Probab. Th. Rel. Fields 72, 559–587 (1986)

Probability Theory and Related Fields

Self-Intersections of 1-Dimensional Random Walks

David J. Aldous*

Department of Statistics, University of California, Berkeley, CA 94720, USA

Summary. Consider a random walk S_n on the integers, where the steps ξ_i have mean 0 and variance σ^2. Let T be the time of first self-intersection of the random walk. It is shown that, as $\sigma \to \infty$, T grows at rate $\sigma^{2/3}$. More precisely, $T/\sigma^{2/3}$ has a non-degenerate limit distribution which can be described in terms of Brownian motion local time.

1. Introduction and Summary

For an integer $K > 1$ consider the random walk on the integers

$$S_0 = 0; \qquad S_n = \sum_{i=1}^{n} \xi_i^{(K)}$$

where $(\xi_i^{(K)})$ are independent and uniform on the integers $\{-K, -K+1, \ldots, K-1, K\}$. Let $T_1^{(K)}$ be the time of the first self-intersection of this random walk:

$$T_1^{(K)} = \min\{n\colon\ S_n = S_m \text{ for some } 0 \leqq m < n\}. \tag{1.1}$$

Certainly $ET_1^{(K)}$ is finite, for $T_1^{(K)} \leqq \min\{n\colon\ \xi_n = 0\}$ implies $ET_1^{(K)} \leqq 2K+1$. What is the behavior of $T_1^{(K)}$ as $K \to \infty$? Pollard (1979) raised this problem in the context of a computer algorithm, and suggested $ET_1^{(K)} \sim c K^{2/3}$ for some constant c. Shepp and Steele (unpublished) have shown that $ET_1^{(K)}$ is asymptotically in an interval $(c_1 K^{2/3}, c_2 K^{2/3})$ for certain constants $0 < c_1 < c_2 < \infty$. Our main result, Theorem 1.8, implies that the limit

$$c = \lim_{K \to \infty} K^{-2/3} ET_1^{(K)} \tag{1.2}$$

exists, and that $c \approx 0.99$.

* Research supported by National Science Foundation Grant MCS 80-02698.

It is easy to see informally why 2/3 is the correct exponent. The *mean* number $m(K,n)$ of self-intersections up until time n is

$$m(K,n)=\sum_{0\leqq i<j\leqq n}\sum P(S_i=S_j)=\sum_{1\leqq i\leqq n}(n+1-i)\,P(S_i=0). \tag{1.3}$$

Let $\sigma_K^2=\operatorname{var}(\xi_1^{(k)})\sim 3^{-1}K^2$. Now $\operatorname{var}(S_i)=i\sigma_K^2$ so the naive Normal approximation suggests $P(S_i=0)\propto i^{-\frac{1}{2}}\sigma_K^{-1}$ (where $\propto$ means "is proportional to") and then the sum (1.3) gives

$$m(K,n)\propto n^{3/2}\,\sigma_K^{-1}.$$

Thus $m(K,n)=1$ for some n of order $\sigma_K^{2/3}$, which suggests that the time of first self-intersection is of this order.

The purpose of this paper is to prove the natural limit theorem underlying (1.2). As the sketch above suggests, the uniformity of the distribution of $\xi_1^{(K)}$ is not essential. Suppose that for each $K\geqq 1$ we have an i.i.d. sequence $\xi_1,\xi_2,\ldots$ of integer-valued random variables (the subscript K will be suppressed). Suppose $E\xi_1=0$, $E\xi_1^2=\sigma^2<\infty$ for each K, and suppose $\sigma\to\infty$ as $K\to\infty$. Let $S_n=\sum_{i=1}^{n}\xi_i$. Define the normalized partial sum process $S^*(t)$, $0\leqq t<\infty$, by

$$S^*(t)=\sigma^{-4/3}\,S_{[t\sigma^{2/3}]}. \tag{1.4}$$

Under the condition

$$\lim_{H\to\infty}\limsup_{K\to\infty} E\left(\frac{\xi_1}{\sigma}\right)^2 1_{(|\xi_1|>H\sigma)}\to 0 \quad \text{as} \quad K\to\infty; \tag{1.5}$$

we can apply the "weak convergence" version of the Lindeberg-Feller Central Limit Theorem (Billingsley 1968, p. 77) to conclude

$$(S^*(t),\ 0\leqq t<\infty)\xrightarrow{\mathcal{D}}(W(t),\ 0\leqq t<\infty) \quad \text{as} \quad K\to\infty, \tag{1.6}$$

where $W(t)$ is Brownian motion, in the usual sense of weak convergence on $D[0,\infty)$.

It is convenient to study the entire process of self-intersections:

$$T_0=0, \qquad T_i=\min\{n>T_{i-1}:\ S_n=S_m \text{ for some } 0\leqq m<n\}.$$

Define normalized variables

$$T_i^*=\sigma^{-2/3}\,T_i.$$

We want to extend (1.6) to show that $(S^*;T_1^*,T_2^*,\ldots)$, a random element of $D[0,\infty)\times[0,\infty)^\infty$, converges as $K\to\infty$ to some limit $(W;U_1,U_2,\ldots)$, where the (U_i) are some process of pseudo self-intersections of the Brownian motion $W(t)$ (necessarily *pseudo* because of course Brownian motion has *real* self-intersections almost everywhere). Let $L(t,x)$ be local time for $W(t)$. Then $L(t,W(t))$ indicates the "density" of past time that W has spent at its present position. Define $0<U_1<U_2<\ldots$ by

conditional on W, the times (U_i) are the times of the events of a non-homogenous Poisson process of rate $L(t,W(t))$. (1.7)

Informally, this means $P(\text{some } U_i \in (t, t+dt) \mid W(u),\ 0 \leqq u \leqq t) = L(t, W(t))\,dt$. Here is the result of this paper.

Theorem 1.8. *Under technical conditions* (1.5) *and* (4.2),

$$(S^*, T_1^*, T_2^*, \ldots) \xrightarrow{\mathcal{D}} (W, U_1, U_2, \ldots) \quad as \quad K \to \infty.$$

The extra technical condition (4.2) is a type of "uniform non sublattice" condition on the distributions $\xi_1^{(K)}$, needed to obtain local limit estimates. The following sketch should make this plausible. Fix t_0 and condition on S_n, $n \leqq t_0 \sigma^{2/3}$. The chance of a self-intersection during $t_0 \sigma^{2/3} \leqq t \leqq (t_0+\delta)\sigma^{2/3}$ is about

$$\delta\sigma^{2/3} \times \text{density of points } \{S_n\colon\ n \leqq t_0 \sigma^{2/3}\} \text{ around } S_{[t_0\sigma^{2/3}]}. \tag{1.9}$$

The normalized path $S^*(t)$, $t \leqq t_0$, of (1.6) approximates some Brownian path $W(t)$ which has local time density $L(t_0, W(t_0))$ around $W(t_0) \approx S^*(t_0)$. Allowing for the space and time rescalings which relate $S^*(t)$ to S_n, we see that the density in (1.9) is about $\sigma^{2/3} L(t_0, W(t_0))/(\sigma^{4/3})$. So the whole quantity (1.9) is about $\delta L(t_0, W(t_0))$, and this represents the probability of some T_i^* during $(t_0, t_0+\delta)$ given $(S^*(t),\ t \leqq t_0)$. Since S^* converges to W, this suggests Theorem 1.8.

In Sects. 4–6 this sketch is turned into an honest proof. Our proof is a rather complicated assembly of standard ideas - local Normal approximations, weak convergence, and construction techniques. Rick Durrett (personal communication) observed that for certain special sequences of distributions $\xi_1^{(K)}$, e.g. simple symmetric random walk stopped at geometric times, Theorem 1.8 can be deduced simply from known results about convergence of local times of random walks. This is described in Sect. 3.

The most notable consequence of Theorem 1.8 is the asymptotic distribution of T_1.

Corollary 1.10. *Under the hypotheses of Theorem* 1.8

$$\sigma^{-2/3} T_1 \xrightarrow{\mathcal{D}} U_1 \quad as \quad K \to \infty.$$

For the special case where $\xi_1^{(K)}$ is uniform on $\{-K, \ldots, K\}$, the hypotheses of Theorem 1.8 are readily verified, and then

$$3^{1/3} K^{-2/3} T_1 \xrightarrow{\mathcal{D}} U_1. \tag{1.11}$$

Note also that

$$\begin{aligned} P(T_1 > m+n \mid T_1 > m) &\leqq P(S_m, S_{m+1}, \ldots, S_{m+n} \text{ all different} \mid T_1 > m) \\ &= P(T_1 > n). \end{aligned}$$

This subexponentiality property implies convergence of all moments in (1.10), in particular $\sigma^{-2/3} ET_1 \to EU_1$. To investigate the distribution of U_1 we introduce the random variable

$$Y = \int_{-\infty}^{\infty} L^2(1, x)\,dx, \tag{1.12}$$

which is studied in Sect. 2. From the definition of (U_i),

$$P(U_1 > t \mid W) = \exp\left(-\int_0^t L(s, W(s))\, ds\right). \tag{1.13}$$

Lemmas 2.2 and 2.3 imply

$$\int_0^t L(s, W(s))\, ds \stackrel{\mathcal{D}}{=} \tfrac{1}{2} t^{3/2} Y. \tag{1.14}$$

Thus the distribution function of U_1 can be expressed in terms of the Laplace transform of Y:

$$P(U_1 > t) = E \exp(-\tfrac{1}{2} t^{3/2} Y). \tag{1.15}$$

Unfortunately no formula for this Laplace transform is known, although the variable Y has been studied by Borodin (1982). An expression for EU_1 can be obtained by writing $EU_1 = \int_0^\infty P(U_1 > t)\, dt$ and using (1.15) and the fact (change of variables)

$$\int_0^\infty \exp(-a t^{3/2} y)\, dt = a^{-2/3} \Gamma(5/3)\, y^{-2/3}.$$

This gives

$$EU_1 = 2^{2/3} \Gamma(5/3)\, EY^{-2/3} = 1.433 \ldots EY^{-2/3}.$$

Proposition 2.4 establishes the values of EY and EY^2, and a Taylor series expansion (2.12) leads to an approximation

$$EY^{-2/3} \approx 1.00 \tag{1.16}$$

which is supported by computer simulations. Thus

$$EU_1 \approx 1.43. \tag{1.17}$$

In the case where $\xi_1^{(K)}$ is uniform on $\{-K, \ldots, +K\}$, we see from (1.11) and (1.17) that

$$\lim_{K \to \infty} K^{-2/3} ET^{(K)} = 3^{-1/3} EU_1 \approx 0.99.$$

Other natural questions concern the position S_{T_1} of the first self-intersection, and the range of the random walk before the first self-intersection:

$$M_+^{(K)} = \max_{i \leq T_1} S_i, \qquad M_-^{(K)} = \min_{i \leq T_1} S_i.$$

Theorem 1.8 yields

Corollary 1.15. *Under the hypotheses of Theorem* 1.8

$$\sigma^{-4/3}(S_{T_1}^{(K)}, M_+^{(K)}, M_-^{(K)}) \xrightarrow{\mathcal{D}} (W_{U_1}, M_+, M_-) \quad as \quad K \to \infty,$$

where $M_+ = \sup_{t \leq U_1} W(t), \qquad M_- = \inf_{t \leq U_1} W(t).$

Since U_1 is a (randomized) stopping time for $W(t)$, the martingale optional sampling theorem and maximal inequalities give some information about the limit distributions: for instance,

$$EW_{U_1}=0; \qquad EW_{U_1}^2=EU_1$$
$$E\{\max(M_+, M_-)\}^2 \leqq 4EU_1.$$

The distribution of M_+ is in principle susceptible of exact analysis by diffusion techniques, but I am unable to carry through this analysis to an explicit conclusion.

In Sect. 7 we discuss three topics related to Theorem 1.8.

(a) Iterates of random functions.

(b) Self-avoiding 1-dimensional random walks.

(c) First self-intersections of random walks in d dimensions, and on abstract groups.

2. Some Distributions Related to Local Time

We define local time $L(t,x)$ as occupation density:

$$\int_{-\infty}^{y} L(t,x)\,dx = \int_0^t 1_{(W(s)\leqq y)}\,ds,$$

thereby disagreeing by a factor of 2 from some other definitions. Define

$$Y_t = \int_{-\infty}^{\infty} L^2(t,x)\,dx \tag{2.1}$$

so that Y_1 is the variable Y of (1.12). Lemmas 2.2 and 2.3 were used to establish (1.14). We remark that the process $L(t, W(t))$ has been studied for different reasons by Barlow (1982).

Lemma 2.2. $$\int_0^t L(s, W(s))\,ds = \tfrac{1}{2} Y_t.$$

Proof. Fix ω. Let $t_1 < t_2$. Let

$$(a,b) = (\inf_{t_1 \leqq s \leqq t_2} W(s), \sup_{t_1 \leqq s \leqq t_2} W(s)).$$

Then $Y_{t_2} - Y_{t_1} = \int \{L(t_2,x) - L(t_1,x)\}\{L(t_2,x)+L(t_1,x)\}\,dx$ where we need only integrate over (a,b) because $L(t_2,x) = L(t_1,x)$ outside that interval. Since $\int \{L(t_2,x) - L(t_1,x)\}\,dx = t_2 - t_1$, we obtain

$$2 \inf_{a\leqq x\leqq b} L(t_1,x) \leqq (Y_{t_2} - Y_{t_1})/(t_2 - t_1) \leqq 2 \sup_{a \leqq x \leqq b} L(t_2,x).$$

As $t_1, t_2 \to t$ we have $a, b \to W(t)$ and so by joint continuity of L

$$\frac{dY_t}{dt} = 2L(t, W(t)),$$

establishing the lemma.

Lemma 2.3. $$Y_t \stackrel{\mathcal{D}}{=} t^{3/2}\, Y_1.$$

Proof. This is just the scaling property that Y_t inherits from $W(t)$. Fix $a>0$. Let $\hat{W}(t)=a^{-\frac{1}{2}}\, W(a\,t)$, so $\hat{W}(t)$ is another Brownian motion. Then $\hat{W}$ has local time $\hat{L}(t,x)=a^{-\frac{1}{2}}\, L(a\,t, a^{\frac{1}{2}}\, x)$, and so

$$\begin{aligned}\hat{Y}_t &= \int \hat{L}^2(t,x)\,dx\\ &= a^{-1}\int L^2(a\,t, a^{\frac{1}{2}}\,x)\,dx\\ &= a^{-3/2}\int L^2(a\,t,z)\,dz = a^{-3/2}\, Y_{at}.\end{aligned}$$

But $\hat{Y}_t \stackrel{\mathcal{D}}{=} Y_t$, so setting $a=t^{-1}$ gives the result.

Now set $Y=Y_1$. The next result has been obtained independently by Borodin (1982) using Fourier methods, but it seems interesting to give a more probabilistic proof. The process $L(1,x)$ has been studied by Perkins (1982a) for different reasons.

Proposition 2.4.

(a) $EY=(32/9\pi)^{\frac{1}{2}}=1.064\ldots$

(b) $EY^2=11/9$.

Proof. Brownian motion has the time-reversibility property

$$(W(s),\ 0\leqq s\leqq t) \stackrel{\mathcal{D}}{=} (W(t-s)-W(t),\ 0\leqq s\leqq t) \tag{2.5}$$

which implies that for fixed t

$$\begin{aligned}L(t, W(t)) &\stackrel{\mathcal{D}}{=} L(t,0)\\ &\stackrel{\mathcal{D}}{=} |W(t)| \quad \text{by the Lévy representation of } L(t,0).\end{aligned}$$

So

$$\begin{aligned}EY &= 2\int_0^1 EL(t, W(t))\,dt \quad \text{by Lemma 2.2}\\ &= 2\int_0^1 E|W(t)|\,dt\\ &= 2\int_0^1 (2t/\pi)^{\frac{1}{2}}\,dt,\end{aligned}$$

which gives (a). For (b),

$$\begin{aligned}Y_t^2 &= \int_0^t 2\,Y_s\,dY_s/ds\,ds\\ &= \int_0^t 4\,Y_s\,L(s, W(s))\,ds \quad \text{by Lemma 2.2.}\end{aligned}$$

Taking expectations and differentiating,

$$\frac{d}{dt}(E Y_t^2) = 4E Y_t\, L(t, W(t)).$$

But Lemma 2.3 implies

$$E Y_t^2 = t^3 E Y_1^2. \tag{2.6}$$

Taking derivatives,

$$\begin{aligned} E Y_t^2 &= t/3 \frac{d}{dt}(E Y_t^2) \\ &= 4t/3\, E Y_t L(t, W(t)) \\ &= 4t/3\, E Y_t L(t, 0) \end{aligned} \tag{2.7}$$

using the time-reversibility property (2.5). Now let T_λ be an exponential, rate λ, variable independent of W. From (2.6) and (2.7) we get the scaling property

$$Y_t L(t,0) \stackrel{\mathcal{D}}{=} t^2 Y_1 L(1,0)$$

and so

$$E Y_{T_\lambda} L(T_\lambda, 0) = 2\lambda^{-2} E Y_1 L(1,0),$$

whence (2.7) gives

$$E Y^2 = \tfrac{2}{3}\lambda^2 E Y_{T_\lambda} L(T_\lambda, 0) = \tfrac{2}{3}\lambda^2 E \int_{-\infty}^{\infty} L^2(T_\lambda, x)\, L(T_\lambda, 0)\, dx. \tag{2.8}$$

To evaluate this, fix x and let H_x be the first hitting time of $W(t)$ on x. Define

$$\begin{aligned} p &= P(H_x < T_\lambda) \\ &= E \exp(-\lambda H_x) \quad \text{by conditioning on } H_x \\ &= \exp(-x\sqrt{2\lambda}) \end{aligned} \tag{2.9}$$

where the last identity is classical (Williams (1979) p. 85). A standard result from local time theory is

$$L(H_x, 0) \text{ has exponential distribution, rate } (2x)^{-1}.$$

A proof is given in Williams (1974) Theorem 4.2; the same argument applied to Brownian motion killed at rate λ yields

$$L(H_x \wedge T_\lambda, 0) \text{ has exponential distribution, rate } q = (1-p^2)^{-1}\sqrt{2\lambda}. \tag{2.10}$$

Now consider the accumulations of local time at 0 and x over the interval up to the first hit on x, then the interval until the next hit on 0, then the interval until the next hit on x, and so on. We see from (2.9), (2.10) and the strong Markov property that

$$(L(T_\lambda, 0), L(T_\lambda, x)) \stackrel{\mathcal{D}}{=} \Big(\sum_{\substack{1 \leqq i \leqq N \\ i \text{ odd}}} X_i, \sum_{\substack{1 \leqq i \leqq N \\ i \text{ even}}} X_i \Big)$$

where (X_i) are independent exponentials, rate q, and N is independent of (X_i) with $P(N=n) = (1-p)p^{n-1}$, $n \geqq 1$. Routine tedious calculations enable us to evaluate $EL^2(T_\lambda, x)\, L(T_\lambda, 0)$, and then to obtain (b) from (2.8).

We can use Proposition 2.4 to estimate $EY^{-2/3}$, needed at (1.16). For a random variable Y with mean μ and variance s^2, and for a smooth function f, the Taylor series expansion

$$f(y)=f(\mu)+(y-\mu)f'(\mu)+\tfrac{1}{2}(y-\mu)^2 f''(\mu)+r(y) \tag{2.11}$$

gives

$$Ef(Y)=f(\mu)+\tfrac{1}{2}s^2 f''(\mu)+E\,r(Y).$$

In our setting, with $f(y)=y^{-2/3}$ and the numerical values given in Proposition 2.4 we get

$$EY^{-2/3}=1.002\ldots+E\,r(Y). \tag{2.12}$$

Computer simulations give $EY^{-2/3}=1.01\pm 0.02$. Jensen's inequality gives the rigorous lower bound

$$EY^{-2/3}\geqq (EY)^{-2/3}=0.96\ldots$$

but a good rigorous upper bound seems harder to find.

3. A Special Case of the Theorem

Brownian motion local time can be obtained as a limit of simple random walk local times. Rick Durrett (personal communication) suggested this could be used to establish Theorem 1.3 for the special case of simple symmetric random walks stopped at geometric times: this section shows how. Unfortunately, it seems impossible to use this method more generally, even for the uniform case.

Let (η_i) be independent, $P(\eta_i=1)=P(\eta_i=-1)=\frac{1}{2}$. Let $S_n=\sum_{i=1}^{n}\eta_i$. For each K define

$$L^{(K)}(i/K, j/K^{\frac{1}{2}})=K^{-\frac{1}{2}}\sum_{s=0}^{i}1_{(S_s=j)}$$
$$W^{(K)}(t)=K^{-\frac{1}{2}}S_{[Kt]}. \tag{3.1}$$

By interpolating, for each ω we can define $L^{(K)}(t,x)$ as a continuous function on $[0,\infty)\times(-\infty,\infty)$, so that $L^{(K)}$ becomes a random element of $C([0,\infty)\times(-\infty,\infty))$. It is a standard result (Perkins 1982b; Borodin 1981) that

$$(W^{(K)}, L^{(K)})\xrightarrow{\mathcal{D}}(W,L) \tag{3.2}$$

where $L=L(t,x)$ is local time for Brownian motion $W=W(t)$.

Let $(A_{K,i})$ be independent events with $P(A_{K,i})=K^{-3/4}$, independent of (η_i). Let $X_{K,n}$ be the time of occurrence of the n^{th} event of $(A_{K,i}\colon i\geqq 1)$. For future reference, note that the L^2 martingale maximal inequality implies

$$\sup_{n\leqq K^{3/8}}|K^{-1}X_{K,n}-K^{-\frac{1}{4}}n|\xrightarrow[p]{}0 \quad \text{as} \quad K\to\infty. \tag{3.3}$$

Now define $S_n^{(K)}=S_{X_{K,n}}$. For each K, $(S_n^{(K)};\ n\geqq 0)$ is a random walk whose step distribution $\xi^{(K)}$ is that of the simple random walk stopped at the geometric time $X_{K,1}$. So $E\xi^{(K)}=0$ and

$$\sigma_K^2=\operatorname{var}(\xi^{(K)})=EX_{K,1}=K^{3/4}. \tag{3.4}$$

Now define

$$\hat{T}_K = \min\{n: \text{ there exists } m<n,\ S_m = S_n, \text{ and } A_{K,m} \text{ and } A_{K,n} \text{ occur}\}.$$

We shall prove

$$(W^{(K)}, K^{-1}\hat{T}_K) \xrightarrow{\mathscr{D}} (W, U) \tag{3.5}$$

where U is the variable U_1 of (1.7). But the first self-intersection time $T_1^{(K)}$ for $(S_n^K: n \geqq 0)$ is such that

$$\hat{T}_K = X_{K, T_1^{(K)}},$$

and so from (3.5) and (3.3) we get

$$(W^{(K)}, K^{-\frac{1}{4}} T_1^{(K)}) \xrightarrow{\mathscr{D}} (W, U).$$

In view of (3.4) this is just the assertion

$$(W^{(K)}, \sigma_K^{-2/3} T_1^{(K)}) \xrightarrow{\mathscr{D}} (W, U).$$

The full form of Theorem 1.8 can be proved similarly.

To prove (3.5), fix t and consider the event $\{\hat{T}_K > tK\}$. This is the intersection over j of the events:

> there do not exist times $0 \leqq m < n \leqq tK$ such that $S_m = S_n = j$ and such that events $A_{K,m}$ and $A_{K,n}$ occur.

Now conditional on the path (S_i), these events are independent as j varies. So, conditional on the path,

$$\begin{aligned} P(\hat{T}_K > tK) &= \prod_j P(\text{there do not exist } 0 \leqq m < n \leqq tK \text{ such} \\ &\qquad \text{that } S_m = S_n = j \text{ and } A_{K,m} \text{ and } A_{K,n} \text{ occur}) \\ &= \prod_j P(Z_j \leqq 1) \text{ for } Z_j \overset{\mathscr{D}}{=} \text{Binomial}(N_j, K^{-3/4}),\ N_j = \sum_{s=0}^{tK} 1_{(S_s = j)} \\ &= \prod_j \{1 - \tfrac{1}{2}(K^{-3/4} N_j)^2 \phi(K^{-3/4} N_j)\} \end{aligned}$$

where $\phi(x) \to 1$ as $x \downarrow 0$. Now $N_j = K^{\frac{1}{2}} L^{(K)}(t, j/K^{\frac{1}{2}})$, so

$$P(\hat{T}_K > tK \mid W^{(K)}) = \prod_j \{1 - \tfrac{1}{2} K^{-\frac{1}{2}} L^{2(K)}(t, j/K^{\frac{1}{2}})\, \phi(K^{-\frac{1}{4}} L^{(K)}(t, j/K^{\frac{1}{2}}))\}.$$

But $\prod_j \{1 - \frac{1}{2} x_j^2\, \phi(x_j)\} = \exp(-\frac{1}{2}(1-\delta) \sum x_j^2)$ where $\delta = \delta((x_j)) \to 0$ as $\max |x_j| \to 0$. And (3.2) implies $\max_j K^{-\frac{1}{4}} L^{(K)}(t, j/K^{\frac{1}{2}}) \xrightarrow[p]{} 0$ as $K \to \infty$, so

$$P(\hat{T}_K > tK \mid W^{(K)}) = \exp(-\tfrac{1}{2}(1 - \Delta_K(t))\, Y^{(K)}(t)) \tag{3.6}$$

where $Y^{(K)}(t) = K^{-\frac{1}{2}} \sum_j L^{2(K)}(t, j/K^{\frac{1}{2}})$ and $\Delta_K(t) \xrightarrow[p]{} 0$ as $K \to \infty$.

From (3.2),

$$(W^{(K)}, L^{(K)}, Y^{(K)}) \xrightarrow{\mathscr{D}} (W, L, Y) \tag{3.7}$$

where $Y = Y(t)$ is defined in Sect. 2. By (1.13) and Lemma 2.2

$$P(U>t \mid W) = \exp(-\tfrac{1}{2} Y(t))$$

and then (3.6) and (3.7) imply (3.5).

4. Some Estimates for the Random Walks

Sections 4–6 contain the proof of the general case of Theorem 1.8. In this section we specify technical assumptions on the random walks, and give local limit theorems, conditioned limit theorems and bounds on mean numbers of self-intersections. These results are fairly standard, so the proofs are merely outlined.

Here are the hypotheses for Theorem 1.8, as stated in Sect. 1. For each $K \geqq 1$ we have an i.i.d. integer-valued sequence $\xi_i^{(K)}$ with $E\xi_1^{(K)} = 0$ and $\operatorname{var}(\xi_1^{(K)}) = \sigma_K^2$, and with partial sums $S_n^{(K)} = \sum_{i=1}^{n} \xi_i^{(K)}$. (The superscripts K will generally be suppressed.) Assume that as $K \to \infty$ we have $\sigma \to \infty$ and (repeating (1.5))

$$\lim_{H\to\infty} \limsup_{K\to\infty} E(\xi_1/\sigma)^2 \, 1_{(|\xi_1| > H\sigma)} = 0. \tag{4.1}$$

We also impose the technical hypotheses

$$\sigma^{2/3} \max_j P(\xi_1 = j) \to 0 \quad \text{as} \quad K \to \infty; \tag{4.2a}$$

there exist $\alpha, \beta > 0$ such that for all K

$$\min_{|j| \leqq \beta\sigma} P(\xi_1 = j) \geqq \alpha/\sigma. \tag{4.2b}$$

To see why, first consider self-intersections of the form $S_{m+1} = S_m$; it is easy to see that

$$\sigma^{2/3} P(\xi_1 = 0) \to 0 \quad \text{as} \quad K \to \infty$$

is a necessary condition for Theorem 1.8, so hypothesis (4.2a) seems reasonable. Second, Theorem 1.8 will not hold if ξ_1 is supported on some sublattice of the integers, and our proof uses a local *CLT*, Lemma 4.3 below. Hypothesis (4.2b) is used only in establishing Lemma 4.3, and could be weakened (at the expense of requiring more complicated Fourier analysis to establish Lemma 4.3).

Let $\phi(1, x)$ be the standard Normal density.

Lemma 4.3. $\sup_j |\sigma n^{\frac{1}{2}} P(S_n = j) - \phi(1, j\sigma^{-1} n^{-\frac{1}{2}})| \to 0$ *as* $K \to \infty$, $n \to \infty$.

Outline of Proof. By easy Fourier analysis, the Lemma is true in the special case where $\xi_1^{(K)}$ is uniform on $\{-r_K, -r_k+1, \ldots, r_K\}$, for some $r_K \to \infty$ as $K \to \infty$. Second, in the presence of the global *CLT*, the local *CLT* (4.3) is equivalent to the "local smoothness" property

$$\sigma n^{\frac{1}{2}} |P(S_n = j_1) - P(S_n = j_2)| \to 0$$
$$\text{as } K \to \infty, \ n \to \infty, \qquad \sigma^{-1} n^{-\frac{1}{2}} |j_1 - j_2| \to 0. \tag{4.4}$$

Hypothesis (4.2b) implies we can write

$$\mathscr{L}(\xi_i) = \hat{\alpha} \mathscr{L}(\xi_{1,i}) + (1-\hat{\alpha}) \mathscr{L}(\xi_{2,i}), \qquad \hat{\alpha} = 2\alpha\beta,$$

where $\xi_{1,i}$ is uniform on $\{-[\beta\sigma], \ldots, [\beta\sigma]\}$. So we can write $S_n = S_{1,N} + S_{2,n-N}$, where $S_{q,m} = \sum_{i=1}^{m} \xi_{q,i}$ and N has Binomial$(n, \hat{\alpha})$ distribution. By conditioning on N and $(S_{2,m};\ m \geqq 0)$, we see that the smoothness property (4.4) for $S_{1,n}$ (which holds by the "special case" above) implies the same property holds for S_n, thus establishing the lemma.

For the next result, let $\phi(s, x)$ be the density of $W(s)$ and let

$$G(t, x) = \int_0^t \phi(s, x)\, ds = EL(t, x)$$

where L is local time for Brownian motion $W(s)$. Let

$$m(K, n, j) = \sum_{i=1}^{n} P(S_i = j). \tag{4.5}$$

Lemma 4.6. *Fix* $0 < s < t$.

(a) $\sup_{s\sigma^{2/3} \leqq i \leqq t\sigma^{2/3}} \sup_j |\sigma^{4/3} P(S_i = j) - \phi(i\sigma^{-2/3}, j\sigma^{-4/3})| \to 0$ *as* $K \to \infty$.

(b) $\sup_{u \leqq t} \sup_j |\sigma^{2/3} m(K, u\sigma^{2/3}, j) - G(u, j\sigma^{-4/3})| \to 0$ *as* $K \to \infty$.

Outline of Proof. Assertion (a) is just a reformulation of Lemma 4.3. To prove (b), sum (a) over $s\sigma^{2/3} < i \leqq t\sigma^{2/3}$ to get

$$\sup_j |\sigma^{2/3}\{m(K, t\sigma^{2/3}, j) - m(K, s\sigma^{2/3}, j)\} - \{G(t, j\sigma^{-4/3}) - G(s, \sigma^{-4/3})\}| \to 0.$$

From this and monotonicity of $m(K, \cdot, j)$ it suffices to prove

$$\lim_{s \downarrow 0} \limsup_{K \to \infty} \sup_j \sigma^{2/3} m(K, s\sigma^{2/3}, j) = 0. \tag{4.7}$$

But we can use Lemma 4.3 to show that, for any $i(K) \to \infty$ as $K \to \infty$, $i(K) \leqq s\sigma^{2/3}$,

$$\lim_{s \downarrow 0} \limsup_{K \to \infty} \sup_j \sigma^{2/3} \{m(K, s\sigma^{2/3}, j) - m(K, i(K), j)\} = 0.$$

Then taking $i(K) \to \infty$ sufficiently slowly,

$$\sigma^{2/3} m(K, i(K), j) \leqq \sigma^{2/3} i(K) \max_j P(\xi_1 = j)$$
$$\to 0 \quad \text{by hypothesis (4.2a)},$$

and this establishes the lemma.

Our next result is a conditioned limit theorem. Recall the definition (1.4) of the renormalized random walk $S^*(t)$. Let $\tilde{d}$ be a metrization of weak convergence on $D[0, \infty)$.

Lemma 4.8. *Let $\Theta(K, i, j)$ be the conditional law of $(S^*(t), 0 \leqq t < \infty)$ given $S_i = j$. Let $\Phi(t, y)$ be the conditional law of $(W(s), 0 \leqq s < \infty)$ given $W(t) = y$. Then for fixed $0 < s_0 < s_1$ and $J < \infty$,*

$$\lim_{K \to \infty} \sup_{s_0 \leqq i\sigma^{-2/3} \leqq s_1} \sup_{|j\sigma^{-4/3}| \leqq J} \tilde{d}(\Theta(K, i, j), \Phi(i\sigma^{-2/3}, j\sigma^{-4/3})) = 0.$$

Outline of Proof. It suffices to prove that, if $i = i(K)$ and $j = j(K)$ satisfy

$$i\sigma^{-2/3} \to t_0 > 0, \quad j\sigma^{-4/3} \to y$$

then

$$\Theta(K, i, j) \to \Phi(t_0, y) \quad \text{weakly as } K \to \infty.$$

Write $\hat{S}_i$ for the conditioned random walks, $\hat{S}^*(t)$ for their normalized versions, and write $\hat{W}$ for the conditioned Brownian motions. The basic weak convergence result (1.6) shows

$$(\hat{S}^*(t), t \geqq i\sigma^{-2/3}) \xrightarrow{\mathscr{D}} (\hat{W}(t), t \geqq t_0)$$

and so the issue is proving

$$(\hat{S}^*(t), t \leqq i\sigma^{-2/3}) \xrightarrow{\mathscr{D}} (\hat{W}(t), t \leqq t_0).$$

But convergence of finite-dimensional distributions here can be deduced from the local *CLT* (4.3). And tightness follows because the processes $(\hat{S}_u; u \geqq i)$ have exchangeable increments, and for such processes tightness in D is a consequence of tightness of f.d.d.'s (this last fact can be deduced from Billingsley (1968) Theorem 24.2).

The final results of this section give bounds on mean numbers of (pseudo) self-intersections. We first treat the Brownian motion case. Fix t_0 and a continuous function $w(t)$, $0 \leqq t \leqq t_0$ (which is to be thought of as a typical Brownian path) such that

$$w(\cdot) \text{ has an occupation density } f(x) \text{ such that } F \equiv \sup f(x) < \infty. \tag{4.9}$$

Let $W(t)$, $t \geqq 0$ be Brownian motion with $W(0)$ arbitrary. Define $N(t)$, $t \geqq 0$, by: conditional on W, the process N is the non-homogeneous Poisson process of rate $f(W(t))$.

Lemma 4.10. (a) $EN(t) \leqq Ft$.

(b) $EN(t) \leqq P(N(t) \geqq 1) + (Ft)^2$.

Proof. Assertion (a) is clear, since $f(W(t)) \leqq F$, so N is stochastically dominated by the Poisson process of rate F. For the same reason we have

$$E(N(t) - N(s) \mid W(u), u \leqq s) \leqq Ft; \quad s \leqq t.$$

By conditioning on the time of the first event of N,

$$E(N(t) - 1 \mid N(t) \geqq 1) \leqq Ft.$$

So

$$EN(t) - P(N(t) \geqq 1) = E(N(t)-1)\,1_{(N(t)\geqq 1)}$$
$$\leqq F\,t\,P(N(t) \geqq 1)$$
$$\leqq (F\,t)^2 \quad \text{by (a)},$$

and this establishes (b).

We now give the corresponding random walk result. For each K let s_i, $i \leqq t_0\,\sigma^{2/3}$ be a sequence of integers (to be thought of as a typical path of the random walk). Let $A = \{s_i\}$. Let

$$s^*(t) = \sigma^{-4/3}\, s_{[t\sigma^{2/3}]} \qquad t \leqq t_0,$$

be the normalized path. Let S_i be the random walk with S_0 arbitrary, and let $M(K, n) = |\{i\colon 1 \leqq i \leqq n,\ S_i \in A\}|$ be the number of visits of S_i to the set A.

Lemma 4.11. *Suppose $s^*(\cdot) \to w(\cdot)$ in $D[0, t_0]$ as $K \to \infty$, for some $w(\cdot)$, F as at (4.9). Then*

(a) $\limsup_{K\to\infty} EM(K, t\,\sigma^{2/3}) \leqq F\,t.$

(b) $\limsup_{K\to\infty} \{EM(K, t\,\sigma^{2/3}) - P(M(K, t\,\sigma^{2/3}) \geqq 1)\} \leqq (F\,t)^2.$

Proof.

$$EM(K, t\,\sigma^{2/3}) \leqq \sum_{i \leqq t_0\sigma^{2/3}} m(K, t\,\sigma^{2/3}, s_i - S_0)$$

for m as at (4.5);

$$\leqq \sum_{i \leqq t\sigma^{2/3}} \sigma^{-2/3}\, G(t, \sigma^{-4/3}(s_i - S_0)) + \delta_K$$

where $\delta_K \to 0$, by Lemma 4.6(b);

$$= \sum_{i \leqq t\sigma^{2/3}} \sigma^{-2/3}\, G(t, s^*(i\,\sigma^{-2/3}) - a_K) + \delta_K, \quad \text{for some } a_K;$$

$$= \int_0^t G(t, s^*(u) - a_K)\, du + \delta_K.$$

Now $s^*(\cdot) \to w(\cdot)$ and $G(t, \cdot)$ is continuous. So

$$\limsup_{K\to\infty} EM(K, t\,\sigma^{2/3}) \leqq \sup_a \int_0^t G(t, w(u) - a)\, du$$
$$\leqq \int_{-\infty}^{\infty} G(t, x)\, F\, dx$$

since $w(\cdot) - a$ has occupation density $f(x-a)$ bounded by F;

$$= F\,t.$$

This is assertion (a), and part (b) follows in the same way as in Lemma 4.10.

5. A Construction

The Skorokhod representation theorem says that, given random variables X, Y whose distributions are close in the sense of weak convergence, then we can construct a joint distribution which makes the variables close in probability. More sharply, let S be a Polish space with bounded complete metric d. For distributions μ, v on S let

$$\tilde{d}(\mu, v) = \inf E\, d(X, Y) \tag{5.1}$$

where the infimum is taken over all joint distributions (X, Y) with marginals μ, v. Then $\tilde{d}$ is a metrization of weak convergence. Further, the inf in (5.1) is attained. See Pollard (1984) for discussion.

In the proof of Theorem 1.8 we encounter a complication. Associated with X (resp. Y) are events (A_i) (resp. (B_i)), and we need to construct a joint distribution such that not only are X and Y close, but also the events A_i and B_i almost coincide for each i. Proposition 5.2 gives conditions under which we can make such a construction.

Notation. In this section X, Y denote S-valued random variables; A, B, Δ denote events; μ, v denote (sub)probability distributions on S; $\mathscr{L}(X)$ and $\mathscr{L}(X|A)$ denote distribution and conditional distribution.

Proposition 5.2. *Suppose we are given families* $(X;\ A_1, A_2, \ldots)$ *and* $(Y;\ B_1, B_2, \ldots)$. *Suppose that for some* $\alpha, \varepsilon, \lambda, \theta > 0$.

(i) $(1+\varepsilon)^{-1} \leqq P(A_i)/P(B_i) \leqq 1+\varepsilon$ *for all i.*

(ii) $\sum P(A_i) \leqq \theta;\quad \sum P(B_i) \leqq \theta.$

(iii) $\sum P(A_i) \leqq P(\bigcup A_i) + \lambda;\quad \sum P(B_i) \leqq P(\bigcup B_i) + \lambda.$

(iv) $\tilde{d}(\mathscr{L}(X|A_i), \mathscr{L}(Y|B_i)) \leqq \alpha$ *for all i.*

Let $0 < \eta < \frac{1}{2}$. *Then we can construct* $(\hat{X}, \hat{Y}, \hat{A}_i, \hat{B}_i, \Delta)$ *such that*

(a) $(\hat{X}, \hat{A}_1, \hat{A}_2, \ldots) \overset{\mathscr{D}}{=} (X, A_1, A_2, \ldots);\quad (\hat{Y}, \hat{B}_1, \hat{B}_2, \ldots) \overset{\mathscr{D}}{=} (Y, B_1, B_2, \ldots).$

(b) *Outside* Δ *we have:* $\hat{A}_i = \hat{B}_i$ *for each i, and these sets are disjoint as i varies.*

(c) $P(\Delta) \leqq 4(2\eta+\varepsilon)\theta + 4(2+\varepsilon)\lambda\eta^{-1} + 2\lambda.$

(d) $E\, d(\hat{X}, \hat{Y})\, 1_{\Delta^c} \leqq \tilde{d}(\mathscr{L}(X), \mathscr{L}(Y)) + 2(1-2\eta)^{-1}\theta\alpha.$

Remarks. (a) We apply this in the next section, where X and Y will be segments of Brownian motion and the rescaled random walk, and the events A_i and B_i will be the events that self-intersection occurs in some small space-time region.

(b) When doing constructions we assume we can find random variables independent of any given family of random objects. The pedantic reader may add "enlarging the probability space if necessary".

(c) The rest of this section is devoted to the proof of Proposition 5.2. We first state six lemmas; the first three are obvious, and the second three will be proved later.

Lemma 5.3. "Family extension". *Let* $(X_\gamma\colon \gamma \in \Gamma)$ *and* $(Y_\gamma\colon \gamma \in \Gamma_0 \subset \Gamma)$ *be such that* $(X_\gamma\colon \gamma \in \Gamma_0) \overset{\mathscr{D}}{=} (Y_\gamma\colon \gamma \in \Gamma_0)$. *Then we can construct* $(Y_\gamma\colon \gamma \in \Gamma \setminus \Gamma_0)$ *such that* $(X_\gamma\colon \gamma \in \Gamma) \overset{\mathscr{D}}{=} (Y_\gamma\colon \gamma \in \Gamma)$.

Lemma 5.4. "Domain extension". *Let* μ *be a probability distribution. Let* X, *defined on* $\Omega_0 \subset \Omega$, *be such that* $P(\Omega_0, X \in \cdot) \leqq \mu(\cdot)$. *Then we can define* X *on* $\Omega \setminus \Omega_0$ *such that* X *has distribution* μ.

Lemma 5.5. "Splicing". *Let* (X_k, Y_k), $k \geqq 1$, *be such that* $Y_k \overset{\mathscr{D}}{=} Y_1$ *for each* k. *Then we can construct* $(\hat{X}_1, \hat{X}_2, \ldots, \hat{Y})$ *such that* $(\hat{X}_k, \hat{Y}) \overset{\mathscr{D}}{=} (X_k, Y_k)$ *for each* k.

Lemma 5.6. *Suppose we are given random variables* X, Y *and events* $A_0 \subset A$, $B_0 \subset B$ *such that* $P(A_0^c | A) \leqq \eta$ *and* $P(B_0^c | B) \leqq \eta$. *Then we can construct* $A_1 \subset A_0$ *and* $B_1 \subset B_0$ *such that*

(a) $P(A_1) = P(B_1) \geqq (1 - 2\eta) \min(P(A), P(B))$.

(b) $\tilde{d}(\mathscr{L}(X | A_1), \mathscr{L}(Y | B_1)) \leqq (1 - 2\eta)^{-1} \tilde{d}(\mathscr{L}(X | A), \mathscr{L}(Y | B))$.

Lemma 5.7. *Suppose variables* X, Y *are defined on* $A, B \subset \Omega$ *respectively, and suppose* $P(A, X \in \cdot) \leqq \mu(\cdot)$ *and* $P(B, Y \in \cdot) \leqq \nu(\cdot)$ *for probability distributions* μ, ν. *Then we can extend* X, Y *to all* Ω *and construct* Δ *such that*

(a) $\mathscr{L}(X) = \mu$; $\quad \mathscr{L}(Y) = \nu$.

(b) $A \cup B \subset \Delta$ *and* $P(\Delta) \leqq P(A) + P(B)$.

(c) $E\, d(X, Y)\, 1_{\Delta^c} \leqq \tilde{d}(\mu, \nu)$.

For the final lemma, note that we can extend $\tilde{d}$ to subprobability distributions μ, ν with equal total mass $|\mu| = |\nu| = p$ as follows: $\tilde{d}(\mu, \nu) = \inf E\, d(X, Y)\, 1_A$, where the infimum is taken over all (X, Y) defined on some $A \subset \Omega$ such that $P(A) = p$ and $P(A, X \in \cdot) = \mu(\cdot)$, $P(A, Y \in \cdot) = \nu(\cdot)$.

Lemma 5.8. *Suppose the probability distributions* μ, ν *can be written as* $\mu = \mu_1 + \mu_2$, $\nu = \nu_1 + \nu_2$, *where* $\mu_1, \mu_2, \nu_1, \nu_2$ *are subprobability distributions with* $|\mu_j| = |\nu_j|$. *Then*

(a) $\tilde{d}(\mu, \nu) \leqq \tilde{d}(\mu_1, \nu_1) + \tilde{d}(\mu_2, \nu_2)$.

(b) $\tilde{d}(\mu_2, \nu_2) \leqq \tilde{d}(\mu, \nu) + \tilde{d}(\mu_1, \nu_1)$.

Proof of Proposition 5.2. Let I be the set of i such that $P(\bigcup_{j<i} A_j | A_i) \leqq \eta$ and $P(\bigcup_{j<i} B_j | B_i) \leqq \eta$. For $i \in I$ let $A_{i,0} = A_i \setminus \bigcup_{j<i} A_j$, and similarly for $B_{i,0}$. By Lemma 5.6 and hypothesis (iv), for $i \in I$ we can construct $A_{i,1} \subset A_{i,0}$, $B_{i,1} \subset B_{i,0}$ such that

$$\tilde{d}(\mathscr{L}(X | A_{i,1}), \mathscr{L}(Y | B_{i,1})) \leqq (1 - 2\eta)^{-1} \alpha \tag{5.9}$$

$$P(A_{i,1}) = P(B_{i,1}) \geqq (1 - 2\eta) \min(P(A_i), P(B_i)). \tag{5.10}$$

By construction the events $(A_{i,1})$ are disjoint. Let $(J_i;\ i \in I)$ be disjoint events with $P(J_i) = P(A_{i,1}) = P(B_{i,1})$. For $i \in I$ construct $(\hat{X}, \hat{Y})$ on J_i such that

$$\begin{gathered} \mathscr{L}(\hat{X} | J_i) = \mathscr{L}(X | A_{i,1}); \quad \mathscr{L}(\hat{Y} | J_i) = \mathscr{L}(Y | B_{i,1}); \\ E(d(\hat{X}, \hat{Y}) | J_i) = \tilde{d}(\mathscr{L}(X | A_{i,1}), \mathscr{L}(Y | B_{i,1})). \end{gathered} \tag{5.11}$$

Let $J = \bigcup_I J_i$. By disjointness,

$$\mathscr{L}((\hat{X}; J_i, i \in I | J) = \mathscr{L}((X; A_{i,1}, i \in I) | \bigcup_I A_{i,1})$$

and similarly for $\hat{Y}, B_{i,1}$. By Lemmas 5.3 and 5.4 we can construct $\hat{A}_i \supset J_i$, $\hat{B}_i \supset J_i$ and extend the domain of $\hat{X}$ to $\hat{A} = \bigcup_{i \geqq 1} \hat{A}_i$ and the domain of $\hat{Y}$ to $\hat{B}$

$=\bigcup_{i\geqq 1}\hat{B}_i$ such that

$$\mathscr{L}((\hat{X}, J_i, \hat{A}_i, i\geqq 1)|\hat{A})=\mathscr{L}((X, A_{i,1}, A_i, i\geqq 1)|\bigcup_{i\geqq 1} A_i) \tag{5.12}$$

and similarly for $\hat{Y}$ and the B's. Interpret J_i, $A_{i,1}$, $B_{i,1}$ as empty for $i\notin I$. Note $J\subset\hat{A}\cap\hat{B}$. Applying Lemma 5.7 conditionally on J^c, we can extend $\hat{X}$ and $\hat{Y}$ to the remainder of Ω such that

$$\mathscr{L}(\hat{X}|J^c)=\mathscr{L}(X|(\bigcup_I A_{i,1})^c); \qquad \mathscr{L}(\hat{Y}|J^c)=\mathscr{L}(Y|(\bigcup_I B_{i,1})^c) \tag{5.13}$$

and such that for some $J^c\supset\Delta_1\supset J^c\cap(\hat{A}\cup\hat{B})$ we have

$$E\, d(\hat{X}, \hat{Y})\, 1_{J^c\cap\Delta_1^c}\leqq\tilde{d}(P(J^c, X\in\cdot), P(J^c, Y\in\cdot)) \tag{5.14}$$

$$P(\Delta_1)\leqq P(J^c\cap\hat{A})+P(J^c\cap\hat{B}). \tag{5.15}$$

From (5.12) and (5.13) we get conclusion (a). Define Δ to be the union of Δ_1, the events $A_i\triangle B_i$ for $i\geqq 1$, and the events $A_i\cap A_j$ for $i\neq j$ and the events $B_i\cap B_j$ for $i\neq j$. Then conclusion (b) holds by definition. We now have to estimate the quantities in (c) and (d).

First, by hypothesis (ii),

$$P(J)\leqq\sum P(J_i)\leqq\sum P(A_i)\leqq\theta. \tag{5.16}$$

Next, set $N_A=\sum 1_{A_i}$, and similarly for N_B. By definition of Δ,

$$P(\Delta)\leqq P(\Delta_1)+\sum_{I^c}[P(A_i)+P(B_i)]+\sum_I P(A_i\triangle B_i)+P(N_A\geqq 2)+P(N_B\geqq 2). \tag{5.17}$$

But by hypothesis (iii),

$$P(N_A\geqq 2)\leqq EN_A-P(N_A\geqq 1)\leqq\lambda \tag{5.18}$$

and similarly for N_B. Next, since $J_i\subset\hat{A}_i\cap\hat{B}_i$,

$$P(\hat{A}_i\triangle\hat{B}_i)\leqq P(\hat{A}_i\backslash J_i)+P(\hat{B}_i\backslash J_i). \tag{5.19}$$

For each $i\in I$

$$\begin{aligned} P(\hat{A}_i\backslash J_i)&=P(A_i)-P(A_{i,1})\\ &\leqq P(A_i)-(1-2\eta-\varepsilon)\,P(A_i) \quad \text{by (5.10) and hypothesis (i)}\\ &\leqq(2\eta+\varepsilon)\,P(A_i),\end{aligned}$$

and so using (5.16)

$$\sum_I P(\hat{A}_i\backslash J_i)\leqq(2\eta+\varepsilon)\,\theta, \tag{5.20}$$

and similarly for $\hat{B}_i$. Next,

$$P(J^c\cap\hat{A})\leqq\sum_{I^c}P(\hat{A}_i)+\sum_I P(\hat{A}_i\backslash J_i) \tag{5.21}$$

and similarly for $\hat{B}_i$. By hypothesis (iii),

$$\lambda \geqq \sum P(A_i) - P(\bigcup A_i)$$
$$= \sum_{i \geqq 1} P(A_i)\, P(\bigcup_{j<i} A_j | A_i)$$
$$\geqq \eta \sum_{I_A^c} P(A_i)$$

where I_A is the set of i such that $P(\bigcup_{j<i} A_j | A_i) \leqq \eta$. Thus $\sum_{I_A^c} P(A_i) \leqq \lambda \eta^{-1}$, and similarly $\sum_{I_B^c} P(B_i) \leqq \lambda \eta^{-1}$. But $I^c = I_A^c \cup I_B^c$, and so using hypothesis (i)

$$\sum_{I^c} P(A_i) \leqq \lambda \eta^{-1}(2+\varepsilon) \tag{5.22}$$

and similarly for B_i. Combining (5.15), (5.17–5.22) gives conclusion (c). To estimate (d),

$$E\, d(\hat{X}, \hat{Y})\, 1_{\Delta^c} \leqq E\, d(\hat{X}, \hat{Y})\, 1_J + E\, d(\hat{X}, \hat{Y})\, 1_{J^c \cap \Delta^c}. \tag{5.23}$$

$$E\, d(\hat{X}, \hat{Y})\, 1_J = \sum_I E\, d(\hat{X}, \hat{Y})\, 1_{J_i}$$
$$\leqq \sum_I (1-2\eta)^{-1} \alpha P(J_i) \quad \text{by (5.11) and (5.9)}$$
$$\leqq (1-2\eta)^{-1} \alpha \theta \quad \text{using (5.16).} \tag{5.24}$$

And

$$E\, d(\hat{X}, \hat{Y})\, 1_{J^c \cap \Delta^c} \leqq \tilde{d}(P(J^c, \hat{X} \in \cdot), P(J^c, \hat{Y} \in \cdot)) \quad \text{by (5.14), since } \Delta \supset \Delta_1,$$
$$\leqq \tilde{d}(\mathscr{L}(\hat{X}), \mathscr{L}(\hat{Y}))$$
$$+ \tilde{d}(P(J, \hat{X} \in \cdot), P(J, \hat{Y} \in \cdot)) \quad \text{by Lemma 5.8}$$
$$\leqq \tilde{d}(\mathscr{L}(X), \mathscr{L}(Y)) + E\, d(\hat{X}, \hat{Y})\, 1_J. \tag{5.25}$$

Combining (5.23–5.25) gives conclusion (d).

Proof of Lemma 5.6. Construct $(\hat{X}, \hat{Y})$ with marginals $\mathscr{L}(X|A)$, $\mathscr{L}(Y|B)$, and with joint distribution such that $E d(\hat{X}, \hat{Y}) = \tilde{d}(\mathscr{L}(X|A), \mathscr{L}(Y|B))$. Construct $\hat{A}_0, \hat{B}_0$ such that $\mathscr{L}(\hat{X}, \hat{A}_0) = \mathscr{L}((X, A_0|A)$ and $\mathscr{L}(\hat{Y}, \hat{B}_0) = \mathscr{L}((Y, B_0)|B)$. Then

$$P(\hat{A}_0 \cap \hat{B}_0) \geqq 1 - P(\hat{A}_0^c) - P(\hat{B}_0^c) = 1 - P(A_0^c|A) - P(B_0^c|B)$$
$$\geqq 1 - 2\eta \tag{5.26}$$

by hypothesis. Without loss of generality, suppose $P(A) \leqq P(B)$. Let $\hat{A}_1 = \hat{A}_0 \cap \hat{B}_0$ and $\hat{B}_1 = \hat{A}_0 \cap \hat{B}_0 \cap \hat{D}$, where $\hat{D}$ is independent of everything previous, $P(\hat{D}) = P(A)/P(B)$. Then

$$\tilde{d}(\mathscr{L}(\hat{X}|\hat{A}_1), \mathscr{L}(\hat{Y}|\hat{B}_1)) = \tilde{d}(\mathscr{L}(\hat{X}|\hat{A}_0 \cap \hat{B}_0), \mathscr{L}(\hat{Y}|\hat{A}_0 \cap \hat{B}_0))$$
$$\leqq E(d(\hat{X}, \hat{Y})|\hat{A}_0 \cap \hat{B}_0)$$
$$\leqq E\, d(\hat{X}, \hat{Y})/P(\hat{A}_0 \cap \hat{B}_0)$$
$$\leqq (1-2\eta)^{-1}\, \tilde{d}(\mathscr{L}(X|A), \mathscr{L}(Y|B)) \tag{5.27}$$

using (5.26). Finally, construct $A_1 \subset A_0$ and $B_1 \subset B_0$ such that

$$\mathscr{L}((X, A_0, A_1)|A) = \mathscr{L}(\hat{X}, \hat{A}_0, \hat{A}_1) \quad \text{and} \quad \mathscr{L}((Y, B_0, B_1)|B) = \mathscr{L}(\hat{Y}, \hat{B}_0, \hat{B}_1).$$

Then (5.27) gives conclusion (b). And

$$P(A_1) = P(A_1|A)\,P(A) = P(\hat{A}_1)\,P(A) = P(\hat{A}_0 \cap \hat{B}_0)\,P(A) \geqq (1-2\eta)\,P(A)$$

by (5.26); similarly

$$P(B_1) = P(\hat{A}_0 \cap \hat{B}_0 \cap \hat{D}) = P(\hat{A}_0 \cap \hat{B}_0)\,P(\hat{D})\,P(B) = P(A_1),$$

and this gives conclusion (a).

Proof of Lemma 5.7. Construct $(\hat{X}, \hat{Y})$ with marginals μ, ν and joint distribution such that $E\,d(\hat{X}, \hat{Y}) = \tilde{d}(\mu, \nu)$. Construct $\hat{A}, \hat{B}$ such that $\mathscr{L}(\hat{X}|\hat{A}) = \mathscr{L}(X|A)$ and $\mathscr{L}(\hat{Y}|\hat{B}) = \mathscr{L}(Y|B)$. Take $\hat{\Delta} \supset \hat{A} \cup \hat{B}$ such that $P(\hat{\Delta}) = P(\hat{A}) + P(\hat{B})$; unless this sum is greater than 1, in which case the Lemma is trivially true with $\Delta = \Omega$. Construct $\Delta \supset A \cup B$ such that $P(\Delta) = P(\hat{\Delta})$. Construct X and Y on Δ^c such that $\mathscr{L}((X, Y)|\Delta^c) = \mathscr{L}((\hat{X}, \hat{Y})|\hat{\Delta}^c)$. Now X has been defined on $A \cup \Delta^c$, and $P(A \cup \Delta^c, X \in \cdot) = P(\hat{A} \cup \hat{\Delta}^c, \hat{X} \in \cdot) \leqq \mu(\cdot)$, so by Lemma 5.4 we can extend X to all of Ω to make $\mathscr{L}(X) = \mu$. Similarly for Y. Then

$$E\,d(X, Y)\,1_{\Delta^c} = E\,d(\hat{X}, \hat{Y})\,1_{\hat{\Delta}^c} \leqq E\,d(\hat{X}, \hat{Y}) = \tilde{d}(\mu, \nu),$$

which gives the conclusion of the Lemma.

Proof of Lemma 5.8. Conclusion (a) is straightforward. To prove (b), observe first that any distribution can be approximated by distributions uniform on finite sets. Thus we may suppose $\mu_1, \mu_2, \nu_1, \nu_2$ are uniform on $\{x_1, \ldots, x_M\}$, $\{x_{M+1}, \ldots, x_N\}$, $\{y_1, \ldots, y_M\}$, $\{y_{M+1}, \ldots, y_N\}$ respectively. In this case,

$$\tilde{d}(\mu, \nu) = \min\left\{N^{-1} \sum_{i=1}^{N} d(x_i, y_{\pi(i)}) \colon \pi \text{ a permutation of } \{1, \ldots, N\}\right\},$$

because every doubly-stochastic matrix is an average of permutation matrices. Similarly,

$$\tilde{d}(\mu_1, \nu_1) = \min\left\{N^{-1} \sum_{i=1}^{M} d(x_i, y_{\rho(i)}) \colon \rho \text{ a permutation of } \{1, \ldots, M\}\right\}$$

$$\tilde{d}(\mu_2, \nu_2) = \min\left\{N^{-1} \sum_{i=M+1}^{N} d(x_i, y_{\sigma(i)}) \colon \sigma \text{ a permutation of } \{M+1, \ldots, N\}\right\}.$$

Thus it suffices to show that, given permutations π of $\{1, \ldots, N\}$ and ρ of $\{1, \ldots, M\}$, we can construct a permutation σ of $\{M+1, \ldots, N\}$ such that

$$\sum_{i=M+1}^{N} d(x_i, y_{\sigma(i)}) \leqq \sum_{i=1}^{N} d(x_i, y_{\pi(i)}) + \sum_{i=1}^{M} d(x_i, y_{\rho(i)}). \tag{5.28}$$

To do the construction, fix $M < i \leqq N$. Define inductively

$$j_{i,1}=i$$
$$j_{i,2n}=\pi(j_{i,2n-1}); \qquad j_{i,2n+1}=\rho^{-1}(j_{i,2n}),$$

terminating at $t_i=\min\{2n: j_{i,2n}>M\}$. Let $\sigma(i)=j_{i,t_i}$. It is easy to check

(i) σ is indeed a permutation of $\{M+1,\ldots,N\}$.

(ii) The ordered pairs $(j_{i,2n-1}, j_{i,2n})$, $2n\leqq t_i$, $M<i\leqq N$ are distinct elements of the set $\{(i,\pi(i)): 1\leqq i\leqq N\}$.

(iii) The ordered pairs $(j_{i,2n}, j_{i,2n+1})$, $2n<t_i$, $M<i\leqq N$, are distinct elements of the set $\{(i,\rho^{-1}(i)): 1\leqq i\leqq M\}$.

But this establishes (5.28), since by the triangle inequality

$$d(x_i, y_{\sigma(i)})\leqq \sum_{2n\leqq t_i} d(x_{j_{i,2n-1}}, y_{j_{i,2n}})+\sum_{2n<t_i} d(y_{j_{i,2n}}, x_{j_{i,2n+1}}).$$

6. Proof of Theorem 1.8

We shall construct Brownian motion $W(t)$ and the normalized partial sum processes $S^*(t)$, together with their self-intersections U_i, T_i^*, such that as $K\to\infty$ the sample paths $W(\cdot)$ and $S^*(\cdot)$ become close, and the positions and times U_i, T_i^* of self-intersections become close. The construction will be done inductively over successive intervals of length δ. That is, having constructed $(W(t), S^*(t);\ t\leqq m\delta)$ we shall use Proposition 5.2 to describe a joint distribution for $(W(t), S^*(t);\ m\delta\leqq t\leqq(m+1)\delta)$ and their self-intersections conditioned on the past processes $w(\cdot)$, $S^*(\cdot)$, on $[0, m\delta]$. As a preliminary, we describe the conditional increments of the processes separately.

Let $\delta>0$, $m\geqq 1$ and $K\geqq 1$ be fixed until further notice. Let $\sigma=\sigma_K$. To ease typography, we write $\delta\sigma^{2/3}$ where we should write the integer $[\delta\sigma^{2/3}]$.

(6.1) *The Conditional Increment of* (W, U_i)

Fix a path $w(\cdot)$ on $[0, m\delta]$. Condition on $W(\cdot)=w(\cdot)$ on $[0, m\delta]$, and let $(\hat{W}(t),\ 0\leqq t\leqq\delta;\ \hat{U}_1, \hat{U}_2, \ldots)$ denote the conditional distribution of $W(m\delta+t)$ and of the pseudo self-intersections $U-m\delta$ for which $U\in[m\delta, (m+1)\delta]$.

Of course $\hat{W}$ is just Brownian motion started with $\hat{W}(0)=w(m\delta)$. The point process $\hat{U}$ can be regarded as the superposition of two independent processes U^0, U^+, where U^0 represents the pseudo-intersections of $\hat{W}$ with $w(\cdot)$, and U^+ represents the pseudo self-intersections of $\hat{W}$. Let $f(x)$, $-\infty<x<\infty$, be the occupation density (i.e. local time at time $m\delta$) of $w(\cdot)$, and write

$$F=\sup_x f(x). \tag{6.2}$$

Conditional on $\hat{W}$, the process U^0 is a Poisson process of rate $f(\hat{W}(t))$. Now we can write $f=\sum f_i$, where $f_i(\cdot)$ is the occupation density of the path segment $(w(t);\ i\sigma^{-2/3}\leqq t\leqq(i+1)\sigma^{-2/3})$. Thus we can regard U^0 as the superposition of processes U^i, where conditional on $\hat{W}$ the processes U^i are independent Poisson processes of rates $f_i(\hat{W}(t))$.

Let I_K be the index set of all (i,j) such that $0 \leqq i \leqq m\,\delta\sigma^{2/3}$ and $0 \leqq j \leqq \delta\sigma^{2/3}$. For such (i,j) let $B_{i,j}$ be the event that some point of the process U^i occurs during the interval $[j\,\sigma^{-2/3}, (j+1)\,\sigma^{-2/3}]$.

(6.3) *The Conditional Increment of* (S, T_i)

Fix a path $s(i)$, $0 \leqq i \leqq m\,\delta\sigma^{2/3}$, which is a possible path of the random walk S_i. Condition on $S_i = s(i)$, $i \leqq m\,\delta\sigma^{2/3}$, and let $(\hat{S}(i),\ 0 \leqq i \leqq \delta\sigma^{2/3};\ \hat{T}_1, \hat{T}_2, \ldots)$ denote the conditional distribution of $S(m\,\delta\sigma^{2/3}+i)$ and of the self-intersections $T - m\,\delta\sigma^{2/3}$ of S for which $T \in [m\,\delta\sigma^{2/3}, (m+1)\,\delta\sigma^{2/3}]$.

Of course $\hat{S}$ is just the random walk started with $\hat{S}(0) = s(m\,\delta\sigma^{2/3})$. The point process $\hat{T} = (\hat{T}_i)$ is the superposition of two processes T^0, T^+, where T^0 represents the intersections of $\hat{S}$ with $s(\cdot)$ and T^+ represents the self-intersections of $\hat{S}$.

For $(i,j) \in I_K$ let $A_{i,j}$ be the event that $\hat{S}_j = s(i)$. Let $s^*(t)$ and $S^*(t)$ denote the normalized paths, as at (1.4).

(6.4) *The Conditional Joint Increments*

Now suppose we are given a path $w(\cdot)$ as in (6.2), and for each K we are given a path $s(i)$ as in (6.3), such that the normalized paths $s^*(t)$ satisfy

$$s^*(\cdot) \to w(\cdot) \quad \text{in } D[0, m\,\delta] \text{ as } K \to \infty. \tag{6.5}$$

We shall use Proposition 5.2 to construct for each K a joint distribution for $(\hat{S};\ A_{i,j},\ (i,j) \in J_K)$ and $(\hat{W};\ B_{i,j},\ (i,j) \in J_K)$, where $J_K \subset I_K$ will be specified later (6.14), and events Δ_1 (depending on K) such that:

the marginal distributions are as specified in (6.1) and (6.3); (6.6) (a)

outside Δ_1 we have: $A_{i,j} = B_{i,j}$ for $(i,j) \in J_K$, and these events are disjoint as (i,j) varies; (b)

$$\limsup_{K \to \infty} P(\Delta_1) \leqq 16(\delta F)^{3/2} + 2(\delta F)^2 \tag{c}$$

for F defined at (6.2);

$$\lim_{K \to \infty} E\, d(\hat{S}^*, \hat{W})\, 1_{\Delta_1^c} = 0 \tag{d}$$

where d is a bounded metric on $D[0, \delta]$.

The details of this construction will be given later (6.14).

Next, by Lemma 5.5 we can splice these joint distributions into one collection encompassing $\hat{W}$ and all the random walks $\hat{S}^{(K)}$, $K \geqq 1$. Next, for each K the family $(\hat{W};\ B_{i,j},\ (i,j) \in J_K)$ constructed above can be extended by Lemma 5.3 to a family $(\hat{W};\ B_{i,j},\ (i,j) \in I_K;\ \hat{U}_1, \hat{U}_2, \ldots)$ with the distribution prescribed in (6.1). Similarly, for each K we can extend $(\hat{S};\ A_{i,j},\ (i,j) \in J_K)$ to a family $(\hat{S};\ A_{i,j},\ (i,j) \in I_K;\ \hat{T}_1, \hat{T}_2, \ldots)$ with the distribution prescribed in (6.3).

This finishes the construction of the conditional joint increments. We now wish to estimate the distance between the conditioned processes $\hat{S}^*$ and $\hat{W}$, and between the conditioned times $\hat{U}_i$ and $\hat{T}_i^*$. Here are some Lemmas, to be proved later, of similar format: we define a "bad" event Δ (depending on K), and give a bound on $P(\Delta)$ as $K \to \infty$.

Lemma 6.7. *Let* $\Delta_2 = \bigcup_{(i,j)\in I_K \setminus J_K} (A_{i,j} \cup B_{i,j})$. *Then* $P(\Delta_2) \to 0$ *as* $K \to \infty$.

Lemma 6.8. *Let* Δ_3 *be the event that either the process* T^+ *of* (6.3) *or the process* U^+ *of* (6.1) *has at least one point. Then* $\limsup_{K\to\infty} P(\Delta_3) \leqq 2\delta^{3/2}$.

Lemma 6.9. *Let* Δ_4 *be the event that the process* U^0 *of* (6.1) *has at least two points during an interval* $[j\sigma^{-2/3}, (j+1)\sigma^{-2/3}]$, *for some* $j \leqq \delta\sigma^{2/3}$. *Then* $P(\Delta_4) \to 0$ *as* $K \to \infty$.

Finally, it is immediate from (6.6)(d) that we can find events Δ_5 such that as $K \to \infty$ we have $d(S^*, W)\, 1_{\Delta_1^c \cap \Delta_5^c} \to 0$ a.s. and $P(\Delta_5) \to 0$. Now set $\Delta = \bigcup_{q=1}^{5} \Delta_q$. Combining the estimates in (6.6)–(6.9) gives the following bounds on the distance between the conditioned processes.

Outside Δ we have $|T_i^* - U_i| \leqq \sigma^{-2/3}$ for all i; (6.10)(i)

$d(S^*, W)\, 1_{\Delta^c} \to 0$ a.s. as $K \to \infty$; (ii)

$\limsup_{K\to\infty} P(\Delta) \leqq 16(\delta F)^{2/3} + 2(\delta F)^2 + 2\delta^{3/2}$. (iii)

Here (ii) and (iii) are clear. Property (i) comes from (6.6)(b). For matching events $A_{i,j}$ and $B_{i,j}$ is equivalent to pairing off self-intersection times T^*, U such that each pair falls into some interval $[j\sigma^{-2/3}, (j+1)\sigma^{-2/3}]$; and the other Δ_q are defined so that outside Δ no other self-intersections occur.

(6.11) *The Entire Processes*

In the previous section we have constructed Brownian motion $W(t)$ and its pseudo self-intersections on the time interval $m\delta \leqq t \leqq (m+1)\delta$; the random walks S_i and their self-intersections on the interval such that the normalized walks $S^*(t)$ have time interval $m\delta \leqq t \leqq (m+1)\delta$; all this conditional on past paths $s^*(t)$, $w(t)$, $0 \leqq t \leqq m\delta$ and supposing these paths were convergent as $K \to \infty$. Now by (6.10)(ii) the paths $S^*(\cdot)$ converge to $W(\cdot)$ on $m\delta \leqq t \leqq (m+1)\delta$ if we except certain "bad" sets Δ. Thus by induction on m we can construct the processes over all time $0 \leqq t < \infty$ (the first step being similar to the general step). Then (6.10) gives the following estimates.

Proposition 6.12. *Let* $\delta > 0$ *be fixed. We can construct the Brownian motion process* $(W; U_1, U_2, \ldots)$; *the normalized random walk processes* $(S^*; T_1^*, T_2^*, \ldots)$ *for all* $K \geqq 1$; *and events* $\Delta(K, m)$ *such that*

(i) *outside* $\Delta(K, m)$ *we have* $|T_i^* - U_i| \leqq \sigma^{-2/3}$ *for all* i *such that* $U_i \leqq m\delta$;

(ii) $\sup_{t\leqq m} |S^*(t) - W(t)|\, 1_{\Delta^c(K,m)} \to 0$ a.s. *as* $K \to \infty$ *(*m *fixed)*;

(iii) $\limsup_{K\to\infty} P(\Delta(K, m+1) \mid \mathscr{F}_{\delta m})\, 1_{\Delta^c(K,m)} \leqq \alpha(\delta, m\delta)$ a.s.,

where $\alpha(\delta, t) = 16(\delta F_t)^{3/2} + 2(\delta F_t)^2 + 2\delta^{3/2}$, $F_t = \sup_x L(t, x)$, *and* $\mathscr{F}_t$ *is the* σ*-field generated by the processes* S^*, W *on* $[0, t]$;

(iv) $\Delta(K, m) \in \mathscr{F}_{\delta m}$ *and* $\Delta(K, m) \subset \Delta(K, m+1)$.

This Proposition almost completes the proof of Theorem 1.8. Fix t. By considering (i) and (ii) above for $m=[t/\delta]$, we see that to prove Theorem 1.8 it suffices to prove

$$\lim_{\delta\downarrow 0}\limsup_{K\to\infty} P(\Delta^c(K,[t/\delta]))=0. \tag{6.13}$$

(Note that $\Delta(K,m)$ depends also on δ, though the notation suppresses this.) To prove (6.13), fix $\varepsilon>0$. By (iii) and the fact that $\alpha(\delta, m\delta)\in\mathscr{F}_{\delta m}$,

$$\limsup_{K\to\infty} P(\Delta(K,m+1)\setminus\Delta(K,m),\,\alpha(\delta,m\delta)\leqq\varepsilon\delta\,|\,\mathscr{F}_{\delta m})\leqq\varepsilon\delta.$$

Since $\alpha(\delta,t)$ is increasing in t we have, for $m\leqq t/\delta$,

$$\limsup_{K\to\infty} P(\Delta(K,m+1)\setminus\Delta(K,m),\,\alpha(\delta,t)\leqq\varepsilon\delta)\leqq\varepsilon\delta.$$

Summing over $m\leqq t/\delta$,

$$\limsup_{K\to\infty} P(\Delta(K,[t/\delta]),\,\delta^{-1}\alpha(\delta,t)\leqq\varepsilon)\leqq m\varepsilon\delta\leqq\varepsilon t.$$

But from the definition of α we see that $\delta^{-1}\alpha(\delta,t)\to 0$ as $\delta\to 0$, so

$$\lim_{\delta\downarrow 0}\limsup_{K\to\infty} P(\Delta(K,[t/\delta]))\leqq\varepsilon t.$$

Since ε is arbitrary, this proves (6.13).

(6.14) *Details of the Construction of Joint Increments*

The proof of Theorem 1.8 is now complete except for the details of the construction which yields properties (6.6) and Lemmas 6.7-6.9. We return to the setting of (6.4). We are given paths $s^*(\cdot)\to w(\cdot)$ in $D[0,m\delta]$, and we have the conditional increments of the separate processes prescribed by (6.1) and (6.3). Fix $L<\infty$ and $0<\tau<\delta$. Let J_K be the set of $(i,j)\in I_K$ such that

$$j\geqq\tau\sigma^{2/3}\quad\text{and}\quad |s^*(i\sigma^{-2/3})-s^*(m\delta)|\leqq L. \tag{6.15}$$

Proposition 6.16. *As* $K\to\infty$,

(i) $\hat{S}^*\xrightarrow{\mathscr{D}}\hat{W}$ *on* $D[0,\delta]$.

(ii) $\sup_{J_K}|P(A_{i,j})/P(B_{i,j})-1|\to 0$.

(iii) $\sup_{J_K}\tilde{d}(\mathscr{L}(\hat{S}^*\,|\,A_{i,j}),\mathscr{L}(\hat{W}\,|\,B_{i,j}))\to 0$.

(iv) $\limsup P(\bigcup_{I_K\setminus J_K}A_{i,j})\leqq P(\sup_{t\leqq\delta}|W(t)|\geqq L)+F\tau$.

(v) $\limsup P(\bigcup_{I_K\setminus J_K}B_{i,j})\leqq P(\sup_{t\leqq\delta}|W(t)|\geqq L)+F\tau$.

(vi) $\limsup\sum_{I_K}P(A_{i,j})\leqq P(\bigcup_{I_K}A_{i,j})+(F\delta)^2$.

(vii) $\limsup\sum_{I_K}P(B_{i,j})\leqq P(\bigcup_{I_K}B_{i,j})+(F\delta)^2$.

(viii) $\limsup \sum_{I_K} P(A_{i,j}) \leqq F\,\delta$.

(ix) $\limsup \sum_{I_K} P(B_{i,j}) \leqq F\,\delta$.

We defer the proof for a moment. Proposition 6.16 holds with J_K defined at (6.15) for fixed L, τ. So we can redefine J_K by taking $L_K \to \infty$ and $\tau_K \to 0$ sufficiently slowly, and the assertions of the Proposition will remain true except that (iv), (v) are improved to

(x) $$P\Big(\bigcup_{I_K \setminus J_K} A_{i,j}\Big) \to 0; \qquad P\Big(\bigcup_{I_K \setminus J_K} B_{i,j}\Big) \to 0.$$

This gives Lemma 6.7. And we can apply Proposition 5.2 to $(\hat{S}; A_{i,j}, (i,j) \in J_K)$ and $(\hat{W}: B_{i,j}, (i,j) \in J_K)$. For Proposition 6.16 shows that the hypotheses of Proposition 5.2 hold with (as $K \to \infty$)

$$\varepsilon \to 0$$
$$\lambda \leqq (F\,\delta)^2$$
$$\theta \leqq F\,\delta$$
$$\alpha \to 0.$$

Applying Proposition 5.2 with $\eta = (F\,\delta)^{\frac{1}{2}}$, we obtain a construction for the joint distribution which satisfies (6.6). (Proposition 5.2 requires $\eta < \frac{1}{2}$: if not, assertion (6.6) is trivial for $\Delta_1 = \Omega$.)

Proof of Proposition 6.16. Assertion (i) follows from (1.6). Assertions (vi)-(ix) are rephrasings or easy consequences of Lemmas 4.10 and 4.11.

Proof of (iv). $\bigcup_{I_K \setminus J_K} A_{i,j} \subset D_1 \cup D_2$, where

$$D_1 = \{\sup_{t \leqq \delta} |\hat{S}^*(t) - \hat{S}^*(0)| \geqq L\}$$
$$D_2 = \{\hat{S}_j \in G \text{ for some } 1 \leqq j \leqq \tau\,\sigma^{2/3}\}$$
$$G = \{s(i):\ 0 \leqq i \leqq m\,\delta\,\sigma^{2/3}\}.$$

Now $P(D_1) \to P(\sup_{t \leqq \delta} |W(t)| \geqq L)$ by (1.6). And $\limsup_{K \to \infty} P(D_2) \leqq F\,\tau$ by Lemma 4.11(a).

Proof of (v). Similarly, $\bigcup_{I_K \setminus J_K} B_{i,j} \subset D_1 \cup D_2$, where

$$D_1 = \{\sup_{t \leqq \delta} |\hat{W}(t) - \hat{W}(0)| \geqq L - 2d_K\}$$
$$D_2 = \{\text{some point of } U^0 \text{ occurs before time } \tau\}$$
$$d_K = \sup_{t \leqq m\delta} |s^*(t) - w(t)|.$$

Now $P(D_1) \to P(\sup_{t \leqq \delta} |W(t)| > L)$ since $d_K \to 0$. And $\limsup_{K \to \infty} P(D_2) \leqq F\,\tau$ by Lemma 4.10(a).

Proof of (ii). $P(A_{i,j})=P(\hat{S}_j=s(i))=P(S_j=s(i)-s(m\,\delta\,\sigma^{2/3}))$.

So using Lemma 4.6(a),

$$\sup_{J_K}|\sigma^{4/3}P(A_{i,j})-\phi(j\,\sigma^{-2/3},\sigma^{-4/3}(s(i)-s(m\,\delta\,\sigma^{2/3})))|\to 0.$$

Since $s^*(\cdot)\to w(\cdot)$,

$$\sup_{J_K}|\sigma^{4/3}P(A_{i,j})-\phi(j\,\sigma^{-2/3},w(i\,\sigma^{-2/3})-w(m\,\delta))|\to 0. \tag{6.17}$$

Now consider $B_{i,j}$, the event that the point process U^i of (6.1) has some point in the time interval $j\sigma^{-2/3}\leqq t\leqq(j+1)\sigma^{-2/3}$. Since U^i has rate bounded by F, we easily get

$$P(B_{i,j})\leqq m_{i,j}\leqq P(B_{i,j})/(1-F\,\sigma^{-2/3}), \tag{6.18}$$

where $m_{i,j}$ is the mean number of points in that interval. But

$$m_{i,j}=\int_{-\infty}^{\infty}dx\int_{j\sigma^{-2/3}}^{(j+1)\sigma^{-2/3}}dt\,f_i(x)\,\phi(t,x-w(m\,\delta)). \tag{6.19}$$

As $K\to\infty$ the probability measure with density $\sigma^{2/3}f_i(x)\,dx$ approximates the probability measure degenerate at $w(i\,\sigma^{-2/3})$; and the probability measure with density $\sigma^{2/3}\,1_{(j\sigma^{-2/3}\leqq t\leqq(j+1)\sigma^{-2/3})}dt$ approximates the measure degenerate at $j\,\sigma^{-2/3}$. Taking $K\to\infty$ in (6.19), a weak convergence argument shows

$$\sup_{J_K}|\sigma^{4/3}m_{i,j}-\phi(j\,\sigma^{-2/3},w(i\,\sigma^{-2/3})-w(m\,\delta))|\to 0. \tag{6.20}$$

Putting together (6.18), (6.17) and (6.20) gives

$$\sup_{J_K}\sigma^{4/3}|P(A_{i,j})-P(B_{i,j})|\to 0.$$

This yields (ii), since (6.17) shows $\liminf_{K\to\infty}\inf_{J_K}\sigma^{4/3}P(A_{i,j})>0$.

Proof of (iii). The proof is similar to the proof of (ii) above, using Lemma 4.8.

This concludes the proof of Proposition 6.16.

Proof of Lemma 6.9. The process U^0 in (6.1) is stochastically dominated by the homogeneous Poisson process of rate F. So for fixed j, the chance of more than one point of U^0 during $j\sigma^{-2/3}\leqq t\leqq(j+1)\sigma^{-2/3}$ is at most $(F\sigma^{-2/3})^2$. Thus

$$P(\Lambda_4)\leqq\delta\,\sigma^{2/3}[F\,\sigma^{-2/3}]^2\to 0 \quad \text{as } K\to\infty.$$

Proof of Lemma 6.8. The process U^+ of (6.1) is distributed as the process U of (1.7), restricted to the time interval $(0,\delta)$. So the chance of at least one point of U^+ existing is bounded by the mean number of points

$$\begin{aligned} E\int_0^{\delta}L(t,W(t))\,dt&=\tfrac{1}{2}\delta^{3/2}EY &&\text{by Lemmas 2.2 and 2.3}\\ &\leqq\delta^{3/2} &&\text{using Proposition 2.4.}\end{aligned}$$

Similarly, the process T^+ of (6.3) is distributed as the process T of self-intersections of the original random walk S, restricted to the time interval $0 \leqq i \leqq \delta \sigma^{2/3}$. So the probability of some self-intersection is bounded by the mean number of self-intersections, which can be bounded by

$$\delta \sigma^{2/3} m(K, \delta \sigma^{2/3}, 0) \to \delta G(\delta, 0) \quad \text{as } K \to \infty, \text{ by Lemma 4.6(b)}$$
$$= (2/\pi)^{\frac{1}{2}} \delta^{3/2} \leqq \delta^{3/2}.$$

7. Related Topics

(A) Iterates of Random Functions

Given a function f from a set S to itself, and $x \in S$, consider the iterates of f:

$$x_0 = x$$
$$x_{i+1} = f(x_i).$$

Let $t = t(x, f) = \infty$ if the sequence (x_i) has all elements distinct,

$$= \min\{t\colon\ x_t = x_r \text{ for some } r < t\} \text{ otherwise,}$$

and in the latter case let $r(x, f)$ be the integer r such that $x_r = x_t$. Plainly when $t < \infty$ the sequence (x_i) eventually cycles through $(x_r, x_{r+1}, \dots, x_{t-1})$. Call t the *cycle time*. If we now consider a random function F and fixed x, then $X_i = F(X_{i-1})$, $T = t(x, F)$, $R = r(x, F)$ are random variables.

For a finite set S, $\# S = N$ say, we can consider the case where F is uniform over the set of all N^N functions from S to S. In this case the problem of studying the cycle time T $(= T^{(N)})$ is just the classical "Birthday Problem", since it makes no difference if we take X_0 to be uniform on S, and then $X_0, X_1, X_2, \dots$ are i.i.d. uniform on S until $T = \min\{t\colon\ X_t = X_r \text{ for some } r < t\}$. So

$$P(T^{(N)} = t) = N^{-1} t \prod_{i=1}^{t-1} (1 - i/N)$$

$$P(R^{(N)} = r \mid T^{(N)} = t) = 1/t, \qquad 0 \leqq r < t.$$

And as $N \to \infty$ we have $N^{-\frac{1}{2}}(R^{(N)}, T^{(N)}) \xrightarrow{\mathscr{D}} (R^*, T^*)$, where

$$P(T^* > t) = \exp(-t^2/2) \tag{7.1}$$

and conditional on T^*, R^* is uniform on $(0, T^*)$.

Suppose now we wish to study random functions on the integers. Then we cannot pick a function uniformly. However, for fixed K we can define a random function $F\colon \mathbb{Z} \to \mathbb{Z}$ such that the variable $F(i)$ is uniform on $\{i-K, \dots, i+K\}$, and independent for different i. Informally, F is uniform on the set of f with $|f(i) - i| \leqq K$ for all i. In this setting the iterates $X_i = F(X_{i-1})$, $X_0 = 0$, from precisely the random walk (S_i) studied in this paper, with $\xi_i^{(K)}$ uniform, up until the time of first self-intersection. Thus the cycle time $T^{(K)}$ is distributed as the first self-intersection time at (1.11), and so

$$3^{1/3} K^{-2/3} T^{(K)} \xrightarrow{\mathscr{D}} U_1. \tag{7.2}$$

Many different problems concerning the uniform random function on a finite set have been studied by different authors: see e.g. Knuth (1981), p. 8 and 518–520, Pavlov (1981), Pittel (1983). The same questions can be asked about the random functions $F^{(K)}$ of the integers defined above. Some of these questions may be attacked using the methods of this paper: we shall state some results without giving details.

One simple result concerns the analogue of (7.1).

Proposition 7.3. $3^{1/3} K^{-2/3}(R^{(K)}, T^{(K)}) \xrightarrow{\mathscr{D}} (R^*, U_1)$, *where the limiting joint distribution is specified by* (1.13) *and*

$$P(R^* \leqq t \mid W, U_1) = L(t, W(U_1))/L(U_1, W(U_1)), \qquad 0 \leqq t \leqq U_1.$$

In other words, conditional on (W, U_1), R^* is uniform with respect to local time at $W(U_1)$ on the time interval $(0, U_1)$.

Another type of problem concerns components of random functions. For a function f and initial point x, the *component* $C(x, f)$ is the set of points y such that the sequence of iterates of f started from y has the same ultimate cycle as the sequence started from x. In other words, it is the component of the directed graph which has an edge (i, j) if $f(i) = j$. Let $C^{(N)}$ be the component of the uniform random function F of $\{1, \ldots, N\}$ which contains 1. Then

$$\begin{aligned} &\text{(a)} \quad P(j \in C^{(N)}) \to 2/3 \quad \text{as } N \to \infty \ (j \neq 1) \\ &\text{(b)} \quad N^{-1} E \# C^{(N)} \to 2/3 \text{ as } N \\ &\text{(c)} \quad N^{-1} \# C^{(N)} \xrightarrow{\mathscr{D}} C^*, \quad \text{where } C^* \text{ has density } f(c) = \tfrac{3}{2}(1-c)^{\frac{1}{2}},\ 0 < c < 1. \end{aligned} \tag{7.4}$$

(These results are sketched in Aldous (1985a), Sect. 11, but probably have appeared previously in the literature.) Now let $C^{(K)}$ be the component of the random function $F^{(K)}$ of the integers containing 0. We can say a little about the asymptotic behavior of $C^{(K)}$, analogous to (7.4).

Fix $x \in \mathbb{R}$. Let W_1, W_2 be independent Brownian motions started at 0, x, and let L_1, L_2 be their local times. Conditional on (W_1, W_2) let $T_{i,j}$ $(i, j = 1, 2)$ be independent, with $T_{i,j}$ distributed as the first event of a Poisson process of rate $L_j(t, W_i(t))$. Let

$$\phi(x) = P(T_{1,1} > T_{1,2} \text{ or } T_{2,2} > T_{2,1}).$$

Proposition 7.5.

(a) $P(j \in C^{(K)}) \to \phi(x)$ *as* $K \to \infty$, $3^{\frac{1}{2}} K^{-4/3} j \to x$.

(b) $3^{\frac{1}{2}} K^{-4/3} E \# C^{(K)} \to \int_{-\infty}^{\infty} \phi(x)\, dx < \infty$.

It seems hard to obtain any quantitative information about ϕ. Simulations suggest $\phi(0) \simeq 0.3$. Also, it is not clear how to express the limiting distribution (which presumably exists analogously to (7.4c)) of $K^{-4/3} \# C^{(K)}$ in terms of Brownian motions. Heuristically, there is a limit process in which an infinite number of independent Brownian motions start simultaneously from each point $x \in \mathbb{R}$ and interact with each other's local times (c.f. Arratia (1981) for simpler models in this spirit).

(B) Self-Avoiding Walks

There is a large literature on self-avoiding walks, and similar models for polymers, in d dimensions - see e.g. Freed (1981). The 1-dimensional case, though physically less interesting, is still mathematically challenging - see Westwater (1984). In the 1-dimensional setting of Theorem 1.8, let the step distribution be uniform on $\{-K, \ldots, K\}$. Let S_n^A have the distribution of S_n conditioned on $\{T_1 > n\}$. Given $c > 0$, let W_t^A have the distribution of $W(t)$ given $\{U_1^c > t\}$, where U_1^c is the first event of the process (1.7) with rate $c\,L(t, W(t))$.

In the case $K=1$, obviously $|S_n^A| = n$. I *conjecture* that the other conditioned processes also grow linearly; that is,

$$n^{-1} S_n^A \xrightarrow{\mathscr{D}} a_K \varepsilon \quad \text{as } n \to \infty,\ K \text{ fixed,} \tag{7.6}$$

where a_K is constant, $P(\varepsilon = \pm 1) = \frac{1}{2}$;

$$t^{-1} W_t^A \xrightarrow{\mathscr{D}} c^{1/3} \varepsilon \quad \text{as } t \to \infty,\ c \text{ fixed;} \tag{7.7}$$

$$\text{the constants } a_K \text{ in (1) satisfy } a_K \sim 3^{-1/3} K^{2/3} \text{ as } K \to \infty. \tag{7.8}$$

Let us make a few comments on these conjectures. If $\xi_1^{(K)}$ is bounded, the self-avoiding walk $S_0, S_1, S_2, \ldots$ can cross a level L only finitely many times, so it is easy to see that $n^{-1} S_n^A$ *cannot* converge to 0 as $n \to \infty$. This suggests (7.6). I do not see how to prove (7.6), though subadditive ergodic theory might be applicable.

For (7.7) there is a "physicist's argument". Let $(\tilde{W}(s),\ 0 \leqq s \leqq t)$ have the distribution of $(W(s)\colon 0 \leqq s \leqq t)$ given $U^c > t$. By scaling, $(t^{-\frac{1}{2}} \tilde{W}(st)\colon 0 \leqq s \leqq 1)$ has the distribution of $(W(s),\ 0 \leqq s \leqq 1)$ given $U^{ct^{3/2}} > 1$. In physicist's language, Brownian motion $(W(s),\ 0 \leqq s \leqq 1)$ has density (proportional to) $\exp\left\{-\frac{1}{2} \int_0^1 |f'(s)|^2\, ds\right\}$. Arguing as at (1.13), the effect of the conditioning is to multiply this density by $\exp\{-\frac{1}{2} c\, t^{3/2} Y_f\}$, where $Y_f = \int_{-\infty}^{\infty} L_f^2(1, x)\, dx$ and L_f is the occupation density of f. So

$$(t^{-\frac{1}{2}} \tilde{W}(st);\ 0 \leqq s \leqq 1) \quad \text{has density} \propto \exp\left\{-\frac{1}{2} \int_0^1 |f'(s)|^2\, ds - \tfrac{1}{2} c\, t^{3/2}\, Y_f\right\}.$$

Now Y_f scales as $Y_{af} = a^{-1} Y_f$. So dividing by $t^{-\frac{1}{2}}$,

$$(t^{-1} \tilde{W}(st) = 0 \leqq s \leqq 1) \quad \text{has density} \propto \exp\{-\tfrac{1}{2} t\, A(f)\}, \tag{7.9}$$

where

$$A(f) = \int_0^1 |f'(s)|^2\, ds + c\, Y_f.$$

The functional $A(f)$ is minimized by $f_{\pm}(s) = \pm c^{1/3} s$, $0 \leqq s \leqq 1$. So (7.9) suggests that as $t \to \infty$, $(t^{-1} \tilde{W}(st);\ 0 \leqq s \leqq 1)$ converges in distribution to the process degenerate at f_+, f_-, and this gives (7.7). I presume this argument can be formalized, perhaps using the techniques of Westwater (1984).

Finally, Theorem 1.8 says that for fixed K we can approximate $(\sigma^{-4/3} S_{[\sigma^{2/3}t]}, \sigma^{-2/3} T_1)$ by (W_t, U_1^1), where $\sigma = 3^{-\frac{1}{2}} K$. Thus we can approximate $(\sigma^{-4/3} S^A_{[\sigma^{2/3}t]};\ t \geqq 0)$ by $(W_t^A;\ t \geqq 0)$, with $c=1$. Strictly, this approximation holds on any $[0, t_0]$ as $K \to \infty$; if the approximation held uniformly on $[0, \infty)$, then (7.6) and (7.7) would imply (7.8).

(C) Discrete Groups

It is natural to ask what happens to Theorem 1.8 if the random walks are d-dimensional, $d \geqq 2$. More generally, suppose for each K we have a random walk $S_n^{(K)}$ on a discrete group $G^{(K)}$, and suppose that the time $T^{(K)}$ of first self-intersection satisfies $T^{(K)} \xrightarrow[p]{} \infty$. Then we can ask whether there exist normalizing constants a_K such that $T^{(K)}/a_K \xrightarrow{\mathscr{D}} U$ (U non-degenerate). These questions are investigated in Aldous (1985b). Surprisingly, these more abstract problems are usually much *simpler* than the 1-dimensional case treated here. Typically the limit U is either exponential (e.g. in the d-dimensional analogue of Theorem 1.8), or Rayleigh (for rapidly mixing random walks on finite groups).

Acknowledgements. My thanks to Mike Steele for suggesting this problem; to Rick Durrett for the idea in Sect. 3; to Martin Barlow for helpful remarks about the scaling properties of local time: to Persi Diaconis; and to a referee for his careful reading.

References

Aldous, D.J.: Exchangeability and related topics, in: Ecole d'Ete St. Flour 1983, Springer Lecture Notes **1117**. Berlin, Heidelberg, New York: Springer 1985a

Aldous, D.J.: Self-intersections of Random walks on discrete groups. Math. Proc. Camb. Philos. Soc. **98**, 155–177 (1985b)

Arratia, R.: Limiting point processes for rescaling of coalescing and annihilating Random walks on Z^d. Ann. Probab. **9**, 909–936 (1981)

Barlow, M.T.: $L(B_t, t)$ is Not a Semimartingale, in: Seminaire de Probabilites XVI, 209–211, Springer Lecture Notes in Mathematics **920**. Berlin, Heidelberg, New York: Springer 1982

Billingsley, P.: Convergence of probability measures. New York: Wiley 1968

Borodin, A.N.: On the asymptotic behavior of local times of recurrent Random walks with finite variance. Theory Probab. Appl. **26**, 758–772 (1981)

Borodin, A.N.: Distribution of integral functions of the Brownian motion process. LOMI **119**, 19–38 (1982) (Russian. English translation in J. Soviet Math. **27**, 3005–3022)

Freed, K.F.: Polymers as self-avoiding walks. Ann. Probab. **9**, 537–556 (1981)

Gnedenko, B.V., Kolmogorov, A.N.: Limit distributions for sums of independent Random variables. Reading-London: Addison-Wesley 1954

Knuth, D.E.: The Art of computer programming, Vol. 2, 2nd edition. Reading: Addison-Wesley 1981

Pavlov, Y.L.: Limit theorem for a characteristic of a Random mapping. Theory Probab. Appl. **26**, 829–834 (1981)

Perkins, E.: Local time is a semimartingale. Z. Wahrscheinlichkeitstheor. Verw. Geb. **60**, 79–117 (1982a)

Perkins, E.: Weak invariance principles for local time. Z. Wahrscheinlichktstheor. Verw. Geb. **60**, 437–451 (1982b)

Pittel, B.: On the distributions related to transitive classes of Random finite mappings. Ann. Probab. **11**, 428–441 (1983)

Pollard, D.: Convergence of stochastic processes. Berlin, Heidelberg, New York: Springer 1984

Pollard, J.M.: On not storing the path of a Random walk. BIT **19**, 545–548 (1979)

Westwater, J.: On Edwards' model for polymer chains. Proc. of the 4th Bielefeld Conference on Mathematical Physics. Singapore: World Scientific Publishing 1984

Williams, D.: Path decomposition and continuity of local time for one-dimensional diffusions, I. Proc. Lond. Math. Soc. **28**, 738–768 (1974)

Williams, D.: Diffusions, Markov processes and Martingales. New York: Wiley 1979

Received March 7, 1985