PROB. THEORY REL. FIELDS **92**, 465–492 (1992)

Local Times on Curves and Uniform Invariance Principles

Richard F. Bass [1]

Davar Khoshnevisan

Department of Mathematics
University of Washington
Seattle, Washington 98195

Abstract. Sufficient conditions are given for a family of local times $\{L_t^\mu\}$ of d–dimensional Brownian motion to be jointly continuous as a function of t and μ. Then invariance principles are given for the weak convergence of local times of lattice valued random walks to the local times of Brownian motion, uniformly over a large family of measures. Applications include some new results for intersection local times for Brownian motions on $\mathbb{R}^2$ and $\mathbb{R}^3$.

Short title: Local times on curves

[1] Research partially supported by NSF grant DMS–8822053

PROB. THEORY REL. FIELDS **92**, 465–492 (1992)

1. Introduction.

For local times of one–dimensional Brownian motion, there is a huge body of literature for both the modulus of joint continuity and for invariance principles. However, when one turns to d–dimensional Brownian motion, much less is known. Local times at points do not exist, and the appropriate analogue to study is additive functionals L_t^μ corresponding to certain measures μ. For continuity, there are a few results concerning joint continuity in t and μ, such as [B] and [Y]. There are some results on the convergence of functionals of random walks to a single additive functional (see [Dy]), but nothing, as far as we know, on uniform convergence to a family of additive functionals.

The purpose of this paper is to study continuity properties and invariance principles which are uniform over large families $\mathfrak{M}$ of measures μ. We use the term "local times on curves" instead of "additive functionals" because (1) most of the examples we look at have μ's supported on curves and (2) the term "additive functional" is strongly associated with probabilistic potential theory; we make no use of this deep subject, but instead rely on stochastic calculus methods.

Our first set of results concerns the continuity of L_t^μ as a function of t and μ. If $\mathfrak{M}$ is a family of measures μ, each of which satisfies a very mild regularity condition, we show that L_t^μ is jointly continuous in t and μ, even when $\mathfrak{M}$ is a very large family. Largeness, here, is measured by the metric entropy of $\mathfrak{M}$ with respect to a certain metric for the space of measures on $\mathbb{R}^d$ with the topology of weak convergence.

The majority of the paper is concerned with invariance principles. We suppose that $X_1, X_2, \ldots$ is a sequence of mean 0, lattice valued i.i.d. random variables with finite variance, and possibly satisfying additional moment conditions. We let S_n denote the partial sums. We suppose that for each $\mu \in \mathfrak{M}$, there is a sequence of measures μ_n converging weakly to μ. Since the X_i are lattice valued, we suppose the μ_n are supported on $n^{-1/2}\mathbb{Z}^d$. Then, if the μ_n satisfy the same mild regularity condition as we imposed on the μ and the metric entropy of the μ_n is suitably bounded, then the local time for $S_j/\sqrt{n}$ corresponding to μ_n converges weakly to L_t^μ, uniformly over $\mu \in \mathfrak{M}$. The size of the family $\mathfrak{M}$ that is allowed is determined by the number of moments of the X_i.

Although our theorems are quite general, they also seem to be quite powerful, as a number of examples show. For example, in the case of classical additive functionals, where the μ's have densities with respect to Lebesgue measure, we get continuity results and invariance principles over a large class of functions, with minimal smoothness assumptions. If μ is a measure supported on a curve and we approximate μ by curves containing the support of $S_j/\sqrt{n}$, we get an invariance principle that is uniform over a large family of curves.

One of the most interesting examples is that of intersection local times. If $\alpha(x,s,t)$ is the intersection local time of two independent Brownian motions, then α measures the amount of time that the two Brownian motions differ by x, $x \in \mathbb{R}^2$. LeGall [LG] and Rosen [Ro] have shown that the number of intersections of two independent random walks converges to the intersection local time of two independent Brownian motions at a single level x when the random walk has two moments. This result can also be obtained as a corollary of our methods. In addition, if the random walk has $2+\rho$ moments for some $\rho > 0$, we get the new result that weak convergence holds uniformly at all levels x. LeGall and Rosen also have results for invariance principles for k–multiple points. Again, with $2+\rho$ moments, we can get the corresponding uniform invariance principle.

To get some idea of the relative sharpness of our theorems, we look at the case of local times of one–dimensional Brownian motion. A problem that has been studied by a number of people is the question of an invariance principle that is uniform over all the levels x; see [Bo2] and the references therein. As an immediate corollary of our theorems, we get an invariance principle, uniform over all levels x, provided the X_i have $2+\rho$ moments for some $\rho > 0$. The reader should compare this with the results of [Bo1]; there, using techniques highly specific to one–dimensional Brownian motion, the uniform invariance principle is obtained under the assumption of finite variance.

Our results on the joint continuity of local times of curves with respect to t and the measure μ are given in Section 2. We also remark there that many of the results have analogues for symmetric stable processes.

In Section 3 we prove a local central limit theorem. The theorem is that of Spitzer [S]; we derive an estimate of the error term that may be of independent interest.

In Sections 4 and 5, we derive exponential estimates for the tails of the difference of two local times for the random walk. Some of these ideas seem likely to have applications elsewhere: the theme is that if one wants weak convergence or exponential estimates for additive functionals, one only has to compute first moments.

In Section 6, we give our invariance principles, with different versions depending on how well–behaved the tails of the X_i are. The fewer moments, the smaller the family $\mathfrak{M}$ that is allowed. If one has only finite variance, one can still get convergence of the finite dimensional distributions if $d \le 3$, but not (by our techniques) uniform results.

Finally, we give our examples, already discussed above, in Section 7.

2. Construction and Joint Continuity.

Let Z_t be Brownian motion on d–dimensional Euclidean space $\mathbb{R}^d$. Let g be the Green function of Z_t if $d \geq 3$. If $d = 1$ or 2, g shall denote the 1–potential density of Z_t. So

$$g(x,y) = \begin{cases} \int_0^\infty p_s(x,y)ds & \text{if } d \geq 3 \\ \int_0^\infty e^{-s} p_s(x,y)ds & \text{if } d\text{=1 or 2} \end{cases}$$

where $p_s(x,y)$ is the transition function of Z. We define the potential of a measure μ by

$$g\mu(x) = \int g(x,y)\mu(dy).$$

Then it is well–known [Br] that if the map $x \mapsto g\mu(x)$ is bounded and continuous, then there is a continuous additive functional $\{L_t^\mu\}$ so that

$$M_t^\mu = g\mu(Z_t) - g\mu(Z_0) + L_t^\mu \tag{2.1}$$

is a mean zero martingale. L_t^μ is called a local time of Z on the support of μ.

If $\mathfrak{M}$ is a family of positive measures on $\mathbb{R}^d$, define

$$d_G(\mu,\nu) = \sup_{x \in \mathbb{R}^d} |g\mu(x) - g\nu(x)|, \qquad \mu,\nu \in \mathfrak{M}.$$

Define $H_G(\epsilon) = H_G^{\mathfrak{M}}(\epsilon)$ to be the metric entropy of $\mathfrak{M}$ with respect to the norm d_G. In other words, $H_G(\epsilon) = \log N_G(\epsilon)$, where $N_G(\epsilon)$ is the minimum number of d_G–balls of radius ϵ required to cover $\mathfrak{M}$. If

$$H_G(x) \leq c_{2.1} x^{-r}, \qquad x < 1, \tag{2.2}$$

for some r, we say that the exponent of metric entropy of H_G is $\leq r$.

We then have

Proposition 2.1. *If $g\mu$ is bounded and continuous for each $\mu \in \mathfrak{M}$ and if $d_G(\mu,\nu) \leq 1$, then*

$$\mathbb{P}^y(\sup_{t \leq 1} |L_t^\mu - L_t^\nu| \geq \lambda) \leq c_{2.2} \exp(-\lambda / c_{2.3}\sqrt{d_G(\mu,\nu)}), \qquad \mu,\nu \in \mathfrak{M}, y \in \mathbb{R}^d,$$

where $c_{2.3}$ depends only on $\sup_{\mu \in \mathfrak{M}} \|g\mu\|_\infty$.

Proof. Let $U_t^\mu = g\mu(Z_t) - g\mu(Z_0)$ and similarly for U_t^ν. Note $|U_t^\mu - U_t^\nu| \leq 2d_G(\mu,\nu)$. Write N_t for $M_t^\mu - M_t^\nu$. Applying Itô's formula,

$$(U_t^\mu - U_t^\nu)^2 = 2\int_0^t (U_s^\mu - U_s^\nu)\, dN_s - 2\int_0^t (U_s^\mu - U_s^\nu)\, d(L_s^\mu - L_s^\nu) \\ + [U^\mu - U^\nu, U^\mu - U^\nu]_t.$$

Since $[U^\mu - U^\nu, U^\mu - U^\nu]_t$ is $[N,N]_t$, we take expectations to get

$$\begin{aligned}\mathbb{E}^y N_\tau^2 = \mathbb{E}^y [N,N]_\tau &\le 4(d_G(\mu,\nu))^2 + 2d_G(\mu,\nu)\mathbb{E}^y(L_\tau^\mu + L_\tau^\nu) \\ &\le c_{2.4} d_G(\mu,\nu) \end{aligned} \tag{2.3}$$

for bounded stopping times τ.

Consider arbitrary bounded stopping times $T \ge S$. Then denoting the shift operator by θ_t,

$$\begin{aligned}\mathbb{E}^y\left\{|N_T - N_S| \,|\, \mathcal{F}_S\right\} \le \left[\mathbb{E}^y\left\{|N_T - N_S|^2 | \mathcal{F}_S\right\}\right]^{1/2} &= \left[\mathbb{E}^y\left\{[N,N]_T - [N,N]_S | \mathcal{F}_S\right\}\right]^{1/2} \\ &\le \left[\mathbb{E}^y\left\{[N,N]_\infty \circ \theta_S | \mathcal{F}_S\right\}\right]^{1/2} \le \left[\sup_x \mathbb{E}^x [N,N]_\infty\right]^{1/2} \\ &\le c_{2.4}^{1/2} (d_G(\mu,\nu))^{1/2} \qquad \text{(by (2.3).)}\end{aligned}$$

Using (2.1), we get

$$\mathbb{E}^y\{|(L_T^\mu - L_T^\nu) - (L_S^\mu - L_S^\nu)| \,|\mathcal{F}_S\} \le (4 + c_{2.4}^{1/2})(d_G(\mu,\nu))^{1/2}.$$

An application of [DM] p. 193, completes the proof. ☐

Theorem 2.2. *Let $\mathfrak{M}$ be a family of positive measures on $\mathbb{R}^d$. Suppose*

(i) $\sup_{x\in\mathbb{R}^d} \sup_{\mu\in\mathfrak{M}} g\mu(x) < \infty$, *and for all $\mu \in \mathfrak{M}$, $x \mapsto g\mu(x)$ is continuous;*

(ii) *H_G has exponent of metric entropy $< r < 1/2$.*

Then $(t,\mu) \mapsto L_t^\mu$ is almost surely jointly continuous. Moreover,

$$\limsup_{\delta\to 0} \sup_{0\le t\le 1} \sup_{\substack{\mu,\nu\in\mathfrak{M} \\ d_G(\mu,\nu)\le\delta}} \frac{|L_t^\mu - L_t^\nu|}{\delta^{1/2-r}} < \infty, \quad a.s.$$

Remark 2.3. One could give an integral condition that H_G needs to satisfy and also a more precise modulus of continuity, but even in the case of one–dimensional local times our result here is not sharp. This reflects the fact that Proposition 2.1 yields exponential tails for $L_t^\mu - L_t^\nu$ and not Gaussian ones.

Proof of Theorem 2.2. The theorem follows from the estimates of Proposition 2.1 by a standard metric entropy argument (cf. [Du]). ☐

Define another metric on our family of measures $\mathfrak{M}$ by

$$d_L(\mu,\nu) = \sup_{\psi\in\mathfrak{L}} \left| \int \psi d\mu - \int \psi d\nu \right|, \qquad \mu,\nu \in \mathfrak{M},$$

where $\mathfrak{L}$ is the collection of all functions $\psi : \mathbb{R}^d \mapsto \mathbb{R}^1_+$ such that $\|\psi\|_\infty \vee \|\nabla\psi\|_\infty \leq 1$. It is not hard to show that the d_L metric metrizes weak convergence of probability measures. (This metric is equal to what is sometimes known as the *bounded Lipschitz* metric.)

Example 2.4. Suppose $d = 1$. Consider point masses, δ_x and δ_y, on x and y, respectively. Then $d_L(\delta_x, \delta_y) = \sup_{\psi\in\mathfrak{L}} |\psi(x) - \psi(y)| \leq |x - y| \wedge 2$. This is actually an equality. To see this, assume without loss of generality that $y = 0$, $x > y$. Then define

$$\psi_0(\alpha) = (\alpha \vee 0) \wedge 2.$$

Then $\psi_0 \in \mathfrak{L}$ and we have that $|\psi_0(x) - \psi_0(y)| = |x - y| \wedge 2$. So $d_L(\delta_x, \delta_y) \geq |\psi_0(x) - \psi_0(y)|$, and hence we obtain

$$d_L(\delta_x, \delta_y) = |x - y| \wedge 2 \qquad x, y \in \mathbb{R}^1.$$

Example 2.5. Fix two maps $F_i : [0, 1] \to \mathbb{R}^d, i = 1, 2$. Define for all Borel sets A, the measures

$$\mu_i(A) = |\{0 \leq t \leq 1 : F_i(t) \in A\}|, \qquad i = 1, 2,$$

where $|\cdot|$ denotes Lebesgue measure.

Choose $\psi \in \mathfrak{L}$. Then

$$\left|\int \psi d\mu_1 - \int \psi d\mu_2\right| = \left|\int_0^1 \psi(F_1(t))dt - \int_0^1 \psi(F_2(t))dt\right|,$$

and so

$$d_L(\mu_1, \mu_2) \leq \int_0^1 (|F_1(t) - F_2(t)| \wedge 2)\, dt,$$

much as in Example 2.4. The right hand side is equivalent to the L_0-metric corresponding to convergence in measure.

Definition 2.6. Let $\mathfrak{M}$ be a family of positive finite measures on $\mathbb{R}^d$ such that for some $\gamma \in \mathbb{R}^1$ and constant $c_{2.5} = c_{2.5}(\gamma)$,

$$\sup_{\mu\in\mathfrak{M}} \sup_{x\in\mathbb{R}^d} \mu\left(B(x, r)\right) \leq c_{2.5} r^{d-2+\gamma}, \qquad r \leq 1.$$

We call the largest such γ the *index* of $\mathfrak{M}$. If $\mathfrak{M} = \{\mu_0\}$, then we say that γ is the index of μ_0.

Proposition 2.7. *If index$(\mathfrak{M}) > 0$ and $\sup_{\mathfrak{M}} \mu(\mathbb{R}^d) < \infty$, then $\|g\mu\|_\infty < \infty$ and $g\mu(\cdot)$ is continuous for each $\mu \in \mathfrak{M}$. In particular, for every $\mu \in \mathfrak{M}$, L_t^μ is a continuous additive functional.*

Proof. The second statement is a well–known consequence of the first. For the first statement, consider the $d \geq 3$ case first. Then $g(x,y) = c_d|x-y|^{2-d}$ for some c_d. So

$$\begin{aligned} g\mu(x) = \int g(x,y)\mu(dy) &= \int\limits_{B(x,1)} g(x,y)\mu(dy) + \int\limits_{B(x,1)^c} g(x,y)\mu(dy) \\ &\leq c_d\mu(\mathbb{R}^d) + \int\limits_{B(x,1)} g(x,y)\mu(dy), \end{aligned}$$

where $B(x,r)$ is the ball of radius r centered about x.

But if $0 < \gamma < index(\mathfrak{M})$,

$$\begin{aligned} \int\limits_{B(x,1)} g(x,y)\mu(dy) &= \sum_{j=0}^\infty \int_{2^{-(j+1)} \leq |x-y| < 2^{-j}} g(x,y)\mu(dy) \\ &\leq c_d \sum_{j=0}^\infty |2^{-(j+1)}|^{2-d} \mu(B(x, 2^{-j})), \\ &\leq c_{2.6} \sum_{j=0}^\infty 2^{-j\gamma} < \infty. \end{aligned}$$

The $d = 2$ case is similar, since $g(x,y) \leq -c_{2.7} \log|x-y|$ for $|x-y| < 1$. The $d = 1$ case is also easy and is done in a similar fashion.

To show continuity, consider the $d \geq 3$ case again. Then for $\epsilon > 2|x-y|$,

$$\begin{aligned} |g\mu(x) - g\mu(y)| \leq \Big| &\int\limits_{B(x,\epsilon)} g(x,y)\mu(dy) - \int\limits_{B(x,\epsilon)} g(y,z)\mu(dz)\Big| \\ &+ \Big| \int\limits_{B(x,\epsilon)^c} \big(g(x,z) - g(y,z)\big)\mu(dz)\Big| = I_{2.4} + II_{2.4}. \end{aligned} \tag{2.4}$$

The second term is estimated as follows.

$$\begin{aligned} II_{2.4} &\leq \int\limits_{B(x,\epsilon)^c} |g(x,z) - g(y,z)| \\ &\leq c_{2.8}|x-y| \int\limits_{B(x,\epsilon)^c} \big(|x-z| \vee |y-z|\big)^{1-d} \mu(dz) \leq c_{2.9}\epsilon^{1-d}|x-y|. \end{aligned} \tag{2.5}$$

For the first term of (2.4),

$$\begin{aligned}
I_{2.4} &\le \int_{B(x,\epsilon)} (g(x,y)+g(y,z))\,\mu(dz) \\
&= \sum_{j=0}^{\infty} \int_{2^{-(j+1)}\epsilon\le|x-z|<2^{-j}\epsilon} (g(x,z)+g(y,z))\,\mu(dz) \\
&\le 2\sup_{\alpha} \sum_{j=-1}^{\infty} \int_{2^{-(j+1)}\epsilon\le|\alpha-z|<2^{-j}\epsilon} \int g(\alpha,z)\mu(dz) \\
&\le c_{2.10}\sup_{\alpha} \sum_{j=-1}^{\infty} \left(2^{-(j+1)}\epsilon\right)^{2-d} \mu\left(B(\alpha,2^{-j}\epsilon)\right) \\
&\le c_{2.11} \sum_{j\ge -1} 2^{-j\gamma}\epsilon^{\gamma} \le c_{2.12}\epsilon^{\gamma}. \qquad (2.6)
\end{aligned}$$

Putting (2.4), (2.5), and (2.6) together gives the existence of a constant $c_{2.13}$ such that

$$\sup_{|x-y|\le\delta} |g\mu(x)-g\mu(y)| \le c_{2.13}|\delta\epsilon^{1-d}+\epsilon^{\gamma}| \qquad \epsilon > 2\delta.$$

Therefore letting $\epsilon = c_{2.14}\delta^{1/(\gamma+d-1)}$, we get,

$$\begin{aligned}
\sup_{|x-y|\le\delta} |g\mu(x)-g\mu(y)| &\le c_{2.15}\delta^{\gamma/(d+\gamma-1)} \\
&\to 0 \qquad \text{as } \delta\to 0.
\end{aligned}$$

This proves the proposition for $d\ge 3$. The cases when $d\le 2$ are quite similar. □

The following relates the two metrics, d_L and d_G:

Proposition 2.8. *If μ and ν are two positive finite measures on $\mathbb{R}^d$ so that index(μ,ν) $>\gamma>0$, then for some constant $c_{2.16}$ depending only on γ,*

$$d_G(\mu,\nu) \le c_{2.16}[d_L(\mu,\nu)]^{\ell}$$

where $\ell = \gamma/(d+\gamma-1)$.

Proof. Take $d\ge 3$:

$$\begin{aligned}
d_G(\mu,\nu) &= \sup_{x\in\mathbb{R}^d} |g\mu(x)-g\nu(x)| \\
&\le \sup_{x\in\mathbb{R}^d} \Big| \int_{B(x,\epsilon)} g(x,y)(\mu-\nu)(dy)\Big| + \sup_{x\in\mathbb{R}^d} \Big| \int_{B(x,\epsilon)^c} g(x,y)(\mu-\nu)(dy)\Big| \\
&= I_{2.7} + II_{2.7}. \qquad (2.7)
\end{aligned}$$

We proceed to estimate each term on the right hand side of (2.7) separately. Consider the second term first.

$$II_{2.7} = \sup_{x\in\mathbb{R}^d} c_d \Big| \int\limits_{B(x,\epsilon)^c} |x-y|^{2-d}(\mu-\nu)(dy)\Big|.$$

But $\psi_0(y) \equiv |x-y|^{2-d} \wedge \epsilon^{2-d}$ satisfies $\|\psi_0\|_\infty \le \epsilon^{2-d}$ and $\|\nabla\psi_0\|_\infty \le c_{2.17}\epsilon^{1-d}$ for a constant $c_{2.17}$. So

$$\begin{aligned} II_{2.7} &\le c_{2.18} \sup_{x\in\mathbb{R}^d} \left\{ \left| \int \psi d(\mu-\nu)\right| : \|\psi\|_\infty \le \epsilon^{2-d},\ \|\nabla\psi\|_\infty \le c_{2.17}\epsilon^{1-d}\right\} \\ &\le c_{2.19}\epsilon^{1-d} d_L(\mu,\nu), \qquad \epsilon \le 1. \end{aligned} \tag{2.8}$$

We estimate the first term that appears in (2.7) exactly as in (2.6) to get

$$I_{2.7} \le c_{2.20}\epsilon^\gamma. \tag{2.9}$$

Putting (2.9), (2.8), and (2.7) together, and letting $\epsilon = d_L(\mu,\nu)^{1/(d+\gamma-1)}$,

$$\begin{aligned} d_G(\mu,\nu) &\le c_{2.21}\left[\epsilon^\gamma + d_L(\mu,\nu)\epsilon^{1-d}\right] \\ &= c_{2.22}\left[d_L(\mu,\nu)\right]^\ell. \end{aligned}$$

The cases when $d \le 2$ are much the same. □

Now let $H_L(\epsilon)$ be the metric entropy with respect to metric d_L. Then Proposition 2.8 and Theorem 2.1 together yield the following

Theorem 2.9. *Let $\mathfrak{M}$ be a family of positive finite measures on $\mathbb{R}^d$. Assume that index$(\mathfrak{M}) > \gamma$, and let $\ell = \gamma/(\gamma+d-1)$. If the exponent of metric entropy of H_L is $< r < \ell/2$, then almost surely,*

$$\limsup_{\delta\to 0} \sup_{0\le t\le 1} \sup_{\substack{\mu,\nu\in\mathfrak{M} \\ d_L(\mu,\nu)\le\delta}} \frac{|L_t^\mu - L_t^\nu|}{\delta^{\ell/2-r}} < \infty.$$

Remark 2.10. Theorem 2.2 holds for many other Markov processes as well as for Brownian motion. For example, if Z_t is a symmetric stable process of order α the statements and proofs of Proposition 2.1 and Theorem 2.2 go through with only minor changes.

In the stable case, $g(x) = c_{2.23}|x|^{\alpha-d}$. Just as above, $g\mu$ will be continuous and bounded if $\mu(B(x,r)) \le c_{2.24}r^{d-\alpha+\gamma}$ uniformly for $x \in \mathbb{R}^d, r \le 1$. Proposition 2.8 still holds provided we here define ℓ by $\ell = \gamma/(\gamma+d-\alpha+1)$. Similarly, with this change in the definition of ℓ, Theorem 2.9 holds as well.

3. A local central limit theorem.

In this section, we derive a local central limit theorem, which is that of Spitzer [S] pp. 76–78, but we use the additional moments to get better estimates of the error terms. We apply this to the problem of estimating the potential kernel of a random walk (cf. [NS]).

Let $X_1, X_2, \cdots$ be i.i.d. $\mathbb{R}^d$–valued random vectors. Here $X_j = (X_j^1, \ldots, X_j^d)$. We consider the case $d \geq 3$ for a random walk, $S_n = \sum_{j=1}^n X_j$. Assume the X_i's take values in $\mathbb{Z}^d$, are mean 0, have the identity for covariance matrix, are strongly aperiodic, and have finite third moments. Let $\phi(u) = \mathbb{E}\exp(iu \cdot X_1)$, where $a \cdot b$ is the usual inner product. Also let $p_n(x, y) = \mathbb{P}^x\{S_n = y\}$. Then we have the following local central limit theorem:

Proposition 3.1. *There is a constant $c_{3.1}$ such that for all n,*

$$\sup_x |p_n(x,0) - (2\pi n)^{-d/2} e^{-|x|^2/2n}| \leq c_{3.1}(1 + \mathbb{E}|X_1|^3) n^{-(d+1)/2} (\log^+ n)^{(d+3)/2}.$$

(Here $\log^+(n) = \log(n) \vee 1$.).

Proof. We follow the proof of P9 given in [S] pp. 76–78 closely. Let

$$E(n,x) = |p_n(0,x) - (2\pi n)^{-d/2} e^{-|x|^2/2n}|.$$

Then

$$\sup_x (2\pi n)^{d/2} E(n,x) \leq (2\pi)^{-d/2} \sum_{j=1}^4 I_j^{(n)}$$

where

$$I_1^{(n)} = \sup_x \left| \int_{|\alpha| \leq A_n} \left(\phi^n(\alpha n^{-1/2}) - e^{-|\alpha|^2/2} \right) e^{-ix\cdot\alpha/\sqrt{n}} d\alpha \right|,$$

$$I_2^{(n)} = \sup_x \left| \int_{|\alpha| \geq An} e^{-|\alpha|^2/2 - ix\cdot\alpha n^{-1/2}} d\alpha \right|,$$

$$I_3^{(n)} = \sup_x \left| \int_{A_n \leq |\alpha| \leq r\sqrt{n}} \phi^n(\alpha n^{-1/2}) e^{-ix\cdot\alpha/\sqrt{n}} d\alpha \right|, \qquad \text{and}$$

$$I_4^{(n)} = \sup_x \left| \int_{\substack{|\alpha| > r\sqrt{n} \\ \alpha \in \sqrt{n}\mathcal{C}}} \phi^n(\alpha n^{-1/2}) e^{-ix\cdot\alpha/\sqrt{n}} d\alpha \right|.$$

Here $\mathcal{C} = \{x \in \mathbb{R}^d : \max_{i \leq d} |x^i| \leq \pi\}$ is the unit cube of side π. Furthermore, $A_n = \sqrt{2\beta \log n}$ for some large β and $r > 0$ is a constant that is small. We proceed to estimate each term separately. Take $n > 1$.

$$I_1^{(n)} \leq n \int_{|\alpha| \leq A_n} \left| \phi(\alpha n^{-1/2}) - e^{-|\alpha|^2/2n} \right| d\alpha$$

since for all $a, b \in \mathbb{R}^d$,

$$||a|^n - |b|^n| \le n|a-b| \left(|a| \vee |b|\right)^{n-1}.$$

By definition, $A_n/\sqrt{n} \to 0$. So a Taylor expansion implies

$$\phi(\alpha n^{-1/2}) = 1 - \frac{|\alpha|^2}{n} + E_1(\alpha, n)$$
$$e^{-|\alpha|^2/2n} = 1 - \frac{|\alpha|^2}{n} + E_2(\alpha, n)$$

and for all $|\alpha| \le A_n$, $E_i(\alpha, n) \le c_{3.2}(1 + \mathbb{E}|X_1|^3)(|\alpha|^3/n^{3/2})$, $\qquad i = 1, 2$.

Therefore there exists a constant $c_{3.3}$, independent of $\alpha \in \{x \in \mathbb{R}^d : |x| \le A_n\}$, so that

$$\begin{aligned}\sup_{|\alpha|\le A_n} \left|\phi(\alpha n^{-1/2}) - e^{-|\alpha|^2/2n}\right| &\le c_{3.3}(1 + \mathbb{E}|X_1|^3) A_n^3 n^{-3/2} \\ &\le c_{3.3}\sqrt{8}\beta^{3/2}(1 + \mathbb{E}|X_1|^3) n^{-3/2}(\log n)^{3/2}.\end{aligned}$$

Therefore,

$$\begin{aligned}I_1^{(n)} &\le c_{3.4}(1 + \mathbb{E}|X_1|^3) n A_n^d n^{-3/2}(\log n)^{3/2}|B(0,1)| \\ &= c_{3.5}(1 + \mathbb{E}|X_1|^3) n^{-1/2}(\log n)^{(d+3)/2}. \end{aligned} \tag{3.1}$$

Next,

$$I_2^{(n)} \le \int_{|\alpha|\ge A_n} e^{-|\alpha|^2/2} d\alpha \le c_{3.6} n^{-\beta}. \tag{3.2}$$

The upper bound for $I_3^{(n)}$ and $I_4^{(n)}$ is done exactly as in [S]: for r small enough

$$I_3^{(n)} \le 2n^{-\beta}. \tag{3.3}$$

Also, for r small enough, there exists $\delta \in (0,1)$ so that

$$\begin{aligned}I_4^{(n)} &\le (1-\delta)^n \left|\left\{\alpha : |\alpha| > r\sqrt{n};\ \alpha \in \sqrt{n}\mathcal{C}\right\}\right| \\ &\le c_{3.7} n^{-\beta}, \end{aligned} \tag{3.4}$$

where $|B|$ is the Lebesgue measure of the Borel set B. Putting (3.1)–(3.4) together, the proposition is proved. $\square$

Recall that X_1 is subgaussian if there exists $r > 0$ such that for all $t > 0$,

$$\mathbb{E}e^{t|X_1|} \le 2e^{t^2 r}. \tag{3.5}$$

Define the potential kernel for the random walk:

$$G(x,y) = \sum_n p_n(x,y).$$

Recall $d \geq 3$ and hence G is well–defined and finite. Also recall that for some constant c_d, $g(x,y) = c_d|x-y|^{2-d}$. Then the above proposition implies:

Proposition 3.2. *Assume the X's are subgaussian. Then there is a constant $c_{3.8}$ so that for every $x, y \in \mathbb{Z}^d$,*

$$|G(x,y) - g(x,y)| \leq c_{3.8}|x-y|^{1-d}\left(\log^+|x-y|\right)^{1+d}.$$

Proof. By translation, it is enough to do this for $y = 0$. Clearly we shall only need to consider the case $|x| \geq 1$. By Chebyshev's inequality and (3.5), for all $x \in \mathbb{R}^d$,

$$\mathbb{P}\left\{|S_n| \geq |x|\right\} \leq 2\exp(-|x|^2/4nr). \tag{3.6}$$

Let $f(x) = [\frac{|x|^2}{k\log|x|}]$, k large. Then

$$\sum_{n=1}^{f(x)} E(n,x) \leq \sum_{n=1}^{f(x)} p_n(0,x) + \sum_{n=1}^{f(x)} (2\pi n)^{-d/2} e^{-|x|^2/2n}. \tag{3.7}$$

We bound each term on the right hand side of (3.7) separately:

$$\begin{aligned}
\sum_{n=1}^{f(x)} p_n(0,x) &\leq \sum_{n=1}^{f(x)} \mathbb{P}\{|S_n| \geq |x|\} \\
&\leq \sum_{n=1}^{f(x)} 2e^{-|x|^2/4nr} \qquad \text{(by (3.6))} \\
&\leq 2f(x)\exp(-\frac{|x|^2}{4rf(x)}) \leq c_{3.9}\frac{|x|^2}{k\log|x|}\cdot|x|^{-k/4r} \\
&\leq |x|^{1-d} \qquad \text{if } k \text{ is large enough.}
\end{aligned} \tag{3.8}$$

Similarly,

$$\sum_{n=1}^{f(x)} (2\pi n)^{-d/2} e^{-|x|^2/2n} \leq |x|^{1-d} \qquad \text{if } k \text{ is large enough,} \quad |x| \geq 1. \tag{3.9}$$

Then (3.7), (3.8), and (3.9) imply that if k is large enough,

$$\sum_{n=1}^{f(x)} E(n,x) \leq 2|x|^{1-d}. \tag{3.10}$$

Now we estimate $\sum_{n=f(x)}^{\infty} E(n,y)$ as follows:

$$\sum_{n\geq f(x)} E(n,x) \leq c_{3.10} \sum_{n\geq f(x)} n^{-(d+1)/2} (\log^+ n)^{(d+3)/2} \qquad \text{(Proposition 3.1)}$$
$$\leq c_{3.11} |x|^{1-d} (\log^+ |x|)^{d+1}. \tag{3.11}$$

Putting (3.11) and (3.10) together implies

$$\sum_{n\geq 1} E(n,x) \leq c_{3.12} |x|^{1-d} (\log^+ |x|)^{d+1}.$$

This in turn implies

$$\left| G(0,x) - \sum_{n\geq 1} (2\pi n)^{-d/2} e^{-|x|^2/2n} \right| \leq 1 + c_{3.12} |x|^{1-d} (\log^+ |x|)^{d+1}.$$

However, it is easy to show that

$$|\sum_{n\geq 1} (2\pi n)^{-d/2} e^{-|x|^2/2n} - g(x,0)| \leq c_{3.13} |x|^{1-d} (\log^+ |x|)^{d+1}.$$

This proves the proposition. □

Corollary 3.3. *If $d \geq 3$ and the X_i's are subgaussian,*
(a) $G(0,x) \leq c_{3.14}(1 \wedge |x|^{2-d})$;
(b) For each $\beta \in (0,1)$, there exists $c_{3.15} = c_{3.15}(\beta)$ such that

$$|G(0,x) - G(0,y)| \leq c_{3.15} \frac{|x-y|}{(|x| \wedge |y|)^{d-1}} + c_{3.15} \frac{|x-y|^{1-\beta}}{(|x| \wedge |y|)^{d-1-\beta}}. \tag{3.12}$$

Proof.

$$G(0,0) = p_0(0,0) + \sum_{n=1}^{\infty} p_n(0,0) \leq 1 + \sum_{1}^{\infty} c_{3.16} n^{-d/2} \leq c_{3.17}.$$

by Proposition 3.1. So part (a) follows by this equation if $x = 0$ and by Proposition 3.2 if $|x| \geq 1$.

Note that part (b) is trivial if $x = y$ or if $x = 0$ or $y = 0$. So let us exclude these cases. By Proposition 3.2, if $\beta \in (0,1)$,

$$|G(0,x) - g(0,x)| \leq c_{3.18} |x|^{-(d-1-\beta)},$$

and similarly for $|G(0,y) - g(0,y)|$. But

$$|g(0,x) - g(0,y)| \le c_{3.19}|y-x|/(|x| \wedge |y|)^{1-d}.$$

Since $|x-y| \ge 1$, part (b) follows by the triangle inequality. $\square$

Remark 3.4. Note that if $\mathrm{Cov}(X_1) = Q$ for Q any positive definite matrix, Corollary 3.3 still holds. To see this, one merely needs to replace the proofs of Propositions 3.1 and 3.2 with ones where the identity matrix I is replaced by Q (cf. [S]).

Remark 3.5. The assumption that the random walk be strongly aperiodic may be removed by the method of Spitzer.

4. Moment Bounds.

In this section we consider the analogues of some of the results of Section 2 with random walk in place of Brownian motion. Fix n. Let μ_n be a finite measure supported on $n^{-1/2}\mathbb{Z}^d$. Let $\mathfrak{M}_n$ be a family of such measures.

Let us define $index_n(\mathfrak{M}_n)$ to be the largest γ such that there exists $c_{4.1}$ with

$$\mu_n(B(x,s)) \le c_{4.1}s^{d-2+\gamma}, \qquad x \in \mathbb{R}^d, s \in [1/2\sqrt{n}, 1], \mu_n \in \mathfrak{M}_n. \tag{4.1}$$

Note, taking $x \in n^{-1/2}\mathbb{Z}^d$ and $s = 1/2\sqrt{n}$, then in particular

$$\mu_n(\{x\}) \le c_{4.1}n^{1-(d+\gamma)/2} \le c_{4.1}n^{1-d/2}. \tag{4.2}$$

Define

$$L_k^{n,\mu_n} = n^{d/2-1}\sum_{j=0}^{k} \mu_n(\{S_j/\sqrt{n}\}). \tag{4.3}$$

Proposition 4.1. *If $index_n(\mu_n) > \gamma$, then $\sup_x \mathbb{E}^x L_\infty^{n,\mu_n} \le c_{4.2}$, where $c_{4.2}$ depends only on $\mu_n(\mathbb{R}^d), \gamma$, and the constant $c_{4.1}$ of (4.1).*

Proof. By translation invariance, it suffices to suppose $x = 0$.

$$\begin{aligned}
\mathbb{E}^0 L_\infty^{n,\mu_n} &= n^{d/2-1}\sum_{j=0}^{\infty}\sum_{y\in\mathbb{Z}^d} \mu_n(\{y/\sqrt{n}\})p_j(0,y) \\
&= n^{d/2-1}\sum_{y\in\mathbb{Z}^d} G(0,y)\mu_n(\{y/\sqrt{n}\}) \\
&\le c_{4.3}n^{d/2-1}\sum_{k=0}^{\infty}\sum_{2^k\le|y|<2^{k+1}} |y|^{2-d}\mu_n(\{y/\sqrt{n}\}) + n^{d/2-1}G(0,0)\mu_n(\{0\}) \\
&\le c_{4.4}n^{d/2-1}\sum_{k=0}^{\infty} 2^{k(2-d)}\mu_n(B(0,2^{k+1}/\sqrt{n})) + n^{d/2-1}G(0,0)\mu_n(\{0\}) \\
&= I_{4.5} + II_{4.5}
\end{aligned} \tag{4.5}$$

$II_{4.5}$ is bounded using (4.2) and Corollary 3.3(a). For $I_{4.5}$, note

$$\sum_{k=0}^{\infty} 2^{k(2-d)} \mu_n(B(0, 2^{k+1}/\sqrt{n})) = \sum_{2^k \le \sqrt{n}} + \sum_{2^k > \sqrt{n}}$$
$$\le c_{4.1} \sum_{2^k \le \sqrt{n}} 2^{k(2-d)} (2^{k+1}/\sqrt{n})^{d-2+\gamma} + c_{4.5} \sum_{2^k > \sqrt{n}} 2^{k(2-d)}$$
$$\le c_{4.6} n^{1-d/2}. \tag{4.6}$$

□

Corollary 4.2. *Let $y \in n^{-1/2}\mathbb{Z}^d$. If μ_n^r is μ_n restricted to $B(y,r) - \{y\}$, then $\mathbb{E}^y L_\infty^{n,\mu_n^r} \le c_{4.7} r^\gamma$.*

Proof. The proof is very similar to that of Proposition 4.1, except that we may omit $II_{4.5}$ and in estimating $I_{4.5}$ in (4.6), we need only look at $\sum_{2^k \le r\sqrt{n}}$. □

Recalling the definition of d_L from Section 2, notice that

$$d_L(\mu_n, \nu_n) = \sup\{\sum_{y \in \mathbb{Z}^d} \psi(y/\sqrt{n})(\mu_n - \nu_n)(\{y/\sqrt{n}\}) : \psi \in \mathfrak{L}\}.$$

Taking $\psi = c_{4.8}/\sqrt{n}$ at $x \in n^{-1/2}\mathbb{Z}^d$ and 0 on $B(x, 1/2\sqrt{n})^c$, we see

$$\mu_n(\{x\}) - \nu_n(\{x\}) \le c_{4.9} \sqrt{n} d_L(\mu_n, \nu_n). \tag{4.7}$$

Lemma 4.3. *Suppose $\|\psi\|_\infty \le 1$, $\mu(\mathbb{R}^d), \nu(\mathbb{R}^d) \le c_{4.10}$, and $|\psi(x) - \psi(y)| \le |x-y|^\alpha$. Then $|\int \psi(y)(\mu - \nu)(dy)| \le c_{4.11} (d_L(\mu, \nu))^\alpha$.*

Proof. Let φ be a smooth, nonnegative, radially symmetric function with compact support and $\int_{\mathbb{R}^d} \varphi(x)\,dx = 1$. Let $\varphi_\epsilon(x) = \epsilon^{-d}\varphi(x/\epsilon)$, $\psi_\epsilon = \psi * \varphi_\epsilon$ for $\epsilon > 0$.

First,

$$|\psi_\epsilon(x) - \psi(x)| = |\int [\psi(x-y) - \psi(x)]\varphi_\epsilon(y)\,dy|$$
$$\le \int |y|^\alpha \varphi_\epsilon(y)\,dy = \epsilon^\alpha \int |y|^\alpha \varphi(y)\,dy \le c_{4.12}\epsilon^\alpha.$$

Next, let u be a unit vector, $\nabla_u f = \nabla f \cdot u$. Without loss of generality, we may assume $u = (1, 0, \ldots, 0)$. Then

$$\int \nabla_u \varphi_\epsilon(y)\,dy = \int \cdots \int \nabla_u \varphi_\epsilon(y^1, \ldots, y^d)\,dy^1 \ldots dy^d$$
$$= \int \cdots \int 0\,dy^2 \ldots dy^d = 0$$

since φ_ϵ has compact support. So

$$\begin{aligned} |\nabla_u \psi_\epsilon(x)| = |\int \psi(x-y)\nabla_u\varphi_\epsilon(y)\,dy| &= |\int [\psi(x-y)-\psi(x)]\nabla_u\varphi_\epsilon(y)\,dy| \\ &\le \int |y|^\alpha \epsilon^{-(d+1)}\varphi(y/\epsilon)\,dy \le c_{4.13}\epsilon^{\alpha-1}. \end{aligned}$$

Hence,

$$\begin{aligned} |\int \psi d(\mu-\nu)| &\le \int |\psi-\psi_\epsilon|\,d(\mu+\nu) + |\int \psi_\epsilon\, d(\mu-\nu)| \\ &\le c_{4.12}\epsilon^\alpha(\mu(\mathbb{R}^d)+\nu(\mathbb{R}^d)) + c_{4.13}\epsilon^{\alpha-1}d_L(\mu,\nu). \end{aligned}.$$

Now take $\epsilon = d_L(\mu,\nu)$. □

Let $\mathfrak{M}_n$ be a family of measures supported on $n^{-1/2}\mathbb{Z}^d$ with $index_n(\mathfrak{M}_n) > \gamma$. We now obtain:

Proposition 4.4. *For each $\beta > 0$,*

$$\sup_{\mu_n,\nu_n\in\mathfrak{M}_n}\sup_x \mathbb{E}^x(L_\infty^{n,\mu_n} - L_\infty^{n,\nu_n}) \le c_{4.14}(d_L(\mu_n,\nu_n))^{\ell_\beta},$$

where

$$\ell_\beta = \gamma(1-\beta)/(d+\gamma-1-\beta) \tag{4.8}$$

and $c_{4.14}$ depends on $c_{4.1}$, γ, β, and $\sup_{\mathfrak{M}_n}\mu_n(\mathbb{R}^d)$.

Proof. By translation invariance we may suppose $x=0$. Write δ for $d_L(\mu_n,\nu_n)$, ℓ for ℓ_0. Let $G_K(y) = G(0,y)\wedge K$. As in (4.5),

$$\begin{aligned} &\mathbb{E}^0 L_\infty^{n,\mu_n} - \mathbb{E}^0 L_\infty^{n,\nu_n} \\ &\quad = n^{d/2-1}\sum_{y\neq 0}[G(0,y)-G_K(y)](\mu_n-\nu_n)(\{y/\sqrt{n}\}) + n^{d/2-1}\sum_y G_K(y)(\mu_n-\nu_n)(y/\sqrt{n}) \\ &\qquad\quad + n^{d/2-1}[G(0,0)-G_K(0)](\mu_n-\nu_n)(\{0\}) \\ &\quad = I_{4.9} + II_{4.9} + III_{4.9} \end{aligned} \tag{4.9}.$$

Note $G(0,y)-G_K(y) = 0$ if $|y| > c_{4.15}K^{1/(2-d)}$. So, writing $\zeta = K^{1/(2-d)}$,

$$\begin{aligned} I_{4.9} &\le n^{d/2-1}\sum_{0<|y|<c_{4.15}\zeta} G(0,y)(\mu_n+\nu_n)(\{y/\sqrt{n}\}) \\ &= 2(\mathbb{E}^0 L_\infty^{n,\mu_n^r} + \mathbb{E}^0 L_\infty^{n,\nu_n^r}) \le 4c_{4.16}r^\gamma, \end{aligned}$$

where $r = c_{4.15}\zeta/\sqrt{n}$.

¿From Corollary 3.3 it follows easily that if $x, y \in n^{-1/2}\mathbb{Z}^d$, then

$$|G_K(x\sqrt{n}) - G_K(y\sqrt{n})| \le c_{4.17} n^{1-d/2} \left(\frac{|x-y|^{1-\beta}}{\zeta^{d-1-\beta}} + \frac{|x-y|}{\zeta^{d-1}} \right)$$

and

$$|G_K(x\sqrt{n})| \le c_{4.18} n^{1-d/2} \zeta^{2-d}.$$

Define $\psi(x) = n^{d/2-1} G_K(x\sqrt{n})$ for $x \in n^{-1/2}\mathbb{Z}^d$ and define $\psi(x)$ by some suitable interpolation procedure if $x \notin n^{-1/2}\mathbb{Z}^d$. Looking at the cases $\zeta > |x-y|$ and $\zeta \le |x-y|$ separately, we see

$$|\psi(x) - \psi(y)| \le c_{4.19} \left(\frac{|x-y|^{1-\beta}}{\zeta^{d-1-\beta}} \wedge \zeta^{2-d} \right).$$

Applying Lemma 4.3 to $c_{4.19}^{-1} \zeta^{d-1-\beta} \psi$, we get

$$II_{4.9} \le c_{4.20} \frac{\delta^{1-\beta}}{\zeta^{d-1-\beta}}.$$

Finally, by (4.7) and (4.2)

$$III_{4.9} \le c_{4.21} \left(n^{d/2-1/2} \delta \wedge n^{-\gamma/2} \right). \tag{4.10}$$

Looking at the cases when $n^{(d-1)/2}$ is greater than and less than $n^{-\gamma/2}$ separately,

$$III_{4.9} \le c_{4.22} \delta^{\ell}.$$

Choose K so that $\zeta = (\delta^{1-\beta} n^{\gamma/2})^{1/(\gamma+d-1-\beta)}$. So

$$\begin{aligned} I_{4.9} + II_{4.9} + III_{4.9} &\le c_{4.23} n^{(\gamma/2)[\gamma/(d-1+\gamma-\beta)-1]} \delta^{\gamma(1-\beta)/(d-1+\gamma-\beta)} + c_{4.22} \delta^{\ell} \\ &\le c_{4.24} \delta^{\ell_\beta}, \end{aligned}$$

since $n \ge 1$ and $\gamma/(d-1+\gamma-\beta) - 1 < 0$. □

5. Martingale Calculus estimates.

As in Sections 2–4, we restrict attention to the case $d \ge 3$. Let $\mathfrak{M}_n$ be as in Section 4. Fix $\mu_n, \nu_n \in \mathfrak{M}_n$ and let

$$A_k^n = L_k^{n,\mu_n} - L_k^{n,\nu_n}, \qquad U^n(x) = \mathbb{E}^x A_\infty^n.$$

There exists a mean 0 martingale M_k^n so that

$$U_k^n \equiv U^n(S_k) - U^n(S_0) = M_k^n - A_k^n.$$

Define $B_m^n = \max_{k\le m}|A_k^n|$. Fixing n allows us to drop the n subscript. We proceed to estimate $\mathbb{E}^y B_\infty^2$ with this convention in mind.

Proposition 5.1. *There is a constant $c_{5.1} = c_{5.1}(\mathfrak{M}_n)$ so that*

$$\sup_{n\ge 1} \sup_{\substack{\mu_n,\nu_n\in\mathfrak{M}_n \\ d_L(\mu_n,\nu_n)\le\delta}} \mathbb{E}^y |B_\infty^n|^2 \le c_{5.1}\delta^{\ell_\beta}$$

where ℓ_β is defined in (4.8).

Proof. Fix n. We shall temporarily drop the n superscripts. Notice that

$$\begin{aligned} |A_k|^2 &\le 2|U_k|^2 + 2|M_k|^2 \\ &\le 2c_{4.14}^2\delta^{2\ell_\beta} + 2|M_k|^2 \qquad \text{(Proposition 4.4)}. \end{aligned}$$

where $\delta = d_L(\mu,\nu)$. Hence

$$\begin{aligned} \mathbb{E}^y|B_\infty|^2 &\le 2c_{4.14}^2\delta^{2\ell_\beta} + 2\mathbb{E}^y \sup_k |M_k|^2 \\ &\le 2c_{4.14}^2\delta^{2\ell_\beta} + 8\mathbb{E}^y|M_\infty|^2 \qquad \text{(Doob's inequality)}. \end{aligned} \tag{5.1}$$

But

$$\begin{aligned} |M_\infty|^2 &\le 2|U_\infty|^2 + 2|A_\infty|^2 \\ &\le 2c_{4.14}^2\delta^{2\ell_\beta} + 2|A_\infty|^2 \qquad \text{(Proposition 4.4)}. \end{aligned}$$

Therefore (5.1) yields

$$\mathbb{E}^y|B_\infty|^2 \le 18c_{4.14}^2\delta^{2\ell_\beta} + 16\mathbb{E}^y|A_\infty|^2. \tag{5.2}$$

Letting $\Delta A_k \equiv A_{k+1} - A_k$, note that

$$\begin{aligned} A_\infty^2 &= \sum_{k=0}^{\infty}(A_{k+1}^2 - A_k^2) = \sum_k \Delta A_k(A_{k+1} + A_k) \\ &= \sum_k \Delta A_k\,(2A_{k+1} - \Delta A_k) \\ &= 2\sum_k A_{k+1}\Delta A_k - \sum_k (\Delta A_k)^2 \end{aligned}$$

So

$$
\begin{aligned}
A_\infty^2 &= 2\sum_k (A_\infty - A_{k+1})\Delta A_k + \sum_k (\Delta A_k)^2. \qquad (5.3)\\
&= I_{5.3} + II_{5.3}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}^y I_{(5.3)} &= \sum_k \mathbb{E}^y \left\{ \mathbb{E}(A_\infty - A_{k+1} | \hat{\mathcal{F}}_{k+1})\Delta A_k \right\} \qquad (\hat{\mathcal{F}}_i = \sigma\{X_1, \dots, X_i\})\\
&= \sum_k \mathbb{E}^y \left\{ \mathbb{E}^{S_{k+1}}[A_\infty]\Delta A_k \right\}\\
&\leq 2c_{4.14}\delta^{\ell_\beta} \sum_k \mathbb{E}^y |\Delta A_k| \qquad \text{(Proposition 4.4)}\\
&\leq 2c_{4.14}\delta^{\ell_\beta}\mathbb{E}^y (L_\infty^{n,\mu} + L_\infty^{n,\nu})\\
&\leq 4c_{4.14}c_{4.1}\delta^{\ell_\beta}. \qquad \text{(Proposition 4.1)}. \qquad (5.4)
\end{aligned}
$$

Next, using (4.2) and (4.7),

$$
\sup_x |(\mu_n - \nu_n)(\{x\}) \leq c_{5.2}(\sqrt{n}\delta \wedge n^{1-(d+\gamma)/2}).
$$

So

$$
\begin{aligned}
\mathbb{E}^y II_{5.3} &= n^{d-2}\mathbb{E}^y \sum_j [(\mu_n - \nu_n)(\{S_{j+1}/\sqrt{n}\})]^2\\
&\leq n^{d/2-1} \sup_x |(\mu_n - \nu_n)(\{x\})\mathbb{E}^y n^{d/2-1} \sum_j (\mu_n + \nu_n)(\{S_{j+1}/\sqrt{n}\}))\\
&\leq c_{5.2}[n^{(d-1)/2}\delta \wedge n^{-\gamma/2}][\mathbb{E}^y L_\infty^{n,\mu_n} + \mathbb{E}^y L_\infty^{n,\nu_n}].
\end{aligned}
$$

By the argument following (4.10) and Proposition 4.1,

$$
\mathbb{E}^y II_{5.3} \leq c_{5.3}\delta^\ell.
$$

Adding, we get Proposition 5.1. □

Using this, we prove the following exponential estimate:

Proposition 5.2. *For all $x \in (0,\infty)$, all $\beta \in (0,1)$,and all $\delta \leq 1$,*

$$
\sup_{n\geq 1} \sup_{\substack{\mu_n,\nu_n \in \mathfrak{M}_n \\ d_L(\mu_n,\nu_n)\leq\delta}} \mathbb{P}^y \left\{ \sup_k |L_k^{n,\mu_n} - L_k^{n,\nu_n}| \geq x \right\} \leq 2\exp\left\{ -\frac{x}{\sqrt{c_{5.4}\delta^{\ell_\beta}}} \right\}.
$$

Proof. Define $A^n(t) = A^n_{[t]}$ and $B^n_t = \sup_{s<t} |A^n(t)|$. Then $t \mapsto B^n_t$ is predictable and increasing. Since $t \mapsto B^n_t$ is also a sub-additive functional, Proposition 5.1 and Cauchy–Schwarz show that

$$\mathbb{E}^y \{B^n_\infty - B^n_T | \mathcal{F}_T\} \le \sqrt{c_{5.1}\delta^{\ell_\beta}}.$$

Therefore, by [DM] p. 193, for every $\delta \le 1$, all $x > 0$, and $\lambda \in (0, (c_{5.1}\delta^{\ell_\beta})^{-1/2}/8)$,

$$\sup_{n\ge 1} \sup_{\substack{\mu_n,\nu_n\in\mathfrak{M}_n \\ d_L(\mu_n,\nu_n)\le\delta}} \mathbb{E}^y e^{\lambda|B^n_\infty|} \le \left(1-\lambda\sqrt{c_{5.1}\delta^{\ell_\beta}}\right)^{-1}.$$

Hence

$$\sup_{n\ge 1} \sup_{\substack{\mu_n,\nu_n\in\mathfrak{M}_n \\ d_D(\mu_n,\nu_n)\le\delta}} \mathbb{P}^y \{|B^n_\infty| \ge x\} \le e^{-\lambda x}\left(1-\lambda\sqrt{c_{5.1}\delta^{\ell_\beta}}\right)^{-1}.$$

Letting $\lambda = 1/16\sqrt{c_{5.1}\delta^{\ell_\beta}}$, we get the result. □

6. Invariance Principles.

Throughout this section assume that $X_1, X_2, \cdots$ are mean zero random vectors taking values in $\mathbb{Z}^d$ and that $\mathrm{Cov}(X_1) = I$, the identity matrix. Further moment conditions will be imposed later. Let $S_n = \sum_{j=1}^n X_j$. Let $\mathfrak{M}$ be a family of positive measures on $\mathbb{R}^d$. Suppose for each $\mu \in \mathfrak{M}$ there exists a sequence of positive measures, $\mu_n = \mu(n)$ converging weakly to μ, and for each n, μ_n is supported on $n^{-1/2}\mathbb{Z}^d$. Let

$$\mathfrak{M}_n = \{\mu(n) \ : \ \mu \in \mathfrak{M}\}.$$

Hypothesis 6.1.

(a) *There exists $c_{6.1}$, independent of n, such that $\mu_n(\mathbb{R}^d) \le c_{6.1}$, $\mu_n \in \mathfrak{M}_n$;*

(b) *for some $\gamma > 0$, there exists $c_{6.2} \in (0,\infty)$, independent of n, such that*

$$\sup_x \mu_n(B(x,s)) \le c_{6.2}s^{d-2+\gamma} \qquad \text{if } 1/2\sqrt{n} \le s \le 1,\ n \ge 1,\ \mu_n \in \mathfrak{M}_n;$$

(c_β) *there exists $c_{6.3}$ and $\epsilon > 0$, independent of n, such that if H^n_L is the metric entropy of $\mathfrak{M}_n$ with respect to d_L, then*

$$H^n_L(x) \le c_{6.3}x^{-(\ell_\beta/2-\epsilon)}, \qquad H_L(x) \le c_{6.3}x^{-(\ell_\beta/2-\epsilon)}, \qquad x \in (0,1).$$

In what follows we will formulate a number of invariance principles. See [Bi] for the appropriate definitions concerning weak convergence on metric spaces. But perhaps the simplest way to describe what converging weakly uniformly over a family means is to say: one can find a probablility space supporting a Brownian motion Z_t and a random walk with the same distribution as the S_n's such that $S_{[nt]}/\sqrt{n}$ converges uniformly to Z_t, $t \in [0,1]$, a.s., and $L^{n,\mu_n}_{[nt]}$ converges uniformly to L^μ_t, $t \in [0,1]$, $\mu \in \mathfrak{M}$, a.s.

A. Subgaussian case.

In this subsection, assume $d \geq 3$ and assume that the X_i's are subgaussian. The following proposition follows from Proposition 5.2 just as Theorem 2.2 followed from Proposition 2.1, by standard metric entropy arguments.

Proposition 6.2. *If Hypothesis 6.1 holds for some $\beta \in (0,1)$, then for each $\eta > 0$*

$$\lim_{\delta\to 0}\sup_{n\geq 1}\sup \mathbb{P}\left\{\sup_{k\geq 1}\sup_{\substack{\mu_n,\nu_n\in\mathfrak{M}_n\\ d_L(\mu_n,\nu_n)\leq\delta}} |L^{n,\mu_n}_k - L^{n,\nu_n}_k| \geq \eta\right\} = 0.$$

Theorem 6.3. *If Hypothesis 6.1 holds for some $\beta \in (0,1)$, then the process*

$$\{(n^{-1/2}S_{[nt]}, L^{n,\mu_n}_{[nt]}) \ : \ 0 \leq t \leq 1 \ , \ \mu \in \mathfrak{M}\}$$

converges weakly to the process $\{(X_t, L^\mu_t) \ : \ 0 \leq t \leq 1 \ , \ \mu \in \mathfrak{M}\}$.

Proof. In order to keep things as simple as possible, we will prove that $L^{n,\mu_n}_{[nt]} \Rightarrow L^\mu_t$. A standard modification to our argument will show the joint convergence of the local time together with the random walk.

We start by showing the convergence of the finite dimensional distributions. We give the proof for the one dimensional marginals, the general case being entirely analogous.

Define φ and φ_ϵ as in the proof of Lemma 4.3. Recall $\varphi_\epsilon * \mu_n(x) = \int \varphi_\epsilon(x-y)\mu_n(dy)$. Define μ^ϵ_n to be the measure on $n^{-1/2}\mathbb{Z}^d$ that puts mass $n^{-d/2}\varphi_\epsilon * \mu_n(\{z/\sqrt{n}\})$ on the point $z/\sqrt{n}$, $z \in \mathbb{Z}^d$.

First, we show $d_L(\mu, \mu * \varphi_\epsilon) \to 0$ as $\epsilon \to 0$. We write

$$\begin{aligned} d_L(\mu, \mu * \varphi_\epsilon) &= \sup_{\psi\in\mathcal{L}} |\int \psi(y)\mu(dy) - \int \psi(y)\mu * \varphi_\epsilon(y)dy| \\ &= \sup_{\psi\in\mathcal{L}} |\int \psi(y)\mu(dy) - \int \psi * \phi_\epsilon(y)\mu(dy)|. \end{aligned} \tag{6.1}$$

Since $\psi \in \mathcal{L}$, $\psi * \varphi_\epsilon$ converges uniformly to ψ as $\epsilon \to 0$. Hence the right hand side of (6.1) tends to 0. A similar argument shows that $d_L(\mu_n, \mu_n * \varphi_\epsilon) \to 0$ as $\epsilon \to 0$, uniformly in n.

Secondly, we calculate, using Hypothesis 6.1(b),

$$\begin{aligned}\mu_n(B(x,s)) &= n^{-d/2}\sum_{|z-x|\le s\sqrt{n}}\sum_y \varphi_\epsilon(y/\sqrt{n})\mu_n(\{(z-y)/\sqrt{n}\})\\ &\le c_{6.4}n^{-d/2}\sum_y \varphi_\epsilon(y/\sqrt{n})s^{d-2+\gamma}\\ &\le c_{6.4}n^{-d/2}\epsilon^{-d}||\varphi||_\infty s^{d-2+\gamma}\#\{y\ :\ y/\epsilon\sqrt{n}\in \text{support}(\varphi)\}\\ &\le c_{6.5}s^{d-2+\gamma}, \end{aligned}\tag{6.2}$$

if $1/2\sqrt{n}\le s\le 1$. A similar calculation shows that $\sup_n \mu_n^\epsilon(\mathbb{R}^d)\le c_{6.6}$, independently of n and ϵ.

Thirdly, we show that for each $\epsilon>0$, μ_n^ϵ converges to $\mu*\varphi_\epsilon$ uniformly on compacts, as $n\to\infty$. Since $\mu_n^\epsilon(x)=\int\varphi_\epsilon(x-y)\mu_n(dy)$, the μ_n are uniformly bounded, and φ_ϵ is smooth, then $\{\mu_n^\epsilon\ :\ n\ge 1\}$ is an equicontinuous family of functions of x. For each fixed x, $\mu_n^\epsilon(x)\to\int\varphi_\epsilon(x-y)\mu(dy)=\mu*\varphi_\epsilon(x)$, since $\mu_n\xrightarrow{w}\mu$.

In view of (6.1), (6.2), and Propositions 5.2, 2.1, and 2.8, to show $\{L_{[nt]}^{n,\mu_n}\ :\ 0\le t\le 1\}$ converges weakly to $\{L_t^\mu\ :\ 0\le t\le 1\}$, it suffices to show that $L_{[nt]}^{n,\mu_n}\Rightarrow L_t^{\mu*\varphi_\epsilon}$ for each ϵ. But

$$L_{[nt]}^{n,\mu_n^\epsilon}=n^{-1}\sum_{j=0}^{[nt]}\varphi_\epsilon*\mu_n(n^{-1/2}S_j).\tag{6.3}$$

Since $\varphi_\epsilon*\mu_n$ converges to $\varphi_\epsilon*\mu$ uniformly on compacts, the desired convergence follows immediately by Donsker's theorem.

To complete the proof, it remains to establish tightness of $\{L_{[nt]}^{n,\mu_n}\ :\ 0\le t\le 1,\ \mu_n\in\mathfrak{M}_n\}$. But this follows from Proposition 6.2. $\square$

B. 3+ρ moments.

We still assume $d\ge 3$, but now only require that $\mathbb{E}|X_1|^{3+\rho}<\infty$, for some $\rho>0$.

Theorem 6.4. *If Hypothesis 6.1 holds for some $\beta\in(0,1)$, then the conclusion of Theorem 6.3. is still valid.*

Proof. Let $\alpha=1/8$, $a_n=n^{1/2-\alpha}$. If $X_j=(X_j^1,\dots,X_j^d)$, define $\tilde{X}_j=(\tilde{X}_j^1,\dots,\tilde{X}_j^d)$ by

$$\tilde{X}_j^i=X_j^i 1_{(|X_j^i|<a_n)},\qquad i=1,\dots,d.$$

Let $e_i=\mathbb{E}\tilde{X}_1^i$ and define X_j' by $(X_j^i)'=\tilde{X}_j^i-Y_j^i$, where Y_j^i is a random variable independent of the X's that takes the value $[a_n]\text{sgn}(e_i)$ with probability $|e_i|/[a_n]$, and the value 0 with probability $1-|e_i|/[a_n]$.

Since $\mathbb{E}X_1^i = 0$,

$$|e_i| = |\int_{[-a_n,a_n]^c} x\mathbb{P}(X_1^i \in dx)| \le 2\int_{a_n}^{\infty} \mathbb{P}(|X_1^i| \ge x)dx$$
$$\le 2a_n^{-(2+\rho)} \int_{a_n}^{\infty} x^{2+\rho}\mathbb{P}(|X_1^i| \ge x)dx$$
$$\le c_{6.7}a_n^{-(2+\rho)}\mathbb{E}|X_1^i|^{3+\rho}. \tag{6.4}$$

If n is large enough, $|e_i| < 1$.

Note that the X_j' are mean 0, have finite $3+\rho$ moments (with a bound independent of n), have covariance close to the identity matrix, are bounded by $2a_n$ for n large, and still take values in $\mathbb{Z}^d$ (which is why we did not simply define X' by $\tilde{X} - \mathbb{E}\tilde{X}$). We have by Chebyshev's inequality

$$\mathbb{P}(X_j' \ne X_j) \le \mathbb{P}(\tilde{X}_j \ne X_j) + \sum_{j=1}^{d} |e_j|/[a_n] \le c_{6.8}a_n^{-(3+\rho)}\mathbb{E}|X_j|^{3+\rho}$$
$$= o(1/n). \tag{6.5}$$

Let $S_k' = \sum_{j=1}^{k} X_j'$. By Bernstein's inequality,

$$\mathbb{P}(|S_n'| \ge |x|) \le 2\mathrm{exp}\left(\frac{-|x|^2}{2n + 4a_n|x|/3}\right), \qquad |x| \ge 1.$$

The expression on the right hand side is largest when n is the largest, and so if $f(x) = [|x|^2/k\log|x|]$,

$$\sum_{j=1}^{f(x)} \mathbb{P}(|S_n'| \ge |x|) \le 2f(x)\exp\left(\frac{-|x|^2}{2f(x) + 4a_{f(x)}|x|/3}\right)$$
$$\le c_{6.9}|x|^{1-d} \tag{6.6}$$

if k is large enough.

We now use (6.3) in place of (3.8), and proceeding exactly as in the proofs of Proposition 3.2 and Corollary 3.3, we conclude that

$$|G'(0,x)| \le c_{6.10}(1 \wedge |x|^{2-d}) \tag{6.7}$$

and that for each $\beta \in (0,1)$, there exists a $c_{6.11} = c_{6.11}(\beta)$ such that

$$|G'(0,x) - G'(0,y)| \le c_{6.11}\left(\frac{|x-y|^{1-\beta}}{(|x| \wedge |y|)^{d-1-\beta}} + \frac{|x-y|}{(|x| \wedge |y|)^{d-1}}\right), \tag{6.8}$$

where G' is defined in terms of X' just as G was defined in terms of X.

All the estimates of Sections 4 and 5 and of Proposition 6.1 are still valid, provided we replace L_k^{n,μ_n} by $(L_k^{n,\mu_n})' = n^{d/2-1}\sum_{j=1}^k \mu_n(\{S_j'/\sqrt{n}\})$. Then for all $\eta > 0$,

$$\begin{aligned}\mathbb{P}(\sup_{t\le 1} \sup_{\substack{\mu_n,\nu_n\in\mathfrak{M}_n\\ d_L(\mu_n,\nu_n)\le\delta}} &|L_{[nt]}^{n,\mu_n} - L_{[nt]}^{n,\nu_n}| \ge \eta)\\ &\le \mathbb{P}(\sup_{t\le 1} \sup_{\substack{\mu_n,\nu_n\in\mathfrak{M}_n\\ d_L(\mu_n,\nu_n)\le\delta}} |(L_{[nt]}^{n,\mu_n})' - (L_{[nt]}^{n,\nu_n})'| \ge \eta)\\ &\quad + \mathbb{P}(X_j \ne X_j' \text{ for some } j \le n).\end{aligned} \tag{6.9}$$

The first term on the right hand side of (6.9) can be made small, uniformly in n, by taking δ small and using Proposition 6.1 (applied to L'). To bound the second term, we write

$$\mathbb{P}(X_j \ne X_j' \text{ for some } j \le n) \le n\mathbb{P}(X_1 \ne X_1') \to 0$$

as $n \to \infty$ by (6.5). Tightness follows readily.

The proof of the convergence of the f.d.d.'s given in Theorem 6.3 goes through without change. □

C. $2+\rho$ moments.

Still assuming $d \ge 3$, we now assume only that $\mathbb{E}|X_1|^{2+\rho} < \infty$, for some $\rho > 0$.

Theorem 6.5. *Suppose Hypothesis 6.1 holds with $\beta = 1-\rho$. Then the conclusion of Theorem 6.3 holds.*

Proof. Let $\alpha = \rho/8$, $a_n = n^{1/2-\alpha}$. Define $\tilde{X},\ X'$ as in subsection B. As in the proof of Theorem 6.4,

$$|e_i| \le c a_n^{-(1+\rho)}\mathbb{E}|X_1|^{2+\rho},$$

and in place of (6.5) we get

$$\mathbb{P}(X_j' \ne X_j) \le \mathbb{P}(\tilde{X}_j \ne X_j) + \sum_{j=1}^d |e_j|/[a_n] \le c a_n^{-(2+\rho)}\mathbb{E}|X_1|^{2+\rho} = o(1/n). \tag{6.10}$$

Using Bernsteins's inequality, we get (6.6) as before. However,

$$\begin{aligned}\mathbb{E}|\tilde{X}_j^i|^3 &\le 3\int_0^{a_n} x^2\mathbb{P}(|X| \ge x)dx\\ &\le 3a_n^{1-\rho}\int_0^{a_n} x^{1+\rho}\mathbb{P}(|X| \ge x)dx \le c_{6.12}a_n^{1-\rho}\end{aligned}$$

and

$$\mathbb{E}|Y_j^i|^3 = [a_n]^3 |e_i| / [a_n] \le c_{6.13} a_n^{1-\rho}.$$

So

$$\mathbb{E}|X_j'|^3 \le c_{6.14} a_n^{1-\rho}.$$

Hence in Proposition 3.1 we can only conclude

$$\sup_{x \in \mathbb{Z}^d} |\mathbb{P}^x(S_n' = 0) - (2\pi n)^{-d/2} e^{-|x|^2/2n}| \le c_{6.15} n^{-(d+\rho)/2-\epsilon}$$

for some small $\epsilon > 0$. Using this estimate in (3.11),

$$\sum_{n \ge f(x)} E(n,x) \le c_{6.16} |x|^{-d-\rho+2}. \tag{6.11}$$

Following the proofs of Proposition 3.2 and Corollary 3.3, but using (6.6) in place of (3.8) and (6.11) in place of (3.11), we get (6.7) and (6.8) with $\beta = 1 - \rho$.

As in the $3 + \rho$ moment case, using (6.10), we get tightness. No changes are needed to the proofs of the convergence of the f.d.d.'s. □

D. Second moments.

When the X_j's have only finite second moments, our methods do not give uniform invariance principles. But we still can prove the convergence of the f.d.d.'s when $d = 3$.

Theorem 6.6. *Suppose $d = 3$, $\mathbb{E}|X_j|^2 < \infty$, and Hypothesis 6.1(a), (b) hold. For measures $\mu^1, \dots, \mu^N \in \mathfrak{M}$,*

$$(L_{[nt]}^{n,\mu_n^i} : \ 0 \le t \le 1, \ i = 1, \dots, N)$$

converges weakly to $(L_t^{\mu^i} : 0 \le t \le 1, i = 1, \dots, N)$.

Proof. We give the argument for $N = 1$, the general case being analogous. Examining the proof of Theorem 6.3, we see that we need only show that for each $\eta > 0$

$$\mathbb{P}(\sup_{k \le n} |L_k^{n,\mu_n} - L_k^{n,\mu_n^\epsilon}| \ge \eta) \to 0 \tag{6.12}$$

as $\epsilon \to 0$, uniformly for $n \ge n_0(\eta)$.

Let $\theta > 0$, $\zeta_n = \theta n^{1/2}$, and $K_n = \zeta_n^{2-d} = (\theta n^{1/2})^{2-d}$. As in the proof of Proposition 4.4, define

$$\psi_n(x) = n^{d/2-1} G_{K_n}(x n^{1/2})$$

for $x \in n^{-1/2}\mathbb{Z}^d$ and by a suitable interpolation procedure for $x \notin n^{-1/2}\mathbb{Z}^d$. Write ν_n for μ_n^ϵ. By Spitzer [S], $|G(0,z)| \le c_{6.17}(1 \wedge |z|^{2-d})$ and $G(0,z) = g(0,z)(1+o(1))$, as $|z| \to \infty$. So given a, there exists M_1 such that $|G(0,x) - g(0,x)| \le a|x|^{2-d}$ if $x \in \mathbb{Z}^d$, $|x| \ge M_1$. Hence if $|w|,\ |z| \ge M_1$, $w, z \in \mathbb{Z}^d$,

$$|G(0,w) - G(0,z)| \le \frac{c_{6.18}|w-z|}{(|w| \wedge |z|)^{d-1}} + \frac{2a}{(|w| \wedge |z|)^{d-2}}. \tag{6.13}$$

Now $|\psi_n(y) - \psi_n(z)|$ will be largest if $|y|, |z| \ge c_{6.19}\theta$, $y, z \in n^{-1/2}\mathbb{Z}^d$ for some constant $c_{6.19}$ independent of n. So suppose $|y|, |z| \ge c_{6.19}\theta$. Then for n large enough, $|y|\sqrt{n}, |z|\sqrt{n} \ge M_1$. So by (6.13), for n large enough,

$$\begin{aligned}|\psi_n(y) - \psi_n(z)| &\le n^{d/2-1}\left[\frac{c_{6.18}|y-z|\sqrt{n}}{(|y| \wedge |z|)^{d-1}n^{(d-1)/2}} + \frac{2a}{(|y| \wedge |z|)^{d-2}n^{(d-2)/2}}\right]\\ &\le \frac{c_{6.20}|y-z|}{\theta^{d-1}} + \frac{2a}{\theta^{d-2}}.\end{aligned} \tag{6.14}$$

Let $b > 0$. Choose θ small enough so that $\theta^\gamma < b$. Since the sequence $\{\mu_n\}$ is tight, we choose M_2 large so that $\mu_n(B(0, M_2)^c) \le b$. By the estimate (6.14),

$$\begin{aligned}|\psi_n(y) - \psi_n * \varphi_\epsilon(y)| &\le \int |\psi_n(y) - \psi_n(y - \epsilon x)|\varphi(x)dx\\ &\le \frac{c_{6.20}\epsilon|x|}{\theta^{d-1}} + \frac{2a}{\theta^{d-2}} \le b\end{aligned}$$

if we take a and ϵ small, and n sufficiently large. Therefore,

$$|\int \psi_n(y)(\mu_n - \nu_n)(dy)| = |\int [\psi_n(y) - \psi_n * \varphi_\epsilon(y)]\mu_n(dy)| \le c_{6.21}b.$$

As in the proof of Proposition 4.4 (see (4.9)),

$$|\mathbb{E}^x L_\infty^{n,\mu_n} - \mathbb{E}^x L_\infty^{n,\nu_n}| \le \theta^\gamma + |\int \psi_n(y)(\mu_n - \nu_n)(dy)| + c_{6.22}(d_L(\mu_n, \nu_n))^\ell. \tag{6.15}$$

So taking ϵ smaller if necessary, we can make the right hand side of (6.15) less than $(2 + c_{6.21})b$. Plugging the estimate (6.15) into the proof of Proposition 5.2 and using Chebyshev's inequality, we get finally

$$\begin{aligned}\mathbb{P}^0(\sup_k |L_k^{n,\mu_n} - L_k^{n,\mu_n^\epsilon}| \ge \eta) &\le \eta^{-2}\mathbb{E}^0[\sup_k |L_k^{n,\mu_n} - L_k^{n,\mu_n^\epsilon}|^2]\\ &\le c_{6.23}\eta^{-2}b\end{aligned}$$

if n is sufficiently large, which is precisely what we wanted. $\square$

E. $d = 1, 2$.

The results for $d = 1, 2$ follow by the usual projection argument.

Theorem 6.7. *Theorems 6.3, 6.4, 6.5, and 6.6 hold for $d = 1$ and 2.*

Proof. Fix $M > 0$. Given μ defined on $\mathbb{R}^d$, $d = 1$ or 2, define $\hat{\mu}$ on $\mathbb{R}^3$ by

$$\hat{\mu}(A \times B) = \mu(A)|B \cap B(0, M)|, \qquad\qquad A \subseteq \mathbb{R}^d,\ B \subseteq \mathbb{R}^{3-d},$$

$B(0, M)$ the ball in $\mathbb{R}^{3-d}$. Similarly, given μ_n defined on $n^{-1/2}\mathbb{Z}^d$, define $\hat{\mu}_n$ on $n^{-1/2}\mathbb{Z}^3$.

Define $\hat{X}_j = (X_j, Y_j)$, where Y_j is simple random walk on $\mathbb{Z}^{3-d}$, independent of the X_j's. Define

$$\hat{L}_k^{n,\hat{\mu}_n} = n^{1/2} \sum_{j=1}^{k} \hat{\mu}_n(\{\hat{S}_j/\sqrt{n}\}).$$

Then by Theorem 6.3, 6.4, or 6.6, $\hat{L}_{[nt]}^{n,\hat{\mu}_n}$ converges weakly to $\hat{L}_t^{\hat{\mu}}$, where $\hat{L}$ is the additive functional associated to $\hat{\mu}$.

But it is clear that for all μ_n, $\hat{L}_k^{n,\hat{\mu}_n} = L_k^{n,\mu_n}$ up until the first time $n^{-1/2}|\sum_{i=1}^k Y_j|$ exceeds M, and for all μ, $\hat{L}_t^{\hat{\mu}} = L_t^{\mu}$ up until the first time $(3-d)$–dimensional Brownian motion exceeds M in absolute value. Since M is arbitrary, the weak convergence of $L_{[nt]}^{n,\mu_n}$ to L_t^{μ} follows easily. $\square$

7. Examples.

A. Classical additive functionals – L^p functionals.

Suppose $p > d/2$, and $p^{-1} + q^{-1} = 1$. Let $\mathfrak{F}$ be a subset of $\{f \in L^p(B(0,1)) :\ f \geq 0\}$. Let H_p denote the metric entropy of $\mathfrak{F}$ with respect to $d_p(f_1, f_2) = ||f_1 - f_2||_p$. Note in what follows we do not assume our f's are continuous.

Theorem 7.1. *If $\sup_{f \in \mathfrak{F}} ||f||_p < \infty$ and the exponent of metric entropy of H_p is less than $1/2$, then $\int_0^t f(Z_s)ds$ is jointly continuous in $t \in [0, 1]$ and $f \in \mathfrak{F}$ (with respect to the d_p metric.)*

Proof. Here $\mathfrak{M} = \{\mu :\ \mu$ has a density $f(x)$ with respect to Lebesgue measure, $f \in \mathfrak{F}\}$, and $L_t^{\mu} = \int_0^t f(Z_s)ds$. By Hölder's inequality,

$$\mu(\mathbb{R}^d) = \int_{B(0,1)} f(x)dx \leq c_{7.1}||f||_p,$$

and $\mu(B(x,s)) = \int_{B(0,1)} 1_{B(x,s)}(y)f(y)dy \le ||1_{B(x,s)}||_q \; ||f||_p \le c_{7.2}s^{d/q}$, for $s \le 1$ and $f \in \mathfrak{F}$.

So the total mass of the μ's is uniformly bounded and the index of $\mathfrak{M}$ is $d/q - d + 2 = 2 - d/p > 0$. If $\mu(dx) = f(x)dx$ and $\nu(dx) = h(x)dx$, then

$$d_G(\mu,\nu) = \sup_x |\int g(0,x)[f(x) - h(x)]dx| \le ||f-h||_p \; ||g(0,\cdot)||_q \le c_{7.3} d_p(f,h),$$

since $g \in L^q(B(0,1))$ when $p > d/2$.

Our result now follows by Theorem 2.2. □

Since changing f on a set of measure 0 does not affect L_t^μ (here $\mu(dx) = f(x)dx$), but can have a drastic effect on $n^{-1}\sum_j f(n^{-1/2}S_j)$, for an invariance principle one must have some additional regularity for f (cf. the next example).

B. Classical additive functionals – indicators.

Let $\mathfrak{A}$ be a subset of $\{A : A \subseteq B(0,1)\}$. Suppose that for almost every $y \in \mathbb{R}^d$, $r \in (0,1]$, and $A \in \mathfrak{A}$, as $n \to \infty$,

$$n^{-d/2}\#\{n^{-1/2}\mathbb{Z}^d \cap A \cap B(y,r)\} \to |A \cap B(y,r)|. \tag{7.1}$$

Define $d_S(A,B) = |A \triangle B|$.

Theorem 7.2. *Suppose the X_i satisfy the assumptions of Section 6 and have $2+\rho$ moments. Let $\beta = 1-\rho$ and let ℓ_β be defined by (4.8). Suppose $\mathfrak{A}$ satisfies (7.1) and the exponent of metric entropy of $\mathfrak{A}$ with respect to d_S is less than $\ell_\beta/2$. Then $n^{-1}\sum_{i=1}^{[nt]} 1_A(S_j/\sqrt{n})$ converges weakly to $\int_0^t 1_A(Z_s)ds$, uniformly over $t \in [0,1]$ and $A \in \mathfrak{A}$.*

Proof. For $A \in \mathfrak{A}$, define μ_A by $\mu_A(dx) = 1_A(dx)$. Define $\mu_{A,n}$ by $\mu_{A,n}(\{n^{-1/2}x\}) = n^{-d/2}1_A(n^{-1/2}x)$. That $\mu_{A,n}$ converges to μ_A follows by (7.1) and [Bi]. That Hypothesis 6.1 (a) and (b) hold is easy. Hypothesis 6.1 (c) follows from the crude estimate

$$|\int \psi(x)[1_A(x) - 1_B(x)]dx| \le |A \triangle B|, \qquad \psi \in \mathcal{L},$$

and a similar formula for $d_L(\mu_{A,n}, \mu_{B,n})$. Now apply Theorem 6.5. □

C. Local times on curves.

This example works for hypersurfaces in $\mathbb{R}^d$ for any dimension d, but for simplicity we restrict ourselves to $d=2$ and the curves of form

$$C = \{(t, f(t)):\ t \in [0,1],\ ||f||_\infty \leq c_{7.4}\}. \tag{7.2}$$

We will use C to denote the graph of C. Let $\mathfrak{C}$ be a collection of such curves. Let $\mu_C(A) = |\{t:\ (t, f_C(t)) \in A\}|$.

For such $C \in \mathfrak{C}$, we let $f_{C,n}$ be a function from $[0,1]$ to $[-2c_{7.4}, 2c_{7.4}]$, such that $f_{C,n}$ takes values in $n^{-1/2}\mathbb{Z}$, has jumps only at t's in $n^{-1/2}\mathbb{Z}$, and $f_n \to f$ in L^1–norm. Denote the curve and graph of $\{(t, f_{C,n}(t)):\ t \in [0,1]\}$ by C_n. If C^1 and C^2 denote two curves of the form (7.2) (corresponding to f_1, and f_2, resp.), let $d_C(f_1, f_2) = ||f_{C^1} - f_{C^2}||_1$.

Theorem 7.3. *Suppose the X_i satisfy the assumptions of Section 6 and have $2+\rho$ moments. Suppose that for some $c_{7.4}$ and ϵ independent of n*

$$H_C^n(x) \leq c_{7.4} x^{-(\ell_\beta/2-\epsilon)}, \qquad H_C(x) \leq c_{7.4} x^{-(\ell_\beta/2-\epsilon)}, \qquad x \in (0,1),$$

wher $H_C^n(x)$ (resp. $H_C(x)$) is the metric entropy of $\mathfrak{C}$ (resp. $\mathfrak{C}_n = \{C_n:\ C \in \mathfrak{C}\}$) with respect to d_C. Then $n^{-1/2}\sum_{i=1}^{[nt]} 1_{C_n}(n^{-1/2}S_j)$ converges weakly to $L_t^{\mu_C}$, uniformly over $t \in [0,1]$, $C \in \mathfrak{C}$.

Proof. Define $\mu_{C,n}(A) = n^{-1/2}\#\{k \leq \sqrt{n}:\ (n^{-1/2}k, f_{C,n}(n^{-1/2}k)) \in A\}$. Since $f_{C,n} \to f$ in L^1, $\mu_{C,n} \xrightarrow{w} \mu_C$. Note that $\mu_{C,n}(\mathbb{R}^d) \leq 1$, while

$$\mu_{C,n}(B(x,r)) \leq c_{7.5} r,$$

so the index of $\mathfrak{C}_n$ is 1. The result follows from Theorem 6.5. □

D. Local times in $\mathbb{R}^1$.

Even for local times in $\mathbb{R}^1$, our results are fairly strong. For $x \in \mathbb{R}^1$, let μ_x be the point mass at x. The $L_t^{\mu_x}$ (usually written as L_t^x) is just local time at x. Clearly the μ_x are uniformly bounded with index 1. By Example 2.2, $H_L(\delta) \leq c_{7.6}|\log(\delta)|$.

Define $\Gamma(n,x)$ to be $n^{-1/2}$ times the unique integer lying in the interval $[x\sqrt{n}, x\sqrt{n}+1]$.

Theorem 7.4. *If the X_i have finite $2+\rho$ moments for some $\rho > 0$, and are as in Section 6, then $n^{-1/2}\sum_{j=1}^{[nt]} 1_{[\sqrt{n}x, \sqrt{n}x+1)}(S_j)$ converge weakly to L_t^x, uniformly over all levels x.*

Proof. It suffices to prove the result uniformly over $x \in [-M, M]$ for each M. Define $\mu_{n,x}$ to be point mass at $\Gamma(n,x)$. Clearly $\mu_{n,x} \xrightarrow{w} \mu_x$ as $n \to \infty$, the $\mu_{n,x}$ are uniformly bounded, have index 1, and entropy is of order $|\log(x)|$. The result follows by Theorem 6.5 and the observation that $S_j \in [\sqrt{n}x, \sqrt{n}x+1]$ if and only if $\mu_{n,x}(\{n^{-1/2}S_j\}) = 1$. $\square$

The question of invariance principles for local time has a long history, dating back to [CH]. Using techniques highly specific to one–dimensional Brownian motion, Borodin [Bo1] has proved Theorem 7.4 when the X_i's have finite second moments. For a slightly different notion of local time, [P] has a uniform invariance principle if the X_i have $1+\sqrt{2} \approx 2.732$ moments.

E. Fractals.

For simplicity, we confine ourselves to $d = 2$ and fractals of the following form: let $F_0 = [0,1]^2$, let F_1 be the union of R closed squares of with sides ofsize a, such that the interiors are pairwise disjoint. To form F_2, replace each of the squares making up F_1 by replicas of F_1, and continue.

To be more precise, if S is any square, let Ψ_S be the orientation preserving affine map that takes S to F_0. Let

$$F_2 = \bigcup\{\ \Psi_S^{-1}(F_1) :\ S \text{ is one of the } R \text{ squares with sides of size } a \text{ making up } F_1\},$$
$$F_{k+1} = \bigcup\{\ \Psi_S^{-1}(F_1) :\ S \text{ is one of the } R^k \text{ squares with sides of size } a^k \text{ making up } F_k\}.$$

Let $F = \bigcap_{k=0}^{\infty} F_k$.

For example, if $F_1 = [0,1]^2 - (1/3,2/3)^2$, F will be the Sierpinski carpet. If $F_1 = (\ [0,1/3] \cup [2/3,1]\ )^2$, we get the 2–dimensional Cantor set.

Let μ be the Hausdorff–Besicovitch measure on F, normalized to have total mass 1.

Theorem 7.5. *If the Hausdorff dimension of $F > 0$,*

$$\frac{1}{|F_n|}\int_0^t 1_{F_n}(Z_s)ds \overset{\text{a.s.}}{\to} L_t^{\mu}.$$

Remark 7.6. The convergence in probability is a consequence of results in [B].

Proof. It is not hard to see that $\sup_x \mu(B(x,s)) \le c_{7.7}s^{\gamma}$, where γ is the Hausdorff dimension of F.

Suppose $\psi \in \mathcal{L}$, and let $S_1, \ldots, S_R$ be the squares making up F_1. Let $\mu_n(dx) = |F_n|^{-1}1_{F_n}(x)dx$. Let x_i be the lower left corner of S_i.

Since μ_2, μ_1 both have total mass 1,

$$\int_{S_i} \psi(x)[\mu_2(dx) - \mu_1(dx)] = |F_1|^{-1} \int_{S_i} [\psi(x) - \psi(x_i)][|F_1|^{-1} 1_{F_2}(x) - 1_{F_1}(x)]dx$$
$$= |F_1|^{-1} a^2 \int_{F_0} \psi_i(x)[\mu_1(dx) - \mu_2(dx)], \tag{7.2}$$

where $\psi_i(x) = \psi \circ \Psi_{S_i}^{-1}(x) - \psi \circ \Psi_{S_i}^{-1}(x_i)$. Since $|\nabla \psi_i(x)| \le a$, and $\psi_i(0) = 0$, then $||\psi_i||_\infty \le \sqrt{2}a$. So the right hand side of (7.2) is bounded above by $|F_1|^{-1}\sqrt{2}a^3 d_L(\mu_0, \mu_1)$. Summing over i, and taking the supremum over $\psi \in \mathcal{L}$,

$$d_L(\mu_2, \mu_1) \le \sqrt{2}a^3 R|F_1|^{-1} d_L(\mu_2, \mu_1) = \sqrt{2}a d_L(\mu_1, \mu_0).$$

By an induction argument,

$$d_L(\mu_{k+1}, \mu_k) \le (\sqrt{2}a)^k d_L(\mu_1, \mu_0). \tag{7.3}$$

If $R = 1$, so that F_1 is a single square, then F is a single point. This case is ruled out by the assumption that the dimension of F is strictly bigger than 0. So $R > 1$, and hence $a < 1/2$.

Let $\mathfrak{M} = \{\mu_n\}_{n=1}^\infty \cup \{\mu\}$. To cover $\mathfrak{M}$ with d_L–balls of radius δ, first put a ball B of radius δ around μ. Since $\mu_k \xrightarrow{w} \mu$, (7.3) shows that $d_L(\mu_k, \mu) \le c_{7.8}(\sqrt{2}a)^k$. So B covers all but $|\log(\delta/c_{7.8})/\log(\sqrt{2}a)| + 1$ of the μ_n's. So at most $c_{7.9}|\log(\delta)|$ balls are needed, hence $H_L(\delta) \approx |\log|\log(\delta)||$.

By Theorem 2.2, L_t^ν is continuous with respect to d_L, for $\nu \in \mathfrak{M}$. This implies our result. □

F. Intersection local time – double points.

Let S_n^1 and S_n^2 be two independent identically distributed random walks converging in law to two independent Brownian motions, Z_t^1, and Z_t^2. By redefining these processes on a suitable probability space, we may assume that the convergence is almost sure.

Define $\mu_{u,x}(A) = |\{t \in [0,u] : \ Z_t^2 + x \in A\}|$. In [BK], it is shown that $\alpha(x, s, u) = L_s^{\mu_{u,x}}$ is the intersection local time for (Z^1, Z^2). Let us consider the corresponding invariance principle. We discuss the case $d = 3$ first. (If $d \ge 4$, the paths of Z^1 and Z^2 do not intersect.)

If $x = (x^1, x^2)$, let $\Gamma_2(n, x) = (\Gamma(n, x^1), \Gamma(n, x^2))$, where Γ is defined in subsection D. Define

$$\mu_{u,x,n}(A) = n^{-1} \sum_{k=1}^{[nu]} 1_A(S_j^2/\sqrt{n} + \Gamma_2(n, x)).$$

Lemma 7.7. *There exists $\gamma > 0$ such that for each M, with probability one,*

$$\mu_{u,x,n}(B(y,s) - \{y\}) \le c_{7.10} s^{1+\gamma}, \qquad x, y \in B(0,M),\ s \le 1,$$

where $c_{7.10}$ depends on M and ω.

Proof. For simplicity, we prove this when $x = 0$, the general case being similar.

$$\begin{aligned}\mathbb{E}^z \mu_{\infty,0,n}(B(y,s) - \{y\}) &\le n^{-1} \sum_{w \ne y} G(z,w) 1_{B(y,s)}(n^{-1/2} w) \\ &\le n^{-1} \sum_{k=0}^{\infty} \sum_{2^k \le |w-z| < 2^{k+1}} 2^{-k} n^{-1/2} \#\{B(y\sqrt{n}, s\sqrt{n}) \cap \mathbb{Z}^d \cap [B(z, 2^{k+1}) - B(z, 2^k)]\} \\ &\le c_{7.11} s^{1+\gamma},\end{aligned}$$

for $\gamma = 1/2$.

This estimate is uniform in z, hence the potential of $\mu_{\infty,0,n}(B(y,s) - \{y\})/a$ is bounded above by 1, where $a = \sup_z \mathbb{E}^z \mu_{\infty,0,n}(B(y,s) - \{y\})$. By [DM], p. 193,

$$\mathbb{P}^z\{\mu_{\infty,0,n}(B(y,s) - \{y\}) > c_{7.11} s^{9/8}\} \le c_{7.12} \exp(-c_{7.13} s^{-1/8}).$$

For each k, we can choose $N_k = c_{7.14} 2^{3k}$ balls, each of radius 2^{-k+2}, so that for every $y \in B(0,M)$ and every $s \le 2^{-k+1}$, $B(y,s)$ is covered by one of these N_k balls. Hence,

$$\begin{aligned}\mathbb{P}^z\{\mu_{\infty,0,n}(B(y,s) - \{y\}) > c_{7.11} s^{9/8} \text{ for some } y \in B(0,M) \text{ and some } s \in [2^{-k}, 2^{-k+1}]\} \\ \le N_k c_{7.12} \exp(-c_{7.13} 2^{k/8}).\end{aligned}$$

Summing over k and using the Borel–Cantelli lemma, we conclude

$$\sup_{\substack{y \in B(0,M) \\ 0 < s \le 1}} \mu_{\infty,0,n}(B(y,s) - \{y\}) \le c_{7.14} s^{9/8}, \qquad \text{a.s.}$$

□

Theorem 7.8. *Let X_i^1, X_i^2 be two independent sequences of i.i.d. r.v.'s, identically distributed, and satisfying the assumptions of Section 6 with $2 + \rho$ moments. If $d = 3$, $L_{[ns]}^{n,\mu_{u,n,x}}$ converges weakly to $\alpha(x,s,u)$, uniformly over $x \in \mathbb{R}^3$, $s, u \in [0,1]$.*

Proof. We apply Theorem 6.5. Since $\sup_{y,n} \mathbb{P}^y\{\sup_{j \le n} |S_j^2/\sqrt{n}| \ge M\} \to 0$ as $M \to \infty$, it suffices to look at the $\mu_{u,n,x}$ restricted to $B(0,M)$.

For each u, the metric entropy of $\{\mu_{u,x,n} : x \in B(0,M)\}$ is bounded above by $c_{7.15} \delta^{-3}$. For each x, the total variation of $\mu_{u_2,x,n} - \mu_{u_1,x,n}$ is bounded above by $u_2 - u_1$.

So Hypothesis 6.1 (c_β) holds for every $\beta > 0$. Hypothesis 6.1 (a) is clear and 6.1 (b) is Lemma 7.7.

Since $S_n^2/\sqrt{n}$ converges uniformly to Z_t^2, $\mu_{u,n,x} \xrightarrow{w} \mu_{u,x}$. The result follows. □

To handle the case $d = 2$, we use the projection technique of Section 6 E, and get Theorem 7.8 for the case $d = 2$ as well.

For both the $d = 2$ and $d = 3$ cases, weak convergence at a single level x follows by Theorem 6.6 or 6.7 under the assumption of finite variance only.

G. Intersection local time – multiple points.

In [BK], we gave a method for constructing intersection local time for the intersection of $k+1$ independent Brownian motions in $\mathbb{R}^2$ from the intersection local time of k independent planar Brownian motions. A completely analogous construction can be made for the number of intersections of k random walks. We then can get the analogue of Theorem 7.8: for $d = 2$ only, the number of intersections converges weakly to the k–tuple intersection local time, uniformly over all the variables, provided the X's have $2 + \rho$ moments. As in the proof of Theorem 7.8, the only work is in finding the index of the family of measures, and as in [BK], the estimates needed for $k + 1$–intersection local time follow from those obtained for k–intersection local time.

For multiple points, we cannot use a projection argument, and must work with 2–dimensional random walks killed off at a geometric rate. So it is necessary to rework the results of Section 3 for $d = 2$ with G replaced by the λ–resolvent of S_n. We leave the (numerous) details to the interested reader.

References.

[B] Bass, R. F.: Joint continuity and representations of additive functionals of d-dimensional Brownian motion. Stoch. Proc. Applic. **17**, 211–227 (1984)

[BK] Bass, R. F., Khoshnevisan, D.: Intersection local times and Tanaka formulas. In preparation

[Bi] Billingsley, P.: Convergence of Probability Measures. New York: Wiley 1968

[Bo1] Borodin, A. N.: On the asymptotic behavior of local times of recurrent random walks with finite variance. Theor. Prob. and Applic. **26**, 758–772 (1981)

[Bo2] Borodin, A. N.: Brownian local time. Russian Math. Surveys **44**, 1–51 (1989)

[Br] Brosamler, G. A.: Quadratic variation of potentials and harmonic functions. Trans. Amer. Math. Soc. **149**, 243–257 (1970)

[CH] Chung, K. L., Hunt, G. A.: On the zeros of $\sum_1^n \pm 1$. Ann. of Math. **50**, 385–400 (1949)

[DM] Dellacherie, C., Meyer, P.-A.: Probabilités et Potentiel: Théorie des Martingales. Paris: Hermann 1980

[Du] Dudley, R. M.: Sample functions of the Gaussian process. Ann. Probab. **1**, 66–103 (1973)

[Dy] Dynkin, E. B.: Self–intersection gauge for random walks for Brownian motion. Ann. Probab. **16**, 1–57 (1988)

[LG] LeGall, J.-F.: Propriétés d'intersection des marches aléatoires I: convergence vers le temps local d'intersection. Comm. Math. Phys. **104**, 471–507 (1986)

[NS] Ney, P. E., Spitzer, F.: The Martin boundary for random walks. Trans. Amer. Math. Soc. **121**, 116–132 (1966)

[P] Perkins, E.: Weak invariance principles for local time. Z. f. Wahrschein. **60**, 437–451 (1982)

[Ro] Rosen, J.: Random walks and intersection local time. Ann. Probab. **18**, 959–977 (1990)

[S] Spitzer, F.: Principles of Random Walk. Berlin: Springer 1964

[Y] Yor, M.: Sur la transformée de Hilbert des temps locaux Browniens et une extension de la formule d'Itô. In: Azèma, J., Yor, M. (eds.) Séminaire de Probabilités XVI, pp. 238–247. Berlin: Springer 1982