An asymptotic formula for integer points on Markoff-Hurwitz surfaces

Alex Gamburd, Michael Magee, and Ryan Ronan

September 8, 2017

Abstract

We establish an asymptotic formula for the number of integer solutions to the Markoff-Hurwitz equation

\[ x_1^2 + x_2^2 + \ldots + x_n^2 = ax_1x_2 \ldots x_n + k. \]

When \( n \geq 4 \) the previous best result is by Baragar (1998) that gives an exponential rate of growth with exponent \( \beta \) that is not in general an integer when \( n \geq 4 \). We give a new interpretation of this exponent of growth in terms of the unique parameter for which there exists a certain conformal measure on projective space.

Contents

1 Introduction 2
1.1 Connection to simple closed curves and character varieties . . . . . . . . . . . . 7
1.2 Paper organization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9
1.3 Notation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9

2 Markoff-Hurwitz tuples and moves 10
2.1 Basic properties of the Markoff-Hurwitz equation . . . . . . . . . . . . . . . . . 10
2.2 The polynomial semigroup . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15
2.3 Multiplicities . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
2.4 Increasing the size of \( K \) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18

3 Converting the linear count to the nonlinear count 19
3.1 Acceleration . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
3.2 The renewal equation for \( M \) . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
3.3 Iteration . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 22
3.4 Comparison to the linear count . . . . . . . . . . . . . . . . . . . . . . . . . . . . 26
3.5 Using the linear semigroup count to prove Proposition 19 . . . . . . . . . . . . 29

∗A. Gamburd was supported in part by NSF award DMS-1603715.
†M. Magee was supported in part by NSF award DMS-1701357.
4 The linear semigroup count
4.1 Renewal (again) ................................................................. 33
4.2 Spectral theory of the transfer operator .......................... 35
4.3 Proofs of Theorem 9 and 11 given the spectral theorems .... 39
4.4 Consequences of uniformly contracting dynamics .......... 40
4.5 Behavior of the eigenvalue .............................................. 41

5 Proof of uniform contraction .......................... 42
5.1 Setup ................................................................. 42
5.2 Overview of the proof of Proposition 41 .................... 44
5.3 Proof of equations (5.1)-(5.10) ................................. 45

1 Introduction
For integer parameters \( n \geq 3, a \geq 1, \) and \( k \in \mathbb{Z} \) consider the Diophantine equation
\[
x_1^2 + x_2^2 + \ldots + x_n^2 = ax_1x_2\ldots x_n + k. \tag{1.1}
\]
We call this the generalized\(^1\) Markoff-Hurwitz equation. In this paper we count solutions to (1.1) in integers, which we call Markoff-Hurwitz tuples. More precisely, let \( V \) be the affine subvariety of \( \mathbb{C}^n \) cut out by (1.1). We are interested in the asymptotic size of the set
\[
V(\mathbb{Z}) \cap B(R)
\]
where \( B(R) \) is the ball of radius \( R \) in the \( \ell^\infty \) norm on \( \mathbb{R}^n \subset \mathbb{C}^n \).

When \( n = 3, a = 3 \) and \( k = 0 \) solutions to (1.1) in positive integers are called Markoff triples, and the numbers that appear therein are called Markoff numbers\(^2\). The Markoff numbers are intimately connected with Diophantine properties of the rationals via the Markoff spectrum [Bom07], and also with hyperbolic geometry and free groups [Aig13].

The question of counting \( |V(\mathbb{Z}) \cap B(R)| \) for Markoff triples was first investigated in the thesis of Gurwood [Gur76] who established an asymptotic formula using the correspondence between Markoff and Farey trees. An improved error term was obtained by Zagier in [Zag82, pg. 711], and a very clean proof of a slightly weaker result can be found in Belyi [Bel01]. The current best result is due to McShane and Rivin [MR95]:

**Theorem 1 (McShane-Rivin).** The number \( M(R) \) of Markoff triples \( (x, y, z) \) with \( x \leq y \leq z \leq R \) is given by
\[
M(R) = C(\log R)^2 + O(\log R \log \log R)
\]
as \( R \to \infty \), with \( C > 0 \).

Perhaps somewhat surprisingly, the asymptotic growth for \( n \geq 4 \) is not of the order \( (\log R)^{n-1} \), as was first noticed by Baragar [Bar94a], who obtained the following result

\(^1\) Normally \( k = 0 \) is considered.
\(^2\) A long standing conjecture of Frobenius asserts that each Markoff number appears as the maximal entry of only one triple, up to reordering. If one assumes this conjecture, then the problems of counting Markoff triples and numbers are the same.
Theorem 2 (Baragar). There is a number $\beta = \beta(n)$ such that when $k = 0$, if $V(\mathbb{Z}) - \{(0, 0, 0)\}$ is nonempty then for every $\varepsilon > 0$

$$\Omega((\log R)^{\beta(n)-\varepsilon}) \leq |V(\mathbb{Z}) \cap B(R)| \leq O((\log R)^{\beta(n)+\varepsilon}). \quad (1.2)$$

This was strengthened by Baragar in [Bar98] under the same hypotheses to

$$|V(\mathbb{Z}) \cap B(R)| = (\log R)^{\beta + o(1)}. \quad (1.3)$$

In [Bar98] the following bounds for the exponents $\beta(n)$ were also obtained

$$\beta(3) = 2,$$
$$\beta(4) \in (2.430, 2.477),$$
$$\beta(5) \in (2.730, 2.798),$$
$$\beta(6) \in (2.963, 3.048),$$

and in general

$$\frac{\log(n - 1)}{\log 2} < \beta(n) < \frac{\log(n - 1)}{\log 2} + o(n^{-0.58}).$$

In 1995 [Sil95], it was asked by Silverman whether in the setting of $k = 0$

1. there is a true asymptotic formula for $|V(\mathbb{Z}) \cap B(R)|$ with main term proportional to $\log(R)\beta$, and
2. furthermore, $\beta(n)$ is irrational?

The irrationality of $\beta$ remains a tantalizing open question and one may wonder whether it is even algebraic. On the other hand, our methods do give some further insight into the nature of this mysterious number (cf. Theorem 9 below). The main goal of this paper is to extend Baragar’s exponential rate of growth estimate to a true asymptotic formula$^3$.

When $k > 0$ there are certain exceptional families of solutions to (1.1) that have a different quality of growth. We describe these families in Definition 13 and for fixed $k, a, n$ we write $\mathcal{E}$ for the set of exceptional tuples. We obtain the following theorem for the asymptotic number of Markoff-Hurwitz tuples.

Theorem 3. For each $(n, a, k)$ with $V(\mathbb{Z}) - \mathcal{E}$ infinite, there is a positive constant $c = c(n, a, k)$ such that

$$|(V(\mathbb{Z}) - \mathcal{E}) \cap B(R)| = c(\log R)^\beta + o((\log R)^\beta).$$

Here $\beta$ is the same constant as Theorem 2.

Remark 4. We explain in Section 2.1 that removing $\mathcal{E}$ is necessary in Theorem 3 since the exceptional families have $|\mathcal{E} \cap B(R)| \geq cR$, $c > 0$ for $R \geq R_0(n, a, k)$ when they are non-empty. On the other hand, $\mathcal{E}$ is non-empty only when $k - n + 2$ or $k - n - 1$ is a square.

$^3$The techniques in [Bar98] “were inspired in part by Boyd’s work on the Apollonian packing problem [Bov71, Boy73, Boy82].” Boyd’s result was extended to a true asymptotic formula in the work of Kontorovich and Oh [KO11].
Remark 5. The issue of the existence and infinitude of integral solutions for general \(a, k\), even for \(n = 3\), is quite subtle: see [Mor53, SM57]. In recent work of Ghosh and Sarnak [GS17], the Hasse principle is established to hold for Markoff-type cubic surfaces \(x_1^2 + x_2^2 + x_3^2 - x_1 x_2 x_3 = k\) for almost all \(k\).

Remark 6. In Theorem 9 we give a new characterization of \(\beta\) as the unique parameter for which there exists a conformal measure for the action of a linear semigroup on projective space.

Our counting arguments, as in [Zag82] and [Bar94a, Bar98], depend on an infinite descent for solutions to (1.1) that goes back to Markoff [Mar80] in the case of Markoff triples and Hurwitz [Hur07] in the higher dimensional setting of \(n > 3, k = 0\). In Section 2.1 we explain how the counting problem for \(V(\mathbb{Z})\) can be related to the analogous one for \(V(\mathbb{Z}_+)\), where \(\mathbb{Z}_+\) are the positive integers.

Given \(x \in V(\mathbb{Z}_+)\), fixing all of the coordinates of \(x\) except \(x_j\) and viewing (1.1) as a quadratic polynomial in \(x_j\), the other root is given by

\[
x'_j = a \prod_{i \neq j} x_i - x_j.
\]

Therefore for each \(j\) one has the **Markoff-Hurwitz move**

\[
m_j(x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, a \prod_{i \neq j} x_i - x_j, \ldots, x_n)
\]

that yields a new solution to (1.1). Infinite descent for the Markoff-Hurwitz equation says that any unexceptional tuple in \(V(\mathbb{Z}_+)\) can be reduced to one in a compact set \(K_0 = K_0(n, a, k)\) by a series of Markoff-Hurwitz moves (cf. Corollary 15).

After renormalizing (1.1), the Markoff-Hurwitz moves \(\{m_j\}\) induce the moves

\[
\lambda_j(z_1, \ldots, z_n) = \left( z_1, \ldots, \hat{z}_j, \ldots, z_n, \prod_{i \neq j} z_i - z_j \right), \quad 1 \leq j \leq n - 1,
\]

on ordered tuples, where \(\hat{\cdot}\) denotes omission. If enough of the \(z_i\) are large, the move \(\lambda_j\) can be approximated by

\[
z \mapsto \left( z_1, \ldots, \hat{z}_j, \ldots, z_n, \prod_{i \neq j} z_i \right)
\]

to high accuracy relative to the largest entries of \(z\). At the level of logarithms this corresponds to

\[
(\log z_1, \log z_2, \ldots, \log z_n) \mapsto \left( \log z_1, \ldots, \hat{\log z}_j, \ldots, \log z_n, \sum_{i \neq j} \log z_i \right).
\]

Thus one is naturally led to study the linear semigroup generated by linear maps

\[
\gamma_j(y_1, y_2, \ldots, y_n) = \left( y_1, \ldots, \hat{y}_j, \ldots, y_n, \sum_{i \neq j} y_i \right) \quad (1.4)
\]
on ordered \( n \)-tuples \((y_1, \ldots, y_n)\). Indeed, this is the approach of Zagier [Zag82] in the setting of Markoff triples and Baragar [Bar94a] for general \( n, a \) with \( k = 0 \). Let
\[
\Gamma = \langle \gamma_1, \ldots, \gamma_{n-1} \rangle
\]
where we have written a ‘+’ to indicate we are generating a semigroup, not a group.

The crucial new idea in this work that explains why we are able to make progress on the counting problem is that we replace\(^4\) the generators of \( \Gamma \) with the countably infinite generating set
\[
T_\Gamma = \{ \gamma_n^A \gamma_j : A \in \mathbb{Z}_{\geq 0}, 1 \leq j \leq n-2 \}
\]
and then consider the semigroup
\[
\Gamma' = \langle T_\Gamma \rangle_+.
\]
Both \( \Gamma \) and \( \Gamma' \) are freely generated by their respective generating sets. Notice that \( \Gamma \) and \( \Gamma' \) preserve the nonnegative ordered hyperplane
\[
\mathcal{H} \equiv \left\{ (y_1, \ldots, y_n) \in \mathbb{R}_{\geq 0}^n : y_1 \leq y_2 \leq \cdots \leq y_n, \sum_{j=1}^{n-1} y_j = y_n \right\} \subset \mathbb{R}_{\geq 0}^n \quad (1.5)
\]
and that any element of \( \Gamma \) maps ordered tuples in \( \mathbb{R}_{\geq 0}^n \) into \( \mathcal{H} \). Therefore the study of orbits of \( \Gamma \) and \( \Gamma' \) on ordered tuples boils down to the study of orbits in \( \mathcal{H} \).

**Example 7.** When \( n = 3 \), the linear map \( \sigma : \mathcal{H} \to \mathcal{H} \) defined by
\[
\sigma(a, b, a + b) = \text{order} (b - a, a, b), \quad (1.6)
\]
where \textit{order} puts a tuple in ascending order from left to right, is such that for \( j = 1, 2 \) we have
\[
\sigma \gamma_j y = y
\]
for all \( y \in \mathcal{H} \). Repeatedly applying the map \( \sigma \) to a triple \((a, b, a + b)\) with \( a \leq b \in \mathbb{Z} \) performs the Euclidean algorithm on \( a, b \). However, one application of \( \sigma \) corresponds in general to less than one step of the algorithm. Replacing \( \Gamma \) with \( \Gamma' \) corresponds to speeding this up so one whole step of the Euclidean algorithm corresponds to one semigroup generator. As for counting, the orbit of \((0,1,1)\) under \( \Gamma \) is precisely those \((a, b, a + b)\) with \((a, b) = 1\) and thus can be counted by elementary methods. This is exploited in Zagier’s paper [Zag82].

We can use the basis
\[
e_j = (0, \ldots, 0, 1_j, 0, \ldots, 0, 1)
\]
for the subspace spanned by \( \mathcal{H} \). This basis clarifies the action of \( \Gamma' \).

\(^4\)See our discussion in Section 3.1 about the benefits of this replacement. It is inspired by the ‘Time Acceleration Machine’ described by Zorich in [Zor06, Section 5.3].

\(^5\)This follows from a similar argument to the proof of Lemma 16 we give below.
Figure 1: When $n = 4$, the semigroup elements map $\Delta = \mathcal{H}/\mathbb{R}_+$ into a strictly smaller subset. After iteration this leads to more and more empty space (see also Figure 2). This doesn’t occur when $n = 3$, as one can also see from the picture: the action of the group elements $\gamma_2$ and $\gamma_3$ on the vertical coordinate axis is a copy of the $n = 3$ dynamics.

**Example 8.** When $n = 3$ the semigroup $\Gamma'$ is generated by the

$$g_A := \gamma_2^A \gamma_1 = \begin{pmatrix} 0 & 1 \\ 1 & A + 1 \end{pmatrix}$$

with respect to the basis $\{e_1, e_2\}$. These generators are classically connected with continued fractions by the formulae

$$\begin{pmatrix} 0 & 1 \\ 1 & A_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & A_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & A_k \end{pmatrix} = \begin{pmatrix} \ast & b \\ \ast & d \end{pmatrix}, \quad \frac{b}{d} = \frac{1}{A_1 + \frac{1}{A_2 + \cdots + \frac{1}{A_k}}}.$$

So our semigroups $\Gamma$ and $\Gamma'$ are natural extensions of the Euclidean algorithm and continued fractions semigroup to higher dimensions. We write $\Delta = \mathcal{H}/\mathbb{R}_+$ and we can view $\Delta$ as a subset of $\mathbb{R}^{n-2}$ (see Section 5 for details). The key distinction that appears when $n \geq 4$ is that

$$\Delta \neq \bigcup_{j=1}^{n-1} \gamma_j(\Delta)$$

and so the induced dynamics on $\mathcal{H}/\mathbb{R}_+$ has ‘holes’ as we illustrate in Figure 1.

We get a new characterization of the parameter $\beta$ in terms of the action of $\Gamma'$ on $\mathcal{H}/\mathbb{R}_+$.

---

6See [Zor06] for the discussion of such an extension in the context of translation surfaces.
Theorem 9. The $\beta$ from Theorem 2 is the unique parameter in $(1, \infty)$ such that there exists a probability measure $\nu_\beta$ on $\Delta = \mathcal{H}/\mathbb{R}_+$ with the property

$$
\int_{w \in \Delta} f(w) \, d\nu_\beta(w) = \sum_{\gamma \in T} \int_{w \in \Delta} f(\gamma.w) |\text{Jac}_w(\gamma)|^{\frac{\beta}{n-1}} \, d\nu_\beta(w)
$$

for all $f \in C^0(\Delta)$. We call $\nu_\beta$ a conformal measure.

Remark 10. Theorem 9 can be viewed as a partial analog of the connection between the exponent of growth of a finitely generated Fuchsian group and the Hausdorff dimension of its limit set as a result of Patterson-Sullivan theory [Pat76, Sul79, Sul84]. In our setting, the lack of any symmetric space means the parameter $\beta$ is not in any obvious way connected to the Hausdorff dimension of the compact $\Gamma'$-invariant subset of $\Delta$.

In Section 3.4 we reduce Theorem 3 to a counting theorem for orbits of the semigroup $\Gamma'$. The relevant counting quantity is defined by

$$
N(y, a) \equiv \sum_{\gamma \in \Gamma' \cup \{e\}} 1\{\log(\gamma.y)_n - \log(y)_n \leq a\} \tag{1.7}
$$

for $y \in \mathcal{H} - 0$ and $a \geq 0$. We prove

Theorem 11. There is a positive bounded $C^1$ function $h$ on $\mathcal{H}$ that is invariant under the action of $\mathbb{R}_+$ and such that

$$
N(y, a) = h(y)e^{\beta a}(1 + o_{a \to \infty}(1))
$$

for all $y \in \mathcal{H} - 0$, where the implied function in the small $o$ does not depend on $y$. Moreover, $h$ satisfies the recursion

$$
\sum_{\gamma \in T_F} \left(\frac{\gamma.y}_n\right)^{-\beta} h(\gamma.y) = h(y). \tag{1.8}
$$

The constant $\beta$ is the same as in Theorem 2.

Remark 12. The embedding of the $(n-1)$-dimensional version of $\mathcal{H}$ inside the $n$-dimensional version implies by Theorem 11 that $\beta(n) \geq \beta(n-1)$ and in particular that $\beta(n) \geq 2$ for all $n \geq 3$.

1.1 Connection to simple closed curves and character varieties

Theorem 1 can be rephrased as a counting result for the number of simple\footnote{This means there are no self crossings.} closed geodesics of length $\leq \log R$ on the modular torus. This is the topological once-punctured torus that is uniformized by the quotient of the hyperbolic plane by the group

$$
\left\langle \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \right\rangle \leq \text{PSL}_2(\mathbb{R}).
$$

McShane and Rivin [MR95] actually obtain the analogous counting result to Theorem 1 for simple closed geodesics on arbitrary hyperbolic once punctured tori, by use of a special norm
Figure 2: In the same setting \((n = 4)\) of Figure 1, we show in black the images of \(\Delta\) under the action of all words of length 10 in the generators \(\{\gamma_1, \gamma_2, \gamma_3\}\).
on the first homology of the surface. Mirzakhani proved in [Mir08] an asymptotic counting result, without explicit error term, for simple closed geodesics on any finite area complete Riemann surface. These asymptotics have recently been extended by Mirzakhani [Mir16] to more general orbits of the mapping class group. In Mirzakhani’s results the exponents of growth are dimensions of Teichmüller spaces. It is interesting to compare this to our characterization of Theorem 9.

In [HN13], Huang and Norbury showed that when \( n = a = 4 \) and \( k = 0 \), \( V(\mathbb{R}_+) \) is a parametrization of the Teichmüller space of finite area hyperbolic structures on \( \mathbb{R}P^2 \) minus three points, and moreover the coordinates of points on \( V(\mathbb{R}_+) \) are functions of the lengths of one-sided\(^8\) simple closed geodesics in the relevant hyperbolic structure. From these facts they deduce from Baragar’s Theorem 2 that the number \( n_j^{(1)}(L) \) of one sided simple closed geodesics of length \( \leq L \) in a hyperbolic structure \( J \) on \( \mathbb{R}P^2 \) minus three points satisfies

\[
\lim_{L \to \infty} \frac{\log n_j^{(1)}(L)}{\log L} = \beta(4).
\]

The second author (Magee) of this paper has recently shown [Mag17] that the methods here can be extended to prove that \( n_j^{(1)}(L) \) is asymptotic to \( cL^\beta \), for some \( c = c(J) > 0 \), somewhat in analogy to Mirzakhani’s results.

We also mention the recent work of Hu, Tan and Zhang [HPZ15] that describes some regions in \( \mathbb{C}^n \) where the group of automorphisms of (1.1) acts properly discontinuously. This extends previous work of Goldman [Gol03] that describes ranges of \( k \) in the case of \( n = 3 \) where the group \( \text{Aut}(V) \) act ergodically or properly discontinuously (or some combination thereof, on different components of the variety). Quite strikingly, for certain ranges of \( k \) the action of \( \text{Aut}(V) \) is ergodic on \( V(\mathbb{R}) \) yet preserves the infinite discrete subset \( V(\mathbb{Z}) \). In [HPZ15] the authors also prove a ‘McShane identity’ that gives a closed form expression for 1 in terms of an infinite sum over an orbit of the semigroup; see [McS91, McS98] for McShane’s original identity.

### 1.2 Paper organization

We prove our theorems in the order we have stated them with earlier parts of the paper depending on later parts. In Section 2 we describe the passage from \( V(\mathbb{Z}) \) to \( V(\mathbb{Z}_+) \) and describe in full the action of the Markoff-Hurwitz moves on \( V(\mathbb{Z}_+) \). At the end of Section 2 we have fixed a large compact region of \( V(\mathbb{Z}_+) \) outside of which the orbits of the action of Markoff-Hurwitz generators are a disjoint union of a finite number of orbits that we understand well. In Section 3 we fit the counting of these orbits to certain counts for the linear semigroup \( \Gamma' \). Using Theorem 11 as a black box, we prove Theorem 3. In Section 4, we prove Theorems 11 and 9 given Proposition 41 that says the action of \( \Gamma' \) on projective space is contracting. It is at this point we establish the connection with Baragar’s exponent of growth that we call \( \beta \). Finally, in Section 5 we prove Proposition 41.

### 1.3 Notation

For the reader’s convenience we describe the notation we use in this paper. We will use 1 for an indicator function. A vector with an entry \( \hat{\bullet} \) with a hat means that that entry is

\(^8\)This means a thickening of the geodesic is homeomorphic to a Möbius band.
omitted. We use Vinogradov notation $O, o, \ll, \gg$ in the standard way. Any implied constants may depend on $n, a, k$ that we view as fixed throughout much of the paper. If there is any dependence of an implied constant on a variable we denote this as a subscript e.g. $\ll_\epsilon$, and we also use subscripts to indicate which variable is tending to a limit, e.g. $o_{a \to \infty}$. For the sake of convenience, we take the liberty of applying functions to vectors, which means we apply the function component-wise, and we write inequalities between vectors to mean that the inequality holds at every component. For a set $S$ in a semigroup we may write $S^{(k)}$ for the $k$-fold product of the set with itself. We also write $\mathbb{R}_+, \mathbb{R}_{\geq 0}$ for the sets of positive (resp. nonnegative) real numbers, and similar for integers. We write $\{x\}$ for the fractional part of a real number $x$, that is, $x = n + \{x\}$ for $n \in \mathbb{Z}$ and $\{x\} \geq 0$.

Acknowledgements

We would like to thank Peter Sarnak, Giulio Tiozzo, and Peter Whang for helpful conversations about this work.

2 Markoff-Hurwitz tuples and moves

2.1 Basic properties of the Markoff-Hurwitz equation

The automorphism group

By an automorphism of $V$ we mean a polynomial automorphism of $V(\mathbb{C})$. We write $\text{Aut}(V)$ for the group of all such maps. By results of Horowitz \cite{Hor75} when $n = 3$ and Hu, Tan and Zhang \cite[Theorem 1.1]{HPZ15} for $n \geq 4$, one has

$$\text{Aut}(V) = \mathcal{G} \rtimes (N \rtimes S_n)$$

where

1. $N$ is the group of transformations that change the sign of an even number of variables. Hence $|N| = 2^{n-1}$.

2. $S_n$ is the symmetric group on $n$ letters that acts by permuting the coordinates of $\mathbb{C}^n$.

3. $\mathcal{G}$ is the nonlinear group generated by the Markoff-Hurwitz moves $m_j$ discussed in the Introduction.

One important corollary of this classification is that $V(\mathbb{Z})$ is invariant under $\text{Aut}(V)$.

Exceptional solutions

For $a = 1$ and $a = 2$ there are certain exceptional families of points in $V(\mathbb{Z})$ whose growth rate is totally different from the points we wish to count\textsuperscript{9}. These appear only for certain values of $k$ and we describe them now.

\textbf{Definition 13.} We say that $x \in V(\mathbb{Z}_+)$ is a \textit{fundamental exceptional solution} if it belongs to one of the following two families

\textsuperscript{9}See Silverman \cite{Sil89} for a discussion of a phenomenon of surfaces containing curves that have many more integral points than one would expect from the surface as a whole.
1. One has $a = 1$ and after reordering the coefficients of $x$ so that $x_1 \leq x_2 \leq \ldots \leq x_n$

$$x_1 = x_2 = \ldots = x_{n-3} = 1, \quad x_{n-2} = 2.$$ 

In this case $x$ is a Markoff-Hurwitz tuple if and only if

$$(x_{n-1} - x_n)^2 = k - n - 1. \quad (2.1)$$

2. One has $a = 2$ and after reordering the coefficients of $x$ so that $x_1 \leq x_2 \leq \ldots \leq x_n$

$$x_1 = x_2 = \ldots = x_{n-2} = 1.$$ 

In this case $x$ is a Markoff-Hurwitz tuple if and only if

$$(x_{n-1} - x_n)^2 = k - n + 2. \quad (2.2)$$

We say that $x \in V(\mathbb{Z})$ is an exceptional solution if $x$ is in the $\text{Aut}(V)$-orbit of a fundamental exceptional solution. We write $\mathcal{E}$ for the collection of exceptional solutions in $V(\mathbb{Z})$. If $x \in V(\mathbb{Z})$ is not an exceptional solution we say it is an unexceptional solution.

Note that if (2.1) or (2.2) occur then they occur in an infinite family for that given $n, a, k$. In either case, all sufficiently large positive integers appear as the maximal entry of some fundamental exceptional solution and this maximal entry determines the tuple up to reordering. Therefore for some $c > 0$ there are $cR + O(1)$ fundamental exceptional solutions with maximal entry $\leq R$. This is not the type of growth we want to study (cf. Theorem 3). It is also clear, but useful to note, that the property of being exceptional (respectively, unexceptional) in $V(\mathbb{Z})$ is $\text{Aut}(V)$-invariant.

**Passage from $V(\mathbb{Z})$ to $V(\mathbb{Z}_+)$**

We now describe the relationship between asymptotic counting of $V(\mathbb{Z}) - \mathcal{E}$ and $V(\mathbb{Z}_+) - \mathcal{E}$. Recall that $n \geq 3$, $a \geq 1$ and $k$ are fixed integers, and $N$ is the group of automorphisms of $V = V_{n,a,k}$ that change the sign of an even number of the coordinates. We decompose the action of $N$ on $V(\mathbb{Z}) - \mathcal{E}$ as follows.

Let $X_0$ be the elements of $V(\mathbb{Z}) - \mathcal{E}$ with at least one coordinate equal to 0. If $k < 0$ then $X_0$ is empty, and if $k \geq 0$ then one obtains for $(x_1, \ldots, x_n) \in X_0$ the equation

$$x_1^2 + \ldots + x_n^2 = k$$

from which it is apparent that $X_0$ is finite, with a bound on its size depending on $n$ and $k$. To indicate this we write $|X_0| = O_{n,k}(1)$.

Now let $X(R) = (V(\mathbb{Z}) - \mathcal{E} - X_0) \cap B(R)$, the unexceptional elements of $V(\mathbb{Z})$ with norm $\leq R$ and no zero coordinate. The group $N$ acts freely on $X(R)$. Therefore

$$2^{n-1}|N \setminus X(R)| = |X(R)|.$$ 

The orbits of $N$ on $X(R)$ fall into two categories, according to which we decompose

$$N \setminus X(R) = Y_+(R) \sqcup Y_-(R)$$
where \( Y_+(R) \) are orbits with a unique representative with all coordinates positive, and \( Y_- (R) \) the remaining orbits, which have a unique representative with \( x_1 < 0 \) and \( x_i > 0 \) for \( i \geq 2 \).

We now argue that \( |Y_-(R)| \) is bounded independently of \( R \). To see this, consider \( N, x \in Y_-(R) \), where \( x \) is the representative described before with \( x_1 \) the only negative coordinate. Let \( \tilde{x}_1 = -x_1 \) and \( \tilde{x}_i = x_i \) for \( i \geq 2 \) be the coordinates of \( \tilde{x} \). The parametrization \( x \to \tilde{x} \) is obviously 1:1 and

\[
\tilde{x}_1^2 + \ldots + \tilde{x}_n^2 + a\tilde{x}_1\tilde{x}_2 \ldots \tilde{x}_n = k.
\]

Because all the \( \tilde{x}_i > 0 \) and \( a \geq 1 \), this equation has no solutions when \( k \leq 0 \) and only finitely many when \( k > 0 \), with a bound depending only on \( n \) and \( k \). In any case, this shows \( |Y_-(R)| = O_{n,k}(1) \).

Since \( Y_+(R) \) is parametrized 1:1 by \( (V(\mathbb{Z}_+) - \mathcal{E}) \cap B(R) \), the previous arguments combine to show

\[
| (V(\mathbb{Z}) - \mathcal{E}) \cap B(R) | = |X(R)| + |X_0 \cap B(R)| = 2^{n-1}|N \setminus X(R)| + O_{n,k}(1)
\]

\[
= 2^{n-1}(|Y_+(R)| + |Y_-(R)|) + O_{n,k}(1)
\]

\[
= 2^{n-1}(|V(\mathbb{Z}_+) - \mathcal{E}) \cap B(R) | + O_{n,k}(1).
\]

**Infinite descent**

The following proposition says that outside of a compact set, the effects of the moves \( m_i \) on the maximal entries of unexceptional Markoff-Hurwitz tuples are at least somewhat predictable. This is a very special feature of the Diophantine equation (1.1) that will allow us to count solutions.

**Proposition 14.** Suppose \( k \in \mathbb{Z} \). There is a compact set \( K_0 = K_0(n,a,k) \) such that for unexceptional \( x \in V(\mathbb{Z}_+) - K_0 \) the following hold:

1. If \( x_j \) is the largest coordinate of \( x \) then the largest entry of \( m_j(x) \) is smaller than \( x_j \), that is, \( (m_j(x))_i < x_j \) for all \( i \).
2. The largest entry of \( x \) appears in exactly one coordinate.
3. If \( x_j \) is not the largest coordinate of \( x \) then it becomes the largest after the move \( m_j \), that is, \( (m_j(x))_j > (m_j(x))_i \) for all \( i \neq j \). (This property holds for all \( x \in V(\mathbb{Z}_+) \).)
4. If \( x_j \) is not the largest coordinate of \( x \), then the number of distinct entries of \( m_j(x) \) is at least the number of distinct entries of \( x \). In particular, if \( x \) has distinct entries then \( m_j(x) \) has distinct entries.
5. Every move \( m_j \) maps \( V(\mathbb{Z}_+) - K_0 \) into \( V(\mathbb{Z}_+) \).

The compact \( K_0 \) can be taken to be a closed ball about the origin in the \( \ell^\infty \) norm on \( \mathbb{R}^n \), and the result still holds after increasing the radius of \( K_0 \).

**Proof of Proposition 14. Part 1.** Suppose without loss of generality that \( x_1 \leq x_2 \leq \ldots \leq x_{n-1} \leq x_n \). Adapting a proof of Cassels from [Cas57, pg. 27], consider the quadratic polynomial in \( x_n \) given by

\[
f(T) = T^2 - ax_1 x_2 \ldots x_{n-1} T + x_1^2 + x_2^2 + \ldots + x_{n-1}^2 - k.
\]
Then $f$ has roots at $x_n$ and $x'_n$ where $x'_n$ is the last entry of $m_n(x)$. The conclusion of Part 1 holds unless

$$x_{n-1} \leq x_n \leq x'_n$$

or

$$x'_n < x_{n-1} = x_n.$$

In either case, since the coefficient of $T^2$ is positive it follows that $f(x_{n-1}) \geq 0$. Then

$$0 \leq f(x_{n-1}) = -ax_1x_2\ldots x_{n-2}^2 + x_1^2 + x_2^2 + \ldots + 2x_{n-1}^2 - k.$$

so

$$ax_1x_2\ldots x_{n-2} + \frac{k}{x_{n-1}^2} \leq n.$$

By an easy argument (cf. Section 2.4) it is possible to increase the radius of $K_0$ so that for $x \in V(\mathbb{Z}^+) - K_0$ ordered as we assume, $x_{n-1} \geq \left(\frac{x_n}{2a}\right)^{\frac{1}{n-1}}$. In particular, we can increase the radius of $K_0$ so that under the ongoing assumptions on $x$, $x_{n-1}^2 > |k|$. Then, since $a, n, x_1, x_2, \ldots, x_{n-1}$ are positive integral, we have

$$ax_1x_2\ldots x_{n-2} \leq n.$$

This means there are a finite number of possibilities for $x_1, x_2, \ldots, x_{n-2}$.

In the case $x'_n \geq x_n$ this directly implies

$$ax_1x_2\ldots x_{n-2}x_{n-1} - x_n \geq x_n$$

so

$$ax_1x_2\ldots x_{n-2}x_{n-1}x_n \geq 2x_n^2.$$

Then from (1.1)

$$x_n^2 \leq x_1^2 + \ldots + x_{n-2}^2 + x_{n-1}^2 - k$$

and it follows that

$$(x_n + x_{n-1})(x_n - x_{n-1}) \leq x_1^2 + \ldots + x_{n-2}^2 - k.$$ 

If $x_n - x_{n-1} > 0$ then the finite number of possibilities for $x_1, x_2, \ldots, x_{n-2}$ yield a finite number of possible $x$.

The alternative is that $x_n = x_{n-1}$, and the following logic also applies to the case $x'_n < x_{n-1} = x_n$. Then $x_n$ is a root of one of finitely many quadratic polynomials

$$(2 - ax_1 \ldots x_{n-2})x_n^2 + x_1^2 + \ldots + x_{n-2}^2 - k = 0.$$

Again, this yields finitely many possibilities for $x$ aside from those having $x_1, \ldots, x_{n-2}$ such that $2 - ax_1 \ldots x_{n-2} = 0$ and $x_1^2 + \ldots + x_{n-2}^2 - k = 0$. Note that if $k \leq 0$ we have exhausted the possibilities. Otherwise we must have either $a = 1$ and $k = (n - 3)1 + 4$ in which case

$$x_1 = x_2 = \ldots = x_{n-3} = 1, \quad x_{n-2} = 2,$$
or \( a = 2 \) and \( k = n - 2 \), in which case

\[ x_1 = x_2 = \ldots = x_{n-2} = 1. \]

These are precisely the fundamental exceptional solutions that are ruled out by hypothesis. Therefore for any given \( n, a, k \) only finitely many unexceptional \( x \) do not satisfy Part 1 of the Proposition.

Part 2. If the largest entry of \( x \) is not unique then performing the move at one of the largest entries does not decrease the largest entry, contradicting Part 1.

Part 3. Suppose \( x_1 \leq x_2 \leq \ldots < x_n \) and let \( x' = (x'_1, \ldots, x'_n) = m_j(x) \) with \( j < n \). The coefficient \( x'_j \) is given by

\[
x'_j - x_n = a \prod_{i \neq j} x_i - x_j - x_n = x_n \left( a \prod_{i \neq j, n} x_i - 1 \right) - x_j.
\]

If \( a \geq 2 \) then the right hand side is \( \geq x_n - x_j > 0 \) so we are done. If \( a = 1 \) and \( x_{n-2} \geq 2 \) then we are also done by a similar argument.

The remaining scenario is \( a = 1 \) and \( x_1 = x_2 = \ldots = x_{n-2} = 1 \). In this case \( x \) satisfies the equation

\[
x_{n-1}^2 + x_n^2 - x_{n-1}x_n = k - n + 2.
\]

The form on the left hand side is positive definite so only finitely many possible solutions exist for \( (x_{n-1}, x_n) \) given \( n \) and \( k \). Add these to the compact set of Part 1.

Part 4. This follows from Part 3 since if \( x' = m_j(x) \) as in the Proposition, then all the entries of \( x'_i \) with \( i \neq j \) are distinct, but \( x'_j \) is larger than all of these.

Part 5. By Part 3 it suffices to check that we can increase the radius of \( K_0 \) so that for \( x \in V(\mathbb{Z}_+) - K_0 \) with \( x_1 \leq x_2 \leq \ldots \leq x_n, m_n(x)_n > 0 \). If not, one obtains \( ax_1 \ldots x_{n-1} - x_n \leq 0 \) from which it follows \( ax_1x_2 \ldots x_n \leq x_n^2 \). The Markoff-Hurwitz equation then gives

\[
x_1^2 + \ldots + x_{n-1}^2 \leq k.
\]

As in Part 1, we can increase the radius of \( K_0 \) so that under the ongoing assumptions on \( x \), \( x_{n-1}^2 \geq |k| \). It follows then that (2.3) cannot occur outside of \( K_0 \).

\[
\square
\]

Corollary 15 (Infinite descent). Any unexceptional Markoff-Hurwitz tuple can be algorithmically reduced to one in the compact set \( K_0 \) by a series of Markoff-Hurwitz moves that strictly decrease maximal entries.

Corollary 15 was established by Markoff [Mar80] in the case \( n = a = 3 \) and \( k = 0 \). In that case, every Markoff triple can be reduced to \((1,1,1)\) by a series of Markoff moves. Hurwitz [Hur07] showed the analogous result for \( n = a > 3 \) and \( k = 0 \) and showed more generally that when \( k = 0 \), the Markoff-Hurwitz tuples can be reduced to a finite set of fundamental solutions. These fundamental solutions were characterized by Baragar in [Bar94b] whenever \( a \geq 2(n - 1)^{1/2} \); he also presented two different constructions yielding sequences of equations whose sets of fundamental solutions grow without bound.
2.2 The polynomial semigroup

We now perform a normalization that allows us to treat all parameters $a, k$ with a semigroup action that only depends on $n$. For $x \in V(Z_+)$ let

$$z = a^{\frac{1}{n-2}}x.$$ 

Note that $a^{\frac{1}{n-2}} \geq 1$ with equality if and only if $a = 1$. Then $z = (z_1, \ldots, z_n)$ satisfies the equation

$$z_1^2 + z_2^2 + \ldots + z_n^2 = z_1z_2\ldots z_n + k'$$  \hspace{1cm} (2.4)

where

$$k' = ka^{\frac{2}{n-2}}.$$ 

Say that $z$ is exceptional/unexceptional if $x$ has the corresponding property. We will also work with ordered tuples $z$ so that

$$z_1 \leq z_2 \leq \ldots \leq z_n.$$ 

Write $\mathcal{M}$ for the set of all such ordered tuples $z \in a^{\frac{1}{n-2}}Z_+^n$ satisfying (2.4). Counting

$$\mathcal{M} \cap B(R)$$ 

is not equivalent to counting $V(Z_+) \cap B(a^{-\frac{1}{n-2}}R)$ due to the presence of elements with duplicate entries. We will return to treat this point in Section 2.3. Let

$$K = a^{\frac{1}{n-2}}K_0$$  \hspace{1cm} (2.5)

where $K_0$ is the compact set from Proposition 14.

The Markoff-Hurwitz moves $\{m_j\}$ induce the moves

$$\lambda_j(z_1, \ldots, z_n) = \left(z_1, \ldots, \hat{z_j}, \ldots, z_n, \prod_{i \neq j} z_i - z_j \right), \quad 1 \leq j \leq n - 1,$$  \hspace{1cm} (2.6)

where $\bullet$ denotes omission$^{10}$. Since $K$ is a closed ball about 0 in the $\ell^\infty$ norm, Part 3 of Proposition 14 implies that the $\{\lambda_j\}$ preserve $\mathcal{M} - K$. Let $\Lambda = \Lambda(n)$ denote the semigroup of piecewise polynomial self-maps of $C^n$ generated by the $\lambda_j$. In Section 2.3 we will reduce Theorem 3 to an orbital counting estimate. For $z_0 \in \mathcal{M} - K$ let

$$\Lambda.z_0 \subset \mathcal{M} - K$$

denote the orbit of $z_0$ under $\Lambda$.

$^{10}$The reason we now have $n - 1$ moves instead of $n$ is that we never perform the move that will decrease the maximal entry, therefore moving us towards $K$. This eliminates backtracking from our ‘random walk’.
Lemma 16. If \( z_0 \in M - K \) has distinct entries then the map \( \Lambda \to M - K \) given by

\[
\lambda \mapsto \lambda(z_0)
\]

is injective. It follows that the semigroup \( \Lambda \) is free\(^{11}\) on the generators \( \{\lambda_j\} \).

Proof. For the first part, if the map is not injective then at some point there must be \( \lambda_1 \in \Lambda \) and some \( j \neq j_2 \) such that

\[
\lambda_{j_1}, \lambda_1(z_0) = \lambda_{j_2}, \lambda_1(z_0).
\]

(2.7)

Since by Proposition 14, Part 4 the entries of \( \lambda_1(z_0) \) are distinct we find \( z = \lambda_1(z_0) \) with distinct entries so that \( \lambda_{j_1}z = \lambda_{j_2}z \). But this cannot be the case since e.g. the sets \( \{z_1, \ldots, \hat{z}_{j_1}, \ldots, z_n\} \) and \( \{z_1, \ldots, \hat{z}_{j_2}, \ldots, z_n\} \) are not the same.

For the second part it is enough to find some \( a \) and \( k \) so that there is a point in \( V(Z_+) - K \) with all entries distinct. Then freeness of \( \Lambda \) follows from the first part of the proof.

Given \( n \) we first choose some \( a \) and \( k \) so that \( V(Z_+) \) contains an infinite orbit. For example, the orbit of \((1,1,\ldots,1)\) in the case \( a = n \) and \( k = 0 \) is infinite and contains no exceptional points. Then we may find \( z_0 \) outside \( K \) with distinct entries, since it is possible to increase the number of distinct entries by application of \( \lambda_i \), using Proposition 14, Part 3.

\[ \square \]

2.3 Multiplicities

In the rest of the paper we will count in orbits of the free semigroup \( \Lambda \). It is extremely useful to be able to work with a fixed free semigroup for each \( n \). The cost of this, however, is that \( \Lambda \) acts on ordered tuples. Since the original problem was to count points in \( V(Z_+) \) we therefore need to take into account the multiplicity of the order map \( V(Z_+) \to M \).

This is best done in relation to the moves \( m_j \). Given \( x \in V(Z_+) - K_0 \), we say that a sequence

\[ j_1, j_2, j_3, \ldots, j_l, \ldots \]

is admissible for \( x \) if for all \( l \), \( j_l \) is not the largest coordinate of

\[ x^{(l-1)} = m_{j_{l-1}} m_{j_{l-2}} \cdots m_{j_2} m_{j_1} x. \]

Notice then that the largest entries of \( x^{(l)} \) are increasing in \( l \) and therefore \( x^{(l)} \in V(Z_+) - K_0 \) for all \( l \geq 1 \). Also, a sequence is admissible if and only if \( j_1 \) is not the largest coordinate and \( j_l \neq j_{l-1} \) for any \( l \leq 2 \). Write \( \Sigma^*(x) \) for the set of all finite admissible sequences for \( x \).

Lemma 17. Given \( x \in V(Z_+) - K_0 \) the map \( \phi_x : \Sigma^*(x) \to V(Z_+) \) given by

\[
\phi_x(j_1, j_2, j_3, \ldots, j_l) = m_{j_1} m_{j_1-1} m_{j_{l-2}} \cdots m_{j_2} m_{j_1} x
\]

is injective. Note this is regardless of whether \( x \) has duplicate entries. Moreover, for any \( x, x' \in V(Z_+) - K_0 \), the images of \( \phi_x \) and \( \phi_{x'} \) are disjoint unless either \( x' \in \text{image}(\phi_x) \) or \( x \in \text{image}(\phi_{x'}) \).

\(^{11}\) As a semigroup of polynomial maps.
Proof. It is clear from Proposition 14, Part 3 that the $m_{j_1}x$ with $j_1$ admissible are distinct. It is then enough to show $\phi_x$ is injective to show that there are no $x \neq x' \in V(\mathbb{Z}_+) - K_0$ and $j, j'$ admissible for the respective $x, x'$ so that $m_j(x) = m_j'(x')$. But since $m_j(x)$ has a distinct largest entry by Proposition 14 Part 2, it has to be the case that $j = j'$. Then applying $m_j$ gives $x = x'$.

Now suppose $x' \notin \text{image}(\phi_x)$ and $x \notin \text{image}(\phi_{x'})$. If $\text{image}(\phi_x) \cap \text{image}(\phi_{x'}) \neq \emptyset$ then at some point there must have been $x^{(3)} \neq x^{(4)} \in V(\mathbb{Z}_+) - K_0$ and $j, j'$ admissible for $x^{(3)}, x^{(4)}$ respectively so that $m_j(x^{(3)}) = m_{j'}(x^{(4)})$. But we have already established this cannot happen. \hfill \Box

Lemma 18. Let $x \in V(\mathbb{Z}_+) - K_0$ and $z = \text{order}(a^{-\frac{1}{2}}x)$ the corresponding element of $\mathcal{M} - K$. There exists a bijection

$$\Theta_x : \Sigma^*(x) \rightarrow \Lambda$$

that is an intertwiner for the map $x' \mapsto z(x') = \text{order}(a^{-\frac{1}{2}}x')$ in the sense that

$$\Theta_x(j_1, j_2, \ldots, j_i) \cdot z(x) = z(\phi_x(j_1, j_2, j_3, \ldots, j_i))$$

for all $(j_1, j_2, \ldots, j_i) \in \Sigma^*(x)$.

Proof. We’ll show for all $x'$ there is a one to one correspondence between the admissible sequences $(j)$ of length 1 and $\{\lambda_j : 1 \leq j \leq n - 2\}$ so that $\Theta_x(j) \cdot z(x) = z(\phi_x(j))$. This is clear if $x_1' \leq x_2' \leq \ldots < x'_n$ is ordered (send $j \mapsto \lambda_j$). Otherwise pick an ordering of $x'$. The general result follows by repeating this process. \hfill \Box

Lemma 17 implies that the set $V(\mathbb{Z}_+) - \mathcal{E}$ decomposes into the finite set $K_0$ and a finite number of orbits of the form

$$\phi_{x^{(0)}}(\Sigma^*(x^{(0)})).$$

Each one of these orbits has either all its points exceptional or unexceptional. Since we assume throughout the rest of the paper that $V(\mathbb{Z}) - \mathcal{E}$ is infinite, it follows that the collection $\mathcal{U}$ of unexceptional basepoints $x^{(0)}$ is finite and nonempty. Summing up,

$$V(\mathbb{Z}_+) - \mathcal{E} - K_0 = \prod_{x^{(0)} \in \mathcal{U}} \phi_{x^{(0)}}(\Sigma^*(x^{(0)})),$$

so

$$|(V(\mathbb{Z}_+) - \mathcal{E}) \cap B(R)| = O_{n,a,k}(1) + \sum_{x^{(0)} \in \mathcal{U}} \sum_{s \in \Sigma^*(x^{(0)})} 1\{\max(\phi_{x^{(0)}}(s)) \leq R\} = O_{n,a,k}(1) + \sum_{x^{(0)} \in \mathcal{U}} \sum_{s \in \Sigma^*(x^{(0)})} 1\{z(\phi_{x^{(0)}}(s)) \leq a^{-\frac{1}{2}}R\}.$$ 

Applying Lemma 18 to the above sum, one obtains

$$O_{n,a,k}(1) + \sum_{x^{(0)} \in \mathcal{U}} \sum_{\lambda \in \Lambda} 1\{(\lambda \cdot z(x^{(0)})) \leq a^{-\frac{1}{2}}R\}.$$ 

Therefore, Theorem 3 will follow from asymptotic estimates for the quantity
\[ \sum \mathbbm{1}_{\lambda \in \Lambda} \left\{ (\lambda, z^{(0)})_n \leq R \right\} \]  

(2.8)

where \( z^{(0)} \in z(U) \subset M - K \). These estimates are taken up in the next section. We draw the reader’s attention to the fact that the count is over \( \Lambda \) and not over \( M \).

### 2.4 Increasing the size of \( K \)

Before we begin the count we increase the size of \( K \). Recall that \( K \) and \( K_0 \) are balls with center 0 in the \( \ell^{\infty} \) norm with radii coupled by (2.5) and that we are free to increase their radii (maintaining the relationship (2.5)). The following can be thought of as regularizing the dynamics of \( M \) at a fixed scale depending on \( n, a, k \). We state our requirements in terms of \( z = (z_1, \ldots, z_n) \).

First we make sure \( z_{n-1} \) is reasonably large compared to \( z_n \). Suppose \( z_{n-1} \leq cz_n^{-\frac{1}{n-1}} \). Then \( z_1 \leq z_2 \leq \ldots \leq z_{n-1} \leq cz_n^{-\frac{1}{n-1}} \). Then (2.4) gives

\[ z_n^2 \leq c^{n-1} z_n^2 + k' \]

which is a contradiction for \( c < 1 \) and \( z_n \) large enough depending on \( k' \). We increase the radius of \( K \) so that

\[ z_{n-1} \geq \frac{1}{2} z_n^{-\frac{1}{n-1}} \]  

(2.9)

for all \( z \in M - K \).

Now we make sure \( z_n \) is large enough so certain inequalities hold. Note that

\[ \frac{(n - 1) \log(1 - 2z_n^{-1/(n-1)}) - (n - 1) \log 2}{\log z_n} \]  

(2.10)

tends to 0 as \( z_n \to \infty \). So we increase the radius of \( K \) so that

\[ (2.10) \geq -1/2 \]  

(2.11)

for all \( z \in M - K \). It will also be convenient for the sake of simplifying arguments to assume that

\[ z_n \geq 10 \]  

(2.12)

for all \( z \in M - K \). Furthermore by increasing the radius of \( K \), using (2.9) we can also ensure

\[ z_{n-1} > 2 \]  

(2.13)

and

\[ z_1^2 + \ldots + z_{n-1}^2 - k' \geq 0 \]  

(2.14)

for \( z \in M - K \).
3 Converting the linear count to the nonlinear count

3.1 Acceleration

In the last Section 2 we reduced our Main Theorem 3 to obtaining an asymptotic for the count

$$\sum_{\lambda \in \Lambda} 1\left\{ (\lambda.z(0))_n \leq R \right\}$$  \hspace{1cm} (3.1)

where $z(0)$ is one of a finite set of unexceptional points in $\mathcal{M} - K$. For the rest of the paper we view $z(0)$ as fixed.

There is a general framework in which to count over the tree-like $\Lambda$, called the renewal method. This was first used in counting by Lalley [Lal88] in the setting of self-similar fractals and subsequently extended by him [Lal89] to the setting of Schottky groups. The essence of the method is a recursion over $\Lambda$.

Our departure from other uses of renewal in counting problems is that we perform what we call acceleration. Concretely, we replace the generators $\{\lambda_j : 1 \leq j \leq n\}$ of $\Lambda$ with the countably infinite set of generators

$$S = S_\Lambda = \{\lambda_{n-1}^A\lambda_j : A \in \mathbb{Z}_{\geq 0}, \ 1 \leq j \leq n - 2\}.$$  

It is easy to see that $S_\Lambda$ are free generators for the subsemigroup

$$\Lambda' = \bigcup_{j=1}^{n-2} \Lambda \lambda_j \subset \Lambda$$

that contains the words beginning with $\lambda_j$, $1 \leq j \leq n - 2$. This acceleration is crucial for our method and has two advantages:

1. The quality of our fitting the nonlinear count for $\Lambda$ to a linear count to $\Gamma$ depends on the size of the quantity

$$\alpha(z) = \prod_{j=1}^{n-2} z_j,$$

cf. Lemma 23 below. This quantity can be small for long words with respect to the generators $\{\lambda_j\}$, because $\lambda_{n-1}$ does not alter $\alpha(z)$. On the other hand, we prove in Lemma 22 that $\alpha(z)$ grows doubly exponentially in the word length with respect to the generators $S_\Lambda$.

2. When we eventually arrive at the dynamics of $\Gamma'$ on $P(\mathbb{R}^n_{\geq 0})$, the unaccelerated system would be non-uniformly contracting and therefore we could not expect there to be a finite invariant measure for this system. On the other hand, the acceleration we perform leads to uniformly contracting dynamics (cf. Proposition 41) and in turn to a nice description of the invariant measure and leading eigenfunction for the transfer operator in the Ruelle-Perron-Frobenius Theorem (Theorem 35).

Now, the orbit $\Lambda.z(0)$ breaks up into the countable union of orbits

$$\Lambda.z(0) = \bigcup_{A_0=0}^{\infty} \Lambda'.\lambda_{n-1}^{A_0} z(0).$$  \hspace{1cm} (3.2)
It is clear that an asymptotic formula for (3.1) is equivalent to an asymptotic formula for

\[ M_0(z, a) \equiv \sum_{\lambda \in \Lambda \cup \{e\}} 1\{\log \log (\lambda z)_n - \log \log z_n \leq a\} \tag{3.3} \]

when \( z = z^{(0)} \). On the other hand, our methods can prove an asymptotic formula for the following quantity

\[ M(z, a) = \sum_{\lambda \in \Lambda' \cup \{e\}} 1\{\log \log (\lambda z)_n - \log \log z_n \leq a\} \tag{3.4} \]

for arbitrary unexceptional \( z \in \mathcal{M} - K \). Precisely, we will obtain the following proposition.

**Proposition 19.** For all unexceptional \( z \in \mathcal{M} - K \) there is a positive constant \( c_* \) such that as \( a \to \infty \),

\[ M(z, a) = e^{\beta a}(c_*(z) + o(1)), \]

where \( \beta > 1 \) is the constant from Theorem 2 and the rate of decay in the small \( o \) does not depend on \( z \). Moreover, the \( c_*(z) \) have a uniform bound depending only on \( n \).

The proof of Proposition 19 will occupy the rest of this Section. Before beginning, we show how Proposition 19 implies our main Theorem 3. This passage relies on the following elementary lemma.

**Lemma 20.** For unexceptional \( z \in \mathcal{M} - K \) we have

\[ (\lambda_{n-1}^A z)_n \geq 2^A z_n. \]

**Proof.** One can calculate easily that for \( z = (z_1, \ldots, z_n) \), \( \lambda_{n-1}^A z \) is obtained by \( A \) applications of the matrix

\[ g_{\alpha(z)} = \begin{pmatrix} 0 & 1 \\ -1 & \alpha(z) \end{pmatrix} \]

to the last two entries of \( z \), where \( \alpha(z) = \prod_{j \leq n-2} z_j \). This quantity will appear repeatedly in the rest of the paper. If \( z = z(x) \) with \( x_1 \leq x_2 \leq \ldots \leq x_n \) then

\[ \alpha(z) = ax_1 x_2 \ldots x_{n-2} \in \mathbb{Z}_+. \]

If \( \alpha(z) = 1 \) then this matrix is torsion and this contradicts the maximal entries of \( \lambda_{n-1}^A z \) growing with \( A \) (since \( z \in \mathcal{M} - K \)). If \( \alpha(z) = 2 \) then \( z \) must be an exceptional solution. Otherwise \( \alpha(z) \geq 3 \) and if we let \( Z_A = (\lambda_{n-1}^A z)_n \) then the \( Z_A \) satisfy the recurrence

\[ Z_{A+1} = \alpha(z)Z_A - Z_{A-1} \geq 2Z_A. \]

Therefore \( (\lambda_{n-1}^A z)_n \geq 2^A z_n. \)

**Proof of Theorem 3 given Proposition 19.** By our previous discussion it suffices to prove an asymptotic formula for \( M_0(z^{(0)}, a) \) for a fixed \( z^{(0)} \). But using (3.2) gives
\[ M_0(z^{(0)}, a) = \sum_{A_0=1}^{\infty} M(\lambda_{n-1}^{A_0}z^{(0)}, a - \log \log (\lambda_{n-1}^{A_0}z^{(0)})_n + \log \log z^{(0)}). \] (3.5)

By using Lemma 20, the value \( A_0 = A_{\text{max}} \) where \( a - \log \log (\lambda_{n-1}^{A_0}z^{(0)})_n + \log \log z^{(0)} \) first becomes negative is bounded by

\[ A_{\text{max}} \leq \frac{\log z^{(0)}e^a}{\log 2}. \]

Let the small \( o \) term in Proposition 19 be bounded in absolute value by a positive function \( F(a) \) that tends to 0 as \( a \to \infty \). Let \( \kappa \) be a small positive constant to be chosen. The \( A_0 \) such that \( a - \log \log (\lambda_{n-1}^{A_0}z^{(0)})_n + \log \log z^{(0)} \geq \kappa a \) contribute

\[ \log z^{(0)}e^{\beta a} \sum_{A_0: a - \log \log (\lambda_{n-1}^{A_0}z^{(0)})_n + \log \log z^{(0)} \geq \kappa a} \frac{c_* (\lambda_{n-1}^{A_0}z^{(0)})}{(\log (\lambda_{n-1}^{A_0}z^{(0)})_n)^\beta} (1 + O(\sup_{a' \geq \kappa a} F(a))). \]

to (3.5) by Proposition 19. Furthermore, by Lemma 20,

\[ \sum_{A_0: a - \log \log (\lambda_{n-1}^{A_0}z^{(0)})_n + \log \log z^{(0)} \geq \kappa a} \frac{c_* (\lambda_{n-1}^{A_0}z^{(0)})}{(\log (\lambda_{n-1}^{A_0}z^{(0)})_n)^\beta} \leq \sum_{A_0} \frac{c_* (\lambda_{n-1}^{A_0}z^{(0)})}{(A_0 \log 2)^\beta} \]

converges to some limit \( c_\infty(z^{(0)}) \) as \( a \to \infty \), using \( \beta > 1 \). Therefore the terms we have discussed so far give a contribution of

\[ \log z^{(0)}c_\infty(z^{(0)})e^{\beta a} (1 + o(1)) \]

to \( M_0(z^{(0)}, a) \) via (3.5).

For the remaining \( A_0 \) such that \( a - \log \log (\lambda_{n-1}^{A_0}z^{(0)})_n + \log \log z^{(0)} < \kappa a \) we use Proposition 19 in a coarser way to get \( M(z, a) \leq Ce^{\beta a} \) for some constant \( C \), uniformly over unexceptional \( z \in \mathcal{M} - K \). Then any remaining \( A_0 \) contributes at most \( Ce^{\beta \kappa a} \) to (3.5). Therefore the remaining contributions are in total at most

\[ A_{\text{max}}Ce^{\beta \kappa a} \leq \frac{\log z^{(0)}Ce^{(1+\beta \kappa) a}}{\log 2} \]

which is negligible when \( 1 + \beta \kappa < \beta \), and we can find such a \( \kappa \) since \( \beta > 1 \).

\[ \square \]

### 3.2 The renewal equation for \( M \)

We now take up the proof of Proposition 19. While the statement of Proposition 19 is uniform over all unexceptional \( z \in \mathcal{M} - K \), our previous arguments show that the unexceptional elements of \( \mathcal{M} - K \) break up into finitely many orbits of \( \Lambda \). Therefore it is sufficient for us to establish Proposition 19 for \( z = \lambda_0z^{(0)} \), where \( z^{(0)} \in z(U) \) is a fixed unexceptional basepoint and \( \lambda_0 \) is an arbitrary element of \( \Lambda \). We therefore view \( z^{(0)} \) as fixed from now on, and we will prove Proposition 19 for \( z = \lambda_0z^{(0)} \), with uniformity over \( \lambda_0 \in \Lambda \).

We now describe the renewal equation, for which we need some new concepts. Define the shift \( s: \Lambda' \to \Lambda' \cup \{e\} \) by
Now extend this definition so that $s(\lambda \lambda_0) = s(\lambda)\lambda_0$ for all $\lambda \in \Lambda'$ and $\lambda_0 \in \Lambda \cup \{e\}$. We define the \textit{distortion function} $\tau_* : \Lambda'.(\Lambda \cup \{e\}) \to \mathbb{R}_{\geq 0}$ by

$$
\tau_*(\lambda) \equiv \log \log(\lambda.z^{(0)})_n - \log \log(s(\lambda).z^{(0)})_n.
$$

This depends on the constant $z^{(0)}$. One also has the iterated version of distortion

$$
\tau_*^N(\lambda) = \sum_{p=0}^{N-1} \tau_*(s^p(\lambda)) = \log \log(\lambda.z^{(0)})_n - \log \log(s^N(\lambda).z^{(0)})_n.
$$

for any $\lambda \in s^{-N}(\Lambda)$. The \textit{renewal equation} for $M$ is then

$$
M(\lambda z^{(0)}, a) = \sum_{\lambda' \in S_{\lambda}} M(\lambda' \lambda z^{(0)}, a - \tau_*(\lambda' \lambda)) + 1\{0 \leq a\}
$$

for all $\lambda \in \Lambda$. Note that the summation above is finite since the $\lambda'$ act to strictly increase maximal entries in $M$.

### 3.3 Iteration

The eventual goal is to compare the asymptotics of $M(\lambda z^{(0)}, a)$ to those of an analogous quantity for the linear semigroup $\Gamma$ introduced in the Introduction. Before this happens, a regularization must occur. In our approach\footnote{In Zagier’s approach in \cite{Zag82} for the case $n = a = 3$, there is a special mapping arising from the close connection between the Markoff equation and hyperbolic geometry. This mapping offers a much better fit to the linear semigroup count than is available in general. See footnote 16 for more on this.}, the quality of the comparison to the linear semigroup depends on the size of

$$
\alpha(z_1, \ldots, z_n) = \prod_{j=n-2} z_j.
$$

It is clear that no $\lambda \in \Lambda$ decreases $\alpha(z)$. To pass to the case that $\alpha(\lambda'.z^{(0)})$ is large, we iterate the renewal equation (3.7) $L$ times. This yields

$$
M(\lambda z^{(0)}, a) = \sum_{\lambda' : s^L(\lambda') = \lambda} M(\lambda' z^{(0)}, a - \tau_*^L(\lambda')) + \sum_{l=1}^{L-1} \sum_{\lambda' : s^l(\lambda') = \lambda} 1\{\tau_*^l(\lambda') \leq a\} + 1\{0 \leq a\}
$$

recalling the definition of $\tau_*^L$ from (3.6). We now show that for suitable $L$ the last two summations in (3.8) are negligible. The following lemma is used at several points in the rest of the paper.

**Lemma 21.** There are constants $c_0$ and $c_1$ depending only on $n$ such that for all $L \in \mathbb{N}$, $x \geq 0$

$$
\sum_{\lambda' : s^L(\lambda') = \lambda} 1\{\tau_*^L(\lambda') \leq x\} \leq c_1(c_0 + x)^L e^x.
$$

\footnote{In Zagier’s approach in \cite{Zag82} for the case $n = a = 3$, there is a special mapping arising from the close connection between the Markoff equation and hyperbolic geometry. This mapping offers a much better fit to the linear semigroup count than is available in general. See footnote 16 for more on this.}
As a consequence, for any \( \delta > 0 \), there is \( c = c(\delta) > 0 \) such that when \( L = \left\lfloor \frac{ca}{\log a} \right\rfloor \) one has
\[
\sum_{l=1}^{L-1} \sum_{\lambda' : s'((\lambda')) = \lambda} \{ \tau'_l(\lambda') \leq a \} = O(e^{(1+\delta)a}) \tag{3.10}
\]
and
\[
c'_l(c_0 + x)^L \leq e^{\delta x} \tag{3.11}
\]
for all \( x \geq a/2 \).

Proof. For the first part of this proof, let \( \tilde{\lambda} \) denote an arbitrary element of \( \Lambda' \), and \( z := \tilde{\lambda}z^{(0)} \). The proof of Lemma 20 can be easily adapted to show that for arbitrary unexceptional \( z' \in M - K \)
\[
(\lambda^A_{n-1}z')_n \geq (\alpha(z') - 1)^A z'_n.
\]
This gives, setting \( z' = \lambda_jz \)
\[
\tau_s(\lambda^A_{n-1}\lambda_j\tilde{\lambda}) = \log \log(\lambda^A_{n-1}\lambda_jz)_n - \log \log z_n
\]
\[
\quad \geq \log \log((\alpha(\lambda_j(z)) - 1)^A(\lambda_jz)_n) - \log \log z_n.
\]
Now,
\[
\alpha(\lambda_j(z)) = \prod_{1 \leq i \leq n-1, i \neq j} z_i = a \prod_{1 \leq i \leq n-1, i \neq j} x_i
\]
where \( x \) is an integer solution to (1.1) corresponding to \( z \). By using (2.9) we get \( \alpha(\lambda_j(z)) \geq z_{n-1} \geq \frac{1}{2}z_n^\frac{1}{n-1} \) and hence using \( (\lambda_jz)_n \geq z_n \),
\[
\tau_s(\lambda^A_{n-1}\lambda_j\tilde{\lambda}) \geq \log \log\left(\frac{1}{2A}z_n^{A/(n-1)}(1 - 2z_n^{-1/(n-1)}A z_n) - \log \log z_n\right)
\]
\[
\quad \geq \log \left(1 + \frac{A}{n-1}\right)\left(1 + \frac{(n-1)\log(1 - 2z_n^{-1/(n-1)} - (n-1)\log 2)}{\log z_n}\right)
\]
\[
\quad \geq \log \left(1 + \frac{A}{2(n-1)}\right), \tag{3.12}
\]
where the last inequality is by the previously prepared (2.11).

Now, if \( \lambda = \lambda^A_{n-1}\lambda_{j_1}\lambda^A_{n-1}\lambda_{j_1-1} \cdots \lambda^A_{n-1}\lambda_{j_2}\lambda^A_{n-1}\lambda_{j_1} \) then by \( l \) applications of (3.12) we get
\[
\tau_s^l(\lambda) = \sum_{p=0}^{l-1} \tau_s(\lambda^p(\lambda)) \geq \sum_{q=1}^l \log \left(1 + \frac{A_q}{2(n-1)}\right).
\]
Therefore the number of \( \lambda' \) that can contribute to (3.9) is bounded by the size of the set
\[
\left\{ (A_1, A_2, A_3, \ldots, A_L) \in \mathbf{Z}_{\geq 0}^L : \sum_{q=1}^L \log \left(1 + \frac{A_q}{2(n-1)}\right) \leq x \right\}. \tag{3.13}
\]

This times the number of possible choices for \( j_1, \ldots, j_L \). The latter can be crudely bounded by \( (n-2)^L \).
Claim. The size of the set in (3.13) is bounded by $(2(n-1)(c_0 + x))^{L}e^x$ for some positive constant $c_0$.

Proof of Claim. We prove this by induction on $L$. The base case ($L = 1$) is clear. For the induction, after choosing the first $A_1$ the remaining $A_2, \ldots, A_L$ must satisfy

$$\sum_{q=2}^{L} \log \left(1 + \frac{A_q}{2(n-1)}\right) \leq x - \log \left(1 + \frac{A_1}{2(n-1)}\right).$$

So the size of the set in (3.13) is bounded by

$$\sum_{A_1=1}^{[2(n-1)e^x]} (2(n-1))^{L-1} \left(c_0 + x - \log \left(1 + \frac{A_1}{2(n-1)}\right)\right)^{L-1} e^x \frac{1}{1 + \frac{A_1}{2(n-1)}} \leq \sum_{A_1=1}^{[2(n-1)e^x]} \frac{1}{2(n-1) + A_1}.$$ 

The final sum is within a constant $c_0$ of $x$. This completes the proof of the Claim.

So in total we obtain that the sum in (3.9) is bounded by $c_1(c_0 + x)L^{e^x}$ with $c_1 = 2(n-2)(n-1)$. As for the stated consequence, we get

$$\sum_{l=1}^{L-1} \sum_{\lambda : s^l(\lambda') = \lambda} 1 \{\tau^l_* (\lambda') \leq a\} \ll c_1^L (c_0 + a)^L e^a.$$ 

If we choose $L \approx ca/\log(1+a)$ with $c$ small enough depending on $\delta$ we obtain our result. \(\square\)

Since we expect $M(\lambda z(0), a) \approx e^{\beta a}$ with $\beta = \beta(n) > 1$, choosing parameters as in Lemma 21 gives

$$M(\lambda z(0), a) = \sum_{\lambda' : s^L(\lambda') = \lambda} M(\lambda' z(0), a - \tau^L_*(\lambda')) + O(e^{(1+\delta)a}) \quad (3.14)$$

and the big $O$ term is truly an error term when $\delta$ is small. The benefits to our iteration in (3.14) can be quantified by the following result.

Lemma 22. There is some $C > 0$ such that for all $\lambda \in \Lambda \cup \{e\}$ and $\lambda'$ such that $s^L(\lambda') = \lambda$, we have both

$$\alpha(\lambda' z(0)) \geq \frac{1}{2} \exp(C\phi^L) \quad (3.15)$$

and

$$(\lambda' z(0))_n \geq \exp(C\phi^L) \quad (3.16)$$

where $\phi = \frac{1 + \sqrt{5}}{2} > 1$ is the golden ratio.

Proof. For $1 \leq j \leq n-2$

$$(\lambda_j z)_n = \prod_{i \neq j} z_i - z_j = z_n z_{n-1} \prod_{i \neq j, n-1, n} z_i - z_j \geq (z_n - 1) z_{n-1}$$
since \( z_i \geq 1 \) for all \( i \) and \( z_{n-1} \geq z_j \). So then for any \( A \geq 0 \)
\[
(\lambda_{n-1}^A z)_n \geq (\lambda_j z)_n \geq (z_n - 1)z_{n-1}.
\]
Then
\[
(\lambda_{n-1}^{A_2} \lambda_j \lambda_{n-1}^{A_1} z)_n \geq (((\lambda_{n-1}^{A_1} \lambda_j)z)_n - 1)(\lambda_{n-1}^{A_1} \lambda_j z)_{n-1}
\]
\[
\geq ((\lambda_{n-1}^{A_1} \lambda_j z)_n - 1)z_n
\]
using the inequality \((\lambda z)_{n-1} \geq z_n\) for any \( \lambda \in \Lambda \). Therefore the numbers
\[
Z_p = (\lambda_{n-1}^{A_p} \lambda_{j_p} \cdots \lambda_{n-1}^{A_2} \lambda_{j_2} \lambda_{n-1}^{A_1} \lambda_{j_1} z)_n \geq 10
\]
(cf. (2.12)) satisfy the two stage recursive estimate \( Z_p \geq (Z_{p-1} - 1)Z_{p-2} \) for \( p \geq 2 \). Then an elementary argument gives the existence of \( C \) such that
\[
Z_p \geq \exp(C\phi^p).
\]
This gives the required (3.16).

On the other hand
\[
\alpha(\lambda_{n-1}^A z) \geq \alpha(\lambda_j z) \geq z_{n-1} \geq \frac{1}{2}z_n^{\frac{1}{n-1}}
\]
where the last inequality is by (2.9). The result (3.15) now follows after replacing \( C \) with a suitable smaller constant. \( \square \)

In the sequel we choose
\[
L = \left\lceil e \frac{a}{\log a} \right\rceil
\]
so that (3.14) and (3.11) hold with\(^{13}\)
\[
\delta = \min \left( \frac{1}{10}, \frac{\beta - 1}{2} \right).
\] (3.17)
Then for all \( \lambda' z^{(0)} \) appearing in (3.14) we have
\[
\alpha(\lambda' z^{(0)}) \geq \frac{1}{2} \exp(C\phi^{ca/\log a})
\] (3.18)
by Lemma 22.

\(^{13}\)We know by Remark 12 that \( \beta \geq 2.\)
3.4 Comparison to the linear count

Now we relate the terms \( M(\lambda'z^{(0)}, a) \) appearing in (3.14) to orbital counting for \( \Gamma \), the linear semigroup defined in the Introduction. We begin with the expression for \( M(\lambda'z^{(0)}, a) \) in (3.4).

Denoting by \( S^{(N)} \) the \( N \)-fold product of the countable generating set \( S \) for \( \Lambda' \), then we can write

\[
M(\lambda'z^{(0)}, a) = \sum_{N=0}^{\infty} \sum_{\lambda^{(2)} \in S^{(N)}} 1\{r^N_{\tau}(\lambda^{(2)}\lambda') \leq a\}. \tag{3.19}
\]

We will proceed by

1. Matching \( \lambda'z^{(0)} \) with some element of \( H \subset \mathbb{R}^n_+ \) that is very close to \( \log(\lambda'z^{(0)}) \).
2. Matching each \( \lambda^{(2)} \) with an element \( \gamma^{(2)} \) of \( \Gamma \) in the obvious way.

With Part 1 in mind, we define for \( z \in \mathcal{M} \)

\[
f(z) \equiv (\log z_1, \log z_2, \ldots, \log z_n-1, \sum_{j=1}^{n-1} \log z_i).
\]

The reason to use this map over just taking log of coordinates is that we expect \( \log(z) \) to be very close to the hyperplane \( H \) defined in (1.5), so we just go ahead and fit \( \log(z) \) to this plane. The following lemma (cf. Lemma 2 in [Zag82]) says that when \( \alpha(z) \) is big, \( f(z) \) is a good fit to \( \log(z) \). In this paper, we write inequalities between vectors to mean they hold at every coordinate.

**Lemma 23.** There are constants \( C_1 \) and \( C_2 \) depending only on \( n \) such that when \( z \in \mathcal{M} - K \) with \( \alpha(z) > C_1 \)

\[
\log(z) \leq f(z) \leq \log(z) + C_2\alpha(z)^{-2}(0, 0, 0, \ldots, 0, 1). \tag{3.20}
\]

**Proof.** Since \( z \) satisfies the equation (2.4), and \( z_n \) is always the larger of the two quadratic roots of the resulting quadratic in \( z_n \), we have

\[
z_n = \frac{A(z)}{2} \left( 1 + \sqrt{1 - \frac{4C(z) - k'}{A(z)^2}} \right)
\]

where

\[
A(z) = \prod_{i=1}^{n-1} z_i, \quad C(z) = \sum_{i=1}^{n-1} z_i^2
\]

and \( k' \geq 0 \) is the constant from (2.4). Now the first inequality of (3.20) follows from (2.14).

\[\text{\footnote{14}{That is, } S^{(N)} \text{ is the elements of } \Lambda' \text{ that are a product of } N \text{ generators. We extend this definition to } S^{(0)} = \{e\}.}}\]

\[\text{\footnote{15}{When we write } \log \text{ of a vector we always mean take log of each coordinate.}}\]

\[\text{\footnote{16}{Although our } f \text{ is not even close to being as good as Zagier’s function } f \text{ from } [Zag82]: \text{ the quality of fit of Zagier’s } f \text{ improves with the size of } z_{n-1} \text{ whereas we need } z_{n-2} \text{ to be big. This is one reason we must accelerate.}}\]
For the second inequality we estimate
\[ \frac{C(z) - k'}{A(z)^2} \leq \sum_{i=1}^{n-1} \frac{z_{n-1}^2}{\prod_{j \neq n} z_j^2} \leq (n-1) \frac{1}{\prod_{j \leq n-2} z_j^2} = (n-1)\alpha(z)^{-2}. \]

We can then choose \( C_1 \) large enough so that when \( \alpha(z) > C_1 \) we have
\[ z_n = A(z)(1 + O_n(\alpha(z)^{-2})), \]
by increasing \( C_1 \) again if necessary we obtain
\[ \log(z_n) = \log(A(z)) + O_n(\alpha(z)^{-2}) = f(z_n) + O_n(\alpha(z)^{-2}). \]

The following adapts an idea of Zagier from [Zag82, Proof of Lemma 3] to our setting. While the strength of approximation is different, we take the same approach in noting that if \( f(z) \) is close to \( y \) then \( f(\lambda_j z) \) will be close to \( \gamma_j y \). Of course this is designed to be iterated.

**Lemma 24.** There are \( C_1, C_2 \) depending only on \( n \) such that for all \( \epsilon > 0 \), for \( z \in \mathcal{M} - K \), \( \alpha(z) > \max(C_1, 2C_2^{1/2} \epsilon^{-1/2}) \), and for \( y^{(1)}, y^{(2)} \in \mathcal{H} \), if
\[ y^{(1)} + \epsilon(0, 0, \ldots, 0, \frac{1}{2}, \frac{1}{2}, 1) < f(z) \leq y^{(2)} \]
then
\[ \gamma_jy^{(1)} + \epsilon(0, 0, \ldots, 0, \frac{1}{2}, \frac{1}{2}, 1) < f(\lambda_j z) \leq \gamma_jy^{(2)} \]
for all \( 1 \leq j \leq n - 1 \).

**Proof.** We first prove the upper bound for \( f(\lambda_j z) \) from (3.22). The inequality \( f(z) \leq y^{(2)} \) implies that \( \log(z_i) \leq y^{(2)}_i \) for \( i \leq n - 1 \). By Lemma 23 we get \( \log(z_n) \leq f(z)_n \leq y_n \) as well. Then \( f(\lambda_j z) \leq \gamma_jy^{(2)} \) follows.

For the other inequality, \( f(z) > y^{(1)} + \epsilon(0, 0, \ldots, 0, 1/2, 1/2, 1) \) implies \( \log(z_i) > y^{(1)}_i + \epsilon/2 \) for all \( i \leq n - 3 \) and \( \log(z_i) > y^{(1)}_i + \epsilon/2 \) for \( i = n - 2, n - 1 \). By Lemma 23, \( \log(z_n) \geq f(z)_n - C_2\alpha(z)^{-2} \geq y^{(1)}_n + \epsilon - C_2\alpha(z)^{-2} \). Since \( \alpha(z) > 2C_2^{1/2} \epsilon^{-1/2} \) we get
\[ \log(z_n) \geq y^{(1)}_n + \epsilon/4. \]

When \( i \leq n - 3 \) we have \( f(\lambda_j z)_i \geq (\gamma_j y^{(1)}_i)_i \) quite clearly. If \( j \leq n - 2 \) we have \( f(\lambda_j z)_{n-2} = \log z_{n-2} \geq y^{(1)}_{n-1} + \epsilon/2 = (\gamma_j y^{(1)})_{n-2} + \epsilon/2 \) and if \( j = n - 1 \) then \( f(\lambda_j z)_{n-2} = \log z_{n-2} \geq y^{(1)}_{n-2} + \epsilon/2 = (\gamma_j y^{(1)})_{n-2} + \epsilon/2 \). At the \( (n-1) \)st coordinate we have \( f(\lambda_j z)_{n-1} = \log z_{n-1} \geq y^{(1)}_{n-1} + 3\epsilon/4 = (\gamma_j y^{(1)})_{n-1} + 3\epsilon/4 \) which is sufficient. It remains to check the last coordinate. Here,
\[ f(\lambda_j z)_n = \sum_{i \neq j} \log z_i \geq \sum_{i \neq j} y^{(1)}_i + 5\epsilon/4 = (\gamma_j y^{(1)})_n + 5\epsilon/4. \]

The inequality above is due to the fact that at least one of \( \log z_{n-2}, \log z_{n-1} \) appear on the left hand side (giving \( \epsilon/2 \)) and \( \log z_n \) also appears (giving \( 3\epsilon/4 \)).
We can now accomplish Parts 1 and 2 of our plan above. Recall we have some fixed \( z^{(0)} \in \mathcal{M} - K \). For each given \( \lambda' \in \Lambda \) (in particular, those that occur in (3.14)) we define
\[
y(\lambda') = f(\lambda' z^{(0)}).
\]
We choose our parameters as follows: let \( C_2 \) be the constant from Lemma 24 and set
\[
\epsilon = \epsilon(a) = 16C_2 \exp(-2C_2^\epsilon \phi \alpha / \log \alpha).
\] (3.23)
so that by (3.18)
\[
4C_2 \alpha (\lambda' z^{(0)})^{-2} \leq \epsilon
\]
for all \( \lambda' \) appearing in (3.14).

**Lemma 25** (Completing Part 1). We have
\[
(1 - \epsilon)y(\lambda') + \epsilon(0, 0, \ldots, 0, 1, 1) < f(\lambda' z^{(0)}) = y(\lambda').
\]

*Proof.* For any nonidentity map \( \lambda' \in \Lambda \),
\[
(\lambda' z^{(0)})_{n-2} \geq (z^{(0)})_{n-1} > 2,
\]
using (2.13). Therefore \( f(\lambda' z^{(0)})_{n-2} \geq \log(2) > 1/2 \). Since \( f(\lambda' z^{(0)}) \in \mathcal{H} \) it follows that
\[
\epsilon f(\lambda' z^{(0)}) \geq \epsilon(0, 0, \ldots, 0, 1, 1),
\]
from which the lemma is a direct consequence.

Now for each
\[
\lambda^{(2)} = \lambda^{A_N}_{n-1} \lambda^{A_{n-1}^{-1}}_{j_N} \ldots \lambda^{A_2}_{n-1} \lambda^{A_1^{-1}}_{j_1} \in S^{(N)}, \quad 1 \leq j_{i} \leq n-2 \forall i
\]
appearing in (3.19), we set
\[
\gamma^{(2)} = \gamma^{(2)}(\lambda^{(2)}) = \gamma^{A_N}_{n-1} \gamma^{A_{n-1}^{-1}}_{j_N} \ldots \gamma^{A_2}_{n-1} \gamma^{A_1^{-1}}_{j_1} \in \Gamma' \cup \{e\}.
\] (3.24)
This is the matching of Part 2. Since \( \Lambda' \) and \( \Gamma' \) are free, this gives a bijective correspondence.

The key point now is that by iterating Lemma 24 we obtain for all coupled \( \lambda^{(2)}, \gamma^{(2)} \),
\[
(1 - \epsilon)\gamma^{(2)} y(\lambda') + \epsilon(0, 0, \ldots, 0, 1, 1) < f(\lambda^{(2)} \lambda' z^{(0)}) \leq \gamma^{(2)} y(\lambda')
\]
where we have used the linearity of the action of \( \Gamma \) to pull out the factor of \( (1 - \epsilon) \). Using Lemma 23 we get
\[
\log(\lambda^{(2)} \lambda' z^{(0)})_n \leq f(\lambda^{(2)} \lambda' z^{(0)})_n \leq (\gamma^{(2)} y(\lambda'))_n
\]
and
\[
\log(\lambda^{(2)} \lambda' z^{(0)})_n \geq f(\lambda^{(2)} \lambda' z^{(0)})_n - \frac{\epsilon}{4} \geq (1 - \epsilon)(\gamma^{(2)} y(\lambda'))_n.
\]
Then taking logarithms gives
\[ \log \log(\lambda(2)^2 \lambda' z(0))_n \leq \log(\gamma(2) y(\lambda'))_n \leq \log \log(\lambda(2)^2 \lambda' z(0))_n + 2\epsilon \] (3.25)
using \(2\epsilon + \log(1 - \epsilon) > 0\) for \(\epsilon \ll 1\).

Note (3.25) also holds when \(\gamma(2) = e, \lambda(2) = e\). Now we claim we can reasonably compare each of the \(M(\lambda' z(0), a - \tau^L_*(\lambda'))\) from (3.14) to \(N(y(\lambda'), a')\) defined in (1.7) with \(a'\) very close to \(a - \tau^L_*(\lambda')\).

**Lemma 26.** We have
\[ N(y(\lambda'), a - \tau^L_*(\lambda') - \epsilon) \leq M(\lambda' z(0), a - \tau^L_*(\lambda')) \leq N(y(\lambda'), a - \tau^L_*(\lambda') + \epsilon). \]

**Proof.** We write out
\[ N(y, a') = \sum_{\gamma(2) \in \Gamma \cup \{\epsilon\}} 1 \{ \log(\gamma(2) y(\lambda'))_n - \log y(\lambda')_n \leq a' \} \]
and compare to
\[ M(\lambda' z(0), a - \tau^L_*(\lambda')) = \sum_{\lambda(2) \in \Lambda \cup \{\epsilon\}} 1 \left\{ \log \log(\lambda(2)^2 \lambda' z(0))_n - \log \log(\lambda' z(0))_n \leq a - \tau^L_*(\lambda') \right\} \]
term by term, matching \(\gamma(2)\) with \(\lambda(2)\) as in (3.24). By (3.25) we have
\[ \log(\gamma(2) y(\lambda'))_n - \log y(\lambda')_n \leq \log \log(\lambda(2)^2 \lambda' z(0))_n - \log \log(\lambda' z(0))_n + 2\epsilon \]
and
\[ \log \log(\lambda(2)^2 \lambda' z(0))_n - \log \log(\lambda' z(0))_n - 2\epsilon \leq \log(\gamma(2) y(\lambda'))_n - \log y(\lambda')_n \]
from which the result follows. \(\square\)

### 3.5 Using the linear semigroup count to prove Proposition 19

We now use Theorem 11, whose proof will be deferred to Section 4. Let \(y' = y(\lambda') = f(\lambda' z(0))\).

**Lemma 27.** Let \(\delta\) be the small constant from (3.17). We have
\[ M(\lambda z(0), a) = (1 + o(1)) e^{\beta a} \sum_{\lambda': s \in (\lambda') = \lambda} e^{-\beta \tau^L_*(\lambda')} h(y') + O\left( \exp(\beta a^\delta + (1 + \delta) a) \right). \]
The big and small \(o\) terms have implied constant and decay rates that are independent of \(\lambda z(0)\).

**Proof.** Using Lemma 26 in the expression (3.14) gives that up to a negligible \(O(e^{(1+\delta) a})\),
\[ \sum_{\lambda': s \in (\lambda') = \lambda} N(y(\lambda'), a - \epsilon - \tau_*(\lambda')) \leq M(\lambda z(0), a) \leq \sum_{\lambda': s \in (\lambda') = \lambda} N(y(\lambda'), a + \epsilon - \tau_*(\lambda')) \] (3.26)
where \( y(\lambda') = f(\lambda' z(0)) \).

We want to carefully use Theorem 11 that says that along with \( h, \beta \) there is some function \( F(a) \) such that

\[
|N(y, a) - e^{\beta a} h(y)| \leq F(a) e^{\beta a} h(y)
\]

and \( F(a) \to 0 \) as \( a \to \infty \). The minor problem with using this in (3.26) is that there may be terms with \( a' = a \pm \epsilon - \tau_*(\lambda') \) close to zero, or less than zero. Letting \( \delta \) be the same small parameter as before, we note that if \( a' \leq a^\delta \) then there is some constant \( C_3 \geq 1 \) such that

\[
|N(y, a') - e^{\beta a'} h(y)| \leq C_3 e^{\beta a'}
\]

which follows from Theorem 11 when \( 0 \leq a' \leq a^\delta \) and is trivial when \( a' < 0 \) since then \( N(y, a') = 0 \).

Therefore, working with the right hand inequality of (3.26) we get

\[
M(\lambda z(0), a) \leq \sum_{\lambda': s^L(\lambda') = \lambda} (e^{\beta a'} h(y') + 1\{ a' \leq a^\delta \} C_3 e^{\beta a'} + 1\{ a' > a^\delta \} F(a') e^{\beta a'} h(y'))
\]

where we write \( a' = a'(\lambda') = a + \epsilon - \tau_*^L(\lambda') \) and \( y' = y(\lambda') \). Therefore

\[
M(\lambda z(0), a) \leq \left( 1 + \sup_{b \geq a^\delta} F(b) \right) \sum_{\lambda': s^L(\lambda') = \lambda} e^{\beta a'} h(y') + C_3 \sum_{\lambda': s^L(\lambda') = \lambda, a' \leq a^\delta} e^{\beta a'}.
\] (3.27)

For the first term in (3.27) note that

\[
\sum_{\lambda': s^L(\lambda') = \lambda} e^{\beta a'} h(y') = e^{\beta a} \sum_{\lambda': s^L(\lambda') = \lambda} e^{\beta \epsilon - \beta \tau_*^L(\lambda')} h(y')
\]

\[
= (1 + O(\exp(-2C_1 \phi \log a))) e^{\beta a} \sum_{\lambda': s^L(\lambda') = \lambda} e^{-\beta \tau_*^L(\lambda')} h(y').
\]

The last term in (3.27) can be bounded by

\[
\ll e^{\beta a} \sum_{\lambda': s^L(\lambda') = \lambda, \tau_*^L(\lambda') \geq a + \epsilon - a^\delta} e^{-\beta \tau_*^L(\lambda')}.
\]

The contributions to the sum above from \( M - 1 \leq \tau_*^L(\lambda') \leq M \) are bounded by

\[
\sum_{\lambda': s^L(\lambda') = \lambda, \tau_*^L(\lambda') \geq M - 1} 1\{ M \geq \tau_*^L(\lambda') \geq M - 1 \} e^{-\beta \tau_*^L(\lambda')} \leq c_1^L (c_0 + M)^L e^{M} e^{-\beta(M-1)}
\]

by Lemma 21, equation (3.9). Summing this quantity over natural numbers from \( M_0 = [a - a^\delta - 1] \) to infinity, using the bound (3.11) to replace \( c_1^L (c_0 + M)^L \) by \( e^{M} \), gives

\[
\sum_{\lambda': s^L(\lambda') = \lambda, \tau_*^L(\lambda') \geq a + \epsilon - a^\delta} e^{-\beta \tau_*^L(\lambda')} \ll e^{-(\beta - 1 - \delta)(a - a^\delta)};
\]

30
so we get for the last term in (3.27)

\[
\sum_{\lambda':s^L(\lambda')=\lambda, a' \leq a^\delta} e^{\beta a'} \ll \exp((\beta - 1 - \delta)a^\delta + (1 + \delta)a).
\]

Therefore it can be absorbed into the error stated in the lemma. The lower bound for
\( M(\lambda z(0), a) \) is similar. Notice that our constants and rates of decay do not depend on \( \lambda z(0) \).

Proposition 19 will now follow from Lemma 27 and the following proposition.

**Proposition 28.** For fixed \( \lambda \) and \( z(0) \) there is a constant \( c_*(\lambda z(0)) \) such that

\[
a_L(\lambda z(0)) = \sum_{\lambda':s^L(\lambda')} h(y(\lambda')) e^{-\beta \tau^L_*(\lambda')} = c_*(\lambda z(0)) + o(1)
\]
as \( L \to \infty \), with a rate of decay that is independent of \( \lambda \). The values \( c_*(\lambda z(0)) \) are bounded by some constant independent of \( \lambda \).

**Proof.** We are going to prove the sequence is Cauchy with a very fast rate. Consider the difference of consecutive terms. Again we write \( y' = y(\lambda') \). For \( \lambda'' \in S_\Lambda \) we write \( y'' = y''(\lambda''', \lambda') = f(\lambda''\lambda' z(0)) \). We suppress the dependence of these variables on others to improve readability.

We obtain

\[
a_{L+1} - a_L = \sum_{\lambda: s^L(\lambda') = \lambda} h(y'') e^{-\beta \tau^L_{s^L(\lambda')}(\lambda')} - \sum_{\lambda': s^L(\lambda') = \lambda} h(y') e^{-\beta \tau^L_{s^L(\lambda')}(\lambda')}
\]
\[= \sum_{\lambda: s^L(\lambda') = \lambda} e^{-\beta \tau^L_{s^L(\lambda')}(\lambda')} \left( \sum_{\lambda'' \in S_\Lambda} h(y'') e^{-\beta \tau^L_{s^L(\lambda')}(\lambda'') - \tau^L_{s^L(\lambda')}(\lambda')} \right) - h(y') \]
\[= \sum_{\lambda: s^L(\lambda') = \lambda} e^{-\beta \tau^L_{s^L(\lambda')}(\lambda')} \left( \sum_{\lambda'' \in S_\Lambda} h(y'') e^{-\beta \tau^L_{s^L(\lambda')}(\lambda'')} \right) - h(y') \]
\[= \sum_{\lambda: s^L(\lambda') = \lambda} e^{-\beta \tau^L_{s^L(\lambda')}(\lambda')} \left( \sum_{\lambda'' \in S_\Lambda} h(y'') \left( \frac{\log(\lambda' z(0))_n}{\log(\lambda'' \lambda' z(0))_n} \right)^{\beta} \right) - h(y') \quad (3.28)
\]

The point is that the terms in parentheses should be close to zero by the recursion (1.8) satisfied by \( h \) over \( \Gamma' \). We will use Lemma 23 which gives a bound when \( \alpha(\lambda' z(0)) > C_1 \). On the other hand by Lemma 22 there is some \( L_0 \) such that when \( L \geq L_0 \) and \( s^L(\lambda') = \lambda \) then \( \alpha(\lambda' z(0)) > C_1 \).

We use the natural bijection

\[
S_\Lambda \rightarrow T_{\Gamma}, \quad \lambda'' \mapsto \gamma(\lambda'').
\]

When \( L > L_0 \), repeating the arguments of the previous section leading up to (3.25) gives the bounds

\[31\]
\[
\log(\lambda' z'(0))_n \leq y'_n \leq (1 + O(\alpha(\lambda' z'(0))^{-2})) \log(\lambda' z'(0))_n
\]  
(3.29)

\[
\log(\lambda'' \lambda' z'(0))_n \leq (\gamma(\lambda'')y')_n \leq (1 + O(\alpha(\lambda' z'(0))^{-2})) \log(\lambda'' \lambda' z'(0))_n
\]  
(3.30)

where the implied constant depends only on \(n\). Moreover, using Lemma 23 gives

\[
\log(\lambda'' \lambda' z'(0)) \leq y'' \leq \log(\lambda'' \lambda' z'(0)) + C_2 \alpha(\lambda' z'(0))^{-2}(0, 0, \ldots, 0, 1)
\]  
(3.31)

whenever \(L > L_0\).

Suppose \(L > L_0\). We must estimate the cost of replacing \(y''\) by \(\gamma(\lambda'')y'\) and \(\left(\frac{\log(\lambda' z'(0))_n}{\log(\lambda'' \lambda' z'(0))_n}\right)^\beta\)

by \(\left(\frac{y'_n}{\gamma(\lambda'')y'}\right)^\beta\) in (3.28). Since using (3.30) and (3.31) gives that \(y''\) is within \(O(\alpha(\lambda' z'(0))^{-2})\log(\lambda'' \lambda' z'(0))_n\) of \(\gamma(\lambda'')y'\) and \(h\) is \(C^1\), we get

\[
h(y'') = h(\gamma(\lambda'')y') + O(\alpha(\lambda' z'(0))^{-2}) \log(\lambda'' \lambda' z'(0))_n).
\]

Using (3.29) and (3.30) gives

\[
\left(\frac{y'_n}{(\gamma(\lambda'')y')_n}\right)^\beta (1 + O(\alpha(\lambda' z'(0))^{-2}))^{-\beta} \leq \left(\frac{\log(\lambda' z'(0))_n}{\log(\lambda'' \lambda' z'(0))_n}\right)^\beta \leq \left(\frac{y'_n}{\gamma(\lambda'')y'}\right)^\beta (1 + O(\alpha(\lambda' z'(0))^{-2}))^\beta.
\]

Using that \(h\) and \(\left(\frac{\log(\lambda' z'(0))_n}{\log(\lambda'' \lambda' z'(0))_n}\right)^\beta\), \(\left(\frac{y'_n}{(\gamma(\lambda'')y')_n}\right)^\beta\) are bounded we get

\[
\sum_{\lambda'' \in S_\Lambda} h(y'') \left(\frac{\log(\lambda' z'(0))_n}{\log(\lambda'' \lambda' z'(0))_n}\right)^\beta = \sum_{\gamma'' \in T_\Gamma} h(\gamma(\lambda'')y') \left(\frac{y'_n}{\gamma(\lambda'')y'}\right)^\beta + O(\alpha(\lambda' z'(0))^{-2})
\]

\[
\sum_{\lambda'' \in S_\Lambda} h(y'') \left(\frac{\log(\lambda' z'(0))_n}{\log(\lambda'' \lambda' z'(0))_n}\right)^\beta = h(y') + O(\alpha(\lambda' z'(0))^{-2})
\]

where the last equality uses the recursion (1.8). Therefore for \(L \geq L_0\)

\[
|a_{L+1} - a_L| \ll \alpha(\lambda' z'(0))^{-2} \sum_{\lambda'' : \gamma''(\lambda'') = \lambda} e^{-\beta \tau''(\lambda')}.
\]

It is possible to use a fortiori estimates to prove the sum above is universally bounded, for example by using the work of Baragar [Bar94a] in the case of \(k = 0\). To keep things self contained, since we only need a coarse bound we instead use Lemma 21 to prove

\[
\sum_{\lambda'' : \gamma''(\lambda'') = \lambda} e^{-\beta \tau''(\lambda')} \ll \exp(C_4L^{1+\eta})
\]  
(3.32)

for some constant \(C_4\) and small \(\eta\). However, \(\alpha(\lambda' z'(0))^{-2}\) is much smaller than this: by Lemma 22 we have \(\alpha(\lambda' z'(0))^{-2} \ll \exp(-2C\phi L)\) where \(\phi > 1\) so not only is

\[
|a_{L+1} - a_L| \ll \exp(C_4L^{1+\eta} - 2C\phi L)
\]

32
very small but we can sum the differences to get a Cauchy sequence. Indeed $C_4 L^{1+\eta} - 2C_3 \phi^L \leq C_5 - C_6 \phi^L$ for some $C_5, C_6 > 0$. Therefore for $L_1 \geq L_0$

$$\sum_{L=L_1}^{\infty} |a_{L+1} - a_L| \ll \sum_{L=L_1}^{\infty} \exp(-C_6 \phi^L) = o_{L_1 \to \infty}(1) \quad (3.33)$$

so the sequence converges at a uniform rate to its limit $c_\ast(\lambda z^{(0)})$. The uniform boundedness of $c_\ast(\lambda z^{(0)})$ will follow from the uniform boundedness of $a_{L_0}(\lambda z^{(0)})$ given (3.33), and $a_{L_0}(\lambda z^{(0)})$ is uniformly bounded by using that $h$ is bounded and the already established (3.32). This finishes the proof.

Putting Proposition 28 and Lemma 27 together proves Proposition 19 given Theorem 11. In the rest of the paper we prove Theorem 11.

4 The linear semigroup count

4.1 Renewal (again)

Now we discuss renewal for the quantity $N(y,a)$ that appears in Theorem 11. The renewal equation for $N(y,a)$ says

$$N(y,a) = \sum_{\gamma \in \Gamma} N(\gamma.y,a - \log(\gamma.y)^n + \log y^n) + 1\{0 \leq a\}. \quad (4.1)$$

Notice from its Definition in (1.7) that the function $N(y,a)$ is invariant under multiplication of the $y$ variable by $R^+$. With this in mind, we are going to consider

$$P(R_{n \geq 0}^n) = R_{n \geq 0}^n/R^+,$$

the quotient of $R_{n \geq 0}^n$ by the multiplicative action of positive real numbers. Let $\Delta \subset P(R_{n \geq 0}^n)$ denote the projection of $H$. We will from now on use a coordinate

$$w = (w_1, w_2, \ldots, w_{n-1}, 1)$$

with $w_1 \leq w_2 \leq \ldots \leq w_{n-1}$ and $\sum_{j=1}^{n-1} w_j = 1$ to uniquely represent a point in $\Delta$. We now view $N(w,a)$ as a function on $\Delta \times R_{\geq 0}$. Note that equation (4.1) descends to $(w,a) \in \Delta \times R_{\geq 0}$.

Now, for the first time in the paper, we start the full argument of the renewal method\textsuperscript{17}. This begins with taking a Laplace transform which we define for general $f$ of suitable decay by

$$\hat{f}(s) = \int_{-\infty}^{\infty} e^{-sx} f(x) \, dx.$$ 

The outcome of taking a Laplace transform of the renewal equation (4.1) in the $a$ variable, ignoring issues of convergence\textsuperscript{18}, is that

\textsuperscript{17}Previously we just used an iteration of a renewal equation to perform a linearization. 
\textsuperscript{18}These issues are worked out in Lemma 31.
\[ \hat{N}(w, s) = \sum_{\gamma \in T_\Gamma} \left( \frac{w_n}{\gamma(w)_n} \right)^s \hat{N}(\gamma.w, s) + \frac{1}{s} \]  

(4.2)

for all \( w \in \Delta \), where \( \hat{N}(w, s) \) is the Laplace transform \( \hat{N}(w, \bullet) \) in the \( a \) variable. Thus \( s \) is a frequency parameter dual to the counting parameter \( a \). Notice that the function \( (\gamma, w) \mapsto w_n(\gamma.w)_n \) descends from \( T_\Gamma \times H \) to a well defined real valued function on \( T_\Gamma \times \Delta \).

Now we introduce the transfer operator that will play a crucial role in this section. For a function \( f \) on \( \Delta \) we define

\[ L_s[f](w) = \sum_{\gamma \in T_\Gamma} \left( \frac{w_n}{\gamma(w)_n} \right)^s f(\gamma.w) \]  

(4.3)

whenever the sum is pointwise absolutely convergent on \( \Delta \). Then (4.2) can be rephrased as

\[ \hat{N}(\bullet, s) = s^{-1}(1 - L_s)^{-1}1, \]  

(4.4)

whenever the resolvent operator \( (1 - L_s)^{-1} \) exists in such a way it can act on the constant function \( 1 \).

There is a procedure due to Lalley to convert (4.4) together with a sufficiently complete description of the spectrum of \( L_s \) on a suitable Banach space into Theorem 11. More specifically we will appeal to the perturbation theory and Fourier analysis developed in [Lal89, Sections 7 and 8]. In the next section we will lay out the necessary spectral theory of \( L_s \).

Before that, let us calculate explicitly the sum in (4.3).

**Lemma 29.** An element \( \gamma^A_n \gamma_j \) of \( T_\Gamma \) acts on \( \Delta \) by

\[ \gamma^A_{n-1} \gamma_j[w_1, \ldots, w_{n-1}, 1] = [w_1, \ldots, \hat{w}_j, \ldots, w_{n-1}, 1 + A(1 - w_j), 1 + (A + 1)(1 - w_j)]; \]  

(4.5)

in particular,

\[ (\gamma^A_{n-1} \gamma_j, (w_1, \ldots w_{n-1}, 1))_n = 1 + (A + 1)(1 - w_j). \]  

(4.6)

**Proof.** This is a direct calculation.

**Example 30** (Gauss map). When \( n = 3 \), the only inverse branches are of the form

\[ \gamma^A_2 \gamma_1(w_1, w_2) = (1 + (A + 1)w_2)^{-1}(w_2, 1 + Aw_2). \]

With the change of variables \( x = w_1/w_2 \), these are precisely the inverse branches of the Gauss map \( x \mapsto \{1 \over x\} \):

\[ \gamma^A_2 \gamma_1 : x \mapsto \frac{1}{x + A + 1}, \quad A \in \mathbb{Z}_{\geq 0}. \]
4.2 Spectral theory of the transfer operator

In this section, we give a full account of the spectral theory of $L_s$. A good reference for the spectral theory of transfer operators is the book of Baladi [Bal00]. We begin with the following lemma.

**Lemma 31.** When $\Re(s) > 1$ the summation in the defining equation (4.3) of $L_s$ is absolutely and uniformly convergent on $\Delta$ and so gives a well defined continuous map of Banach spaces\(^{19}\)

$$L_s : C^0(\Delta) \rightarrow C^0(\Delta).$$

**Proof.** Substituting Lemma 29, equation (4.6) in the Definition (4.3), the summation amounts to

$$L_s[f](w) = \sum_{j \in [n-2]} \sum_{A \in \mathbb{N}} \frac{1}{(1 + (A + 1)(1 - w_j))^s} f(\gamma_n^{A} \gamma_{j} w).$$

(4.7)

Since $w_j \leq 1/2$ for $j \in [n-2]$ and $f$ is bounded, each sum in $L$ converges uniformly absolutely on $\Delta$ for $\Re(s) > 1$. The limit is then continuous and bounded by a constant multiple, depending on $s$, of $\|f\|\infty$. \qed

We obtain the following consequence of Lemma 31 by a standard application of the Schauder-Tychonoff Theorem.

**Corollary 32** (Existence of eigenmeasures). Let $L_s^*$ denote the dual of $L_s$. For each real $s > 1$ there is a number $\lambda_s > 0$ and a probability measure $\nu_s$ such that $L_s^* \nu_s = \lambda_s \nu_s$.

**Example 33** (Transfer operator for the Gauss map). Let $n = 3$. Carrying on from Example 30, we have in the coordinate $x = w_1/w_2$

$$L_s[f](x) = \sum_{A \in \mathbb{N}} \frac{(x + 1)^s}{(x + A + 2)^s} f \left( \frac{1}{x + 1 + A} \right).$$

This is not the usual transfer operator for the Gauss map. However, letting $M_{(x+1)^s}$ denote the operator of multiplication by $(x + 1)^s$, we get

$$M_{(x+1)^s}^{-1} L_s M_{(x+1)^s}[f](x) = \sum_{A \in \mathbb{N}} \frac{1}{(x + A + 1)^s} f \left( \frac{1}{x + 1 + A} \right) = L_s^{\text{Gauss}}[f](x),$$

the classical transfer operator for the Gauss map. This coincides with the Perron-Frobenius operator for the Gauss map when $s = 2$. The leading eigenfunction of $L_2^{\text{Gauss}}$ corresponds to a multiplicity 1 eigenvalue 1 and eigenfunction

$$h(x) = \frac{1}{1+x}.$$ 

This eigenfunction was known to Gauss [Gau], and its invariance property was formally proved by Kuzmin [Kuz32]. Correspondingly, the leading eigenfunction of $L_2$ is $[M_{(x+1)^2}h](x) = (x + 1) = \frac{1}{w_2}$ with eigenvalue 1.

\(^{19}\) $C^0$ is the Banach space of continuous functions with the supremum norm.
Our functional analysis takes place on the Banach space $C^1(\Delta)$ which consists of continuously differentiable functions on $\Delta$ with the norm
\[ \|f\|_{C^1} = \|f\|_\infty + \|\nabla f\|_\infty. \]
We use the standard Euclidean metric on $\Delta$ given by the coordinates $w_1, \ldots, w_{n-1}$.

**Lemma 34.** In the region $\Re(s) > 1$, the mapping $s \mapsto L_s$ gives a holomorphic family of bounded operators on the Banach space $C^1(\Delta)$. In particular, for $\Re(s) > 1$, $L_s$ is bounded on $C^1(\Delta)$.

We will prove the following version of the Ruelle-Perron-Frobenius Theorem.

**Theorem 35** (Ruelle-Perron-Frobenius). Let $s \in (1, \infty)$ be a real parameter for the transfer operator $L_s : C^1(\Delta) \to C^1(\Delta)$.

1. The eigenvalue $\lambda_s$ is multiplicity one and the rest of the spectrum of $L_s$ is contained in a ball of radius $R(s)$ strictly less than $\lambda_s$. For any compact interval $I \subset (1, \infty)$ there is an $\epsilon(I) > 0$ such that $\lambda_s - R(s) \geq \epsilon$ for $s \in I$.

2. There is a unique probability measure $\nu_s$ such that $L_s^* \nu_s = \lambda_s \nu_s$.

3. The unique eigenfunction $h_s \in C^1(\Delta)$ for the eigenvalue $\lambda_s$ with $\nu_s(h_s) = 1$ is positive.

In the case of the Gauss map, a version of Theorem 35 was first proved by Wirsing [Wir74].

It is well-known that Theorem 35 follows from eventually contracting dynamics for example, by the use of Birkhoff cones and contraction of a Hilbert projective metric as in the paper of Liverani [Liv95]. The only thing that is possibly nonstandard about our setting is the presence of both countably many branches and a semigroup action for which we expect the invariant set to have non full Hausdorff dimension (cf. Figures 1 and 2). We explain the proof of Lemma 34 and Theorem 35 in Section 4.4.

These proofs depend crucially on our dynamics being uniformly contracting, which we make precise in Proposition 41. We freely make use of this property henceforth. Let $T_{\Gamma}^{\mathbb{Z}^+}$ denote the set of all positively indexed sequences $(\gamma^{(1)}, \gamma^{(2)}, \ldots)$ with each $\gamma^{(j)} \in T_{\Gamma}$. Because the elements of $T_{\Gamma}$ uniformly contract $\Delta$, one obtains for any fixed $w_0 \in \Delta$ a map
\[ \lim : T_{\Gamma}^{\mathbb{Z}^+} \to \Delta, \quad \lim(\gamma^{(1)}, \gamma^{(2)}, \ldots) := \lim_{j \to \infty} \gamma^{(1)} \cdots \gamma^{(j)} w_0; \]

in fact, this map does not depend on the choice of $w_0$. The image of this map is the attractor of the iterated function system given by the elements of $T_{\Gamma}$, which we also call the limit set of $\Gamma'$, and denote it by $\mathcal{R}(\Gamma')$. Then $\mathcal{R}(\Gamma')$ is a compact $\Gamma'$-invariant subset of $\Delta$.

The Ruelle-Perron-Frobenius Theorem is not enough for input to Lalley’s framework of complex analysis. One must also know that there is some non trivial spectral bound for $L_s$ on the vertical line $s = \beta + it$, the trivial bound being that the spectral radius is no greater than $\lambda_\beta$. In the context of subshifts of finite type, this was investigated by Pollicott in [Pol84] who found a cohomological criterion for a nontrivial spectral bound. We make the following Definition as in Pollicott [Pol84, pg. 139], adapted to the current setting.
Definition 36. We say that a function \( f = u + iv \) with
\[
u, v : T_\Gamma \times \Delta \to \mathbb{R}
\]
is regular if there is no \( r \in \mathbb{R} \) and bounded\(^{20}\) function \( G : \mathfrak{K}(\Gamma') \to \mathbb{R} \) such that
\[
v(\gamma, w) - G(\gamma, w) + G(w) - r \in 2\pi \mathbb{Z}
\]
for all \( \gamma \in T_\Gamma \) and \( w \in \mathfrak{K}(\Gamma') \). In other words, there is no \( r \in \mathbb{R} \) so that \( v - r \) is cohomologous on \( \mathfrak{K}(\Gamma') \) to a \( 2\pi \mathbb{Z} \)-valued function.

The following theorem can be viewed as an extension of a result of Wielandt \([\text{Wie50}]\) on the spectrum of finite dimensional complex matrices. It was proved by Pollicott \([\text{Pol84, Theorem 2}]\) in the context of shifts of finite type in symbolic dynamics. The proof goes through perfectly well in our context\(^{21}\) to give

**Theorem 37** (Wielandt’s Theorem, after Pollicott). If
\[
F_s(\gamma, w) \equiv -s \log \left( \frac{\gamma.w_n}{w_n} \right) \in C^1(\Delta; \mathbb{C}) \quad (4.8)
\]
is regular, \( \Im(s) \neq 0 \), and \( \Re(s) > 1 \) then the spectral radius of the operator \( L_s : C^1(\Delta) \to C^1(\Delta) \) is strictly less than \( \lambda_{\Re(s)} \).

This is applicable in the present setting:

**Proposition 38.** For all \( s \in \mathbb{C} - \mathbb{R} \), the function in \((4.8)\) is regular.

**Proof.** It is enough to show that for
\[
\tau(\gamma, w) = \log \left( \frac{\gamma.w_n}{w_n} \right) = \log(\gamma.w)_n - \log w_n
\]
there is no bounded \( G \) on \( \mathfrak{K}(\Gamma') \) such that the values of
\[
\tau'(\gamma, w) := \tau(\gamma, w) - G(\gamma, w) + G(w)
\]
for \( (\gamma, w) \in T_\Gamma \times \mathfrak{K}(\Gamma') \) are contained in a translate of a discrete subgroup of \( \mathbb{R} \). So it is also enough to show that for any such \( \tau' \), the gaps between distinct values of \( \tau' \) are not bounded below.

The fundamental simple fact we use is that for \( \gamma \in T_\Gamma \) and \( w \) such that \( \gamma.w = w \), (from which it follows \( w \in \mathfrak{K}(\Gamma') \))
\[
\tau'(\gamma, w) = \tau(\gamma, w) - G(\gamma, w) + G(w) = \tau(\gamma, w).
\]

\(^{20}\) It is possible to impose more regularity on \( G \) in this definition but it is not necessary for our purposes.

\(^{21}\) The main point is that our definition of regular function is strong enough to rule out \( L_s \) having an eigenvalue of modulus \( \lambda_{\Re(s)} \). This fact is supplemented by compactness arguments relying on the Ionescu Tulcea-Mariesc type inequality that we establish in Lemma 42.
Then it remains to show that gaps between distinct values of $\tau$ on the fixed points of $\gamma \in T\Gamma$ are not bounded below. We compute that

$$\gamma^A_{n-1} \gamma_{n-2} = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & 0 \\ A & \cdots & A & 0 & A & 1 \\ A + 1 & \cdots & A + 1 & 0 & A + 1 & 1 \end{pmatrix},$$

so (using the block lower triangular structure)

$$\det(\gamma^A_{n-1} \gamma_{n-2} - TI_n) = (1 - T)^{n-3}(-T)(T^2 - (A + 1)T - 1).$$

Consequently, the eigenvalues aside from 0 and 1 are

$$T = A + 1 \pm \sqrt{(A + 1)^2 + 4}.$$

Let $T_+$ be the largest, that is, $T_+ = \frac{A + 1 + \sqrt{(A + 1)^2 + 4}}{2} > A + 1$. One can find an eigenvector $v_+$ for $T_+$ where

$$v_+ = (0, 0, \ldots, 0, 1, T_+, T_+(T_+ - A)) > 0,$$

moreover, $v_+ \in \mathcal{H}$. Now, $(\gamma.v_+)_n/(v_+)_n = T_+$ and so $\tau(\gamma^A_{n-1} \gamma_{n-2}, [v_+]) = \log T_+$. Writing $T_+ = T_+(A)$, we have

$$\log T_+(A + 1) - \log T_+(A) = \log \left( \frac{A + 2 + \sqrt{(A + 2)^2 + 4}}{A + 1 + \sqrt{(A + 1)^2 + 4}} \right) = \log \left( \frac{1 + \frac{1}{A + 1}}{1 + \frac{1}{A + 1} + \sqrt{1 + \frac{4}{(A + 1)^2}}} \right) \to 0$$

as $A \to \infty$. As the terms are easily seen to be non-zero, this completes the proof.

The contour shifting argument of Lalley hinges on the behavior of the eigenvalue $\lambda_s$ and, in particular, on the location of the possible real value $\beta$ such that $\lambda_\beta = 1$. Since our dynamics is suitably uniformly contracting, if such a value exists it is unique:

**Proposition 39.** The eigenvalue $\lambda_s$ is a real analytic function of $s$ that is strictly decreasing on $(1, \infty)$. We have $\lambda_s < 1$ for sufficiently large $s$. As such, any value $\beta_0 \in (1, \infty)$ such that $\lambda_{\beta_0} = 1$ is unique, and if no such $\beta_0$ exists then $\lambda_s < 1$ for all $s \in (1, \infty)$.

As we will discuss momentarily, such a $\beta_0$ does exist, and it coincides with Baragar’s $\beta$ from Theorem 2. Note that when $s = \beta$ we obtain from Theorem 35 a unique measure such that $\mathcal{L}_{\beta}^* \nu_\beta = \nu_\beta$. Then we will show $\nu_\beta$ is the conformal measure of Theorem 9. Proposition 39 will be proved in Section 4.5.
4.3 Proofs of Theorem 9 and 11 given the spectral theorems

Here we make a sketch of the passage from the spectral theory outlined in Section 4.2 to Theorems 9 and 11 via (4.4) and the techniques of Lalley from [Lal89]. Firstly, if there is no value $\beta_0$ such that $\lambda_{\beta_0} = 1$ then Proposition 39 together with Lemma 34 imply that the resolvent $(1 - L_s)^{-1}$ exists as a holomorphic family of bounded operators on $C^1(\Delta)$ in the region $\Re(s) > 1$. This would imply by standard contour shifting arguments in combination with (4.4) that for any $\eta > 0$\n\n$$N(w, a) = O(e^{(1+\eta)a}). \quad (4.9)$$\n
But this can be used along with the arguments of Section 3 to show for some $z$ in an infinite orbit of $\Lambda$ that\n\n$$N(w, a) = O(e^{(1+\eta)a}).$$\n
And this contradicts Baragar’s result (Theorem 2) when $\eta$ is small. Here we use the fact that for any $n$, there is an infinite orbit in $V(\mathbb{Z}_+)$ when $n = a$ and $k = 0$ coming from the tuple $(1, 1, \ldots, 1)$. In fact, for small $\eta$, (4.9) is already in contradiction to some of Baragar’s results from [Bar94a] on orbits of the linear semigroup $\Gamma$.

Now suppose there is such a $\beta_0 > 1$ as in Proposition 39. Then Lalley’s method of proof of his analog of Theorem 11 is by a contour shifting argument involving control on meromorphic behavior of $(1 - L_s)^{-1}$ in the following two ways:

1. By standard results in Linear Perturbation Theory [Kat76, Sections 4.3 and 7.1], Lemma 34 and Part 1 of Theorem 35 imply that the functions\n\n$$s \mapsto \lambda_s, \ s \mapsto h_s, \ s \mapsto \nu_s$$\n
extend to holomorphic functions on a neighborhood of the real line segment $(1, \infty)$ in $\Re(s) > 1$ such that\n\n$$\lambda_s \neq 0, \ L_s h_s = \lambda_s h_s, \ L_s^* \nu_s = \lambda_s \nu_s, \ \nu_s(h_s) = 1.$$\n
By suitable spectral decomposition of $L_s$, one finds a neighborhood $U$ of $s = \beta_0$ and an operator $L'_s$ such that $(1 - L'_s)^{-1}$ is a holomorphic family of bounded operators on $C^1(\Delta)$ for $s \in U$ and moreover\n\n$$(1 - L_s)^{-1} g = (1 - \lambda_s)^{-1} \nu_s(g) h_s + (1 - L'_s)^{-1} g$$\n
for $s \in U \setminus \{\beta_0\}$. This is the analog of [Lal89, Proposition 7.2].

2. By use of Theorem 37 along with its supplement Proposition 38, we obtain that\n\n$$s \mapsto (1 - L_s)^{-1}$$\n
is holomorphic in a neighborhood of every $s$ with $\Re(s) = \beta_0$, with the exception of $s = \beta_0$.

The outcome of Lalley’s argument is that\n\n$$N(w, a) = h_{\beta_0}(w) e^{\beta_0 a} + o(e^{\beta_0 a})$$\n
where the decay in the small $o$ does not depend on $w$. Our argument of Section 3.4 converts this into a version of Theorem 3 with $\beta$ replaced by $\beta_0$. Finally, this contradicts Baragar’s Theorem 2 unless $\beta = \beta_0$. Then Theorem 11 is proved, assuming the theorems of Section 4.2.

Theorem 9 is now a direct consequence of the following fact:
Lemma 40. For all \( \gamma \in \Gamma \) we have

\[
\frac{(\gamma \cdot w)_n}{w_n} = |\text{Jac}_w(\gamma)|^{-\frac{1}{n-1}}
\]

where \( |\text{Jac}_w(\gamma)| \) is the absolute value of the Jacobian determinant of \( \gamma \) acting on \( \Delta = \mathcal{H}/\mathbb{R}_+ \) at the point \( w \).

This can be checked by a direct calculation on general grounds as in [Pol, Lemma 2.1], or by using explicit formulae that appear later in this paper, e.g. by calculating the determinants of total derivatives we calculate in Section 5.

4.4 Consequences of uniformly contracting dynamics

The spectral theorems of the previous section all rely on the action of \( \Gamma' \) on \( \Delta \) being by contractions. That can be summarized in the following proposition.

Proposition 41. There are constants \( D > 0 \) and \( \rho < 1 \) such that for all \( \gamma^{(1)}, \gamma^{(2)}, \ldots, \gamma^{(N)} \in T\Gamma \) we have

\[
\|d_w[\gamma^{(1)} \gamma^{(2)} \ldots \gamma^{(N)}]\|_{\text{op}} \leq D\rho^N.
\]

Here we view \( \gamma^{(1)} \gamma^{(2)} \ldots \gamma^{(N)} \) as self-maps of \( \Delta \), using the fixed Euclidean metric on \( \Delta \), \( d_w \) is the total derivative of the map at \( w \in \Delta \), and \( \|\bullet\|_{\text{op}} \) is the operator norm of the map between tangent spaces (using the \( \ell^2 \) norms coming from the metric).

We will prove Proposition 41 in Section 5. The dynamical Proposition 41 gets brought into play by the following two-norm inequality with origins in the work of Ionescu Tulcea and Marinescu [ITM50].

Lemma 42. There is \( C > 0 \) such that for any \( Q \in \mathbb{N} \) and \( \Re(s) > 1 \)

\[
\|\nabla L_s^Q[f](w)\|_2 \leq C|s|L_s^Q[|f|(w)] + D\rho^Q L_s^Q[\|\nabla f\|_2](w)
\]

for all \( w \in \Delta \). We write \( \|\bullet\|_2 \) for the pointwise \( \ell^2 \) norm in an individual tangent space fiber.

Proof. This is standard given Proposition 41: it essentially boils down to the chain rule. The only thing to take care with are the infinite sums that arise, but these are all absolutely convergent when \( \Re(s) > 1 \).

We can now prove Lemma 34.

Proof of Lemma 34. We are proving \( s \mapsto L_s \) is a holomorphic mapping to bounded operators on \( C^1(\Delta) \). If we truncate the summation going into the expression (4.7) for \( L_s \) at some fixed \( L \) to form

\[
L_s^{(L)} = \sum_{j \in \mathbb{N} - 2} \sum_{A \leq L} \frac{1}{(1 + (A + 1)(1 - w_j))^s} f(\gamma_{n-1}^A \gamma_j \cdot w);
\]
the resulting \( L_s^{(L)} \) is easily seen to be holomorphic by taking a complex derivative. So it remains to show that \( L_s^{(L)} \to L_s \) uniformly on compact sets, say in the norm topology. But the tail consists of \( n - 2 \) terms of the form

\[
(L_s - L_s^{(L)})[f](w) = \sum_{A > L} \frac{1}{(1 + (A + 1)(1 - w_j))} f(\gamma_{n-1}^{A} \gamma_j w).
\]

Then \( \| L_s - L_s^{(L)} \|_{C^0} \to 0 \) as \( L \to \infty \) and this is uniform for \( s \) in \( W \), a compact subset of \( \Re(s) > 1 \). On the other hand, the proof of Lemma 42 also applies to \( L_s - L_s^{(L)} \), so applying it when \( Q = 1 \) gives

\[
\| \nabla (L_s - L_s^{(L)})[f] \|_\infty \leq C|s| \| (L_s - L_s^{(L)})[f] \|_\infty + D\rho \| (L_s - L_s^{(L)}) [\| \nabla f \|_2] \|_\infty.
\]

This implies

\[
\| L_s - L_s^{(L)} \|_{C^1(\Delta)} \ll W \| L_s - L_s^{(L)} \|_{C^0(\Delta)},
\]

which we’ve established goes to zero uniformly on \( W \).

The proof of the Ruelle-Perron-Frobenius Theorem 35 now proceeds either via use of Birkhoff cones as in Liverani’s paper [Liv95] or by a more direct approach as in Pollicott [Pol, Lemma 2.3]. The classical proof of this Theorem for subshifts of finite type can be found in [PP90, Theorem 2.2]. In any approach Lemma 42 is the key input. The uniform spectral gap stated in Part 1 of Theorem 35 is a consequence of the uniformity of Lemma 42 for \( s \) in a fixed compact subinterval of \( (1, \infty) \).

4.5 Behavior of the eigenvalue

In this section we prove Proposition 39. The statement that \( \lambda_s \) is real analytic on \( (1, \infty) \) follows from the fact we noted in the previous Section 4.2 that by perturbation theory in combination with Theorem 35 Part 1

\[
s \mapsto \lambda_s
\]

is holomorphic in a neighborhood of \( (1, \infty) \) in \( \Re(s) > 1 \). Notice that we have the bound

\[
L_s[f](w) = \sum_{j \in [n-2]} \sum_{A \in \mathbb{N}} \frac{1}{(1 + (A + 1)(1 - w_j))^s} f(\gamma_{n-1}^{A} \gamma_j w) \\
\leq (n - 2) \| f \|_\infty \sum_{A \in \mathbb{N}} \frac{1}{(1 + \frac{1}{2}(A + 1))^s} \leq 2(n - 2) \| f \|_\infty \sum_{A \in \mathbb{N}} \frac{1}{(3 + A)^s}.
\]

Letting \( f = h_s \) and \( w \) such that \( h_s(w) = \| h_s \|_\infty \) gives

\[
\lambda_s \leq 2(n - 2) \sum_{A \in \mathbb{N}} \frac{1}{(3 + A)^s}
\]

so \( \lambda_s \to 0 \) as \( s \to \infty \).

It remains to show that \( \lambda_s \) is strictly decreasing in \( s \). Let \( I \) be a fixed compact subinterval of \( (1, \infty) \). By Theorem 35 \( \lambda_{s-N} L_s^{N} \) converges in \( C^1 \) norm to \( h_s \) and this convergence is uniform for \( s \in I \). This implies

\[
\log \lambda_s = \frac{\log \left( L_s^N [1](w) \right)}{N} + o(1)
\]

(4.10)
where the error is uniform in $s \in I$ and $w \in \Delta$. We calculate

$$L_s^N[1](w) = \sum_{\gamma \in (T_\Gamma)^N} \left( \frac{\gamma . w}{w_n} \right)_n^{-s}, \quad \frac{d}{ds} L_s^N[1](w) = \sum_{\gamma \in (T_\Gamma)^N} -\log \left( \frac{\gamma . w}{w_n} \right) \left( \frac{\gamma . w}{w_n} \right)^{-s}.$$ 

Now we make the **Claim**: There is some $c > 0$ such that

$$\log \left( \frac{\gamma . w}{w_n} \right) \geq cN.$$

for all $\gamma \in (T_\Gamma)^N$. Assuming the Claim we get

$$\frac{d}{ds} L_s^N[1](w) \leq -cNL_s^N[1](w)$$

and hence

$$\frac{d}{ds} \log L_s^N[1](w) \leq -cN.$$

This means $\log \lambda_s$ is a uniform limit of functions with derivatives bounded above by a negative constant, so $\lambda_s$ must be strictly decreasing as required.

To prove the Claim it is enough to show (by expanding $\log(\gamma . w) - \log w_n$ as a telescoping sum) that for all $w \in \Delta$ and $\gamma' = \gamma_{n-1} \gamma_j \in T_\Gamma$

$$\frac{(\gamma'. w)_n}{w_n} = 1 + (A + 1)(1 - w_j) \geq c.$$

This is true with $c = 3/2$ since $w_j \leq 1/2$. This completes the proof of Proposition 39.

## 5 Proof of uniform contraction

In this section we prove Proposition 41 asserting that the elements of $T_\Gamma$ eventually uniformly contract $\Delta$.

### 5.1 Setup

We define the sets

$$\Delta \equiv \{(w_1, w_2, \ldots, w_{n-2}, w_{n-1}) : 0 \leq w_1 \leq w_2 \leq \ldots \leq w_{n-2} \leq w_{n-1} \leq 1, \sum_{i \in [n-1]} w_i = 1\},$$

$$\Delta_{\text{core}} \equiv \{(w_1, w_2, \ldots, w_{n-2}, w_{n-1}) \in \Delta : 0 \leq w_{n-1} - \sum_{j \in [n-2]} w_j \leq w_{n-2}\},$$

and

$$\Delta_{\text{cusp}} \equiv \{(w_1, w_2, \ldots, w_{n-2}, w_{n-1}) \in \Delta : w_{n-1} - \sum_{j \in [n-2]} w_j \geq w_{n-2}\}$$

where we use the notation $[N] = \{1, 2, \ldots, N\}$. We also define the set

$$\Delta_0 \equiv \Delta_{\text{core}} \cup \Delta_{\text{cusp}}.$$
Recall that the elements of $T_{\Gamma}$ are all of the form $\gamma = \gamma_{n-1}^L \gamma_i$ where $L \in \mathbb{N}$ and $i = 1, 2, \ldots, n-2$. Note that for each $w \in \Delta$, we have $\gamma_i(w) \in \Delta_{\text{core}}$ for $i = 1, 2, \ldots, n-2$ and $\gamma_{n-1}(w) \in \Delta_{\text{cusp}}$. In particular, $\gamma(w) \in \Delta_0$ for all $\gamma \in T_{\Gamma}$ and $w \in \Delta$.

From now on, we choose to use $n-2$ coordinates in $\Delta$ instead of $n-1$, using the relationship $w_{n-1} = 1 - \sum_{i \in [n-2]} w_i$.

Note that on $\Delta_0$ we have $\sum_{j \in [n-2]} w_j \leq \frac{1}{2}$ by combining the conditions that

$$w_{n-1} \geq \sum_{j \in [n-2]} w_j$$

and

$$1 - w_{n-1} = \sum_{j \in [n-2]} w_j.$$ 

Similarly, it is easy to show that on $\Delta_{\text{core}}$ we have $w_{n-2} \leq \frac{1}{2}$ and $w_{n-3} \leq \frac{1}{4}$, while on $\Delta_{\text{cusp}}$ we have $w_{n-2} \leq \frac{1}{3}$ and $w_{n-3} \leq \frac{1}{5}$.

**Remark 43.** It is clear that Proposition 41 can be proved with the local $\ell^2$ operator norms replaced by local $\ell^1$ norms, since the norms are equivalent, possibly at the expense of a larger $N$. 

43
5.2 Overview of the proof of Proposition 41

We will now prove Proposition 41 (the $\ell^1$ norm variant). We will appeal to the following bounds.

\[
\|d\gamma_i\|_1 \leq \frac{2}{2 - w_i} \leq \begin{cases} \frac{6}{5} & \text{on } \Delta_{\text{cusp}}, 1 \leq i \leq n - 3 \\
\frac{4}{5} & \text{on } \Delta_{\text{core}} \end{cases} \tag{5.1}
\]

\[
\|d\gamma_{n-1}\|_1 = \frac{1 + 2(w_1 + w_2 + \ldots + w_{n-2}) - 2w_1}{(1 + w_1 + w_2 + \ldots + w_{n-2})^2} \leq 1 \text{ on } \Delta_0 \tag{5.2}
\]

\[
\|d(\gamma_i \circ \gamma_j)\|_1 \leq \frac{2}{4 - 2w_j - w_i} \leq \frac{4}{5} \text{ on } \Delta_0, 1 \leq i < j \leq n - 2 \tag{5.3}
\]

\[
\|d(\gamma_i \circ \gamma_j)\|_1 \leq \frac{2}{4 - 2w_j - w_{i+1}} \leq \frac{4}{5} \text{ on } \Delta_0, 1 \leq j \leq i < n - 2 \tag{5.4}
\]

\[
\|d(\gamma_{n-2} \circ \gamma_j)\|_1 \leq \frac{4 + 2(w_1 + \ldots + w_{n-2}) - 2w_1 - 3w_j}{3 + (w_1 + \ldots + w_{n-2}) - 2w_j} \leq \frac{4}{5} \text{ on } \Delta_0, 1 \leq j \leq n - 2 \tag{5.5}
\]

\[
\|d(\gamma_{n-1} \circ \gamma_i)\|_1 \leq \frac{2}{3 - 2w_i} \leq \begin{cases} \frac{10}{13} & \text{on } \Delta_{\text{cusp}}, 1 \leq i \leq n - 3 \\
\frac{4}{5} & \text{on } \Delta_{\text{core}} \end{cases} \tag{5.6}
\]

\[
\|d(\gamma_{n-1} \circ \gamma_{n-2})\|_1 \leq \frac{2}{3 - 2w_{n-2}} \leq \begin{cases} \frac{6}{7} & \text{on } \Delta_{\text{cusp}} \\
1 & \text{on } \Delta_{\text{core}} \end{cases} \tag{5.7}
\]

\[
\|d(\gamma_i \circ \gamma_{n-1} \circ \gamma_{n-2})\|_1 \leq \frac{2}{6 - 4w_{n-2} - w_i} \leq \frac{4}{7} \text{ on } \Delta_0, 1 \leq i \leq n - 3 \tag{5.8}
\]

\[
\|d(\gamma_{n-2} \circ \gamma_{n-1} \circ \gamma_{n-2})\|_1 \leq \frac{7 + 2(w_1 + \ldots + w_{n-2}) - 2w_1 - 6w_{n-2}}{5 + (w_1 + \ldots + w_{n-2}) - 4w_{n-2}} \leq \frac{32}{49} \text{ on } \Delta_0 \tag{5.9}
\]

\[
\|d(\gamma_{n-1} \circ \gamma_{n-1} \circ \gamma_{n-2})\|_1 \leq \frac{2}{4 - 3w_{n-2}} \leq \begin{cases} \frac{3}{5} & \text{on } \Delta_{\text{cusp}} \\
\frac{4}{5} & \text{on } \Delta_{\text{core}} \end{cases} \tag{5.10}
\]

We will prove these bounds below by direct calculation. Using these bounds we can prove the following result for any $n \geq 3$ which implies Proposition 41 via Remark 43.

**Lemma 44.** Given the bounds (5.1)-(5.10), \( \|d(\gamma_{n-1}^L \circ \gamma_i \circ \gamma_{n-1}^K \circ \gamma_j)\|_{\Delta_0} \|_1 \leq \frac{24}{25} \) for each \( L, K \in \mathbb{N} \), and each \( i, j = 1, 2, \ldots, n - 2 \).

**Proof.** Throughout this proof, we repeatedly use the fact that \( \gamma_k(w) \in \Delta_{\text{core}} \) for \( k = 1, 2, \ldots, n-2 \) and \( \gamma_{n-1}(w) \in \Delta_{\text{cusp}} \). We distinguish 3 cases.

**Case I: \( L \geq 1, K \geq 1 \):**

Using equations (5.2), (5.6) and (5.7), we have

\[
\|d(\gamma_{n-1}^L \circ \gamma_i \circ \gamma_{n-1}^K \circ \gamma_j)\|_{\Delta_0} \|_1 \leq \|d_{\gamma_{n-1}}^{L-1}\|_{\Delta_{\text{cusp}}} \|_1 \|d(\gamma_{n-1} \circ \gamma_i)\|_{\Delta_0} \|_1 \|d_{\gamma_{n-1}}^{K-1}\|_{\Delta_{\text{cusp}}} \|_1 \|d(\gamma_{n-1} \circ \gamma_j)\|_{\Delta_0} \|_1 \leq 1 \cdot \frac{6}{7} \cdot 1 \cdot 1 \cdot \frac{24}{25} < \frac{24}{25}.
\]

**Case II: \( L \geq 0, K = 0 \):**

Using equations (5.2), (5.3), (5.4), (5.5), we have

\[
\|d(\gamma_{n-1}^L \circ \gamma_i \circ \gamma_j)\|_{\Delta_0} \|_1 \leq \|d_{\gamma_{n-1}}^{L-1}\|_{\Delta_{\text{cusp}}} \|_1 \|d(\gamma_i \circ \gamma_j)\|_{\Delta_0} \|_1 \leq 1 \cdot \frac{4}{5} < \frac{24}{25}.
\]
Case III: $L = 0, K \geq 1$:
We first suppose that $j \leq n - 3$. Then by equations (5.1), (5.2), (5.6) we have
$$
\left\| d(\gamma_i \circ \gamma_{n-1} \circ \gamma_j) |_{\Delta_0} \right\|_1 \leq \left\| d\gamma_i |_{\Delta_{cusp}} \right\|_1 \left\| d\gamma_{n-1} |_{\Delta_{cusp}} \right\|_1 \left\| d(\gamma_{n-1} \circ \gamma_j) |_{\Delta_0} \right\|_1 \leq \frac{6}{5} \cdot 1 \cdot \frac{4}{5} = \frac{24}{25}.
$$
Finally, if $j = n - 2$ we are left with two subcases. If $K = 1$, then by equations (5.8) and (5.9) we have
$$
\left\| d(\gamma_i \circ \gamma_{n-1} \circ \gamma_{n-2}) |_{\Delta_0} \right\|_1 \leq \frac{32}{49} < \frac{24}{25}.
$$
Otherwise, we have $K \geq 2$ and by equations (5.1), (5.2), (5.10) we have
$$
\left\| d(\gamma_i \circ \gamma_{n-1} \circ \gamma_{n-2}) |_{\Delta_0} \right\|_1 \leq \frac{6}{5} \cdot 1 \cdot \frac{4}{5} = \frac{24}{25}.
$$
In the remainder of this section we prove equations (5.1)-(5.10) by direct calculation. In all following sections, we define
$$
w \equiv (w_1, w_2, \ldots, w_{n-2}, 1 - \sum_{k=1}^{n-2} w_k, 1),
$$
and
$$
\beta(w) \equiv \sum_{k=1}^{n-2} w_k.
$$
Also recall that the $\|\cdot\|_1$ of a matrix is equal to the maximum over columns of the matrix of the sum of the absolute values of the column. From now on, we call such a sum an absolute column sum.

5.3 Proof of equations (5.1)-(5.10)

Proof of equation (5.1)

For $i = 1, 2, \ldots, n - 2$ we have
$$
\gamma_i(w) = (w_1, w_2, \ldots, \widehat{w_i}, \ldots, w_{n-2}, 1 - \beta(w), 1, 2 - w_i),
$$
which, after projectivizing and removing the placeholder components, gives
$$
\gamma_i(w) = \left( \frac{w_1}{2 - w_i}, \frac{w_2}{2 - w_i}, \ldots, \frac{\widehat{w_i}}{2 - w_i}, \ldots, \frac{w_{n-2}}{2 - w_i}, \frac{1 - \beta(w)}{2 - w_i} \right)
$$
which is a function in $(n - 2)$ variables with $(n - 2)$ components. The $(n - 2) \times (n - 2)$ total derivative $d\gamma_i$ is given by the following matrix
where the row and column indices are indicated to the left and above respectively. Each of these partial derivatives is immediate, except for the \((n-2, i)\) entry which follows from an application of the quotient rule. Note that the sign of entry \((n-2, i)\) is negative on \(\Delta_0\). The signs of the other entries are self-evident.

The absolute column sum for each column \(k\) with \(k \neq i\) is

\[ C_k = \frac{2}{2 - w_i} \]

For column \(k = i\) the absolute column sum is

\[ C_i = \frac{1 + 2\beta(w) - 2w_i}{(2 - w_i)^2} \]

We must compute which absolute column sum is maximal on \(\Delta_0\). Note that on \(\Delta_0\) we have \(\beta(w) \leq \frac{1}{2}\). Furthermore, we have the following equivalences:

\[ C_i \leq C_k, k \neq i \quad \Leftrightarrow \quad 1 + 2\beta(w) - 2w_i < 4 - 2w_i \quad \Leftrightarrow \quad \beta(w) < \frac{2}{3} \]

Any column \(k \neq i\) is maximal and \(\|d\gamma_i\|_1 \leq \frac{2}{2 - w_i}\) on \(\Delta_0\). For each \(i\), we have \(w_i \leq \frac{1}{3}\) on \(\Delta_{\text{cusp}}\), and \(w_i \leq \frac{1}{2}\) on \(\Delta_{\text{core}}\). This gives the bound that \(\|d\gamma_i\|_1 \leq \frac{6}{5}\) on \(\Delta_{\text{cusp}}\) and \(\leq \frac{4}{3}\) on \(\Delta_{\text{core}}\), proving equation (5.1).

**Proof of equation (5.2)**

We have

\[ \gamma_{n-1}(w) = (w_1, w_2, \ldots, w_{n-2}, 1, 1 + \beta(w)) \]

which after projectivizing and removing placeholder components becomes

\[ \gamma_{n-1}(w) = \left( \frac{w_1}{1 + \beta(w)}, \frac{w_2}{1 + \beta(w)}, \ldots, \frac{w_{n-2}}{1 + \beta(w)} \right) \]
The \((n - 2) \times (n - 2)\) total derivative \(d\gamma_{n-1}\) is given by the following matrix

\[
\begin{pmatrix}
1 & 2 & 3 & \ldots & n-2 \\
\frac{1}{(1 + \beta(w))^2} & \frac{-w_1}{1 + \beta(w)} & \frac{-w_1}{1 + \beta(w)} & \cdots & \frac{-w_1}{1 + \beta(w)} \\
\frac{-w_2}{(1 + \beta(w))^2} & \frac{1 + \beta(w) - w_2}{(1 + \beta(w))^2} & \frac{-w_2}{(1 + \beta(w))^2} & \cdots & \frac{-w_2}{(1 + \beta(w))^2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{-w_{n-2}}{(1 + \beta(w))^2} & \frac{-w_{n-2}}{(1 + \beta(w))^2} & \frac{-w_{n-2}}{(1 + \beta(w))^2} & \cdots & \frac{1 + \beta(w) - w_{n-2}}{(1 + \beta(w))^2}
\end{pmatrix}
\]

For each column \(k = 1, 2, \ldots, n - 2\) we have the absolute column sum

\[C_k = \frac{1 + 2\beta(w) - 2w_k}{(1 + \beta(w))^2}.
\]

Since \(w_1 \leq w_2 \leq \ldots \leq w_{n-2}\) on \(\Delta_0\), we have column \(C_1\) is maximal. Hence, \(|d\gamma_{n-1}|_1 = \frac{1 + 2\beta(w) - 2w_1}{(1 + \beta(w))^2}\) on \(\Delta_0\).

To bound this norm, observe

\[
\frac{1 + 2\beta(w) - 2w_1}{(1 + \beta(w))^2} = \frac{1 + 2\beta(w) - 2w_1}{1 + 2\beta(w) + \beta(w)^2} \leq \frac{1 + 2\beta(w) - 2w_1}{1 + 2\beta(w) - 2w_1} = 1.
\]

Hence \(|d\gamma_{n-1}|_1 \leq 1\) on \(\Delta_0\), proving equation (5.2).

**Proof of equations (5.3) and (5.4)**

We first prove equation (5.3). Assume first that \(1 \leq i < j \leq n - 2\). We have

\[
\gamma_i \circ \gamma_j(w) = (w_1, w_2, \ldots, \hat{w}_i, \ldots, \hat{w}_j, \ldots, w_{n-2}, 1 - \beta(w), 1, 2 - w_j, 4 - 2w_j - w_i).
\]

Define \(\psi(w) \equiv 4 - 2w_j - w_i\). Then, after projectivizing and removing placeholder components we have

\[
\gamma_i \circ \gamma_j(w) = \left(\frac{w_1}{\psi(w)}, \frac{w_2}{\psi(w)}, \ldots, \frac{\hat{w}_i}{\psi(w)}, \ldots, \frac{\hat{w}_j}{\psi(w)}, \ldots, \frac{w_{n-2}}{\psi(w)}, \frac{1 - \beta(w)}{\psi(w)}, \frac{1}{\psi(w)}\right).
\]

The \((n - 2) \times (n - 2)\) total derivative \(d(\gamma_i \circ \gamma_j)\) is given by the matrix
and the latter inequality holds since 

\[ \beta \]

Thus 

\[ \gamma \]

which is nonnegative on \( \Delta_0 \). Thus 

\[ C_i \]

and the absolute column sum of row \( j \) is 

\[ C_j \]

Subtracting \( C_i \) from \( C_j \) we obtain 

\[ \frac{2\beta(w) - w_j - w_i}{(\psi(w))^2} \]

which is nonnegative on \( \Delta_0 \). Thus 

\[ C_j \geq C_i \]

Furthermore \( C_k \geq C_j \) for each \( k \neq i, j \) since 

\[ 4 + 4\beta(w) - 4w_j - 3w_i \leq 2(4 - 2w_j - w_i) \]

and the latter inequality holds since \( \beta(w) \leq \frac{1}{2} \) and \( w_i \geq 0 \) on \( \Delta_0 \).

Thus 

\[ \gamma_i \circ \gamma_j \]

on \( \Delta_0 \). Using the bound that each \( w_k \leq \frac{1}{2} \) on \( \Delta_0 \) we have 

\[ ||d(\gamma_i \circ \gamma_j)||_1 \leq \frac{4}{5} \] proving equation (5.3).

To prove equation (5.4), we now consider \( \gamma_i \circ \gamma_j \) for \( 1 \leq j \leq i < n - 2 \). We have 

\[ \gamma_i \circ \gamma_j (w) = (w_1, w_2, \ldots, w_j, \ldots, \hat{w}_{i+1}, \ldots, w_{n-2}, 1 - \beta(w), 1, 2 - w_j, 4 - 2w_j - w_{i+1}) \]

The remainder of the proof for equation (5.4) is nearly identical after careful bookkeeping of indices (for example, column \( C_{i+1} \) for \( 1 \leq j \leq i < n - 2 \) plays the role of \( C_i \) for \( 1 \leq i < j \leq n - 2 \)).
Proof of equation (5.5)

We have

\[ \gamma_{n-2} \circ \gamma_j(w) = (w_1, w_2, \ldots, \hat{w}_j, \ldots, w_{n-2}, 1, 2 - w_j, 1) \]

We define \( \kappa(w) \equiv 3 + \beta(w) - 2w_j \). Then, after projectivizing and removing placeholder components, we have

\[ \gamma_{n-2} \circ \gamma_j(w) = \left( \frac{w_1}{\kappa(w)}, \frac{w_2}{\kappa(w)}, \ldots, \frac{w_j}{\kappa(w)}, \ldots, \frac{w_{n-2}}{\kappa(w)}, \frac{1}{\kappa(w)} \right). \]

The \((n-2) \times (n-2)\) total derivative, \(d(\gamma_{n-2} \circ \gamma_j)\) is the matrix

\[
\begin{bmatrix}
1 & -w_1 & 2 & \cdots & j-1 & w_1 & j & j+1 & \cdots & n-2 \\
3+\beta(w)-2w_j-w_1 & -w_2 & \cdots & (\kappa(w))^{j-1} & (\kappa(w))^{j+1} & \cdots & (\kappa(w))^2 \\
(\kappa(w))^2 & 3+\beta(w)-2w_j-w_2 & \cdots & (\kappa(w))^2 & (\kappa(w))^2 & \cdots & (\kappa(w))^2 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
(\kappa(w))^2 & (\kappa(w))^{j-1} & \cdots & (\kappa(w))^{j+1} & (\kappa(w))^2 & \cdots & (\kappa(w))^2 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
(\kappa(w))^2 & (\kappa(w))^{j-1} & \cdots & (\kappa(w))^{j+1} & (\kappa(w))^2 & \cdots & (\kappa(w))^2 \\
(\kappa(w))^2 & (\kappa(w))^{j-1} & \cdots & (\kappa(w))^{j+1} & (\kappa(w))^2 & \cdots & (\kappa(w))^2 \\
\end{bmatrix}
\]

The absolute column sum for column \(k \neq j\) is

\[ C_k = \frac{4 + 2\beta(w) - 3w_j - 2w_k}{(\kappa(w))^2} \]

and the absolute column sum for column \(j\) is

\[ C_j = \frac{1 + \beta(w) - w_j}{(\kappa(w))^2}. \]

Note that \(w_1 \leq w_k\) for all \(k\), so \(C_1 \leq C_k\) for each \(k \neq j\). Furthermore, subtracting column sum \(C_j\) from \(C_1\) and using the trivial bound \(w_k \leq \frac{1}{2}\) on \(\Delta_0\) for all \(k\) we obtain

\[ \frac{3 + \beta(w) - 2w_j - 2w_1}{(\kappa(w))^2} \geq \frac{1 + \beta(w)}{(\kappa(w))^2} \geq 0. \]

Hence, \(C_1 \geq C_j\), and \(C_1\) is maximal. We have \(\|d(\gamma_{n-2} \circ \gamma_j)\|_1 \leq \frac{4 + 2\beta(w) - 3w_j - 2w_1}{(\kappa(w))^2}\) on \(\Delta_0\). Separately bounding the numerator and denominator on \(\Delta_0\) we have

\[ 4 + 2\beta(w) - 3w_j - 2w_i \leq 4 + 2\left(\frac{1}{2}\right) = 5, \]

\[ \kappa(w) = 3 + \beta(w) - 2w_j \geq 3 + (\beta(w) - w_j) - w_j \geq 3 - \frac{1}{2} = \frac{5}{2}. \]

Thus \(\|d(\gamma_{n-2} \circ \gamma_j)\|_1 \leq \frac{5}{\left(\frac{5}{2}\right)^2} = \frac{4}{5}\) on \(\Delta_0\). This proves equation (5.5).
Proof of equation (5.6) and (5.7)

For \( i = 1, 2, \ldots, n - 2 \) we have

\[
\gamma_{n-1} \circ \gamma_i(w) = (w_1, w_2, \ldots, \widehat{w_i}, \ldots, w_{n-2}, 1 - \beta(w), 2 - w_i, 3 - 2w_i)
\]

which, after projectivizing and removing placeholder components, becomes

\[
\gamma_{n-1} \circ \gamma_i(w) = \left( \frac{w_1}{3 - 2w_i}, \frac{w_2}{3 - 2w_i}, \ldots, \frac{w_i}{3 - 2w_i}, \ldots, \frac{w_{n-2}}{3 - 2w_i}, \frac{1 - \beta(w)}{3 - 2w_i} \right).
\]

The \((n - 2) \times (n - 2)\) total derivative \(d(\gamma_{n-1} \circ \gamma_i)\) is given by

\[
\frac{1}{3 - 2w_i} \begin{bmatrix}
1 & 2 & \ldots & i-1 & i & i+1 & \ldots & n-2 \\
0 & 0 & \ldots & 0 & \frac{2w_1}{(3 - 2w_i)^2} & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \frac{2w_{i-1}}{3 - 2w_i} & \frac{(3 - 2w_i)^2}{2w_{i+1}} & 1 & \ldots & 0 \\
0 & 0 & \ldots & 0 & \frac{(3 - 2w_i)^2}{2w_{n-2}} & 0 & \ldots & \frac{1}{3 - 2w_i} \\
\end{bmatrix}.
\]

The absolute column sum for column \( k \neq i \) is

\[
C_k = \frac{2}{3 - 2w_i}
\]

and the absolute column sum for column \( i \) is

\[
C_i = \frac{1 + 4\beta(w) - 4w_i}{(3 - 2w_i)^2}
\]

Each column \( C_k \) with \( k \neq i \) is maximal since

\[
C_k \geq C_i \quad \Leftrightarrow \quad 2(3 - 2w_i) \geq 1 + 4\beta(w) - 4w_i \quad \Leftrightarrow \quad 5 \geq 4\beta(w)
\]

and \( \beta(w) \leq \frac{1}{2} \) on \( \Delta_0 \). Thus \( \|d(\gamma_{n-1} \circ \gamma_i)\|_1 \leq \frac{2}{3 - 2w_i} \) on \( \Delta_0 \).

When \( i = n - 2 \), we have \( w_{n-2} \leq \frac{1}{2} \) on \( \Delta_{\text{core}} \) and \( w_{n-2} \leq \frac{1}{3} \) on \( \Delta_{\text{cusp}} \). Thus \( \|d(\gamma_{n-1} \circ \gamma_{n-2})\|_1 \leq 1 \) on \( \Delta_{\text{core}} \) and \( \leq \frac{6}{7} \) on \( \Delta_{\text{cusp}} \). This proves equation (5.7).

For \( i \leq n - 3 \), we have the stronger bound \( w_i \leq \frac{1}{3} \) on \( \Delta_{\text{core}} \) and \( w_i \leq \frac{1}{5} \) on \( \Delta_{\text{cusp}} \). This gives \( \|d(\gamma_{n-1} \circ \gamma_i)\|_1 \leq \frac{4}{3} \) on \( \Delta_{\text{core}} \) and \( \leq \frac{10}{13} \) on \( \Delta_{\text{cusp}} \). This proves equation (5.6).
Proof of equation (5.8)
For each $i \leq n - 3$ we have

$$\gamma_i \circ \gamma_{n-1} \circ \gamma_{n-2}(w) = (w_1, w_2, \ldots, \hat{w_i}, \ldots, w_{n-3}, 1 - \beta(w), 2 - w_{n-2}, 3 - 2w_{n-2}, 6 - 4w_{n-2} - w_i)$$

which, after projectivizing and removing placeholder components, becomes

$$\gamma_i \circ \gamma_{n-1} \circ \gamma_{n-2}(w) = \left( \frac{w_1}{\mu(w)}, \ldots, \frac{\hat{w_i}}{\mu(w)}, \frac{w_{n-3}}{\mu(w)}, \frac{1 - \beta(w)}{\mu(w)}, \frac{2 - w_{n-2}}{\mu(w)} \right)$$

where $\mu(w) \equiv 6 - 4w_{n-2} - w_i$. Then the $(n-2) \times (n-2)$ total derivative $d(\gamma_i \circ \gamma_{n-1} \circ \gamma_{n-2})$ is

$$
\begin{bmatrix}
1 & \ldots & i-1 & i & i+1 & \ldots & n-3 & n-2 \\
\frac{1}{\mu(w)} & \ldots & 0 & \frac{w_1}{(\mu(w))^2} & 0 & \ldots & 0 & \frac{4w_1}{(\mu(w))^2} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & \ldots & \frac{1}{\mu(w)} & \frac{w_{i-1}}{(\mu(w))^2} & 0 & \ldots & 0 & \frac{4w_{i-1}}{(\mu(w))^2} \\
0 & \ldots & 0 & \frac{1}{(\mu(w))^2} & \mu(w) & \ldots & 0 & \frac{4w_{i+1}}{(\mu(w))^2} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & \ldots & 0 & \frac{w_{n-3}}{(\mu(w))^2} & 0 & \ldots & \frac{1}{\mu(w)} & \frac{4w_{n-3}}{(\mu(w))^2} \\
\frac{1}{\mu(w)} & \ldots & \frac{1}{\mu(w)} & \frac{-5\beta(w) + 4w_{n-2} + w_i}{(\mu(w))^2} & \frac{-1}{\mu(w)} & \ldots & \frac{-1}{\mu(w)} & \frac{-2 - 4\beta(w) + 4w_{n-2} + w_i}{(\mu(w))^2} \\
0 & \ldots & 0 & \frac{2 - w_{n-2}}{(\mu(w))^2} & 0 & \ldots & 0 & \frac{2 + w_i}{(\mu(w))^2}
\end{bmatrix}
$$

The absolute column sum of column $k \neq i, n - 2$ is

$$C_k = \frac{2}{\mu(w)} = \frac{2(6 - 4w_{n-2} - w_i)}{(\mu(w))^2},$$

whereas the absolute column sum of column $i$ is

$$C_i = \frac{7 + 2\beta(w) - 2w_i - 6w_{n-2}}{(\mu(w))^2}$$

and the absolute column sum of column $n - 2$ is

$$C_{n-2} = \frac{4 + 8\beta(w) - 8w_{n-2} - 4w_i}{(\mu(w))^2}.$$

Subtracting $C_{n-2}$ from $C_i$ we obtain

$$\frac{3 - 6\beta(w) + 2w_{n-2} + 2w_i}{(\mu(w))^2} \geq \frac{3 - 6(\frac{1}{7})}{(\mu(w))^2} \geq 0$$

on $\Delta_0$. This shows $C_i \geq C_{n-2}$.
In fact, each column $C_k$ with $k \neq i, n-2$ is maximal since
$$C_k \geq C_i, k \neq i, n-2 \iff 2(6-4w_{n-2}-w_i) \geq 7+2\beta(w)-2w_i-6w_{n-2} \iff 5 \geq 2\beta(w)+2w_{n-2}$$
and $\beta(w)$, $w_{n-2} \leq \frac{1}{2}$ on $\Delta_0$. Hence $\|d(\gamma_i \circ \gamma_{n-1} \circ \gamma_{n-2})\|_1 \leq \frac{2}{6-4w_{n-2}-w_i}$ on $\Delta_0$. The denominator is bounded by
$$6 - 4w_{n-2} - w_i \geq 6 - 5\left(\frac{1}{2}\right) = \frac{7}{2}$$
so $\|d(\gamma_i \circ \gamma_{n-1} \circ \gamma_{n-2})\|_1 \leq \frac{4}{7}$ on $\Delta_0$. This proves equation (5.8).

**Proof of equation (5.9)**

We have
$$\gamma_{n-2} \circ \gamma_{n-1} \circ \gamma_{n-2}(w) = (w_1, w_2, \ldots, w_{n-3}, 2-w_{n-2}, 3-2w_{n-2}, 5+\beta(w)-4w_{n-2})$$
which, after projectivizing and removing placeholder components, becomes
$$\gamma_{n-2} \circ \gamma_{n-1} \circ \gamma_{n-2}(w) = \left(\frac{w_1}{\theta(w)}, \frac{w_2}{\theta(w)}, \ldots, \frac{w_{n-3}}{\theta(w)}, \frac{2-w_{n-2}}{\theta(w)}\right)$$
where $\theta(w) := 5+\beta(w)-4w_{n-2}$. Then the $(n-2) \times (n-2)$ total derivative $d(\gamma_{n-2} \circ \gamma_{n-1} \circ \gamma_{n-2})$ is

$$
\begin{bmatrix}
1 & 2 & 3 & \cdots & n-3 & n-2 \\
\frac{5+\beta(w)-4w_{n-2}-w_1}{(\theta(w))^2} & -\frac{w_1}{(\theta(w))^2} & -\frac{w_1}{(\theta(w))^2} & \cdots & -\frac{w_1}{(\theta(w))^2} & -\frac{3w_1}{(\theta(w))^2} \\
-\frac{w_2}{(\theta(w))^2} & \frac{5+\beta(w)-4w_{n-2}-w_2}{(\theta(w))^2} & \frac{w_2}{(\theta(w))^2} & \cdots & \frac{w_2}{(\theta(w))^2} & \frac{3w_2}{(\theta(w))^2} \\
-\frac{w_3}{(\theta(w))^2} & \frac{5+\beta(w)-4w_{n-2}-w_3}{(\theta(w))^2} & \frac{w_3}{(\theta(w))^2} & \cdots & \frac{w_3}{(\theta(w))^2} & \frac{3w_3}{(\theta(w))^2} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
n-3 & -\frac{w_{n-3}}{(\theta(w))^2} & -\frac{w_{n-3}}{(\theta(w))^2} & -\frac{w_{n-3}}{(\theta(w))^2} & \cdots & \frac{3w_{n-3}}{(\theta(w))^2} \\
-\frac{2-w_{n-2}}{(\theta(w))^2} & -\frac{2-w_{n-2}}{(\theta(w))^2} & -\frac{2-w_{n-2}}{(\theta(w))^2} & \cdots & \frac{3w_{n-3}}{(\theta(w))^2} & \frac{1+\beta(w)+w_{n-2}}{(\theta(w))^2}
\end{bmatrix}
$$

The absolute column sum for each column $k \leq n-3$ is
$$C_k = \frac{7 + 2\beta(w) - 2w_k - 6w_{n-2}}{(\theta(w))^2}$$
and the absolute column sum for column $n-2$ is
$$C_{n-2} = \frac{1 + 2\beta(w) - 2w_{n-2}}{(\theta(w))^2}.$$
Subtracting $C_{n-2}$ from $C_k$ we obtain
\[
\frac{6 - 2w_k - 4w_{n-2}}{(\theta(w))^2} \geq 0
\]
on $\Delta_0$ (using the bound $w_k \leq \frac{1}{2}$ for all $k$). Hence, $C_j \leq C_k$ for each $k \neq j$. Of the remaining column sums, $C_1$ is maximal since $w_1 \leq w_k$ for each $k$ on $\Delta_0$. Thus, \[\|d(\gamma_{n-2} \circ \gamma_{n-1} \circ \gamma_{n-2})\|_1 \leq \frac{7 + 2\theta(w) - 2w_1 - 6w_{n-2}}{(5 + \theta(w) - 4w_{n-2})^2}.
\]
Separately bounding the numerator and the denominator we have
\[7 + 2\beta(w) - 2w_1 - 6w_{n-2} \leq 7 + 1 \leq 8
\]
\[5 + \beta(w) - 4w_{n-2} \geq 5 + (\beta(w) - w_{n-2}) - 3w_{n-2} \geq 5 - 3\left(\frac{1}{2}\right) = \frac{7}{2}
\]
so \[\|d(\gamma_{n-2} \circ \gamma_{n-1} \circ \gamma_{n-2})\|_1 \leq \frac{32}{49}\] on $\Delta_0$, proving equation (5.9).

**Proof of equation (5.10)**

We have
\[\gamma_{n-1} \circ \gamma_{n-1} \circ \gamma_{n-2}(w) = (w_1, w_2, \ldots, w_{n-3}, 1 - \beta(w), 3 - 2w_{n-2}, 4 - 3w_{n-2})
\]
which, after projectivizing and removing placeholder components, becomes
\[\gamma_{n-1} \circ \gamma_{n-1} \circ \gamma_{n-2}(w) = \left(\frac{w_1}{4 - 3w_{n-2}}, \frac{w_2}{4 - 3w_{n-2}}, \ldots, \frac{w_{n-3}}{4 - 3w_{n-2}}, \frac{1 - \beta(w)}{4 - 3w_{n-2}}\right).
\]
The $(n - 2) \times (n - 2)$ total derivative $d(\gamma_{n-1} \circ \gamma_{n-1} \circ \gamma_{n-2})$ is
\[
\begin{bmatrix}
1 & 2 & \ldots & n-3 & n-2 \\
1 & 0 & \ldots & 0 & 3w_1 \\
0 & 1 & \ldots & 0 & (4 - 3w_{n-2})^2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
n-3 & 0 & \ldots & 1 & 3w_{n-3} \\
n-2 & -1 & \ldots & -1 & (4 - 3w_{n-2})^2 \\
4 - 3w_{n-2} & 4 - 3w_{n-2} & \ldots & 4 - 3w_{n-2} & (4 - 3w_{n-2})^2 \\
4 - 3w_{n-2} & 4 - 3w_{n-2} & \ldots & 4 - 3w_{n-2} & (4 - 3w_{n-2})^2 \\
4 - 3w_{n-2} & 4 - 3w_{n-2} & \ldots & 4 - 3w_{n-2} & (4 - 3w_{n-2})^2
\end{bmatrix}
\]
The absolute column sum for each column $k \leq n - 3$ is
\[C_k = \frac{2}{4 - 3w_{n-2}}
\]
and the absolute column sum for column $n - 2$ is
\[C_{n-2} = \frac{1 + 6\beta(w) - 6w_{n-2}}{(4 - 3w_{n-2})^2}.\]
Each column $C_k$ for $k \neq n - 2$ is maximal because
\[C_k \geq C_{n-2}, k \neq n - 2 \quad \iff \quad 2(4 - 3w_{n-2}) \geq 1 + 6\beta(w) - 6w_{n-2} \quad \iff \quad 7 \geq 6\beta(w).
\]
Thus, \[\|d(\gamma_{n-1} \circ \gamma_{n-1} \circ \gamma_{n-2})\|_1 \leq \frac{2}{4 - 3w_{n-2}}\] on $\Delta_0$. Since $w_{n-2} \leq \frac{1}{2}$ on $\Delta_0$ we have
\[\|d(\gamma_{n-1} \circ \gamma_{n-1} \circ \gamma_{n-2})\|_1 = \frac{4}{5},\] proving equation (5.10).
References

[Aig13] M. Aigner. *Markov’s theorem and 100 years of the uniqueness conjecture*. Springer, Cham, 2013. A mathematical journey from irrational numbers to perfect matchings.

[Bal00] V. Baladi. *Positive transfer operators and decay of correlations*, volume 16 of *Advanced Series in Nonlinear Dynamics*. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[Bar94a] A. Baragar. Asymptotic growth of Markoff-Hurwitz numbers. *Compositio Math.*, 94(1):1–18, 1994.

[Bar94b] A. Baragar. Integral solutions of Markoff-Hurwitz equations. *J. Number Theory*, 49(1):27–44, 1994.

[Bar98] A. Baragar. The exponent for the Markoff-Hurwitz equations. *Pacific J. Math.*, 182(1):1–21, 1998.

[Bel01] G.V. Belyi˘ı. Markov’s numbers and quadratic forms. *J. Math. Sci., New York*, 106(4):3087–3097, 2001.

[Bom07] E. Bombieri. Continued fractions and the Markoff tree. *Expo. Math.*, 25(3):187–213, 2007.

[Boy71] D.W. Boyd. The disk-packing constant. *Aequationes Math.*, 7:182–193, 1971.

[Boy73] D.W. Boyd. Improved bounds for the disk-packing constant. *Aequationes Math.*, 9:99–106, 1973.

[Boy82] D.W. Boyd. The sequence of radii of the Apollonian packing. *Math. Comp.*, 39(159):249–254, 1982.

[Cas57] J.W.S. Cassels. *An introduction to Diophantine approximation*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 45. Cambridge University Press, New York, 1957.

[Gau] C.F. Gauss. Brief an Laplace vom 30 Jan. 1812 Werke X1 pp 371-374.

[Gol03] W.M. Goldman. The modular group action on real SL(2)-characters of a one-holed torus. *Geom. Topol.*, 7:443–486, 2003.

[GS17] A. Ghosh and P. Sarnak. Integral points on Markoff type cubic surfaces. *arXiv://1706.06712 v1*, October 2017.

[Gur76] C. Gurwood. *Diophantine approximation and the Markov chain*. PhD thesis, New York University, 1976.

[HN13] Y. Huang and P. Norbury. Simple geodesics and Markoff quads. *arXiv:1312.7089*, December 2013.

[Hor75] R.D. Horowitz. Induced automorphisms on Fricke characters of free groups. *Transactions of the American Mathematical Society*, 208:41–50, 1975.
[HPZ15] H. Hu, S. Peow Tan, and Y. Zhang. Polynomial automorphisms of $C^n$ preserving the Markoff-Hurwitz polynomial. arXiv://1501.06955, January 2015.

[Hur07] A. Hurwitz. Über eine Aufgabe der unbestimmten Analysis. Archiv. Math. Phys., 3:185–196, 1907. Also: Mathematisch Werke, Vol. 2, Chapt. LXX (1933 and 1962), 410–421.

[ITM50] C. T. Ionescu Tulcea and G. Marinescu. Théorie ergodique pour des classes d’opérations non complètement continues. Ann. of Math. (2), 52:140–147, 1950.

[Kat76] T. Kato. Perturbation theory for linear operators. Springer-Verlag, Berlin-New York, second edition, 1976. Grundlehren der Mathematischen Wissenschaften, Band 132.

[KO11] A. Kontorovich and H. Oh. Apollonian circle packings and closed horospheres on hyperbolic 3-manifolds. J. Amer. Math. Soc., 24(3):603–648, 2011. With an appendix by Oh and Nimish Shah.

[Kuz32] R.O. Kuzmin. On a problem of Gauss. Atti del Congresso Internazionale dei Matematici, Bologna, 6:83–89, 1932.

[Lal88] S.P. Lalley. The packing and covering functions of some self-similar fractals. Indiana Univ. Math. J., 37(3):699–710, 1988.

[Lal89] S.P. Lalley. Renewal theorems in symbolic dynamics, with applications to geodesic flows, non-Euclidean tessellations and their fractal limits. Acta Math., 163(1-2):1–55, 1989.

[Liv95] C. Liverani. Decay of correlations. Ann. of Math. (2), 142(2):239–301, 1995.

[Mag17] M. Magee. Counting one sided simple closed geodesics on Fuchsian thrice punctured projective planes. arXiv:1705.09377, May 2017.

[Mar80] A. Markoff. Sur les formes quadratiques binaires indéfinies. Math. Ann., 17(3):379–399, 1880.

[McS91] G. McShane. A remarkable identity for lengths of curves. ProQuest LLC, Ann Arbor, MI, 1991. Thesis (Ph.D.)–University of Warwick (United Kingdom).

[McS98] G. McShane. Simple geodesics and a series constant over Teichmüller space. Invent. Math., 132(3):607–632, 1998.

[Mir08] M. Mirzakhani. Growth of the number of simple closed geodesics on hyperbolic surfaces. Ann. of Math. (2), 168(1):97–125, 2008.

[Mir16] M. Mirzakhani. Counting Mapping Class group orbits on hyperbolic surfaces. arXiv:1601.03342, January 2016.

[Mor53] L. J. Mordell. On the integer solutions of the equation $x^2 + y^2 + z^2 + 2xyz = n$. J. London Math. Soc., 28:500–510, 1953.

[MR95] G. McShane and I. Rivin. A norm on homology of surfaces and counting simple geodesics. Internat. Math. Res. Notices, (2):61–69 (electronic), 1995.
[Pat76] S. J. Patterson. The limit set of a Fuchsian group. *Acta Math.*, 136(3-4):241–273, 1976.

[Pol] M. Pollicott. Statistical properties of the Rauzy-Veech-Zorich map. Available at http://homepages.warwick.ac.uk/~masdbl/teichmuller-asip.pdf.

[Pol84] M. Pollicott. A complex Ruelle-Perron-Frobenius theorem and two counterexamples. *Ergodic Theory Dynam. Systems*, 4(1):135–146, 1984.

[PP90] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque*, (187-188):268, 1990.

[Sil89] J.H. Silverman. Integral points on curves and surfaces. In *Number theory (Ulm, 1987)*, volume 1380 of *Lecture Notes in Math.*, pages 202–241. Springer, New York, 1989.

[Sil95] J.H. Silverman. Counting integer and rational points on varieties. *Astérisque*, (228):4, 223–236, 1995. Columbia University Number Theory Seminar (New York, 1992).

[SM57] H. Schwartz and H. T. Muhly. On a class of cubic Diophantine equations. *J. London Math. Soc.*, 32:379–382, 1957.

[Sul79] D. Sullivan. The density at infinity of a discrete group of hyperbolic motions. *Inst. Hautes Études Sci. Publ. Math.*, (50):171–202, 1979.

[Sul84] D. Sullivan. Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups. *Acta Math.*, 153(3-4):259–277, 1984.

[Wie50] H. Wielandt. Unzerlegbare, nicht negative Matrizen. *Math. Z.*, 52:642–648, 1950.

[Wir74] E. Wirsing. On the theorem of Gauss-Kusmin-Lévy and a Frobenius-type theorem for function spaces. *Acta Arith.*, 24:507–528, 1973/74. Collection of articles dedicated to Carl Ludwig Siegel on the occasion of his seventy-fifth birthday, V.

[Zag82] D. Zagier. On the number of Markoff numbers below a given bound. *Math. Comp.*, 39(160):709–723, 1982.

[Zor06] A. Zorich. Flat surfaces. In *Frontiers in number theory, physics, and geometry. I*, pages 437–583. Springer, Berlin, 2006.