J. Math. Kyoto Univ. (JMKYAZ)
19-1 (1979) 91-103

Strong approximation of diffusion processes by transport processes

By

Luis G. GOROSTIZA* and Richard J. GRIEGO**

(Communicated by Prof. S. Watanabe, Dec. 16, 1977)

1. Introduction.

It is well known that under certain conditions neutron transport phenomena, represented by linear transport processes (i.e., with piecewise linear paths), can be approximated by diffusion-type phenomena (see, e.g. [2]). These conditions require that the mean free path goes to zero, which means essentially that the collisions between the particles undergoing transport occur more rapidly. The result is that we have diffusions as limits of linear transport processes. Conversely, the limiting diffusions may be viewed as being approximated by the transport processes, and this leads naturally to the question of whether general diffusions can be approximated by linear transport processes. In this paper we give an affirmative answer to this question, for one-dimensional diffusions.

In [7] it was shown that one-dimensional Brownian motion can be approximated in a strong sense by a sequence of linear transport processes $\{x_n\}$ of a particularly simple type, namely x_n represents the motion of a particle on the real line changing velocities between n and $-n$ at the ends of successive random time intervals that are independent and exponentially distributed with parameter n^2. Here we extend the result of [7] to general diffusions; this extension is not immediate, because any strong approximation must depend on the local behavior of the diffusion, which in the case of Brownian motion is always the same.

It is of interest to mention other results related to approximations of diffusions. Kac [8] noted that the transport process x_n described above is associated with the telegraph equation $n^{-2}u_{tt}+2u_t=u_{xx}$. As $n\to\infty$, the equations converge to the diffusion equation for Brownian motion, $u_t=u_{xx}/2$, and hence one might expect that the x_n converge in some way to Brownian motion. Convergence in distribution could probably be obtained by Kac's methods, although this was not explicitly noted (in [7] it is shown that the convergence is pathwise, uniformly on compact intervals, with probability one). Other weak approximations of

Centro de Investigación del I. P. N., Mexico and The University of New Mexico., U. S. A.
* Research supported in part by CONACYT grant PNCB 000063.
** Research supported in part by NIH Grant No. RR-08139.

certain diffusions are obtained as special cases of results of Pinsky [11], Watanabe [17, 18], and Gorostiza [6]. Weak convergence of general classes of transport processes in several dimensions to certain classes of diffusions has been obtained by Papanicolaou [10] (in contrast with [10], we are interested in approximating general diffusions by simple transport processes). There are other strong approximations in one dimension, e.g. Knight [9] (random walks to Brownian motion), Quiring [12] (jump processes to diffusion), and Stone [14] (continuous dependence of diffusions on the speed and killing measures).

Besides the purely mathematical question of approximating general diffusions by transport processes, our motivation for this work was to develop a method of simulating by computer, transport processes that are of interest in applied sciences (e.g. neutron transport, seismic motion). Clearly, for applications, rate of convergence results are needed, and we hope to obtain such results in future work.

The proof of our approximation, once one has some insight into how to construct the transport process, can be done using standard techniques.

2. Results.

We will consider regular diffusions on $(-\infty, \infty)$, starting from 0 at time 0, and on natural scale. The term "transport process" shall mean a process with piecewise linear paths, in which the time intervals corresponding to the linear pieces are exponentially distributed, all with the same parameter, but the slopes may be random. The purpose of this paper is to prove the following result.

Theorem. *Given a diffusion Y, there exist transport processes Z_n, $n=1, 2, \cdots$, on the same space as Y, such that $Z_n \to Y$ uniformly on compact time intervals as $n \to \infty$, with probability one.*

Suppose the diffusion Y has absolutely continuous speed measure with density $\dot{m}$. Then we can give an explicit construction of Z_n: The consecutive slopes are

$$\beta_1^n = k_1 n [2/\dot{m}(0)]^{1/2}$$

$$\beta_i^n = k_i n \left[2/\dot{m}\left(\sum_{j=1}^{i-1} k_j \xi_j^n\right)\right]^{1/2}, \quad i>1,$$

where $k_1, k_2, \cdots$, are independent and take the values 1 and -1 with probability 1/2 each, ξ_1^n is exponentially distributed with parameter $n[2\dot{m}(0)]^{1/2}$; ξ_j^n, $j>1$, is exponentially distributed with (random) parameter

$$n\left[2\dot{m}\left(\sum_{i=1}^{j-1} k_i \xi_i^n\right)\right]^{1/2},$$

and independent of k_j; the time durations of the linear pieces are independent and exponentially distributed with parameter $2n^2$, and independent of the slopes.

The construction of Z_n above may have several consecutive up-going (or down-going) linear pieces. A simpler transport approximation $\tilde{Z}_n$ can be obtained from Z_n by interpolating linearly each run of up-going (or down-going) pieces. Then, in $\tilde{Z}_n$ the positive and negative slopes follow one another in succession, and their durations are independent and exponentially distributed with parameter n^2. Let N be the first time there is a change of sign in the slopes β_i^n, $i=1,2,\cdots$, of Z_n, and let L_n be the absolute slope corresponding to the interpolation of the first N pieces of Z_n; i.e., L_n is the first absolute slope of $\tilde{Z}_n$. The expected value and second moment of L_n are related to arithmetic means of the β_i^n as follows:

$$EL_n=E\frac{1}{N}\sum_{i=1}^{N}|\beta_i^n|, \quad \text{and} \quad EL_n^2=E\frac{2}{N(N+1)}\sum_{i=1}^{N}|\beta_i^n|\sum_{j=1}^{i}|\beta_j^n|.$$

Similarly for the next slopes.

In the case of Brownian motion, $\dot{m}\equiv 2$, $\tilde{Z}_n$ has slopes $\pm n$, and it coincides with the approximation of [7].

From the construction of Z_n it is clear that each linear piece depends only on the past of Z_n and new independent random variables; moreover, Z_n is produced without any recourse to the diffusion paths; this is because the transport paths are not obtained in a deterministic way from diffusion paths; the diffusion is used only in a probabilistic way, via the Skorohod embedding, to represent certain random quantities. This aspect of the transport process is convenient for computer generation of approximate diffusion paths, and approximate evaluation of stochastic integrals by the method of Wong and Zakai [19].

3. Proofs.

The theorem is proved in two parts: Lemma 1 shows that general diffusions can be approximated by diffusions of a special type, and Lemma 2 shows that certain diffusions, including those of the special type in Lemma 1, can be approximated by transport processes.

We will use the following few facts about diffusions (see e.g. [1], [4]). On a probability space $(\Omega, \mathcal{F}, P)$, where Ω is the set of continuous real functions on $[0, \infty)$, let Y be a regular diffusion on $(-\infty, \infty)$, starting from 0 at time 0, on natural scale, with speed measure m. There is a standard Brownian motion X on Ω such that $Y(t, w)=X(T(t, w), w)$, $t\geqq 0$, $w\in\Omega$, where the random time transformation $T(t, w)$ is the inverse of $M(t, w)=\frac{1}{2}\int_{-\infty}^{\infty}\lambda(x, t, w)\,m(dx)$, and λ is Brownian local time. $\lambda(x, t, w)$ is continuous in (x, t), and $M(\cdot, w)$ and $T(\cdot, w)$ are continuous and strictly increasing to ∞, with probability one. If m has density $\dot{m}$ (with respect to Lebesgue measure), $\dot{m}$ is strictly positive, and $M(t)=\frac{1}{2}\int_0^t \dot{m}(X(s))\,ds$. A Borel measure on $(-\infty, \infty)$ is a diffusion speed measure if and only if it is strictly positive (positive on nonempty open sets), and locally finite (finite on compact sets).

Lemma 1. *Given a diffusion Y, there exist diffusions Y_n, $n=1, 2, \cdots$, on the same space as Y, with speed measures that have densities infinitely differentiable with compact supports, such that $Y_n \to Y$ uniformly on finite intervals as $n \to \infty$, with probability one.*

Proof. Stone's general Theorem 1 [14] can be used in part here, but we give an easy argument for completeness. Let m_n, $n=1, 2, \cdots$, be strictly positive and locally finite Borel measures on $(-\infty, \infty)$, and denote m the speed measure of Y. Suppose $m_n \to m$ weakly, in the sense that

$$\int_{-\infty}^{\infty} f(x)\, m_n(dx) \to \int_{-\infty}^{\infty} f(x)\, m(dx) \qquad \text{as} \quad n \to \infty,$$

for all real functions f on $(-\infty, \infty)$, continuous with compact support. Then for each $w \in \Omega$,

$$M_n(t, w) = \frac{1}{2}\int_{-\infty}^{\infty} \lambda(x, t, w)\, m_n(dx) \to \frac{1}{2}\int_{-\infty}^{\infty} \lambda(x, t, w)\, m(dx) = M(t, w),$$

because $\lambda(\cdot, t, w)$ is continuous and vanishes outside a finite interval. Consequently,

$$T_n(t, w) = M_n^{-1}(t, w) \to M^{-1}(t, w) = T(t, w),$$

and therefore

$$Y_n(t, w) = X(T_n(t, w), w) \to X(T(t, w), w) = Y(t, w).$$

Y_n is a diffusion on the same space as Y, with speed measure m_n, and $Y_n \to Y$ uniformly on finite intervals a. s., because X has continuous paths, hence uniformly continuous on finite intervals, and because $T_n \to T$ pointwise and the fact that T_n and T are strictly increasing a. s. implies that $T_n \to T$ uniformly on $[0, \infty)$ a. s.

Therefore we have to construct measures m_n such that $m_n \to m$ weakly, and m_n has density $\dot{m}_n$ infinitely differentiable with compact support. Let $m|_{[-n, n]}$ denote the restriction of m to the interval $[-n, n]$ and let

$$\dot{m}_n(x) = \int_{-\infty}^{\infty} \varphi_n(x-y)\, m|_{[-n, n]}(dy), \quad -\infty < x < \infty,$$

where $\{\varphi_n\}$ is a delta-function sequence, infinitely differentiable with compact supports that do not increase with n. Then $\dot{m}_n$ is infinitely differentiable (see [16], p. 289), and since the function φ_n and the measure $m|_{[-n, n]}$ have compact supports, so does $\dot{m}_n$. The measure

$$m_n(dx) = \dot{m}_n(x)\, dx$$

is strictly positive and locally finite. We now show that $m_n \to m$ weakly. Let f be a real continuous function on $(-\infty, \infty)$, with compact support, then

$$\int_{-\infty}^{\infty} f(x)\, m_n(dx) = \int_{-\infty}^{\infty} f(x)\, \dot{m}_n(x)\, dx$$

$$= \int_{-\infty}^{\infty} f(x) \Big(\int_{-\infty}^{\infty} \varphi_n(x-y)\, m|_{[-n,n]}(dy) \Big) dx$$

$$= \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} f(x) \varphi_n(x-y)\, dx \Big) m|_{[-n,n]}(dy).$$

Since f and all φ_n have supports contained in a bounded set, then for n sufficiently large

$$\int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} f(x) \varphi_n(x-y)\, dx \Big) m|_{[-n,n]}(dy) = \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} f(x) \varphi_n(x-y)\, dx \Big) m(dy).$$

Since $\{\varphi_n\}$ is a delta-function sequence,

$$\int_{-\infty}^{\infty} f(x) \varphi_n(x-y)\, dx \to f(y),$$

and again because f and all φ_n have supports in a bounded set, the functions $\int_{-\infty}^{\infty} f(x) \varphi_n(x-\cdot)\, dx$ are uniformly bounded by a (finite) function with compact support; so, finally, since m is finite on compact sets, by the dominated convergence theorem

$$\int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} f(x) \varphi_n(x-y)\, dx \Big) m(dy) \longrightarrow \int_{-\infty}^{\infty} f(y)\, m(dy),$$

which finishes the proof.

Lemma 2. *Given a diffusion Y with speed measure that has a twice continuously differentiable density, there exist transport processes Z_n, $n=1, 2, \cdots$, on the same space as Y, such that $Z_n \to Y$ uniformly on compact intervals as $n \to \infty$, with probability one.*

Proof. We prove the lemma for the time interval $[0, 1]$.

Let $\dot{m}$ denote the speed measure density of Y. We will suppose first that $\dot{m}$ is bounded away from zero and $\dot{m}'$, $\dot{m}''$ are bounded.

Fix a positive integer n. Let $S_0^n=0$, and S_i^n, $i=1, 2, \cdots$, be random variables non-decreasing in i, and β_i^n, $i=0, 1, \cdots$, non-zero random variables. These random variables will be appropriately chosen shortly. Let

$$\gamma_i^n = |\beta_i^n|^{-1} |Y(S_i^n) - Y(S_{i-1}^n)|, \quad i=1, 2, \cdots,$$

and

$$\Gamma_0^n = 0, \qquad \Gamma_i^n = \sum_{j=1}^{i} \gamma_j^n, \quad i=1, 2, \cdots.$$

We define the n-th transport process Z_n by

$$Z_n(\Gamma_i^n) = Y(S_i^n), \quad i=0, 1, \cdots,$$

with linear interpolation between the consecutive random times Γ_i^n, $i=0, 1, \cdots$. It is clear that in the interval $[\Gamma_{i-1}^n, \Gamma_i^n]$ the process Z_n has slope $|\beta_i^n|$ or $-|\beta_i^n|$. Using the random change of time T from the Brownian motion X, and denoting $W_i^n=T(S_i^n)$, we have

$$Z_n(\Gamma_i^n)=X(W_i^n), \quad i=0, 1, \cdots,$$

and

$$\gamma_i^n=|\beta_i^n|^{-1}|X(W_i^n)-X(W_{i-1}^n)|, \quad i=1, 2, \cdots.$$

Also,

$$S_i^n=\sum_{j=1}^{i} s_j^n, \quad i=1, 2, \cdots,$$

where

$$s_j^n=\frac{1}{2}\int_{W_{j-1}^n}^{W_j^n} \dot{m}(X(s))\,ds, \quad j=1, 2, \cdots.$$

Now we choose the S_i^n, equivalently the W_i^n, as follows. By the Skorohod representation ([13], p. 163), we take the W_i^n so that $X(W_i^n)-X(W_{i-1}^n)$ has the same distribution as $k_i\xi_i^n$, where k_i takes the values 1 and -1, each with probability 1/2, ξ_i^n is exponentially distributed with random parameter $n[2\dot{m}(X(W_{i-1}^n))]^{1/2}$, and independent of k_i; the k_i are independent, and independent of X, but ξ_i^n depends on X via its parameter (which is determined by the Skorohod barrier that X hits first in the $(i-1)$-th representation). Observe that for each i,

$$E[k_i\xi_i^n|X(W_{i-1}^n)]=0,$$

$$\mathrm{Var}[k_i\xi_i^n|X(W_{i-1}^n)]=E[(\xi_i^n)^2|X(W_{i-1}^n)]=(n^2\dot{m}(X(W_{i-1}^n)))^{-1},$$

and hence

$$E[W_i^n-W_{i-1}^n|X(W_{i-1}^n)]=(n^2\dot{m}(X(W_{i-1}^n)))^{-1}.$$

Notice that $X(W_i^n)-X(W_{i-1}^n)$, $i=1, 2, \cdots$, are not independent in general.

Now we choose the β_i^n. Let

$$\beta_i^n=k_i n[2/\dot{m}(X(W_{i-1}^n))]^{1/2}, \quad i=1, 2, \cdots.$$

Then γ_i^n, which is distributed as $\xi_i^n|\beta_i^n|^{-1}$, is exponentially distributed with parameter $2n^2$ conditioned on $X(W_{i-1}^n)$, and it is easy to verify that γ_i^n, $i=1, 2, \cdots$, are independent.

Observe that in the interval $[\Gamma_{i-1}^n, \Gamma_i^n]$ the process Z_n has slope β_i^n.

With this setting, we can state the basic idea of the proof. For each n, consider the random variables

$$A_n=\max_{1\le i\le 2n^2} |\Gamma_i^n-i/2n^2|,$$

and

$$B_n=\max_{1\le i\le 2n^2} |S_i^n-i/2n^2|.$$

If we show that $A_n \to 0$ and $B_n \to 0$ as $n \to \infty$, with probability one, then, due to the way in which Z_n is defined, and to the uniform continuity of diffusion paths on compact intervals, it will follow that

$$\lim_{n\to\infty} \max_{0\le t\le 1} |Z_n(t)-Y(t)|=0 \text{ a.s.},$$

thus proving the theorem.

Proof that $A_n \to 0$ a.s.

Let $\mathcal{F}_i^n$ denote the σ-algebra generated by $X(W_j^n)$, $j\leqq i$. Then Γ_i^n is $\mathcal{F}_i^n$-measurable, and since $\Gamma_i^n=\Gamma_{i-1}^n+\gamma_i^n$, and γ_i^n is exponential with parameter $2n^2$ conditional on $\mathcal{F}_{i-1}^n$, we have

$$E[\Gamma_i^n \mid \mathcal{F}_{i-1}^n]=\Gamma_{i-1}^n+1/2n^2.$$

Therefore $\{\Gamma_i^n-i/2n^2, \mathcal{F}_i^n\}_i$ is a martingale, and by the martingale inequality (see [3]), for $\varepsilon>0$ we have

$$\begin{aligned} P[A_n>\varepsilon] &\leqq \varepsilon^{-2}E(\Gamma_{2n^2}^n-1)^2 \\ &=\varepsilon^{-2}\Big\{\sum_{i=1}^{2n^2} E(\gamma_i^n-1/2n^2)^2 \\ &\quad +2\sum_{i=1}^{2n^2-1}\sum_{j=i+1}^{2n^2} E(\gamma_i^n-1/2n^2)(\gamma_j^n-1/2n^2)\Big\}, \end{aligned}$$

but

$$\begin{aligned} E(\gamma_i^n-1/2n^2)^2 &= EE[(\gamma_i^n-1/2n^2)^2 \mid X(W_{i-1}^n)] \\ &=1/(2n^2)^2=1/4n^4, \end{aligned}$$

and since γ_i^n and γ_j^n are independent for $i<j$,

$$E(\gamma_i^n-1/2n^2)(\gamma_j^n-1/2n^2)=0,$$

hence

$$P[A_n>\varepsilon]\leqq 1/\varepsilon^2 2n^2,$$

and the Borel-Cantelli Lemma implies that $A_n \to 0$ a.s.

Proof that $B_n \to 0$ a.s.

First we will write S_i^n in a different way. Let

$$f(x)=\int_0^x\Big(\int_0^y \dot{m}(z)\,dz\Big)dy, \quad -\infty<x<\infty.$$

Then $f''(x)=\dot{m}(x)$, and Ito's formula ([5](p. 29) gives

$$\frac{1}{2}\dot{m}(X(s))\,ds=df(X(s))-\Big(\int_0^{X(s)}\dot{m}(x)\,dx\Big)dX(s),$$

where $dX(s)$ is the Ito differential; hence

$$s_j^n=\frac{1}{2}\int_{W_{j-1}^n}^{W_j^n}\dot{m}(X(s))\,ds$$

$$=\int_{X(W_{j-1}^n)}^{X(W_j^k)}\left(\int_0^y\dot{m}(x)\,dx\right)dy-\int_{W_{j-1}^n}^{W_j^n}\left(\int_0^{X(s)}\dot{m}(x)\,dx\right)dX(s)$$

$$=\int_{X(W_{j-1}^n)}^{X(W_j^n)}\left(\int_{X(W_{j-1}^n)}^y\dot{m}(x)\,dx\right)dy-\int_{W_{j-1}^n}^{W_j^n}\left(\int_{X(W_{j-1}^n)}^{X(s)}\dot{m}(x)\,dx\right)dX(s)$$

(we subtracted and added $\int_0^{X(W_{j-1}^n)}\dot{m}(x)\,dx\,[X(W_j^n)-X(W_{j-1}^n)]$). Let

$$c_j^n=\int_{X(W_{j-1}^n)}^{X(W_j^n)}\left(\int_{X(W_{j-1}^n)}^y\dot{m}(x)\,dx\right)dy,\quad j=1,2,\cdots,$$

$$d_j^n=\int_{W_{j-1}^n}^{W_j^n}\left(\int_{X(W_{j-1}^n)}^{X(s)}\dot{m}(x)\,dx\right)dX(s),\quad j=1,2\cdots,$$

and

$$C_i^n=\sum_{j=1}^i c_j^n,\quad D_i^n=\sum_{j=1}^i d_j^n,\quad i=1,2,\cdots.$$

So

$$S_i^n=C_i^n-D_i^n,\quad i=1,2,\cdots.$$

We will show that

$$\max_{1\le i\le 2n^2}|C_i^n-i/2n^2|\longrightarrow 0\quad \text{a.s.}$$

and

$$\max_{1\le i\le 2n^2}|D_i^n|\longrightarrow 0\quad \text{a.s.},$$

which implies the desired result.

Using Taylor's theorem,

$$c_j^n=\frac{1}{2}\dot{m}(X(W_{j-1}^n))[X(W_j^n)-X(W_{j-1}^n)]^2$$

$$+\frac{1}{6}\dot{m}'(X(W_{j-1}^n))[X(W_j^n)-X(W_{j-1}^n)]^3$$

$$+\frac{1}{24}\dot{m}''(q_j)[X(W_j^n)-X(W_{j-1}^n)]^4,$$

where q_j is an appropriate point in an appropriate interval. Let

$$g_j^n=\frac{1}{2}\dot{m}(X(W_{j-1}^n))[X(W_j^n)-X(W_{j-1}^n)]^2$$

$$+\frac{1}{6}\dot{m}'(X(W_{j-1}^n))[X(W_j^n)-X(W_{j-1}^n)]^3,\quad j=1,2,\cdots,$$

$$h_j^n=\frac{1}{24}\dot{m}''(q_j)[(X(W_j^n)-X(W_{j-1}^n)]^4, \quad j=1, 2, \cdots,$$

and

$$G_i^n=\sum_{j=1}^{i} g_j^n, \quad H_i^n=\sum_{j=1}^{i} h_j^n, \quad i=1, 2, \cdots.$$

So

$$C_i^n=G_i^n+H_i^n, \quad i=1, 2, \cdots,$$

and we will show that

$$\max_{1\leq i\leq 2n^2} |G_i^n-i/2n^2| \longrightarrow 0 \quad \text{a. s.}$$

and

$$\max_{1\leq i\leq 2n^2} |H_i^n| \longrightarrow 0 \quad \text{a. s.}$$

To work with G_i^n, again denoting $\mathcal{F}_j^n$ the σ-algebra generated by $X(W_k^n)$, $k\leqq j$, we compute first

$$E[g_j^n|\mathcal{F}_{j-1}^n]=\frac{1}{2}\dot{m}(X(W_{j-1}^n))E[(X(W_j^n)-X(W_{j-1}^n))^2|\mathcal{F}_{j-1}^n]$$

$$+\frac{1}{6}\dot{m}'(X(W_{j-1}^n))E[(X(W_j^n)-X(W_{j-1}^n))^3|\mathcal{F}_{j-1}^n]$$

$$=\frac{1}{2}\dot{m}(X(W_{j-1}^n))/n^2\dot{m}(X(W_{j-1}^n))+0=1/2n^2;$$

hence also $Eg_j^n=1/2n^2$, and

$$E(g_j^n-1/2n^2)^2=E(g_j^n)^2-Eg_j^n/n^2+1/4n^4$$
$$=E(g_j^n)^2-1/4n^4.$$

Now,

$$E(g_j^n)^2=\frac{1}{4}E\{\dot{m}(X(W_{j-1}^n))^2E[(X(W_j^n)-X(W_{j-1}^n))^4|X(W_{j-1}^n)]\}$$

$$+\frac{1}{36}E\{\dot{m}'(X(W_{j-1}^n))^2E[(X(W_j^n)-X(W_{j-1}^n))^6|X(W_{j-1}^n)]\}$$

$$+\frac{1}{6}E\{\dot{m}(X(W_{j-1}^n))\dot{m}'(X(W_{j-1}^n))E[(X(W_j^n)-X(W_{j-1}^n))^5|X(W_{j-1}^n)]\}$$

$$=\frac{1}{4}E\{\dot{m}(X(W_{j-1}^n))^2 4!/n^4[2\dot{m}(X(W_{j-1}^n))]^2\}$$

$$+\frac{1}{36}E\{\dot{m}'(X(W_{j-1}^n))^2 6!/n^6[2\dot{m}(X(W_{j-1}^n))]^3\}+0,$$

and since $\dot{m}$ is bounded away from zero and $\dot{m}'$ is bounded, then $E(g_j^n-1/2n^2)^2\leqq Kn^{-4}$, for all j, where K is a constant.

The above computations show that $\{G_i^n - i/2n^2, \mathcal{F}_i^n\}_i$ is a martingale, and therefore by the martingale inequality, using $E[g_j^n - 1/2n^2 | \mathcal{F}_{j-1}^n] = 0$, for $\varepsilon > 0$ we have

$$P[\max_{1 \le i \le 2n^2} |G_i^n - i/2n^2| > \varepsilon] \le \varepsilon^{-2} \sum_{i=1}^{2n^2} E(g_i^n - 1/2n^2)^2 \le 2K/\varepsilon^2 n^2,$$

and by the Borel-Cantelli Lemma we obtain the desired result for G_i^n.

For H_i^n, using Chebyshev's inequality we have, for $\varepsilon > 0$,

$$P[\max_{1 \le i \le 2n^2} |H_i^n| > \varepsilon] \le P\left[\sum_{i=1}^{2n^2} |h_i^n| > \varepsilon\right]$$

$$\le \varepsilon^{-1} \sum_{i=1}^{2n^2} \frac{1}{24} E\{|\dot{m}''(q_j)| (X(W_j^n) - X(W_{j-1}^n))^4\},$$

but $\dot{m}''$ is bounded, and

$$E(X(W_j^n) - X(W_{j-1}^n))^4 = 4!\, n^{-4} E(2\dot{m}(X(W_{j-1}^n)))^{-2},$$

and $\dot{m}$ is bounded away from zero, so

$$P[\max_{1 \le i \le 2n^2} |H_i^n| > \varepsilon] \le K n^{-2},$$

where K is a constant. Again the Borel-Cantelli Lemma yields the desired result for H_i^n.

It remains to show that $\max_{1 \le i \le 2n^2} |D_i^n| \to 0$ a.s. Let $\mathcal{D}_i^n$ denote the σ-algebra generated by $X(t)$, $t \le W_i^n$. Then $\{D_i^n, \mathcal{D}_i^n\}_i$ is a martingale, because $E[d_i^n | \mathcal{D}_{i-1}^n] = 0$ (see [5], p. 30). Hence, by the martingale inequality, using $E[d_i^n | \mathcal{D}_{i-1}^n] = 0$, for $\varepsilon > 0$ we have

$$P[\max_{1 \le i \le 2n^2} |D_i^n| > \varepsilon] \le \varepsilon^{-2} \sum_{i=1}^{2n^2} E(d_i^n)^2.$$

Now (see [5], p. 28),

$$E(d_i^n)^2 = E\int_{W_{j-1}^n}^{W_j^n} \left(\int_{X(W_{j-1}^n)}^{X(s)} \dot{m}(x)\,dx\right)^2 ds$$

$$\le KE\{(W_j^n - W_{j-1}^n)(X(W_j^n)^+ - X(W_j^n)^-)^2\},$$

where K is a constant, because $\dot{m}$ is bounded, and $X(W_j^n)^-$ and $X(W_j^n)^+$ are respectively the lower and upper Skorohod barriers corresponding to the j-th representation. By the Schwarz inequality,

$$E\{(W_j^n - W_{j-1}^n)(X(W_j^n)^+ - X(W_j^n)^-)^2\}$$

$$\le [E(W_j^n - W_{j-1}^n)^2]^{1/2} [E(X(W_j^n)^+ - X(W_j^n)^-)^4]^{1/2},$$

but, since we are on the interval [0, 1],

$$E(X(W_j^n)^+-X(W_j^n)^-)^4\leqq E(\sup_{0\leq t\leq 1} X(t)-\inf_{0\leq t\leq 1} X(t))^4<\infty,$$

and (see [13], p. 163), there is a constant M such that

$$\begin{aligned} E(W_j^n-W_{j-1}^n)^2 &= EE[(W_j^n-W_{j-1}^n)^2\,|\,X(W_{j-1}^n)] \\ &\leqq ME[(\xi_j^n)^4\,|\,X(W_{j-1}^n)]=M4!\,n^{-4}E(2\dot{m}(X(W_{j-1}^n)))^{-2} \\ &\leqq Nn^{-4}, \end{aligned}$$

where N is a constant, the last inequality because $\dot{m}$ is bounded away from zero. In conclusion,

$$P[\max_{1\leq i\leq 2n^2}|D_i^n|>\varepsilon]\leqq Ln^{-2},$$

where L is a constant, and once more the Borel-Cantelli Lemma gives the desired result.

We have proved the lemma under the assumption that $\dot{m}$ is bounded away from zero and $\dot{m}'$, $\dot{m}''$ are bounded. Now we drop this assumption, so Y is now a diffusion with speed measure density $\dot{m}$. Observe that $\dot{m}$ is strictly positive on $(-\infty, \infty)$. Define a new diffusion Y_n on the same space as Y, with speed measure density given by

$$\dot{m}_n(x)=\begin{cases} \dot{m}(x), & x\in[-n^2, n^2], \\ \dot{m}(-n^2), & x<-n^2, \\ \dot{m}(n^2), & x>n^2, \end{cases}$$

except if necessary a little to the left of $-n^2$, and a little to the right of n^2, so as to make $\dot{m}_n$ twice continuously differentiable and bounded away from zero on $(-\infty, \infty)$. By the first part of the lemma, Y_n can be approximated uniformly a.s. by transport processes $Z_{n,k}$, $k=1, 2, \cdots$. On the other hand, the construction of the approximating transport processes in the first part of the lemma can also be carried out for Y, even if $\dot{m}$ is not bounded away from zero; let $\tilde{Z}_k$, $k=1, 2, \cdots$, be these transport processes. Because of their construction from the same Brownian motion X, the processes $Z_{n,k}$ and $\tilde{Z}_k$ will differ on the time interval $[0, 1]$ only if X leaves the interval $[-n^2, n^2]$ on $[0, 1]$. Then, using Chebyshev's inequality,

$$P[\sup_{0\leq t\leq 1}|Z_{n,k}(t)-\tilde{Z}_k(t)|>0]\leqq P[\sup_{0\leq t\leq 1}|X(t)|>n^2]\leqq n^{-2}K,$$

where $K=E\sup_{0\leq t\leq 1}|X(t)|<\infty$. Therefore, by the Borel-Cantelli Lemma, $Z_{n,k}=\tilde{Z}_k$ on $[0, 1]$ for all sufficiently large n, with probability one, for all k. Now, clearly $m_n\to m$ weakly, and therefore $Y_n\to Y$ uniformly on $[0, 1]$ a.s. (see the first part of the proof of Lemma 1). Fix $w\in\Omega$ in a set of probability one where all approximations hold, and take $\varepsilon>0$. Let n be sufficiently large so that

$\sup_{0\leq t\leq 1} |Y_n(t, w)-Y(t, w)|<\varepsilon$. Then let k be sufficiently large so that $\sup_{0\leq t\leq 1} |Z_{n,k}(t, w) -Y_n(t, w)|<\varepsilon$. Taking n larger if necessary, we then also have $\sup_{0\leq t\leq 1} |\tilde{Z}_k(t, w) -Y_n(t, w)|<\varepsilon$ for sufficiently large k. It follows that $\sup_{0\leq t\leq 1} |\tilde{Z}_k(t, w)-Y(t, w)|<2\varepsilon$ for all sufficiently large k, so $\tilde{Z}_n\to Y$ uniformly on $[0, 1]$ as $n\to\infty$, with probability one.

The proof is finished.

Remark 1. It is well-known that

$$\int_0^{X(t)}\left(\int_0^s \dot{m}(x)\,dx\right)ds-\int_0^t\left(\int_0^{X(s)}\dot{m}(x)\,dx\right)ds, \quad t\geqq 0$$

is a martingale (see [15]). This actually plays a role in the proof of Lemma 2, but not in a direct way, because we must look at s_j^n in the form $s_j^n=c_j^n-d_j^n$.

Remark 2. The probability space was actually enlarged to accommodate the new random variables brought in for the proof.

Lemmas 1 and 2 prove the theorem, because the approximating diffusions of Lemma 1 satisfy the conditions of Lemma 2.

The explicit construction of the transport approximation Z_n given right after the theorem is taken directly from the proof of Lemma 2.

We now give some indications on the proof of the statements made about the transport approximation $\tilde{Z}_n$, using the notation in Lemma 2. The time of the first change of sign in the slope of Z_n, or equivalently, the time of the first change of slope in $\tilde{Z}_n$, can be written as $\tau_1^n=\Gamma_N^n$, where N is an independent random variable that takes the value k with probability 2^{-k}, for $k=1, 2, \cdots$. Using the fact that γ_i^n is exponentially distributed with parameter $2n^2$ conditional on γ_k^n, $X(W_k^n)$, $k\leqq i-1$, it is easy to show (e.g. using characteristic functions), that τ_1^n is exponential with parameter n^2. If τ_i^n is the time of the i-th slope change in $\tilde{Z}_n$, then, similarly, $\tau_i^n-\tau_{i-1}^n$ is exponential with parameter n^2. The first absolute slope of $\tilde{Z}_n$ is

$$L_n=|\tilde{Z}_n(\tau_1^n)|/\tau_1^n=\sum_{j=1}^N \gamma_i^n|\beta_i^n|/\sum_{j=1}^N \gamma_i^n.$$

Somewhat long but straightforward calculations yield the given values of EL_n and EL_n^2 (first, condition on N, then condition on the $X(W_i^n)$, and use Calculus).

CENTRO DE INVESTIGACIÓN DEL I. P. N., MEXICO
THE UNIVERSITY OF NEW MEXICO, U. S. A.

References

[1] Breiman, L. (1968). *Probability*. Addison-Wesley, Reading, Mass.
[2] Case, K. M. and Zweifel, P. F. (1967). *Linear Transport Theory*. Addison-Wesley, Reading, Mass.
[3] Doob, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[4] Freedman, D. (1971). *Brownian Motion and Diffusion*. Holden-Day, San Francisco.

[5] Gikhman, I.I. and Skorokhod, A.V. (1968). *Stochastic Differential Equations*. Springer-Verlag, New York.

[6] Gorostiza, L.G. (1973). An invariance principle for a class of d-dimensional polygonal random functions. Trans. Amer. Math. Soc. Vol. 177, 413-445.

[7] Griego, R.J., Heath, D. and Ruiz-Moncayo, A. (1971). Almost sure convergence of uniform transport processes to Brownian motion. Ann. Math. Stat., Vol. 42, No. 3, 1129-1131.

[8] Kac, M. (1956). *Some Stochastic Problems in Physics and Mathematics*. Magnolia Petroleum Co. Lectures in pure and applied science. No. 2. Reprinted in part as: (1974). A stochastic model related to the telegrapher's equation. Rocky Mtn. J. Math., Vol. 4, No. 3, 497-509.

[9] Knight, F.B. (1962). On the random walk and Brownian motion. Trans. Amer. Math. Soc., Vol. 103, 218-228.

[10] Papanicolaou, G. (1975). Asymptotic analysis of transport processes. Bull. Amer. Math. Soc., Vol. 81, No. 2, 330-392.

[11] Pinsky, M. (1968). Differential equations with a small parameter and the central limit theorem for functions defined on a finite Markov chain. Z. Wahrschein. und Verw. Geb., Vol. 9, 101-111.

[12] Quiring, D. (1972). Random evolutions on diffusion processes. Z. Wahrschein. und Verw. Geb., Vol. 23, 230-244.

[13] Skorokhod, A.V. (1965). *Studies in the Theory of Random Processes*. Addison-Wesley, Reading, Mass.

[14] Stone, C.J. (1963). Limit theorems for random walks, birth and death processes, and diffusion processes. Illinois J. Math., Vol. 7, 638-660.

[15] Stroock, D.W. and Varadhan, S.R.S. (1969). Diffusion processes with continuous coefficients: I, II. Comm. Pure Appl. Math. II, 22, 345-400; 479-530.

[16] Treves, F. (1967). *Topological Vector Spaces, Distributions, and Kernels*. Academic Press, New York.

[17] Watanabe, T. (1968). Approximation of uniform transport process on a finite interval to Brownian motion. Nagoya Math. J., Vol. 32, 297-314.

[18] Watanabe, T. (1969). Convergence of transport process to diffusion. Proc. Japan Acad., Vol. 45, 470-472.

[19] Wong, E. and Zakai, M. (1969). Riemann-Stieltjes approximations of stochastic integrals. Zeit. Wahrs., Vol. 12, 87-97.