ALGEBRAIC SYSTEMS OF GENERATING FUNCTIONS AND RETURN PROBABILITIES FOR RANDOM WALKS

STEVEN P. LALLEY

1. Introduction

The purpose of these lectures is to set out a method for analyzing the principal singularities of certain systems of generating functions $F_i(z)$ that are interrelated by functional equations of the form

$$F_i(z) = zQ_i(F_1(z), F_2(z), \dots). \tag{1.1}$$

Such systems occur in a variety of combinatorial and probabilistic contexts, several of which are discussed below. Our primary interest in them stems from their occurrence in random walk problems, especially for random walks on homogeneous trees and treelike structures. We shall see that the asymptotic behavior of *return probabilities* for such random walks is governed by the leading singularity of systems of the form (1.1).

1.1. **The Lagrange Inversion Formula.** The scalar form of the system (1.1), in which there is a single generating function $F(z)$ that satisfies a functional equation

$$F(z) = zQ(F(z)), \tag{1.2}$$

has a history that dates to Lagrange (or perhaps even earlier). Lagrange discovered what is now known as the *Lagrange Inversion Formula*, which gives an explicit formula for the coefficients of the power series $F(z) = \sum_{n\geq 0} a_n z^n$ representing $F(z)$ in terms of the link function $Q(w)$ and its powers.

Theorem 1.1. *(Lagrange Inversion Formula) Let $Q(w) = \sum_{n\geq 0} b_n w^n$ be a formal power series in w with constant coefficient $b_0 \neq 0$. There is a unique formal power series $F(z) = \sum_{n\geq 0} a_n z^n$ that satisfies the functional equation (1.2). Its coefficients are given by*

$$a_n = n^{-1}[w^{n-1}]Q(w)^n. \tag{1.3}$$

Here $[w^{n-1}]Q(w)^n$ denotes the coefficient of w^{n-1} in the power series expansion of the function Q^n.

Various proofs of Lagrange's theorem are known: see [24], [10], and [23] for some of the standard ones. A more intricate but also more interesting proof, based on a form of Spitzer's Combinatorial Lemma ([6], Ch. XII, Section 6), is given in [20]: this proof shows that the Lagrange Formula is intimately connected with the combinatorics of lattice paths. The essence of this proof is distilled in section 1.2 below.

1991 *Mathematics Subject Classification.* Primary 60F05 Secondary 60F99.

Key words and phrases. regular language, algebraic function, Perron-Frobenius theory, Lyapunov-Schmidt method.

Supported by NSF grant DMS-0071970.

The validity of the Lagrange formula does not require convergence of the formal power series $Q(w)$. However, if Q has a nonzero radius of convergence, then so will $F(z)$. (EXERCISE!) In this event, the Cauchy Integral Formula may be used in conjunction with (1.2) to relate the coefficients a_n of $F(z)$ with the values of $Q(w)$ on any contour γ surrounding the origin such that $Q(w)/w$ is analytic everywhere inside γ except at $w = 0$:

$$a_n = \frac{1}{2\pi i n} \oint_\gamma \frac{Q(w)^n}{w^n} \, dw. \tag{1.4}$$

This formula is useful for asymptotic calculations. Assume that the coefficients b_n of the power series $Q(w)$ are nonnegative (this will be the case in most combinatorial and probabilistic applications). If there exists $\rho > 0$ less than the radius of convergence of $Q(w)$ such that $\rho Q'(\rho) = Q(\rho)$, then $w = \rho$ is a *saddle point* for the integrand in (1.4), and so the saddle point method of asymptotic evaluation ([5], section 2.2) may be applied. In particular, unless $Q(w)$ is a linear function, the second derivative at the saddlepoint $w = \rho$ will be positive, and so

$$a_n \sim \frac{1}{\sigma R^n n^{3/2}} \tag{1.5}$$

for a suitable constant $\sigma > 0$, where

$$R = \rho / Q(\rho), \tag{1.6}$$

provided that $|Q(w)| < Q(\rho)$ for all $w \neq \rho$ such that $|w| = \rho$. (This last proviso will be satisfied unless the set of indices n such that $q_n \neq 0$ is contained in a coset of a proper subgroup of the integers.) In particular, it follows that R is the radius of convergence of the power series $F(z)$.

1.2. **Hitting Times for Left-Continuous 1D Random Walks.** Let $\{q_x\}_{x=0,1,\dots}$ be a probability distribution on the nonnegative integers, and let $\xi_1, \xi_2, \dots$ be independent, identically distributed random variables with distribution $\{q_x\}$. Define

$$S_n = \sum_{j=1}^{n} (\xi_j - 1) \; ; \tag{1.7}$$

the sequence S_n is a *left-continuous* random walk on the integers. Define τ to be the first time n that $S_n = -1$ (or ∞ if there is no such n). One may easily check (EXERCISE!) that the generating function $F(z) := Ez^\tau$ satisfies the functional equation (1.2), with

$$Q(w) = \sum_{m=0}^{\infty} q_m w^m.$$

The Lagrange Inversion Formula (1.3) translates as

$$P\{\tau = n\} = n^{-1} P\{S_n = -1\}. \tag{1.8}$$

EXERCISE: Give a direct proof of (1.8) using Spitzer's Combinatorial Lemma. A special case of Spitzer's Lemma may be stated as follows:

Lemma 1.2. *Let $x_1, x_2, \dots, x_n$ be a sequence of integers ≥ -1 with sum -1. Then there is a unique cyclic permutation π of the integers $1, 2, \dots, n$ such that*

$$\sum_{j=1}^{k} x_{\pi(j)} \geq 0 \;\; \forall \, k = 1, 2, \dots, n-1. \tag{1.9}$$

The proof of Lemma 1.2 is another EXERCISE. The trick is to guess where the cycle must begin.

Now consider the asymptotic behavior of the probabilities $P\{\tau = n\}$. There are two trivial cases: (A) If $q_0 + q_1 = 1$, then the random walk S_n makes no jumps to the right: it stays at the starting point 0 for a geometrically-distributed number of steps, then jumps to -1, and so τ has a geometric-plus-one distribution. (B) If $q_0 = 0$ then the jumps of the random walk are all nonnegative, and so $\tau = \infty$. When $q_0 > 0$ and $q_0+q_1 < 1$, the generating function $Q(w)$ is strictly convex. There are several possibilities, depending on whether the mean $Q'(1)$ of the distribution $\{q_x\}$ is greater, or less, or equal to 1: (C) If $Q'(1) > 1$ then there exists $0 < \rho < 1$ such that $\rho Q'(\rho) = Q(\rho)$; in this case $R = \rho/Q(\rho) > 1$. (D) If $Q'(1) < 1$, and if the radius of convergence of $Q(w)$ is infinite, then there exists $1 < \rho < \infty$ such that $\rho Q'(\rho) = Q(\rho)$, and once again $R = \rho/Q(\rho) > 1$. (E) If $Q'(1) = 1$ then the saddlepoint is at $\rho = 1$, and $R = 1$. In cases (C), (D), and (E), if the distribution $\{q_x\}$ is *nonlattice* (that is, not supported by a coset of a proper subgroup of $\mathbb{Z}$) then $|Q(w)| < Q(\rho)$ for all $w \neq \rho$ on the contour $|w| = \rho$, and thus, by relation (1.5),

$$P\{\tau = n\} \sim \frac{1}{\sigma R^n n^{3/2}}. \tag{1.10}$$

(Note: In case (E), if the radius of convergence of $Q(w)$ is 1 then the Laplace expansion of the integral (1.4) requires in addition that $Q''(1) < \infty$, that is, that the distribution $\{q_x\}$ has a finite second moment.) Finally, if $Q'(1) < 1$ but the radius of convergence of $Q(w)$ is finite, then there may be no saddlepoint, in which case the decay of $P\{\tau = n\}$ as $n \to \infty$ may be considerably more complicated.

1.3. **Size of a Galton-Watson Tree.** Let N be the *size* of a Galton-Watson tree for which the offspring distribution is $\{q_x\}_{x=0,1,\dots}$, that is, N is the total number of individuals (including the progenitor) ever born. This will be finite if and only if the mean offspring number $\mu = \sum xq_x$ is no greater than 1. The random variable N may be decomposed as a sum $1 + N_1 + N_2 + \cdots + N_Z$, where Z is the number of offspring of the progenitor and $N_1, N_2, \dots$ are the sizes of the Galton-Watson trees engendered by the members of the first generation. These are conditionally independent, given Z, each with the same distribution as N. Therefore, the probability generating function $F(z) = Ez^N$ satisfies the functional equation (1.2), where $Q(w)$ is the probability generating function of the offspring distribution $\{q_x\}$. Because solutions to the Lagrange equation (1.2) are unique, it follows that the generating function $F(z)$ coincides with that of the random variable τ discussed in section 1.2 above; hence, the random variables N and τ have the same distribution. It makes an interesting EXERCISE to prove this directly, without using generating functions.

In the critical case, when $\mu = 1$, the solution ρ of the saddlepoint equation $\rho Q'(\rho) = Q(\rho)$ is at $\rho = 1$ (since $Q(1) = 1$) and $R = 1$. Consequently, by (1.10), if the offspring distribution is nonlattice then for a suitable constant $C > 0$,

$$P\{N = n\} \sim C/n^{3/2}. \tag{1.11}$$

1.4. **Multitype Galton-Watson Trees.** Next, consider a *multitype* Galton-Watson tree with finitely many types $i = 1, 2, \dots, I$. For each ordered type i, there is a probability distribution $q_i = \{q_i(m_1, m_2, \dots, m_I)\}$ on the set $\mathbb{Z}_+^I$ of $I-$vectors of nonnegative integers that governs the numbers of offspring of different types produced by a type-i individual in one generation. Denote by N the size of the tree, that is, the number of individuals of *all* types ever born. Clearly, the distribution of N will depend on the type of the progenitor.

Denote by $F_i(z)$ the probability generating function of N when the progenitor has type i; then by a one-step analysis entirely analogous to that used for unitype Galton-Watson trees above, one finds that

$$F_i(z) = zQ_i(F_1(z), F_2(z), \dots, F_I(z)) \tag{1.12}$$

where $Q_i(w_1, w_2, \dots, w_I)$ is the (multivariate) generating function of the distribution q_i. This system is evidently of type (1.1).

The first serious attempt to analyze *systems* of functional equations of type (1.1) was made by I. J. Good in the 1950s (see [9]). The primary impetus for Good's study seems to have been the problem we have just discussed – the size of the multitype Galton-Watson tree. Good discovered, among other things, a remarkable generalization of Lagrange's Inversion Formula that has itself engendered a substantial literature. Unfortunately, Good's formula involves determinants, and so, because of the cancellations in the determinants, asymptotic analysis using the saddlepoint method cannot be carried out in the same manner as in the scalar case. Thus, for the purpose of asymptotic analysis, Good's formula proves to be a wrong turn. In sections 3 and 5 below, we shall present a method for asymptotic analysis of the coefficients of power series $F_i(z)$ that are interrelated by functional equations of type (1.1). Before coming to this, however, we shall discuss the occurrence of such systems of functional equations in random walk problems.

2. Random walks on trees and free products

Let $\mathcal{T}^d$ be the infinite homogeneous tree of degree $d \geq 3$, with a vertex $\emptyset$ designated as the *root*. Assume that the edges of the tree are assigned *colors* from the set $[d] = \{1, 2, \dots, d\}$ in such a way that every vertex of the tree is incident to exactly one edge of each color. Given a probability distribution $\{p_i\}_{i\in[d]}$ on the set of colors, a *nearest-neighbor random walk* on $\mathcal{T}^d$ may be performed as follows: At each step, choose a color i at random from the step distribution $\{p_i\}$, independently of all previous choices, and move across the unique edge incident to the current vertex with color i.

The reader will recognize that the random walk just described may also be described as a nearest-neighbor (right) random walk on the free product $\Gamma = \mathbb{Z}_2^{*d}$ of d copies of the two-element group $\mathbb{Z}_2$. (See [3] or [25] for a formal definition of the free product of a collection of groups.) The group elements are finite reduced words whose letters are elements of the set $[d]$ of colors, that is, words in which no color i appears twice consecutively. Note that there is a one-to-one correspondence between reduced words and vertices of $\mathcal{T}^d$: the letters of the word representing a given vertex indicate the unique path from the root $\emptyset$ to the vertex. Multiplication in the group consists of concatenation followed by reduction (successive elimination of adjacent matching pairs of letters). The location X_n of the random walker after n steps is given by

$$X_n = \xi_1, \xi_2, \dots, \xi_n, \tag{2.1}$$

where $\xi_1, \xi_2, \dots$ are i.i.d. with distribution $\{p_i\}$. One may also use equation (2.1) to define *non-nearest-neighbor* random walks: for such walks, the common distribution of the steps ξ_n is no longer restricted to words of length 1. We shall say that a random walk is *finite-range* if the step distribution $\{p_x\}$ is supported by a finite subset of Γ.

Our primary interest is in the asymptotic behavior of the transition probabilities of the random walk X_n, and in particular on the behavior of the return probabilities $P\{X_n = \emptyset\}$. The method of analysis that we shall develop applies generally to all finite-range random walks on $\mathcal{T}^d$, and even to a large class of infinite-range random walks. As we shall see, the method extends also to *countable* free products. However, to present the method in its simplest form, we shall for now restrict attention to the nearest-neighbor case. To avoid complications stemming from periodicity, we shall also assume that the step distribution attaches positive probability $p_\emptyset$ to the empty word $\emptyset$.

Remark. In the nearest-neighbor ccase, there is another approach to the analysis of transition probabilities: see [8] and [2]. This approach does not generalize to the non-nearest-neighbor case, unfortunately.

2.1. **The Green's function.** The Green's function of a random walk on a discrete group Γ is defined to be the generating function of the return probabilities:

$$G(z) = \sum_{n=0}^{\infty} P\{X_n = \emptyset\} z^n. \tag{2.2}$$

In this section, we shall establish some fundamental properties of the Green's function and introduce a denumerable family of auxiliary generating functions to which the Green's function is related by a system of algebraic functional equations.

For any element $x \in \Gamma$, define the generating functions

$$G_x(z) = \sum_{n=0}^{\infty} P\{X_n = x\} z^n \qquad \text{and} \tag{2.3}$$

$$F_x(z) = \sum_{n=1}^{\infty} P\{\tau_x = n\} z^n, \tag{2.4}$$

where

$$\tau_x = \min\{n \geq 0 \,:\, X_n = x\}. \tag{2.5}$$

Note that $G = G_\emptyset$. Also, if the random walk is irreducible (that is, for any two states x, y there is at least one positive probability path leading from x to y) then the sums of the series (2.3) and (2.4) are strictly positive (possibly $+\infty$) for all positive arguments z. Because the coefficients of the power series defining the functions G_x and F_x are probabilities, each has radius of convergence at least one; moreover, the first-passage generating functions $F_x(z)$ are bounded in modulus by 1 for all $|z| \leq 1$. In fact, all of the the functions $G_x(z)$ have common radius of convergence R, as we shall see below.

The functions $G_x(z)$ can be expressed in terms of the first-passage generating functions $F_x(z)$ by an application of the Markov property. Conditioning on (i) the first step of the random walk, and then (ii) the value of the first-passage time τ_x, one obtains the relations

$$G(z) = 1 + p_\emptyset z G(z) + \sum_{x \neq \emptyset} p_x z F_{x^{-1}}(z) G(z) \qquad \text{and} \tag{2.6}$$

$$G_x(z) = F_x(z) G(z) \;\; \forall\, x \neq \emptyset. \tag{2.7}$$

The first of these may be solved for $G(z)$:

$$G(z) = \left\{ 1 - p_\emptyset z - \sum_{x \neq \emptyset} p_x z F_{x^{-1}}(z) \right\}^{-1} \tag{2.8}$$

2.2. **The Lagrangian System.** The arguments used thus far have been ccompletely general, and do not depend on the particular structure of the group Γ. Assume now that Γ is a free product (finite or denumerable) of copies of $\mathbb{Z}_2$, and that X_n is a nearest-neighbor random walk on Γ. Under this assumption, the first-passage generating functions, which by equation (2.6) determine the Green's function, are themselves interrelated by a system of functional equations that derive from the Markov property. Because the random walk is nearest-neighbor, the only values of x that occur in the relation (2.8) are words with a single letter i; consequently, we shall consider only these.

For those random paths that ultimately visit the state i are three possibilities for the first step: X_1 could be $\emptyset$, it could be i, or it could be a one-letter word j for some $j \neq i$. In the last case, the path must first return to $\emptyset$ before visiting i, and so we may condition on the first time at which this happens. This leads to the equations

$$F_i(z) = z \left\{ p_i + p_\emptyset F_i(z) + \sum_{j \neq i} p_j F_j(z) F_i(z) \right\} \tag{2.9}$$

We shall refer to this system as the *Lagrangian system* of the random walk. Observe that it has the generic form (1.1). In addition, the equations may be iterated, by successive re-substitutions on the right sides. If the random walk is irreducible, as we shall assume henceforth, then for any two colors i, j, some iterate of the equation for $F_i(z)$ includes on the right side a term in which $F_j(z)$ occurs as a factor with a positive coefficient. Consequently, all of the functions $F_i(z)$ have the same radius of convergence R, and have the same type of singularity at $z = R$. It follows, by equations (2.7) and (2.8), that all of the functions $G_x(z)$ have a common radius of convergence ρ, and that $\rho \leq R$.

2.3. **Finiteness of G at the radius of convergence.** Because our primary interest is in the asymptotic behavior of the coefficients of the Green's function $G(z)$, it behooves us to look further into the relation between its radius ρ of convergence and the radius R of convergence of the first-passage generating functions $F_i(z)$. In principle, the relevant information is already embedded in the functional equations (2.8) and (2.9); however, it is easier for us to exploit a structural property of the group Γ to deduce some restrictions on the nature of the singularities. The structural property is this: Any irreducible random walk on the the free product Γ of three or more copies of $\mathbb{Z}_2$ must be transient. This follows from a stronger theorem of Kesten [15], which asserts that the *spectral radius* R of the transition kernel of an irreducible random walk must be greater than 1.

Proposition 2.1. *Let R be the common radius of convergence of the first-passage generating functions $F_i(z)$ and let ρ be the radius of convergence of the Green's function $G(z)$. Then $R = \rho$, and*

$$G(R) < \infty. \tag{2.10}$$

Proof. The proof uses the fact that any irreducible random walk on Γ must be transient. First note that for any two elements $x, y \in \Gamma$ there is a positive probability path leading from x to y. Let m be the length of such a path and $c_m(x, y) > 0$ its probability; then for every $n \geq 0$,

$$P\{X_{n+m} = y\} \geq c_m(x, y) P\{X_n = x\}.$$

Consequently, for every positive argument s of the Green's functions,

$$G_y(s) \geq c_m(x, y) s^m G_x(s). \tag{2.11}$$

Suppose now that $G(R) = \infty$; then by the "Harnack" inequalities (2.11), the ratios $G_x(s)/G(s)$ remain bounded, and bounded away from zero, as $s \to R-$. Hence, by a diagonal argument, there is a sequence $s_n \to R-$ such that

$$\varphi_x = \lim_{n\to\infty} \frac{G_x(s_n)}{G(s_n)}$$

exists. The Harnack inequalities guarantee that the function φ is everywhere positive. Furthermore, by equations (2.7), the function φ is $\Gamma-$invariant, that is, the ratios φ_{xy}/φ_x depend only on y. Most important, the function $x \mapsto \varphi_x$ is $R-$harmonic, that is, for every $x \in \Gamma$,

$$\varphi_x = R \sum_y \varphi_y P\{\xi_1 = y^{-1}x\}.$$

This follows from the hypothesis that $G(s) \to \infty$ as $s \to R-$, since by the Markov property the Green's function satisfies the relations

$$G_x(z) = \delta_{0,x} + z \sum_y G_y(z) P\{\xi_1 = y^{-1}x\}.$$

Since φ is $\Gamma-$invariant and $R-$harmonic, the Doob $h-$transform of the transition probability kernel of the random walk X_n by the ratios $R\varphi_x/\varphi_y$ is $\Gamma-$invariant. The $h-$transform is the transition kernel defined by

$$q(x, y) := RP\{\xi_1 = y^{-1}x\} \frac{\varphi_x}{\varphi_y} = q(\emptyset, x^{-1}y).$$

The iterates of this transition probability kernel satisfy

$$q^{(n)}(\emptyset, \emptyset) = R^n P\{X_n = \emptyset\}.$$

Since φ_x is $\Gamma-$invariant, the kernel $q(x, y)$ is the transition probability kernel of a right random walk on Γ. The hypothesis that $G(R) = \infty$ and the relation between $q(x, y)$ and the transition kernel of the original random walk now implies that the random walk with transition probability kernel q is recurrent. This contradicts the fact that all irreducible random walks on Γ are transient. □

Corollary 2.2. *For an irreducible random walk on* Γ*,*

$$p_\emptyset + \sum_{x \neq \emptyset} p_x F_{x^{-1}}(R) < 1/R. \tag{2.12}$$

2.4. **Behavior off the real axis.**

Proposition 2.3. *For any* aperiodic*, irreducible random walk on a nonamenable discrete group, the Green's function* $G(z)$ *is regular at every point on the circle of convergence* $|z| = R$ except $z = R$*.*

Proof. Exercise. See [1] for the solution. □

3. Lagrangian Systems of Functional Equations

3.1. **A theorem of Flajolet and Odlyzko.** We have already seen in the case of a scalar Lagrange equation (1.2) that the asymptotic behavior of the coefficients of the solution $F(z)$ is controlled by the behavior of F near its smallest positive singularity $z = R$. The same is true for Lagrangian systems (1.1). To extract the asymptotics from the behavior at the singularity, we shall call upon a Tauberian theorem of Flajolet and Odlyzko [7].

Theorem 3.1. *(Flajolet & Odlyzko) Let $G(z) = \sum_{n=0}^{\infty} a_n z^n$ be a power series with radius of convergence R. Suppose that G has an analytic continuation to the* Pac-Man *domain*

$$\Delta_{\rho,\phi} = \{z : |z| < \rho \quad \text{and} \quad |\arg(z - R)| > \phi\}, \tag{3.1}$$

where $\rho > R$ and $\phi < \pi/2$, and suppose that as $z \to R$ in $\Delta_{\rho,\phi}$,

$$G(R) - G(z) \sim K(R - z)^{\alpha} \tag{3.2}$$

for some $K \neq 0$ and $\alpha \notin \{0, 1, 2, \dots\}$. Then as $n \to \infty$,

$$a_n \sim \frac{K}{\Gamma(-\alpha) R^n n^{\alpha+1}}. \tag{3.3}$$

Thus, to determine the asymptotic behavior of the coefficients it suffices to determine the nature of the lead singularity.

3.2. **Lagrangian Systems.** Suppose now that $F(z)$ is a Banach-space valued analytic function that satisfies a functional equation of the form

$$F(z) = zQ(F(z)), \tag{3.4}$$

where Q is a holomorphic mapping of the Banach space to itself such that $Q(0) \neq 0$. (A mapping $Q : B \to B$ is said to be *holomorphic* if it is infinitely differentiable [in the sense of [28], section 4.5], and if for every analytic function $F : \mathbb{C}^m \to B$ the composition $Q(F)$ is analytic.) We shall refer to systems of the form (3.4) as *Lagrangian systems*. Much of the theory extends, with very little change, to the more general class of functional equations

$$F(z) = Q(z, F(z));$$

however, in the interest of simplicity we shall restrict attention to systems of the more special type (3.4). When the system of generating functions $F_i(z)$ satisfying the equations (1.1) is finite, with (say) I members, then the Banach space of interest is finite-dimensional, to wit, $\mathbb{C}^I$,with the usual norm — this is the case, for instance, for the multitype Galton-Watson tree discussed in section 1.4, and also for nearest-neighbor random walks on *finite* free products of $\mathbb{Z}_2$. More examples will be considered in section 4.2.1 below. For random walks on *infinite* free products, random walks on *finite* free products whose step distribution has infinte support, and multitype Galton-Watson trees with denumerably many types, it is necessary to consider infinite systems (1.1). In such cases the choice of an appropriate Banach space will not always be obvious. See section 6 below for an extended example.

3.3. **Analytic continuation of a Banach space-valued function.** Let B be a (complex) Banach space, and let $Q : B \to B$ be a holomorphic mapping. Under what conditions will the functional equation (3.4) have an analytic solution $F(z)$ in a neighborhood of the complex plane containing the origin $z = 0$? Clearly, the equation admits the solution $F(0) = 0$ at $z = 0$. By the Implicit Function Theorem ([28], ch. 4) for Banach-valued functions, this solution admits an analytic continuation to a neighborhood of $z = 0$, as the linearized system

$$dF = Q(F)\,dz + z\frac{\partial Q}{\partial F}\,dF \tag{3.5}$$

is solvable for dF in terms of dz when $z = 0$. (Here and throughout the notes $\partial Q/\partial F$ or $\partial Q/\partial w$ will denote the Jacobian operator of the mapping Q.) Furthermore, analytic continuation of the solution $F(z)$ is possible along any curve in the complex plane starting at $z = 0$ on which the linear operator

$$I - z\frac{\partial Q}{\partial F} \tag{3.6}$$

remains invertible. This will be the case as long as the spectral radius of the operator

$$z\mathcal{L}(z) := z\left(\frac{\partial Q}{\partial F}\right)_{F(z)} \tag{3.7}$$

remains less than one. Thus, singular points of the analytic function $F(z)$ can only occur at those points where the spectral radius of $z\mathcal{L}(z)$ attains or exceeds the value 1.

In general, the dependence of the spectrum of the operator $(\partial Q/\partial F)_{F(z)}$ on the parameter z may be quite complicated, even though the mapping Q and the function $F(z)$ are holomorphic. Moreover, in general the spectrum need not be purely discrete. These complications arise only in the infinite-dimensional case. Therefore, we shall first discuss in detail the finite-dimensional case.

3.4. **The finite-dimensional case.** We now restrict attention to the the finite-dimensional case, where the Banach space $B = \mathbb{C}^I$ for some integer $I < \infty$. Discussion of the infinite-dimensional case will be resumed in section 5 below. Assume that the components Q_i of the mapping $Q(w) = Q(w_1, w_2, \dots, w_I)$ are given by convergent power series with nonnegative coefficients. Note that this has been the case in all the examples encountered thus far; moreover, all of the component functions $F_i(z)$ have also been defined by power series with nonnegative coefficients. The following proposition shows that this is no accident.

Proposition 3.2. *If the components $Q_i(w)$ of the mapping Q are given by convergent power series with nonnegative coefficients then the system (3.4) has a unique analytic solution in a neighborhood of $z = 0, F(z) = 0$ whose components $F_i(z)$ are given by convergent power series with nonnegative coefficients.*

Proof. We have already seen that there is a unique analytic solution $F(z)$ in a neighborhood of the origin, by the Implicit Function Theorem, so what must be proved is the nonnegativity of the power series coefficients of (the components of) $F(z)$. For the purposes of this proof, $F \le G$ will mean that the power series coefficients of $F(z)$ are dominated by those of $G(z)$. Define a series of approximate solutions to (3.4) by setting $F_0(z) = 0$ and for each $n \ge 0$,

$$F_{n+1}(z) = zQ(F_n(z)). \tag{3.8}$$

Because the mapping $Q(w)$ is holomorphic, its Jacobian $\partial Q/\partial w$ is uniformly bounded in (sup) norm for $\|w\| \le C$, for any $C < \infty$; consequently, the mapping which sends a function $H(z)$ to the function $zQ(H(z))$ is contractive for small z and H. Therefore, the sequence of functions F_n defined by (3.8) is uniformly norm convergent for $|z| \le \varepsilon$, provided $\varepsilon > 0$ is sufficiently small, and the limit function $F(z)$ is the solution to (3.4).

Since the mapping Q is nonnegative, $F_1 \geq F_0 = 0$, and hence, by induction, $F_{n+1} \geq F_n$ for every $n \geq 0$. In particular, each $F_n(z)$ has nonnegative power series coefficients. Since the functions $F_n(z)$ converge uniformly in norm to $F(z)$, the Cauchy Integral Formula implies that $F(z)$ has nonnegative power series coefficients. □

Corollary 3.3. *Under the hypotheses of Proposition 3.2, the entries of the matrix-valued function $\mathcal{L}(z) = (\partial Q/\partial w)_{F(z)}$ are given by convergent power series with nonnegative coefficients. Therefore, for positive arguments s, the entries of $\mathcal{L}(s)$ are nonnegative, nondecreasing, and convex in s. Moreover, the spectral radius $\lambda(s)$ is nondecreasing and continuous in s for nonnegative arguments s. Finally,*

$$|\lambda(z)| \leq \lambda(|z|), \tag{3.9}$$

and so the function $F(z)$ has an analytic continuation to any disc $|z| \leq r$ such that $r\lambda(r) < 1$.

For positive arguments s, the matrix $\mathcal{L}(s)$ has nonnegative entries, and so its spectral radius coincides with its largest nonnegative eigenvalue. In general, this eigenvalue need not be simple, nor need it be the only eigenvalue on the circle of radius $\lambda(s)$ centered at 0. However, if $\mathcal{L}(s)$ is a *Perron-Frobenius* matrix (a nonnegative matrix M such that some power M^n has all entries positive), then by the Perron-Frobenius theorem ([22]) the eigenvalue $\lambda(s)$ must be simple, and there can be no other eigenvalue of modulus $\lambda(s)$.

Define the *dependency graph* of the system (3.4) to be the directed graph $\mathcal{G}$ with vertex set $[I] := \{1, 2, \dots, I\}$ such that for each pair (i, j) of vertices there is a directed edge from j to i if and only if the variable w_j occurs as a factor in a term of $Q_i(w)$ with positive coefficient. Let A be the incidence matrix of $\mathcal{G}$, that is, the $I \times I$ matrix whose (i, j)th entry is 1 if there is a directed edge from i to j and 0 otherwise. Say that the system (3.4) is *irreducible aperiodic* if the incidence matrix A of its dependency graph is Perron-Frobenius. Clearly, the (i, j)th entry of $\mathcal{L}(s)$ is positive for $s > 0$ if and only if the (i, j)th entry of A is positive, and so $\mathcal{L}(s)$ is a Perron-Frobenius matrix if and only if A is.

Proposition 3.4. *Assume that the components $Q_i(w)$ of the mapping Q are given by convergent power series with nonnegative coefficients, and assume further that the Lagrangian system (3.4) is irreducible and aperiodic. Then the matrix $\mathcal{L}(s)$ is Perron-Frobenius for all $s > 0$ such that the spectral radius $\lambda(s) \leq 1/s$. Furthermore, the lead eigenvalue $\lambda(z)$ and the corresponding right and left eigenvalues $h(z)$ and $w(z)$, normalized so that*

$$w(z)^T h(z) = w(z)^T \mathbf{1} = 1, \tag{3.10}$$

are analytic in a neighborhood of the segment $[0, R)$ and continuous on $[0, R]$, where

$$R = \min\{s : s\lambda(s) = 1\}. \tag{3.11}$$

Finally, if $Q_i(\mathbf{0}) > 0$ for some index i then $R < \infty$ and $F(R)$ is finite.

Proof. That $\mathcal{L}(s)$ is Perron-Frobenius for $0 < s \leq R$ folows from the discussion preceding the statement of the proposition. Because the mapping $z \mapsto \mathcal{L}(z)$ is analytic for $|z| < R$, and because the lead eigenvalue $\lambda(s)$ is simple for $0 < s \leq R$, by the Perron-Frobenius theorem, results of regular perturbation theory (see [21], section XII.2) imply that the mappings

$$\begin{aligned} &z \mapsto \lambda(z) \\ &z \mapsto h(z), \qquad \text{and} \\ &z \mapsto w(z) \end{aligned}$$

have analytic continuations to a neighborhood of $(0, R)$, and are continuous at $z = R$. Now suppose that $Q_i(\mathbf{0}) > 0$ for some i; then by (3.5), the derivative $F_i'(0) > 0$, and so $F_i(s) > 0$ for all $0 < s \leq R$. Since the system (3.4) is irreducible and aperiodic, it follows that $F_j(s) > 0$ for all indices j and $0 < s \leq R$. Consequently, the Jacobian matrix $\mathcal{L}(s) = (\partial Q/\partial w)_{F(s)}$ is not identically zero, and so the spectral radius of $s\mathcal{L}(s)$ cannot remain bounded below 1 for all $s > 0$; thus, $R < \infty$. Finally, $F(s)$ must remain bounded as $s \to R-$ because (a) if $Q(w)$ is linear then the components of $F(s)$ grow at most quadratically in s, by (3.4); and (b) if $Q(w)$ is nonlinear then $(\partial Q/\partial w)_{F(s)}$ has entries that grow at least linearly in $F(s)$, and so it would be impossible for $F(s)$ to grow unboundedly without having the spectral radius $s\lambda(s)$ exceed the value 1. □

We are now in a position to determine the nature of the singularity at $s = R$:

Theorem 3.5. *Assume that the components $Q_i(w)$ of the mapping Q are given by convergent power series with nonnegative coefficients, and assume further that the Lagrangian system (3.4) is irreducible and aperiodic. If $Q_i(\mathbf{0}) > 0$ for some index i, and if the mapping $Q(w)$ is* not *linear, then for each index i there exists $C_i > 0$ such that*

$$F_i(R) - F_i(z) \sim C_i\sqrt{R - z} \qquad \text{as } z \to R. \tag{3.12}$$

Proof. Denote by h and w the right and left Perron-Frobenius eigenvectors of $R\mathcal{L}(R)$; by the Perron-Frobenius theorem, the entries of h and w are strictly positive. Define projection operators P_U, P_V by

$$\begin{aligned} P_U x &= (w^T x)h \\ P_V x &= x - P_U x. \end{aligned}$$

Denote by $\mathcal{V}$ the range of the projection P_V. Observe that because all points of the spectrum of $R\mathcal{L}(R)$ other than the simple eigenvalue 1 are of modulus < 1, the restriction of the operator $I - P_V R\mathcal{L}(R)$ to the subspace $\mathcal{V}$ is invertible. We will show that as $z \to R$,

$$P_U(F(R) - F(z)) = \{C\sqrt{R - z}\}h + o(\sqrt{R - z}) \qquad \text{and} \tag{3.13}$$

$$P_V(F(R) - F(z)) = O(R - z) \tag{3.14}$$

where $C > 0$. The desired relations (3.12) will then follow, by the strict positivity of the eigenvectors w and h, since the functions $F_i(s)$ are positive for positive arguments s.

The key to relations (3.13)–(3.14) is that the linearization (3.5) of the Lagrangian equations (3.4) becomes singular at $z = R$: in particular, it cannot be solved for dF in terms of dz. Hence, we shall look at higher order terms in the expansion of (3.4) around $z = R$. Write

$$\begin{aligned} V(z) &= P_V F(z), & \Delta V &= V(R) - V(z), \\ U(z) &= P_U F(z) = (w^T F(z))h, & \Delta U &= U(R) - U(z), \\ u(z) &= w^T F(z), & \Delta F &= F(R) - F(z), \\ &\text{and} & \Delta z &= R - z. \end{aligned}$$

The functional equation (3.4) and Taylor's theorem (in its multivariate form) imply that

$$\Delta F = (F(R)/R)\Delta z + R\mathcal{L}(R)\Delta F - \mathcal{H}(\Delta F, \Delta F) + \text{higher order terms}, \tag{3.15}$$

where each component of $\mathcal{H}$ is a symmetric bilinear quadratic form. Because the power series coefficients of the mappng Q are nonnegative, so are the coefficients of $\mathcal{H}$, and because Q is nonlinear, at least some of these coefficients are positive. Applying the projection operators P_U, P_V to equation (3.15) and using the bilinearity of $\mathcal{H}$ yields

(3.16)
$$\Delta U = (U(R)/R)\Delta z + P_U R\mathcal{L}(R)\Delta U - \mathcal{H}(\Delta U, \Delta U) + O(|\Delta U||\Delta V| + |\Delta V|^2) \qquad \text{and}$$

(3.17)
$$\Delta V = (V(R)/R)\Delta z + P_V R\mathcal{L}(R)\Delta V + O((\Delta U)^2) + O(|\Delta U||\Delta V| + |\Delta V|^2).$$

Since P_U is the projection onto the Perron-Frobenius eigenspace, the second term on the right side of (3.16) cancels the left side, and so the equation may be rewritten as

$$\kappa(\Delta u)^2 = (u(R)/R)\Delta z + \text{higher order terms.} \tag{3.18}$$

Observe that the higher order terms are $O(|\Delta u||\Delta V| + |\Delta V|^2)$ or $o(\Delta z)$; also, because the entries of the eigenvector w are positive, $u(R) > 0$. Because the restriction of $I - P_V R\mathcal{L}(R)$ to the subspace $\mathcal{V}$ is invertible, the linearization of (3.17) can be solved for ΔV in terms of Δz, and so

$$|\Delta V| = O(|\Delta z| + |\Delta u|^2). \tag{3.19}$$

Thus, to prove relations (3.13) and (3.14), it suffices to show that the constant κ in (3.18) is positive. But this follows from the fact, noted above, that the coefficients of the Hession $\mathcal{H}$ are positive, because each entry ΔF_i of ΔF has the form

$$\Delta F_i = (\Delta u)h_i + \text{a linear function of } \Delta V.$$

$\square$

The Tauberian theorem of Flajolet and Odlyzko now yields the following corollary.

Corollary 3.6. *Under the hypotheses of Theorem 3.5, the coefficients of the power series $F_i(z) = \sum_n a_{n,i} z^n$ obey the asymptotic relations*

$$a_{n,i} \sim C_i / R^n n^{3/2}. \tag{3.20}$$

3.5. **Consequence: Local limit theorem for NNRW on $\mathcal{T}^d$.** As a first example, consider nearest-neighbor random walk on a homogeneous tree $\mathcal{T}^d$, or equivalently, on the free product of d copies of $\mathbb{Z}_2$. Assume that the holding probability $p_\emptyset > 0$, so that the random walk is aperiodic; then the Green's function $G(z)$ is regular at every point on the circle of convergence $|z| = R$ except at $z = R$. Assume also that the nearest-neighbor transition probabilities $p_i > 0$, so that the random walk, and the corresponding Lagrangian system (2.9), are irreducible. Then by Theorem 3.5, the first-passage generating functions $F_i(z)$ have square-root singularities (3.12), and so by Corollary 3.6 their coefficients satisfy

$$P\{\tau_i = n\} \sim C_i / R^n n^{3/2}. \tag{3.21}$$

Furthermore, by Proposition 2.1, the Green's function is finite at $z = R$, and so by (2.8), it too must have a square-root singularity, that is,

$$G(R) - G(z) \sim \kappa\sqrt{R - z}. \tag{3.22}$$

Therefore, by Corollary 3.6,

$$P\{X_n = \emptyset\} \sim C / R^n n^{3/2}. \tag{3.23}$$

This was first proved by Gerl and Woess [8], using somewhat different methods.

4. Random walks on regular languages

4.1. **Definition.** A random walk on a free product takes values in the set of all finite words whose letters are elements of the set of indices of the groups in the free product. It evolves by successive modifications of the letters at the end of the word (in the nearest-neighbor case, these are just deltetions or adjunctions), where the modifications are determined by random inputs (in the nearest-neighbor case, words of length 0 or 1). The rules that determine how a particular input is used to modify the current state of the random walk are just the group multiplication laws. These rules are *local*: for nearest-neighbor random walks, they use the last letter of the current state; and for finite-range random walks, where the step distribution is suported by the set of words of length $\leq K$, the rules use only the last K letters of the current state.

A natural and useful generalization is to Markov chains on the set of finite words over a finite or denumerable alphabet which evolve in the manner just described, *except* that the rules determining how inputs are used to modify the current state are no longer given by a group multiplication law. We shall call such Markov chains *random walks on regular languages*, for reasons that we shall explain shortly. Roughly, a random walk on a regular language is a Markov chain on the set of all finite words from a finite alphabet A whose transition probabilities obey the following rules: (1) Only the last two letters of a word may be modified in one jump, and at most one letter may be adjoined or deleted. (2) Probabilities of modification, deletion, and/or adjunction depend only on the last two letters of the current word. More precisely, a RWRL is a Markov chain whose transition probabilities satisfy:

$$
\begin{gathered}
P(X_{n+1} = x_1x_2\dots x_m a'b' \,|\, X_n = x_1x_2\dots x_m ab) = p(a'b'|ab) \\
P(X_{n+1} = x_1x_2\dots x_m a' \,|\, X_n = x_1x_2\dots x_m ab) = p(a'|ab) \\
P(X_{n+1} = x_1x_2\dots x_m a'b'c \,|\, X_n = x_1x_2\dots x_m ab) = p(a'b'c|ab) \\
P(X_{n+1} = a'b' \,|\, X_n = a) = p(a'b'|a) \\
P(X_{n+1} = a' \,|\, X_n = a) = p(a'|a) \\
P(X_{n+1} = \emptyset \,|\, X_n = a) = p(\emptyset|a) \\
P(X_{n+1} = a \,|\, X_n = \emptyset) = p(a|\emptyset) \qquad \text{and} \\
P(X_{n+1} = \emptyset \,|\, X_n = \emptyset) = p(\emptyset|\emptyset).
\end{gathered} \tag{4.1}
$$

Here $\emptyset$ denotes the empty word. We do *not* assume that the transition probabilities are all positive, nor do we assume that there exist positive-probability paths connecting any two words. Later, however, we shall impose certain further restrictions regarding aperiodicity and irreducibility.

Denote by $\mathcal{L}$ the set of all words that are accessible from the root $\emptyset$ by positive probability paths. It is not difficult to see that $\mathcal{L}$ is a *regular language*, that is, there is a finite-state automaton $\mathcal{M}$ such that $\mathcal{L}$ is the set of all words accepted by $\mathcal{M}$. (See [13] for background on finite-state automata and regular languages.) This explains the term *random walk on regular language.*

4.2. **Examples.**

4.2.1. *Finite-range reflecting random walks on $\mathbb{Z}_+$.* Consider a Markov chain on $\mathbb{Z}_+$ with transition probabilities

$$\begin{aligned} &P(X_{n+1} = x + k \,|\, X_n = x) = p_k && \text{for } |k| \leq K \text{ and } x \geq K; \\ &P(X_{n+1} = x + k \,|\, X_n = x) = p_{x,k} && \text{for } 0 \leq x < K \text{ and } -x \leq k \leq K. \end{aligned} \tag{4.2}$$

Such a process behaves as a (homogeneous) random walk with increments bounded by K outside a finite neighborhood $\{0, 1, \cdots, K-1\}$ of the origin. It is not difficult to see that a Markov chain governed by transition probabilities (4.2) is equivalent to a random walk on the regular language

$$\mathcal{L} = \{a^m b \,|\, m \in \mathbb{Z}_+ \text{ and } b \in \{0, 1, 2, \cdots, K-1\}\}.$$

The word $a^m b$ corresponds to the integer $mK + b$. Since the increments in X_n are of magnitude $\leq K$, only the last two letters of the representing word $aa \cdots ab$ need be modified to account for any possible change of state. Local limit theorems for Markov chains with transition probabilities (4.2) were proved, by Wiener-Hopf methods, in [17].

4.2.2. *Finite-range random walks on homogeneous trees.* Let X_n be a finite-range random walk on the free product $\Gamma = \mathbb{Z}_2^d$, as defined in section 2 above. Assume that the step distribution has support contained in the set of all words of length $\leq K$ from the alphabet $[d]$. Then the random walk X_n admits a description as a random walk on the regular language consisting of finite words $a_1 a_2 \ldots a_m b$, with $m \geq 0$, such that (i) each a_i is a word of length K in the letters $[d]$; (ii) b is a word of length $0 \leq |b| \leq K-1$; and (iii) no cancellations between successive a_i, a_{i+1} (or between a_m, b) are possible.

Observe that if the transition probabilities are modified at finitely many elements of Γ, then the resulting Markov chain, a *perturbed random walk*, although no longer a "random walk" (in the usual sense, that is, that the transition probabilities are group-invariant), is nevertheless still a random walk on a regular language. We will see that the transition probabilities of such modified random walk satisfy local limit theorems, but not necessarily with the same $3/2-$power law as in (3.23).

4.2.3. *Random walk on the modular group $PSL(2, \mathbb{Z})$.* The *modular group* is the quotient group $PSL(2, \mathbb{Z}) = SL(2, \mathbb{Z}) / \{\pm I\}$ where $SL(2, \mathbb{Z})$ is the (multiplicative) group of 2×2 matrices with integer entries and determinant 1, and I is the 2×2 identity matrix. (For information about the modular group and its connections with the theory of elliptic modular forms, and also for specific facts used in the subsequent discussion, see [19], especially Chapter XI.) A *(right) random walk* on $PSL(2, \mathbb{Z})$ is a sequence

$$X_n = \xi_1 \xi_2 \ldots \xi_n \tag{4.3}$$

where $\xi_1, \xi_2, \ldots$ are i.i.d. random variables valued in $PSL(2, \mathbb{Z})$ whose distribution has finite support. We will show that any such random walk has a description as a random walk on a regular language:

The modular group is finitely generated, with generators

$$U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad U^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \qquad \text{and } T = T^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{4.4}$$

(More accurately, the group is generated by the *equivalence classes* $\langle T\rangle := \{\pm T\}$ and $\langle U\rangle := \{\pm U\}$. Henceforth, we shall not distinguish between matrices and their equivalence classes in $PSL(2,\mathbb{Z})$.) These generators satisfy the relations

$$T^2 = (TU)^3 = I. \tag{4.5}$$

Thus, every element of $PSL(2,\mathbb{Z})$ may be written as a finite word in the letters T, U, U^{-1} in which no two T's appear consecutively, and no U is adjacent to a U^{-1}. It follows that the modular group is also generated by T, W, W^2, where $W = TU$. The relations (4.5) translate as $T^2 = W^3 = I$. Every element $M \in PSL(2,\mathbb{Z})$ has a representation as a finite word

$$M = T^a W^{n_1} T W^{n_2} T W^{n_3} \cdots T W^{n_k} T^b \tag{4.6}$$

where $a, b = 0$ or 1 and each n_i is either 1 or 2.

The modular group contains free normal subgroups of finite index, and so it has a representation as a regular language. Following is a sketch of how such a representation may be obtained (see [19], section XI.3E for details). Let Γ' be the *commutator subgroup* of $PSL(2,\mathbb{Z})$, that is, the subgroup consisting of all finite products of *commutators.* (The *commutators* of any group are the elements of the form $aba^{-1}b^{-1}$.) The commutator subgroup of any finite or countable group is a normal subgroup, and so Γ' is a normal subgroup of $PSL(2,\mathbb{Z})$. The group Γ' is free (see [19], Theorem XI.3E) , with generators

$$\begin{aligned} X &= TWTW^2, & & & X^{-1} &= WTW^2T, \\ Y &= TW^2TW, & &\text{and} & Y^{-1} &= W^2TWT. \end{aligned} \tag{4.7}$$

The index of Γ' in $PSL(2,\mathbb{Z})$ (that is, the number of distinct cosets of Γ') is 6. Thus, since Γ' is a normal subgroup, one may choose 6 distinct elements $A_1, A_2, \dots, A_6$ of $PSL(2,\mathbb{Z})$, one from each coset, such that every element of $PSL(2,\mathbb{Z})$ has a representation

$$M = WA_i, \tag{4.8}$$

where W is a reduced word in the (free) generators X, X^{-1}, Y, Y^{-1}. Observe that there may be some cancellation of letters at the end of W; but since there are only 6 possibilities for A_i, there are only finitely many possible types of cancellation, involving at most a bounded number of letters at the end of W. Thus, there is a finite subset $\mathcal{B} \subset PSL(2,\mathbb{Z})$ such that every element of $PSL(2,\mathbb{Z})$ has a representation as a *reduced* word

$$M = W'B_j, \tag{4.9}$$

where W' is a reduced word in X, X^{-1}, Y, Y^{-1} and $B_j \in \mathcal{B}$. This exhibits $PSL(2,\mathbb{Z})$ as a regular language $\mathcal{L}$, with alphabet $\mathcal{A} = \left\{X, X^{-1}, Y, Y^{-1}\right\} \cup \mathcal{B}$.

Now consider a right random walk (4.3) whose jumps ξ_i are i.i.d. from a distribution on $PSL(2,\mathbb{Z})$ with finite support. Because there are only finitely many possible values of ξ_{n+1}, there are only finitely many different ways that multiplication by ξ_{n+1} can modify the end of the word X_n. In particular, there is a finite integer K such that all positive-probability transitions $X_n \to X_{n+1}$ involve only the last K letters of X_n, and such that the transition probabilities depend only on the last K letters of the current state X_n. By replacing the regular language $\mathcal{L}$ by a regular language $\mathcal{L}'$ whose alphabet consists of the words of length $\leq K$ in $\mathcal{L}$, one obtains a representation of the random walk X_n as a random walk on a regular language.

A local limit theorem for *nearest-neighbor* random walk on the modular group was proved by Woess [26], using different and somewhat more difficult arguments than those employed here. Later [27], Woess also proved that, for any nearest-neighbor random walk on a discrete group containing a free normal subgroup of finite index, the Green's function is algebraic; however, his arguments do not allow one to deduce the type of the leading singularity.

4.3. **The Green's function of a RWRL.** Assume now that X_n is a random walk on a regular language $\mathcal{L}$. For $x, y \in \mathcal{L}$, define the *Green's function(s)* $G_{xy}(z)$ as follows:

$$G_{xy}(z) := \sum_{n=0}^{\infty} P^x\{X_n = y\} z^n = E^x \left(\sum_{n=0}^{\infty} z^n \mathbf{1}\{X_n = y\} \right). \tag{4.10}$$

Thus, $G_{xy}(z)$ is the expected discounted number of visits to y given that the initial state is x, where z is the discount factor. Since the coefficients in the power series are all probabilities, the series converge absolutely and uniformly in the closed unit disc $|z| \leq 1$, and so the radius of convergence is at least 1. Later we shall prove that, under suitable irreducibility hypotheses, the power series of all of the Green's functions $G_{xy}(z)$ have the same radius R of convergence.

The Green's functions $G_{xy}(z)$ are determined by a finite system of generating functions $H_{ab,c}(z)$, the *reentry* generating functions, indexed by words ab and c of lengths 1 and 2, respectively. These are defined as follows:

$$H_{ab,c}(z) := E^{ab} z^T \mathbf{1}\{X_T = c\} = \sum_{n=1}^{\infty} z^n P^{ab}\{T = n \text{ and } X_n = c\} \tag{4.11}$$

where

$$T := \min\{n : |X_n| < |X_0|\}. \tag{4.12}$$

Observe that some — or even all — of the generating functions $H_{ab,c}(z)$ may be identically zero, as there may be no positive-probability paths from ab to c through words of length ≥ 2. Those that are not identically zero we shall refer to as the *nondegenerate* reentry generating functions.

The relationship between the Green's functions and the reentry generating functions $H_{ab,c}(z)$ is a routine consequence of the Markov property. Let x and y be words of lengths $|x|, |y| \leq 1$. If $X_0 = x$, then X_1 must be a word of length $0, 1$, or 2; if X_1 is a word of length 2, then there can be no more visits to y until after the Markov chain X_n returns from the set of words of length ≥ 2. Hence, by the Markov property, if $|x| \leq 1$ and $|y| \leq 1$, then

$$G_{xy}(z) = \delta_x(y) + z \sum_{w:|w|\leq 1} p(w|x) G_{wy}(z) + z \sum_{ab} \sum_{c} p(ab|x) H_{ab,c}(z) G_{cy}(z), \tag{4.13}$$

where $\sum_{ab}$ is over all two-letter words, and $\sum_c$ is over all one-letter words.

The equations (4.13) may be written compactly in matrix form as follows. Define $\mathbf{G}(z)$ to be the matrix-valued function of z with entries $G_{xy}(z)$, indexed by words x, y of lengths ≤ 1, and define $\mathbf{P}$ to be the matrix of transition probabilities $p(y|x)$, where again x and y are words of length ≤ 1. Define $\mathbf{K}(z)$ to be the matrix-valued function of z with entries $K_{xy}(z)$, indexed by words x, y of length ≤ 1, defined by

$$\begin{aligned} K_{xy}(z) &= \sum_{ab} p(ab|x) H_{ab,y}(z) \qquad && \text{if } |y| = 1; \\ &= 0 && \text{if } |y| = 0. \end{aligned} \tag{4.14}$$

Then equations (4.13) are equivalent to the matrix equations

$$\begin{aligned} \mathbf{G}(z) &= \mathbf{I} + z\mathbf{P}\mathbf{G}(z) + z\mathbf{K}(z)\mathbf{G}(z) \qquad \Longleftrightarrow \\ \mathbf{I} &= (\mathbf{I} - z\mathbf{P} - z\mathbf{K}(z))\,\mathbf{G}(z) \end{aligned} \tag{4.15}$$

where $\mathbf{I}$ is the identity matrix. Notice that for all z of sufficiently small modulus $|z|$, the matrix $(z\mathbf{P}+z\mathbf{K}(z))$ has norm less than one, and so the matrix $(\mathbf{I}-z\mathbf{P}-z\mathbf{K}(z))$ is invertible; thus,

$$\boxed{\mathbf{G}(z) = (\mathbf{I} - z\mathbf{P} - z\mathbf{K}(z))^{-1}} \tag{4.16}$$

for all z in a disc centered at $z = 0$.

4.4. **The Lagrangian system of a RWRL.** Recall (see equations 4.1)) that, for a random walk on a regular language, only the last two letters of the current word play a role in the transition to the next word. Consequently, for any initial state $\xi = x_1x_2\dots x_mab \in \mathcal{L}$, the letters $x_1x_2\dots x_m$ are ignored up to time T. It follows that for any word $\xi = x_1x_2\dots x_mab \in \mathcal{L}$, and any $c \in A$,

$$H_{ab,c}(z) = E^{\xi}z^{T}\mathbf{1}\{X_T = x_1x_2\dots x_mc\}. \tag{4.17}$$

The homogeneity relations (4.17) lead directly to a system of quadratic equations relating the functions $H_{ab,c}$. These equations derive from the Markov property: Starting from any word w, the random walk must make an initial jump to another state. If this first jump is to a word whose length is less than that of the initial word w, then $T = 1$; otherwise, the initial jump must be either to a word w' of the same length, or to a word w'' with one additional letter. In the latter case, T is the first time after the initial jump that *two* letters are removed; in the former, T is the first time that a single letter is removed. Summing over all possible initial jumps and applying the identity (4.17) yields the following system of equations:

$$\begin{aligned} H_{ab,c}(z) = zp(c|ab) + zp(ab|ab)H_{ab,c}(z) + z\sum_{d,e\in A} p(de|ab)H_{de,c}(z) \\ +z\sum_{d,e,f\in A} p(def|ab)\sum_{g\in A} H_{ef,g}(z)H_{dg,c}(z), \end{aligned} \tag{4.18}$$

We shall abbreviate this system of equations as

$$\boxed{\mathbf{H}(z) = z\mathbf{Q}(\mathbf{H}(z))} \tag{4.19}$$

where $\mathbf{H}(z)$ is the vector of *nondegenerate* first-passage generating functions $H_{ab,c}(z)$ (that is, those that are *not* identically zero), and $\mathbf{Q}$ is the vector of quadratic polynomials $Q_{ab,c}$ in the variables $H_{ab,c}$ occurring on the right sides of the equations (4.18) above. Observe that $\mathbf{Q}$ is quadratic (that is, each of its components is quadratic) and its coefficients are nonnegative. Consequently, the solution set of (4.19) is an *algebraic curve*, and so the functions $H_{ab,c}(z)$ are *algebraic functions*. Most important, the system (4.19) is of the generic form (3.4), and so the results of section 3 apply.

The hypotheses of Theorem 3.5 require that the Lagrangian system be irreducible and aperiodic, that is, that the incidence matrix of the dependency graph be a Perron-Frobenius matrix. Following is a relatively simple set of conditions on the transition probabilities of the RWRL that will guarantee that the Lagrangian system (4.18) is irreducible and aperiodic.

These are not in any sense minimal, nor do they cover all cases of the examples discussed in section 4.2 above.

Assumption 4.1 (Aperiodicity). *For all words $x_1x_2\dots x_m$,*

$$P(X_{n+1} = x_1x_2\dots x_m \,|\, X_n = x_1x_2\dots x_m) > 0 \tag{4.20}$$

This is equivalent to assuming that all of the transition probabilities $p(ab\,|\,ab), p(a\,|\,a)$, and $p(\emptyset\,|\,\emptyset)$ are positive.

Assumption 4.2 (Irreducibility-A). *Let $\mathcal{L}$ be the set of all words that may be reached from $\emptyset$ via positive-probability paths. Then $\forall w \in \mathcal{L}$ there is a positive-probability path from w to $\emptyset$.*

Assumption 4.3 (Irreducibility-B). *For every word $x = x_1x_2\dots x_m \in \mathcal{L}$ of length 2 (or longer) and every triple abc of letters there exist words*

$$w_1 = x_1x_2\dots x_m y_1y_2\dots y_n ab \in \mathcal{L} \quad \textit{and}$$
$$w_2 = x_1x_2\dots x_m y_1y_2\dots y_n c \in \mathcal{L}$$

and positive probability paths

$$x \longrightarrow w_1 \longrightarrow w_2 \longrightarrow x$$

through words prefaced by x, w_1, and x, respectively.

Proposition 4.4. *If the transition probabilities of a RWRL satisfy Assumptions 4.1, 4.2, and 4.3 above, then the Lagrangian system (4.18) is irreducible and aperiodic. Furthermore, the mapping $\mathbf{Q}$ will be nonlinear and will satisfy the restriction $\mathbf{Q}(\mathbf{0}) \neq \mathbf{0}$.*

Proof. Assumption 4.1 ensures that the incidence matrix of the dependency graph will have positive entries on the diagonal (because the second term on the right side of (4.18) will be positive). Hence, the system will be aperiodic. Assumption 4.3 guarantees that, for each pair of triples $a'b'c'$ and abc it is possible to make a succession of substitutions in 4.18) (in the last sum) so as to obtain an equation for $H_{a'b',c'}(z)$ in which $H_{ab,c}(z)$ appears as a factor in a term with positive coefficient. Thus, the system will be irreducible. Moreover, the right side of (4.18) will include quadratic terms with positive terms, so $\mathbf{Q}$ will be nonlinear. Finally, Assumption 4.2 implies that at least some of the transition probabilities $p(c|ab)$ will be positive, so $\mathbf{Q}(\mathbf{0}) \neq \mathbf{0}$ (because the first term on the right side of at least one of the equations (4.18) will be positive). □

Corollary 4.5. *Under Assumptions 4.1, 4.2, and 4.3, for every triple of letters abc, there are constants $C_{ab,c} > 0$ so that*

$$P^{ab}\{T = n\,;\, X_T = c\} \sim \frac{C_{ab,c}}{\varrho^n n^{3/2}}, \tag{4.21}$$

where $\varrho \geq 1$ is the smallest positive singularity of the Lagrangian system (4.18).

4.5. **Local limit theorems for RWRL.** For nearest-neighbor random walks on homogeneous trees, the spectral radius R (the radius of convergence of the Green's function) coincides with the smallest positive singularity ϱ of the associated Lagrangian system. For random walks on regular languages, this need not be the case. Furthermore, the lead singularity of the Green's function need not be a square-root singularity, as it is for the Lagrangian system, and consequently the return probabilities need not obey a $3/2-$power law, as they must for random walks on homogeneous trees. In fact, there are five distinct possibilities. These are enumerated in the following theorem.

Theorem 4.6. *Assume that the transition probabilities of the RWRL X_n satisfy the assumptions 4.1, 4.2, and 4.3. Then the spectral radius R of the random walk is bounded above by the lead singularity ϱ of the associated Lagrangian system. The random walk may be positive recurrent, null recurrent, or transient. In the null recurrent case, $R = 1$ and for any two words $x, y \in \mathcal{L}$ there exists a positive constants C_{xy} such that*

$$P^x \{X_n = y\} \sim C_{xy} n^{-1/2}. \tag{4.22}$$

In the transient case, $R > 1$, and for suitable positive constants C_{xy} one of the following laws holds for all pairs $x, y \in \mathcal{L}$:

$$P^x \{X_n = y\} \sim C_{xy} R^{-n}; \text{ or} \tag{4.23}$$

$$P^x \{X_n = y\} \sim C_{xy} R^{-n} n^{-1/2}; \text{ or} \tag{4.24}$$

$$P^x \{X_n = y\} \sim C_{xy} R^{-n} n^{-3/2}. \tag{4.25}$$

All five cases are possible. This may be verified by considering nearest-neighbor reflecting random walks on the nonnegative integers $\mathbb{Z}_+$, whose Green's functions may be obtained in closed form by solving quadratic equations (EXERCISE!).

Proof of Theorem 4.6 (Sketch). We shall consider only words $x, y \in \mathcal{L}$ of length 1, and leave as an exercise for the reader to complete the argument. First, observe that the Green's functions $G_{xy}(z)$ all have the same radius of convergence. This follows from the irreducibility assumptions 4.2–4.3, which guarantee that there exist positive-probability paths from x to y and back, because this implies that there is a positive constant c and integers k, l such that for all $n \geq 0$,

$$\begin{aligned}
P^x\{X_{n+k} = y\} &\geq cP^x\{X_n = x\}, \\
P^x\{X_{n+k} = y\} &\geq cP^y\{X_n = y\}, \\
P^x\{X_{n+l} = x\} &\geq cP^x\{X_n = y\}, \qquad \text{and} \\
P^x\{X_{n+k} = x\} &\geq cP^y\{X_n = x\}.
\end{aligned}$$

Recall that the matrix $\mathbf{G}(z)$ of Green's functions $G_{xy}(z)$ indexed by words of length 1 is determined by the reentry generating functions by equations (4.16). The matrix $\mathbf{P}$ in (4.16) is constant, and the entries of $\mathbf{K}(z)$ are positive linear combinations of the reentry generating functions $H_{ab,c}(z)$; consequently, the radius R of convergence of $\mathbf{G}(z)$ cannot exceed the radius ϱ of convergence of $\mathbf{H}(z)$. In particular, if the spectral radius of $\varrho\mathbf{P}+\varrho\mathbf{K}(\varrho)$ is less than one, then (4.16) implies that $\mathbf{G}(z)$ must have the same type of singularity — a square-root singularity – at $z = \varrho$ as does $\mathbf{H}(z)$. This implies that $R = \varrho \geq 1$ and that the asymptotic relations (4.25) must hold, by the Tauberian theorem of Flajolet and Odlyzko.

It is also possible that the spectral radius of $s\mathbf{P} + s\mathbf{K}(s)$ attains the value 1 at some $s = R < \varrho$. If $R < \varrho$, then since $\mathbf{K}(z)$ is analytic at $z = \varrho$, the function $\mathbf{G}(z)$ will have a pole at $z = R$. This pole must be simple, because $\mathbf{P} + \mathbf{K}(z)$ is a Perron-Frobenius matrix at $z = R$. Thus, in this case, the asymptotic relations (4.23) hold, possibly with $R = 1$. (If $R = 1$ then the RWRL is positive recurrent.)

The last possibility is that the spectral radius of $s\mathbf{P} + s\mathbf{K}(s)$ attains the value 1 at $s = R = \varrho$. In this case the singularity of $\mathbf{G}(z)$ at $z = R$ is such that for each entry $G_{xy}(z)$,

$$G_{xy}(R) - G_{xy}(z) \sim \frac{c_{xy}}{\sqrt{R-z}}. \tag{4.26}$$

The proof of this is somehat delicate, and uses the fact that the matrix $R\mathbf{P} + R\mathbf{K}(R)$ is a Perron-Frobenius matrix. See [18] for details. In brief, the reasoning is as follows: by (4.16), the function $\mathbf{G}(z)^{-1}$ behaves near $z = R$ like

$$\mathbf{M}\sqrt{R-z} + \mathbf{N}(z)$$

where $\mathbf{M}$ is a rank-one constant matrix and $\mathbf{N}(z)$ is analytic near $z = R$, and so the inverse of $\mathbf{G}(z)$ blows up like $1/\sqrt{R-z}$. Theorem 3.1 implies that if the Green's function $G_{xy}(z)$ behaves like (4.26) at the singularity $z = R$, then the coefficients must satisfy (4.24).

It remains to show that if the transition probabilities obey the asymptotic law (4.25) then $R > 1$. Note that (4.25) implies that the Markov chain X_n is transient. Consequently, there is at least one two-letter word ab such that $\sum_y H_{ab,y}(1) < 1$ (that is, so that the probability of return to the set of one-letter words from ab is less than one). The irreducibility assumptions ensure that the number of steps until X_n ends with the letters ab has an exponentially decaying tail (uniformly in all starting states, since the alphabet is finite). At each such time, the RWRL may choose to never the set of words of length $< |X_n|$: this occurs with positive (conditional) probability independent of the letters of X_n that precede the closing ab. Thus, the probability that $|X_n| \leq 1$ must decay at least exponentially, that is $R > 1$.

5. Infinite-dimensional Lagrangian systems

In section 3 we showed that for a large class of *finite-dimensional* Lagrangian systems the lead singularity is of the square-root type (3.12). This, together with the Tauberian theorem of Flajolet and Odlyzko, implies that the power series coefficients of the component functions satisfy a $3/2-$power law. In this section, we consider the simplest infinite-dimensional case, where the spectrum of the Jacobian operator $\partial Q/\partial F$ near the spectral radius consists of eigenvalues of finite multiplicity. In section 6 below we will show how this analysis may be used to obtain local limit theorems for random walks on *infinite* free products.

5.1. **Lyapunov-Schmidt reduction.** Recall that by the Implicit Function Theorem, the Lagrangian system

$$F(z) = Q(F(z)) \tag{5.1}$$

has a solution $z = 0$, $F(0) = 0$ that admits an analytic continuation along any curve on which the spectral radius of the operator

$$z\mathcal{L}(z) = z\left(\frac{\partial Q}{\partial w}\right)_{w=F(z)} \tag{5.2}$$

remains less than 1. Unfortunately, the spectral radius of a continuous, operator-valued function need not be itself continuous (although it must be at least lower semi-continuous), and so the singular points of the solution $F(z)$ of (3.4) may not in general be located by searching for the points z where the spectrum of $z\mathcal{L}(z)$ includes the value 1. However, in certain problems the spectral radius of $\mathcal{L}(s)$ will for positive s be an isolated eigenvalue of finite multiplicity, and in such cases the spectral radius of $\mathcal{L}(z)$ will vary continuously with z near the positive axis $s > 0$. When this happens, the singular behavior of $F(z)$ at $z = R$ will essentially be determined by a finite-dimensional section, and classical methods of algebraic geometry (the Weierstrass preparation theorem, Newton diagrams, Puiseux expansions, etc.) may be used. This program is known in nonlinear analysis as *Lyapunov-Schmidt reduction*; it is commonly used in bifurcation theory (see, for instance, [28], chapter 8).

Hypothesis 5.1. *There exists $R \in (0, \infty)$ such that*
- (a) *the spectral radius of $s\mathcal{L}(s)$ is less than 1 for all $s \in [0, R)$;*
- (b) *$\lim_{z\to R} F(z) = F(R)$ exists and is finite;*
- (c) *1 is an isolated eigenvalue of $R\mathcal{L}(R)$ with finite multiplicity.*

Condition (a) guarantees that $F(z)$ has a unique analytic continuation along the line segment $[0, R)$, and condition (b) implies that the Jacobian operator $\mathcal{L}(z)$ has a limit $\mathcal{L}(R)$ as $z \to R$ (from inside the disk). Condition (c) implies that the Banach space B may be decomposed as a direct sum

$$B = V \oplus W, \tag{5.3}$$

where V is the space of eigenvectors of $R\mathcal{L}(R)$ with eigenvalue 1, and that there is a projection operator $P_V : B \to V$ with range V that commutes with $\mathcal{L}(R)$, given by

$$P_V = \frac{1}{2\pi i} \oint (\zeta - R\mathcal{L}(R))^{-1}\, d\zeta, \tag{5.4}$$

where the contour integral extends over a circle in the complex plane surrounding the point $\zeta = 1$ that contains no other points of the spectrum of $R\mathcal{L}(R)$ (see [14], Theorem 6.17). It then follows that

$$P_W = I - P_V \tag{5.5}$$

is a projection operator with range W that also commutes with $R\mathcal{L}(R)$. That 1 is an isolated point of the spectrum of $R\mathcal{L}(R)$ implies that 1 is *not* in the spectrum of the restriction to W of $P_W R\mathcal{L}(R)$, equivalently, $I - P_W R\mathcal{L}(R)$ is invertible on W.

Proposition 5.2. *Assume that Hypothesis 5.1 holds, and let P_V be the projection operator defined by (5.4). Then in some neighborhood of $z = R$, the $V-$valued function $P_V F(z)$ satisfies a functional equation of the form*

$$P_V F(z) = zQ_V(z, P_V F(z)), \tag{5.6}$$

where $Q_V(z, v)$ is a function of $\dim V + 1$ complex variables that is holomorphic in a neighborhood of $z = R$, $v = P_V F(R)$.

Proof. Consider the mapping $K : \mathbb{C}\times V\times W \to W$ defined by $K(z, v, w) = w-zP_W Q(v+w)$. This is holomorphic, and takes the value 0 at the point $\omega := (R, P_V F(R), P_W F(R))$. I claim that there is a holomorphic mapping $W(z, v)$, valued in W and defined in a neighborhood of $z = R$, $v = P_V F(R)$, such that for all z, v in this neighborhood,

$$W(z, v) = zP_W Q(v + W(z, v)),$$

and such that all zeros of $K(z, v, w)$ near ω are of the form $(z, v, W(z, v))$. This follows from the Implicit Function Theorem for the mapping K, because the Jacobian of the mapping

$w \mapsto K(z, v, w)$ is

$$I - zP_W(\partial Q/\partial F)_{v+w}$$

and this is nonsingular at $z = R$, $v = P_V F(R)$, since $P_W R\mathcal{L}(R) = R\mathcal{L}(R)P_W$ and $I - P_W R\mathcal{L}(R)$ is invertible on W. Now consider the functional equation (3.4). Applying the projections P_V and P_W to both sides, and using the relation $I = P_V + P_W$, one obtains

$$P_V F(z) = zP_V Q(P_V F(z) + P_W F(z)) \qquad \text{and}$$
$$P_W F(z) = zP_W Q(P_V F(z) + P_W F(z)).$$

The solution to the second of these equations must be at $P_W F(z) = W(z, P_V F(z))$, and so the first may be rewritten in the form

$$P_V F(z) = zP_V Q(P_V F(z) + W(z, P_V F(z))).$$

Since the function $W(z, v)$ is holomorphic in its arguments, this is the desired functional equation. □

5.2. **Simple eigenvalues and square-root singularities.** Hypothesis 5.1 requires only that the spectrum of the Jacobian operator $\mathcal{L}(R)$ have an isolated lead eigenvalue of finite multiplicity. If the lead eigenvalue is simple, as will often be the case, the conclusions of Proposition 5.2 can be substantially strengthened.

Hypothesis 5.3. *There exists $R \in (0, \infty)$ such that*

(a) *the spectral radius of $s\mathcal{L}(s)$ is less than 1 for all $s \in [0, R)$;*
(b) $\lim_{z\to R} F(z) = F(R) \neq 0$ *exists and is nonzero and finite;*
(c) 1 *is an isolated eigenvalue of $R\mathcal{L}(R)$ with multiplicity* 1.
(d) *The projection $P_V F(R)$ of $F(R)$ on the $1-$eigenspace V is* nonzero.

When Hypothesis 5.3 holds, the projection P_V defined by (5.4) will have a one-dimensional range, and so there will exist nonzero elements $h \in B$ and $\nu \in B^*$ (here B^* denotes the Banach space dual to B) such that for all $\varphi \in B$,

$$P_V \varphi = \langle \nu, \varphi \rangle h. \tag{5.7}$$

Note that since P_V is a projection and h is in the range of P_V,

$$\langle \nu, h \rangle = 1. \tag{5.8}$$

Proposition 5.4. *Assume that Hypothesis 5.3 holds, and let h and ν be such that (5.7) holds. Define*

$$v(z) = \langle \nu, F(z) \rangle. \tag{5.9}$$

Then there exist an integer $m \geq 1$ and a function $A(\zeta)$ holomorphic in a neighborhood of $\zeta = 0$, satisfying $A(0) = 0$, such that in some neighborhood of $z = R$,

$$v(R) - v(z) = A((R - z)^{1/m}). \tag{5.10}$$

Remark: The representation (5.10) implies that the function $v(z)$ has an analytic continuation to a slit disk centered at $z = R$ of the form

$$\{z : |z - R| < \varepsilon \quad \text{and} \quad |\arg(z - R)| > 0\} \tag{5.11}$$

and that its asymptotic behavior near $z = R$ is of the form (3.2) for some rational α.

Proof. Set $u(\xi) = v(R) - v(R-\xi)$. By Proposition 5.2, the function $P_V F(z) = v(z)h$ satisfies a functional equation (5.6) in a neighborhood of $z = R$. This may be rewritten in ξ and u as a functional equation for $u(\xi)$ of the form

$$K(\xi, u(\xi)) = 0 \tag{5.12}$$

where $K(\xi, u)$ is a function of two variables (ξ, u) that is holomorphic in a neighborhood of the origin in $\mathbb{C}^2$, and $K(0,0) = 0$. We may assume that the power series expansion of the function $K(\xi, u)$ contains a term au^d, where $a \neq 0$, that is not divisible by ξ, for if this were not the case then K could be replaced by K/ξ^m for some m. (The power series for K must include terms divisible by u, because the functional equation (5.12) derives from (5.6), which holds for $z < R$ near R.) Thus, by the Weierstrass Preparation Theorem ([11], Chapter 1) there exists a Weierstrass polynomial

$$\Phi(\xi, u) = u^d + \sum_{j=1}^{d} a_j(\xi) u^{d-j} \tag{5.13}$$

such that in some neighborhood of $(0,0) \in \mathbb{C}^2$ the zero sets of K and Φ coincide. (NOTE: The definition of a Weierstrass polynomial requires that each of the coefficients $a_j(\xi)$ be holomorphic in ξ.)

Since the ring of holomorphic functions near $(0,0)$ is a unique factorization domain ([11], Chapter 1), the polynomial Φ may be factored into irreducible Weierstrass polynomials:

$$\Phi(\xi, u) = \prod_{i=1}^{r} \Phi_i(\xi, u). \tag{5.14}$$

Here each Φ_i is irreducible in the ring of locally holomorphic functions; this means that in some neighborhood of the origin, $\Phi_i(\xi, u)$ is, for each fixed $\xi \neq 0$, irreducible in the ring $\mathbb{C}[u]$ of polynomials in the variable u. Consequently, the derivative $\partial\Phi_i/\partial u$ cannot have a zero in common with Φ_i (except when $\xi = 0$), and so by the Implicit Function Theorem, for each $\xi \neq 0$ and each u such that $\Phi_i(\xi, u) = 0$, the equation $\Phi_i(\xi', u') = 0$ defines a branch of an analytic function $u(\xi')$ near $\xi' = \xi$ and $u' = u$. For at least one of the indices i (say $i = 1$), one of the branches of the analytic function $u(\xi)$ defined by $\Phi_i = 0$ coincides with the function $u(\xi) := v(R) - v(R-\xi)$ for $\xi > 0$ small.

The result (5.10) now follows by a standard argument in the theory of algebraic functions of a single variable (see, for instance, [12], vol. 2, section 12.2). If the branch $u(\xi)$ is followed around a small contour $\xi \in \gamma$ surrounding 0, then after a finite number m of circuits $u(\xi)$ will return to the original branch (because for each ξ the polynomial $\Phi_1(\xi, u)$ has only m roots, where m is the degree of Φ_1 in the variable u). Consequently, u is an analytic function of ξ^m in a neighborhood of $\xi = 0$ (observe that $\xi = 0$ is a removable singularity because all roots of $\Phi_1(\xi, u) = 0$ approach zero as $\xi \to 0$), and the representation (5.10) follows. □

Let $\mathcal{H}(v, w)$ be the Hessian (second differential) form of the mapping Q at the point $F(R)$, that is, the symmetric bilinear form $\mathcal{H} : B \times B \to B$ defined by

$$\mathcal{H}(v, v) = \left(\frac{d^2}{d\varepsilon^2} Q(F(R) + \varepsilon v) \right)_{\varepsilon=0}. \tag{5.15}$$

Theorem 5.5. *Let $F(z)$ be the solution of the Lagrangian equation (3.4), and assume that Hypothesis 5.3 holds. Assume that the Hessian $\mathcal{H}(h, h)$ has a nonzero projection on the* *subspace V, that is,*

$$\langle \nu, \mathcal{H}(h,h)\rangle \neq 0. \tag{5.16}$$

Then the function $v(z) := \langle \nu, F(z)\rangle$ has a square-root singularity at $z = R$, that is, for some nonzero constant C,

$$v(R) - v(z) \sim C\sqrt{R-z} \tag{5.17}$$

as $z \to R$ in a slit domain (5.11). Furthermore, the function $P_W F(z)$ satisfies

$$P_W F(R) - P_W F(z) = O(|R-z|) \tag{5.18}$$

near $z = R$.

Remark. Theorem 5.5 shows that under Hypothesis 5.3 and (5.16), the function $v(z)$ satisfies most of the hypotheses of the Flajolet-Odlyzko Transfer Theorem, with $\alpha = 1/2$: the only hypothesis that requires separate verification is that v has no singularity on the circle $|z| = R$ except that at $z = R$. Puiseux expansions for other linear functionals of $F(z)$ can be deduced from (5.17)and (5.18).

Proof of Theorem 5.5. The proof is modelled on that of Theorem 3.5 above. Proposition 5.4 implies that $v(z)$ has a Puiseux expansion in powers of $(R-z)^\alpha$ for some positive rational number α. Since $P_W F(z) = W(z, v(z)h)$ where $W(z,v)$ is the holomorphic function constructed in the proof of Proposition 5.2, it follows that $P_W F(z)$ also has a Puiseux expansion in powers of $(R-z)^\alpha$. We must show that α is a multiple of $1/2^k$ for some $k \geq 1$, and that the first nonzero terms in the expansions are as advertised.

Write

$$\begin{aligned} W(z) &= P_W F(z), & \Delta z &= R - z, \\ V(z) &= P_V F(z) = v(z)h & \Delta F &= F(R) - F(z), \end{aligned}$$

etc. Recall that $v(R) \neq 0$, by (d) of Hypothesis 5.3. The functional equation (3.4) and Taylor's theorem ([28], section 4.6) imply that

$$\Delta F = (F(R)/R)\Delta z + R\mathcal{L}(R)\Delta F - \frac{1}{2}\mathcal{H}(\Delta F, \Delta F) + \mathcal{R}_3 \tag{5.19}$$

where the remainder $\mathcal{R}_3$ is $o(\Delta z + \|\Delta F\|^2)$ as $\Delta z \to 0$. Apply the projections P_V, P_W to both sides and use the bilinearity of the Hessian form to obtain

$$\Delta W = (W(R)/R)\Delta z + P_W R\mathcal{L}(R)\Delta W - P_W \mathcal{H}(\Delta V, \Delta V)/2 + \mathcal{R}_W \quad \text{and} \tag{5.20}$$

$$\Delta V = (V(R)/R)\Delta z + P_V R\mathcal{L}(R)\Delta V - P_V \mathcal{H}(\Delta V, \Delta V)/2 + \mathcal{R}_V \tag{5.21}$$

where $\mathcal{R}_W, \mathcal{R}_V = O(\|\Delta V\|\|\Delta W\| + \|\Delta W\|^2) + o(\Delta z)$. Since P_V is the projection onto the $1-$eigenspace of $R\mathcal{L}(R)$, the second term on the right side of (5.21) cancels the left side, and so the equation reduces to

$$(\Delta v)^2 \langle \nu, \mathcal{H}(h,h)\rangle = 2(v(R)/R)\Delta z + \mathcal{R}'_V \tag{5.22}$$

where $\mathcal{R}_V = \mathcal{R}'_V h$. Recall that the operator $P_W - P_W R\mathcal{L}(R)$ is invertible on the Banach space W, with inverse (say) $\mathcal{M}$, so equation (5.20) may be rewritten as

$$\Delta W = \mathcal{M}(W(R)/R)\Delta z - (\Delta v)^2 \mathcal{M}(\mathcal{H}(h,h)) + \mathcal{R}_W. \tag{5.23}$$

There are now two possibilities: (a) $\|\Delta W\| = o(|\Delta v|)$, or (b) there is a sequence of points $z_n \to R$ along which $\|\Delta w\| \geq \varepsilon|\Delta v|$ for some $\varepsilon > 0$. I claim that possibility (b) cannot occur: If so, equation (5.23) would imply that $\|\Delta W\| = O(\Delta z)$; this would imply that the remainder term in equation (5.22) is $o(\Delta z)$; and this would make (5.22) impossible, since $v(R) \neq 0$ and $\langle \nu, \mathcal{H}(h,h)\rangle \neq 0$. Thus, (a) must hold. But (a) implies that the remainder term $\mathcal{R}'_V$ is of smaller order of magnitude than the other terms in equation (5.22). It follows that $(\Delta v)^2 \sim C\Delta z$ for some nonzero constant C. This implies that the leading nonconstant term in the Puiseux expansion of $v(z)$ is $C\sqrt{R-z}$ for some nonzero C. That $P_W F(z)$ has Puiseux expansion (5.18) with first nonzero term proportional to $(R-z)$ now follows from (5.5) together with (5.23).

6. Application: Local Limit Theorem for RW on Infinite Free Products

We now turn our attention to random walks on *infinite* free products. For simplicity, we shall consider only the nearest-neighbor case. In the nearest-neighbor case, the methods of [8] provide another approach to the study of transition probabilities; however, this approach breaks down in the non-nearest-neighbor case, whereas the method we shall develop may be adapted to finite-range random walks, and even to some infinite-range random walks.

Nearest-neighbor random walk on a countable free product Γ of copies of $\mathbb{Z}_2$ was defined in section 2 above. Recall that the Green's function of a nearest-neighbor random walk (defined by equation (2.2)) is determined by the *first-passage* generating functions — see equation (2.8). The first-passage generating functions are themselves interrelated by the Lagrangian system (2.9) of quadratic equations. This system involves infinitely unknowns $F_i(z)$, and so analysis of the lead singularity requires the use of the machinery of section 5.

6.1. Holomorphic character of the link function. Theorem 5.5 requires that the function $F(z)$ take values in a Banach space B, that the link function $Q : B \to B$ be holomorphic, and that the conditions of Hypothesis 5.3 be satisfied. Of these, the most critical is that 1 should be an isolated, simple eigenvalue of the Jacobian operator $R\mathcal{L}(R)$, and the proof of this will occupy most of the argument.

Recall that, by Corollary 2.2, the function $F(z) = (F_i(z))$ takes values in the Banach space B of bounded sequences (with sup norm), for all $|z| \leq R$. The Lagrangian system (2.9) has the special form

$$F(z) = z(M(F(z)) + N(F(z)) \times F(z)) \tag{6.1}$$

where $M, N : B \to B$ are bounded linear operators and the operation $\times$ is coordinatewise multiplication. Now any linear operator $L : B \to B$ is clearly holomorphic, and by the product rule ([28], sec. 4.3) the coordinatewise product of holomorphic mappings is holomorphic, so the implied link function Q for the system (2.9) is holomorphic on B.

6.2. Spectrum of the Jacobian Operator $\mathcal{L}(s)$. Recall from equation (3.7) that the operator $\mathcal{L}(s)$ is the Jacobian $\partial Q/\partial w$ evaluated at $w = F(s)$. In our case the Lagrangian system specializes to (2.9); the function $F(z)$ has coordinates $F_i(z)$, indexed by $i \in \mathbb{N}$. Thus, by (2.9), the Jacobian in matrix form has entries

$$\begin{aligned} \mathcal{L}(s)_{ij} &= p_j F_i(s) && \text{for} \quad j \neq i; \qquad \text{and} \\ &= p_\emptyset + \sum_{j \neq i} p_j F_j(s) && \text{for} \quad j = i. \end{aligned} \tag{6.2}$$

This is clearly a positive operator for $s \geq 0$, as all entries in the matrix are nonnegative. Moreover, since the functions $F_i(s)$ are nondecreasing in s, the norms and spectral radii of the operators $\mathcal{L}(s)$ are monotone in s.

The key to the spectral analysis of the Jacobian operator is that it decomposes as the sum of a compact operator and a scalar multiple of the identity. To see this, observe that

$$\mathcal{L}(s) = \mathcal{K}(s) + \mathcal{D}(s) + H(s)I \tag{6.3}$$

where

$$\begin{aligned} \mathcal{K}(s)_{ij} &= p_j F_i(s), \\ \mathcal{D}(s)_{ij} &= -2p_i F_i(s)\delta_{ij}, \quad \text{and} \\ H(s) &= p_\emptyset + \sum_j p_j F_j(s). \end{aligned} \tag{6.4}$$

Observe that by equation (2.8),

$$H(z) = (G(z) - 1)/zG(z) \tag{6.5}$$

where $G(z)$ is the Green's function; note for future reference that $(G(s) - 1)/sG(s)$ is nondecreasing in s for $0 \leq s \leq R$, and that

$$H(R) < 1/R \tag{6.6}$$

by Corollary 2.2.

Lemma 6.1. *Let R be the dcommon radius of convergence of the Green's function $G(z)$ and the first-passage generating functions $F_i(z)$. For each $s \in [0, R]$, the operators $\mathcal{K}(s)$ and $\mathcal{D}(s)$ are compact.*

Proof. The operator $\mathcal{K}(s)$ has rank one and so is trivially compact. The operator $\mathcal{D}(s)$ is diagonal, with diagonal entries $-2p_i F_i(s)$, and so it maps the unit ball of B (the sequences with entries bounded in absolute value 1) into the set of sequences with entries bounded by $2p_i F_i(R)$. This set is compact, since the values $F_i(R)$ are bounded and $\sum_i p_i < \infty$. □

Corollary 6.2. *For each $s \in [0, R]$, the operator $\mathcal{L}(s)$ and its adjoint $\mathcal{L}(s)^*$ have purely discrete spectrum. Each element of the spectrum not equal to $H(s)$ is an eigenvalue of finite multiplicity (the same for both $\mathcal{L}(s)$ and $\mathcal{L}(s)^*$), and $H(s)$ is the only possible accumulation point of the spectrum. The spectral radius of $\mathcal{L}(s)$ varies continuously with s for $0 \leq s \leq R$. There is a unique $\sigma \in (0, R]$ such that the spectral radius of $\sigma\mathcal{L}(\sigma)$ is 1, and at least for $s > \sigma - \varepsilon$, for some $\varepsilon > 0$, the spectral radius of $\mathcal{L}(s)$ is an eigenvalue of finite multiplicity.*

Note: Later it will be shown that $\sigma = R$.

Proof. The Riesz-Schauder theory of compact operators (see, for example, [14], Ch. 4) asserts that the spectrum of a compact operator consists of at most countably many isolated eigenvalues of finite multiplicity (the same multiplicity for both the operator and its adjoint) that accumulate only at 0. (For a compact operator on an infinite-dimensional Banach space, 0 must also be an element of the spectrum). Since addition of a scalar multiple of the identity to an operator merely shifts its spectrum, the statements about the spectrum of a particular $\mathcal{L}(s)$ follow.

To see that the spectral radius $\rho(s)$ of $\mathcal{L}(s)$ must vary continuously with s, observe that either $\rho(s) = H(s)$ or $\rho(s)$ is an isolated eigenvalue of finite multiplicity. Isolated eigenvalues of finite multiplicity must vary continuously ([14], Chapter 6), but so does the Green's function; hence, the spectral radius is continuous in s. The spectral radius of $s\mathcal{L}(s)$ cannot remain bounded away from 1 for $s \in [0, R]$, because if so then by the

Implicit Function Theorem, the function $F(z)$ would have an analytic continuation to a neighborhood of $z = R$. Thus, by the Intermediate Value Theorem of calculus, for some $\sigma \in [0, R]$ the spectral radius of $\sigma\mathcal{L}(\sigma)$ attains the value 1.

Recall now that the function $H(s)$ is nondecreasing and is bounded above by $1/R$, by (6.6). Consequently, when the spectral radius of $s\mathcal{L}(s)$ attains (or exceeds) the value 1 it must coincide with an eigenvalue of finite multiplicity. Since the spectral radius of $s\mathcal{L}(s)$ is nondecreasing in s, the last assertion of the corollary follows. □

6.3. **Simplicity of the Lead Eigenvalue.**

Proposition 6.3. *The eigenvalue* 1 *of the operator* $T = \sigma\mathcal{L}(\sigma)$ *has multiplicity one. Furthermore, there are* strictly *positive right and left eigenvectors h and ν.*

The proof will consist of several lemmas. By Corollary 6.2, 1 is an eigenvalue of T with finite multiplicity, and no elements of the spectrum have modulus greater than one. *A priori* it is possible that 1 is not the only element of the spectrum on the unit circle; however, Corollary 6.2 implies that the spectrum of T has only finitely many points on the unit circle, and all are eigenvalues of finite multiplicity.

Lemma 6.4. *There exist positive right and left eigenvectors h and ν for T with eigenvalue* 1.

Proof. Because the operator T is positive, if $T\varphi = \varphi$ then $T|\varphi| \geq |\varphi|$. Since V contains nonzero elements, it follows that there is a nonnegative vector g such that $Tg \geq g$. By the positivity of T, the sequence $g_n := T^n g$ is nondecreasing, and so in particular each g_n is positive. If it could be shown that the sequence g_n converges in norm, then the limit h would be a positive eigenvector of T with eigenvalue 1.

Claim: The sequence g_n converges in $B-$norm.

Proof. Let $\lambda_1, \lambda_2, \ldots, \lambda_r$ be the eigenvalues of modulus 1, with $\lambda_1 = 1$, and let $V_1, V_2, \ldots, V_r$ be the corresponding eigenspaces. Define projections $P_1, P_2, \ldots, P_r$ by

$$P_j = \frac{1}{2\pi i} \oint_{\gamma_j} (\zeta I - T)^{-1} \, d\zeta,$$

where γ_j is a circle surrounding λ_j that encircles no other points of the spectrum. Each P_j is a projection whose range is the eigenspace V_j (see [14], Theorem 6.17). The projections P_j commute with each other and with T (since they are linear combinations of powers of T). The range U of complementary projection

$$P_U = I - P_1 - P_2 - \cdots - P_r$$

is an invariant subspace of T, and the spectral radius of TP_U is strictly less than one. Now

$$\begin{aligned} T^n g &= \sum_{j=1}^{r} T^n P_j g + T^n P_U g \\ &= \sum_{j=1}^{r} \lambda^n P_j g + T^n P_U g. \end{aligned}$$

Since each λ_j has absolute value 1, any increasing subsequence of the integers has a subsequence along which λ_j^n converges; consequently, along any such subsequence, $T^n g$ converges in norm. The limit is of necessity a positive element of B. Finally, there can be only one possible limit, because since $g_n \leq g_{n+1}$ for all n, if there were two subsequential limits ψ_A and ψ_B then $\psi_A \leq \psi_B$ and $\psi_B \leq \psi_A$. □

A similar argument shows that there is a positive left eigenvector. □

The following lemma will imply that positive eigenvectors of T and T^* are *strictly* positive, and this in turn will imply that the lead eigenspaces are one-dimensional.

Lemma 6.5. *Let $\mathcal{M}_m$ be the finite section of $\mathcal{L}(s)$ indexed by pairs i, j such that $1 \leq i, j \leq m$. Then for each $m = 3, 4, \dots$ and every $s \in (0, R]$, the matrix $\mathcal{M} = \mathcal{M}_m$ is a Perron-Frobenius matrix, that is, some positive power of $\mathcal{M}$ has strictly positive entries.*

Proof. By our standing hypotheses on the transition probabilities of the random walk, $p_i > 0$ for each index i and $p_\emptyset > 0$, and so all of the entries of $\mathcal{M}_m$ are positive. □

Corollary 6.6. *Positive eigenvectors of T and T^* are* strictly *positive.*

Proof. Suppose that $Th = h$ is a positive eigenvector. Since h is not identically 0, some entry, say the $(j; y)$ entry, is strictly positive. By Lemma 6.5, for each $m \geq 1$ some positive power $\mathcal{M}_m^n$ of the finite section $\mathcal{M}_m$ of T has all entries positive; consequently, if $m \geq j$ then all entries of $T^n h$ indexed by $(i; x)$, where $i \leq m$, are positive. Since $m \geq j$ is arbitrary, it follows that h is strictly positive. A similar argument shows that positive eigenvectors of T^* must be strictly positive. □

Proof of Proposition 6.3. Denote by V and V^* the 1−eigenspaces of the operators T and T^*. By Corollary 6.2, V and V^* are finite-dimensional subspaces of B and B^* of the same dimension. By Corollary 6.6, there are strictly positive right and left eigenvectors $h \in V$ and $\nu \in V^*$.

Claim 1. $\varphi, \psi \in V \implies |\varphi \vee \psi| \in V$.

Proof. Since the operator T is positive,

$$T|\varphi \vee \psi| \geq |T\varphi| \vee |T\psi| = |\varphi| \vee |\psi|. \tag{6.7}$$

To see that the inequality is actually an equality, apply the positive linear functional ν to both sides:

$$\langle \nu, |\varphi \vee \psi| \rangle = \langle T^* \nu, |\varphi \vee \psi| \rangle = \langle \nu, T|\varphi \vee \psi| \rangle \geq \langle \nu, |\varphi \vee \psi| \rangle.$$

Since ν is *strictly* positive, it must be that equality holds in (6.7). □

Claim 2. *Let $\varphi \in V$ be real-valued. Then φ is strictly positive, or strictly negative, or identically* 0.

Proof. Suppose first that there are indices $\alpha = (i; x)$ and $\beta = (j; y)$ such that $\varphi(\alpha) > 0 > \varphi(\beta)$. By the irreducibility of T (Lemma 6.5), there exists and integer $n \geq 1$ such that the

(α, β) entry of T^n is positive. Now since $\varphi \in V$ it is an eigenvector of T^n with eigenvalue 1; by Claim 1, the positive part of φ is also an eigenvector with eigenvalue 1. Thus,

$$\varphi(\alpha) = \sum_{\gamma \in J} T^n_{\alpha,\gamma}\varphi(\gamma) \qquad \text{and}$$
$$\varphi(\alpha) = \sum_{\gamma \in J_+} T^n_{\alpha,\gamma}\varphi(\gamma)$$

where J_+ denotes the set of indices γ such that $\varphi(\gamma) \geq 0$ and J the set of *all* indices γ. The two sums can be equal only if $\sum_J$ contains no negative terms. But n was chosen so that there would be at least one negative term, namely, the term $\gamma = \beta$. This proves that either $\varphi \geq 0$ or $\varphi \leq 0$.

A similar argument shows that if $\varphi \in V$ is nonnegative, then either $\varphi \equiv 0$ or φ is strictly positive. □

Let $\varphi \in V$ be any right eigenvector of T with eigenvalue 1; I will show that φ is a scalar multiple of h. Without loss of generality, assume that φ is real-valued (because if $\varphi \in V$ then its real and imaginary parts are both in V, as T is positive) and positive (by Claim 2). Define $a \geq 0$ to be the supremum of all nonnegative real numbers b such that $b\varphi \leq h$. If the vector $a\varphi$ is not identically equal to h then there is an index $\alpha = (i; x)$ such that $a\varphi(\alpha) < h(\alpha)$, and so for $b > a$ sufficiently near a it must be the case that $b\varphi(\alpha) < h(\alpha)$. But then $h - b\varphi$ is a real-valued element of V with both negative and positive entries, contradicting Claim 2.

6.4. **Singularity of F at $z = \sigma$.**

Lemma 6.7. *Let $\mathcal{H}$ be the Hessian of the link mapping Q associated with the Lagrangian system (2.9), and let h and ν be the positive eigenvectors of $T = \sigma\mathcal{L}(\sigma)$ and T^* with eigenvalue* 1*. Then*

$$\langle \nu, \mathcal{H}(h,h)\rangle > 0. \tag{6.8}$$

Proof. The Hessian operator is the matrix of second partial derivatives of the link mapping Q. The first partial derivatives are given in equations (6.2); from these equations it is apparent that the nonzero second partials are

$$\frac{\partial^2 Q_i}{\partial F_i \partial F_j} = p_j \qquad \text{for} \quad j \neq i \tag{6.9}$$

Hence,

$$\mathcal{H}(h,h)_{i;x} = \sum_{j \neq i} h_i h_j p_j. \tag{6.10}$$

Since the eigenvector h has strictly positive entries, this sum is strictly positive for all i. Since ν is positive, (6.8) follows. □

Corollary 6.8. *For every positive $\mu \in B^*$, the function $\langle \mu, F(z)\rangle$ has a square-root singularity at $z = \sigma$. Consequently,*

$$\sigma = R. \tag{6.11}$$

Proof. Proposition 6.3 and Lemma 6.7 imply that the hypotheses of Theorem 5.5 are satisfied. Consequently, the function $\langle \nu, F(z)\rangle$ has a square-root singularity at $s = \sigma$. But the common radius of convergence of the power series $F_i(z)$ is R; therefore, it must be that $\sigma = R$. Consequently, For any positive $\mu \in B^*$,

$$\langle \mu, F(z)\rangle = \langle \mu, h\rangle\langle \nu, F(z)\rangle + \langle \mu, P_W F(z)\rangle.$$

By Theorem 5.5, the projection $P_W F(z)$ has a Puiseux series around $z = R$ whose first term is linear in $R-z$, and by Proposition 6.3, the lead eigenvectors ν and h are *strictly* positive. Thus, $\langle \mu, h\rangle > 0$, and so the Puiseux expansion of $\langle \mu, F(z)\rangle$ has as its first nonconstant term a nonzero multiple of $\sqrt{R-z}$. □

6.5. **Singularity of the Green's function at** $z = R$**.**

Proposition 6.9. *Let $G(z)$ be the Green's function of an aperiodic, irreducible, quasi-nearest-neighbor random walk on a countable free product of finite groups, and let R be its radius of convergence. Then $G(z)$ has a square-root singularity at $z = R$, that is, it has a Puiseux expansion in a neighborhood of $z = R$ of the form*

$$G(z) - G(R) = \sum_{n=2^k}^{\infty} g_n (R-z)^{n/2^{k+1}} \tag{6.12}$$

whose first nonconstant term is $g_{2^k}\sqrt{R-z}$, with $g_{2^k} < 0$.

Proof. The relation (2.8) exhibits G as a meromorphic function of a positive linear combination $\mathcal{F}$ of the first-passage generating functions F_i. By Proposition 2.1, $G(z)$ is finite at $z = R$, and so near $z = R$ the function $G(z)$ is in fact an *analytic* function of $\mathcal{F}(z)$. By Corollary 6.8, every positive linear functional of $F(z)$ has a square-root singularity at $z = R$; consequently, the same is true of G. □

6.6. **Local Limit Theorem.** By Proposition 2.3 the Green's function $G(z)$ has no singularity on the circle of convergence except at $z = R$, and by the preceding proposition at $z = R$ the function $G(z)$ admits the expansion (6.12). Therefore, the Flajolet-Odlyzko Transfer Theorem implies the following Local Limit Theorem.

Theorem 6.10. *For nearest-neighbor random walk on a countable free product of copies of $\mathbb{Z}_2$ with positive holding probability $p_\emptyset$, the return probabilities satisfy*

$$P\{X_n = \emptyset\} \sim \frac{C}{R^n n^{3/2}}. \tag{6.13}$$

6.7. **Remarks.** The crux of the preceding argument is the analysis of section 6.2, which establishes that the Jacobian operator $\mathcal{L}(s)$ has discrete spectrum. Only in this part of the argument does the detailed structure of the Lagrangian system (2.9) play a significant role, and only the decomposition (6.3) of the Jacobian as a sum of a compact operator and a multiple of the identity is needed. It is not difficult to show that a similar decomposition holds for the corresponding Lagrangian systems of non-nearest-neighbor random walks, and of random walks on free products of arbitrary finite groups. Other applications of the results of section 5 will be published elsewhere.

References

[1] D. Cartwright (1991) Singularities of the Green's function of a random walk on a discrete group.
[2] D. Cartwright & P. M. Soardi (1986) Random walks on free products, quotients and amalgams. *Nagoya Math. J.* **102** 163–180.
[3] P. de la Harpe (2000) *Topics in Geometric Group Theory.* University of Chicago Press, Chicago, IL.
[4] M. Drmota (1997) Systems of functional equations. *Random Structures and Algorithms* **10** 103–124.
[5] A. Erdelyi (1956) *Asymptotic expansions.* Dover Publications, NY
[6] W. Feller (1965) *An Introduction to Probability Theory and its Applications*, volume 2. Wiley, NY
[7] P. Flajolet & A. Odlyzko (1990) Singularity analysis of generating functions. *SIAM J. Discrete Math.* **3** 216–240.
[8] P. Gerl & W. Woess (1986) Local limits and harmonic functions for nonisotropic random walks on free groups. *Probability Theory & Reated Fields*, **71**, 341–355.
[9] I. J. Good (1960) Generalizations to several variables of Lagrange's expansion, with applications to stochastic processes. *Proc. Cambridge Philos. Soc.* **56**, 367–380.
[10] I. P. Goulden & D. M. Jackson (1983) *Combinatorial Enumeration.* Wiley & Sons, NY.
[11] P. Griffiths & J. Harris (1978) *Principles of Algebraic Geometry.* Wiley, NY.
[12] E. Hille (1959) *Analytic Functions*, volumes 1–2. Ginn, Boston.
[13] J. Hopcraft & J. Ullman (1979) *Introduction to Automata Theory, Languages, and Computation.* Addison-Wesley, Reading MA
[14] T. Kato (1966) *Perturbation Theory for Linear Operators.* Springer, NY
[15] H. Kesten (1959) Full Banach mean values on countable groups. *Math. Scand.* **7**, 146–156.
[16] S. Lalley (1993) Finite range random walk on free groups and homogeneous trees. *Annals of Probability* **21** 2087–2130.
[17] S. Lalley (1995) Return probabilities for random walk on a half-line. *J. Theoretical Probability* **8** 571–599.
[18] S. Lalley (2001) Random walks on regular languages and algebraic systems of generating functions. *Algebraic Methods in Statistics and Probability (Notre Dame, IN, 2000).* Amer. Math. Soc., Providence, RI
[19] J. Lehner (1964) *Discontinuous Groups and Automorphic Forms.* American Math. Soc., Providence RI
[20] M. Lothaire *et al* (1983) *Combinatorics on words.* Addison-Wesley, Reading MA
[21] M. Reed & B. Simon (1978) *Methods of Mathematical Physics IV: Analysis of Operators.* Academic Press, NY
[22] E. Seneta (1981) *Nonnegative Matrices and Markov Chains*, 2nd ed., Springer, NY.
[23] R. Stanley (1999) *Enumerative combinatorics*, volume 2. Cambridge University Press, Cambridge.
[24] H. Wilf (1990) *generatingfunctionology.* Academic Press, Boston, MA.
[25] W. Woess (2000) *Random walks on infinite graphs and groups.* Cambridge University Press, Cambridge
[26] W. Woess (1982) A local limit theorem for random walks on certain discrete groups. *Probability measures on groups (Oberwolfach, 1981)*, pp. 467–477, Lecture Notes in Math. **928**, Springer, Berlin-New York.
[27] W. Woess (1987) Context-free languages and random walks on groups. *Discrete Math.* **67** 81–87.
[28] E. Zeidler (1986) *Nonlinear Functional Analysis and its Applications. I: Fixed Point Theorems* Springer, NY

University of Chicago, Department of Statistics, USA
E-mail address: `lalley@galton.uchicago.edu`