Journal of Statistical Physics, Vol. 54, Nos. 3/4, 1989

Exponential Convergence to Equilibrium for a Class of Random-Walk Models

Alan D. Sokal[1] **and Lawrence E. Thomas**[2]

Received June 16, 1988; final October 7, 1988

We prove exponential convergence to equilibrium (L^2 geometric ergodicity) for a random walk with inward drift on a sub-Cayley rooted tree. This random-walk model generalizes a Monte Carlo algorithm for the self-avoiding walk proposed by Berretti and Sokal. If the number of vertices of level N in the tree grows as $c_N \sim \mu^N N^{\gamma-1}$, we prove that the autocorrelation time τ satisfies $\langle N \rangle^2 \lesssim \tau \lesssim \langle N \rangle^{1+\gamma}$.

KEY WORDS: Markov chain; random walk; geometric ergodicity; dynamic critical phenomena; Monte Carlo; self-avoiding walk.

1. INTRODUCTION

The study of dynamic critical phenomena in statistical mechanical model systems is of interest for two reasons. First, and most obviously, to the extent that the mathematical dynamics is a reasonable model of a real physical dynamics, the conclusions are of direct physical interest. A second and more subtle reason arises out of the widespread use of dynamic Monte Carlo methods as a tool for studying the *static* properties of statistical mechanical systems.[1,2] Monte Carlo studies of critical phenomena have been greatly hampered by critical slowing down: the autocorrelation time τ of the Monte Carlo stochastic process grows to infinity as the critical point is approached, which leads to a corresponding growth in the statistical error bars.[3] The rate of growth of τ is thus a crucial factor in determining the statistical efficiency of the Monte Carlo algorithm.

[1] Department of Physics, New York University, New York, New York 10003.
[2] Department of Mathematics, University of Virginia, Charlottesville, Virginia 22903.
[3] Each block of data of length $\approx 2\tau$ can be considered, roughly speaking, to contribute one "statistically independent" data point. Therefore, the "effective sample size" from a Monte Carlo run of length n is $\approx n/2\tau$, resulting in statistical error bars of order $(\tau/n)^{1/2}$. For a more detailed treatment, see ref. 1, Sections 1.2.3 and 1.2.4, and ref. 3, Section 4.1.

The present paper is one of a series by the authors and their collaborators[3-12] aimed at studying the dynamic critical behavior of Monte Carlo algorithms in statistical mechanics and quantum field theory. In this paper we study a problem—random walk with inward drift on a countable rooted graph—that is a generalization of a Monte Carlo algorithm for the self-avoiding walk (SAW) proposed by Berretti and Sokal[3] and used subsequently by several groups.[13-17]

The Berretti–Sokal algorithm generates self-avoiding walks with one endpoint fixed at the origin and the other endpoint free, in a *variable-length* ensemble controlled by a monomer activity β; as β approaches the critical point β_c, the average walk length $\langle N \rangle$ tends to infinity. The elementary moves of the algorithm are to delete the last bond of the walk ($\Delta N = -1$) or to append one bond to the end of the walk ($\Delta N = +1$); the relative probabilities of these two moves are chosen so as to leave invariant the Gibbs measure π_β.

Now the space of all self-avoiding walks (of arbitrary length) starting at the origin and ending anywhere has the structure of a *rooted tree*: the root is the zero-step walk, and a walk ω' is declared to be a child of ω if it is a one-step extension of ω. It is then easy to see that the Berretti–Sokal algorithm is precisely a *random walk with inward drift* on this tree.

We can now abstract the situation: given an arbitrary countable rooted graph G (satisfying certain growth restrictions) and an "activity" β, we define on G the "Gibbs" measure π_β and the corresponding random walk with inward drift. The question is now: For which graphs G does this random walk have exponential convergence to equilibrium ($\tau < \infty$) for all $\beta < \beta_c$? And if there is exponential convergence to equilibrium, how does the autocorrelation time τ behave as $\beta \to \beta_c$?

We are unable to answer these questions in general, but for a very interesting class of graphs—the *sub-Cayley rooted trees*—we prove rigorous upper and lower bounds on τ which are close to, but not quite, sharp. A connected rooted graph G is said to be sub-Cayley if, for each vertex $x \in G$, the rooted graph of descendants of x (with x as its root) is isomorphic to a rooted subgraph of G. The key fact is that *the space of all SAWs is a sub-Cayley rooted tree*: this expresses the fact that any segment of a self-avoiding walk must itself be self-avoiding. We are thus able to analyze a class of Markov chains which includes as a special case the Berretti–Sokal algorithm for SAWs.

To state our main result, assume for simplicity that the number of vertices of level N in the tree grows as $c_N \sim \mu^N N^{\gamma-1}$; here γ is a "critical exponent" and the sub-Cayley property implies that $\gamma \geqslant 1$. Then we prove that

$$\langle N \rangle^2 \lesssim \tau \lesssim \langle N \rangle^{1+\gamma} \tag{1.1}$$

(Note that in the SAW case γ is believed to be quite close to 1: $\gamma = 43/32$ in $d = 2$, ≈ 1.16 in $d = 3$, and $= 1$ in $d \geqslant 4$.) We had originally hoped to prove the Berretti–Sokal[3] conjecture $\tau \sim \langle N \rangle^2$, but in fact we are *not* able to do so,[4] and this for a very good reason: it is probably *false*![18] The exact dynamic critical behavior in this model is thus an open question.

For general graphs G, we can offer only some partial results: a lower bound $\tau \gtrsim \langle N \rangle^2$, and an upper bound which in general does not extend all the way to the critical point. In fact, we give an example—a "maximally unbalanced tree"[7]—in which exponential convergence to equilibrium breaks down well *before* the critical point. Thus, upper bounds on τ near the critical point require some structural hypothesis on the graph G; being a sub-Cayley rooted tree is sufficient but presumably far from necessary. It is an open problem to find a weaker sufficient condition.

Finally, we mention that very similar results (more general but slightly weaker) have been obtained concurrently by Lawler and Sokal,[7] using very different methods.

2. PRELIMINARIES

In this section we review briefly the theory of discrete-time Markov chains on a countable state space,[19] with emphasis on the L^2 spectral properties of the transition probability operator. Most of this theory can be generalized to Markov chains on an arbitrary (measurable) state space[20] and to continuous-time Markovian jump processes, but we shall not need this here.

Consider a Markov chain on a countable state space S, with transition probability matrix $P = \{p(i \to j)\}_{i,j \in S} = \{p_{ij}\}_{i,j \in S}$. Let $p_{ij}^{(n)} = (P^n)_{ij}$ be the n-step transition probability from i to j. Then the Markov chain is said to be *irreducible* if for each pair $i, j \in S$ there exists an $n \geqslant 0$ for which $p_{ij}^{(n)} > 0$. The chain is said to be *irreducible and aperiodic* if for each pair $i, j \in S$ there exists an $n_0 = n_0(i, j)$ such that $p_{ij}^{(n)} > 0$ for *all* $n \geqslant n_0(i, j)$. All the chains to be considered in this paper are irreducible.

Let $m_{ij} \equiv E_i[\tau_j]$ be the mean hitting time to state j starting in state i. (If $j = i$, this is the mean return time from state i to itself.) For an irreducible chain, it can be shown[19] that if $m_{ii} < \infty$ for at least one state i, then in fact $m_{ij} < \infty$ for all pairs i, j. A Markov chain with this property is said to be *positive-recurrent*.

[4] An early claim by one of us[11] to have proven $\tau \sim \langle N \rangle^2$ turned out, on closer inspection, to be mistaken.

A probability measure $\pi = \{\pi_i\}_{i \in S}$ on S is said to be a *stationary distribution* for the Markov chain if

$$\sum_i \pi_i p_{ij} = \pi_j \tag{2.1}$$

for all $j \in S$. The basic limit theorem for Markov chains is the following (see, e.g., ref. 19):

Theorem 2.1. Let P be the transition matrix for an irreducible Markov chain. Then:

(a) A stationary distribution exists if and only if the chain is positive-recurrent. In this case the stationary distribution is unique, and it is given by $\pi_i = 1/m_{ii}$.

(b) $\lim_{N \to \infty} (1/N) \sum_{n=1}^{N} p_{ij}^{(n)} = \pi_j$ for all i, j.

(c) If the chain is aperiodic, then $\lim_{n \to \infty} p_{ij}^{(n)} = \pi_j$ for all i, j.

Thus, an irreducible positive-recurrent Markov chain converges to equilibrium as time goes to infinity, irrespective of the initial state. Much work has been done on the rate of convergence to equilibrium (see, e.g., ref. 20 and the references cited there); the present paper is a further contribution to that question.

From now on we assume that the Markov chain is irreducible and positive-recurrent, and we let π denote the unique stationary distribution. Let $l^p(\pi)$, $1 \leqslant p \leqslant \infty$, denote the Banach space of complex-valued functions on the state space S having finite norm

$$\|f\|_p \equiv \begin{cases} [\sum_i \pi_i |f(i)|^p]^{1/p} & \text{if} \quad 1 \leqslant p < \infty \\ \sup_i |f(i)| & \text{if} \quad p = \infty \end{cases} \tag{2.2}$$

In particular, $l^2(\pi)$ is a Hilbert space with inner product

$$(f, g) \equiv \sum_i \pi_i f(i)^* g(i) \tag{2.3}$$

Let Π be the expectation operator

$$(\Pi f)(i) \equiv \sum_j \pi_j f(j) \qquad \text{for all } i \tag{2.4}$$

On each space $l^p(\pi)$, Π is an idempotent contraction with range equal to the constant functions; on $l^2(\pi)$ it is self-adjoint (hence an orthogonal projection). Now define the action of the transition probability matrix on functions by

$$(Pf)(i) = \sum_j p_{ij} f(j) \tag{2.5}$$

It is not hard to prove the following facts[21]:

Proposition 2.2. Let P be the transition probability for an irreducible positive-recurrent Markov chain with stationary distribution π. Fix $p \in [1, \infty]$, and consider the action of P on the space $l^p(\pi)$. Then:

(a) The operator P is a contraction. (In particular, its spectrum lies in the closed unit disk.)

(b) 1 is a simple eigenvalue of P, as well as of its adjoint P^*, with eigenvector equal to the constant function $\mathbf{1}$. (In particular, $P\Pi = \Pi P = \Pi$.)

(c) If the Markov chain is aperiodic, then 1 is the only eigenvalue of P (and of P^*) on the unit circle.

The goal of this paper is to prove, for certain Markov chains, that the spectrum of $P \upharpoonright \mathbf{1}^\perp$ (or equivalently $P - \Pi$) on $l^2(\pi)$ stays strictly away from the unit circle. As will now be explained, this corresponds to a uniform exponential decay of all autocorrelation functions (for L^2 observables).

Consider the Markov chain started in its equilibrium distribution π; let $X_0, X_1, \ldots \in S$ be the successive states of the Markov chain. Let f be a real-valued function in $l^2(\pi)$. Then $\{f(X_t)\}$ is a stationary stochastic process with mean

$$\mu_f \equiv \langle f(X_t) \rangle = \sum_i \pi_i f(i) \tag{2.6}$$

and unnormalized autocorrelation function[5]

$$\begin{aligned} C_{ff}(t) &\equiv \langle f(X_s)\, f(X_{s+t}) \rangle - \mu_f^2 \\ &= \sum_{i,j} f(i) [\pi_i p_{ij}^{(|t|)} - \pi_i \pi_j]\, f(j) \\ &= (f, (P - \Pi)^{|t|} f) \end{aligned} \tag{2.7}$$

Typically, $C_{ff}(t)$ decays exponentially ($\sim e^{-|t|/\tau}$) for large t; we define the *exponential autocorrelation time*

$$\tau_{\exp, f} = \limsup_{t \to \infty} \frac{t}{-\log |C_{ff}(t)|} \tag{2.8}$$

[5] In the statistics literature, this is called the autocovariance function.

and

$$\tau_{\exp} = \sup_{f \in l^2(\pi)} \tau_{\exp, f} \tag{2.9}$$

Thus, $\tau_{\exp}$ is the relaxation time of the slowest mode in the system (it might be $+\infty$). We now show that $\tau_{\exp}$ is directly related to the spectral radius of $P - \Pi$:

Proposition 2.3. $r(P - \Pi) = \exp(-1/\tau_{\exp})$.

Proposition 2.3 is an immediate consequence of (2.7) together with the following generalization of the spectral radius formula:

Proposition 2.4. Let X be a complex Banach space, and let A be a bounded linear operator on X. Then

$$r(A) = \lim_{n \to \infty} \|A^n\|^{1/n} = \inf_{n \geqslant 1} \|A^n\|^{1/n} \tag{2.10a}$$

$$= \sup_{x \in X} \limsup_{n \to \infty} \|A^n x\|^{1/n} \tag{2.10b}$$

$$= \sup_{x \in X, l \in X^*} \limsup_{n \to \infty} |\langle l, A^n x \rangle|^{1/n} \tag{2.10c}$$

If X is a Hilbert space, then also

$$r(A) = \sup_{x \in X} \limsup_{n \to \infty} |(x, A^n x)|^{1/n} \tag{2.10d}$$

Proof. The first line is the well-known spectral radius formula. [*Sketch of proof*: An analyticity argument shows that

$$\limsup_{n \to \infty} \|A^n\|^{1/n} \leqslant r(A) \leqslant \inf_{n \geqslant 1} \|A^n\|^{1/n}$$

(ref. 22, pp. 235–237); alternatively, one can use submultiplicativity to show that $\lim_{n \to \infty} \|A^n\|^{1/n} = \inf_{n \geqslant 1} \|A^n\|^{1/n}$ and then continue with the rest of the analyticity argument (ref. 23, pp. 124–125).]

Clearly, for all $x \in X$ and $l \in X^*$,

$$\limsup_{n \to \infty} |\langle l, A^n x \rangle|^{1/n} \leqslant \limsup_{n \to \infty} \|A^n x\|^{1/n} \leqslant \limsup_{n \to \infty} \|A^n\|^{1/n}$$

On the other hand, if

$$\lambda > c(A) \equiv \sup_{x \in X, l \in X^*} \limsup_{n \to \infty} |\langle l, A^n x \rangle|^{1/n}$$

then the sequence $\{\lambda^{-n} \langle l, A^n x \rangle\}_{n=1}^{\infty}$ is bounded for all $x \in X$ and $l \in X^*$. By the uniform boundedness theorem, it follows that the sequence

$\{\lambda^{-n}\|A^n\|\}_{n=1}^{\infty}$ is bounded, hence that $\limsup_{n\to\infty}\|A^n\|^{1/n}\leqslant\lambda$. Since $\lambda>c(A)$ was arbitrary, we conclude that $\limsup_{n\to\infty}\|A^n\|^{1/n}\leqslant c(A)$.

If X is a *complex* Hilbert space, we have by the polarization identity

$$(x, A^n y)=\frac{1}{4}\left[((x+y), A^n(x+y))-((x-y), A^n(x-y))\right]$$
$$-\frac{i}{4}\left[((x+iy), A^n(x+iy))-((x-iy), A^n(x-iy))\right] \quad (2.11)$$

so that

$$\sup_{x, y\in X}\ \limsup_{n\to\infty}|(x, A^n y)|^{1/n}\leqslant\sup_{x\in X}\ \limsup_{n\to\infty}|(x, A^n x)|^{1/n} \quad \blacksquare$$

Remark. This proof is implicit in Halmos (ref. 24, pp. 232–233) and probably other places.

In general the supremum in (2.10d) cannot be restricted to a dense subset $D\subset X$: take, for example, $X=l^2(\mathbf{Z})$, $D=$ vectors of finite support, and $A=$ shift. However, if A is self-adjoint, we have:

Proposition 2.5. Let X be a complex Hilbert space and let A be a bounded self-adjoint operator on X. Then

$$\|A\|=r(A)=\sup_{x\in D}\ \limsup_{n\to\infty}|(x, A^n x)|^{1/n} \quad (2.12a)$$

for any dense set $D\subset X$.

If, in addition, X is equipped with a distinguished complex conjugation[6] and A is reality-preserving, then

$$\|A\|=r(A)=\sup_{x\in D}\ \limsup_{n\to\infty}|(x, A^n x)|^{1/n} \quad (2.12b)$$

for any dense set $D\subset X_{\text{real}}$.

Proof. It is well known that $\|A\|=r(A)$. Now let $E_A(\cdot)$ be the spectral projections for A. By definition of $r(A)$ we have $E_A(S_\varepsilon)\neq 0$ for all $\varepsilon>0$, where

$$S_\varepsilon\equiv[-r(A), -r(A)+\varepsilon)\cup(r(A)-\varepsilon, r(A)]$$

So fix $\varepsilon>0$ and choose $x\in D$ such that $E_A(S_\varepsilon)x\neq 0$. Then

$$(x, A^n x)=\int\lambda^n\,(x, E_A(d\lambda)x) \quad (2.13)$$

[6] A *complex conjugation* is an antilinear map $C: X\to X$ which is involutive ($C^2=I$). An element $x\in X$ is called *real* if $Cx=x$; we let X_{real} be the set of all real elements of X. A linear operator A on X is called *reality-preserving* if Ax is real whenever x is.

which for n even is

$$\geqslant [r(A)-\varepsilon]^n \, (x, E_A(S_\varepsilon)x) \tag{2.14}$$

with $(x, E_A(S_\varepsilon)x) = \|E_A(S_\varepsilon)x\|^2 > 0$. Since ε is arbitrary, the claim (2.12a) follows.

Now assume that X has a distinguished complex conjugation and that A is self-adjoint and reality-preserving. Then, if $x = y + iz$ with y, z real, we have

$$(x, A^n x) = (y, A^n y) + (z, A^n z) \tag{2.15}$$

from which (2.12b) easily follows. ▮

Remark. Proposition 2.5 is implicit in a paper of Holley on stochastic Ising models (ref. 25, Lemma 1.13).

Finally, a Markov chain is called *reversible*[7] (with respect to the measure π) if

$$\pi_i p_{ij} = \pi_j p_{ji} \tag{2.16}$$

for all $i, j \in S$. [Summing (2.16) over i, we see that the measure π is necessarily a stationary distribution for P.] Reversibility is equivalent to the self-adjointness of P as an operator on the space $l^2(\pi)$. Thus, for reversible Markov chains the spectrum of P lies in the interval $[-1, 1]$, and τ_{exp} is determined by the spectrum of $P \upharpoonright \mathbf{1}^\perp$ closest to either 1 or -1. For many purposes, however, the spectrum near -1 is of little importance; only the spectrum near 1 matters.[8] Therefore, it is worth defining a *modified autocorrelation time* τ'_{exp} based on the spectrum near $+1$ only:

$$\tau'_{\text{exp}} \equiv \frac{-1}{\log[\sup \operatorname{spec}(P - \Pi)]} \tag{2.17}$$

(compare Proposition 2.3).

[7] For the physical significance of this term, see Kemeny and Snell (ref. 26, Section 5.3) or Iosifescu (ref. 27, Section 4.5).

[8] For example, in Monte Carlo work, the statistical errors are proportional to $(1+\lambda)/(1-\lambda)$, where λ is in the spectrum of P [see, e.g., ref. 6, Eqs. (2.19) and (2.23); and see ref. 28 for a rigorous central limit theorem]. Thus, spectrum near -1 is actually *helpful*; only spectrum near $+1$ corresponds to harmful critical slowing down. Another way of seeing that spectrum near -1 is harmless is to note that replacing P by $(I+P)/2$ removes all spectrum near -1; algorithmically, this corresponds to randomly deciding at each time step either to use P or else to do nothing (each with probability 1/2). Of course, from a practical point of view such an algorithm would be rather silly—it is just the original algorithm with half of the time wasted doing nothing—but the fact that it has rapid convergence to equilibrium implies, by a kind of *reductio ad absurdum*, that the original algorithm must also have rapid convergence to equilibrium for all *practical* purposes.

3. RANDOM WALK WITH INWARD DRIFT ON A COUNTABLE ROOTED GRAPH

In this section we define our random-walk model and analyze its properties. Our graph-theoretic terminology generally follows Essam and Fisher,[29] except that our graphs need not be finite.

Let $G = (V, E, \mathbf{0})$ be a countable, connected, rooted graph with vertex set V, edge set E, and a distinguished vertex $\mathbf{0}$, called the *root*. The *level* of a vertex x, denoted $|x|$, is the number of edges in the shortest path which connects x to the root. We write c_N for the number of vertices of level N ($N = 0, 1, 2,...$). If x is adjacent to y, then $|y|$ must be either $|x| - 1$, $|x|$, or $|x| + 1$; we call y a *parent*, *sibling*, or *child* of x, respectively, and write $p(x)$, $s(x)$, and $c(x)$ for the number of parents, siblings, and children of x. Each vertex other than the root must have at least one parent. We remark that G is a tree if and only if each vertex other than the root has precisely one parent and no siblings. Finally, we say that y is a *descendant* of x (and that x is an *ancestor* of y), denoted $x \preccurlyeq y$, if there exists a path of length $|y|$ that contains y, x, and the root. Equivalently, y is a descendant of x iff it is either x itself, or a child of x, or a child of a child of x, etc. We denote by V_x the set of all descendants of x, and by $G_x = (V_x, E_x, x)$ the associated rooted graph with x as the root.

Rooted graphs $G = (V, E, \mathbf{0})$ and $G' = (V', E', \mathbf{0}')$ are said to be *isomorphic* if there is an isomorphism of (V, E) onto (V', E') which takes $\mathbf{0}$ onto $\mathbf{0}'$. A *rooted subgraph* of $G = (V, E, \mathbf{0})$ is a rooted graph $G_1 = (V_1, E_1, \mathbf{0})$, where (V_1, E_1) is a subgraph of (V, E) containing $\mathbf{0}$.

A connected rooted graph $G = (V, E, \mathbf{0})$ is said to be *Cayley* (resp. *sub-Cayley*) if, for each $x \in V$, the rooted graph $G_x = (V_x, E_x, x)$ is isomorphic to G (resp. to a rooted subgraph of G). One example is the *Cayley rooted tree of order* q, in which the root has q children, each of these has q children, and so on indefinitely. Several important examples of sub-Cayley rooted trees will be given below. Note that every sub-Cayley rooted graph satisfies the submultiplicativity condition $c_{M+N} \leqslant c_M c_N$ (all M, N).

We now restrict attention to graphs satisfying

$$\sup_x p(x) \leqslant M_p < \infty \tag{3.1}$$

$$\sup_x c(x) \leqslant M_c < \infty \tag{3.2}$$

It follows that

$$\mu \equiv \limsup_{N \to \infty} c_N^{1/N} \leqslant M_c < \infty \tag{3.3}$$

We call μ the *growth factor* of the rooted graph G. We define for $\beta \geqslant 0$ the generating function

$$Z(\beta) \equiv \sum_{x \in V} \beta^{|x|} = \sum_{N=0}^{\infty} c_N \beta^N \tag{3.4}$$

$Z(\beta)$ is finite for $0 \leqslant \beta < \mu^{-1}$ and infinite for $\beta > \mu^{-1}$.

Fix now a countable, connected, rooted graph $G = (V, E, \mathbf{0})$ satisfying (3.1)–(3.2), fix $\beta > 0$, and fix $M \geqslant M_0 \equiv \sup_x [p(x) + \beta c(x)]$. We can then define a discrete-time Markov chain with state space V having transition probabilities

$$P(x, y) = \begin{cases} 1/M & \text{if } \; y \text{ is a parent of } x \\ \beta/M & \text{if } \; y \text{ is a child of } x \\ \{M - [p(x) + \beta c(x)]\}/M & \text{if } \; y = x \\ 0 & \text{otherwise} \end{cases} \tag{3.5}$$

We call this Markov chain the *standard discrete-time random walk on G with parameters* (β, M). It is an irreducible reversible Markov chain with invariant measure

$$\pi(x) = \text{const} \times \beta^{|x|} \tag{3.6}$$

π is finite iff $Z(\beta) < \infty$; in this case the Markov chain is positive-recurrent, and we normalize π to be a probability measure

$$\pi(x) = Z(\beta)^{-1} \, \beta^{|x|} \tag{3.7}$$

Our goal in this paper is to prove bounds on the spectrum of $P \restriction \mathbf{1}^{\perp}$ considered as a (self-adjoint) operator on $l^2(\pi)$. Of particular interest is the behavior of the spectral gap $m \equiv 1 - \sup \operatorname{spec}(P \restriction \mathbf{1}^{\perp})$ as β approaches the *critical point* $\beta_c \equiv \mu^{-1}$.

Similarly, we can define a continuous-time Markovian jump process with state space V having transition rates

$$J(x, y) = \begin{cases} 1 & \text{if } \; y \text{ is a parent of } x \\ \beta & \text{if } \; y \text{ is a child of } x \\ 0 & \text{otherwise} \end{cases} \tag{3.8}$$

We call this process the *standard continuous-time random walk on G with parameter β*. It is an irreducible reversible Markov process with invariant measure π, and is positive-recurrent iff $Z(\beta) < \infty$. Our proofs for the discrete-time problem can easily be adapted to prove bounds on the spectrum of $\tilde{J} \upharpoonright \mathbf{1}^{\perp}$, where $\tilde{J}$ is the infinitesimal generator defined by

$$(\tilde{J}f)(x) = \sum_{y} J(x, y)[f(x) - f(y)] \tag{3.9}$$

Examples. 1. Let V be the nonnegative integers $\mathbf{Z}_+$, let E be the nearest-neighbor bonds, and let the root be 0. This rooted graph is in fact the Cayley rooted tree of order 1. The Markov chain (3.5) is the random walk with constant drift on $\mathbf{Z}_+$, with elastic boundary conditions at 0. Fix $0 < \beta < 1 \equiv \mu^{-1}$ and for simplicity take $M = M_0 \equiv 1 + \beta$. Then the spectrum of $P \upharpoonright \mathbf{1}^{\perp}$ is the interval $[-\lambda, \lambda]$, where

$$\lambda = 2\beta^{1/2}/(1+\beta) = 1 - \frac{1}{8}\langle N \rangle^{-2} + O(\langle N \rangle^{-3})$$

$$\langle N \rangle \equiv \sum_{x} \pi(x)\,|x| = \frac{d}{d\beta} \log Z(\beta)$$

(see ref. 3, Appendix A, or ref. 30). For other properties of this example, see refs. 8, 31, and 32.

2. Let $\mathscr{L}$ be a regular lattice[9] with coordination number q (e.g., $\mathscr{L} = \mathbf{Z}^d$ with $q = 2d$), and let V be the set of all walks on $\mathscr{L}$ (of arbitrary length) starting at the origin and ending anywhere. We give V the structure of a rooted tree by declaring the zero-step walk to be the root and declaring ω' to be a child of ω if it is a one-step extension of ω. This tree is precisely the Cayley rooted tree of order q. Some of the properties of the random walk (3.5) in this case have been computed by Berretti and Sokal (ref. 3, Appendix A).

3. Same as Example 2, but now let V be the set of all *self-avoiding* walks on $\mathscr{L}$ which start at the origin and end anywhere. This is a *sub-Cayley* rooted tree: every descendant $\tilde{\omega}$ of ω can be written uniquely as[10] $\tilde{\omega} = \omega \circ \omega'$, where $\omega, \omega' \in V$, since every segment of a self-avoiding walk must itself be self-avoiding. However, this is *not* a Cayley rooted tree, since not every walk of the form $\omega \circ \omega'$ with $\omega, \omega' \in V$ is self-avoiding. The discrete-time random walk (3.5) (with $M = 1 + \beta q$) is the transition matrix of a Monte Carlo algorithm for self-avoiding walks first proposed by Berretti and Sokal.[(3)]

[9] A *regular lattice* is, by definition, a countable Abelian group which is endowed with a translation-invariant graph structure. The *coordination number* of a regular lattice is the number of vertices adjacent to any given vertex; we assume that this number is finite.

[10] The symbol $\circ$ denotes concatenation. That is, if $\omega = (\omega_0, \ldots, \omega_M)$ and $\omega' = (\omega'_0, \ldots, \omega'_N)$ with $\omega_0 = \omega'_0 = 0$, then $\omega \circ \omega' = (\omega_0, \ldots, \omega_M, \omega_M + \omega'_1, \ldots, \omega_M + \omega'_N)$.

One rather crude intuition about the behavior of the random walk (3.5) [or (3.8)] was set forth by Berretti and Sokal.[3] They argued that if one looks only at the level $|x|$, then this quantity executes a random walk with drift on the nonnegative integers. (This random walk is not precisely Markovian, nor are its transition probabilities precisely those of Example 1 above, but no matter.) They then argued, by analogy with Example 1, that the hitting time to the root should typically be of order $\langle N \rangle^2$; and since each visit to the root (or any other chosen state, for that matter) erases "memory" of the past, the autocorrelation time τ should be of order $\langle N \rangle^2$.

A more careful way of expressing this intuition is to consider the *aggregated Markov chain*,[33,9] in which all the states of a given level are lumped into a single state. The transition matrix of the aggregated chain can be thought of as (3.5) followed by a randomization operation which redistributes the walker uniformly around the states of its current level. This randomization would intuitively be expected to accelerate the convergence to equilibrium, and this can in fact be proven: the spectral gap of the aggregated chain is a rigorous upper bound on the spectral gap of the original chain. An analysis of the aggregated chain (under the assumption $c_N \sim \mu^N N^{\gamma-1}$) then yields the rigorous *lower bound* $\tau \geqslant \text{const} \times \langle N \rangle^2$ (see Section 4).

The autocorrelation time τ could, however, be much larger than this lower bound, if there exist modes which relax significantly more slowly than the level $|x|$. Whether or not this occurs depends on the detailed structure of the graph G, and not only on the $\{c_N\}$. One way of seeing this is to note that a walker at a site $x \neq \mathbf{0}$ feels a net drift "inward" (i.e., toward the root) if $\beta c(x) < p(x)$, and a net drift "outward" (i.e., away from the root) if $\beta c(x) > p(x)$. We distinguish three cases:

1. If $\beta c(x) < p(x)$ for *all* $x \neq \mathbf{0}$, then the drift is "uniformly inward," and a relatively straightforward Liapunov-function argument can be employed to prove geometric ergodicity (see Section 5). In particular, this occurs if $\beta < M_c^{-1}$.

2. If $\beta > \mu^{-1}$, then the drift is "on the average outward," and a finite invariant measure π does not exist. (The Markov chain is thus either null-recurrent or transient; it might be amusing to determine which.)

3. If $M_c^{-1} \leqslant \beta < \mu^{-1}$, then the situation is much more delicate. "On the average" the drift is inward—that is why there exists an exponentially decaying invariant measure π—but at certain sites $x \neq \mathbf{0}$ the local drift may be outward. In particular, the graph G may contain large connected regions in which the drift is outward, and this can spoil the geometric ergodicity (Example 4.1). On the other hand, if G is a *sub-Cayley tree*, we shall show that such pathologies cannot occur:

Theorem 3.1. Let G be a sub-Cayley rooted tree satisfying (3.1)–(3.2), and let $\beta < \mu^{-1}$. Define

$$\bar{r} = \sup_{1 < K^* < (\beta\mu)^{-1}} \left[1 - \frac{\beta}{M} \frac{K^* - 1}{Z(\beta K^*) - 1} \right]^{-1} \tag{3.10}$$

(clearly $\bar{r} > 1$). Then the hitting time to the root, τ_0, satisfies

$$E_x[\bar{r}^{\tau_0}] < \infty \tag{3.11}$$

for all x, and the spectrum of $P \upharpoonright \mathbf{1}^\perp$ is contained in the interval

$$[1 - 2M_0/M, \bar{r}^{-1}] \tag{3.12}$$

where $M_0 \equiv \sup_x [p(x) + \beta c(x)]$.

Theorem 3.1 (together with its consequence, Corollary 3.1) is the main result of this paper; we now give a brief outline of its proof.

The main thrust of our proof is to show that the hitting times to the root have an exponentially decaying density, in the sense that $E_x[r^{\tau_0}] < \infty$ for some $r > 1$. (We obtain quantitative bounds on r in terms of β, M, and the function Z.) First we use the fact that G is a sub-Cayley tree to bound hitting times from x to the root in terms of hitting times from a child of the root to the root (Lemmas 3.1–3.3). Next we use a beautiful identity (Lemma 3.4) to relate hitting times from x to the root, averaged over the invariant measure π, to return times from the root to itself. Finally, we use the explicit transition matrix (3.5) to relate return times from the root to itself to hitting times from a child of the root to the root (Lemma 3.5). Putting this all together, we obtain an algebraic inequality for the hitting time from a child of the root to the root in terms of itself (Lemma 3.6). For $r - 1$ sufficiently small, the solution set of this algebraic inequality consists of two disconnected intervals $[1, K_1]$ and $[K_2, +\infty]$ (Lemma 3.7). Our goal is to show that the true value lies in the lower of these two intervals. To do this, we argue as follows: If G is a *finite* sub-Cayley tree, then $E_x[r^{\tau_0}]$ is a *continuous* extended-real-valued function of r (Lemma 3.8); and for $r = 1$ it obviously equals 1. Therefore, by continuity, $E_x[r^{\tau_0}]$ must lie in the lower of the two intervals whenever r is small enough so that the two intervals are disconnected (Lemma 3.9). If G is an *infinite* sub-Cayley tree, then the foregoing argument can be applied to finite "cutoff" trees $G^{(N)}$, yielding bounds on $E_x[r^{\tau_0}]$ which are *uniform* in N; and these uniform bounds can then be carried over to the original tree G (Lemma 3.10). This proves that hitting times to the root decay exponentially; we then use an argument based on Dirichlet boundary conditions to relate this to the L^2 spectrum (Lemmas 3.11 and 3.12).

Let us emphasize the logical structure of this proof. For the first half of the proof (through Lemma 3.7), we obtain bounds relating the expectations $E_x[r^{\tau_0}]$ for different x, but for all we know these expectations might be $+\infty$! Only through the continuity argument for cutoff trees (Lemmas 3.8 and 3.9) do we obtain bounds which are guaranteed to be finite. Moreover, these finite bounds are *uniform*, so the cutoff can be removed.[11] Arguments of this type have been used in other areas of mathematical statistical mechanics and quantum field theory by Brydges *et al.*[34] and by Slade.[35]

We now proceed with the proof of Theorem 3.1. Some of the lemmas below are stated in greater generality than we shall really need; the aim is to show their "natural" context. We need a few definitions: If $G=(V, E)$ is a graph, we say that a Markov chain with state space V is a Markov chain on G if its transition matrix $P=\{p_{ij}\}_{i,j\in V}$ satisfies $p_{ij}=0$ whenever $i\neq j$ with $\{i, j\}\notin E$. (By abuse of language, we shall often fail to distinguish between G and V.) For any Markov chain and any subset A of the state space, we define the hitting time $\tau_A \equiv \min\{t\geqslant 1: X_t\in A\}$; note that $\tau_A\geqslant 1$ always, even if $X_0\in A$. Finally, we assume that the sub-Cayley rooted tree G is nontrivial, i.e., contains at least one vertex other than the root.

Lemma 3.1. Let $P=\{p_{ij}\}_{i,j\in T}$ be the transition matrix for a discrete-time Markov chain on a countable tree T. Let S be a connected subset of T (i.e., a subtree), and define the restricted transition matrix $P^S=\{p^S_{ij}\}_{i,j\in S}$ by

$$p^S_{ij}=\begin{cases} p_{ij} & \text{if } \ i\neq j \\ p_{ii}+\sum_{k\notin S} p_{ik} & \text{if } \ i=j \end{cases} \tag{3.13}$$

Then, for each $x\in S$, there exists a coupled (non-Markovian) stochastic process $X_t=(X^S_t, X^T_t)$ on $S\times T$ such that:

1. $X^S_0=X^T_0=x$.
2. $\{X^S_t\}$ is Markovian with transition matrix P^S.
3. $\{X^T_t\}$ is Markovian with transition matrix P.
4. $X^S_t=X^T_t$ for $t<\tau^{(T)}_{S^c}\equiv\min\{u: X^T_u\notin S\}$.
5. For any set $A\subset S$, $\tau^{(S)}_A\equiv\min\{t\geqslant 1: X^S_t\in A\}\leqslant\tau^{(T)}_A\equiv\min\{t\geqslant 1: X^T_t\in A\}$.

[11] Otherwise put: We state Lemmas 3.1–3.7 for general sub-Cayley trees, but we only *use* them for *finite* sub-Cayley trees. For finite state space it can be proven by general arguments that $E_x[r^{\tau_0}]<\infty$ for *some* $r>1$, but this is insufficient for our purposes, since we need bounds which are *uniform* in N. To this end we employ Lemmas 3.1–3.7 together with the continuity argument of Lemmas 3.8–3.10.

Proof. For each history $\{X_t^T\}$ of the Markov process P (with initial condition x), define $\{X_t^S\}$ to be the subsequence of $\{X_t^T\}$ consisting of those entries which are in S. It is not hard to see that $\{X_t^S\}$ is Markovian with transition probability P^S. (The key fact is that since T is a tree and S is connected, whenever the process $\{X_t^T\}$ leaves S it must reenter S at the same point from which it left.) Facts 4 and 5 are then obvious. ▮

Corollary. The condition

$$E_x^{(P^S)}[r^{\tau_A}] \leqslant E_x^{(P)}[r^{\tau_A}]$$

holds for all $x \in S$, $A \subset S$, and $r \geqslant 1$.

Lemma 3.2. Consider a discrete-time Markov chain whose state space is a countable tree T. Let $x \neq y$ be points in T, and let $x \equiv x_0, x_1, ..., x_n \equiv y$ be the unique self-avoiding path in T from x to y. Then, for any $r > 0$,

$$E_x[r^{\tau_y}] = \prod_{i=0}^{n-1} E_{x_i}[r^{\tau_{x_{i+1}}}] \tag{3.14}$$

Proof. This is an immediate consequence of the strong Markov property. ▮

Lemma 3.3. For the random walk (3.5) on a sub-Cayley rooted tree,

$$E_x[r^{\tau_{\mathrm{parent}(x)}}] \leqslant \max_{|y|=1} E_y[r^{\tau_{\mathbf{0}}}] \equiv K(r) \tag{3.15}$$

and

$$E_x[r^{\tau_{\mathbf{0}}}] \leqslant K(r)^{|x|} \tag{3.16}$$

for any $r \geqslant 1$ and any $x \neq \mathbf{0}$.

Proof. The first inequality is an immediate consequence of the sub-Cayley hypothesis and the "translation invariance" of (3.5), together with Lemma 3.1. The second inequality is an immediate consequence of the first, together with Lemma 3.2. ▮

Lemma 3.4. For the random walk (3.5),

$$E_0[r^{\tau_{\mathbf{0}}}] \leqslant r + (r-1) \sum_{x \neq \mathbf{0}} \beta^{|x|} E_x[r^{\tau_{\mathbf{0}}}] \tag{3.17}$$

for any $r \geqslant 1$; equality holds if the right-hand side is finite.

Proof. For any x we have

$$E_x[r^{\tau_0}] = r \sum_{y \neq \mathbf{0}} P(x, y)\, E_y[r^{\tau_0}] + rP(x, \mathbf{0}) \tag{3.18}$$

(just let the Markov chain take one step starting at x and keep track of where it lands). In particular, if the right-hand side of (3.17) is finite, then $E_x[r^{\tau_0}]$ is finite for all x. Summing (3.18) against the invariant measure $\beta^{|x|}$ and interchanging the order of summation (everything is nonnegative), we obtain

$$\begin{aligned}\sum_x \beta^{|x|} E_x[r^{\tau_0}] &= r \sum_x \sum_{y \neq \mathbf{0}} \beta^{|x|} P(x, y)\, E_y[r^{\tau_0}] + r \sum_x \beta^{|x|} P(x, \mathbf{0}) \\ &= r \sum_{y \neq \mathbf{0}} \beta^{|y|} E_y[r^{\tau_0}] + r \end{aligned} \tag{3.19}$$

where on the second line we twice used $\sum_x \beta^{|x|} P(x, y) = \beta^{|y|}$. Since by hypothesis $\sum_{y \neq \mathbf{0}} \beta^{|y|} E_y[r^{\tau_0}] < \infty$, we can rearrange (3.19) to deduce equality in (3.17). On the other hand, if the right-hand side of (3.17) is infinite, the inequality is obvious. ∎

Remarks. 1. With a little more work, it can be shown that equality holds in (3.17) even if the right-hand side is infinite.

2. Relation (3.17) is a special case of a more general formula holding for irreducible positive-recurrent Markov chains:

$$\sum_{x \in A} \pi(x)\, E_x[f(\tau_A)] = \pi(A)\, f(1) + \sum_{x \notin A} \pi(x)\, E_x[(\Delta f)(\tau_A)] \tag{3.20}$$

where A is an arbitrary nonempty set, f is an arbitrary real-valued function on $\{1, 2, 3,...\}$ satisfying mild regularity conditions (e.g., f monotonic or bounded will do), and $(\Delta f)(n) \equiv f(n+1) - f(n)$. See Cogburn (ref. 36, Lemma 3.4) for a proof. Some important special cases of (3.20) are

$$\sum_{x \in A} \pi(x)\, E_x[\tau_A] = 1 \tag{3.21}$$

$$\sum_{x \in A} \pi(x)\, E_x[\tau_A^2] = 1 + 2 \sum_{x \notin A} \pi(x)\, E_x[\tau_A] \tag{3.22}$$

$$\sum_{x \in A} \pi(x)\, E_x[r^{\tau_A}] = r\pi(A) + (r-1) \sum_{x \notin A} \pi(x)\, E_x[r^{\tau_A}] \tag{3.23}$$

In particular, (3.21) is familiar from renewal theory. See, e.g., Nummelin (ref. 20, Chapter 5) for applications.

Lemma 3.5. For the random walk (3.5),

$$E_0[r^{\tau_0}] = \frac{M - \beta c(\mathbf{0})}{M} r + \frac{\beta}{M} r \sum_{|y|=1} E_y[r^{\tau_0}] \tag{3.24a}$$

$$\geqslant r + \frac{\beta r}{M} [K(r) - 1] \tag{3.24b}$$

for any $r \geqslant 1$.

Proof. The equality is just the special case $x = \mathbf{0}$ of (3.18). The inequality holds because $E_y[r^{\tau_0}] = K(r)$ for at least one child y of the root, and $E_y[r^{\tau_0}] \geqslant r \geqslant 1$ for all the others. ∎

Lemma 3.6. For the random walk (3.5) on a sub-Cayley rooted tree,

$$K(r) \leqslant 1 + \frac{M}{\beta r}(r-1)[Z(\beta K(r)) - 1] \tag{3.25}$$

for any $r \geqslant 1$.

Proof. This is an immediate consequence of Lemmas 3.3–3.5. ∎

Lemma 3.7. Consider the random walk (3.5) on a sub-Cayley rooted tree. Let K^* be any number in the interval $1 < K^* < (\beta\mu)^{-1}$, and define

$$r^* \equiv \left[1 - \frac{\beta}{M} \frac{K^* - 1}{Z(\beta K^*) - 1}\right]^{-1} \tag{3.26}$$

(clearly $r^* > 1$). Then, for any $\varepsilon > 0$, there exists $\delta > 0$ such that $1 \leqslant r \leqslant r^* - \varepsilon$ implies $K(r) \notin (K^* - \delta, K^*]$.

Proof. If $K(r) \leqslant K^*$, then $Z(\beta K(r)) \leqslant Z(\beta K^*) < \infty$. It then follows immediately from (3.25) that $K(r) \leqslant K^* - \delta$, where $\delta \equiv (\varepsilon M/\beta r^{*2})[Z(\beta K^*) - 1]$. ∎

Lemma 3.8. For any irreducible Markov chain on a *finite* state space S, any initial distribution α, and any nonempty set $A \subset S$, the quantity $E_\alpha[r^{\tau_A}]$ is a continuous function of r ($r \geqslant 0$) with values in $[0, +\infty]$.

Proof. An irreducible finite Markov chain is recurrent, so $\tau_A < \infty$ with probability 1. Thus

$$f(r) \equiv E_\alpha[r^{\tau_A}] = \sum_{n=1}^{\infty} a_n r^n$$

with $a_n = P_\alpha[\tau_A = n] \geqslant 0$. Now define $R = \liminf_{n\to\infty} a_n^{-1/n}$. Clearly, $f(r)$ is an increasing function of $r \geqslant 0$ which is finite for $0 \leqslant r < R$ and equal to $+\infty$ for $r > R$. Moreover, f is an analytic (hence continuous) function in the disk $|r| < R$, which has (by the Vivanti–Pringsheim theorem) a singularity at $r = R$.[12] So all that remains to be proven is that $\lim_{r\uparrow R} f(r) = +\infty$. Now

$$\begin{aligned} P_\alpha[\tau_A = n] &= P_\alpha[\tau_A > n-1] - P_\alpha[\tau_A > n] \\ &= \langle \alpha, (PI_{A^c})^{n-1}\mathbf{1}\rangle - \langle \alpha, (PI_{A^c})^n \mathbf{1}\rangle \end{aligned} \tag{3.27}$$

where I_{A^c} is the operator of multiplication by the characteristic function χ_{A^c}. Now multiply both sides of (3.27) by r^n and sum from $n=1$ to infinity: both sides certainly converge for $|r| < 1$, and we obtain there

$$f(r) = \langle \alpha, [(r-1)(I - rPI_{A^c})^{-1} + I]\,\mathbf{1}\rangle \tag{3.28}$$

By well-known matrix theory, the right-hand side of (3.28) is a rational function of r; and by analytic continuation, it must equal $f(r)$ throughout the region of analyticity $|r| < R$. It follows that f has a pole at $r = R$, so that $\lim_{r\uparrow R} f(r) = +\infty$. ∎

Lemma 3.9. Consider the random walk (3.5) on a *finite* sub-Cayley rooted tree. Choose any number $K^* > 1$, and define r^* by (3.26). Then, for all r in the interval $1 \leqslant r \leqslant r^*$, we have $K(r) \leqslant K^*$.

Proof. By Lemma 3.8, $K(r) \equiv \max_{|y|=1} E_y[r^{\tau_0}]$ is a continuous function of r; and clearly $K(1) = 1 < K^*$. By Lemma 3.7, there is a forbidden interval $[K^* - \delta, K^*]$ in which $K(r)$ cannot lie for $1 \leqslant r < r^*$. It follows that $K(r)$ must lie *below* this forbidden interval for all such r. By continuity, the bound $K(r) \leqslant K^*$ holds also for $r = r^*$. ∎

Lemma 3.10. Consider the random walk (3.5) on a sub-Cayley rooted tree (finite or infinite). Let K^* be any number in the interval $1 < K^* < (\beta\mu)^{-1}$, and define r^* by (3.26). Then, for all r in the interval $1 \leqslant r \leqslant r^*$, we have $K(r) \leqslant K^*$.

Proof. For each $N \geqslant 1$, let $G^{(N)}$ be the rooted subgraph of G consisting of all the vertices of level $\leqslant N$ together with the edges connecting them. It is easily seen that $G^{(N)}$ is a finite sub-Cayley rooted tree. Moreover, $Z_{G^{(N)}} \leqslant Z_G$ and hence $r^*_{G^{(N)}} \geqslant r^*_G$. Now let r be in the interval $1 \leqslant r \leqslant r^*_G$. By Lemma 3.9, we have

$$E_y^{(G^{(N)})}[r^{\tau_0}] \leqslant K^*$$

[12] This statement is true also in the case $R = \infty$, by Liouville's theorem (since f is a nonconstant entire function). But this observation is unnecessary, since if $R = \infty$, the proof of the lemma is already complete.

for each vertex y with $|y| = 1$, for all N. Now, since $G^{(1)} \subset G^{(2)} \subset \cdots \subset G$, we can employ the argument in Lemma 3.1 to construct a coupled (non-Markovian) stochastic process with state space $G^{(1)} \times G^{(2)} \times \cdots \times G$ whose marginals are the random walks (3.5) on $G^{(N)}$ and G, and in which $\tau_{\mathbf{0}}^{(G^{(1)})} \leqslant \tau_{\mathbf{0}}^{(G^{(2)})} \leqslant \cdots \leqslant \tau_{\mathbf{0}}^{(G)}$. It then follows by the monotone convergence theorem that

$$E_y^{(G)}[r^{\tau_0}] = \lim_{N \to \infty} E_y^{(G^{(N)})}[r^{\tau_0}] \leqslant K^* \qquad \text{for all } r \leqslant r^* \quad \blacksquare$$

We can now optimize over $K^* \in (1, (\beta\mu)^{-1})$. Define $\bar{r}$ by (3.10); it then follows from Lemmas 3.10 and 3.3 and Eq. (3.24a) that $E_x[\bar{r}^{\tau_0}] < \infty$ for all x. This completes the proof of the first half of Theorem 3.1.

The remainder of the proof is a straightforward argument based on Dirichlet boundary conditions:

Lemma 3.11. Let P_D be the (sub-Markovian) matrix obtained from P by deleting the row and column corresponding to $\mathbf{0}$ [P with Dirichlet conditions imposed at $\mathbf{0}$]. Then P_D is a self-adjoint operator on $l^2(\pi, V\setminus\{\mathbf{0}\})$ with norm $\leqslant \bar{r}^{-1}$.

Proof. The self-adjointness of P_D is obvious. Note next that

$$(P_D^n \mathbf{1})(x) = P_x[\tau_{\mathbf{0}} > n] \leqslant \bar{r}^{-(n+1)} E_x[\bar{r}^{\tau_0}] \tag{3.29}$$

Then, for any ψ supported on a finite set of points,

$$\begin{aligned} |(\psi, P_D^n \psi)_{l^2(\pi, V\setminus\{\mathbf{0}\})}| &\leqslant (|\psi|, P_D^n |\psi|)_{l^2(\pi, V\setminus\{\mathbf{0}\})} \\ &\leqslant \|\psi\|_\infty^2 \sum_{x \in \operatorname{supp} \psi} \pi(x) (P_D^n \mathbf{1})(x) \\ &\leqslant \bar{r}^{-(n+1)} \|\psi\|_\infty^2 \sum_{x \in \operatorname{supp} \psi} \pi(x)\, E_x[\bar{r}^{\tau_0}] \end{aligned} \tag{3.30}$$

where $\|\psi\|_\infty \equiv \sup_x |\psi(x)|$. Since such functions ψ are dense in $l^2(\pi, V\setminus\{\mathbf{0}\})$, it follows from Proposition 2.5 that $\|P_D\| \leqslant \bar{r}^{-1}$. $\blacksquare$

Lemma 3.12. The supremum of the spectrum of $P \upharpoonright \mathbf{1}^\perp$ is $\leqslant \|P_D\|$.

Proof. Let $\phi \in \mathbf{1}^\perp$ and set $c = \phi(\mathbf{0})$. Then

$$\begin{aligned} (\phi, P\phi)_{l^2(\pi)} &= ((\phi - c\mathbf{1}), P(\phi - c\mathbf{1}))_{l^2(\pi)} - |c|^2 \\ &= ((\phi - c\mathbf{1}), P_D(\phi - c\mathbf{1}))_{l^2(\pi, V\setminus\{\mathbf{0}\})} - |c|^2 \\ &\leqslant \|P_D\|\ \|\phi - c\mathbf{1}\|_{l^2(\pi)}^2 - |c|^2 \\ &= \|P_D\| (\|\phi\|_{l^2(\pi)}^2 + |c|^2) - |c|^2 \\ &\leqslant \|P_D\|\ \|\phi\|_{l^2(\pi)}^2 \end{aligned} \tag{3.31}$$

This proves the claim. $\blacksquare$

Remark. Lemma 3.12 can alternatively be proven using the min-max theorem[37] or as a corollary of a more general theorem on Dirichlet boundary conditions.[7]

Completion of the Proof of Theorem 3.1. All that remains to be proven is that the infimum of the spectrum of P is $\geqslant 1 - 2M_0/M$. To see this, note that, by (3.5),

$$P = I + \frac{1}{M} Q \tag{3.32}$$

where Q is self-adjoint and negative-semidefinite. Since $\|P\| = 1$ when $M \geqslant M_0$, we must have $\operatorname{spec}(Q) \subset [-2M_0, 0]$. Hence, $\operatorname{spec}(P) \subset [1 - 2M_0/M, 1]$. ∎

Theorem 3.1 gives a quantitative bound on the spectral gap, and hence on the modified autocorrelation time τ'_{exp}. This bound takes a simple form if we assume that c_N, the number of vertices in the tree at level N, has the asymptotic behavior $c_N \sim \mu^N N^{\gamma - 1}$ for some "critical exponent" γ, as is believed to occur for self-avoiding walks. Note that the submultiplicativity property $c_{M+N} \leqslant c_M c_N$, valid for every sub-Cayley rooted graph, implies[38] that $c_N \geqslant \mu^N$; so γ, if it exists, must be $\geqslant 1$.

Corollary 3.1. Under the hypotheses of Theorem 3.1, we have

$$\tau'_{\text{exp}} \leqslant \frac{M}{\beta} \inf_{1 < K^* < (\beta\mu)^{-1}} \frac{Z(\beta K^*) - 1}{K^* - 1} \tag{3.33}$$

If, in addition, $c_N \leqslant A\mu^N N^{\gamma - 1}$ for N sufficiently large, then there exists a constant C_1 (independent of β) such that

$$\tau'_{\text{exp}} \leqslant C_1 (1 - \beta\mu)^{-\gamma - 1} \tag{3.34}$$

Finally, if the c_N further satisfy

$$\liminf_{N \to \infty} \frac{1}{c_N} [N c_N - (N-1)\, \mu c_{N-1}] = \Gamma > 0 \tag{3.35}$$

then there exists a constant C_2 (independent of β) such that

$$\tau'_{\text{exp}} \leqslant C_2 \langle N \rangle^{1+\gamma} \tag{3.36}$$

for $\beta \uparrow \mu^{-1}$, where

$$\langle N \rangle \equiv Z(\beta)^{-1} \sum_{N=0}^{\infty} N \beta^N c_N = \frac{d}{d\beta} \log Z(\beta) \tag{3.37}$$

Remark. If $c_N \sim \mu^N N^{\gamma-1}$, then (3.35) holds with $\Gamma = \gamma$.

Proof. Inequality (3.33) is basically the translation of Theorem 3.1 into the language of autocorrelation times: it follows from $\exp(-1/\tau'_{\exp}) \leqslant \bar{r}^{-1}$ and the elementary inequality $\log(1-x) \leqslant -x$ for $0 \leqslant x \leqslant 1$.

Next assume that $c_N \leqslant A\mu^N N^{\gamma-1}$ for N sufficiently large. Then

$$\begin{aligned} Z(\beta K^*) &= \sum_{N=0}^{\infty} (\beta K^*)^N c_N \\ &\leqslant A \sum_{N=0}^{\infty} (\beta\mu K^*)^N N^{\gamma-1} + \text{polynomial in } \beta \\ &\leqslant A' \sum_{N=0}^{\infty} (\beta\mu K^*)^N (-1)^N \binom{-\gamma}{N} + \text{polynomial in } \beta \\ &= A'(1-\beta\mu K^*)^{-\gamma} + \text{polynomial in } \beta \\ &\leqslant A''(1-\beta\mu K^*)^{-\gamma} \end{aligned} \tag{3.38}$$

uniformly for $0 \leqslant \beta\mu K^* < 1$. Using this inequality to estimate the right-hand side of (3.33) and computing the resulting infimum, we obtain (3.34).

Finally, by hypothesis (3.35), for any $\varepsilon > 0$ we have

$$\begin{aligned} (1-\beta\mu)\langle N\rangle &= \frac{\sum_{N=0}^{\infty} [Nc_N - (N-1)\,\mu c_{N-1}]\,\beta^N}{\sum_{N=0}^{\infty} c_N \beta^N} \\ &\geqslant (\Gamma - \varepsilon) + \frac{\text{polynomial in } \beta}{\sum_{N=0}^{\infty} c_N \beta^N} \end{aligned} \tag{3.39}$$

But $c_N \geqslant \mu^N$ by the sub-Cayley hypothesis, so $\sum_{N=0}^{\infty} c_N \beta^N \geqslant (1-\beta\mu)^{-1}$; hence

$$\liminf_{\beta \uparrow \mu^{-1}} (1-\beta\mu)\langle N\rangle \geqslant \Gamma \tag{3.40}$$

from which (3.36) follows. ∎

The result of Theorem 3.1 can be extended to a slightly larger class of Markov chains by a well-known comparison principle: Let P_1 and P_2 be transition probability matrices on the same state space which are reversible with respect to the same measure π. Assume that the off-diagonal matrix elements of P_2 dominate those of P_1, i.e., $(P_1)_{xy} \leqslant (P_2)_{xy}$ for all $x \neq y$. Then

$$(f, (I-P_1)f)_{l^2(\pi)} = \frac{1}{2}\sum_{x,y} \pi_x (P_1)_{xy} [f(x)-f(y)]^2 \tag{3.41}$$

$$\leqslant \frac{1}{2}\sum_{x,y} \pi_x (P_2)_{xy} [f(x)-f(y)]^2$$

$$= (f, (I-P_2)f)_{l^2(\pi)} \tag{3.42}$$

for any $f \in l^2(\pi)$, so that $P_1 \geqslant P_2$ in the sense of quadratic forms. The Rayleigh–Ritz principle (or the min-max theorem) then implies that $\sup \operatorname{spec}(P_1 \upharpoonright \mathbf{1}^\perp) \geqslant \sup \operatorname{spec}(P_2 \upharpoonright \mathbf{1}^\perp)$; hence, that $\tau'_{\exp,1} \geqslant \tau'_{\exp,2}$.

In particular, consider two rooted graphs $G_1 = (V, E_1, \mathbf{0})$ and $G_2 = (V, E_2, \mathbf{0})$ with the same vertex set V and with $E_1 \subset E_2$. Assume, in addition, that $|x|_{G_1} = |x|_{G_2}$ for each $x \in V$, i.e., the distance from x to $\mathbf{0}$ is the same for both graphs. [Equivalently, G_2 is obtained from G_1 by adding edges $(E_2 \setminus E_1)$ which connect vertices on the *same* or *neighboring levels.*] In this case we call G_2 a *compatible supergraph* of G_1, and G_1 a *compatible subgraph* of G_2. Both graphs have the same $\{c_N\}$ and hence the same μ. Now fix $0 < \beta < \mu^{-1}$ and $M \geqslant \sup_x [p_2(x) + \beta c_2(x)]$, where $p_2(x)$ and $c_2(x)$ are the number of parents and children of x, respectively, in G_2. Let P_1 and P_2 be the corresponding standard transition probabilities defined by (3.5). Then, clearly, $(P_1)_{xy} \leqslant (P_2)_{xy}$ for all $x \neq y$, and hence $\tau'_{\exp}(G_2) \leqslant \tau'_{\exp}(G_1)$.

Otherwise put, let $G = (V, E, \mathbf{0})$ be a countable rooted graph satisfying (3.1)–(3.2), and fix $0 < \beta < \mu^{-1}$ and $M \geqslant \sup_x [p(x) + \beta c(x)]$. Then

$$\tau'_{\exp}(G) \leqslant \inf_H \tau'_{\exp}(H) \tag{3.43}$$

where the infimum is taken over all compatible subgraphs $H \subset G$. In particular, we have:

Corollary 3.2. The results of Theorem 3.1 and Corollary 3.1 remain valid for any countable rooted graph which contains a sub-Cayley rooted tree as a compatible subgraph.

Proof. It suffices to note that the bound (3.10)/(3.33) depends only on the $\{c_N\}$, which are the same for both graphs. ∎

Unfortunately, we do not know any convenient algorithm for testing whether a given rooted graph G satisfies the hypotheses of Corollary 3.2.

Remarks. 1. A similar result for finite graphs with counting measure is proven in ref. 37.

2. Analogous considerations apply to pairs of continuous-time reversible Markovian jump processes with transition rates satisfying $J_1(x, y) \leqslant J_2(x, y)$ for all $x \neq y$.

4. LOWER BOUNDS ON THE AUTOCORRELATION TIME, BY RAYLEIGH–RITZ

We now turn to the variational lower bound on the autocorrelation time τ'_{exp}. Consider an arbitrary countable, connected, rooted graph G satisfying (3.1)–(3.2). Let e_N be the total number of edges which connect vertices of level $N-1$ to those of level N. Thus,

$$e_N = \sum_{|x| = N} p(x) = \sum_{|x| = N-1} c(x) \tag{4.1}$$

In particular, $e_0 = 0$ and $c_N \leqslant e_N \leqslant M_p c_N$ for $N \geqslant 1$ (so $e_N = c_N$ for $N \geqslant 1$ if G is a tree).

Now define the *aggregated Markov chain*[33,9] to have state space $\mathbf{Z}_+ = \{0, 1, 2,...\}$ and transition probabilities

$$\bar{P}(N, N-1) = \frac{e_N}{Mc_N} \tag{4.2a}$$

$$\bar{P}(N, N+1) = \frac{\beta e_{N+1}}{Mc_N} \tag{4.2b}$$

$$\bar{P}(N, N) = 1 - \frac{e_N}{Mc_N} - \frac{\beta e_{N+1}}{Mc_N} \tag{4.2c}$$

This is an irreducible reversible Markov chain with invariant measure

$$\bar{\pi}(N) = \text{const} \times \beta^N c_N \tag{4.3}$$

$\bar{\pi}$ is a finite measure iff $Z(\beta) < \infty$. Now τ'_{exp} of the original Markov chain is bounded below by τ'_{exp} of the aggregated chain, since for any trial function $\bar{f} \in l^2(\mathbf{Z}_+, \bar{\pi})$ we can define a trial function $f \in l^2(V, \pi)$ by $f(x) = \bar{f}(|x|)$, and

$$(f, \mathbf{1})_{l^2(V,\pi)} = (\bar{f}, \mathbf{1})_{l^2(\mathbf{Z}_+,\bar{\pi})} \tag{4.4a}$$

$$(f, f)_{l^2(V,\pi)} = (\bar{f}, \bar{f})_{l^2(\mathbf{Z}_+,\bar{\pi})} \tag{4.4b}$$

$$\begin{aligned}(f, Pf)_{l^2(V,\pi)} &= (\bar{f}, \bar{f})_{l^2(\mathbf{Z}_+,\bar{\pi})} - \frac{1}{MZ(\beta)} \sum_{N=1}^{\infty} e_N \beta^N [\bar{f}(N) - \bar{f}(N-1)]^2 \\ &= (\bar{f}, \bar{P}\bar{f})_{l^2(\mathbf{Z}_+,\bar{\pi})}\end{aligned} \tag{4.4c}$$

Taking the supremum in (4.4c) over all $\bar{f}$ which are normalized and orthogonal to $\mathbf{1}$, we find that

$$\exp\left(\frac{-1}{\tau'_{\text{exp}}}\right) \equiv \sup \operatorname{spec}(P \restriction \mathbf{1}^\perp) \geqslant \sup_{f \perp \mathbf{1}} \frac{(f, Pf)}{(f, f)} = \exp\left(\frac{-1}{\tau'_{\text{exp}}(\text{aggregated})}\right) \tag{4.5}$$

which implies the above assertion.

On the other hand,

$$\sup \operatorname{spec}(P \restriction 1^{\perp}) \geqslant \sup \text{essential spec}(P) \geqslant \sup \text{essential spec}(\bar{P}) \quad (4.6)$$

and in some cases the supremum of the essential spectrum for the aggregated chain can be determined exactly. For example:

Proposition 4.1. Let $\beta > 0$. Suppose that $c_N, e_N > 0$ for all $N \geqslant 1$ and that

$$\lim_{N \to \infty} \frac{c_{N+1}}{c_N} = \mu \quad (4.7)$$

$$\lim_{N \to \infty} \frac{e_N}{c_N} = a \quad (4.8)$$

Then

$$\sup \text{essential spec}(\bar{P}) = 1 - \frac{a}{M} [1 - (\beta\mu)^{1/2}]^2 \quad (4.9)$$

Remarks. 1. If G is a tree, then $a = 1$.

2. This proposition also covers the case $\beta \geqslant \mu^{-1}$, for what it is worth.

Corollary 4.1. Under the same hypotheses, for $\beta \uparrow \mu^{-1}$ we have

$$\tau'_{\exp} \geqslant \frac{-1}{\log\{1 - (a/M)[1 - (\beta\mu)^{1/2}]^2\}} = \frac{4M}{a} (1 - \beta\mu)^{-2} + O((1 - \beta\mu)^{-1}) \quad (4.10)$$

If, in addition, the c_N satisfy

$$\limsup_{N \to \infty} \frac{1}{c_N} [Nc_N - (N-1)\, \mu c_{N-1}] = \Gamma < \infty \quad (4.11)$$

then

$$\tau'_{\exp} \geqslant C \langle N \rangle^2 \quad (4.12)$$

for some $C > 0$ which is independent of β as $\beta \uparrow \mu^{-1}$.

Proof of Corollary 4.1 Assuming Proposition 4.1. The first inequality follows immediately from Proposition 4.1 together with (4.6). Now assume (4.11). Then, for any $\varepsilon > 0$,

$$\begin{aligned}(1 - \beta\mu)\langle N \rangle &= \frac{\sum_{N=0}^{\infty} [Nc_N - (N-1)\, \mu c_{N-1}]\, \beta^N}{\sum_{N=0}^{\infty} c_N \beta^N} \\ &\leqslant (\Gamma + \varepsilon) + \frac{\text{polynomial in } \beta}{\sum_{N=0}^{\infty} c_N \beta^N} \end{aligned} \quad (4.13)$$

The second term is bounded as $\beta \uparrow \mu^{-1}$, so

$$\limsup_{\beta \uparrow \mu^{-1}} (1-\beta\mu)\langle N\rangle < \infty \tag{4.14}$$

from which (4.12) follows. ∎

Proof of Proposition 4.1. Let $U: l^2(\mathbf{Z}_+, \bar{\pi}) \to l^2(\mathbf{Z}_+$, counting measure) be the unitary mapping defined by

$$(Uf)(N) = (\beta^N c_N)^{1/2} f(N) \tag{4.15}$$

We then compute the matrix elements of $U\bar{P}U^{-1}$:

$$(U\bar{P}U^{-1})(N, N-1) = \frac{\beta^{1/2} e_N}{M c_N^{1/2} c_{N-1}^{1/2}} \tag{4.16a}$$

$$(U\bar{P}U^{-1})(N, N+1) = \frac{\beta^{1/2} e_{N+1}}{M c_N^{1/2} c_{N+1}^{1/2}} \tag{4.16b}$$

$$(U\bar{P}U^{-1})(N, N) = 1 - \frac{e_N}{M c_N} - \frac{\beta e_{N+1}}{M c_N} \tag{4.16c}$$

It follows that

$$U\bar{P}U^{-1} = \frac{a}{M}(\beta\mu)^{1/2} \Delta + \left\{1 - \frac{a}{M}[1-(\beta\mu)^{1/2}]^2\right\} I$$
$$+ \frac{a}{M}[1-(\beta\mu)^{1/2}]\,\delta_0 + Q \tag{4.17}$$

where Δ is the discrete Laplace operator on $\mathbf{Z}_+$,

$$\Delta(N, N-1) = \begin{cases} 1, & N \geqslant 1 \\ 0, & N = 0 \end{cases} \tag{4.18a}$$

$$\Delta(N, N+1) = 1 \tag{4.18b}$$

$$\Delta(N, N) = \begin{cases} -2, & N \geqslant 1 \\ -1, & N = 0 \end{cases} \tag{4.18c}$$

I is the identity matrix, $\delta_0 = \mathrm{diag}(1, 0, 0, \ldots)$, and

$$Q(N, N-1) = \left[\frac{\beta^{1/2} e_N}{M c_N^{1/2} c_{N-1}^{1/2}} - \frac{a}{M}(\beta\mu)^{1/2}\right](1-\delta_{N0}) \tag{4.19a}$$

$$Q(N, N+1) = \frac{\beta^{1/2} e_{N+1}}{M c_N^{1/2} c_{N+1}^{1/2}} - \frac{a}{M}(\beta\mu)^{1/2} \tag{4.19b}$$

$$Q(N, N) = \frac{a(1+\beta\mu)}{M} - \frac{e_N}{M c_N} - \frac{\beta e_{N+1}}{M c_N} - \frac{a}{M}\delta_{N0} \tag{4.19c}$$

The Laplacian is easily diagonalized by the functions $\cos k(n+1/2)$, $k \in [0, \pi]$; its spectrum is the interval $[-4, 0]$. The multiplication operator δ_0 is rank one, hence compact. The operator Q is a sum of three terms, each of which is a bounded operator (the identity, left shift, or right shift) times a function which vanishes at infinity; so Q is compact. It follows that the essential spectrum of $U\bar{P}U^{-1}$, and hence of $\bar{P}$, is the interval

$$\left[\frac{-4a}{M}(\beta\mu)^{1/2} + 1 - \frac{a}{M}[1-(\beta\mu)^{1/2}],\quad 1 - \frac{a}{M}[1-(\beta\mu)^{1/2}]\right] \quad \blacksquare$$

We now give two examples which show that τ'_{exp}(aggregated) can in some cases be a very poor lower bound for τ'_{exp}.

Example 4.1 (Lawler and Sokal[7]). Let $c_0 \equiv 1, c_1, c_2,\ldots$ be an *arbitrary* sequence of positive integers satisfying $\lim_{N\to\infty} c_N = +\infty$ and $\sup_{N\geqslant 0}(c_{N+1}/c_N) < \infty$, and let M_c be any integer $\geqslant \sup_{N\geqslant 0}(c_{N+1}/c_N)$. Then there exists a countable rooted tree $T = (V, E, \mathbf{0})$ such that:

(a) $\#(\{x\colon |x| = N\}) = c_N$

(b) $\sup_x c(x) \leqslant M_c$

(c) $\sup \operatorname{spec}(P \upharpoonright \mathbf{1}^{\perp}) \geqslant \begin{cases} 1-(1-\beta M_c)/M & \text{if } 0 \leqslant \beta < M_c^{-1} \\ 1 & \text{if } M_c^{-1} \leqslant \beta < \mu^{-1} \end{cases}$

Proof. We construct a "maximally unbalanced" tree having the given $\{c_N\}$: the root has c_1 children, which are labeled "eldest," "second-eldest," etc.; these children procreate, beginning with the eldest, each one having the maximum allowable number of children (M_c) until c_2 children have been generated; and so on. [In other words, of the c_N vertices at level N, the $\lfloor c_{N+1}/M_c \rfloor$ eldest of these have M_c children each, the one next-eldest has $c_{N+1} - M_c\lfloor c_{N+1}/M_c \rfloor$ children, and the rest have no children. Moreover, if $|x| = |y|$ and x is "elder" to y, then all the children of x are elder to all the children of y.] Now let x_N^* be the eldest vertex of level N; then the tree of descendants of x_N^* contains a complete M_c-ary rooted tree of $K_N + 1$ generations, where K_N is the largest integer such that $c_{N+k} \geqslant M_c^k$ for all $0 \leqslant k \leqslant K_N$. Therefore, the partial generating function

$$Z_{x_N^*}(\beta) \equiv \sum_{y \in V_{x_N^*}} \beta^{|y|-N} \tag{4.20}$$

satisfies

$$Z_{x_N^*}(\beta) \geqslant \sum_{k=0}^{K_N} (M_c\beta)^k \tag{4.21}$$

Moreover,

$$K_N \geqslant (\inf_{k \geqslant 0} \log c_{N+k})/\log M_c \to +\infty \qquad \text{as} \quad N \to \infty$$

Now consider the trial function

$$f_N = \chi_{V_{x_N^*}} - \text{const}$$

where the constant is chosen to make f orthogonal to $\mathbf{1}$. Then a straightforward calculation shows that

$$\frac{(f, (I-P)f)}{(f, f)} = \frac{1}{MZ_{x_N^*}(\beta)} \frac{Z(\beta)}{Z(\beta) - \beta^N Z_{x_N^*}(\beta)} \xrightarrow[N \to \infty]{} \leqslant \left[M \sum_{k=0}^{\infty} (M_c \beta)^k \right]^{-1} \tag{4.22}$$

Claim (c) follows. ∎

Thus, for $\beta \geqslant M_c^{-1}$ it is impossible to prove the existence of an L^2 spectral gap (much less lower bounds on it) given only the $\{c_N\}_{N=0}^{\infty}$ and M_c; it is necessary to have more detailed information about the structure of the graph G.

Example 4.2. Let G_N be the complete q-ary rooted tree of $N+1$ generations (i.e., levels 0 through N). Let P be its standard transition matrix with parameters (β, M). Fix a vertex x_1 of level 1, and define the trial function f by

$$f(x) = \begin{cases} 0 & \text{if} \quad x = \mathbf{0} \\ 1 & \text{if} \quad x \text{ is a descendant of } x_1 \\ -(q-1)^{-1} & \text{otherwise} \end{cases} \tag{4.23}$$

Then f is orthogonal to $\mathbf{1}$, and

$$(f, f) = \frac{Z(\beta) - 1}{Z(\beta)} \frac{1}{q-1} \tag{4.24}$$

$$(f, (I-P)f) = \frac{\beta}{MZ(\beta)} \frac{q}{q-1} \tag{4.25}$$

where

$$Z(\beta) = \frac{(\beta q)^{N+1} - 1}{\beta q - 1} \tag{4.26}$$

Therefore,

$$\sup \operatorname{spec}(P \restriction \mathbf{1}^{\perp}) \geqslant \frac{(f, Pf)}{(f, f)} = 1 - \frac{1}{M} \left(\frac{\beta q - 1}{(\beta q)^N - 1} \right) \tag{4.27}$$

In particular, if $\beta q > 1$, then the spectral gap is exponentially small when N is large.

On the other hand, the aggregated chain for G_N has the associated quadratic form

$$(F, (I - \bar{P}) F) = \frac{\beta}{MZ(\beta)} \sum_{n=0}^{N-1} (\beta q)^n \, [F(n+1) - F(n)]^2 \tag{4.28}$$

which for $\beta q > 1$ is a random walk on $\{0, 1, \ldots, N\}$ with constant *outward* drift. But on a finite interval, outward and inward drift are equivalent by symmetry; for N large, the spectral gap is approximately equal to (β/M) $[1 - (\beta q)^{-1/2}]^2$, by a calculation similar to that performed in the proof of Proposition 4.1. In particular, the spectral gap does *not* go to zero as $N \to \infty$.

The physics underlying this example is the following: If $\beta q > 1$, there is an outward drift, and the equilibrium measure π is strongly concentrated on the leaves $\{x: |x| = N\}$. But it is difficult for the random walker to move from one leaf to another, because to do so it must pass through the highly improbable states with $|x|$ small (fighting against the drift). On the other hand, this slow equilibration is not reflected in the aggregated chain, which measures only the equilibration *between* sets of fixed $|x|$.

Example 4.3. Let $G = (V, E, \mathbf{0})$ be an arbitrary rooted graph, and let P be its standard transition matrix with parameters (β, M). Let $G_N = (V_N, E_N, \mathbf{0}')$ be the graph defined in Example 4.2. Now form the new graph $\tilde{G} = (V \cup V_N, E \cup E_N \cup \{(\mathbf{0}, \mathbf{0}')\}, \mathbf{0})$ by hooking on the root of G_N as a child of the root of G. Let $\tilde{P}$ be the corresponding transition matrix with parameters (β, M); this is well defined (i.e., has nonnegative matrix elements) provided that $M \geqslant \max[M_0(G) + \beta, 1 + \beta q]$. Now $\tilde{P}$ and P clearly have the same essential spectrum. But the trial function constructed in Example 4.2 (extended to all of $\tilde{G}$ by setting $f = 0$ on G) shows that the spectral gap of $\tilde{P}$ can be made arbitrarily small, if $\beta q > 1$, by taking N large. This class of examples shows that the spectral gap can in general be much smaller than the essential spectral gap.

5. UPPER BOUNDS ON THE AUTOCORRELATION TIME, BY LIAPUNOV FUNCTIONS

Finally, we provide an upper bound on the modified autocorrelation time and a proof of geometric ergodicity, in the case that the drift is uniformly inward. Again the setting is that of standard random walk with parameters (β, M) on a rooted graph $G = (V, E, \mathbf{0})$ satisfying (3.1)–(3.2).

Proposition 5.1. Assume there exists an $\varepsilon > 0$ such that $p(x) - \beta c(x) \geqslant \varepsilon$ for all $x \neq \mathbf{0}$. Then $\sup \operatorname{spec}(P \upharpoonright \mathbf{1}^{\perp}) < 1$, and in fact

$$\sup \operatorname{spec}(P \upharpoonright \mathbf{1}^{\perp}) \leqslant 1 - \frac{\varepsilon}{M} + \frac{2\beta M_c}{M}\left[\left(1 + \frac{\varepsilon}{\beta M_c}\right)^{1/2} - 1\right] \tag{5.1}$$

$$= 1 - \frac{\varepsilon^2}{4\beta M M_c} + O(\varepsilon^3)$$

If, in addition, $M > M_0 \equiv \sup_x [p(x) + \beta c(x)]$, then

$$\inf \operatorname{spec} P \geqslant 1 - \frac{2M_0}{M} > -1 \tag{5.2}$$

so the process is geometrically ergodic.

Remarks. 1. Clearly, $p(x)$ and $c(x)$ can take only finitely many values [cf. (3.1)–(3.2)], so the first condition of the proposition is equivalent to the superficially weaker condition $p(x) > \beta c(x)$ for all x.

2. Summing the hypothesis $p(x) - \beta c(x) \geqslant \varepsilon$ with weight $\beta^{|x|}$ over $x \neq \mathbf{0}$, we obtain

$$\sum_{x \neq \mathbf{0}} \beta^{|x|}[p(x) - \beta c(x)] \geqslant \varepsilon[Z(\beta) - 1] \tag{5.3}$$

Now the sum telescopes:

$$\begin{aligned}\sum_{x \neq \mathbf{0}} \beta^{|x|}[p(x) - \beta c(x)] &= \lim_{N \to \infty} \sum_{\substack{x \neq \mathbf{0} \\ |x| \leqslant N}} \beta^{|x|}[p(x) - \beta c(x)] \\ &= \lim_{N \to \infty} \left[\beta c(\mathbf{0}) - \beta^{N+1} \sum_{|x| = N} c(x)\right] \\ &\leqslant \beta c(\mathbf{0})\end{aligned} \tag{5.4}$$

Therefore, $Z(\beta) \leqslant 1 + (\beta/\varepsilon)\, c(\mathbf{0}) < \infty$, so the Markov chain is positive-recurrent.

Proof of Proposition 5.1. Let

$$\psi_\alpha(x) = \alpha^{|x|} \tag{5.5}$$

with $\alpha \geqslant 1$. Then, for $x \neq \mathbf{0}$,

$$\begin{aligned}(P\psi_\alpha)(x) &= \frac{1}{M}\left[p(x)(\alpha^{-1}-1)+\beta c(x)(\alpha-1)+M\right]\psi_\alpha(x)\\ &\leqslant \frac{1}{M}\left[M+\beta c(x)(\alpha+\alpha^{-1}-2)-\varepsilon(1-\alpha^{-1})\right]\psi_\alpha(x)\\ &\leqslant \frac{1}{M}\left[M+\beta M_c(\alpha+\alpha^{-1}-2)-\varepsilon(1-\alpha^{-1})\right]\psi_\alpha(x)\\ &\equiv \frac{1}{\bar{r}(\alpha)}\psi_\alpha(x)\end{aligned} \tag{5.6}$$

where $\bar{r}(\alpha)>1$ provided that α is sufficiently close to 1. It follows that if $\{X(n)\}$ is the Markov process associated with the transition matrix P, then $\{r^n\psi_\alpha(X(n))\}$ is a supermartingale whenever $1 \leqslant r \leqslant \bar{r}(\alpha)$; it will serve as a kind of Liapunov function in the remainder of the proof.

Consider the process starting at $X(0)=x_0$. Let τ be the first time that the process hits the root, and let τ_N be the first time that it hits the root or $\{x\colon |x| \geqslant N\}$. Then, by the optional stopping theorem for supermartingales,

$$\begin{aligned}\alpha^{|x_0|} = \psi_\alpha(x_0) &\geqslant E_{x_0}(r^{\tau_N}\psi_\alpha(X(\tau_N)))\\ &= E_{x_0}(r^{\tau_N}\chi(X(\tau_N)=\mathbf{0})) + \alpha^N E_{x_0}(r^{\tau_N}\chi(|X(\tau_N)|=N))\\ &\geqslant E_{x_0}(r^{\tau_N})\\ &\underset{N\to\infty}{\to} E_{x_0}(r^{\tau})\end{aligned} \tag{5.7}$$

whenever $\alpha \geqslant 1$ and $1 \leqslant r \leqslant \bar{r}(\alpha)$; the last step used the monotone convergence theorem. This implies that if P_D is the matrix defined in Lemma 3.11, then by the same argument used there,

$$\|P_D\| \leqslant \inf_\alpha \bar{r}(\alpha)^{-1} = 1 - \frac{\varepsilon}{M} + \frac{2\beta M_c}{M}\left[\left(1+\frac{\varepsilon}{\beta M_c}\right)^{1/2} - 1\right] < 1 \tag{5.8}$$

The bound (5.1) follows immediately by Lemma 3.12.

The reverse bound (5.2) was already proven in the paragraph surrounding (3.32). ∎

Translating (5.1) into the language of autocorrelation times, we obtain

$$\tau'_{\exp} \leqslant \frac{4\beta M M_c}{\varepsilon^2} + O\left(\frac{1}{\varepsilon}\right) \tag{5.9}$$

Remarks. 1. Essentially the same Liapunov argument (but without mentioning the word "supermartingale") is given by Nummelin (ref. 20, Proposition 5.21).

2. For the Cayley rooted tree of order q, Propositions 4.1 and 5.1 give *exactly* the same bound; hence we have the *equalities*

$$\begin{aligned} \sup \operatorname{spec}(P \restriction \mathbf{1}^{\perp}) &= \sup \text{essential spec}(P) \\ &= \sup \text{essential spec}(\bar{P}) \\ &= 1 - \frac{1}{M} [1 - (\beta q)^{1/2}]^2 \end{aligned} \tag{5.10}$$

Thus, for the Cayley rooted tree, the aggregated chain $\bar{P}$ does reflect the slowest modes of the original chain P (in contrast to Examples 4.1 and 4.2).

ACKNOWLEDGMENTS

We would like to thank Jean Bricmont and Joel Lebowitz for suggesting that we rewrite the Introduction to make it more comprehensible (they are not to blame if we have failed).

This work was supported in part by NSF grants PHY-8413569 and DMS-8705599.

REFERENCES

1. K. Binder, ed., *Monte Carlo Methods in Statistical Physics* (Springer, Berlin, 1979; 2nd ed., 1986).
2. K. Binder, ed., *Applications of the Monte Carlo Method in Statistical Physics* (Springer, Berlin, 1984; 2nd ed., 1987).
3. A. Berretti and A. D. Sokal, *J. Stat. Phys.* **40**:483 (1985).
4. J. Goodman and A. D. Sokal, *Phys. Rev. Lett.* **56**:1015 (1986).
5. S. Caracciolo and A. D. Sokal, *J. Phys. A* **19**:L797 (1986).
6. N. Madras and A. D. Sokal, *J. Stat. Phys.* **50**:109 (1988).
7. G. F. Lawler and A. D. Sokal, *Trans. Am. Math. Soc.* **309**:557 (1988).
8. A. D. Sokal and L. E. Thomas, *J. Stat. Phys.* **51**:907 (1988).
9. A. D. Sokal and L. E. Thomas, Lower bounds on the autocorrelation time of a reversible Markov chain, with applications to statistical mechanics, in preparation.
10. A. D. Sokal, Geometric ergodicity and L^p spectra for Markov chains, in preparation.
11. A. D. Sokal, in *VIIIth International Congress on Mathematical Physics*, R. Sénéor and M. Mebkhout, eds. (World Scientific, Singapore, 1987).
12. R. G. Edwards and A. D. Sokal, *Phys. Rev. D* **38**:2009 (1988).
13. A. J. Guttmann, T. R. Osborn, and A. D. Sokal, *J. Phys. A* **19**:2591 (1986).
14. Ph. de Forcrand, F. Koukiou, and D. Petritis, *J. Stat. Phys.* **45**:459 (1986).
15. Ph. de Forcrand, F. Koukiou, and D. Petritis, *J. Stat. Phys.* **49**:223 (1987).

16. J. M. Pureza, C. Aragão de Carvalho, and S. L. A. de Quieroz, *J. Phys. A* **20**:4409 (1987).
17. A. J. Guttmann and T. R. Osborn, *J. Phys. A* **21**:513 (1988).
18. G. F. Lawler and A. D. Sokal, in preparation; B. Duplantier, private communication.
19. K. L. Chung, *Markov Chains with Stationary Transition Probabilities*, 2nd ed. (Springer, New York, 1967).
20. E. Nummelin, *General Irreducible Markov Chains and Non-Negative Operators* (Cambridge University Press, Cambridge, 1984).
21. Z. Šidák, *Czech. Math. J.* **14**:438 (1964).
22. W. Rudin, *Functional Analysis* (McGraw-Hill, New York, 1973).
23. E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups* (American Mathematical Society, Providence, Rhode Island, 1957).
24. P. R. Halmos, *A Hilbert Space Problem Book*, 2nd ed. (Springer, New York, 1982).
25. R. Holley, *Ann. Prob.* **13**:72 (1985).
26. J. G. Kemeny and J. L. Snell, *Finite Markov Chains* (Springer, New York, 1976).
27. M. Iosifescu, *Finite Markov Processes and Their Applications* (Wiley, Chichester, 1980).
28. C. Kipnis and S. R. S. Varadhan, *Commun. Math. Phys.* **104**:1 (1986).
29. J. W. Essam and M. E. Fisher, *Rev. Mod. Phys.* **42**:271 (1970).
30. W. G. Sullivan, *Z. Wahrsch. Verw. Gebiete* **67**:387 (1984).
31. L. R. Shenton, *Proc. Camb. Phil. Soc.* **51**:442 (1955).
32. J. G. Mauldon, *Am. Math. Monthly* **94**:423 (1987).
33. J. Keilson, *Markov Chain Models—Rarity and Exponentiality* (Springer, New York, 1979).
34. D. C. Brydges, J. Fröhlich, and A. D. Sokal, *Commun. Math. Phys.* **91**:141 (1983).
35. G. Slade, *Commun. Math. Phys.* **110**:661 (1987). See also T. Hara and G. Slade, to be published.
36. R. Cogburn, *Ann. Prob.* **3**:191 (1975).
37. L. E. Thomas and Z. Yin, *J. Math. Phys.* **27**:2475 (1986).
38. J. M. Hammersley and K. W. Morton, *J. Roy. Stat. Soc. B* **16**:23 (1954); J. M. Hammersley, *Proc. Camb. Phil. Soc.* **53**:642 (1957).