X_n in the level M_n.

Lemma 3.1 Suppose that the P random walk X has starting state 1. Then Y is a Q random walk with starting state 1.

Proof Fix n and consider the conditional distribution of Y_{n+1} given any possible Y-history up to time n with $Y_n = j$, say

$$H = \{Y_0 = j_0, Y_1 = j_1, \ldots, Y_n = j\}.$$

But consider how many X-histories up to time n could have given rise to this Y-history. There are exactly j of them corresponding to the j possible values 1,2,...,j which might be taken by M_n: it is obvious that these and only these values of M_n are compatible with this H for which $Y_n = j$, and once the value of M_n is known the entire X-history up to time n can be recreated from the Y-history H by realising that

$$X_k = 2M_k - Y_k, \qquad 0 \leqq k \leqq n,$$

where M_k is determined by M_n and the Y-history through the relation

$$M_k = \min\{Y_\ell \ , \ k \leqq \ell \leqq n; \ M_n\}, \ 0 \leqq k \leqq n.$$

This relation is easily checked and is obvious from a diagram. But since X is a simple random walk all 2^n possible X-histories up to time n are equally likely, and thus the conditional probability given the Y-history H of each of the possible X-histories compatible with H is just 1/j. But if we further condition on an X-history corresponding to $M_n = j-1$ or less, it is clear that at the trasition from n to n+1 the Y-path increases (decreases) by one according as the X-path decreases (increases) by one, either event having conditional probability 1/2 given the X-history, while given the X-history with $M_n = j$ the Y path increases by one with certainty. Thus the conditional probability of $Y_{n+1} = j+1$ given H is

$$(j-1)/j \cdot \frac{1}{2} \quad + \quad 1/j \cdot \frac{1}{2} \quad = \quad (1/j)(\tfrac{1}{2})(j+1),$$

while the conditional probability of $Y_{n+1} = j-1$ given H is

$$(j-1)/j \cdot \frac{1}{2} \quad + \quad 1/j \cdot 0 \quad = \quad (1/j)(\tfrac{1}{2})(j-1) \;.$$

But this is just to say that Y is Markov with stationary transition probabilities Q.

Let us now return to the framework of Section 2. Define a map $\Phi: C \to C$ by

$$(w\Phi)(t) = 2 \sup_{0 \leq s \leq t} w(s) - w(t).$$

Thus if X now denotes the identity map from C to C then in the notation of Theorem 1.3 we have that

$$\Phi = 2M^X - X.$$

Note that the mapping Φ is continuous. Now, in the language of Section 2 Lemma 3.1 tells us by virtue of (2.5) that for any $\delta > 0$ the P^δ_δ distribution of Φ is Q^δ_δ :

$$P^\delta_\delta \Phi \quad = \quad Q^\delta_\delta \;. \tag{3.2}$$

But as δ converges to zero through the sequence $(1/\sqrt{n},\ n \in \mathbb{N})$ it is an immediate consequence of (2.6) that both $P^\delta_\delta \Rightarrow P^0$ and $Q^\delta_\delta \Rightarrow Q^0$. Thus using (2.7) we have

$$P^\delta_\delta \Phi \Rightarrow P^0, \quad Q^\delta_\delta \Rightarrow Q^0, \tag{3.3}$$

and now (3.2) yields

$$P^0 \Phi \quad = \quad Q^0 \tag{3.4}$$

which is just the assertion (i) of Theorem 1.3. This completes the proof of Theorem 1.3 since as remarked below the theorem parts (ii) and (iii) are immediate consequences of part (i).

Notes. (i) The mappings Φ and V_δ do not commute, and thus unlike the situation corresponding to any of the other connections between BM and BES(3) which we have considered it is not immediately obvious from the diffusion result (3.4) that (3.4) should have such a simple discrete analogue as (3.2). Naturally there are also discrete analogues of parts (ii) and (iii) of Theorem 1.3.

(ii) It can be shown from Theorem 1.3 that there is a continuous time analogue of the result which was the key to the proof of Lemma 3.1, namely that for a BM^0 process X with $Y = 2M^X - X$ the conditional distribution of $M^X(t)$ given $\{Y(s), 0 \leq s \leq t\}$ depends only on Y(t) and is in fact the uniform distribution on $[0,Y(t)]$. If this could be proved directly it would then be possible to establish the time-homogeneous Markov property of the Y process by using this result to show that the conditional distribution of the whole of the future of the Y process beyond t given the past history of the Y process up to time t depended only on Y(t), (though it would still be another matter to recognise the process as BES(3)).

4. Path decomposition of BES(3) and BM.

Theorem 1.3 provides a neat proof of the following path decomposition theorem for BES(3) which generalises results of Williams [18] . For a completely different approach to the extension of the present result to more general diffusions on the line see Jacobsen's paper [9] . Jacobsen's path decomposition for diffusions can be obtained from the present result by the method of random time substitution (cf. Williams [18]).

Theorem 4.1 Fix a $b \geqq 0$ and let $Y = \{Y(t),\ t \geqq 0\}$ be a $BES^b(3)$ process with future minimum process F^Y. Suppose that τ is an a.s. finite stopping time of the bivariate process $(Y, F^Y) = \{(Y(t), F^Y(t)),\ t \geqq 0\}$ such that $Y(\tau) = F^Y(\tau)$ a.s. . Then the post-τ proces

$$\{Y(\tau + t) - Y(\tau),\ t \geqq 0\}$$

is a $BES^0(3)$ process which is independent of the pre-τ process

$$\{Y(t),\ 0 \leqq t \leqq \tau\}.$$

Remark. The prime examples of random times τ which satisfy the conditions of the theorem are

(i) $$\tau = \tau_{min} = \inf\{t: Y(t) = F^Y(t)\},$$

that is to say the first time at which the process attains its overall minimum,(and the theorem shows that this is a.s. the only time at which the minimum is attained since the probability of a $BES^0(3)$ process ever returning to the origin is zero).

(ii) For $c > b$,

$$\tau = \sigma_c' = \inf\{t: Y(t) = F^Y(t) = c\}$$

$$= \sigma_c = \sup\{t: Y(t) = c\} \quad \text{a.s.},$$

(again using the theorem to see that $\sigma_c' = \sigma_c$ a.s.).

It was in these cases that the theorem was established by Williams, and for these times τ Williams also gave descriptions of the pre-τ processes (these will be mentioned later).

<u>Proof</u>. It is only necessary to prove the result for a $BES^0(3)$ process since for a $BES^b(3)$ process the result follows at once from the case $b = 0$ by considering a $BES^0(3)$ process after the time τ_b when it first hits b and using the strong Markov property at b. Suppose therefore that Y is a $BES^0(3)$ process, and let $X = 2F^Y - Y$ be the BM^0 process associated with Y by part (ii) of Theorem 1.3. Now since $Y = 2M^X - X$, $F^Y = M^X$, we see that for each $u \geqq 0$ the σ-fields in the underlying probability space generated by

$$\{(Y(t),F^Y(t)),\ 0 \leqq t \leqq u\} \text{ and } \{X(t),\ 0 \leqq t \leqq u\}$$

must be identical. Now we are assuming that τ is a stopping time of the (Y,F^Y) process, and thus τ is equally a stopping time of the X process. But by the strong Markov property of X at τ the process

$$X^* = \{X(\tau + t) - X(\tau),\ 0 \leqq t < \infty\}$$

is a BM^0 independent of $\{X(t),\ 0 \leqq t \leqq \tau\}$, and hence of $\{Y(t),\ 0 \leqq t \leqq \tau\}$. Thus using now the assumption that $Y(\tau) = F^Y(\tau)$ a.s. we have

$$X(\tau) = M^X(\tau) = Y(\tau) \text{ a.s. },$$

so that a.s.

$$\begin{aligned} Y(\tau + t) - Y(\tau) &= [2M^X(\tau + t) - X(\tau + t)] - X(\tau) \\ &= 2X(\tau) + 2M^{X^*}(t) - X(\tau + t) - X(\tau) \\ &= 2M^{X^*}(t) - X^*(t) \end{aligned}$$

and the assertion now follows from part (i) of Theorem 1.3.

For $\tau = \tau_{min}$ and $\tau = \tau_c$ Williams showed in [18] that the pre-τ process also has a simple structure which may be described as follows: for $\tau = \tau_{min}$ we have for a $BES^b(3)$ process Y that

$$\{Y(t),\ 0 \leqq t \leqq \tau_{min}\} \sim \{X(t),\ 0 \leqq t \leqq \tau_\gamma^X\}$$

where X is a BM^b process, γ is a uniform [0,b] random variable independent of X and

$$\tau_\gamma^X = \inf\ \{t: X(t) = \gamma\},$$

(it is easy to see how this can be obtained from the argument above); for $\tau = \sigma_c$ and a $BES^0(3)$ process Y we have already stated Williams' description of the pre-σ_c process in terms of reversed Brownian motion in Proposition 1.2, and for a $BES^b(3)$ process with $0 < b \leqq c$ the pre σ_c process can easily be pieced together from the pre-τ_{min} process, the post-τ_{min} pre-σ_b process and the post-σ_b pre-σ_c process using these results above and the path decompositions guaranteed by Theorem 4.1. Finally, the reader may like to put the above result for $c = b$ together with Proposition 1.2 and (2.11) or (2.12) to prove the remarkable decomposition of the Brownian path described by Williams in Theorem 2 of [17].

Acknowledgement. I would like to thank both David Williams and Martin Jacobsen for stimulating discussions on the subject of this paper.

References.

[1] BILLINGSLEY, P.: Convergence of probability measures. New York-London. Wiley 1968.

[2] BREIMAN, L.: Probability. Reading, Mass.-London-Ontario. Addison Wesley 1968.

[3] DOOB, J.L.: 'Conditional Brownian motion and the boundary limits of harmonic functions', Bull. Soc. Math. France 85, 431-458 (1957).

[4] DYNKIN, E.B.: Markov processes (English translation in two volumes). Berlin-Heidelberg-New York. Springer 1965.

[5] FELLER, W.: An introduction to probability theory and its applications Vol. 1, 3rd Ed. New York-London. Wiley 1968.

[6] ———— : An introduction to probability theory and its applications Vol. 2, 2nd Ed. New York-London. Wiley 1971.

[7] FREEDMAN, D.: Markov chains. San Francisco-Cambridge-London-Amsterdam. Holden-Day 1971.

[8] ———— : Brownian motion and diffusion. San Francisco-Cambridge-London-Amsterdam. Holden-Day 1971.

[9] JACOBSEN, M.: 'Splitting times for Markov processes and a generalised Markov property for diffusions'. To appear in Z. Wahrscheinlichkeitstheorie verw. Gebiete.

[10] KNIGHT, F.B.: 'Brownian local times and taboo processes', Trans.Amer. Math. Soc. 143, 173-185 (1969).

[11] ITÔ, K., McKEAN, H.P.: Diffusion processes and their sample paths. Berlin-Heidelberg-New York. Springer 1965.