Periodic Coefficients and Random Fibonacci Sequences

by

Karyn Anne McLellan

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
August 2012

DALHOUSIE UNIVERSITY

DEPARTMENT OF MATHEMATICS AND STATISTICS

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "Periodic Coefficients and Random Fibonacci Sequences" by Karyn Anne McLellan in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Dated: August 20, 2012

External Examiner: ____________________

Research Supervisor: ____________________

Examining Committee: ____________________

____

Departmental Representative: ____________________

DALHOUSIE UNIVERSITY

DATE: August 20, 2012

AUTHOR: Karyn Anne McLellan

TITLE: Periodic Coefficients and Random Fibonacci Sequences

DEPARTMENT OR SCHOOL: Department of Mathematics and Statistics

DEGREE: PhD CONVOCATION: October YEAR: 2012

Signature of Author

Table of Contents

List of Tables

List of Figures

Abstract

The random Fibonacci sequence is defined by $t_1 = t_2 = 1$ and

$$t_n = \pm t_{n-1} + t_{n-2},$$

for $n \geq 3$, where each $\pm$ sign is chosen at random with $P(+) = P(-) = \frac{1}{2}$. We can think of all possible such sequences as forming a binary tree T. Viswanath has shown that almost all random Fibonacci sequences grow exponentially at the rate $1.13198824\ldots$. He was only able to find 8 decimal places of this constant through the use of random matrix theory and a fractal measure, although Bai has extended the constant by 5 decimal places. Numerical experimentation is inefficient because the convergence is so slow. We will discuss a new computation of Viswanath's constant which is based on a formula due to Kalmár-Nagy, and uses an interesting reduction R of the tree T developed by Rittaud.

Also, we will focus on the growth rate of the expected value of a random Fibonacci sequence, which was studied by Rittaud. This differs from the almost sure growth rate in that we first find an expression for the average of the n^{th} term in a sequence, and then calculate its growth. We will derive this growth rate using a slightly different and more simplified method than Rittaud, using the tree R and a Pascal-like array of numbers, for which we can further give an explicit formula.

We will also consider what happens to random Fibonacci sequences when we remove the randomness. Specifically, we will choose coefficients which belong to the set $\{1, -1\}$ and form periodic cycles. By rewriting our recurrences using matrix products, we will analyze sequence growth and develop criteria based on eigenvalue, trace and order for determining whether a given sequence is bounded, grows linearly or grows exponentially. Further, we will introduce an equivalence relation on the coefficient cycles such that each equivalence class has a common growth rate, and consider the number of such classes for a given cycle length. Lastly we will look at two ways to completely characterize the trace, given the coefficient cycle, by breaking the matrix product up into blocks.

List of Abbreviations and Symbols Used

$\{x_n\}$ .... general sequence

$\{F_n\}$ .... Fibonacci sequence

ϕ .... golden ratio

$\{t_n\}$ .... random Fibonacci sequence/periodic coefficient sequence

P .... probability

$\{f_n\}$ .... linear random Fibonacci sequence considered by Janvresse et al.

$\{\widetilde{f}_n\}$ .... non-linear random Fibonacci sequence considered by Janvresse et al.

T_1 .... binary tree associated with $\{t_n\}$

T_2 .... binary tree associated with $\{f_n\}$

$\widetilde{T} = T$ .... binary tree associated with $\{\widetilde{f}_n\}$

$\mathbb{E}$ .... expected value

P_n .... product matrix

M_j .... terms in the product matrix, equal to A or B

μ_f .... distribution of matrices in P_n (chosen randomly)

γ_f .... upper Lyapunov exponent

e^{γ_f}, τ .... Viswanath's constant

R .... Rittaud's reduced tree

β^* .... Embree-Trefethen constant

$\sigma(\beta)$ .... growth rate of $\{f_n\}$ with coefficient β

$\gamma_f(p)$....... upper Lyapunov exponent as a function of the probability

σ_n .... coefficient cycle

G .... group generated by matrices A and B

$\{a_n\}$ .... general second order linear recurrence sequence

λ_i, λ_1 .... eigenvalues, dominant eigenvalue where it exists

u .... trace of P_n

v .... negative determinant of P_n

$\mathrm{PSL}(2, F)$ .... projective special linear group over a field F

$\mathrm{PS^*L}(2, \mathbb{Z})$ .... projective special linear group, with ± 1 determinants, over $\mathbb{Z}$

PG .... projective analog of group G, where $PG \leq \mathrm{PS^*L}(2, \mathbb{Z})$

$Z(G)$ .... center of group G

$[M]$ .... equivalence class $\{\pm M\}$

ord(M) .... order of matrix M

$K_n(x_1, \ldots, x_n)$ .... continuant polynomial

σ_p .... primitive cycle

$N(n, q)$ .... number of necklaces with n beads and q colors

$N_s(n, q)$ .... number of necklaces with n beads and q colors, with color swapping

$I(t)$ .... number of objects left invariant under transformation t

$\phi(d)$ .... Euler's totient function

C_n .... cyclic group of order n

R^k .... rotation in the cyclic group

S_n .... symmetric group of order n

$N(c_1, \ldots, c_q)$ .... number of permutations with c_j cycles of length j

$G \times H$ .... direct product of groups

D_{2n} .... dihedral group of order $2n$

$F^t R^k$ .... element of D_{2n}

$F^t R^k \pi$ .... element of $D_{2n} \times S_q$

$B_s(n, q)$ .... number of bracelets with n beads and q colors, with color swapping

$B_{sp}(2, q)$ .... number of equivalence classes of coefficient cycles

$\mu(d)$ .... Möbius function

$T^{(a,b)}$ .... tree T with initial nodes (a, b)

$T(p, \alpha)$ .... tree T with $P(+) = p$ and weight α

τ_n .... n^{th} row of entries of T

ρ_n .... n^{th} row of entries of R

$\tau_n^{(a,b)}$ .... n^{th} row of entries of $T^{(a,b)}$

$c_R(\rho_n)$ .... children of ρ_n in the tree R

$c_T(\rho_n)$ .... children of ρ_n in the tree T

$\mathcal{M}$ .... multiset

$\uplus$ .... multiset sum

ρ_n^-, ρ_n^+ .... left and right elements of ρ_n respectively

$c_R^-(\rho_n), c_R^+(\rho_n)$ .... left and right children, respectively, of ρ_n

$m^i(\rho_n)$ .... elements of τ_n which are descendants of ρ_{n-i}

$t(n,k)$ .... number of copies of ρ_{n-3k} in τ_n

$t(k)$ .... corner numbers in Table 5.2

$S(\tau_n), S(\rho_n)$ .... sums of nodes τ_n and ρ_n respectively

$P(\tau_n), P(\rho_n)$ .... products of (non-zero) nodes τ_n and ρ_n respectively

α .... growth rate of $S(\rho_n)$

$A(m)$ .... coefficients in recurrence for $t(k)$

$A(i,m)$ .... coefficients in recurrence for $t(n,k)$

$s(i,m)$ .... number of dots in the m^{th} shape of the sequence with i dots in the 1-d column

$B_{k,n}$ .... exponential partial Bell polynomial

ρ .... a.s. growth rate of a random Fibonacci sequence in R

B_q .... product matrix with q type I Fibonacci blocks

$\hat{B}_q$ .... product matrix with q type II Fibonacci blocks

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. Karl Dilcher for his exceptional guidance, encouragement and advice over the years. (And for never letting me forget an umlaut!) His enthusiasm and care have been truly inspiring, and have shaped not only my thesis, but my experience as a graduate student. I would also like to thank my readers Dr. Jason Brown and Dr. Keith Johnson, as well as my external examiner Dr. Jeffrey Shallit, for taking the time to read my thesis and for providing excellent feedback and insight. Thanks also to Dr. Shallit for traveling from Waterloo for this task. Funding from both NSERC and the Dalhousie Department of Mathematics and Statistics has been gratefully acknowledged. Finally, to Jason, my family and my friends — thank you for your constant support, and belief that I would someday finish, as much as I may have disagreed at the time! It's been a real journey!

Chapter 1

Introduction

The starting point for this thesis is a 1999 paper by Divakar Viswanath entitled *Random Fibonacci Sequences and the Number 1.13198824...* [72], which was in fact a chapter of his Ph.D. thesis [71].

1.1 The Fibonacci Sequence

Before defining a random Fibonacci sequence, we will start by reviewing the well-known regular Fibonacci sequence.

Definition 1.1. The Fibonacci sequence $\{F_n\}$ has initial terms $F_1 = F_2 = 1$ and is given by the recurrence

$$F_n = F_{n-1} + F_{n-2},$$

for $n \geq 3$. It is conventional to define $F_0 = 0$.

We can also represent this recurrence using matrices as follows:

$$\begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_{n-2} \\ F_{n-1} \end{pmatrix}, \tag{1.1}$$

so that multiplying gives us the equations $F_{n-1} = F_{n-1}$ and $F_n = F_{n-1} + F_{n-2}$. Iterating Equation (1.1) for $n \geq 3$ gives

$$\begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{n-2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{1.2}$$

Using elementary linear algebra or the theory of linear recurrences (Vajda [69, p. 18]), it is easy to find the general solution to the Fibonacci recurrence, which is given by Binet's formula,

$$F_n = \frac{1}{\sqrt{5}} \left(\phi^n - (\phi')^n\right), \tag{1.3}$$

where $\phi = \frac{1+\sqrt{5}}{2} = 1.618033989\ldots$ and is called the *golden ratio*, and $\phi' = \frac{1-\sqrt{5}}{2}$ is the conjugate of ϕ. Using this exact form for the Fibonacci numbers, it is straightforward to prove that

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \lim_{n\to\infty} F_n{}^{\frac{1}{n}} = \phi,$$

i.e., the Fibonacci numbers grow exponentially at the rate ϕ. (We will consider this definition of growth rate in Chapter 2.)

Furthermore, we can generalize the Fibonacci numbers so that the initial values are $G_1 = a$, $G_2 = b$, and for $n \geq 3$ we have

$$G_n = G_{n-1} + G_{n-2}.$$

It can be shown that

$$G_n = bF_{n-1} + aF_{n-2}, \tag{1.4}$$

and that for a and b not both zero, the growth rate of this generalized sequence G_n is also ϕ. For a wonderful collection of information on everything to do with the Fibonacci numbers and the golden ratio, visit Ron Knott's award winning website [45].

1.2 The Random Fibonacci Sequence

Viswanath proceeds to define the following.

Definition 1.2. The random Fibonacci sequence $\{t_n\}$ has initial terms $t_1 = t_2 = 1$ and is given by the recurrence

$$t_n = \pm t_{n-1} \pm t_{n-2}, \tag{1.5}$$

for $n \geq 3$, where each $\pm$ sign is chosen independently with probabilities $P(+) = P(-) = \frac{1}{2}$.

This type of discrete distribution is called the *signed Bernoulli distribution*. A random variable has signed Bernoulli distribution if it takes values in the set $\{+1, -1\}$ with equal probability, as is the case with our coefficients. Similarly an *unsigned Bernoulli distribution* uses values in the same set with probabilities which are not necessarily equal, i.e., $P(+) = p$ and $P(-) = 1 - p$ (Tao [68]). We will consider this type of distribution later. For simplicity we choose to let our random variable be the sign $\pm$, rather than the coefficient ± 1.

The terms of the random Fibonacci sequence seem to bounce around between positive and negative values, but as Hayes [37] nicely states, "the steady growth of the Fibonacci numbers is replaced by fluctuations of increasing amplitude." In fact, Hayes suggested coining the term "Vibonacci numbers" because of the way the sign vibrates back and forth. (Also fitting for their creator!) Viswanath was interested in the growth of the sequence $\{|t_n|\}$ of positive values, and so he redefined the recurrence as

$$t_n = \pm t_{n-1} + t_{n-2}, \tag{1.6}$$

where we only have one choice of sign to make. We can do this because when considering $\{|t_n|\}$, Equation (1.5) gives two instances of $|t_{n-1}+t_{n-2}|$ and two of $|t_{n-1}-t_{n-2}|$, where Equation (1.6) gives one instance of each. Since in both cases the probability of each outcome occurring is $\frac{1}{2}$, the definitions (in absolute value) are equivalent. Alternatively, Viswanath could have defined

$$t_n = t_{n-1} \pm t_{n-2}.$$

Note that when generating the sequence $\{|t_n|\}$, we first generate $\{t_n\}$, then we take the absolute value of all terms.

We can think of the random Fibonacci sequence as being generated by flipping a coin; if it comes up heads, we add the previous two terms to get the next one, and if it comes up tails, we subtract them. Devlin [23] makes an interesting analogy to the weather – today's weather is dependent on the weather of the previous two days, although there is a still a large element of chance. Tables 1.1 and 1.2 give two example random Fibonacci sequences, calculated to one hundred terms using Equation (1.6). Notice that the first sequence reaches much higher values.

The main result of Viswanath's paper deals with the growth rate of sequences $\{|t_n|\}$. For a nice introduction to the problem and its history, see the articles by Devlin [23], Peterson [61] and Hayes [37].

1	−8	19	−131	112	−407	−519	−2860	12071	−19165
2	3	−11	−30	407	−295	−926	3603	19165	−7094
3	−11	30	−101	519	−112	407	743	−7094	−26259
5	−8	−41	71	−112	−407	−519	2860	26259	−33353
−2	−19	71	−30	407	−519	−112	−2117	19165	7094
3	11	30	41	−519	−926	−631	4977	45424	−40447
−5	−8	101	−71	−112	−1445	−743	2860	−26259	−33353
8	19	131	112	−407	−2371	−1374	2117	19165	−73800
−13	11	−30	−183	−519	926	−2117	4977	−45424	40447
−5	8	101	295	112	−1445	743	7094	−26259	−33353

Table 1.1: Sample random Fibonacci sequence 1: Values of $t(n)$ for $2 \leq n \leq 101$.

1	−2	−5	245	226	−19	31	112	−343	555
0	−1	7	402	−157	−69	−50	−81	−555	212
1	−3	−12	−157	69	50	81	31	−898	767
−1	−4	19	245	−88	−19	31	−112	343	−555
0	1	−31	88	−19	31	112	−81	−555	212
−1	−3	50	157	−107	−50	−81	−31	898	−343
−1	−2	19	245	−126	−19	31	−50	343	−131
0	−5	69	−88	19	−31	−112	−81	555	−212
−1	−7	88	157	−107	−50	143	−131	898	−343
1	2	157	69	−88	−81	31	−212	−343	131

Table 1.2: Sample random Fibonacci sequence 2: Values of $t(n)$ for $2 \leq n \leq 101$.

Theorem 1.1. *For almost all random Fibonacci sequences,*

$$\lim_{n\to\infty} |t_n|^{\frac{1}{n}} = 1.13198824\ldots.$$

This tells us that with probability one (or we may say *almost surely* or simply *a.s.*), any Fibonacci sequence chosen at random grows exponentially and does so with a fixed growth rate, namely $1.13198824\ldots$. Viswanath's result is counterintuitive because we might expect that with an equal number of additions and subtractions (on average), the terms would eventually balance out to zero. Or, it could be possible that they jump around so chaotically that no limit in the growth rate is reached.

This mysterious number is known as "Viswanath's constant" ([67, A078416]), a term coined by Embree and Trefethen [24]. Viswanath was only able to calculate eight decimal places of it through extensive computation of upper and lower bounds. In 2007, Bai [2] extended the constant by five decimal places to $1.1319882487943\ldots$. There is no known closed form or analytical expression for Viswanath's constant and nothing else is known about its nature, although it is reasonable to conjecture that it is irrational and also transcendental.

It should be noted, though, that there are plenty of Fibonacci-type sequences with ± 1 coefficients that do not grow at the rate $1.13198824\ldots$. For example, if all $+$ signs are chosen, we end up with the regular Fibonacci sequence, which has growth rate $1.618033989\ldots$. It's also possible to construct sequences that do not grow at all. If we choose signs according to the pattern $++-++-\cdots$, which repeats with period three, the resulting sequence is

$$1, 1, 2, 3, -1, 2, 1, 1, 2, 3, -1, 2, \ldots.$$

This sequence begins to repeat after six terms because we land back at our initial values after a multiple of the cycle defining the pattern. We will consider this idea in more detail in Chapter 2. The important thing is that these exceptional sequences have zero probability of occurring at random, or in other words, they belong to a set of measure zero. We can refer to them as random Fibonacci sequences, but this requires us to specify the almost sure condition when talking about the growth of random Fibonacci sequences in general.

The starting point of this thesis was to study the behaviour of these "non-random" random Fibonacci sequences in order to potentially shed some light on Viswanath's elusive constant. We will remove the aspect of randomness by selecting sequences according to patterns of $\pm$ signs, rather than choosing these signs at random. In essence, we will look at sequences which have probability zero of occurring randomly. These sequences, however, can be used to approximate the set of all possible random Fibonacci sequences.

Numerical experiments can be done to enumerate all possible length-n random Fibonacci sequences, and take the arithmetic average of their growth rates. For $n = 20$ there are over half a million branches and the average growth rate calculated is 1.18 (Hayes [37]). Later we will do similar calculations with geometric means over all possible sequences and obtain a similarly poor approximation of 1.12 for $n = 20$

and 1.19 for $n = 21$ (see column 1 of Table 6.3). This method is very impractical when n gets large. Random sampling proves to be more efficient. One example of a random Fibonacci sequence of length one million calculated by Viswanath was shown to have growth rate 1.132 and reached values of over $10^{50,000}$. This still only gives two decimal places of accuracy! Table 1.3 gives the growth rates of five random Fibonacci sequences up to length one million, and approximately the same accuracy as Viswanath's computation is reached. Further, we can increase the accuracy in our

n	Sequence 1	Sequence 2	Sequence 3	Sequence 4	Sequence 5
10^1	1.245730940	1.174618943	1.174618943	1.116123174	1.116123174
10^2	1.162488171	1.140380133	1.099325710	1.165414115	1.093747433
10^3	1.155237866	1.112871284	1.126232310	1.139393990	1.127914018
10^4	1.131234546	1.131036327	1.125746602	1.132062832	1.139082373
10^5	1.132959794	1.131119861	1.131994075	1.132940636	1.134283491
10^6	1.131816247	1.132325605	1.132482700	1.132311681	1.131873039

Table 1.3: The growth rate $|t(n)|^{1/n}$ for five random Fibonacci sequences, $n \leq 10^6$.

approximations by taking averages of growth rates. The Maple program in Figure A.3 calculates the growth rates of j random Fibonacci sequences of length $n + 2$, then takes the average. For $j = 20$ and $n = 40000$, we obtain the values given in Table 1.4. These actually appear to be slightly closer to Viswanath's constant than those in Table 1.3, despite the smaller sequence length. It is thus quite remarkable that Viswanath and Bai were able to find 8 and 13 correct decimals respectively. The random Fibonacci sequence is an interesting example of how a random process can lead to something deterministic over time.

1.131885031	1.131973850	1.131740524
1.132081608	1.132389153	1.1323296404

Table 1.4: Averages of 20 growth rates of random Fibonacci sequences with length 40002.

In [40], Janvresse et al. expanded the definition of the random Fibonacci sequence by introducing linear and non-linear cases. In the *linear* case, the sequence is defined by initial values $f_1 = f_2 = 1$ and the recurrence

$$f_n = f_{n-1} \pm f_{n-2}. \tag{1.7}$$

This is exactly the alternate way we could have defined $\{t_n\}$. The *non-linear* case on the other hand is defined by initial values $\widetilde{f}_1 = \widetilde{f}_2 = 1$ and the recurrence

$$\widetilde{f}_n = |\widetilde{f}_{n-1} \pm \widetilde{f}_{n-2}|. \tag{1.8}$$

The distinguishing feature of the latter definition is that we take the absolute value at every stage when generating the sequence. This is in contrast to $\{|f_n|\}$, for example, where we take absolute values only after the sequence has been generated. Note that Equation 1.8 does not define a linear recurrence. Janvresse et al. were interested in the growth of the positive sequences $\{\widetilde{f}_n\}$ and $\{|f_n|\}$.

1.3 Distributions and Binary Trees

A note of caution – the recurrences in (1.6), (1.7) and (1.8) do not necessarily generate identical sequences $\{|t_n|\}$, $\{|f_n|\}$ and $\{\widetilde{f}_n\}$. Consider the following example.

Example 1.1. Suppose we choose alternating signs $+ - + - \cdots$ We obtain the following sequences:

$$\begin{aligned}
\{t_n\} &= 1, 1, 2, -1, 1, -2, -1, -1, -2, 1, -1, 2, 1, 1, \ldots, \\
\{|t_n|\} &= 1, 1, 2, 1, 1, 2, \ldots, \\
\{f_n\} &= 1, 1, 2, 1, 3, 2, 5, 3, 8, 5, 13, \ldots = |f_n|, \\
\{\widetilde{f}_n\} &= 1, 1, 2, 1, 3, 2, 5, 3, 8, 5, 13, \ldots.
\end{aligned}$$

Note that it is not necessarily the case that the sequence $\{f_n\}$ remains positive. Not only are two of the sequences distinct in absolute value, but one remains bounded while the other appears to grow exponentially. We can explain this discrepancy by showing that the *distribution* of sequences generated under these three methods is the same.

An excellent tool for visualizing these distributions is the binary tree. We give the root a label of 1 and call this row 1. The root has an only child, also labeled

1, occurring in row 2. Every vertex except the root has two children, the values of which depend on the parent and grandparent, and correspond to the random Fibonacci sequence definition we are considering. The right child is generated by choosing $+$, i.e., we add the parent and grandparent; the left child by choosing $-$, i.e., we subtract the parent and grandparent, taking absolute values if required. Each row n for $n \geq 2$ has 2^{n-2} nodes. Each branch of the tree represents a possible random Fibonacci sequence, where the pattern of lefts/rights corresponds to the pattern of $\pm$ signs used in generating the sequence.

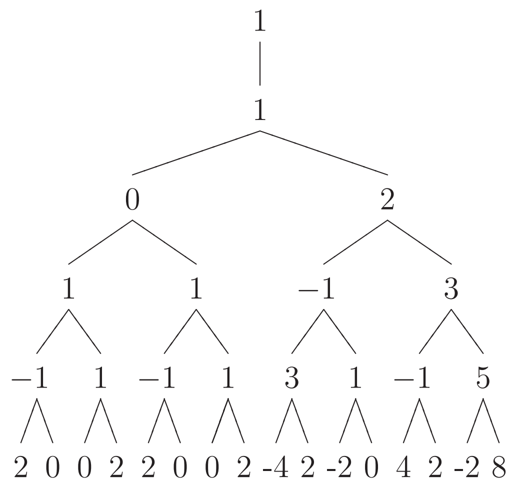
The trees corresponding to the definitions in (1.6), (1.7) and (1.8), which we will denote by T_1, T_2 and $\widetilde{T}$ respectively, are given in Figure 1.1. The first two trees correspond to the linear cases, whereas the third tree corresponds to the non-linear case. If we compare the trees $|T_1|$, $|T_2|$ and $\widetilde{T}$ (where the absolute value of a tree means we take the absolute value of all terms in the tree), the same set of positive sequences occurs in each, although the sequences may be permuted. This remains true regardless of the number of rows we calculate, as is shown in the following theorem.

Theorem 1.2. *The random sequences $\{|t_n|\}$, $\{|f_n|\}$ and $\{\widetilde{f_n}\}$ all have the same distribution, i.e., the probability of a particular n-termed sequence occurring is the same in all three cases.*

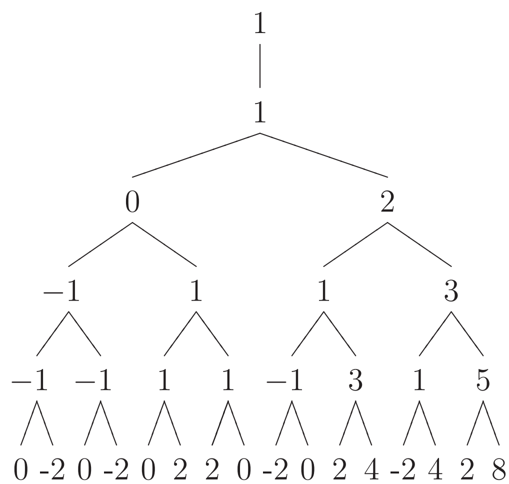
Proof: We can prove this result using induction. By observing all three trees we see that for $n = 3$ we have the same set of positive sequences, $\{(1, 1, 0), (1, 1, 2)\}$, in each case. Notice that for trees T_1 and T_2 nodes can be positive or negative, while in $\widetilde{T}$ nodes are in absolute value. Now suppose that we have the same set of n-termed sequences for all three sequence definitions: $\{|t_n|\}$, $\{|f_n|\}$ and $\{\widetilde{f_n}\}$. Consider any node $|b|$, in row n of each tree $|T_1|$, $|T_2|$ and $\widetilde{T}$, with parent $|a|$ occurring in row $n-1$. Left and right children of b ($|b|$ in the case of $\widetilde{T}$) in row $n+1$ of the trees T_1, T_2 and $\widetilde{T}$ are seen in the tree segments in Figure 1.2.

Now consider b's children in absolute value, i.e., b's children in trees $|T_1|$, $|T_2|$ and $\widetilde{T}$. If a and b have the same sign,

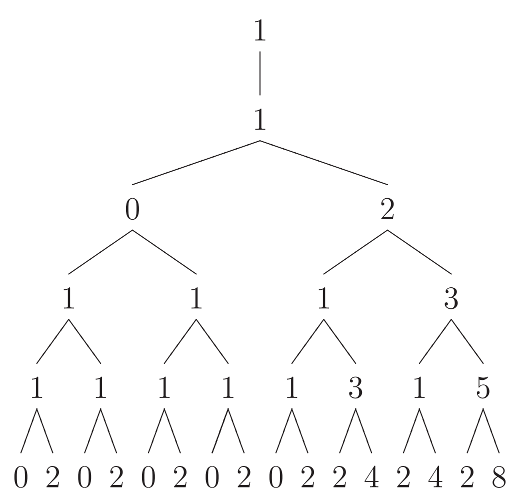
$$|a + b| = |a| + |b|, \quad |a - b| = ||a| - |b||.$$



(a) The tree T_1.

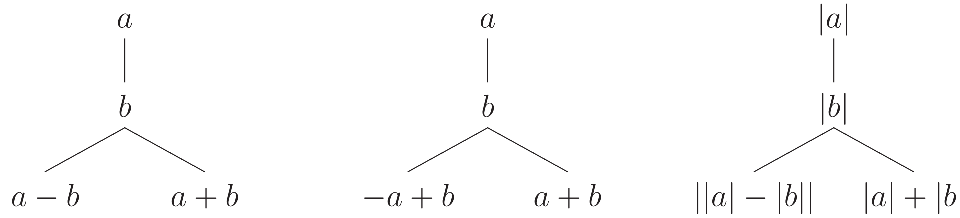


(b) The tree T_2.



(c) The tree $\widetilde{T}$.

Figure 1.1: Trees for different cases of random Fibonacci sequences.



(a) Segment of T_1 (b) Segment of T_2 (c) Segment of $\widetilde{T}$

Figure 1.2: Left and right children in each tree type.

If a and b differ in sign,

$$|a+b| = ||a|-|b||, \quad |a-b| = |a|+|b|.$$

Therefore, the pairs of children are all the same in absolute value, although possibly in different orders. Since a node b with parent a ($|b|, |a|$ in the case of $\widetilde{T}$) in any of the trees T_1, T_2 or $\widetilde{T}$ generates the same pair of children in absolute value, the trees $|T_1|$, $|T_2|$ and $\widetilde{T}$ contain the same set of sequences of a given length n. $\diamond$

Corollary 1.1. *The set of elements in a given row is the same for each of the trees $|T_1|$, $|T_2|$ and $\widetilde{T}$, i.e., the set of n^{th} terms is the same for the sequences $\{|t_n|\}$, $\{|f_n|\}$ and $\{\widetilde{f}_n\}$.*

Proof: The proof is immediate from Theorem 1.2 because each tree $|T_1|$, $|T_2|$ and $\widetilde{T}$ contains the same set of sequences. $\diamond$

Aside from the almost sure growth rate of a random Fibonacci sequence, we will also be interested in the growth rate of the expected value of terms in a given row. Here we simply take the arithmetic mean value of the n^{th} terms in each of the 2^{n-2} possible random Fibonacci sequences, because each sequence occurs with equal probability, and then calculate the growth rate of this mean sequence $\{\mathbb{E}(|t_n|)\}$.

Theorem 1.3. *The sequences $\{|t_n|\}$, $\{|f_n|\}$ and $\{\widetilde{f}_n\}$ have the same almost sure growth rate and the same growth rate of the expected value of terms.*

Proof: This follows directly from the fact that these sequences have the same distribution. In particular, by Corollary 1.1 we know that the sets of n^{th} terms of the sequences $\{|t_n|\}$, $\{|f_n|\}$ and $\{\widetilde{f_n}\}$ are all the same. Therefore, the expected value of the n^{th} term is the same for each sequence and consequently so is its growth rate. $\diamond$

The difference between the almost sure growth rate of a sequence and the growth rate of the expected value of the terms will be discussed later in further detail.

1.4 Determining Viswanath's Constant

We will now take a closer look at the method Viswanath used to calculate his constant. As with the regular Fibonacci sequence, we can represent a random Fibonacci sequence using matrices. We can rewrite the recurrence given in (1.6) as

$$\begin{pmatrix} t_{n-1} \\ t_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} \begin{pmatrix} t_{n-2} \\ t_{n-1} \end{pmatrix}, \tag{1.9}$$

where the choice of $+$ or $-$ sign is represented by the matrices

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix},$$

respectively, and each is chosen with equal probability. We will denote the i^{th} matrix chosen by M_i, the product of n such matrices by P_n and the distribution by μ_f. Iterating Equation (1.9) and including the initial values we can then write

$$\begin{pmatrix} t_{n-1} \\ t_n \end{pmatrix} = M_{n-2}M_{n-3}\cdots M_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = P_{n-2} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

for $n \geq 3$. We will consider the details of these matrix products more closely in Chapter 2. Note that the sequence obtained from Equation (1.7) (linear) can similarly be modeled by a product of two independently and identically distributed (i.i.d.) random matrices although the sequence obtained from Equation (1.8) (non-linear) cannot. It can be modeled by a product of three random matrices; however, their distribution is not i.i.d.

Definition 1.3. There are several possible interpretations of a random matrix:
a) A *random matrix* is a matrix whose entries are *random variables* chosen from some

set with a given distribution.

b) It is also possible to have a matrix whose entries are random variables chosen from possibly distinct sets with distinct distributions. This is used to study properties of distinct distributions relative to one another.

c) Alternately, a *random matrix* is a random variable that takes the form of a matrix. In this case, we are choosing a matrix at random from a given set of matrices according to some distribution.

It is possible to obtain Definition 1.3 c) from Definition 1.3 a) if one simply counts all possible matrices obtained from choosing different matrix entries. In our case, since only one matrix entry is chosen randomly (the bottom right entry is ± 1), both Definitions 1.3 a) and 1.3 c) suit our random matrices equally well, although it is more convenient to think in terms of the latter definition. Further, we will refer to the products P_n as "products of random matrices" rather than "random matrix products" to avoid ambiguity.

The theorem of Furstenberg and Kesten [29] (also proven in Bougerol and Lacroix [11, p. 11]) tells us that for our sequence of i.i.d. random matrices $\{M_n\}$ with distribution μ_f, the *upper Lyapunov exponent* can be defined as

$$\gamma_f = \lim_{n\to\infty} \frac{1}{n} \log \|M_n \cdots M_1\| \tag{1.10}$$

almost surely, where we are using 2-norms for vectors and matrices. Furthermore, according to Bougerol [11], the Lyapunov exponent (which we may shorten simply to Lyapunov exponent) can be written as

$$\lim_{n\to\infty} |t_n|^{\frac{1}{n}} = e^{\gamma_f} \tag{1.11}$$

almost surely, which is the growth rate of the sequence $\{|t_n|\}$. This is equivalent to saying

$$\lim_{n\to\infty} \frac{1}{n} \log |t_n| = \gamma_f.$$

(We will use log to denote the natural logarithm throughout this thesis.) We may also define the exponential growth rate of $\{|t_n|\}$ as the ratio

$$\lim_{n\to\infty} \frac{|t_n|}{|t_{n-1}|} = e^{\gamma_f}.$$

We will further discuss exponential growth, and these equivalent definitions of it, in Chapter 2. We can generalize Theorem 1.1 to the case of $t_1, t_2 \neq 1$.

Proposition 1.1. *For almost all random Fibonacci sequences,*

$$\lim_{n\to\infty} |t_n|^{\frac{1}{n}} = 1.13198824\ldots,$$

where t_1*,* $t_2 \in \mathbb{R}$ *and are not both zero.*

Proof: Equation (1.10) for the upper Lyapunov constant depends only on the product of matrices and not on the initial values of the sequence. Therefore Viswanath's constant is independent of the initial values, as was the case for the regular Fibonacci sequence, with the exception of $t_1 = t_2 = 0$, which produces an all-zero sequence. $\diamond$

Viswanath wanted to find an exact value for γ_f, although very little was known about such a constant. In 1960, Furstenberg and Kesten [29] showed that the upper Lyapunov exponent γ exists under general conditions and in 1963 Furstenberg [30] showed that provided $|\det(M_i) = 1|$, we usually have $\gamma > 0$. Since this determinant condition is true for μ_f, Viswanath concluded that $\gamma_f > 0$ and so almost surely a random Fibonacci sequence grows exponentially. Furthermore, we know that if we choose only plus signs when generating our random Fibonacci sequence, we end up with the regular Fibonacci sequence, which has growth rate $\phi = \frac{1+\sqrt{5}}{2} = 1.618033989\ldots$. This tells us that $1 < e^{\gamma_f} < \phi$ almost surely.

Multiplying P_n by an initial value vector gives us a vector x (containing sequence terms) in $\mathbb{R}^2$. Viswanath parameterized these vectors using slopes m where $m \in [-\infty, \infty)$ to give $\overline{x} = (1, m)^T$ We can now think of our random Fibonacci sequence as a random walk in $\mathbb{R}^2$, where each new matrix added to the product corresponds to a change in direction. In this new setting, we can use Furstenberg's formula [30] (the proof can also be found in Bougerol and Lacroix [11, p. 27–29]), to obtain the following expression for γ_f:

$$\gamma_f = \int \operatorname{amp}(\overline{x})\, d\nu_f(\overline{x}),$$

where $\operatorname{amp}(\overline{x})$ is a smooth function of $\overline{x}$ which gives the average amplification of x in the direction $\overline{x}$ when it is multiplied by A or B and $\nu_f(\overline{x})$ is the unique invariant probability measure over the directions $\overline{x}$ for the random walk. The measure of an interval $[a,b]$ on the real line with $-1 \notin (a,b)$ can be given as

$$\nu_f([a,b]) = \frac{1}{2}\nu_f\left(\left[\frac{1}{-1+b}, \frac{1}{-1+a}\right]\right) + \frac{1}{2}\nu_f\left(\left[\frac{1}{1+b}, \frac{1}{1+a}\right]\right), \tag{1.12}$$

and the amp function, in terms of slope m, is

$$\operatorname{amp}(m) = \frac{1}{4}\log\left(\frac{1+4m^2}{(1+m^2)^2}\right).$$

Next, Viswanath makes clever use of the Stern-Brocot tree (which he discovered independently) to find the invariant measure ν_f. For an excellent introduction to the Stern-Brocot tree, see Graham et al. [33, p. 116]. Viswanath used a variation of the tree, where the nodes are intervals that partition the real line, which represents the set of all slopes m. The root at depth $d = 1$ is given by the interval $[-\infty, \infty] = \left[\frac{-1}{0}, \frac{1}{0}\right]$, with left and right children $\left[\frac{-1}{0}, \frac{0}{1}\right]$ and $\left[\frac{0}{1}, \frac{1}{0}\right]$ respectively. Further, given any node $\left[\frac{a}{b}, \frac{c}{d}\right]$ at depth $d \geq 3$, its left and right children are $\left[\frac{a}{b}, \frac{a+c}{b+d}\right]$ and $\left[\frac{a+c}{b+d}, \frac{c}{d}\right]$ respectively. Any interval in the Stern-Brocot tree can be represented by the sequence of L's and R's (lefts and rights in the tree) required to reach it.

Viswanath found simple rules in terms of the Stern-Brocot intervals (sets of directions) for mapping a direction m to $\frac{1}{m}$, $(1+m)$ or $(-1+m)$ (these mappings can be thought of as Möbius transformations) and used these to write the invariance condition given in Equation (1.12) as an infinite system of linear equations for $\nu_f(I)$, where I is a Stern-Brocot interval written in terms of L's and R's. He was able to guess the solution to this system, which gives recursive relations for $\nu_f(I)$ (for different types of I); this can then be explicitly solved, to give functions of ϕ. When the invariant measure is graphed for increasingly small subintervals of $\mathbb{R}$, a self-similar pattern, repeated at multiple scales, emerges. This is good evidence that the measure ν_f is a fractal, and explains the difficulty in computing it.

Viswanath then used Furstenberg's formula to give upper and lower bounds for γ_f as follows:

$$2\sum_{j=1}^{2^d} \min_{m \in I_j^d} \operatorname{amp}(m)\nu_f(I_j^d) < \gamma_f < 2\sum_{j=1}^{2^d} \max_{m \in I_j^d} \operatorname{amp}(m)\nu_f(I_j^d), \tag{1.13}$$

where I_j^d is the j^{th} Stern-Brocot interval at depth $d+1$ of the tree and $1 \leq j \leq 2^d$. The upper Lyapunov exponent γ_f can be computed to any desired accuracy by finding the upper and lower bounds for large enough d. The problem here lies in computing capability. Viswanath computed his bounds with $d = 28$ using floating point arithmetic and then did a rounding error analysis to show that

$$0.1239755980 < \gamma_f < 0.1239755995.$$

This implies that the growth rate of a random Fibonacci sequence is almost surely

$$e^{\gamma_f} = 1.13198824\ldots$$

by Equation (1.11). Furthermore, it is known that the next digit must be an 8 or a 9. Complete details, as well as Viswanath's program, can be found in [72]. Computer assisted proofs are not uncommon and Viswanath repeated his computation on two completely different systems, assuring its validity.

Work has been done to improve Viswanath's constant as well as the computations required to obtain it. In 2001, Oliveira and De Figueiredo [59] repeated Viswanath's computation using a different and simplified method. They wanted a program that would be accessible to anyone wishing to reproduce Viswanath's result on a desktop machine, as the original computation requires a huge amount of memory. Viswanath used the inequality in (1.13) with careful floating-point calculations to find upper and lower bounds, and then confirmed his result using a rounding-error analysis. Oliveira and De Figueiredo instead used interval arithmetic (also mentioned by Viswanath), which automatically keeps track of rounding errors, to find upper and lower bounds of γ_f as follows:

$$\gamma_f \in 2 \sum_{j=1}^{2^d} \mathrm{AMP}(I_j^d)\, \nu_f(I_j^d),$$

where $\mathrm{AMP}(I_j^d)$ is an interval containing $\mathrm{amp}(I_j^d)$. They obtained the following approximation for e^{γ_f}:

$$1.1319882478 \leq \gamma_f \leq 1.1319882496,$$

which has the same level of accuracy as Viswanath's calculation.

In 2007 Bai [2] extended Viswanath's constant to $1.1319882487943\ldots$, where the next term is 7, 8 or 9. He used a cycle expansion method for the Lyapunov exponent γ_f. In order to use this method he had to map the matrices A and B that Viswanath used to non-negative matrices, and correspondingly change the probability from $\frac{1}{2}$ to $p = \frac{\sqrt{5}-1}{2} = 0.618033989\ldots$. The new Lyapunov exponent γ can be used to calculate the old one from the relation $\gamma_f = \frac{p}{2}\gamma$. Bai's results are numerical. He used the spectrum of an evolution operator which describes the distribution of vector directions under the action of random matrices to compute the convergent components. He then generated an algorithm for cyclic expansion of the spectral determinant that increases efficiency by removing exponentially converging elements. He also noted that Monte Carlo experiments (repeated random sampling, as mentioned earlier) are not very useful, as the convergence is quite slow.

1.5 Lyapunov Exponents and Previous Concepts of Random Sequences

There are very few cases where an exact value for the upper Lyapunov exponent γ can be determined. It is a very difficult problem and there is no general method known for deriving these values.

Viswanath considers the case of the random Fibonacci matrices to be a very natural example where γ can actually be computed. In fact, this is a specific case of the following example given by Furstenberg [30, p. 1]. "Consider the problem of determining the asymptotic behaviour of a random sequence $\{x_n\}$ satisfying

$$x_n = \alpha_n x_{n-1} + \beta_n x_{n-2}, \tag{1.14}$$

where (α_n, β_n) form a sequence of i.i.d. random vectors. In this case we can write

$$\begin{pmatrix} x_{n+1} \\ x_n \end{pmatrix} = M_n M_{n-1} \cdots M_1 \begin{pmatrix} x_1 \\ x_0 \end{pmatrix}, \qquad M_j = \begin{pmatrix} \alpha_{j+1} & \beta_{j+1} \\ 1 & 0 \end{pmatrix},$$

and so the rate of growth of the x is governed by the behaviour of the matrix product $M_n M_{n-1} \cdots M_1$." We will call the general sequence defined in Equation (1.14) a *random Fibonacci-type sequence*. (Note that here the terms in the vectors are switched compared to Equation (1.9), and so the orientation of the matrices is also slightly different.)

Chassaing, Letac and Mora [16] have derived the invariant measure ν_f for several products of 2×2 matrices with positive entries. Viswanath [72] gives an example of such a product of random matrices by choosing $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ with probability p and $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ with probability $1-p$. In these cases the infinite system which defines the invariance requirement is triangular, which means that the value of the measure of an interval depends only on intervals at a lesser depth in the Stern-Brocot tree, unlike Viswanath's case. Some examples of calculable upper Lyapunov exponents are found in Bougerol and Lacroix [11, p. 33] and Chassaing, Letac and Mora [16], although Viswanath's work is much more closely related to the latter. In particular, Bougerol and Lacroix consider the matrix product comprised of matrices $A_n = \left(\begin{smallmatrix} 0 & 1 \\ 1 & \alpha_n \end{smallmatrix}\right)$ where $\alpha_n > 0$ with a specified distribution. This corresponds to the random Fibonacci-type sequence

$$x_n = \alpha_n x_{n-1} + x_{n-2}.$$

Here the invariant measure can be found using Viswanath's techniques when α_n is distributed on the positive integers. Viswanath's random Fibonacci sequence differs because α_n takes on positive and negative values, and it is the only known case with non-positive α_n for which the invariant measure can be found using Stern-Brocot intervals.

We have mentioned that the ideas behind Viswanath's random Fibonacci sequence have roots in the work of Furstenberg [30] and Furstenberg and Kesten [29] on random matrix theory. This field was initiated by Bellman [7] in 1954, who used the example of 2×2 matrices A and B chosen at random, each with probability $\frac{1}{2}$. Bougerol and Lacroix give an excellent and detailed account of random matrix theory in [11]. In general, random recurrences have connections to many different areas of study including dynamical systems, ergodic theory, spectral theory, continued fractions, statistics and physics. Random recurrences can be thought of as a special case of random iterated functions. As well as random recurrences, and their corresponding random sequences, people have also studied random series. Schmuland [65] studied the behaviour of the random harmonic series. He defined

$$X := \sum_{j=1}^{\infty} \frac{\varepsilon_j}{j},$$

where $\varepsilon_j \in \{\pm 1\}$ and is chosen independently and randomly with $P(+1) = P(-1) = \frac{1}{2}$. Here $\mathbb{E}(X) = 0$. Schmuland points out the interesting fact that the binary digits of a number chosen randomly from $[0, 1]$ are equivalent to a sequence of fair coin tosses, and so can be used to model a random sequence $\{\varepsilon_j\}$.

Before Viswanath's work in 1999 there were several different notions of "random Fibonacci sequence" along with other similar ideas about trees and growth. In 1979 Cohn [18] considered a free semigroup whose elements are words comprised of symbols A and B. He generated the group of elements using a binary tree, which he called the "Markoff tree", by concatenation as follows. A pair of words (w_1, w_2) has left child $(w_1, w_1 w_2)$ and right child $(w_2, w_1 w_2)$. This bears resemblance to our binary tree T_1, for example, in that the children of each node are obtained by choosing signs $+$ and $-$, which can also be represented by matrices A and B. Each node can therefore be thought of as a matrix product, or a word, made up of A's and B's. Cohn's tree can be translated to the positive half of the Stern-Brocot tree that was later used by Viswanath by replacing each word w by (a, b), where a and b denote the number of times the symbols A and B occur in w respectively. (Cohn makes no mention of the Stern-Brocot tree.) He points out that the rightmost diagonal of the tree has "Fibonaccian growth". He further provides a characterization of all words in the Markoff tree and also investigates the enumeration of Markoff triples by replacing A and B by specific matrices and considering the traces of the resulting matrix products.

In 1983, Dawson, Gabor, Nowakowski and Wiens [22] defined three different "random Fibonacci-type sequences" (in their words, not according to our definition). First they consider the sequence of positive integers $\{x_n\}$ where the initial p terms are fixed, and subsequent terms are generated by taking the sum of q randomly chosen terms, with replacement, from the list of all existing terms, i.e.,

$$x_{n+1} = \sum_{i=1}^{q} x_{k_i},$$

for $n > p$ where the k_i are chosen at random from $\{1, 2, \ldots, n\}$. Next they consider the case where the k_i are chosen without replacement. Among other things, the expected value of the term x_n is studied. Lastly, they consider the sequence $\{x_n\}$

generated by

$$x_{n+1} = \alpha_n x_n + \beta_{n-1} x_{n-1}, \tag{1.15}$$

where pairs (α_n, β_n) are chosen independently at random and have finite first and second moments. Recall that the k^{th} *moment* of an random variable X is $\mathbb{E}(|X|^k)$ (see Tao [68, p. 15]). The moments of x_n are considered. Note that this is the same general form that Furstenberg defined in Equation (1.14), apart from the conditions on the moments.

Fibonacci famously described his sequence as modeling a rabbit population, and Hayes [37] jokingly describes the random Fibonacci sequence as modeling a population of cannibalistic rabbits. Dawson et al. believe that introducing random variables to the Fibonacci sequence may provide a better model of the growth of certain biological and physical processes.

Interestingly, in 2002, Ben-Naim and Krapivsky [8] defined two new random Fibonacci-type sequences quite similar to that described by Dawson et al. [22], although there was no reference to it. Ben-Naim and Krapivsky's sequences also generate new terms by addition only, in contrast to the random Fibonacci sequence studied by Viswanath. The first sequence has $x_0 = 1$ and

$$x_n = x_{n-1} + x_q,$$

for $n \geq 1$, where q is randomly chosen from the set $\{0, 1, \ldots, n-1\}$. This forces $x_1 = 2$, although the rest of the terms are non-deterministic. In a slight variation, their second sequence has the form

$$x_n = x_p + x_q,$$

where both p and q are chosen randomly from the set. The number of possible sequences increases as $n!$ and $n!^2$ for the two growth models respectively. Also, interestingly, the first model gives monotonically increasing sequences whereas the second increases only in the average. Ben-Naim and Krapivsky use simple recurrences to find the asymptotic growth of $\{\mathbb{E}(x_n)\}$ for their sequences, and note that the growth of the sequence $\{x_n\}$ is slower than that of $\{\mathbb{E}(x_n)\}$. Higher order moments are considered and further, they study the set of all possible sequence values x_n for a given n and find some interesting and complex patterns.

In 2004, Krasikov, Rodgers and Tripp [46] extend the work of Ben-Naim and Krapivsky [8] by considering the random sequence

$$x_n = x_{n-1} + \beta x_q, \tag{1.16}$$

with $\beta > 0$ and q chosen randomly from $\{0, 1, \ldots, n-1\}$ with probability distribution $P_n(q)$. In particular they consider the case for $\beta = 1$, and the case where q has equal probability of taking on any of the values in $\{0, 1, \ldots, n-1\}$, i.e., $P_n(q) = \frac{1}{n}$, where they find an exact solution for $\mathbb{E}(x_n)$ and the divergence of the second moment of x_n, i.e., $\mathbb{E}(x_n^2)$, as functions of n and β. Further Krasikov et al. give conditions on certain sequences for exponential, linear and an intermediate type of growth.

In 1993, Hope [38] used the same random Fibonacci construction as Dawson et al. [22] to generate the sequence with $x_0 = 0$, $x_1 = 1$ and

$$x_n = \alpha_n x_{n-1} + \beta_n x_{n-2},$$

for $n \geq 2$. Here, strictly positive pairs (α_n, β_n) are chosen randomly from a specified probability distribution. Hope showed that subject to some conditions on the pairs (α_n, β_n) for $n \geq 1$, we have

$$\lim_{n\to\infty} \frac{1}{n} \log(x_n) = \Psi$$

almost surely, i.e., the sequence $\{x_n\}$ converges almost surely to a fixed value. (Recall that Furstenberg and Kesten have shown that the upper Lyapunov exponent almost surely exists under general conditions.) Moreover, $\Psi = \mathbb{E}(\log(w))$, where

$$w := \lim_{n\to\infty} w_n = \lim_{n\to\infty} \alpha_1 + \cfrac{\beta_1}{\alpha_2 + \cfrac{\beta_2}{\alpha_3 + \cfrac{\beta_3}{\cfrac{\vdots}{\alpha_{n-1} + \cfrac{\beta_{n-1}}{\alpha_n}}}}}.$$

As an example, Hope assumes that $\beta_n = 1$ for all n and $\alpha_n \in \{1, 2\}$, each chosen with probability $\frac{1}{2}$. The growth constant $\Psi = \mathbb{E}(\log(w))$ is approximated by

$$\frac{1}{2^n} \sum_{\alpha_i \in \{1,2\}} \log([\alpha_1, \alpha_2, \ldots, \alpha_n]) = 0.673,$$

where we have used shorthand notation for the continued fraction and we sum over all possible length n continued fractions with entries 1 or 2, for a given value of n. Therefore, almost all random Fibonacci-type sequences generated in this way have a growth rate of $e^{0.673\ldots} \approx 1.960\ldots$.

Hope furthermore points out the difference between the growth rate of individual sequences $\{x_n\}$ and the growth rate of $\{\mathbb{E}(x_n)\}$. In the latter case we must first find the expected value of the n^{th} term in a random Fibonacci sequence. We can do this by writing

$$\mathbb{E}(x_n) = \mathbb{E}(\alpha_n)\mathbb{E}(x_{n-1}) + \mathbb{E}(\beta_n)\mathbb{E}(x_{n-2}), \tag{1.17}$$

and so

$$\lim_{n\to\infty} \frac{1}{n} \log(\mathbb{E}(x_n)) = \log \Phi,$$

where Φ is the positive root of $X^2 - \mathbb{E}(\alpha_1)X - \mathbb{E}(\beta_1) = 0$, provided $\mathbb{E}(\alpha_1)$ and $\mathbb{E}(\beta_1)$ are finite. The problem with using these ideas on Viswanath's random Fibonacci sequence is that one of Hope's conditions is that $\alpha_n, \beta_n > 0$. In Viswanath's case $\alpha_n, \beta_n \in \{\pm 1\}$. If we calculate $\mathbb{E}(t_n)$ for Viswanath's sequence, we obtain $\mathbb{E}(t_n) = 0$, since $\mathbb{E}(\alpha_n) = \mathbb{E}(\beta_n) = 0$.

Prior to this, Chassaing et al. [16, p. 36] and Bougerol and Lacroix [11, p. 166] have made similar connections between products of random matrices, which may represent random Fibonacci-type sequences, and simple continued fractions. (Note that the continued fractions given by Hope are not simple when $\beta_n \neq 1$ for all n.) Given a real matrix $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $ad - bc \neq 0$, we write its Möbius transformation as

$$M(m) := \frac{am + b}{cm + d},$$

where here, $m \in \mathbb{R} \cup \infty$ is the slope of the vector x mentioned in Viswanath's work. (It is easy to check that $(M_1 M_2)(m) = M_1(M_2(m))$). Now let $\{\alpha_n, \widetilde{\alpha}_n\}$ be a sequence of independent pairs of positive random variables with the same distribution and let

$$M_n := \begin{pmatrix} \alpha_n & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \widetilde{\alpha}_n & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \alpha_n \widetilde{\alpha}_n + 1 & \alpha_n \\ \widetilde{\alpha}_n & 1 \end{pmatrix}.$$

In general it can be shown that the Möbius transformation from a matrix product defines the continued fraction of the random variables:

$$(M_1 M_2 \cdots M_n)(m) = [\alpha_1, \widetilde{\alpha}_1, \ldots, \alpha_n, \widetilde{\alpha}_n, m].$$

Similarly, if we consider the product matrix comprised of the matrices A and B defined by Viswanath (oriented slightly differently), we have

$$(M_1 \cdots M_n)(m) = [\pm 1, \pm 1, \pm 1, \ldots, m], \tag{1.18}$$

as found in Viswanath [72], despite the fact that we are not dealing with positive matrices. So as $n \to \infty$, the distribution of this random continued fraction is in fact the distribution ν_f. Note that for matrices A and B we have

$$M(m) = \frac{1}{m \pm 1},$$

which were the maps defined earlier by Viswanath.

1.6 Generalizations of the Random Fibonacci Sequence

We will now consider some of the many variations and generalizations of the random Fibonacci sequence that have been studied after Viswanath's result.

In 2007, Rittaud [64] finds that the growth rate of the expected value of the non-linear sequence defined by (1.8) is

$$\lim_{n\to\infty} \frac{\mathbb{E}(\widetilde{F}_n)}{\mathbb{E}(\widetilde{F}_{n-1})} = \alpha - 1 = 1.20556943\ldots,$$

where α is the unique real root of $\alpha^3 - 2\alpha^2 - 1 = 0$. We could also write $\alpha - 1$ as $\lim_{n\to\infty} \mathbb{E}(\widetilde{F}_n)^{\frac{1}{n}}$, using the equivalent definition of exponential growth, although this converges much more slowly. It is quite remarkable that this growth rate is an algebraic number of degree 3, since we know nothing about the nature of Viswanath's constant, which we can write as $\lim_{n\to\infty} \mathbb{E}(t_n{}^{\frac{1}{n}})$. Here we are taking the average value of the growth rates of all 2^{n-2} random Fibonacci sequences, whereas Rittaud took the average value of all 2^{n-2} terms in the n^{th} position of the random Fibonacci sequences and then found the growth rate of that average. Rittaud proved his result by introducing the tree R, which is essentially a reduction of the tree $\widetilde{T}$. He found the growth rate of the expected value of terms in R, then translated this back to $\widetilde{T}$. Rittaud also makes note of some of the interesting properties of the tree R, a tree which he believes no one has previously studied. We will discuss the tree R and Rittaud's methods in greater detail in Chapter 5.

Recall that for $p = \frac{1}{2}$ the trees $|T_1|$, $|T_2|$ and $\widetilde{T}$ in Figure 1.1 are all equivalent under reorientation of the branches. Theorem 1.3 tells us that the growth rate of the expected value of n^{th} terms (Rittaud's result) is the same for the absolute values of the random Fibonacci sequences defined by Equations (1.6), (1.7) and (1.8) (corresponding to the three different trees).

Independently of Rittaud, Kalmár-Nagy [43] derived what he calls the "Fibonacci graph". It is very similar in structure to Rittaud's [64] tree R and exhibits the same interesting properties. He starts with the linear random Fibonacci sequence defined by Equation (1.7) and considers the vectors $(x_{n-1}, x_n)^T$ as points in an integer lattice (like Viswanath's [72] random walk description). Noticing symmetry in the lattice, he reduces his maps to $A : (i, j) \mapsto (j, i+j)$ and $B : (i, j) \mapsto (j, |i-j|)$, which is equivalent to the non-linear definition for the random Fibonacci sequence given in Equation (1.8). Kalmár-Nagy unfolds the lattice paths to form a graph, with two colors for directed edges, denoting the two maps. He makes two important observations about this resulting structure. First, he notices that each pair of coordinates is relatively prime, which is equivalent to being a visible point in the lattice (i.e., the straight line connecting it to the origin contains no other lattice points). Further, all relatively prime pairs (rational numbers) appear. Second he notices the graph is made up of loops, resulting from the fact that the map A followed by two B's brings us back to our starting point. This is the key to Rittaud's discovery and analysis of the tree R.

Kalmár-Nagy [44] finds an interesting result which is similar to the tree T_2 derived from the random Fibonacci sequence given by Equation (1.7). He describes a "multiset-valued Fibonacci-type recurrence" where, starting with $\{1\}$ and $\{1\}$, each multiset (each element may occur multiple times) is derived by taking a union (in fact a *multiset sum* because we have repetition) of the Minkowski sums and differences of the previous two multisets. In other words, if we were given the multisets τ_{n-1} and τ_n, we would take the union of all possible pairwise sums and differences occurring between two rows, not just between parent and grandparent. Kalmár-Nagy cleverly derives a closed-form generating function that characterizes the multisets and uses this to show that the growth rate of the geometric mean (in absolute value) of the multisets is $\sqrt{\phi} = 1.272019649\ldots$, where ϕ is the golden ratio.

In 1999 Embree and Trefethen [24] generalized Viswanath's random Fibonacci sequence by incorporating the parameter β as follows:

$$x_n = x_{n-1} \pm \beta x_{n-2}, \tag{1.19}$$

where $x_0 = x_1 = 1$ and the $\pm$ sign is chosen with $p = \frac{1}{2}$. (They consider this a generalization of the case where both signs are chosen independently with $p = \frac{1}{2}$.) This idea was motivated by the fact that when $\beta = \frac{1}{2}$, the sequence $\{x_n\}$ almost surely decays exponentially at the rate 0.929, as opposed to Viswanath's exponential growth rate of $1.1319\ldots$ when $\beta = 1$. They used random matrix products comprised of $A = \left(\begin{smallmatrix} 0 & 1 \\ \beta & 1 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & 1 \\ -\beta & 1 \end{smallmatrix}\right)$.

Embree and Trefethen considered two different methods, the first of which was Monte Carlo experimentation for large n. They refined their results by copying the method Viswanath used, i.e., viewing the random recurrence as a Markov chain and finding the associated invariant measure. (Note that for $\beta = 1$ Viswanath found an exact expression for the invariant measure of an interval whereas Embree and Trefethen approximated the measure numerically.) What they showed was that given $\beta^* \approx 0.70258$, the sequence defined by Equation (1.19) grows exponentially almost surely for $\beta > \beta^*$, and almost surely decays exponentially for $\beta < \beta^*$. This constant has been referred to as the "Embree-Trefethen constant" (see the Online Encyclopedia of Integer Sequences, [67, A118288]). They call the growth rate associated with β the "Lyapunov constant", denoted the function by $\sigma(\beta)$. Furthermore, the value of β resulting in the maximum rate of decay is estimated to be 0.36747, giving a decay rate of 0.8951. Interestingly, a plot of β versus the growth rate appears to be fractal in nature. They further considered the asymptotic behaviour of $\sigma(\beta)$ as $\beta \to 0$ and $\beta \to \infty$, and give the first couple of terms in the expansion. Embree and Trefethen point out that the recurrence

$$x_n = \alpha x_{n-1} \pm \beta x_{n-2},$$

can be reduced to Equation (1.19) via a simple substitution.

In their closing remarks, they suggest considering the sign changes in a random Fibonacci sequence and define the *sign-flip frequency*, $f(\beta)$, to be the proportion of values x_n with $x_n x_{n-1} < 0$ as $n \to \infty$. They also mention the possibilities of $p \neq \frac{1}{2}$

for sign choices, replacing the choice of coefficients $\{-\beta, \beta\}$ by the points $\{e^{i\theta}\beta\}$ on a complex circle with uniform probability distribution (in which case no value of β results in decay) and increasing the number of terms in the recurrence.

In 2005 Makover and McGowan [53] made two contributions to the theory of random Fibonacci sequences, both of which were made using binary trees to represent our set of possible sequences. Their first result is that $\{\mathbb{E}(|t_n|)\}$ grows exponentially; specifically

$$1.12095 \leq (\mathbb{E}(|t_n|))^{\frac{1}{n}} \leq 1.23375, \tag{1.20}$$

where t_n is Viswanath's random Fibonacci sequence. Recall that Rittaud [64] later finds and exact value for this growth rate. Note that if they considered the sequence $\{\mathbb{E}(t_n)\}$, they would have, by linearity of expectation,

$$\begin{aligned}
\mathbb{E}(t_n) &= \mathbb{E}(\pm t_{n-1}) + \mathbb{E}(t_{n-2}) \\
&= \frac{1}{2}(t_{n-1}) + \frac{1}{2}(-t_{n-1}) + \mathbb{E}(t_{n-2}) = \mathbb{E}(t_{n-2}) \\
&= \mathbb{E}(t_{n-4}) = \cdots = \mathbb{E}(t_1) \text{ or } \mathbb{E}(t_2) = 1.
\end{aligned}$$

Note the difference between this expected value and that given by Equation (1.17) for Viswanath's case of two $\pm$ signs. The latter gives $\mathbb{E}(t_n) = 0$, whereas in the case of one $\pm$ sign we obtain $\mathbb{E}(t_n) = 1$.

The proof of their result is completely elementary. Makover and McGowan start with a general term a having children b_1 and b_2, and consider the next levels of the binary tree. Left and right children are given by the absolute value of the sum and difference, respectively, of the parent and grandparent. By taking the absolute values of terms while generating the tree, a reorientation of the leaves occurs (as opposed to taking absolute values after the tree is generated as in $|T_1|$) and the mean is not affected. This idea was further explained in Section 1.3. Also recall that we could have placed the $\pm$ sign in front of t_{n-2} or both t_{n-1} and t_{n-2} with equal results. Considering absolute values, Makover and McGowan were able to find upper and lower bounds for the sum of a's great-grandchildren. Further, they found nice recurrence inequalities in the sums of the rows, $S(n)$, for these upper and lower bounds:

$$4S(n-3) + S(n-2) + S(n-1) \leq S(n) \leq 4S(n-3) + 2S(n-2) + S(n-1).$$

The equations $4S(n-3)+S(n-2)+S(n-1)-S(n)=0$ and $4S(n-3)+2S(n-2)+S(n-1)-S(n)=0$ have the corresponding irreducible cubic polynomials, $x^3-x^2-x-4=0$ and $x^3-x^2-2x-4=0$, which when solved determine bounds on the growth rate of $S(n)$. Dividing by 2 gives the bounds on the growth rate of the expected value of $(|t_n|)$ given in (1.20).

A similar tree is constructed for the recurrence

$$x_n = \pm\beta x_{n-1} + x_{n-2}, \tag{1.21}$$

and again the sum of the great-grandchildren of a is considered. Notice that we can instead use the recurrence in Equation (1.19), given by Embree and Trefethen [24], which is equivalent when considering $|x_n|$ for $p=\frac{1}{2}$. Since this work of Makover and McGowan is based on that of Embree and Trefethen, we will continue to use the coefficient β, rather than use our convention of α for the first coefficient. Considering the different configurations of absolute values, there are six possible sums for the great-grandchildren of a, only one of which contains a subtraction, and so is the only option where exponential decay is possible. The restriction imposed on β implies that decay can only occur for $\beta^2 < \frac{1}{2}$, i.e., $\beta < \frac{1}{\sqrt{2}} \approx 0.7071\ldots$. This is very close to Embree and Trefethen's critical value for which exponential decay can occur, namely, $0.70258\ldots$.

The difficulty in computing Viswanath's constant lies in the fractal nature of the invariant measure required for Furstenberg's formula. In 2000, Wright and Trefethen [75] studied the random Fibonacci-type sequence

$$x_n = x_{n-1} + \beta_n x_{n-2}, \tag{1.22}$$

where the β_n are independent, normally distributed coefficients. In this case the corresponding invariant measure becomes piecewise smooth, and hence much easier to deal with. Approximations to the Lyapunov exponent are calculated and Richardson extrapolation is used to improve the accuracy. Wright and Trefethen showed that the sequence defined by Equation (1.22) almost surely grows exponentially at the rate $1.0574735537\ldots$. Using the same method, they showed that the sequences defined by

$$x_n = \alpha_n x_{n-1} + x_{n-2},$$

where the α_n are also independent, normally distributed coefficients, and

$$x_n = \alpha_n x_{n-1} + \beta_n x_{n-2}$$

have almost sure growth and decay rates of 1.1149200917... and 0.9949018837... respectively.

Sire and Krapivsky [66], like Embree and Trefethen [24], study Equation (1.19). They make the interesting note that for $\beta < 0.70258$, although we have exponential decay for the sequence $\{|x_n|\}$, we have that the expected value is constant, and further, the expected values of the higher order moments grow exponentially. Sire and Krapivsky used perturbation theory and the Riccati variable

$$R_n = \frac{x_{n+1}}{x_n}$$

to extend Embree and Trefethen's [24] asymptotic expansion of $\sigma(\beta)$. (They do this in terms of the Lyapunov exponent $\log(\sigma(\beta))$ rather than the growth rate $\sigma(\beta)$.) They showed that when $\beta < \frac{1}{4}$ the Lyapunov exponent is an analytic function of β and obtain exact non-perturbative results for the $\beta = 1$ case. Sire and Krapivsky also considered the random Fibonacci-type sequence generated by

$$x_{n+1} = x_n + c\beta_n x_{n-1},$$

where the β_n are independent and normally distributed random variables (like that studied by Wright and Trefethen [75]). They again used perturbation theory to give asymptotic expansions of $\log(\sigma(\beta))$.

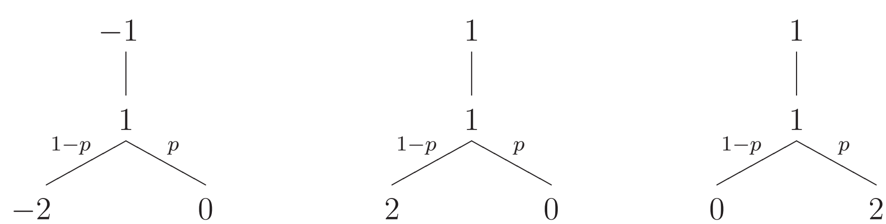
In [3], Bai uses the transfer matrix approach, which comes from statistical physics, to compute Lyapunov exponents. The leading eigenvalue of the transfer matrix (related to the Lyapunov exponent) is of particular interest, and this method gives a more theoretical understanding of the convergence occurring as a result of the methods in Bai [2]. In particular Bai considers the random Fibonacci sequence studied by Embree and Trefethen [24] and Sire and Krapivsky [66] (see Equation (1.19)). He improves some results for $\beta \leq \frac{1}{4}$, including extending the coefficients of the analytic function for $\log(\sigma(\beta))$ with $\beta < \frac{1}{4}$ found in Sire and Krapivsky [66]. Bai introduces and studies a generalized Lyapunov exponent, $\tau(q)$, which is defined as the growth rate of the log of the ensemble average (expected value) of the n^{th} term $|x_n|^q$ for all sequences, i.e., he studies the moments. This quantity takes into account the whole spectrum of sequence behaviour, not just the most probable.

1.7 Changing the Probability, and Further Generalizations

We have considered the general random Fibonacci-type sequence, where terms have coefficients α_n and β_n which are chosen according to some probability distribution. It is possible for α_n to be chosen, for example, from a set of two integers which have unequal probabilities of being selected.

Results analogous to Viswanath's case exist when $P(+), P(-) \neq \frac{1}{2}$. Recall that this is an unsigned Bernoulli distribution. We could think of flipping a biased coin for instance, where $P(+) = p$ and $P(-) = q = 1 - p$. There is a major difference here, however. When $p \neq \frac{1}{2}$, the distributions of our three positive sequences, as described in Theorem 1.2, no longer remain equal. The trees are generated the same way but in this case each branch is not equally likely to occur as it was for $p = \frac{1}{2}$. As a result, the expected value of a term in the n^{th} row is no longer simply the mean value of all terms. We must consider the edges of our trees to be weighted. Let $\widetilde{T}(p, 1)$ denote the tree T where $P(+) = p$. (The 1 denotes a generalization of coefficients which we will soon see.)

Example 1.2. Consider the tree segments in Figure 1.3, which occur in the fourth row of the full binary trees. Negative entries in trees $T_1(p, 1)$ and $T_2(p, 1)$ can cause



(a) $T_1(p, 1)$ (b) $|T_1(p, 1)|$ (c) $\widetilde{T}(p, 1)$

Figure 1.3: Tree segments with weighted edges.

left and right children to be switched compared to $\widetilde{T}(p, 1)$. This can be seen when we compare the positive tree $|T_1(p,1)|$ in Figure 1.3(b) with the tree $\widetilde{T}(p,1)$ in Figure 1.3(c). This permutation of entries changes the distribution of paths in the tree; for example in $|T_1(p,1)|$ the branch $(1,1,0)$ occurs with probability p, whereas in $\widetilde{T}(p,1)$ it occurs with probability $1-p$. Also, the expected value of row 2 of $|T_1(p,1)|$ is $\mathbb{E}(R_2) = 0p + 2(1-p) = 2-2p$ and for $\widetilde{T}(p,1)$ we have $\mathbb{E}(R_2) = 2p + 0(1-p) = 2p$.

We have proved in Theorem 1.3 that Equations (1.6) and (1.7) give sequences (in absolute value) with growth rates equal to Viswanath's constant, when $p = \frac{1}{2}$ and that Equation (1.5) also defines a sequence with the same growth rate. When we change the probability, it is not necessarily true that these variations of the random Fibonacci sequence all behave the same way. Furthermore, the techniques used by Viswanath to evaluate γ_f do not seem to carry over to this case. We can now think of the Lyapunov exponent as being a function of p, which we will denote $\gamma_f(p)$. A result of Peres [60] implies that this function is real and analytic. This may be an indication that there exists a short analytic description of γ_f, but because γ_f is related to a fractal, this seems unlikely.

Consider the random Fibonacci sequence given by Equation (1.5), namely

$$t_n = \pm t_{n-1} \pm t_{n-2},$$

where we choose $+$ with probability p, and each sign is chosen independently. Viswanath shows that $\gamma_f(p)$ is increasing on $[0,1]$. However, Hayes [37] states that numerical results showed that adding a bias, whether toward $+$ or $-$, did cause an increase in growth rate. It appears that for equal probabilities, the growth rate is minimized at $1.13198824\ldots$, i.e., $\min(\gamma_f(p))$ occurs at $p = \frac{1}{2}$, and increasing the bias increases the growth rate up to $1.618033898\ldots$, which corresponds to all-plus or all-minus sequences. This seems to contradict the fact that Viswanath claimed $\gamma_f(p)$ is increasing. The problem lies in which definition of the random Fibonacci sequence we are looking at. Consider the following example.

Example 1.3. Let $p = 0$ so that we choose the $-$ sign each time and our sequences are deterministic. First consider the sequence obtained from Viswanath's original formulation in Equation (1.5):

$$t_n = -t_{n-1} - t_{n-2}.$$

Letting $t_1 = t_2 = 1$, we can calculate the sequence as follows:

$$1, 1, -2, 1, 1, -2, \ldots,$$

where the sequence repeats after 3 terms and does not grow exponentially. Here $\gamma_f(0) = 0$, which is the minimum value. This is in agreement with Viswanath's description of an increasing function $\gamma_f(p)$.

Now if we consider the sequence obtained from Equation (1.6),

$$t_n = -t_{n-1} + t_{n-2},$$

we get

$$1, 1, 0, 1, -1, 2, -3, 5, -8, \ldots,$$

which clearly has growth rate ϕ, and $\gamma_f(0) = \log(\phi)$. This agrees with Hayes' note that we have $\max(\gamma_f(p))$ for $p = 0, 1$. Furthermore, if we consider Equation (1.7) we get

$$f_n = f_{n-1} - f_{n-2},$$

which gives the sequence

$$1, 1, 0, -1, -1, 0, 1, 1, \ldots,$$

which again repeats and has $\gamma_f(0) = 0$. This is in fact the linear case considered by Janvresse et al. [40]. Interestingly, they study properties of the functions $\gamma_f(p)$ and $\widetilde{\gamma}_f(p)$ (corresponding to the recurrences in Equations (1.7) and (1.8) respectively) and show that the former function of p is increasing, which agrees with the growth of the previous sequence.

In [40], Janvresse et al. have shown that in the linear case, for $0 < p \leq 1$, the sequence $\{f_n\}$ grows exponentially at an almost sure rate given by the increasing function

$$\gamma_f(p) = \lim_{n \to \infty} \frac{1}{n} \log |f_n| = \int_0^\infty \log x \, d\nu_\alpha(x),$$

where α is an explicit function of p and ν_α is an explicit probability distribution on $\mathbb{R}^+$. In the non-linear case $\{\widetilde{f}_n\}$ grows exponentially according to the same almost sure expression with a different function α for $\frac{1}{3} \leq p \leq 1$. For $0 \leq p \leq \frac{1}{3}$, the largest Lyapunov exponent, $\widetilde{\gamma}_f(p)$, is zero, i.e.,

$$\widetilde{\gamma}_f(p) = \lim_{n\to\infty} \frac{1}{n} \log \widetilde{f}_n = 0,$$

and so $\{\widetilde{f}_n\}$ has growth rate 1. For $(\frac{1}{3}, 1]$, the function $\widetilde{\gamma}_f(p)$ is increasing.

Janvresse et al. make use of the reduction of random Fibonacci sequences given in Rittaud [64], rather than use Furstenberg's formula, as was done by Viswanath [72]. (The difficult part here lies in the determination of Furstenberg's invariant measure.) We mentioned that Rittaud constructed a subtree R of $\widetilde{T}$ in the non-linear case (or T_2 in the linear case) which removes repetition. This was done by observing the fact that in $\widetilde{T}$, following a path RRL will bring you back to the same edge. In T_2, a similar repetition is uncovered, except here we must take negative values into account. In terms of matrices we have

$$ABBB = -A,$$
$$ABBA = -B.$$

(Note that these matrices, $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$, are slightly different from Viswanath's because here Janvresse et al. are using Equation (1.7) rather than Equation (1.6), i.e., their $\pm$ sign belongs to the second term. Also Janvresse et al. use right multiplication, while Viswanath uses left. This will be discussed further in Chapter 2.) This tree R will be extensively studied in Chapter 5.

After deleting sequences of R's and L's (or equivalently, A's and B's), Janvresse et al. study the survival probability of a term R, and the probability distribution of reduced sequences, i.e., branches in the tree R, which lead to two different functions $\alpha(p)$ for the linear and non-linear cases. Branches of the tree R are divided into right steps (R) and elbows (RL) and interestingly, this decomposition defines the continued fraction given by the ratio of the final two nodes (for some finite branch length). A similar and nice connection between the Stern-Brocot tree and continued fractions is given in Graham et al. [33, p. 305]. Here, a sequence of L's and R's in the Stern-Brocot tree determines the partial quotient of the continued fraction of the rational number reached. The above decomposition also aids in the construction of the Stern-Brocot intervals, on which the probability distribution ν_α, which is required for computation of the Lyapunov exponents $\gamma_f(p)$ and $\widetilde{\gamma}_f(p)$, is defined.

In [42], Janvresse et al. generalize the results in their paper [40] by introducing a coefficient α, to obtain the sequences

$$\begin{aligned} f_n &= \alpha f_{n-1} \pm f_{n-2}, \\ \widetilde{f}_n &= |\alpha \widetilde{f}_{n-1} \pm \widetilde{f}_{n-2}|, \end{aligned} \tag{1.23}$$

for the linear and non-linear cases respectively. We still have that each $\pm$ sign is chosen independently and $+$ is chosen with probability p. Janvresse et al. call this generalization a (p, α)-*random Fibonacci sequence*. This is a special case of the random Fibonacci-type sequence, defined earlier. The previous discussion of [40] by Janvresse et al. concerned $(p, 1)$-random Fibonacci sequences, and Viswanath's case deals with $(\frac{1}{2}, 1)$-random sequences.

Using the methods from [40], Janvresse et al. have shown that for the special case of $\alpha = \lambda_k = 2\cos(\pi/k)$ with integer $k \geq 3$, the upper Lyapunov exponent for $\{|f_n|\}$ (i.e, the log of the growth rate) for $p \in (0, 1]$ is almost surely positive and given by

$$\gamma_f(p, \lambda_k) = \lim_{n\to\infty} \frac{1}{n} \log |f_n| = \int_0^\infty \log x \, d\nu_{k,\rho}(x) > 0,$$

where ρ is an explicit function of p which is dependent on k, and $\nu_{k,\rho}$ is a probability distribution defined on generalized Stern-Brocot intervals. Here we have extended the upper Lyapunov exponent from $\gamma_f(p)$ to $\gamma_f(p, \alpha)$ to denote the new coefficient. (Note that $\gamma_f(\frac{1}{2}, \alpha)$ is a Lyapunov exponent like $\log(\sigma(\beta))$, except the placement of the coefficient differs.) In the non-linear case we have a similar result for the growth of $\{\widetilde{f}_n\}$. For $p \in (\frac{1}{k}, 1]$, the Lyapunov exponent $\widetilde{\gamma}_f(p, \lambda_k)$ is given by the same integral expression, with a slightly different ρ. These results are valid for any positive initial values f_1, f_2 and $\widetilde{f}_1, \widetilde{f}_2$. In fact, for $p \leq \frac{1}{k}$, the behaviour of $\{\widetilde{f}_n\}$ strongly depends on these initial values.

For $\alpha \geq 2$, the linear and non-linear cases are essentially the same because for n large enough we do not need to take absolute values, and it is much easier to determine the exponential growth. We arrive at a similar expression for the Lyapunov exponent, where a different type of probability measure $\mu_{p,\alpha}$ is used.

The connection between random Fibonacci sequences and continued fractions (as seen in Janvresse et al. [40]) still holds for λ_k, and a special type of continued fraction — the Rosen continued fraction — is used. Here partial quotients (of α-Rosen continued fractions) have the form $a_n\alpha$ for $a_n \in \mathbb{Z}\backslash\{0\}$ and $1 \le \alpha < 2$. It is shown that positive real numbers having a finite α-Rosen continued fraction give endpoints on the generalized Stern-Brocot intervals.

Janvresse et al. note that in the non-linear case, for λ_k and $0 \le p \le \frac{1}{k}$, there exists almost surely a bounded subsequence $\{\widetilde{f}_{n_j}\}$ of $\{\widetilde{f}_n\}$. Further, they state necessary and sufficient conditions for the sequence $\{\widetilde{f}_n\}$ to be ultimately periodic. (In Chapter 2, for the $\lambda_k = 1$ and $p = \frac{1}{2}$ case, we will again run across bounded subsequences and conditions for periodicity.) Also, if we choose $p = 0$ in the linear case for λ_k, we obtain a period-k sequence $\{|f_n|\}$. For $p = 1$, in either the linear or the non-linear case, our sequence is again deterministic and grows exponentially with rate

$$\gamma_f(1,\alpha) = \widetilde{\gamma}_f(1,\alpha) = \log\left(\frac{\alpha + \sqrt{\alpha^2+4}}{2}\right).$$

The $p = 1$ case means we are choosing $+$ at every step, so for $\alpha = 1$ it makes sense that our Lyapunov exponent is $\log\left(\frac{1+\sqrt{5}}{2}\right) = \log\phi$.

It is important to note that the even more generalized sequence given by

$$f_n = \alpha f_{n-1} \pm \beta f_{n-2}$$

can be reduced to Equation (1.23) by instead considering the sequence $\{g_n\} = \left\{\frac{f_n}{\beta^{n/2}}\right\}$, and similarly for the non-linear case. Embree and Trefethen [24] study a very similar random Fibonacci sequence, given by Equation (1.19), for $p = \frac{1}{2}$. We can rescale their sequence using $\alpha = \frac{1}{\sqrt{\beta}}$ to obtain that of Janvresse et al. in (1.23) , but the exponential growth is not preserved. Embree and Trefethen showed that their sequence decays exponentially for $\beta < \beta^* \approx 0.70258$. This corresponds to $\alpha > 1.19303$. Janvresse et al. also consider the sign-flip frequency introduced by Embree and Trefethen.

The third paper we consider by Janvresse et al. [41], extends ideas in Rittaud [64] by considering the growth of the expected value of (p,α)-random Fibonacci sequences for the non-linear case. Suppose we have $p_c = (2-\lambda_k)/4$ where $\alpha = \lambda_k = 2\cos(\pi/k)$ with $k \ge 3$, and m_n is the expected value of the n^{th} term of a (p,λ_k)-random Fibonacci sequence, where $p \in [0,1]$. (Note that if $k = 3$, we have $\lambda_k = 1$.) Then, if $p > p_c$,

$$\lim_{n\to\infty} \frac{m_{n+1}}{m_n} = \alpha_k(p)\left(1 + \frac{pq^{k-1}}{\alpha_k(p)^k}\right) > 1,$$

where $\alpha_k(p)$ is the only positive root of the order-$2k$ polynomial

$$P_k(X) := X^{2k} - \lambda_k X^{2k-1} - (2p-1)X^{2k-2} - \lambda_k pq^{k-1}X^{k-1} - p^2q^{2k-2}.$$

Furthermore, if $p = p_c$, then the sequence grows at most linearly and if $p < p_c$ the sequence is bounded. It is easily verified that for $k = 3$ and $p = \frac{1}{2}$ we obtain the result of Rittaud [64], namely, a growth rate of $1.20556943\ldots$. As in Janvresse et al. [42] the case for $\alpha \geq 2$ and $0 < p \leq 1$ can also be considered, giving

$$\lim_{n\to\infty} \frac{m_{n+1}}{m_n} = \frac{\alpha + \sqrt{\alpha^2 + 4(2p-1)}}{2}.$$

To prove these results we need to make use of the generalized tree $\widetilde{T}^{(a,b)}(p,\alpha)$, which has positive initial nodes a and b, and is derived from a (p,α)-random Fibonacci sequence. (Recall the tree $\widetilde{T}(p,1)$ discussed earlier and note that $\widetilde{T}^{(1,1)}(\frac{1}{2},1) = \widetilde{T}$.) A similar tree was considered by Makover and McGowan [53] for $p = \frac{1}{2}$ using Equation (1.21). The probability $p \neq \frac{1}{2}$ gives us (unequally) weighted edges. Given a node z with parent y, its left and right children are $|\alpha z - y|$ and $\alpha z + y$ respectively. Absolute values are taken at each step, as with Makover and McGowan's tree.

The result for the $\alpha \geq 2$ case is much easier to prove. For example, if $b \geq a$, we have that $z \geq y$, i.e., all children are greater than their parents. This removes the need for absolute values in the tree and a linear second order recurrence for the sums of rows (and hence, expected values) is straightforward to derive. For the λ_k case, Janvresse et al. found that the results of Rittaud [64] concerning properties of R generalized in a very natural way to the corresponding subtree $R^{(a,b)}(p,\lambda_k)$ of $\widetilde{T}^{(a,b)}(p,\lambda_k)$. As in [64], the sums of rows of $R^{(a,b)}(p,\lambda_k)$ will be considered, rather than sums of rows of $\widetilde{T}^{(a,b)}(p,\lambda_k)$. Sums in the latter tree can be approximated by partitioning $\widetilde{T}^{(a,b)}(p,\lambda_k)$ into infinitely many copies of the tree $R^{(l_{s+1},l_{s+2})}(p,\lambda_k)$. Here the initial nodes l_{s+1} and l_{s+2} are consecutive nodes in the leftmost branch of $\widetilde{T}^{(a,b)}(p,\lambda_k)$, i.e., the entries in the sequence obtained by taking differences only. We will look at this idea in much more detail in Chapter 5).

We have seen that in Janvresse et al. [42] the linear case was more difficult to deal with than the non-linear case. The same holds true for our expected values. In fact, no results exist for the linear case although it is suspected in Janvresse et al. [41] that similar methods to those contained in this paper may work. Janvresse et al. point out that numerical evidence for the growth rate of the expected value of a random Fibonacci sequence is difficult to obtain!

In [4], Bai uses the transfer matrix approach to study the linear and non-linear $(p, 1)$-random Fibonacci sequences given by Janvresse et al. in [41]. He shows that there exists a critical value $q*$, under which the ensemble average of $|x_n|^q$ is almost surely non-increasing or linearly increasing, and above which the average is almost surely growing exponentially. This number $q*$, as well as the generalized Lyapunov exponent $\tau(q)$ (growth rate of the logarithm of a moment) can be calculated using the transfer operator method. Furthermore, if q is a positive integer, $\tau(q)$ can be exactly determined by a system of polynomial equations. Janvresse et al. determined $\tau(1)$ analytically for the non-linear model, and left the linear model as an open question, where Bai has found results for both models for $q \geq 1$. He also generalizes these ideas to the case of (p, α)-random Fibonacci sequences with coefficient λ_k considered by Janvresse et al. [41]. Bai maps the standard matrices A and B to non-negative matrices using the reduction method employed by Janvresse et al. [41] and Rittaud [64].

Lan [48] and Cureg and Mukherjea [21] look at the growth rates of random Fibonacci-type sequences with $p \neq \frac{1}{2}$, as well as the generalization with coefficient β given by Embree and Trefethen in Equation (1.19), using numerical methods. Both papers looked at the Lyapunov exponent and/or growth rate as a function of β or p. Lan uses a generating function to represent the distribution of direction vectors at each step and makes use of an operator on these functions. The Lyapunov exponent is then represented as a linear functional of a generating function. A new numerical scheme is then used to obtain results. He avoids the need to approximate the fractal measure by using functional iterations which are made smooth. His technique can be applied to random sequences with a one-step memory (in our case using the product of random matrices), and improves the efficiency of previous methods. It may be possible to extend it to sequences with a two-step memory, but anything larger would not be computationally feasible. Lan also considers a coefficient β which is chosen at random from a continuous distribution. He generates figures depicting the dependence of the Lyapunov exponent on the parameters, and also gives asymptotic expansions of $\log(\sigma(\beta))$, as was done by Embree and Trefethen [24] and Sire and Krapivsky [66].

Cureg and Mukherjea follow the method of Embree and Trefethen, and use Furstenberg's formula to study Lyapunov exponents. Since there is no known closed form for the invariant measure when $p \neq \frac{1}{2}$, Cureg and Mukherjea discretize the measure and use numerical methods to evaluate invariant measures and Lyapunov exponents as functions of p and β, giving numerous figures as illustration of their results. Changes in smoothness of the invariant measure can be noted in different cases. They also look at the constant $\gamma_f(p)$ for the Equations (1.6) and (1.7). (Recall that Viswanath only considered this constant, which was shown to be a smooth function of p by Peres [60], for Equation (1.5), which has two choices of $\pm$.) They give evidence that when $\beta \geq 1$, the growth rate, $\gamma_f(p, \beta)$, is always greater than 1, regardless of the value of p. For $\beta < 1$, it seems there is a critical value of p at which the sequence neither grows nor decays, similar to Embree and Trefethen's constant β^*. The paper gives a very excellent and well-written overview of Viswanath's problem and some of the further work done.

Chan [14] considers another random Fibonacci-type sequence, given by $x_{-1} = 0$, $x_0 = 1$, $a_0 = 0$ and

$$x_n = 2^{a_n} x_{n-1} + 2^{a_{n-1}} x_{n-2}, \tag{1.24}$$

where $\{a_n\}$ is an infinite sequence with $a_n \in \mathbb{N}$ for $n \geq 1$. He starts with a randomly chosen $\gamma \in [0, 1)$, and writes a continued fraction in the following form:

$$\gamma = \cfrac{2^{-a_1}}{1 + \cfrac{2^{-a_2}}{1 + \cfrac{2^{-a_3}}{1 + \ddots}}}.$$

The sequence $\{a_n\}$ then gives the exponents of the coefficients in the recurrence given in Equation (1.24). He uses continued fractions and ergodic theory to prove that the growth rate of $\{x_n\}$ is

$$\lim_{n\to\infty} x_n^{\frac{1}{n}} = e^{1.30022988\ldots}$$

for almost all $\gamma \in [0,1)$,

Ergodic theory is a mathematical theory similar to chaos theory, which was developed by Birkhoff, von Neumann, Khinchin and others, and is based on 19$^{\text{th}}$ century physics. Essentially, it says that given a dynamical system subject to certain conditions, with some physical quality P, the time average of P and the space average of P are the same (Chan [14]). Random sequences can be thought of as being generated by certain dynamical systems, for example, Chan's dynamics on continued fractions ([14]).

Chan's result is actually a generalization of the following theorem due to Lévy [52]. If we again consider a randomly chosen $\gamma \in [0,1)$, and associate with it the infinite sequence $\{\alpha_n\}$ of natural numbers obtained from the continued fraction

$$\gamma = \cfrac{1}{\alpha_1 + \cfrac{1}{\alpha_2 + \ddots}},$$

then the random Fibonacci-type sequence defined by $x_{-1} = 0$, $x_0 = 1$ and

$$x_n = \alpha_n x_{n-1} + x_{n-2}$$

has "growth constant" (upper Lyapunov exponent)

$$\lim_{n\to\infty} \frac{1}{n} \log x_n = \frac{\pi^2}{12 \log 2} = 1.186569110\ldots \tag{1.25}$$

for almost all γ. Chan notes that this theorem of Lévy is hard to generalize and requires an analytical closed form of the invariant measure of the random recurrence.

Chan [15] generalizes the results of his paper [14] by considering the random Fibonacci-type sequence defined by $x_{-1} = 0$, $x_0 = 1$, $a_0 = 0$ and

$$x_n = k^{a_n} x_{n-1} + (k-1)k^{a_{n-1}} x_{n-2},$$

where again $\{a_n\}$ is an infinite sequence of natural numbers generated by the continued fraction

$$\gamma = \cfrac{k^{-a_1}}{1 + \cfrac{(k-1)k^{-a_2}}{1 + \cfrac{(k-1)k^{-a_3}}{1 + \ddots}}},$$

for $\gamma \in [0, 1)$ and a fixed k. Chan uses the same method as in [14] to prove that

$$\lim_{n \to \infty} \frac{1}{n} \log x_n = c_k \int_0^1 \frac{\log(1/\gamma)}{(1 + (k-1)\gamma)(k + (k-1)\gamma)} \, d\gamma,$$

where c_k is a function of k. For $k = 2$ we obtain the case given in [14]. We can note that in this case we obtain the integral

$$c_k \int_0^1 \frac{\log(1/\gamma)}{(1+\gamma)(2+\gamma)} \, d\gamma,$$

where $c_k = \frac{1}{\log(4/3)}$. This integral can be evaluated using Maple, for example, as

$$c_k \left(\frac{\pi^2}{12} + \mathrm{Li}_2\left(\frac{3}{2}\right) \right),$$

where Li_2 is the dilog function, defined by

$$\mathrm{Li}_2(z) := \sum_{k=1}^{\infty} \frac{z^k}{k^2} = \int_1^z \frac{\log t}{1-t} \, dt.$$

Note that $\mathrm{Li}_2(1) = \zeta(2) = \pi^2/6$ and that this number also appears in Equation (1.25) in Lévy's result.

Another way to generalize the random Fibonacci sequence is to take sums of more than two terms, as described in Hayes [37]. In the non-random case, the sequence defined by $F_1 = 0$, $F_2 = F_3 = 1$ and

$$F_n = F_{n-1} + F_{n-2} + F_{n-3}$$

for $n \geq 4$ is known as the "tribonacci sequence":

$$0, 1, 1, 2, 4, 7, 13, 24, 44, 81, 149, 274, 504, \ldots.$$

It has growth rate $1.83929\ldots$, which is the unique positive root of $x^3 - x^2 - x - 1 = 0$. The analogous fourth order Fibonacci recurrence gives us the "tetrabonacci" sequence and has growth rate $1.92756\ldots$. In general, the k^{th} order Fibonacci recurrence, which defines the "k-nacci" sequence, is given by

$$F_n = F_{n-1} + F_{n-2} + \cdots + F_{n-k}, \tag{1.26}$$

where initial values are $F_1 = F_2 = \cdots = F_{k-2} = 0$ for $k \geq 3$, $F_{k-1} = F_k = 1$. As k goes to infinity (where $n > k$), we are summing all terms in the sequence to get the next one as follows:

$$F_n = F_{n-1} + F_{n-2} + \cdots + F_1,$$

where we have an infinite number of 0's followed by two 1's as initial values. For all of these extended Fibonacci cases, we can still think of the sequence as starting with $1, 1$. This sequence, the "polynacci sequence" simply gives powers of 2, and hence 2 is the growth rate. As in the regular Fibonacci case, these growth rates hold regardless of the initial values.

If we randomize these sequences, each addition is replaced with $\pm$, with probability $\frac{1}{2}$ of choosing either $+$ or $-$. Numerical experiments show that for the random tribonacci sequence, the growth rate is approximately 1.22 and the random tetrabonacci sequence has growth rate 1.27. It appears that the growth rate increases slowly as we add more terms. Finding the growth rate of the limiting case, the random polynacci sequence defined by

$$t_n = \pm t_{n-1} \pm t_{n-2} \pm \cdots \pm t_1, \tag{1.27}$$

was in fact Viswanath's original problem. He was dealing with random triangular matrices with subdiagonal entries having value ± 1 with probability $\frac{1}{2}$. Each time we generate a new term, we must randomly choose a sign for all previous terms in the recurrence. Numerical results show that the growth rate of this sequence (in absolute value) is approximately 1.32. Due to the problem's difficulty he turned to the two-termed case instead. We can also think of this polynacci recurrence as the random series

$$t_n = \sum_{i=1}^{n-1} \alpha_i t_i,$$

where each α_i is chosen from $\{1, -1\}$ with equal probability for every t_n term generated.

1.8 Applications and Open Questions

There are numerous applications of random matrix products and random Fibonacci-type sequences in a wide range of fields. Problems in one-dimensional disordered systems often reduce to determining the asymptotic properties of products of random matrices, for example, one-dimensional random Ising models (deals with ferromagnetism in statistical mechanics), randomly coupled harmonic oscillators, quantum mechanics of an electron in a one-dimensional disorder potential. Products of random matrices have also been widely used to model discrete stochastic processes such as the evolution of population or investment strategy (Bai [3]). Further, the Lyapunov exponent is a statistical quantity of great physical significance. For example, the mean free energy of a random Ising chain is given by the Lyapunov exponent γ, while the localization length of a wave in a random medium is equal to γ^{-1} (Bai [2]). In general, the calculation of Lyapunov exponents presents a considerable numerical challenge in practice.

Furstenberg and Kesten's work on random matrices has led to new uses of glass, new laser technology and even the development of copper spirals in birth control devices. The theoretical research leading to these applications earned the Nobel Prize for three physicists (Anderson, Mott and van Vleck) in 1977 for their work on "electronic structure of magnetic and disordered systems". Disordered systems are found in non-crystallic materials that have irregular atomic structures. Anderson's contribution to this work was Anderson localization — when a current is passing through a semi-conductor containing impurities, the current will stay localized at certain energies, rather than dispersing. Similarly, with the irregular molecular structure of glass we might expect that light rays will bounce around randomly causing a blurred image, but this does not happen. The repeated random movements actually lead to orderly behaviour of the light ray inside the glass (Devlin [23]). This is the same behaviour we see in the random Fibonacci sequence.

Random Fibonacci sequences have connections to other areas of mathematics apart from random matrix theory. We have seen that our matrices A and B can be seen as Möbius transformations of the complex plane. Also the random walk on slopes m can be thought of as a random dynamical system. Random sequences aid modeling in a variety of areas such as technology, sociology and economics, and examples include modeling transport on a network and income distribution. In particular, random Fibonacci-type sequences are similar to problems that arise when dealing with one-dimensional disordered systems (Krasikov et al. [46]).

Random sequences and even random Fibonacci-type sequences are of growing interest in physics and other applied sciences. As Colman and Rodgers [19] point out, they may be used to "model disordered systems with non-deterministic behaviour, such that after an initial amount of time there are a number of possibilities (with attached probabilities) for the states of the system". Statistical mechanics has used random networks to model complex phenomena, and random sequences with similar properties to the networks have become a useful area of study.

Colman and Rodgers use random sequences in their study of the example of an electrical network taking the form of a binary tree. Each node has one child with probability p and two children with probability $1-p$. He assumes each edge has resistance 1Ω and at n levels all vertices can be connected. He considers the total resistance R_n at level n as a function of p. Interestingly, it must be taken into account that when edges with resistances a and b are connected in series, total resistance is give by $R = a+b$, and total resistance across two parallel edges is given by $R = \frac{1}{a} + \frac{1}{b}$. Colman and Rodgers are interested in the expected value of the sequence R_n as well as higher order moments.

The sequence given by Equation (1.16) is analogous to a localization problem described by a discretized Schrödinger equation on a line with asymmetric hopping rates of particles. The movement of electrons on the physical system is such that at each time step, a particle can either hop to its left with rate β or to its right with rate 1. If it hops to its left, it will always hop to the same fixed location, chosen according to a probability distribution. If it hops to its right, it will always be to the nearest neighbour.

Devlin [23] points out about Viswanath's work that "An easily understood, cute, counter-intuitive result about elementary integer arithmetic can motivate a great many individuals to take a look at an area of advanced mathematics full of deep and fascinating results." An often overlooked application of a neat mathematical result such as this one is that it attracts people to study the area.

In the papers discussed throughout this chapter, there remain some open questions on the generalization of random Fibonacci sequences. Janvresse et al. note that for Equation (1.5), Viswanath's original random Fibonacci sequence with two choices of sign, the problem is not equivalent to that discussed in Janvresse et al. [40] (one choice of $\pm$ sign) when $p \neq \frac{1}{2}$, and no explicit formula for the Lyapunov exponent is known. In Janvresse et al. [41] the question of two coins alternately tossed, one with probability p the other with probability p' is posed. They also consider applying a deterministic rule, such as an irrational rotation around the circle. Further they suggest the study of the variance or other higher order moments for a general p. They even speculate that there is a connection between random Fibonacci sequences and hyperbolic geometry.

Rittaud [64] suggests defining a random Fibonacci sequence by letting

$$\beta = (e^{2i\pi/p})^{Z(w)},$$

where Z is a random variable taking values in $\{0, \ldots, p-1\}$, or some similar variation. This would result in p-ary trees rather than binary trees. He also suggests generalizing to a third order random Fibonacci-type sequence such as

$$g_n = |\alpha g_{n-1} \pm \beta g_{n-2} \pm \gamma g_{n-2}|.$$

He briefly explores a connection between continued fractions and random Fibonacci sequences. Convergents to the continued fraction of an irrational number give finite walks in the tree R, and the irrational number can be thought of as a limit walk in R. The limit walks may be connected for quadratic irrationals with the same continued fraction period. (This continued fraction connection is similar to that found in Graham et al. [33] as mentioned earlier, where irrational numbers are represented by infinite walks in the Stern-Brocot tree and convergents represented by finite ones.) Rittaud further talks about a continued fraction connection for the generalized random Fibonacci sequence $g_n = |\alpha g_{n-2} \pm \beta g_{n-2}|$, studied by Janvresse et al.

in [41].

The major component of this thesis deals with removing the randomness from Viswanath's work. Instead of tackling the growth of a random Fibonacci sequence using random matrix theory or other stochastic methods, the aim is to use matrix products which are not products of random matrices, but products obtained from periodic sequences. This idea will be presented in detail in the following chapters and was derived independently of the work of McGuire [55].

McGuire also had the idea of studying deterministic matrix products, and observing the behaviour of the corresponding sequences. He gave necessary conditions for such a sequence to be periodic, as well as the possible periods of the sequences. In [56], McGuire generalized his results from [55] by considering the random tribonacci, tetranacci, and in general, the k-nacci sequences. McGuire defines the general random k-nacci sequence slightly differently than Equation (1.26), as follows:

$$F_n = F_{n-1} \pm F_{n-2} \pm \cdots \pm F_{n-k},$$

with initial values $F_1 = F_2 = \cdots = F_{k-2} = 0$ for $k \geq 3$ and $F_{k-1} = F_k = 1$. Here the coefficient of F_{n-1} is not chosen at random. In this case, products of $(k-1) \times (k-1)$ matrices are studied. We will restrict our investigation to 2×2 matrices only.

Chapter 2

Growth Types of Periodic Coefficient Sequences

2.1 Non-Random Sequences and Matrix Representations

As mentioned in the Introduction, our aim is to remove the randomness from Viswanath's random Fibonacci sequence by generating sequences according to a fixed pattern. We begin this task by introducing a few definitions.

Definition 2.1. A *coefficient cycle* of length n is an n-tuple $\sigma_n = (s_1, \ldots, s_n)$, where $s_i \in \{+, -\}$ for $1 \leq i \leq n$.

The "non-random Fibonacci sequence" is formulated as follows.

Definition 2.2. A *periodic coefficient sequence* is given by the recursion

$$t_i = \pm t_{i-1} + t_{i-2}, \tag{2.1}$$

for $i \geq 3$, where $t_1 = t_2 = 1$ and each $\pm$ sign is chosen according to σ_n, i.e.,

$$t_i = s_{1+(i-3) \bmod n}\, t_{i-1} + t_{i-2}.$$

This allows our index to take on the values $1, \ldots, n$, and then cycle back through as i increases. A few examples here may be helpful. For simplicity of notation, we will remove the commas form the coefficient cycle σ_n when giving explicit examples.

Example 2.1. Let $\sigma_3 = (++-)$. Generating the periodic coefficient sequence gives

$$\begin{array}{lccccccccc} & + & + & - & + & + & - & + & & \dots \\ t_i = 1 & 1 & 2 & 3 & -1 & 2 & | \, 1 & 1 & & \dots, \end{array}$$

where the bar denotes repetition. Here we have

$$\begin{aligned} t_3 &= s_{1 \bmod 3}\, t_2 + t_1 = s_1(1) + (1) = 1 + 1 = 2, \\ t_4 &= s_{2 \bmod 3}\, t_3 + t_2 = s_2(2) + (1) = 2 + 1 = 3, \\ t_5 &= s_{3 \bmod 3}\, t_4 + t_3 = s_3(3) + (2) = -3 + 2 = -1, \end{aligned}$$

and so on. Notice that the terms t_i begin to repeat after two repetitions of the coefficient cycle. The period of repetition is 6, and our periodic coefficient sequence has bounded growth. Note that the period is a multiple of $n = 3$. Growth type as well as period size will be discussed later in further detail.

In 1966, Whitney [74] used a periodic coefficient in the recurrence relation $g_n = (-1)^{F_{n-2}} g_{n-1}$, where $g_n = L_{F_n}$ and L_n denotes the Lucas numbers. (Recall the Lucas numbers are defined by $L_1 = 2, L_2 = 1$, and $L_n = L_{n-1} + L_{n-2}$ for $n \geq 3$.) The parity of the Fibonacci numbers implies that the coefficients of g_{n-1} follow the pattern $-1, -1, 1, -1, -1, 1, \ldots$ for $n \geq 3$, i.e., we have coefficient cycle $\sigma_3 = (- - +)$.

Example 2.2. Let $\sigma_4 = (+ + - -)$. The periodic coefficient sequence is

$$\begin{array}{ccccccccccccccc} & + & + & - & - & + & + & - & - & + & + & - & - & + & + \\ t_i = 1 & 1 & 2 & 3 & -1 & 4 & 3 & 7 & -4 & 11 & 7 & 18 & -11 & 29 & \end{array}$$

$$\begin{array}{ccccc} - & - & \ldots & & \\ 18 & 47 & -29 & 76 & \ldots. \end{array}$$

Here it appears as if the sequence is unbounded in absolute value. The fact that the growth is actually exponential will be proven shortly.

Example 2.3. Let $\sigma_6 = (+ + + - - -)$. The periodic coefficient sequence is

$$\begin{array}{cccccccccccccc} & + & + & + & - & - & - & + & + & + & - & - & - & + \\ t_i = 1 & 1 & 2 & 3 & 5 & -2 & 7 & -9 & -2 & -11 & -13 & 2 & -15 & \end{array}$$

$$\begin{array}{cccccccc} + & + & - & - & - & \ldots & & \\ 17 & 2 & 19 & 21 & -2 & 23 & -25 & \ldots. \end{array}$$

Considering $|t_n|$, we see that 2 appears as every third term and the remaining values appear to be growing linearly.

We will soon see that these three examples illustrate the three possible types of growth for our periodic coefficient sequences. Upon inspection, there is no obvious connection between the coefficient cycle σ_n and the growth type. Increasing the number of $-$ signs does not seem to slow down growth, and retaining the same “balanced” pattern as seen in Examples 2.2 and 2.3 does not imply the sequences will have similar behaviour. For further evidence, Table 2.1 gives some examples of coefficient cycles for small n, along with the corresponding growth type, where we denote bounded, exponential and linear growth by B, E and L respectively. These growth types can be determined by generating terms of the periodic coefficient sequences, as in the previous examples, although we will soon see a method to verify these calculations.

coefficient cycle	n	growth type
$(+)$	1	E
$(+-)$	2	B
$(++-)$	3	B
$(+-+)$	3	B
$(+++-)$	4	B
$(++--)$	4	E
$(-++-)$	4	E
$(++++-)$	5	E
$(----+)$	5	E
$(+++--)$	5	E
$(++-+-)$	5	E
$(+++++-)$	6	L
$(++++--)$	6	E
$(+++---)$	6	L
$(--++-+)$	6	L

Table 2.1: Growth types of some coefficient cycles.

In order to understand the behaviour of our periodic coefficient sequences we need a different way to represent them. This can be achieved using matrix products. From Equation (1.9) we have that the linear recurrence given in Equation (2.1) can be written as the matrix equation

$$\begin{pmatrix} t_{i-1} \\ t_i \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} \begin{pmatrix} t_{i-2} \\ t_{i-1} \end{pmatrix}, \tag{2.2}$$

for $i \geq 3$, where the vector $(t_{i-2}, t_{i-1})^T$ is multiplied by a member of

$$\left\{ A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},\ B = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \right\},$$

according to the value of the term s_{i-2} in the coefficient cycle σ_n; A for $+$ and B for $-$. Note that in Viswanath's random Fibonacci sequence case, the matrix equation is the same, but we choose one of A or B with probability $\frac{1}{2}$. Also, for the regular Fibonacci sequence, we choose A every time.

Iterating Equation (2.2) and incorporating our initial values gives

$$\begin{pmatrix} t_{i-1} \\ t_i \end{pmatrix} = M_{i-2}M_{i-3}\cdots M_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

for $i \geq 3$, where $M_j \in \{A, B\}$ for $1 \leq j \leq i-2$. From this expression we can evaluate terms t_{i-1} and t_i in our sequence. Now, if we rewrite the above using $i = n+2$, each element in $\sigma_n = (s_1, \ldots, s_n)$ will be used exactly once to select a matrix from $\{A, B\}$ and we can write

$$\begin{pmatrix} t_{n+1} \\ t_{n+2} \end{pmatrix} = M_n M_{n-1}\cdots M_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{2.3}$$

for $n \geq 1$.

Definition 2.3. Given M_j for $1 \leq j \leq n$ in Equation (2.3), we define the *product matrix* P_n associated with the coefficient cycle σ_n to be

$$P_n := M_1 \cdots M_n.$$

We will later see that reversing the order of the matrices in Equation (2.3) to define P_n is permissible when analyzing the growth of the associated sequence, so we write our product matrix with increasing indices so that it better reflects the pattern in the corresponding coefficient cycle. We will still need to use the form in Equation (2.3) when finding sequence terms, however.

Example 2.4. Let $\sigma_3 = (++-)$. We have s_1 and s_2 corresponding to matrix A, and s_3 corresponding to matrix B so that

$$M_3M_2M_1 = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}.$$

If we multiply by our initial value vector we obtain

$$\begin{pmatrix} t_4 \\ t_5 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix},$$

which tells us that the fourth and fifth terms of our periodic coefficient sequence are 3 and -1 respectively, as verified in Example 2.1. Now according to Definition 2.3, we define

$$P_3 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}.$$

Note that $P_3^2 = I$, which corresponds to the fact that in Example 2.1, repeating the coefficient cycle $\sigma_3 = (++-)$ twice brought us back to the initial values $1, 1$. Also notice that reversing the matrices in the product gave us the transpose of the original matrix; we will come back to this fact.

What can we say about matrices P_n? First note that taking determinants gives

$$\det(A) = \det \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = -1,$$

$$\det(B) = \det \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = -1.$$

Therefore, by the multiplicative property of determinants, we must have $\det(P_n) = \pm 1$, i.e., $P(n)$ is *unimodular*. Further, we have that

$$G := \langle A, B \rangle \leq \mathrm{S^*L}(2, \mathbb{Z}),$$

where $\langle A, B \rangle$ is the group generated by the matrices A and B, which we denote by G, and $\mathrm{S^*L}(2, \mathbb{Z})$ is the extension of the special linear group $\mathrm{SL}(2, \mathbb{Z})$ to unimodular matrices. Note that when dealing with integer entries, the invertible matrices are precisely those with determinant ± 1, and so $\mathrm{S^*L}(2, \mathbb{Z}) = \mathrm{GL}(2, \mathbb{Z})$. The following gives a specific value to the determinant of P_n.

Proposition 2.1. *For a product matrix P_n,*

$$\det(P_n) = 1 \iff n \text{ even},$$

$$\det(P_n) = -1 \iff n \text{ odd}.$$

Proof: We have seen that matrices A and B both have determinant -1. Using the fact that determinants are multiplicative, it is then simple to conclude that $\det(P_n) = (-1)^n = 1$ if and only if n is even and $\det(P_n) = -1$ if and only if n is odd. $\diamond$

Before continuing, we should show that G is in fact a group.

Proposition 2.2. *The set of matrices G generated by all possible products of the matrices A and B is a group.*

Proof: We have seen in Example 2.4 that $AABAAB = I$. (We shall see that this is just one product of A's and B's that produces the identity.) G is closed by definition, since the product of any two elements is also comprised of A's and B's. We have associativity from matrix multiplication. Lastly, we must show that our generators A and B have inverses in G. This follows easily from the identity relation because $AABAAB = I$ implies that $ABAAB = A^{-1} = \left(\begin{smallmatrix} -1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $AABAA = B^{-1} = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. $\diamond$

Note that the relation in this proof tells us that G is not a free group. The properties of certain subgroups of matrices of $\mathrm{GL}(2, \mathbb{Z})$ and $\mathrm{SL}(2, \mathbb{Z})$ have been studied extensively. For example, Boca [9] studies the free multiplicative monoid generated by $M_1 = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ and $M_2 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. In particular, he proves an asymptotic formula for the number of elements in the monoid with trace at most n. He also uses continued fractions to show that the elements of the monoid can be uniquely written as products of M_1 and M_2. More generally Kuzmanovich and Pavlichenkov [47] have compiled results on finite groups of integral matrices, including the orders of elements of $\mathrm{GL}(\mathrm{n}, \mathbb{Z})$, and finite subgroups of $\mathrm{GL}(2, \mathbb{Z})$ and $\mathrm{GL}(\mathrm{n}, \mathbb{R})$. The elements in G cannot be written uniquely as products of our A's and B's. For example we have $I = (AB)^6 = (AAB)^2$, and another example that produces a non-identity matrix is $A^2B^2AB = A^4BA^2B^3AB = A^2B^2A^2BA^2BAB$.

In the following theorem we characterize the subgroup K of G with $K \in \mathrm{SL}(2, \mathbb{Z})$, i.e., $K = G \cap \mathrm{SL}(2, \mathbb{Z})$, the set of matrices in G with determinant 1. The set K is closed under matrix multiplication and so is a proper subgroup.

Theorem 2.1. *The group $K < G$ is composed of exactly those elements $P_n = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2, \mathbb{Z})$ with $P_n \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, or $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ when we take $a, b, c, d \pmod 2$.*

Proof: We have that $G = \langle A, B \rangle$ where $A, B = \left(\begin{smallmatrix} 0 & 1 \\ 1 & \pm 1 \end{smallmatrix}\right)$, and $G \leq \mathrm{S^*L}(2, \mathbb{Z})$. Define the map det, which takes any matrix in G to its determinant, as

$$\det : G \to \mathbb{Z}/2\mathbb{Z} = \{1, -1\}.$$

This is a homomorphism because determinants are multiplicative. Let $K = \ker(\det)$, i.e., those elements P_n in G with $\det(P_n) = 1$. K is a normal subgroup of index 2 in G.

We know from Proposition 2.1 that $\det(P_n) = 1$ if and only if n is even. Therefore $P_n \in K$ if and only if it is a product of the matrices $A^2, AB, BA, B^2, A^{-1}B, BA^{-1}$, and their inverses. These six matrices are generators for K. It is a standard result (see Lang [49, p. 4] for example) that $\mathrm{SL}(2,\mathbb{Z})$ is generated by

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

with the relations $S^4 = 1, (ST)^6 = 1$ and $S^2 = (ST)^3$. We can write each of the generators of K in terms of S and T as follows:

$$\begin{aligned} A^2 = TST^2 &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad B^2 = S^3T^2STS = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \\ AB = TS &= \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, \quad BA = S^3TST = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \\ A^{-1}B = STST^2STS^3 &= \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}, \quad BA^{-1} = S^3T^2S = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}. \end{aligned}$$

These elements in $\mathrm{SL}(2,\mathbb{Z})$ generate K as a subgroup of $\mathrm{SL}(2,\mathbb{Z})$. We can write this set of generators in terms of S and T in the simpler form $\{S^2, T^2, ST, TS\}$, i.e., each generator in the displayed list is composed of elements in the above set. We have that

$$S^2 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad T^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad ST = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, \quad TS = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \tag{2.4}$$

Note that this set of generators for K is composed of even-length products of S and T, and further, the set $\{S^2, T^2, ST, TS\}$ generates all sequences of S and T of even length. We also have that $S^{-1} = S^3$ and $T^{-1} = S^3TSTS$, and so sequences of even length in S^{-1} and T^{-1} (generated by $(S^{-1})^2, (T^{-1})^2, S^{-1}T^{-1}$ and $T^{-1}S^{-1}$) also belong to K. The subgroup of $\mathrm{SL}(2,\mathbb{Z})$ containing all even products of S and T is therefore equal to K. Notice that the relations defining $\mathrm{SL}(2,\mathbb{Z})$ are all of even length, and so equivalent representations of a product must have the same parity.

Now consider the coset TK of $\mathrm{SL}(2,\mathbb{Z})$. Elements in this set must contain an odd number of terms. Further, we can show that this coset contains exactly all terms of odd length. Let W be a word of odd length. If it begins with T, then $W \in TK$. From the relation $S^2 = (TS)^3$ we have $S = TSTST$, and so if $W = SW_1$, where $|W_1|$ is even, we have that $W = TSTSTW_1 \in TK$. Similarly, if $W = S^{-1}W_1$ or $W = T^{-1}W_1$, we can write $W = S^3W_1 = (TSTST)^3W_1$ or $W = S^3TSTSW_1 = (TSTST)^3TSTSW_1$ respectively, and both of these words belong to TK because they are of odd length. Therefore K and TK are the only cosets of $\mathrm{SL}(2,\mathbb{Z})$ and K is a subgroup of index 2.

Now, the principal congruence subgroup $\Gamma(2)$ (also called the modular group Λ) is the kernel of the homomorphism

$$\mathrm{mod}_2 : \mathrm{SL}(2,\mathbb{Z}) \to \mathrm{SL}(2,\mathbb{Z}/2\mathbb{Z}),$$

i.e., $\Gamma(2)$ is the set of matrices in $\mathrm{SL}(2,\mathbb{Z})$ which equal I when its entries are taken modulo 2. This mapping is a homomorphism because by properties of modular arithmetic, taking the entries in a product matrix modulo 2 is equivalent to first taking the entries in the matrices A, B modulo 2, and then forming the product. From Lehner [51], for example, we have that $\Gamma(2)$ is generated by

$$S^2 = -I, \quad T^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad U^2 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix},$$

where $U^2 = (T^2)^T = ST \cdot T^2 \cdot TS$. Since these generators are of even length in S and T, we have that $\Gamma(2) \leq K$. Applying the homomorphism mod_2 to K, i.e., $K \to \mathrm{SL}(2,\mathbb{Z})/\Gamma(2) = \mathrm{SL}(2,\mathbb{Z}/2\mathbb{Z})$, gives the image $\mathrm{mod}_2(K) = \{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)\}$, because this is the image of the generators (which are preserved under homomorphism), and no other elements can be created. Further, because subgroups are also preserved by homomorphisms, $\mathrm{mod}_2(K)$ is a subgroup of index 2 in $\mathrm{SL}(2,\mathbb{Z}/2\mathbb{Z})$. Therefore, since K is a subgroup of index 2 in $\mathrm{SL}(2,\mathbb{Z})$, we can conclude that

$$K = \left\{ P_n \in \mathrm{SL}(2,\mathbb{Z}) : \mathrm{mod}_2(P_n) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \right\}.$$

◇

If we extend the homomorphism mod_2 to

$$\mathrm{mod}_2 : \mathrm{S^*L}(2,\mathbb{Z}) \to \mathrm{SL}(2,\mathbb{Z}/2\mathbb{Z}), \tag{2.5}$$

any product matrix $P_n \in G$ under this map is simply a power of A because $\mathrm{mod}_2(A) = \mathrm{mod}_2(B)$. Calculating the first few powers, we see that any $P_n \in G$ must take one of the following forms with entries modulo 2:

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We now have another property of our group G.

Recall from Chapter 1 that we could have used Equation (1.7) or (1.8), i.e., $f_n = f_{n-1} \pm f_{n-2}$ or $\widetilde{f}_n = |\widetilde{f}_{n-1} \pm \widetilde{f}_{n-2}|$, to define the random Fibonacci sequence. These recurrences were designated as linear and non-linear respectively, by Janvresse et al. [40]. We have seen in Theorem 1.3 that for $p = \frac{1}{2}$, both of these recurrences define sequences with the same almost sure growth rate as the random Fibonacci sequence; however, the corresponding matrix recurrences differ.

Instead of using matrices $A, B = \left(\begin{smallmatrix} 0 & 1 \\ 1 & \pm 1 \end{smallmatrix}\right)$, the recurrence in Equation (1.7) can be written as

$$\begin{pmatrix} f_{n-1} \\ f_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \pm 1 & 1 \end{pmatrix} \begin{pmatrix} f_{n-2} \\ f_{n-1} \end{pmatrix}, \tag{2.6}$$

where we will denote $A, B' = \left(\begin{smallmatrix} 0 & 1 \\ \pm 1 & 1 \end{smallmatrix}\right)$. We have already seen this in Section 1.6, however, where Embree and Trefethen [24] use the matrix equation

$$\begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \pm\beta & 1 \end{pmatrix} \begin{pmatrix} x_{n-2} \\ x_{n-1} \end{pmatrix}.$$

to represent the more general recurrence $x_n = x_{n-1} \pm \beta x_{n-2}$, given in Equation (1.19). Alternately, Janvresse et al. [40] use the matrix equation

$$(f_{n-1}, f_n) = (f_{n-2}, f_{n-1}) \begin{pmatrix} 0 & \pm 1 \\ 1 & 1 \end{pmatrix} \tag{2.7}$$

to represent Equation (1.7), and we will denote $A, \hat{B} = \left(\begin{smallmatrix} 0 & \pm 1 \\ 1 & 1 \end{smallmatrix}\right)$. Again we have seen this, in Section 1.7, where the difference between right and left multiplication is pointed out. Note that if we use right multiplication to represent Viswanath's recurrence, $t_n = \pm t_{n-1} + t_{n-2}$, we obtain

$$(t_{n-1}, t_n) = (t_{n-2}, t_{n-1}) \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix},$$

which uses the same matrices A, B as for left multiplication.

It was also mentioned in Chapter 1 that the recurrence in Equation (1.8) requires three matrices in its matrix representation. We can use Equation (2.7) with a slight adjustment. By definition, all terms in the sequence $\{\widetilde{f}_n\}$ are positive and so we must take this into account. We may use

$$(\widetilde{f}_{n-1}, \widetilde{f}_n) = (\widetilde{f}_{n-2}, \widetilde{f}_{n-1}) \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} = (\widetilde{f}_{n-1}, \widetilde{f}_{n-1} - \widetilde{f}_{n-2})$$

when $\widetilde{f}_{n-1} \geq \widetilde{f}_{n-2}$, but if $\widetilde{f}_{n-2} > \widetilde{f}_{n-1}$, we must use

$$(\widetilde{f}_{n-1}, \widetilde{f}_n) = (\widetilde{f}_{n-2}, \widetilde{f}_{n-1}) \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = (\widetilde{f}_{n-1}, \widetilde{f}_{n-2} - \widetilde{f}_{n-1}),$$

where the third matrix $B = \left(\begin{smallmatrix} 0 & 1 \\ 1 & -1 \end{smallmatrix}\right)$ is introduced. This process is equivalent to using only the matrices $A, \hat{B}$, but taking the absolute value of the product matrix after each step. Note that this differs from the case in Equation (2.7) where a random Fibonacci sequence is generated from a product of i.i.d. random matrices, each occurring with probability $\frac{1}{2}$. In the three matrix case, the $\pm$ sign is still chosen with $p = \frac{1}{2}$, where $+$ corresponds to $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, but $-$ corresponds to $\hat{B} = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ or $B = \left(\begin{smallmatrix} 0 & 1 \\ 1 & -1 \end{smallmatrix}\right)$, depending on the value of the terms $\widetilde{f}_{n-1}, \widetilde{f}_{n-2}$. These matrices are neither independent nor identically distributed.

In the linear case of Janvresse et al. [40] we have that

$$\det \begin{pmatrix} 0 & \pm 1 \\ 1 & 1 \end{pmatrix} = 1,$$

and so the group formed from this matrix pair is a subgroup of $\mathrm{SL}(2, \mathbb{Z})$. (The matrix pair A, B' behaves similarly.) In the non-linear case, we also have the third matrix B, which has determinant -1. Therefore the matrix group formed by these three is a subgroup of $\mathrm{S}^*\mathrm{L}(2,\mathbb{Z})$. The matrix $\hat{B}$ has the properties $\hat{B}^3 = -I, \hat{B}^6 = I$ and so the group formed by A and $\hat{B}$ will be quite different from that formed by A and B. Also, in the non-linear case, the matrix group has three generators, $A, \hat{B}$ and B, and so will likely differ radically.

2.2 Growth Types of Linear Recurrences

We have seen three different types of growth, namely, bounded, linear and exponential. Our aim is to deduce the type of growth of a periodic coefficient sequence by analyzing the product matrix P_n. Before continuing, we need to take a closer look at these types of growth. We start by looking at the growth of a *general second order linear recurrence relation* with initial values a_1, a_2, and

$$a_n = ua_{n-1} + va_{n-2},$$

for $n \geq 3$, where $a_1, a_2, u, v \in \mathbb{Z}$. It is easy to extend these values to a larger ring.

The following theorem from Bajaj [5], with clever proof, will help us to define exponential growth.

Theorem 2.2. *For a positive sequence $\{a_n\}$, if the limits*

$$\lim_{n\to\infty} a_n^{\frac{1}{n}} \qquad \text{and} \qquad \lim_{n\to\infty} \frac{a_{n+1}}{a_n}$$

both exist, then they are equal.

Proof: We can use a simple proof by contradiction here. Suppose that the former limit in the statement of the theorem is L and the latter is M, where $L < M$. Now choose a number k such that $L < k < M$, and define the series $\sum_{n=0}^{\infty} b_n$, where $b_n = \frac{a_n}{k^n}$. We then have that

$$\lim_{n\to\infty} b_n^{\frac{1}{n}} = \lim_{n\to\infty} \left(\frac{a_n}{k^n}\right)^{\frac{1}{n}} = \frac{L}{k} < 1.$$

By the root test, the series $\sum_{n=0}^{\infty} b_n$ converges. We also have that

$$\lim_{n\to\infty} \frac{b_{n+1}}{b_n} = \lim_{n\to\infty} \frac{a_{n+1}/k^{n+1}}{a_n/k^n} = \frac{M}{k} > 1.$$

By the ratio test, the series $\sum_{n=0}^{\infty} b_n$ diverges, which is a contradiction. We can argue similarly if $L > M$. $\diamond$

Definition 2.4. We say a sequence $\{a_n\}$ defined by second order linear recurrence has *exponential growth* if

$$\lim_{n\to\infty} |a_n|^{\frac{1}{n}} = c,$$

with $c > 1$. We call c the *growth rate* of the sequence. We can similarly define the growth rate as

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = c,$$

when this limit exists.

This is related to the exponential growth formula $x_n = x_0 e^{kn}$, where k is called the *growth constant*. For $k > 0$ we have exponential growth, $k < 0$ implies exponential decay and $k = 0$ gives a constant sequence. Taking either of the above limits for x_n gives growth rate $c = e^k$. Note that k is analogous to the upper Lyapunov exponent defined in Chapter 1 for matrix products. For example, the Fibonacci sequence F_n has growth rate

$$c = \lim_{n\to\infty} {F_n}^{\frac{1}{n}} = \phi, \tag{2.8}$$

where ϕ is the golden ratio, $\frac{1+\sqrt{5}}{2}$. This is consistent with Binet's formula, given in Equation (1.3).

Definition 2.5. We say a sequence $\{a_n\}$ defined by second order linear recurrence has *linear growth* if the terms can be written as

$$a_n = pn + q,$$

for some $p, q \in \mathbb{Z}$, $p \neq 0$.

Definition 2.6. We say a sequence $\{a_n\}$ defined by second order linear recurrence has *bounded growth* if for all n we have that $|a_n| \leq q$ for some $q \in \mathbb{Z}$.

Now, how do we determine the growth type of a general second order linear recurrence sequence without writing out the terms? We can start by writing our recurrence as a matrix equation, as we did for the random Fibonacci sequence. We want expressions for a_{n-1} and a_n in terms of a_{n-2} and a_{n-1}, which can be done as follows:

$$\begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ v & u \end{pmatrix} \begin{pmatrix} a_{n-2} \\ a_{n-1} \end{pmatrix},$$

so that we obtain $a_n = ua_{n-1} + va_{n-2}$ and $a_{n-1} = a_{n-1}$. The following definition can be found in Vince [70].

Definition 2.7. We call the square matrix $\left(\begin{smallmatrix} 0 & 1 \\ v & u \end{smallmatrix}\right)$ the *companion matrix* for the second order linear recurrence $a_n = ua_{n-1} + va_{n-2}$.

Definition 2.8. The *characteristic equation* in x of a square matrix M is the monic polynomial equation obtained by expanding the expression

$$\det(M - Ix) = 0.$$

If M is a 2×2 matrix, the associated characteristic equation has degree 2. The characteristic equation for the general companion matrix is given by

$$\det \begin{pmatrix} -x & 1 \\ v & u - x \end{pmatrix} = x^2 - ux - v = 0.$$

We could have obtained the characteristic equation directly from the recurrence, by writing

$$a_n - ua_{n-1} - va_{n-2} = 0,$$

and replacing successive sequence terms by powers of x. We can therefore equivalently define the companion matrix of a second order monic polynomial.

Given a second order linear recurrence, $a_n = ua_{n-1} + va_{n-2}$, we denote by λ_1, λ_2 the eigenvalues of the companion matrix, or equivalently the roots of the characteristic equation, called *characteristic roots*. We will assume that $|\lambda_1| \geq |\lambda_2|$. The following definition is taken from Larson, Edwards and Falvo [50, p. 550].

Definition 2.9. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of an $n \times n$ square matrix. We call λ_1 the *dominant eigenvalue* if

$$|\lambda_1| > |\lambda_i|,$$

for all $2 \leq i \leq n$.

It is a well known fact (for example, see Vajda [69, p. 18]) that the closed form expression for the term a_n, with $a_n \geq 1$ and $\lambda_1 \neq \lambda_2$, is given as

$$a_n = \alpha\lambda_1^n + \beta\lambda_2^n. \tag{2.9}$$

We can easily find α and β by solving the pair of equations corresponding to a_1 and a_2 to obtain

$$\alpha = \frac{a_2 - a_1\lambda_2}{\lambda_1(\lambda_1 - \lambda_2)}, \qquad \beta = \frac{a_1\lambda_1 - a_2}{\lambda_2(\lambda_1 - \lambda_2)}.$$

Example 2.5. Consider the case of the Fibonacci numbers F_n, given by the recurrence $F_n = F_{n-1} + F_{n-2}$, where $F_1 = F_2 = 1$. The characteristic equation is therefore $x^2 - x - 1 = 0$, which has roots $\phi, \phi' = \frac{1\pm\sqrt{5}}{2}$ respectively. Substituting the roots and values of α, β into Equation (2.9) gives the well known Binet formula for the Fibonacci numbers,

$$F_n = \frac{1}{\sqrt{5}}(\phi^n - (\phi')^n),$$

as given in Equation (1.3).

In the case of a second order linear recurrence equation with characteristic roots $\lambda_1 = \lambda_2$, the closed form expression for a_n, with $a_n \geq 1$, is given as

$$a_n = (\alpha + \beta n)\lambda_1^n, \tag{2.10}$$

where

$$\alpha = \frac{2a_1\lambda_1 - a_2}{\lambda_1^2}, \qquad \beta = \frac{a_2 - a_1\lambda_1}{\lambda_1^2}$$

can again be easily derived (Vajda [69]). We can use these closed form expressions to verify that there are no other types of growth than the three we have already identified. For example, quadratic growth is not possible.

Theorem 2.3. *A sequence $\{a_n\}$ defined by a second order linear recurrence relation has growth which is bounded, linear or exponential.*

Proof: We can show this by looking at the closed forms for terms in the sequence $\{a_n\}$, as given in Equations (2.9) and (2.10). For $|\lambda_1| > 1$, the term $|\lambda_1|^n$ is responsible for exponentially growing sequences, and for $|\lambda_1| = 1$, the term $\alpha + \beta n$ is responsible for linearly growing sequences. Certain combinations of eigenvalues and coefficients will produce bounded growth, but no other growth types are possible based on the closed forms. In order for non-linear polynomial growth to occur, we must have a term n^k for a fixed $k \geq 2$, which is not the case. $\diamond$

Theorem 2.4 (Kronecker). *Let α be an algebraic integer. If all of the conjugates of α in $\mathbb{C}$ have absolute value 1, then α is a root of unity.*

A proof of this theorem can be found in Greiter [34]. We are dealing with the characteristic equation $x^2 - ux - v = 0$, and so the eigenvalues are algebraic integers. From 2.4 we now know that if $|\lambda_1| = |\lambda_2| = 1$, we are dealing with roots of unity.

We can distinguish among the three growth types of a second order linear recurrence relation by looking at the corresponding eigenvalues, as the following three theorems demonstrate.

2.2.1 Linear Growth

Theorem 2.5. *Given a companion matrix M, the growth rate of the corresponding second order linear recurrence sequence $\{a_n\}$ is linear if and only if the eigenvalues λ_1 and λ_2 of M are equal roots of unity, i.e., they are both 1 or both -1, with the following exceptions. If $\lambda_1 = 1$ with $a_1 = a_2$, $\lambda_1 = -1$ with $a_1 = -a_2$, or $a_1 = a_2 = 0$, the growth is bounded. Furthermore, the growth rate of the absolute value of our sequence $\{a_n\}$ is given by*

$$\lim_{n\to\infty} |a_n|^{\frac{1}{n}} = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = 1.$$

Proof: Suppose we have a sequence $a_n = ua_{n-1} + va_{n-2}$ with $u, v \in \mathbb{Z}$ which grows linearly. By Definition 2.5 we must have $a_n = pn + q$, for $p \neq 0$. We can start by showing that the eigenvalues of M must be equal. If our eigenvalues are not equal, Equation (2.9) tells us that our terms must take the form

$$a_n = \alpha\lambda_1^n + \beta\lambda_2^n,$$

which cannot be equal to $pn + q$. We can see this more clearly by noting that in the above equation, a_n is the sum of two terms each of which is exponentially growing if

$\lambda > 1$ and bounded if $\lambda \leq 1$ (for non-zero coefficients). We must therefore have

$$a_n = (\alpha + \beta n)\lambda_1^n,$$

where λ_1 is a double root. Comparing this $|a_n|$ to $|pn+q|$ tells us that we must have $\beta \neq 0$ and $|\lambda_1| = 1$. By Kronecker's Theorem (2.4), λ_1 must then be a root of unity, and because roots of unity come in conjugate pairs and the coefficients of our recurrence are real, equal roots of unity imply $\lambda_1 = 1$ or -1.

Conversely, suppose we have a double eigenvalue λ_1 which is a root of unity. The solution (in absolute value) to a recurrence of this type is given by

$$|a_n| = |\alpha + \beta n||\lambda_1^n| = |\alpha + \beta n|,$$

which implies our sequence grows linearly in absolute value for $\beta \neq 0$. Note that if $\beta = 0$, we obtain $|a_n| = |\alpha|$, i.e., our sequence is constant in absolute value and hence bounded. We know that $\beta = (a_2 - a_1\lambda_1)/a_1$. If $\lambda_1 = 1$, we have $\beta = 0$ if and only if $a_1 = a_2$ and if $\lambda_1 = -1$ we have $\beta = 0$ if and only if $a_1 = -a_2$. These two cases are the exceptions to linear growth, as well as the trivial case, generated by $a_1 = a_2 = 0$.

Consider the limit $\lim_{n\to\infty} |a_n|^{\frac{1}{n}} = |\alpha + \beta n|^{\frac{1}{n}} = L$. Taking the logarithm of both sides gives

$$\log L = \log \lim_{n\to\infty} |\alpha + \beta n|^{\frac{1}{n}} = \lim_{n\to\infty} \frac{1}{n} \log |\alpha + \beta n| = 0,$$

which implies $L = 1$. The limit of the ratio of terms also gives

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \left| \frac{\alpha + \beta(n+1)}{\alpha + \beta n} \right| = 1,$$

confirming what we know from Definition 2.4. Since we are using the exponential definitions of growth rate here, the fact that the growth rate equals 1, tells us that the growth is not exponential. $\diamond$

2.2.2 Bounded Growth

Theorem 2.6. *Given a companion matrix M, the growth rate of the corresponding second order linear recurrence sequence $\{a_n\}$ is bounded if and only if $|\lambda_1|, |\lambda_2| \leq 1$, where the eigenvalues λ_1, λ_2 cannot be equal roots of unity, or we have one of the* *following combinations of eigenvalues and initial values:* $a_1 = a_2$ *with* λ_1 *or* $\lambda_2 = 1$, $a_1 = -a_2$ *with* λ_1 *or* $\lambda_2 = -1$ *in either the single or double root case,* $a_2 = 0$ *with* λ_1 *or* $\lambda_2 = 0$*, or lastly,* $a_1 = a_2 = 0$.

Proof: Let us first suppose our eigenvalues are equal. Our terms must be of the form

$$|a_n| = |(\alpha + \beta n)\lambda_1^n|.$$

The terms are bounded as $n \to \infty$ if and only if one of two things happens. First we can have $|\lambda_1| < 1$, in which case $|a_n| \to 0$. Second we can have $|\lambda_1| = 1$ with $\beta = 0$. This corresponds to the $\lambda_1 = \pm 1$ exceptions listed in the theorem, as shown in Theorem 2.5.

If we suppose the eigenvalues are not equal, our terms must be of the form

$$|a_n| = |\alpha\lambda_1^n + \beta\lambda_2^n|,$$

The terms are bounded if and only if one of three things happens. First we can have $|\lambda_1|, |\lambda_2| \leq 1$, in which case

$$|a_n| = |\alpha\lambda_1^n + \beta\lambda_2^n| \leq |\alpha||\lambda_1|^n + |\beta||\lambda_2|^n \leq |\alpha| + |\beta|.$$

Therefore all terms in the sequence are bounded above by $|\alpha|+|\beta|$. The other options are $\alpha = 0$ and $|\lambda_2| = 1$, or $\beta = 0$ and $|\lambda_1| = 1$. We have that $\alpha = \frac{a_2 - a_1\lambda_2}{\lambda_1(\lambda_1 - \lambda_2)}$ and $\beta = \frac{a_1\lambda_1 - a_2}{\lambda_2(\lambda_1 - \lambda_2)}$ and therefore these two cases occur precisely for those initial values stated in the theorem. If $a_1 = a_2$ and $\lambda_1 = 1$, we obtain the constant sequence $a_n = \alpha$ and similarly if $\lambda_2 = 1$ we obtain $a_n = \beta$. If $a_1 = -a_2$ and $\lambda_1 = -1$ or $\lambda_2 = -1$, we obtain one of the bounded alternating sequences $a_n = \alpha(-1)^n$ or $a_n = \beta(-1)^n$, implying $|a_n| = |\alpha|$ or $|a_n| = |\beta|$ in absolute value.

For either the equal or non-equal eigenvalue cases we have the trivial bounded case which occurs when $a_1 = a_2 = 0$ and implies $\alpha = \beta = 0$. Similarly, we obtain the all-zero sequence if λ_1 or $\lambda_2 = 0$ and $a_2 = 0$. This forces β or $\alpha = 0$, and hence $a_n = 0$ also. $\diamond$

In the case of bounded growth we cannot calculate the growth rate with the methods used for exponential and linear growth. For example, consider the bounded sequence $a_n = 1, 1, 0, 1, 1, 0, \ldots$. The limit $\lim_{n\to\infty} |a_n|^{\frac{1}{n}}$ does not exist. Similarly, $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$ is not defined for this sequence because it requires dividing by zero an infinite number of times. However, we just consider the growth rate of a bounded sequence to be 1.

2.2.3 Exponential Growth

Finally, we will review, with proof, the following well-known fact about exponential growth.

Theorem 2.7. *The growth of a second order linear recurrence sequence is exponential if and only if $|\lambda_1| > 1$, where λ_1 and λ_2 are the eigenvalues of the corresponding companion matrix, with the exceptions $a_1 = a_2$ with $\lambda_2 = 1$, $a_1 = -a_2$ with $\lambda_2 = -1$, as well as $a_2 = 0$ with $\lambda_2 = 0$, and lastly $a_1 = a_2 = 0$. Furthermore, the growth rate can be written as*

$$\lim_{n\to\infty} |a_n|^{\frac{1}{n}} = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = |\lambda_1|,$$

except in the case of distinct eigenvalues with λ_1 not dominant, where we have

$$\left(\lim_{n\to\infty} |a_{2n}|^{\frac{1}{n}}\right)^{\frac{1}{2}} = \left(\lim_{n\to\infty} \left|\frac{a_{n+2}}{a_n}\right|\right)^{\frac{1}{2}} = |\lambda_1|,$$

or similarly for odd-indexed terms.

Proof: Suppose, w.l.o.g., that $|\lambda_1| > 1$. We must show this implies exponential growth. From Theorems 2.3, 2.5 and 2.6 we can conclude directly that growth isn't linear or bounded and therefore must be exponential. However, we are interested in finding the growth rate of such exponentially growing sequences, and in doing so we will prove that growth is in fact exponential.

We can break our exponential growth into several cases, depending on whether or not the eigenvalues are equal, and whether or not one eigenvalue is dominant. First, we consider the case of λ_1 dominant, which implies unequal eigenvalues. Second, we consider eigenvalues which are distinct, but equal in absolute value (hence there is no dominant eigenvalue) and third, we consider the case of eigenvalues which are equal.

In the first case, since λ_1 is dominant, we have

$$\lim_{n\to\infty} \left(\frac{\lambda_2}{\lambda_1}\right)^n = 0. \tag{2.11}$$

The first definition of growth rate gives us

$$\lim_{n\to\infty} |a_n|^{\frac{1}{n}} = \lim_{n\to\infty} |\alpha\lambda_1^n + \beta\lambda_2^n|^{\frac{1}{n}} = \lim_{n\to\infty} \left|\lambda_1^n \left(\alpha + \beta\left(\frac{\lambda_2}{\lambda_1}\right)^n\right)\right|^{\frac{1}{n}} \tag{2.12}$$

$$= |\lambda_1| \lim_{n\to\infty} \left|\alpha + \beta\left(\frac{\lambda_2}{\lambda_1}\right)^n\right|^{\frac{1}{n}}. \tag{2.13}$$

By Equation (2.11) we have

$$\lim_{n\to\infty} |a_n|^{\frac{1}{n}} = |\lambda_1| \lim_{n\to\infty} |\alpha|^{\frac{1}{n}} = |\lambda_1|,$$

provided $\alpha \neq 0$. (If $\alpha = 0$, our growth rate will be $|\lambda_2|$, which will give exponential growth if its value is greater than 1, otherwise it is an exceptional case.) By Definition 2.4, we can equivalently write

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = |\lambda_1|,$$

which is easy to verify using Equation (2.9).

Now consider the case of distinct eigenvalues, where λ_1 is not dominant. The first limit definition similarly gives

$$\lim_{n\to\infty} |a_n|^{\frac{1}{n}} = |\lambda_1| \lim_{n\to\infty} \left|\alpha + \beta\left(\frac{\lambda_2}{\lambda_1}\right)^n\right|^{\frac{1}{n}}.$$

Here, the ratio of eigenvalues has absolute value 1 and so the term in absolute values is finite for all n. This means that the terms in the limit in the right-hand side will approach 1, even if $\left(\frac{\lambda_2}{\lambda_1}\right)^n$ is alternating. We therefore have $\lim_{n\to\infty} |a_n|^{\frac{1}{n}} = |\lambda_1|$, provided $\{a_n\}$ does not contain infinitely many zeroes. In that case, the limit of n^{th} roots would not converge. Recall that our recurrence $a_n = ua_{n-1} + va_{n-2}$ has $u, v, a_i \in \mathbb{Z}$ and characteristic equation $x^2 - ux - v = 0$. Our eigenvalues, which are a conjugate pair, are equal in absolute value. We may therefore have $\lambda_1 = -\lambda_2$, which implies $\lambda_1, \lambda_2 = \pm c$ or $\pm ic$, where $c \in \mathbb{R}$. This gives us

$$a_n = \lambda_1^n \left(\alpha + \beta\left(\frac{\lambda_2}{\lambda_1}\right)^n\right) = \lambda_1^n(\alpha + \beta(-1)^n).$$

Similarly, it may be the case that the ratio of eigenvalues is $\pm i$. (Recall this ratio must be a root of unity.) In this case, our term a_n is written as

$$a_n = \lambda_1^n \left(\alpha + \beta\left(\frac{\lambda_2}{\lambda_1}\right)^n\right) = \lambda_1^n(\alpha + \beta(\pm i)^n).$$

If $a_1 = 0$ or $a_2 = 0$ it is easy to see that α and β, which must be real, have the same absolute value, and so in the former case (ratio -1) either $a_{2n-1} = 0$ or $a_{2n} = 0$, respectively. In the case of ratio $\pm i$, the term $\left(\frac{\lambda_2}{\lambda_1}\right)^n$ cycles through 4 values, one of which is -1. In either case we have an all-zero subsequence, and so the growth rate of a subsequence $\{a_{2n}\}$ not containing infinitely many zeroes can be found by

$$\lim_{n\to\infty} |a_{2n}|^{\frac{1}{n}} = |\lambda_1|^2,$$

so that the growth of $\{a_n\}$ is

$$\left(\lim_{n\to\infty} |a_{2n}|^{\frac{1}{n}}\right)^{\frac{1}{2}} = |\lambda_1|,$$

or similarly with a_{2n-1}. Note that our characteristic equation can be written as

$$x^2 - (\lambda_1 + \lambda_2)x - \lambda_1\lambda_2 = x^2 - \lambda_1\lambda_2,$$

which gives us the recurrence $a_n = \lambda_1\lambda_2 a_{n-2}$ when $\lambda_1 = -\lambda_2$. This is again proof for two distinct subsequences, corresponding to initial values a_1 and a_2. Therefore the second limit given in Definition 2.4, the limit of the ratio, will not converge. We can again consider the growth rate of $\{a_{2n}\}$ and $\{a_{2n-1}\}$ instead, which gives

$$\lim_{n\to\infty} \left|\frac{a_{n+2}}{a_n}\right| = \lim_{n\to\infty} |\lambda_1^2| \left|\frac{\alpha + \beta\left(\frac{\lambda_2}{\lambda_1}\right)^{n+2}}{\alpha + \beta\left(\frac{\lambda_2}{\lambda_1}\right)^{n}}\right| = |\lambda_1|^2,$$

because the numerator and denominator in the fraction are equal. Therefore, the growth rate of $\{a_n\}$ is

$$\left(\lim_{n\to\infty} \left|\frac{a_{n+2}}{a_n}\right|\right)^{\frac{1}{2}} = |\lambda_1|,$$

The growth rate of $\{a_{2n-1}\}$ is the same since it is based on the same recurrence.

In the third case, that of equal eigenvalues, we have for the first definition

$$\lim_{n\to\infty} |a_n|^{\frac{1}{n}} = \lim_{n\to\infty} |(\alpha + \beta n)\lambda_1^n|^{\frac{1}{n}} = \lim_{n\to\infty} |\lambda_1||\alpha + \beta n|^{\frac{1}{n}} = |\lambda_1|,$$

using the fact that $|\alpha + \beta n|^{\frac{1}{n}} \to 1$. We can similarly use the second limit in Definition 2.4.

We have seen that in all three cases, both definitions of exponential growth agree. The exceptional cases are treated in Theorem 2.6 where it is shown that the growth is bounded.

Conversely, suppose the growth of our recurrence sequence $\{a_n\}$ is exponential, which we know is possible from Theorem 2.3. It is therefore not linear or bounded. By Theorems 2.5 and 2.6, we conclude that we must have either $|\lambda_1| > 1$ or $|\lambda_2| > 1$. Since we are assuming $|\lambda_1| > |\lambda_2|$, we can conclude that $|\lambda_1| > 1$.

$\diamond$

As an example, consider the Fibonacci sequence. Theorem 2.7 tells us that the growth rate of this sequence is given by the largest eigenvalue in absolute value, i.e., $\phi = \frac{1+\sqrt{5}}{2}$.

Note that we can use the Skolem-Mahler-Lech Theorem (Everest et al. [25, p. 25, 31]) in certain cases to show that a sequence is unbounded. One form of the theorem states that for any *non-degenerate* sequence $\{a_n\}$ (i.e., for each pair of distinct eigenvalues, the ratio is not a root of unity, [25, p. 5]) whose characteristic roots are not all roots of unity,

$$\lim_{n\to\infty} |a_n| \to \infty.$$

In the second case in Theorem 2.7, our eigenvalues are distinct but one is not dominant. This leads to a degenerate sequence because the eigenvalues are equal in absolute value. Also, we mentioned earlier the difference between exponential growth and exponential decay. In our case, that of a second order linear recurrence with integral coefficients and initial values, it is not possible for the resulting sequence to approach zero (or any other number) asymptotically and so we cannot have exponential decay.

2.3 Subsequences and the Growth of Periodic Coefficient Sequences

Now that we have considered the details of the growth of a general second order linear recurrence relation, we can consider the case of our periodic coefficient sequence, $t_i = s_{i-2}t_{i-1} + t_{i-2}$. This differs from the former case because the coefficients are not fixed. The key is to look at subsequences of the periodic coefficient sequences. In particular, for sequences generated by a coefficient cycle of length n, we will consider

n subsequences. We will show that these subsequences are second order linear recurrence equations with fixed coefficients, so we can analyze their behaviour. Consider the following motivating example.

Example 2.6. The sequence generated by the coefficient cycle $\sigma_4 = (++--)$ is as follows, where we have listed terms vertically to create four subsequences.

$$\begin{array}{rrrrrl} 1, & -1, & -4, & -11, & -29, & \dots \\ 1, & 4, & 11, & 29, & 76, & \dots \\ 2, & 3, & 7, & 18, & 47, & \dots \\ 3, & 7, & 18, & 47, & 123, & \dots \end{array}$$

It appears as if each subsequence is growing exponentially. Upon inspection we can see that each subsequence is growing according to the recurrence relation

$$t_n = 3t_{n-1} - t_{n-2}. \tag{2.14}$$

The product matrix in this example is $P_4 = AABB = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 3 \end{smallmatrix}\right)$, and is in the form of a companion matrix. (Note this is not always the case.) Reading off the characteristic equation gives $x^2 - 3x + 1 = 0$, which is also what we get from Equation (2.14). Suppose we want to find sequence terms t_i using the product matrix. Recall from Equation (2.3) that we must calculate the reverse product matrix, $M_4M_3M_2M_1 = BBAA = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 3 \end{smallmatrix}\right)$, in order to compute sequence terms. (Note again that reversing the terms results in the transpose matrix.) Starting with $t_1 = t_2 = 1$, we obtain

$$\begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} t_5 \\ t_6 \end{pmatrix},$$
$$\begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} -4 \\ 11 \end{pmatrix} = \begin{pmatrix} t_9 \\ t_{10} \end{pmatrix}.$$

Continuing to multiply each new vector by $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 3 \end{smallmatrix}\right)$, we see that the first two subsequences are generated. How then do we generate the next two subsequences?

If we want to skip ahead one term in the periodic coefficient sequence, we can use $t_2 = 1, t_3 = 2$ as our initial vector and rotate our coefficient cycle one term to the left to obtain $\sigma_4 = (+--+)$. The corresponding matrix product is $P_4 = M_2M_3M_4M_1 =$

$ABBA$, which we must first reverse, in order to calculate sequence terms. Here we have $M_1M_4M_3M_2 = ABBA = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, and the terms generated are

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix} = \begin{pmatrix} t_6 \\ t_7 \end{pmatrix},$$
$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 4 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 7 \end{pmatrix} = \begin{pmatrix} t_{10} \\ t_{11} \end{pmatrix}.$$

We have generated the second and third subsequences. Similarly, using initial terms $t_3 = 2, t_4 = 3$, and coefficient cycle $\sigma_4 = (--++)$ with corresponding product matrix $P_4 = M_3M_4M_1M_2 = BBAA$ and reversal $M_2M_1M_4M_3 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 3 \end{smallmatrix}\right)$ we obtain

$$\begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \end{pmatrix} = \begin{pmatrix} t_7 \\ t_8 \end{pmatrix},$$
$$\begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 7 \\ 18 \end{pmatrix} = \begin{pmatrix} t_{11} \\ t_{12} \end{pmatrix}.$$

Here we have the terms in the third and fourth subsequences.

In this example, we have seen that rotating the coefficient cycle produces n subsequences, each of which grows according to the same linear recurrence. We will prove this result in general, but first a few notes will be useful. If we let $P_n = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, then

$$P_n - Ix = \begin{pmatrix} a - x & b \\ c & d - x \end{pmatrix},$$

which has characteristic equation $(a-x)(d-x)-bc = 0$, i.e., $x^2-(a+d)x+(ad-bc) = 0$. Comparing with the form of the characteristic equation $x^2 - ux - v = 0$, we see that

$$u = a + d = \mathrm{tr}(P_n), \tag{2.15}$$
$$v = -(ad - bc) = -\det(P_n), \tag{2.16}$$

where $u, v \in \mathbb{Z}$ because we have seen the matrices P_n belong to $\mathrm{S^*L}(2, \mathbb{Z})$. Note also that this tells us that $v = \pm 1$, which we have seen.

The following theorem considers what happens when we reverse the terms in a coefficient cycle. This is what allows us to reverse the matrices in the product P_n.

Theorem 2.8. *Given a coefficient cycle $\sigma_n = (s_1, \ldots, s_n)$ with $n \geq 1$ and corresponding product matrix P_n, reversing the terms in the cycle gives $\sigma'_n = (s_n, \ldots, s_1)$ with corresponding product matrix P'_n, where P'_n is the transpose of P_n and has the same characteristic equation as P_n.*

Proof: We will first prove the fact that if the cycle is reversed, the associated product matrix is transposed. Let $P_n = M_1 \cdots M_n$, so that reversing the matrices gives $P'_n = M_n \cdots M_1$. Taking the transpose gives

$$\begin{aligned} P_n^T &= (M_1 \cdots M_n)^T = M_n^T \cdots M_1^T \\ &= M_n \cdots M_1 = P'_n, \end{aligned}$$

because $M_i \in \{A, B\}$ for $1 \leq i \leq n$ and the matrices A and B are both self-transpose.

Now we must show that taking the transpose does not affect the characteristic equation. The characteristic equation of P_n is dependent only on $u = \operatorname{tr}(P_n)$ and $v = -\det(P_n)$. Since the trace is $a + d$ and the determinant is $ad - bc$ for both P_n and P'_n, we have the same equation in both cases. ◇

It is easy to see that Theorem 2.8 is not true for matrix products in general. Similarly, we can check the outcome of the rotation of terms in a product matrix.

Theorem 2.9. *Given a coefficient cycle $\sigma_n = (s_1, \ldots, s_n)$, with $n \geq 1$, rotating the entries (to the right) by k, where $0 \leq k \leq n-1$, gives a new cycle $\sigma'_n = (s_{n-k+1}, \ldots, s_n, s_1, \ldots, s_{n-k})$ with corresponding product matrix having the same characteristic equation as the original product matrix.*

Proof: Suppose σ_n has associated product matrix P_n, which we can write as $P_n = M_1 \cdots M_n$. If we rotate the coefficient cycle by one we get $\sigma'_n = (s_n, s_1, \ldots, s_{n-1})$, with associated product matrix $P'_n = M_n(M_1 \cdots M_{n-1})$. If we suppose $M_1 \cdots M_{n-1} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ (which simply equals I for the $n = 1$ case), we have

$$\begin{aligned} P_n &= (M_1 \cdots M_{n-1})M_n = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} = \begin{pmatrix} b & a \pm b \\ d & c \pm d \end{pmatrix}, \\ P'_n &= M_n(M_1 \cdots M_{n-1}) = \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a \pm c & b \pm d \end{pmatrix}. \end{aligned}$$

The traces of P_n and P_n' are both equal to $b + c \pm d$ and the determinant in both cases is $bc - ad$. We conclude that the characteristic equation is the same for both matrices. We can continue rotating by any number k of places in our product P_n in the same way without changing the characteristic equation. We also have the trivial rotation, where we rotate by 0, or equivalently n terms, again leaving the characteristic equation unchanged. $\diamond$

Unlike Theorem 2.8 for reversal, this theorem for rotation is true for a general 2×2 matrix product. Notice that here we rotated terms in our coefficient cycle to the right, whereas in Example 2.6 we rotated terms to the left. The set of all rotations is equivalent in either case, so direction did not matter for purposes of the proof. The following consequence of Theorem 2.9 will be useful.

Corollary 2.1. *Given a coefficient cycle $\sigma_n = (s_1, \ldots, s_n)$, with $n \geq 1$, the corresponding periodic coefficient sequence can be broken down into n subsequences, each of which grows according to the same second order linear recurrence relation.*

Proof: By rotating a coefficient cycle by k terms, we obtain a product matrix that allows us to compute the $(k+1)$st and $(k+2)$nd subsequences. We can generate all subsequences by rotating by k for $k = 1, \ldots, n-2$. Theorem 2.9 tells us that each product matrix under rotation has the same characteristic equation, hence each subsequence grows according to the same recurrence relation. $\diamond$

We might be tempted to conclude here that each subsequence corresponding to a periodic coefficient sequence must have the same growth type; but this is not always the case, as we have seen that the growth depends not only on the characteristic equation but also on the initial conditions. A few more examples will be helpful in understanding the growth of our subsequences.

Example 2.7. The sequence generated by the coefficient cycle $\sigma_6 = (+++---)$ (see Example 2.3) can be broken down into the following six subsequences:

In absolute value, it appears that four of the six subsequences are growing linearly with difference 8, with the exception of the first two terms of the first subsequence, and the other two subsequences are bounded with absolute value a constant 2. This

$$\begin{array}{rrrr}
1, & 7, & -15, & 23, \ldots \\
1, & -9, & 17, & -25, \ldots \\
2, & -2, & 2, & -2, \ldots \\
3, & -11, & 19, & 27, \ldots \\
5, & -13, & 21, & -29, \ldots \\
-2, & 2, & -2, & 2, \ldots
\end{array}$$

exception comes from the fact that the initial term in the sequence is repeated, and because no other consecutive terms (in absolute value) are repeated, it is not possible for the differences of the first two terms in the first two subsequences to be equal. We could try to deduce the governing recurrence, but we can easily find it from the product matrix P_6. We have $P_6 = A^3B^3 = \left(\begin{smallmatrix} 3 & -4 \\ 4 & -5 \end{smallmatrix}\right)$, which has characteristic equation $x^2 + 2x + 1 = 0$. This implies the recurrence $t_n = -2t_{n-1} - t_{n-2}$, which fits our data, even the exception. (Note that if we rotate our coefficient cycle by $k = 1$, we obtain the matrix $P_6 = BA^3B^2 = \left(\begin{smallmatrix} -1 & 4 \\ 0 & -1 \end{smallmatrix}\right)$, which has the same recurrence as above, as required by Theorem 2.9). Solving, we get a double eigenvalue $\lambda_1 = -1$, which by Theorem 2.5 implies linear growth, but we must also consider the exception. In the third and sixth subsequences we have $\lambda_1 = -1$ and $a_1 = -a_2$, which again by Theorem 2.5 implies bounded growth.

Note that since linear growth requires a double eigenvalue of ± 1, there are only two options for the recurrence, namely

$$t_n = \pm 2t_{n-1} - t_{n-2}.$$

It is clear that if $a_1 = \pm a_2$ (depending on the value of λ_1), the sequence will remain constant in absolute value.

The next example corresponds to Example 2.4 and illustrates subsequences which are all bounded.

Example 2.8. The periodic sequence generated by the cycle $\sigma_3 = (++-)$ of length 3 is broken down into three subsequences as follows:
Each of these sequences is clearly bounded. The product matrix in this case is $P_3 = A^2B = \left(\begin{smallmatrix} 1 & 0 \\ 2 & -1 \end{smallmatrix}\right)$, which has characteristic equation $x^2 - 1 = 0$. The corresponding recurrence is $t_n = t_{n-2}$, which is satisfied by all three subsequences. The eigenvalues

$$\begin{array}{rrrrrr}
1, & 3, & 1, & 3, & 1, & 3, \ldots \\
1, & -1, & 1, & -1, & 1, & -1, \ldots \\
2, & 2, & 2, & 2, & 2, & 2, \ldots
\end{array}$$

of the characteristic equation are $\lambda_1 = 1, \lambda_2 = -1$ and Theorem 2.6 tell us distinct roots of unity imply bounded growth.

We can extend this example by considering what happens when we double the coefficient cycle.

Example 2.9. Let $\sigma_6 = (++-++-)$. We know it must behave the same way as $\sigma_3 = (++-)$ because it generates the same periodic coefficient sequence. This time we obtain six subsequences, each of which is constant, as shown below.

$$\begin{array}{rrr}
1, & 1, & 1, \ldots \\
1, & 1, & 1, \ldots \\
2, & 2, & 2, \ldots \\
3, & 3, & 3, \ldots \\
-1, & -1, & -1, \ldots \\
2, & 2, & 2, \ldots
\end{array}$$

The product matrix is

$$P_6 = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}^2 = I,$$

with characteristic equation $x^2 - 2x + 1 = 0$, recurrence $t_n = 2t_{n-1} - t_{n-2}$ and double eigenvalue $\lambda_1 = 1$. The double eigenvalue would suggest linear growth, but again this falls into the exceptional case because each of our subsequences has $a_1 = a_2 = 1$, ensuring bounded growth. The recurrence was discussed earlier.

There are a couple of interesting things to note about this example. First, it makes sense that the first two subsequences are constant. If we reverse the matrix product P_n to calculate sequence terms, we obtain $BA^2BA^2 = I$, and so

$$\begin{pmatrix} t_7 \\ t_8 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

where the matrix I ensures subsequence values are constant. In order to calculate the next four subsequences using matrices, we must take rotations of the cycle $\sigma_6 = (++-++-)$. Doing this gives $(+-++-+), (-++-++)$, which when reversed have product matrices $ABAABA = AABAAB = I$. It is true in general that if we rotate the matrices in $P_n = I$ to obtain $P_1^{-1}P_nP_1$, then $P_1^{-1}P_nP_1 = P_1^{-1}IP_1 = I$, i.e., this rotation is also equal to I.

Also note that in the previous example, doubling the coefficient cycle produced a set of subsequences, all of which were constant. Given a set of subsequences with bounded growth, it is possible to repeat the coefficient cycle k times, for some k, creating a new coefficient cycle of length nk, which gives us a list of constant subsequences. (Here we have assumed that if a periodic coefficient sequence has bounded growth then it is periodic with some finite period length l, and nowhere have we actually proven this fact. Clearly the converse is true and the complete proof follows shortly.) This is the same as saying that for any product matrix P_n producing bounded growth, we can write $P_n^k = I$ $P_n^k = I$ for some k. We will give a proof of the fact that the growth of a periodic coefficient sequence is bounded if and only if P_n has finite order, but first it will be useful to summarize the different types of growth for subsequences corresponding to product matrices P_n.

Recall that matrices P_n belong to the group $\mathrm{S^*L}(2, \mathbb{Z})$, and so are unimodular. It is a standard result in linear algebra that the product of eigenvalues is the determinant of a matrix. Therefore for matrices P_n, we must have that $\lambda_1\lambda_2 = \pm 1$, implying $|\lambda_1\lambda_2| = 1$. With this information, we can now update Theorems 2.5, 2.6 and 2.7 to the specific case of $M = P_n$. Previously we talked of a companion matrix M corresponding to a second order linear recurrence, but here since P_n is not necessarily a companion matrix, we will have to consider the second order linear recurrence associated with the characteristic equation obtained from P_n. Corollary 2.1 tells us that this recurrence governs the growth of all n subsequences. Note that it is now possible to have $P_n = \pm I$, whereas previously $\pm I$ was not included because it is not in the form of a companion matrix.

Theorem 2.10. *Given a product matrix P_n with eigenvalues λ_1 and λ_2, the growth of a given subsequence is*

1. *linear if and only if $P_n \neq \pm I$ and the eigenvalues λ_1, λ_2 are both 1 or both -1,*
2. *bounded if and only if $P_n = \pm I$ or λ_1, λ_2 are distinct roots of unity,*
3. *exponential if and only if $|\lambda_1| > 1$,*

with the following exceptions: if $\lambda_1 = 1$ with $a_1 = a_2$, $\lambda_1 = -1$ with $a_1 = -a_2$, or $a_1 = a_2 = 0$, the growth is bounded.

Proof: Let us start by considering the case $P_n = \pm I$. We have seen that if $P_n = I$, then all of our subsequences are constant, and so we have $a_1 = a_2$ for each. Also in this case we have $\lambda_1 = \lambda_2 = 1$. Similarly if $P_n = -I$, subsequences will alternate in sign, but be constant in absolute value and we have $\lambda_1 = \lambda_2 = -1$, with $a_1 = -a_2$. These two cases belong to the exceptions in Theorems 2.6 and 2.5. The difference when $P_n = \pm I$ is that all subsequences have this bounded behaviour, instead of just one bounded subsequence among linear subsequences. The exceptional case from Theorem 2.5 remains, but the case from Theorem 2.7 does not occur because it requires an eigenvalue to be ± 1, which does not occur for exponential growth. Similarly the exception $a_2 = 0$ with λ_1 or $\lambda_2 = 0$ is not applicable for product matrices P_n because we cannot have an eigenvalue of 0.

Theorem 2.5 told us that growth was linear if and only if the eigenvalues of P_n were equal roots of unity, and this remains unchanged by the fact that $|\lambda_1 \lambda_2| = 1$. Theorem 2.6 told us that growth was bounded if and only if $|\lambda_1|, |\lambda_2| \leq 1$ as long as the eigenvalues aren't equal roots of unity. Since the eigenvalues must be less than or equal to 1 in absolute value, and their product must have absolute value 1, we conclude that the eigenvalues must be roots of unity. Theorem 2.7 told us that growth was exponential if and only if $|\lambda_1| > 1$. This remains true, although we must now have $|\lambda_2| < 1$, so that the product is 1. Here λ_1 is dominant, and so we cannot have the cases with eigenvalues equal or equal in absolute value. $\diamond$

Note that the Skolem-Mahler-Lech Theorem (mentioned after Theorem 2.7) is only applicable to the exponential case, because this is the only growth type where the eigenvalues are not both roots of unity.

Definition 2.10. We say that a periodic coefficient sequence $\{t_i\}$ with cycle length n grows *exponentially* if all n subsequences grow exponentially, we say it grows *linearly* if at least one subsequence grows linearly and it is *bounded* if all n subsequences are bounded.

We have seen that if some subsequences grow linearly, the rest are bounded (see Example 2.7), and if at least one subsequence grows exponentially, they all do (see Example 2.6). As was the case in Proposition 1.1 for the random Fibonacci sequence, the growth of a periodic coefficient sequence is also independent of its initial values, with the exceptions made in Theorem 2.10. This is because the product matrices, eigenvalues, and hence growth rates do not depend on t_1 and t_2.

Theorem 2.11. *A product matrix P_n has finite order if and only if the associated periodic coefficient sequence has bounded growth.*

Proof: By Definition 2.10, for a periodic coefficient sequence to be bounded, we need all subsequences to be bounded. Theorem 2.10 then tells us that we either have $P_n = \pm I$ or P_n with eigenvalues which are distinct roots of unity. In the former case we clearly have finite order. The key to showing P_n has finite order is to consider our product matrices in Jordan normal form (see Fletcher [28]). Over $\mathbb{C}$, every matrix has a Jordan normal form, and is similar to its Jordan normal form. Similar matrices share trace, determinant, eigenvalues and more importantly order, where similar matrices are conjugate group elements. (Recall matrices A and B are *similar* or *conjugate* if there exists an invertible matrix M such that $A = M^{-1}BM$.) For 2×2 matrices M, there are three different cases.

The first is that we have distinct eigenvalues λ_1 and λ_2, with Jordan normal form

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

The second case is equal eigenvalues λ_1, with $M = \lambda_1 I$, which has Jordan normal form

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}.$$

In this case, since we have $\det P_n = \pm 1$ and integer matrix entries, we must have $\lambda_1 = \pm 1$. The third case is for equal eigenvalues λ_1 with $M \neq \lambda_1 I$, and here the

Jordan normal form is

$$\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}.$$

We are interested in the case of distinct roots of unity (bounded growth), so we can assume our matrix P_n is similar to, and hence has the same order as that in the first case. The order is therefore the smallest number k such that

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^k = \begin{pmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{pmatrix} = I.$$

In other words, the order is the smallest value k such that $\lambda_1^k = \lambda_2^k = 1$. This k exists because we know the eigenvalues are roots of unity.

Conversely, suppose we have a product matrix P_n with finite order, i.e., $P_n^k = I$. If we consider the different cases of the Jordan normal form, P_n must then be similar to $\left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)$, where λ_1 and λ_2 are distinct, and must be roots of unity to ensure finite order, or P_n could be $\pm I$ to begin with. In either case the corresponding periodic coefficient sequence is bounded by Theorem 2.10. Note that if P_n is similar to $\left(\begin{smallmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{smallmatrix}\right)$, the equal roots of unity case, it cannot have finite order because

$$\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}^k = \begin{pmatrix} x^k & kx^{k-1} \\ 0 & x^k \end{pmatrix} \neq I. \tag{2.17}$$

This makes sense because we know the growth in this case is linear. ◇

Corollary 2.2. *A periodic coefficient sequence has bounded growth if and only if it is periodic with some finite period length $l = nk$.*

Proof: Clearly, if a sequence is periodic, then it is bounded. By Theorem 2.11, a product matrix P_n has finite order if and only if the associated periodic coefficient sequence has bounded growth. We mentioned after Example 2.9 that raising a matrix P_n to the power k to obtain I is the same as repeating a length n coefficient cycle k times to obtain a list of nk constant subsequences. This tells us that the n subsequences must have been periodic with period k and that the entire periodic coefficient sequence has period $l = nk$. ◇

McGuire [55] gives a proof of this fact using the pigeonhole principle and also notes that $k|l$. Since k is the order of P_n, where P_n is similar to $\left(\begin{smallmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{smallmatrix}\right)$, with λ_1, λ_2 distinct roots of unity, we have seen that $\lambda_1^k = \lambda_2^k = 1$. Also, since the eigenvalues are a conjugate pair, k must be the smallest power of each of the eigenvalues which produces 1, and so our eigenvalues λ_1 and λ_2 are k^{th} roots of unity, where k is the smallest value for which this occurs.

In Example 2.8 we had $l = 6$, $n = 3$ and $k = 2$. In fact the product matrix $A^2BA^2B = I$ is the shortest product of matrices A and B producing the identity. Also, our eigenvalues were ± 1, which raised to the power $k = 2$ give 1. As a further example, consider the following.

Example 2.10. Consider the coefficient cycle $\sigma_4 = (+++-)$, with product matrix $P_4 = \left(\begin{smallmatrix}2 & -1\\ 3 & -1\end{smallmatrix}\right)$. The characteristic equation is $x^2 - x + 1 = 0$, which gives us eigenvalues $\lambda_1, \lambda_2 = \frac{1\pm\sqrt{-3}}{2}$. Note that $|\lambda_1| = |\lambda_2| = 1$, and so growth is indeed bounded by Theorem 2.10 and Definition 2.10. Using σ_4, the periodic coefficient sequence is $1, 1, 2, 3, 5, -2, 3, 1, 4, -3, 1, -2, -1, -1, -2, -3, -5, 2, -3, -1, -4, 3, -1, 2, 1, 1, \ldots$. This sequence starts to repeat after 24 terms, so we have $l = 24$. Also $n = 4$, and so we should have $k = 6$. We can verify this by raising P_4 to the sixth power obtaining $P_4^6 = I$. Similarly $\left(\frac{1\pm\sqrt{-3}}{2}\right)^6 = 1$. We can therefore write our periodic coefficient sequence as 24 constant subsequences rather than 4 subsequences which repeat after 6 terms.

In his paper *Period of a Linear Recurrence* [70], Vince investigates the period of a general n^{th} order linear recurrence. Unlike our repeating recurrences (which correspond to bounded growth), Vince considers an algebraic number field K, with ring of integers A and constructs a repeating sequence of size n-vectors, X_m, with entries in A, by reducing modulo an ideal in A. The vectors X_m come from

$$X_{m+1} = TX_m,$$

where X_0 is the initial value vector of the linear recurrence and T is the corresponding companion matrix.

In his paper *Periodic Recurrence Relations and Continued Fractions* [13], Carson considers a recurrence of the form

$$\gamma_n x_n = \alpha_n x_{n-1} + \beta_n x_{n-2}, \tag{2.18}$$

and studies the effect of allowing the coefficients $\{\alpha_n\}, \{\beta_n\}$ and $\{\gamma_n\}$ to be bounded sequences which repeat with period k. He proves the following relation for terms x_n appearing k terms apart in the sequence defined by Equation (2.18):

$$\gamma x_{(r+2)k+2} = \alpha x_{(r+1)k+s} + \beta x_{rk+s},$$

where α, β and γ are constant for all integer values of r and s. Carson further applies his results to the numerators and denominators of the convergents of a periodic continued fraction. In [26], Ferguson considers the one-parameter class of linear recurrences given by

$$x_n(t) = \alpha_n x_{n-1} + t\beta_{n-1} x_{n-2}.$$

He gives results (solutions or generating functions) for the cases where $\{a_n\}$ and $\{b_n\}$ are arbitrary sequences, period-2 sequences, and in general, period-k sequences.

Proposition 2.3. *The growth rate of an exponentially growing periodic coefficient sequence $\{t_i\}$ with cycle length n is $|\lambda_1|^{\frac{1}{n}}$, where λ_1 is the dominant eigenvalue of the corresponding product matrix P_n.*

Proof: Theorem 2.10 tells us that the dominant eigenvalue λ_1 exists and gives us the growth rate of any of the n subsequences which are formed by considering every n^{th} term in the periodic coefficient sequence. Taking the n^{th} root of this growth rate allows us to determine the growth rate of the periodic coefficient sequence term by term. $\diamond$

Example 2.11. Continuing Example 2.6 with product matrix $P_4 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 3 \end{smallmatrix}\right)$ and characteristic equation $x^2 - 3x + 1 = 0$, we get eigenvalues $\lambda_1, \lambda_2 = \frac{3\pm\sqrt{5}}{2}$. Here $|\lambda_1| > 1$ and so growth is exponential. By Proposition 2.3, the growth rate of the corresponding periodic coefficient sequence is then $\left(\frac{3+\sqrt{5}}{2}\right)^{\frac{1}{4}}$.

Chapter 3

Trace, Order and Growth Type

3.1 Values of Trace and Determinant

Without explicitly calculating eigenvalues, the characteristic equation $x^2 - ux - v = 0$ can tell us a great deal about the behaviour of our periodic coefficient sequences. In this section we will take a closer look at the values of u and v. Recall from Equations (2.15) and (2.16) that u and v denote $\mathrm{tr}(P_n)$ and $-\det(P_n)$ respectively. Recall also from Proposition 2.1 that $\det(P_n) = \pm 1$ for n even and odd respectively. Since Equation (2.16) tells us $v = -\det(P_n)$, we have that $v = \pm 1$ for n odd and even.

The following theorems relate the parity of the term u with the divisibility of n.

Theorem 3.1. *For a product matrix P_n with trace u,*

$$u \text{ even} \iff 3 \mid n,$$
$$u \text{ odd} \iff 3 \nmid n.$$

Proof: We can consider the entries in our product matrices modulo 2, and look for cyclic behaviour as we increase n. We have seen that the map mod_2 in (2.5) is a homomorphism, and so we can take the take the entries in A, B modulo 2 before we form the product P_n. Start with $n = 0$. This product matrix is trivial and we have $P_0 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, which has $u = 2$. For $n = 1$ we have the matrices $P_1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & \pm 1 \end{smallmatrix}\right)$, which gives $\mathrm{mod}_2(P_1) = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, and therefore $u = 1$. For $n = 2$ we are essentially squaring this matrix, giving

$$\mathrm{mod}_2(P_2) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \equiv \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

which has $u = 1$. For $n = 3$ we obtain

$$\mathrm{mod}_2(P_3) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

which is the same matrix as the $n = 0$ case. As we increase n, our matrices $\mathrm{mod}_2(P_n)$ form a cycle. We can therefore conclude that u is even if and only if $n \equiv 0 \pmod 3$ and u is odd if and only if $n \equiv 1, 2 \pmod 3$. ◇

Recall that in Theorem 2.1 we also looked at matrices with entries modulo 2, when determining properties of matrices in $K < G$. In the case of $n \equiv 0 \pmod 3$ we can make a further distinction.

Theorem 3.2. *For a product matrix P_n with trace u we have*

$$\begin{aligned} u &\equiv 0 \pmod 4 \iff n \equiv 3 \pmod 6, \\ u &\equiv 2 \pmod 4 \iff n \equiv 0 \pmod 6. \end{aligned}$$

Proof: The map $\mathrm{mod}_4 : \mathrm{S^*L}(2, \mathbb{Z}) \to \mathrm{SL}(2, \mathbb{Z}\backslash 4\mathbb{Z})$, like mod_2 is a homomorphism, and so we can follow the proof of Theorem 3.1 and look for cyclic behaviour as we increase n. Start with $n = 0$. The trivial product matrix $P_0 = I$ has $u = 2$. For $n = 1$, we have the following set of matrices modulo 4:

$$\{\mathrm{mod}_4(P_1)\} = \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix} \right\}.$$

We know from Theorem 3.1 that for $n \equiv 1, 2 \pmod 3$, u is odd and hence u is still odd modulo 4. We are only interested in $n \equiv 0 \pmod 3$. For $n = 3$ we get the following set of matrices modulo 4, namely, $\{\mathrm{mod}_4(P_3)\}$:

$$\left\{ \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} \right\}.$$

Here each u value is $0 \pmod 4$. Similarly, for $n = 6, 9$ we have the sets $\{\mathrm{mod}_4(P_6)\}$ and $\{\mathrm{mod}_4(P_9)\}$ respectively:

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \right\},$$

$$\left\{ \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \right\}.$$

In the case of $n = 6$, notice that each matrix has a u value of 2 and for $n = 9$, each has the u value 0. Also notice that the set of matrices for $n = 9$ matches the set for $n = 3$. This tells us that as we continue to increase n, our sets of matrices $\{\mathrm{mod}_4(P_n)\}$ will start to cycle. We can therefore conclude that for all $n \equiv 3 \pmod 6$, we have $u \equiv 0 \pmod 4$, and for all $n \equiv 0 \pmod 6$ we have $u \equiv 2 \pmod 4$.

The converses of these statements are also true. When $u \equiv 0$ or $2 \pmod 4$ we must have $n \equiv 3$ or $0 \pmod 6$ respectively, since for $n \equiv 1, 2, 4, 5 \pmod 6$, u is odd by Theorem 3.1. $\diamond$

Example 3.1. We will continue with Examples 2.4, 2.6 and 2.7 for the coefficient sequences $\sigma_3 = (++-)$, $\sigma_4 = (++--)$ and $\sigma_6 = (+++---)$ respectively. These examples have product matrices $P_3 = \left(\begin{smallmatrix} 1 & 0 \\ 2 & -1 \end{smallmatrix}\right)$, $P_4 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 3 \end{smallmatrix}\right)$ and $P_6 = \left(\begin{smallmatrix} 3 & -4 \\ 4 & -5 \end{smallmatrix}\right)$ respectively. For matrices P_3 and P_6, we have $3|n$ and $|u| = 0$ and 2. This satisfies Theorem 3.1 which says that u is even if and only if $3|n$. Similarly, for P_4 we have $n = 4$ and $u = 3$. Also, Theorem 3.2 is satisfied here because for P_3 we have $u \equiv 0 \pmod 4$ and for P_6 we have $u \equiv 2 \pmod 4$.

The following theorem, which is one of the main results of this thesis, gives the connection between u-value and type of growth. We can use the standard linear algebra fact that the trace of a matrix is equal to he sum of its eigenvalues, so that in our case

$$u = \lambda_1 + \lambda_2.$$

Theorem 3.3. *Given a product matrix P_n with n odd and $P_n \neq \pm I$, the growth of the corresponding periodic coefficient sequence is*

$$\begin{aligned} \textit{exponential} &\iff u \neq 0, \\ \textit{bounded} &\iff u = 0, \end{aligned}$$

and if n is even the sequence growth is

$$\begin{aligned} \textit{exponential} &\iff |u| > 2, \\ \textit{linear} &\iff |u| = 2, \\ \textit{bounded} &\iff |u| = 1. \end{aligned}$$

If $P_n = \pm I$, growth is bounded.

Proof: Recall that the characteristic equation of the corresponding product matrix has form $x^2 - ux \pm 1$. We have seen in Theorem 2.10 that for $P_n = \pm I$, growth is not linear but bounded, despite the fact that $u = \pm 2$. If we first consider the general n odd case, Proposition 2.1 tells us that $v = 1$, i.e., our characteristic equation has the form $x^2 - ux - 1$. Solving for the eigenvalues gives

$$\lambda_1, \lambda_2 = \frac{u \pm \sqrt{u^2 + 4}}{2}.$$

We know from Theorem 2.10 that the growth of a sequence is bounded if and only if the eigenvalues are distinct roots of unity. Here our discriminant is always positive so the eigenvalues cannot be complex. We can therefore conclude that the growth of a sequence is bounded if and only if $\lambda_1 = 1$, $\lambda_2 = -1$. In this case, we have that

$$u = \lambda_1 + \lambda_2 = 0.$$

Conversely if $u = 0$,

$$\lambda_1, \lambda_2 = \frac{u \pm \sqrt{u^2 + 4}}{2} = \frac{\pm\sqrt{4}}{2} = \pm 1,$$

completing the bounded case.

Since the radical is always positive, we can never have a double root, i.e., growth is never linear. Therefore, since it is the only other option, we must have exponential growth if and only if $u \neq 0$.

If we now consider the n even case, Proposition 2.1 tells us that $v = -1$ and so our characteristic equation has form $x^2 - ux + 1$. Solving for the eigenvalues gives

$$\lambda_1, \lambda_2 = \frac{u \pm \sqrt{u^2 - 4}}{2}.$$

Our radical may be positive, negative or zero. We know from Theorem 2.5 that the growth of our sequence is linear if and only if the eigenvalues are both equal to 1 or -1. If we have such a double root, then

$$u = \lambda_1 + \lambda_2 = \pm 2.$$

Conversely if $u = \pm 2$,

$$\lambda_1 = \lambda_2 = \frac{u \pm \sqrt{u^2 - 4}}{2} = \frac{\pm 2}{2} = \pm 1,$$

completing the linear case.

By Theorem 2.6, growth is bounded if and only if we have distinct eigenvalues that are roots of unity. We have that $v = -1$, implying $\det(P_n) = 1$, and therefore $\lambda_1 \lambda_2 = 1$. If λ_1, λ_2 are distinct roots of unity, they must be complex. If the eigenvalues are complex we need $\sqrt{u^2 - 4} < 0$ so the only possibilities for bounded growth are for $u = 0, \pm 1$. But we have seen in Theorem 3.2 that u can take on the value 0 only when $n \equiv 3 \pmod 6$, in which case n is odd. Therefore bounded growth must occur only when $u = \pm 1$. Conversely, when $u = 1, -1$ we get eigenvalues $\frac{1 \pm \sqrt{-3}}{2}$ and $\frac{-1 \pm \sqrt{-3}}{2}$ respectively. These eigenvalues are pairs of distinct roots of unity and therefore by Theorem 2.6 growth is bounded.

The remaining u values, i.e., $|u| > 2$ therefore correspond to sequences with remaining growth type, exponential growth. $\diamond$

Note that the matrices $\pm I$ are an exception because growth must be bounded, but Theorem 3.3 would have us believe that because $\det(\pm I) = 1$ and $\operatorname{tr}(\pm I) = \pm 2$, the growth is linear.

Example 3.2. We will continue Example 3.1 and consider the growth types of the matrices we have been studying. For the n odd case, P_3, we have $u = 0$ and bounded growth by Theorem 3.3. This growth is verified in Example 2.1. For the n even cases, P_4 and P_6, we have $|u| = 3$ and 2 respectively. These trace values correspond to exponential and linear growth, respectively, as given in Theorem 3.3. These growth types are verified in Examples 2.2 and 2.3.

A similar classification system has been used to sort Möbius transformations according to trace. Given a matrix $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{GL}(2, \mathbb{C})$, we will denote its Möbius transformation on $z \in \mathbb{C} \cup \{\infty\}$ by $f_M(z) := \frac{az+b}{cz+d}$. (This notation differs from the definition in Chapter 1, and there our field was restricted to $\mathbb{R}$.) We will also define the *trace* of a Möbius transformation $f_M(z)$ to be $\operatorname{tr}(M)$, and say that two Möbius transformations are *conjugate* if and only if their corresponding matrices are conjugate. The following information can be found in Beardon [6, p. 66, 67]. We call a Möbius transformation $f_M(z) \neq z$ (i.e., non-identity) *parabolic, loxodromic* or *elliptic* if and only if it has exactly one, exactly two, or infinitely many fixed points, respectively. This classification is invariant under conjugation and furthermore two Möbius transformations f and g are conjugate if and only if $\mathrm{tr}^2(f) = \mathrm{tr}^2(g)$. A similar definition classifies our Möbius transformations based on their traces as follows, where the loxodromic case has been broken down into *hyperbolic* and *strictly loxodromic*, depending on whether or not values are real.

Proposition 3.1. *A Möbius transformation f, where $f \neq z$, with $a, b, c, d \in \mathbb{C}$ is*

1. *parabolic* $\iff$ $\mathrm{tr}^2(f) = 4$,
2. *elliptic* $\iff$ $\mathrm{tr}^2(f) \in [0, 4)$,
3. *hyperbolic* $\iff$ $\mathrm{tr}^2(f) \in (4, \infty)$,
4. *strictly loxodromic* $\iff$ $\mathrm{tr}^2(f) \notin [0, \infty)$.

Notice that the strictly loxodromic case only applies to $\mathrm{tr}(f) \notin \mathbb{R}$. The matrices we are concerned with, P_n, belong to $\mathrm{S^*L}(2, \mathbb{Z})$, and hence have real entries, so we can rephrase the above result as follows.

Proposition 3.2. *A Möbius transformation f, where $f \neq z$, with $a, b, c, d \in \mathbb{R}$ is*

1. *parabolic* $\iff$ $|\mathrm{tr}(f)| = 2$,
2. *elliptic* $\iff$ $|\mathrm{tr}(f)| < 2$,
3. *hyperbolic* $\iff$ $|\mathrm{tr}(f)| > 2$,

This classification of real Möbius transformations comes from the classification of conic sections. Note that this is exactly the breakdown of the n even case in Theorem 3.3. Hyperbolic and parabolic transformations correspond to exponential and linear growth respectively, and the elliptic transformation corresponds to bounded growth. Note that for our specific case of product matrices P_n, it is not possible to have trace zero when n is even, by Theorem 3.2. Every Möbius transformation $f_M(z) \neq z$ with $M \in \mathrm{SL}(2, \mathbb{C})$ is conjugate to one of the following Möbius transformations where we let $a = \lambda_1$ be the dominant eigenvalue of M (see Mumford, Series and Wright

[57, p. 83]): $f(z) = z + a$ for the parabolic case (here $a = \pm 1$), $f(z) = a^2 z$ with $|a| > 1$ for the loxodromic case (hyperbolic case if a^2 is real), and $f(z) = a^2 z$ with $|a| = 1$ and $a \neq \pm 1$ for the elliptic case. Recall that because $\det(M) = 1$ we have $\lambda_1 \lambda_2 = 1$, where $\lambda_1 = a$ and $\lambda_2 = \frac{1}{a}$. We can see by examining the eigenvalues (as was done in Chapter 2) that this is equivalent to saying that every non-identity Möbius transformation $f_M(z)$ has $M \in \mathrm{SL}(2, \mathbb{C})$ which is conjugate to one of the following matrices in Jordan normal form (recall the proof of Theorem 2.11): $\left(\begin{smallmatrix} a & 1 \\ 0 & a \end{smallmatrix}\right)$ for the parabolic case, $\left(\begin{smallmatrix} a & 0 \\ 0 & 1/a \end{smallmatrix}\right)$ for the loxodromic case where $|a| > 1$ (hyperbolic for $a^2 \in \mathbb{R}$) and for elliptic the case where and $|a| = 1$ with $a \neq \pm 1$. Note that if $a = \pm 1$ we would obtain the matrix $\pm I$, which has trace ± 2. This case is an exception, however, because we have already seen that the identity matrix does not correspond to linear growth (parabolic).

Now, what if our matrix M has determinant -1, as is possible for matrices $P_n \in \mathrm{S^*L}(2, \mathbb{Z})$? The results in Beardon [6] were true for any non-identity Möbius transformation and so also hold true if $\det(M) = -1$. From Theorem 3.3 we know that we cannot have linear growth when n is odd (i.e., $\det(M) = -1$). If $\mathrm{tr}(M) = 0$ or 1, growth is bounded or exponential respectively, and these cases both fall under elliptic. If $\mathrm{tr}(M) > 2$ we have exponential growth and a hyperbolic Möbius transformation. Growth type does not quite carry over to transformation type in this case. Similar conjugacy maps and matrices can be constructed.

By combining results from Theorems 3.1 and 3.3 we can state some new relations between the divisibility of n and growth type.

Corollary 3.1. *Given a product matrix $P_n \neq \pm I$ we have the following results about the growth of the corresponding periodic coefficient sequence:*

$$\textit{linear} \Rightarrow 6 \mid n,$$

$$n \textit{ or } u \textit{ odd} \Rightarrow \textit{bounded or exponential.}$$

If $P_n = \pm I$, we have $6 \mid n$.

Proof: Suppose we have the product matrix $P_n \neq \pm I$ which has a linearly growing periodic coefficient sequence. By Theorem 3.3 linear growth occurs only when n is even and $|u| = 2$. Theorem 3.1 tells us that u is even if and only if $3 \mid n$, hence $6 \mid n$.

By the contrapositive, if $6 \nmid n$ then growth is not linear. In other words if n or u is odd (i.e., $2 \nmid n$ or $3 \nmid n$) then growth is bounded or exponential. In the $P_n = \pm I$ case, the proof follows as above because $\det(\pm I) = 1$, implying n is even by Proposition 2.1, and $|u| = |\mathrm{tr}(\pm I)| = 2$. $\diamond$

Corollary 3.2. *Given a product matrix P_n with n odd and $P_n \neq \pm I$, we have the following results about the growth of the corresponding periodic coefficient sequence:*

$$\textit{bounded} \Rightarrow 3 \mid n,$$
$$3 \nmid n \Rightarrow \textit{exponential}.$$

Proof: For n odd, Theorem 3.3 tells us that bounded growth implies $u = 0$. Theorem 3.1 then tells us that for u even, $3 \mid n$. The second statement is simply the contrapositive of the first, where we have used the fact from Theorem 3.3 that when n is odd, if growth is not bounded, it must be exponential. $\diamond$

The converse of this theorem does not hold. If n is odd and $3|n$, it is not necessarily the case that the periodic coefficient sequence is bounded.

Example 3.3. If $\sigma_9 = (++++++++-)$ we have $P_9 = \left(\begin{smallmatrix} 21 & -8 \\ 34 & -13 \end{smallmatrix}\right)$ with $\lambda_1 = 4+\sqrt{17}$, and the corresponding sequence grows exponentially.

It would be interesting to consider some of the results in this section using the alternate matrix pair discussed in Chapter 2, namely $A, \hat{B} = \left(\begin{smallmatrix} 0 & \pm1 \\ 1 & 1 \end{smallmatrix}\right)$. We could also use $A, B' = \left(\begin{smallmatrix} 0 & 1 \\ \pm1 & 1 \end{smallmatrix}\right)$, although the fact that $A^T = A$ and $\hat{B}^T = B'$ means that any product matrix composed of A and B' is simply the transpose of the reverse product matrix made up of A and $\hat{B}$. Using a proof similar to that of Theorem 2.8, we can show that reversing a product matrix (in A, $\hat{B}$) does not change its trace, and neither does taking the transpose. Therefore results on growth and trace will be the same for A, B' as for $A, \hat{B}$. The behaviour of product matrices in $A, \hat{B}$ appears to differ, at least slightly, from the behaviour of the product matrices in A, B. For example, the product $AA\hat{B} = \left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$ has n odd, trace 2 and linear growth. For our original product matrices, linear growth could occur only for $n = 0 \pmod 6$. Further, we could look at traces of products of $A, \hat{B}$ and B in the non-linear case.

Lastly, we will look at another characterization of growth type, similar to that in Theorem 3.3, based on continued fractions. For a coefficient cycle $\sigma_n = (s_1, s_2, \ldots, s_n)$, with $s_i \in \{+, -\}$, we will consider the periodic continued fraction $\gamma_n = [\overline{\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_n}]$, where we consider $\hat{s}_i \in \{1, -1\}$. This is not a simple continued fraction (of the form $[a_1, a_2, \ldots]$ with $a_1 \in \mathbb{Z}, a_i \in \mathbb{N}$ for $i \geq 2$) because entries may be negative. We have that (see Olds [58, p. 89] for example) any (simple) periodic continued fraction represents a quadratic irrational and conversely any quadratic irrational can be written as a continued fraction that is periodic from some point on. (Recall that a quadratic irrational is the root of a quadratic equation with integer coefficients, where the discriminant is positive and non-square.) We suspect that a similar result is true for continued fractions containing negative entries (as long as division by zero is not encountered), in which discriminants may be negative, i.e., we may obtain a "complex quadratic" instead of a quadratic irrational.

Conjecture 3.1. *Given a coefficient cycle σ_n with n odd, the growth of the corresponding periodic coefficient sequence is*

$$\begin{aligned} \textit{exponential} &\iff \gamma \textit{ is a quadratic irrational,} \\ \textit{bounded} &\iff \gamma \textit{ is rational or does not exist,} \end{aligned}$$

and if n is even, the sequence growth is

$$\begin{aligned} \textit{exponential} &\iff \gamma \textit{ is a quadratic irrational,} \\ \textit{linear} &\iff \gamma \textit{ is rational or does not exist,} \\ \textit{bounded} &\iff \gamma \textit{ is a complex quadratic.} \end{aligned}$$

Further, it appears that the radical in an irrational γ is the same as the radical in the eigenvalue corresponding to the coefficient cycle, and hence it does not change under rotation. Also it appears that rational values have $\gamma \leq 2$.

The data was computed using Maple, and so where γ "does not exist" (DNE), a division by zero was most likely encountered. To find γ_n we can take the limit of the convergents of the continued fraction (or a subsequence of the convergents to avoid division by zero) or substitute γ_n into itself and solve. These methods, and resulting values of γ_n need to be further investigated. (The author regrettably ran out of time here!)

Example 3.4. Table 3.1 compiles some results on continued fractions and coefficient cycles. Note that rotating the coefficient cycle $\sigma_6 = (+++++-)$ gives the following

coefficient cycle	n	growth	γ_n	dom. eigenvalue
$(+)$	1	E	$(1+\sqrt{5})/2$	$(1+\sqrt{5})/2$
$(+-)$	2	B	$(1-\sqrt{-3})/2$	$(1+\sqrt{-3})/2$
$(++-)$	3	B	DNE	1
$(+--)$	3	B	1	1
$(-++)$	3	B	-1	1
$(+++-)$	4	B	$(3-\sqrt{-3})/2$	$(1+\sqrt{-3})/2$
$(++--)$	4	E	$(3+\sqrt{5})/2$	$(3+\sqrt{5})/2$
$(++++-)$	5	E	$(5+\sqrt{5})/2$	$(1+\sqrt{5})/2$
$(+++++-)$	6	L	2	1
$(++--+-)$	6	L	DNE	1
$(++++++-)$	7	E	$(13+\sqrt{13})/6$	$(3+\sqrt{13})/2$
$(+++-++-)$	7	E	$(1+\sqrt{5})/2$	$(1+\sqrt{5})/2$
$(++++++--)$	8	E	$(21+\sqrt{221})/22$	$(15+\sqrt{221})/2$
$(+++---++-)$	9	B	$3/2$	1
$(++--++--+-)$	10	E	$(1+\sqrt{21})/2$	$(5+\sqrt{21})/2$
$(++++---+--)$	10	B	$(9-\sqrt{-3})/6$	$(1+\sqrt{-3})/2$
$(+++---+++--)$	11	E	$(9+\sqrt{53})/14$	$(-7-\sqrt{53})/2$
$(+++++++--++-)$	12	L	$5/3$	1

Table 3.1: Coefficient cycles and continued fractions.

values of γ_6: $2, 1, DNE, 0, -1, -\frac{1}{2}$, and rotating $\sigma_{12} = (+++++++--++-)$ gives $\frac{5}{3}, \frac{3}{2}, 2, 1, DNE, 0, -1, -\frac{1}{2}, 2, \frac{1}{3}, -\frac{3}{2}, -\frac{2}{5}$. Rotating the coefficient cycle $\sigma_8 = (+++--+-+-)$ gives $2, DNE, DNE, -1, DNE, \frac{1}{2}, -2, DNE$ and the linear coefficient cycle $\sigma_{12} = (++--+--+-++-)$ gives DNE for every rotation.

Recall that Viswanath considered the random infinite continued fraction $[\pm 1, \pm 1, \pm 1, \ldots]$, and in Equation (1.18), we saw a connection between this type of fraction and the Möbius transformation from a product matrix. We suspect that this connection can aid in the explanation of the above results.

3.2 The Order of a Product Matrix

We have seen that product matrices P_n belong to the group $G = \langle A, B\rangle$, which is a subgroup of $\mathrm{S^*L}(2, \mathbb{Z})$. The following definition can be found in Weinstein [73, p. 84], and will help us to study the connection between the order of a product matrix and its growth type.

Definition 3.1. The *projective special linear group*, $\mathrm{PSL}(2, F)$ for a field F is defined as the quotient group

$$\mathrm{PSL}(2, F) := \mathrm{SL}(2, F)/\mathrm{Z}(\mathrm{SL}(2, F)),$$

where $\mathrm{SL}(2, F)$ is the special linear group and $\mathrm{Z}(\mathrm{SL}(2, F))$ is its center.

We can similarly define this group over a ring R. Recall that the center is the set of elements of $\mathrm{SL}(2, F)$ that commute with all other elements. For our particular group $G \leq \mathrm{S^*L}(2, \mathbb{Z})$ we will be interested in the projective group

$$\mathrm{PS^*L}(2, \mathbb{Z}) = \mathrm{S^*L}(2, \mathbb{Z})/\mathrm{Z}(\mathrm{S^*L}(2, \mathbb{Z})).$$

Note that the group $\mathrm{PSL}(2, \mathbb{Z})$ is isomorphic to the modular group (the group of linear fractional transformations $z \to \frac{az+b}{cz+d}$ with $a, b, c, d \in \mathbb{Z}$ and $ad - bc = 1$). With regard to the center, we have the following theorem which can be found in Weinstein [73, p. 83]. We include the proof so that we can subsequently extend the theorem.

Theorem 3.4. *Let R be a commutative ring with identity. Then* $\mathrm{Z}(\mathrm{SL}(2, R)) = \{\pm I\}$.

Proof: To show equality of these two groups, we will show inclusion in both directions. The first direction, $\{\pm I\} \subseteq \mathrm{Z}(\mathrm{SL}(2, R))$, is trivial because $\pm I$ commutes with every matrix in $\mathrm{SL}(2, R)$.

Conversely, we must show that $\mathrm{Z}(\mathrm{SL}(2, R)) \subseteq \{\pm I\}$. First let $Q \in \mathrm{Z}(\mathrm{SL}(2, R))$. We must then show $Q = \pm I$. For any $M \in \mathrm{SL}(2, R)$ we have $MQ = QM$. Now consider the particular matrix $M = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. We have $0, 1 \in R$ and $\det(S) = 1$ so

$M \in \mathrm{SL}(2, R)$. Therefore we must have

$$\begin{aligned} QM &= \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} Q_1 & Q_1 + Q_2 \\ Q_3 & Q_3 + Q_4 \end{pmatrix} \\ = MQ &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix} = \begin{pmatrix} Q_1 + Q_3 & Q_2 + Q_4 \\ Q_3 & Q_4 \end{pmatrix}. \end{aligned}$$

Equating matrix entries gives the relations

$$Q_1 = Q_1 + Q_3 \Rightarrow Q_3 = 0,$$
$$Q_1 + Q_2 = Q_2 + Q_4 \Rightarrow Q_1 = Q_4.$$

Similarly, consider $M^T = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right) \in \mathrm{SL}(2, R)$. Again we can write

$$\begin{aligned} QM^T &= \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} Q_1 + Q_2 & Q_2 \\ Q_3 + Q_4 & Q_4 \end{pmatrix} \\ = M^TQ &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix} = \begin{pmatrix} Q_1 & Q_2 \\ Q_1 + Q_3 & Q_2 + Q_4 \end{pmatrix}. \end{aligned}$$

Equating entries gives

$$Q_1 + Q_2 = Q_1 \Rightarrow Q_2 = 0.$$

Combining all restrictions, we see that any matrix $Q \in \mathrm{Z}(\mathrm{SL}(2, R))$ must be of the form

$$Q = \begin{pmatrix} Q_1 & 0 \\ 0 & Q_1 \end{pmatrix}.$$

But because $Q \in \mathrm{SL}(2, R)$, we know $\det(Q) = Q_1^2 = 1$. Therefore $Q_1 = \pm 1$ and so $Q \in \{\pm I\}$. ◇

The projective special linear group can therefore be defined as

$$\mathrm{PSL}(2, R) = \mathrm{SL}(2, R)/\{\pm I\}.$$

In other words, the group $\mathrm{PSL}(2, R)$ consists of the equivalence classes $[M] = \{\pm M\}$ where $M \in \mathrm{SL}(2, R)$. The identity is the element $[I] = \{\pm I\}$, which we will simply denote by 1. But what about the determinant -1 case? Here, for simplicity, we restrict our ring R to $\mathbb{C}$.

Theorem 3.5. *The center of $S^*L(2,\mathbb{C})$ is $Z(S^*L(2,\mathbb{C})) \leq \{\pm I, \pm iI\}$.*

Proof: The proof is identical to that of Theorem 3.4 up to showing that any matrix $Q \in \mathrm{S^*L}(2,R))$ must be of the form

$$Q = \begin{pmatrix} Q_1 & 0 \\ 0 & Q_1 \end{pmatrix}.$$

We again have that $\det(Q) = Q_1^2$, but here this implies $Q_1^2 = \pm 1$. Solving, we see that Q_1 can take on values ± 1 or $\pm i$, and so $Q \in \{\pm I, \pm iI\}$. $\diamond$

Now, we are interested in the projective group $\mathrm{PS^*L}(2,\mathbb{Z})$. By Theorem 3.5 the center is $\{\pm I, \pm iI\}$, and restricting to $\mathbb{Z}$, we are back to the group $\{\pm I\}$. We can write

$$\mathrm{PS^*L}(2,\mathbb{Z}) = \mathrm{S^*L}(2,\mathbb{Z})/\{\pm I\}.$$

Since G is a subgroup of $\mathrm{S^*L}(2,\mathbb{Z})$, we can define the group PG which is a subgroup of $\mathrm{PS^*L}(2,\mathbb{Z})$ and has elements $[M] = \{\pm M\}$, where $M \in G$. The identity of both PG and $\mathrm{PS^*L}(2,\mathbb{Z})$ is again $[I] = 1$. The next few theorems are important results relating the order of $[M]$ to the trace of M. The following theorem is taken from Weinstein [73, p. 89]. Again, we include the proof so that the theorem can be extended.

Theorem 3.6. *Let $[M] \in \mathrm{PSL}(2,F)$ where F is a field with $\mathrm{char}(F) \neq 2$. Then $\mathrm{ord}([M]) = 2$ if and only if $\mathrm{tr}(M) = 0$.*

Proof: We will prove the forward direction first. Suppose $\mathrm{ord}([M]) = 2$ for a matrix $M = \left(\begin{smallmatrix} M_1 & M_2 \\ M_3 & M_4 \end{smallmatrix}\right)$ in $\mathrm{SL}(2,F)$. This implies $[M]^2 = 1$ by the definition of order. We can then write $[M^2] = 1$, so that $M^2 = \pm I$, and expanding, we have

$$M^2 = \begin{pmatrix} M_1^2 + M_2M_3 & M_2(M_1+M_4) \\ M_3(M_1+M_4) & M_2M_3 + M_4^2 \end{pmatrix} = \pm I. \tag{3.1}$$

Equating entries gives

$$M_2(M_1+M_4) = M_3(M_1+M_4) = 0.$$

We can break this down into two cases, keeping in mind that we are now dealing with a field F rather than a ring R.

Case 1: We have $M_1 + M_4 = 0$, which implies $\mathrm{tr}(M) = 0$ and we are done.

Case 2: We have $M_2 = M_3 = 0$ and $M_1 + M_4 \neq 0$. It follows that

$$M^2 = \begin{pmatrix} M_1^2 & 0 \\ 0 & M_4^2 \end{pmatrix} = \pm I, \tag{3.2}$$

which implies $M_1 = \pm 1$ and $M_4 = \pm 1$ or $M_1 = \pm i$ and $M_4 = \pm i$, if i belongs to the field F. In the former case we cannot have $\mathrm{char}(F) = 2$, or else $M_1 + M_4 = 0$. We know that $M \in \mathrm{SL}(2, F)$ and so $\det(M) = 1$. This means we must have $M_1 M_4 = 1$, which implies $M_1 = M_4 = \pm 1$, or $M_1 = \pm i, M_4 = \mp i$. In the latter case we have $\mathrm{tr}(M) = i - i = 0$, which contradicts the assumptions of Case 2. Therefore we must have

$$M = \begin{pmatrix} M_1 & 0 \\ 0 & M_4 \end{pmatrix} = \pm I.$$

In other words we have $\mathrm{ord}([M]) = 1$, although we have assumed $\mathrm{ord}([M]) = 2$. This is a contradiction and we are done.

To prove the converse we will assume $\mathrm{tr}(M) = 0$ and first show $[M]^2 = 1$. Note that we must have $\mathrm{char}(M) \neq 2$, because otherwise, $\mathrm{tr}(I) = 0$, and $\mathrm{ord}(I) = 1$, which contradicts the statement of the theorem. We can write

$$\begin{aligned} M_1 + M_4 &= 0, \\ M_1^2 + M_1 M_4 &= 0, \\ -M_1^2 &= M_1 M_4. \end{aligned}$$

Now consider the determinant using the above relation:

$$M_1 M_4 - M_2 M_3 = 1, \tag{3.3}$$

$$-M_1^2 - M_2 M_3 = 1,$$

$$M_1^2 + M_2 M_3 = -1. \tag{3.4}$$

Similarly, multiplying the trace equation by M_4, we obtain $-M_4^2 = M_1 M_4$. Applying this to the determinant equation we then get

$$-M_4^2 - M_2 M_3 = 1,$$

$$M_4^2 + M_2 M_3 = -1. \tag{3.5}$$

Substituting Equations (3.4) and (3.5) into the expression for M^2 given in (3.1), and using the fact that $\mathrm{tr}(M) = M_1 + M_4 = 0$ gives

$$M^2 = \begin{pmatrix} M_1^2 + M_2M_3 & M_2(M_1+M_4) \\ M_3(M_1+M_4) & M_2M_3 + M_4^2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I. \tag{3.6}$$

In terms of cosets we have $[M^2] = [-I] = [M]^2$, and therefore $[M]^2 = 1$. It follows that $\mathrm{ord}([M]) = 2$ because if $M = \pm I$ we would have $M^2 = I$. $\diamond$

Again we are interested in $\mathrm{PS^*L}(2,\mathbb{R})$. Ultimately, we are interested in integer matrices only, and the general proof for the $\mathrm{PS^*L}(2,F)$ case may involve much unnecessary use of complex numbers.

Theorem 3.7. *Let* $[M] \in \mathrm{PS^*L}(2,\mathbb{R})$. *Then* $\mathrm{ord}([M]) = 2$ *if and only if* $\mathrm{tr}(M) = 0$.

Proof: The proof follows almost identically to that of the previous theorem. The key is that for both $\mathrm{PS^*L}(2,\mathbb{R})$ and $\mathrm{PSL}(2,F)$, the identity element is $\{\pm I\}$. If $\det(M) = 1$, the proof is that of Theorem 3.6, where we have $F = \mathbb{R}$. If $\det(M) = -1$, we have to make a few considerations. The forward direction remains the same up until Equation (3.2) where we have deduced that $M_1 = \pm 1$ and $M_4 = \pm 1$ or $M_1 = \pm i$ and $M_4 = \pm i$. If we now consider the fact that $\det(M) = -1$, we have $M_1M_4 = -1$ and are reduced to $M_1 = \pm 1, M_4 = \mp 1$ or $M_1 = M_4 = \pm i$. The latter case cannot occur, though, because we have restricted our matrix entries to $\mathbb{R}$. In the former case, we again have a contradiction because we have assumed in *Case 2* that $M_1 + M_2 \neq 0$. We have therefore shown *Case 2* does not occur and so we must have *Case 1*, where $\mathrm{tr}(M) = 0$.

To prove the converse, we follow the steps in Theorem 3.6, except the fact that $\det(M) = -1$ is going to change the sign in Equations (3.4) and (3.5) to

$$M_1^2 + M_2M_3 = 1,$$
$$M_4^2 + M_2M_3 = 1.$$

Since we have assumed $\mathrm{tr}(M) = 0$, Equation (3.6) becomes

$$M^2 = \begin{pmatrix} M_1^2 + M_2M_3 & M_2(M_1+M_4) \\ M_3(M_1+M_4) & M_2M_3 + M_4^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I, \tag{3.7}$$

so that $[M^2] = 1 = [M]^2$. This tells us that $\mathrm{ord}([M]) = 1$ or 2. Suppose $\mathrm{ord}([M]) = 1$. This is true only for $M = \pm I$, in which case we have $\mathrm{tr}(M) = \pm 2$. This is a contradiction because we have assumed $\mathrm{tr}(M) = 0$. Therefore we must have $\mathrm{ord}([M]) = 2$, completing the proof. $\diamond$

Example 3.5. The matrix $P_3 = \left(\begin{smallmatrix} 1 & 0 \\ 2 & -1 \end{smallmatrix}\right)$ we considered in Example 2.4 has $u = \mathrm{tr}(P_3) = 0$ and $P_3^2 = I$. Some other longer examples include $\sigma_9 = (++++-++++-)$ and $\sigma_9 = (+++--+-+-)$, which both have $P_9 = \left(\begin{smallmatrix} 3 & -2 \\ 4 & -3 \end{smallmatrix}\right)$ and hence $u = 0$. Squaring P_9 gives I as required. From Theorem 3.2 we know that $u \equiv 0 \pmod 4$ if and only if $n \equiv 3 \pmod 6$, and therefore $u = 0$ only if $n \equiv 3 \pmod 6$. We can conclude that $n = 9$ is the next instance after $n = 3$ where we may have $\mathrm{ord}([P_n]) = 2$. Note also that the eigenvalues of both P_3 and P_9 are ± 1. These matrices have order 2, and likewise, squaring the eigenvalues gives us 1.

Recall that we are interested in matrices $P_n \in G \leq \mathrm{S^*L}(2, \mathbb{Z})$, or in terms of equivalence classes, $[P_n] \in PG \leq \mathrm{PS^*L}(2, \mathbb{Z})$. The previous two theorems, which tell us about the connection between the order and the trace of a matrix are true for matrices with entries in $\mathbb{R}$ and therefore must also be true for matrices with strictly integer entries. We can be a bit more specific about the order of a matrix $M \in \mathrm{S^*L}(2, \mathbb{R})$, rather than the equivalence class $[M]$.

Corollary 3.3. *Let* $M \in \mathrm{S^*L}(2, \mathbb{R})$ *with* $\mathrm{tr}(M) = 0$. *Then*

$$\det(M) = 1 \iff M^2 = -I \iff \mathrm{ord}(M) = 4,$$
$$\det(M) = -1 \iff M^2 = I \iff \mathrm{ord}(M) = 2.$$

Proof: We know that $\det(M) = \pm 1$. Equation (3.6) in the proof of Theorem 3.6 tells us that if $\det(M) = 1$ then $M^2 = -I$, in which case $\mathrm{ord}(M) = 4$. If $\det(M) = -1$, Equation (3.7) tells us that $M^2 = I$, in which case $\mathrm{ord}(M) = 2$. We cannot have $\mathrm{ord}(M) = 1$, because it would follow that $\mathrm{tr}(M) = 2$. Conversely, if we start with $M^2 = \pm I$, we can work backwards to show that $\det(M) = \mp 1$. Further, if we know $\mathrm{ord}(M) = 2$, it follows by definition that $M^2 = I$. Now suppose that $\mathrm{ord}(M) = 4$.

Since we have assumed $\mathrm{tr}(M) = 0$, we have that

$$M^2 = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}^2 = \begin{pmatrix} a^2+bc & 0 \\ 0 & a^2+bc \end{pmatrix} = (a^2+bc)I.$$

We therefore have that $M^4 = (a^2+bc)^2 I = I$, from which we can deduce that $(a^2+bc) = \pm 1$. This term cannot take on the value 1 because in this case we would have $M^2 = I$, and hence $\mathrm{ord}(M) = 2$. Therefore, we must have $(a^2+bc) = -1$, implying $M^2 = -I$. ◇

Again we can apply this result to the specific case of integer matrices. Note that in Example 3.5, both matrices had determinant -1 and order 2.

In the following, we can take inverses of the elements in the field F, and elements of $\mathrm{PSL}(2,F)$ are also invertible because the matrices are invertible. The next theorem is taken from Weinstein [73, p. 90].

Theorem 3.8. *Let* $[M] \in \mathrm{PSL}(2,F)$*, where* F *is a field with* $\mathrm{char}(F) \neq 3$*. Then* $\mathrm{ord}([M]) = 3$ *if and only if* $\mathrm{tr}(M) = \pm 1$*.*

Proof: We will prove the forward direction first. Assume $\mathrm{ord}([M]) = 3$, i.e., assume $[M^3] = 1$ and $[M] \neq 1$. (If $[M] = 1$ we would have $\mathrm{ord}([M]) = 1$, which is a contradiction.) The matrix M must differ in at least one entry from $\pm I$, and we must have $M^3 = \pm I$, which can be written as $M^2 = \pm M^{-1}$. Using the fact that $\det(M) = 1$, writing out these matrices gives

$$\begin{pmatrix} M_1^2 + M_2M_3 & M_2(M_1+M_4) \\ M_3(M_1+M_4) & M_2M_3 + M_4^2 \end{pmatrix} = \frac{\pm 1}{\det(M)} \begin{pmatrix} M_4 & -M_2 \\ -M_3 & M_1 \end{pmatrix} = \begin{pmatrix} \pm M_4 & \mp M_2 \\ \mp M_3 & \pm M_1 \end{pmatrix}. \tag{3.8}$$

Depending on the values of M_2 and M_3, we have the following three subcases:

Case 1: $M_2 \neq 0$: Equating matrix entries $(M^2)_2$ and $(\pm M^{-1})_2$ (here the subscript denotes position in the matrix, as was the case with M) yields $M_2(M_1+M_4) = \mp M_2$, which implies $(M_1+M_4) = \mathrm{tr}(M) = \mp 1$, as required.

Case 2: $M_3 \neq 0$: Similarly, equating $(M^2)_3$ and $(\pm M^{-1})_3$ yields $M_3(M_1+M_4) = \mp M_3$, again implying $\mathrm{tr}(M) = \mp 1$.

Case 3: $M_2 = M_3 = 0$: Equating $(M^2)_1$ and $(\pm M^{-1})_1$ gives us

$$\begin{aligned} M_1^2 + M_2M_3 &= \pm M_4, \\ M_1^2 &= \pm M_4. \end{aligned} \tag{3.9}$$

Now, our determinant is $\det(M) = M_1M_4 = 1$ and so Equation (3.9) can be written as $M_1^3 = \pm 1$. But we must have $M_1 \neq \pm 1$ because otherwise, the fact that $M_1M_4 = 1$ would imply $M_4 = \pm 1$ also, and thus $M = \pm I$ which is a contradiction. (Note that M_1 must therefore be a complex cube root of unity.) We can write $M^3 \mp 1 = 0$ and factor to give $(M_1 \mp 1)(M_1^2 \pm M_1 + 1) = 0$. Since $M_1 \neq \pm 1$ we must therefore have $M_1^2 \pm M_1 + 1 = 0$. But by Equation (3.9) we can write this as $\pm M_4 \pm M_1 + 1 = 0$, i.e., $\mathrm{tr}(M) = \mp 1$ and we are done.

To show the converse we will assume $\mathrm{tr}(M) = \pm 1$ for $M \in \mathrm{SL}(2, F)$ and show $\mathrm{ord}([M]) = 3$. Since $\mathrm{tr}(\pm I) = \pm 2$ we must have $M \neq \pm I$. (Note that this is because if $2 = 1$ we have $1 = 0$, which is false in any field, and if $2 = -1$ we have $3 = 0$, which is false because $\mathrm{char}(F) \neq 3$.) Therefore $[M]$ is nontrivial, i.e., $\mathrm{ord}([M]) \neq 1$. We now need to show that $M^3 = \pm I$.

By the assumption on the trace, we know $M_1 + M_4 = \pm 1$. Multiplying through by M_1 and using the fact that $\det(M) = M_1M_4 - M_2M_3 = 1$ we have

$$\begin{aligned} M_1^2 + M_1M_4 &= \pm M_1, \\ M_1^2 + 1 + M_2M_3 &= \pm M_1, \\ M_1^2 + M_2M_3 &= \pm M_1 - 1 = \pm M_1 \mp \mathrm{tr}(M), \\ M_1^2 + M_2M_3 &= \pm M_1 \mp M_1 \mp M_4, \\ M_1^2 + M_2M_3 &= \mp M_4. \end{aligned} \tag{3.10}$$

Similarly, if we multiply the trace equation by M_4 we obtain

$$\begin{aligned} M_1M_4 + M_4^2 &= \pm M_4, \\ 1 + M_2M_3 + M_4^2 &= \pm M_4, \\ M_4^2 + M_2M_3 &= \pm M_4 - 1 = \pm M_4 \mp \mathrm{tr}(M), \\ M_4^2 + M_2M_3 &= \pm M_4 \mp M_1 \mp M_4, \\ M_4^2 + M_2M_3 &= \mp M_1. \end{aligned} \tag{3.11}$$

Now we want an expression for M^3. We have seen the following expansion for M^2:

$$M^2 = \begin{pmatrix} M_1^2 + M_2M_3 & M_2(M_1+M_4) \\ M_3(M_1+M_4) & M_2M_3 + M_4^2 \end{pmatrix} = \begin{pmatrix} \mp M_4 & \pm M_2 \\ \pm M_3 & \mp M_1 \end{pmatrix},$$

and have simplified it using Equations (3.10) and (3.11) as well as the trace equation. We can now give an expression for M^3:

$$M^3 = \begin{pmatrix} \mp M_4 & \pm M_2 \\ \pm M_3 & \mp M_1 \end{pmatrix} \begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix} = \begin{pmatrix} \mp M_1M_4 \pm M_2M_3 & 0 \\ 0 & \pm M_2M_3 \mp M_1M_4 \end{pmatrix}.$$

Using the fact that $\det(M) = 1$ we have $M^3 = \mp I$ as required, completing the proof.
$\diamond$

Note that this theorem is true for $M \in \mathrm{SL}(2, F)$ and not necessarily for $M \in \mathrm{S^*L}(2, F)$, i.e, we can apply this theorem only to those $[P_n] \in PG$ with determinant 1. (Matrices in the same equivalence class will have the same determinant.) Also, in the case of $\mathrm{PSL}(2, \mathbb{R})$, *Case 3* cannot occur because it requires M_1 to be complex. Again we can state some further results for matrices $M \in \mathrm{SL}(2, \mathbb{R})$, without the use of equivalence classes.

Corollary 3.4. *For a matrix* $M \in \mathrm{SL}(2, F)$ *we have*

$$\mathrm{tr}(M) = 1 \iff M^3 = -I \Rightarrow \mathrm{ord(M)} = 6,$$
$$\mathrm{tr}(M) = -1 \iff M^3 = I \iff \mathrm{ord}(M) = 3.$$

Proof: In the proof of Theorem 3.8 we state that $M^3 = \pm I$ and following through we see that this implies $\mathrm{tr}(M) = \mp 1$. Conversely, we assume $\mathrm{tr}(M) = \pm 1$ and continue to show that $M^3 = \mp I$, in which case $\mathrm{ord}(M)$ is equal to 6 or 3 respectively. (Note that we cannot have $\mathrm{ord}(M) = 1$, because this would imply $\mathrm{tr}(M) = 2$.) Also, by definition, $\mathrm{ord}(M) = 3$ implies $M^3 = I$, but $\mathrm{ord}(M) = 6$ does not necessarily imply $M^3 = -I$ without any restriction on the trace. $\diamond$

Again, this result only holds for those $P_n \in G$ for which $\det(P_n) = 1$.

Example 3.6. The matrices $P_2 = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $P_4 = \left(\begin{smallmatrix} 2 & -1 \\ 3 & -1 \end{smallmatrix}\right)$ with coefficient cycles $\sigma_2 = (+-)$ and $\sigma_4 = (+++-)$ (from Example 2.6), respectively, both have $u = 1$ and give

$-I$ when cubed, implying order 6. Also, the coefficient cycle $\sigma_8 = (++--+-+-)$ has product matrix $P_8 = \left(\begin{smallmatrix} 1 & -1 \\ 3 & -2 \end{smallmatrix}\right)$. Here we have $u = -1$ and $P_8^3 = I$ as required by Corollary 3.4. Note that for each of these examples the determinant is 1. Recall from Example 2.10 that the eigenvalues of P_4 had the property $\lambda_1^6, \lambda_2^6 = 1$. P_2 has the same eigenvalues $\lambda_1, \lambda_2 = \frac{1+\sqrt{-3}}{2}$. Similarly, the eigenvalues of P_8 are cube roots of unity, namely, $\lambda_1, \lambda_2 = \frac{-1+\sqrt{-3}}{2}$.

Now the question arises — what happens if $M \in \mathrm{S^*L}(2, F)$? We first consider the reverse direction of the previous theorem with $\det(M) = -1$. Again we will consider the case for $F = \mathbb{R}$, since we are ultimately concerned with integer matrices, and the center, $\{\pm I\}$, is easier to work with.

Theorem 3.9. *Let* $[M] \in \mathrm{PS^*L}(2, \mathbb{R})$. *If* $\det(M) = -1$ *and* $\mathrm{tr}(M) = \pm 1$, *then* $\mathrm{ord}([M]) \neq 3$.

Proof: First suppose $\mathrm{tr}(M) = \pm 1$ for $M \in \mathrm{S^*L}(2, F)$ with $\det(M) = -1$. We cannot have $M = \pm I$ because our trace must be ± 1, and so $\mathrm{ord}([M]) \neq 1$. We want to show $M^3 \neq \pm I$. Following the steps in the proof of Theorem 3.8, we can take the trace assumption $M_1 + M_4 = \pm 1$, multiply by M_1 and use the fact that $\det(M) = M_1M_4 - M_2M_3 = -1$ as follows:

$$\begin{aligned} M_1^2 + M_1M_4 &= \pm M_1, \\ M_1^2 + M_2M_3 - 1 &= \pm M_1, \\ M_1^2 + M_2M_3 &= \pm M_1 + 1. \end{aligned} \tag{3.12}$$

(Note that the right-hand side of Equation (3.12) cannot be simplified using the definition of trace as in the $\det(M) = 1$ case. We could write $\pm M_1 + 1 = \pm M_1 \pm \mathrm{tr}(M) = \pm 2M_1 \pm M_4$.) Similarly, multiplying the trace equation by M_4 gives:

$$\begin{aligned} M_1M_4 + M_4^2 &= \pm M_4, \\ M_2M_3 - 1 + M_4^2 &= \pm M_4, \\ M_4^2 + M_2M_3 &= \pm M_4 + 1. \end{aligned} \tag{3.13}$$

Using Equations (3.12) and (3.13) as well as the trace formula we have

$$M^2 = \begin{pmatrix} M_1^2 + M_2M_3 & M_2(M_1 + M_4) \\ M_3(M_1 + M_4) & M_2M_3 + M_4^2 \end{pmatrix} = \begin{pmatrix} \pm M_1 + 1 & \pm M_2 \\ \pm M_3 & \pm M_4 + 1 \end{pmatrix}.$$

We then have the following expression for M^3, which we can simplify, again using Equations (3.12) and (3.13) and the trace equation:

$$\begin{aligned}
M^3 &= \begin{pmatrix} \pm M_1 + 1 & \pm M_2 \\ \pm M_3 & \pm M_4 + 1 \end{pmatrix} \begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix} \\
&= \begin{pmatrix} \pm M_1^2 + M_1 \pm M_2 M_3 & \pm M_1 M_2 + M_2 \pm M_2 M_4 \\ \pm M_1 M_3 \pm M_3 M_4 + M_3 & \pm M_2 M_3 \pm M_4^2 + M_4 \end{pmatrix} \\
&= \begin{pmatrix} \pm(M_1^2 + M_2 M_3) + M_1 & \pm M_2(M_1 + M_4 \pm 1) \\ \pm M_3(M_1 + M_4 \pm 1) & \pm(M_4^2 + M_2 M_3) + M_4 \end{pmatrix} = \begin{pmatrix} 2M_1 \pm 1 & 2M_2 \\ 2M_3 & 2M_4 \pm 1 \end{pmatrix}.
\end{aligned}$$

Now suppose $M^3 = \pm I$. In this case we must have $M_2 = M_3 = 0$. We must also have $2M_1 \pm 1 = 2M_4 \pm 1 = 1$ or -1. Note that if either $M_1 = 0$ or $M_4 = 0$, our matrix M will have at least three zero entries and hence the determinant will be zero. This is a contradiction to our assumption. The only other option is that $M_1 = M_4 = \mp 1$, in which case $M = \mp I$ and again we have a contradiction, since we have shown $M \neq \pm I$. (Even if M could equal $\pm I$ this would prevent us from obtaining an order 3 matrix). We have therefore shown that $M^3 \neq \pm I$, i.e., $\mathrm{ord}([M]) \neq 3$. $\diamond$

We cannot simply deduce that this theorem is true based on Theorems 2.11 and 3.3 because they deal with the specific matrices P_n.

Example 3.7. Consider the coefficient cycle $\sigma_5 = (++++-)$, with corresponding product matrix $P_5 = \left(\begin{smallmatrix} 3 & -1 \\ 5 & -2 \end{smallmatrix}\right)$. We have trace $u = 1$ and $\det(P_5) = -1$, but $P_5^3 = \left(\begin{smallmatrix} 7 & -2 \\ 10 & -3 \end{smallmatrix}\right)$.

Now what happens to the forward direction of Theorem 3.8 when $\det(M) = -1$? This is precisely the contrapositive of Theorem 3.9. Assuming that $\mathrm{ord}([M]) = 3$, we then have that $\mathrm{tr}(M) \neq \pm 1$.

So far, we have been looking at the orders of matrices belonging to the groups $\mathrm{SL}(2, \mathbb{R})$ and $\mathrm{S^*L}(2, \mathbb{R})$. There is more to be said if we consider our specific group $G = \langle A, B \rangle$. We can conclude the following results by combining the results of the previous theorems and corollaries concerning order with earlier results on growth type.

Corollary 3.5. *For a product matrix P_n, we have that* $\text{tr}(P_n) = 0$ *if and only if* $\text{ord}(P_n) = 2$.

Proof: If $\text{ord}(P_n) = 2$ we have by Theorem 3.6 that $\text{tr}(P_n) = 0$. Conversely, suppose that $\text{tr}(P_n) = 0$. We want to discount the possibility in Theorem 3.6 of $\text{ord}(P_n) = 4$, so that we are left with $\text{ord}(P_n) = 2$. By Theorem 3.2, we have that if $\text{tr}(P_n) = 0$ (i.e., $|u| = 0$), then $n \equiv 3 \pmod 6$. But by Corollary 3.3, since n is odd (i.e., $\det(P_n) = -1$), we must have $\text{ord}(P_n) = 2$. $\diamond$

Note that in Example 3.5, all matrices had order 2.

Corollary 3.6. *For a product matrix $P_n \neq \pm I$ with $\det(P_n) = 1$:*

$$\text{ord}([P_n]) = 3 \iff \textit{bounded},$$

For a product matrix $P_n \neq \pm I$ with $\det(P_n) = -1$:

$$\text{ord}(P_n) = 2 \iff \textit{bounded},$$

Proof: Suppose $[P_n]$ has order 2 or 3. By Theorem 2.11 growth must be bounded. Now suppose growth is bounded. Theorem 3.3, for the n even case, tells us that $|u| = |\text{tr}(P_n)| = 1$. By Theorem 3.8, the order of $[P_n]$ must be 3. If we have $\det(P_n) = -1$ (i.e., n odd), Theorem 3.3 tells us that growth is bounded if and only if $\text{tr}(M) = 0$. But by Corollary 3.5, $\text{tr}(P_n) = 0$ if and only if $\text{ord}(P_n) = 2$. $\diamond$

Note that Corollary 3.6 tells us that matrices with $\det(P_n) = -1$ and finite order (i.e., bounded, by Theorem 2.11) can only be of order 2, and so there are no matrices P_n with $\det(P_n) = -1$ and $\text{ord}([P_n]) = 3$. This case was discussed in general in Theorem 3.9, which stated that $\det(M) = -1$ and $\text{tr}(M) = \pm 1$ imply $\text{ord}([M]) \neq 3$. Also, as a follow up to Theorem 3.9, Theorem 3.3 tells us that for $u = \text{tr}(P_n) = \pm 1$ and $\det(P_n) = -1$ odd, we have exponential growth.

We can now completely characterize bounded sequences for any value of n.

Corollary 3.7. *For any product matrix $P_n \neq \pm I$,*

$$\textit{bounded} \iff \text{ord}(P_n) = 2, 3 \textit{ or } 6.$$

Proof: This follows directly from Corollary 3.6, using the fact that $\mathrm{ord}([P_n]) = 3$ implies $\mathrm{ord}(P_n) = 3$ or 6. $\diamond$

Note that it is known that any element $M \in \mathrm{GL}(2,\mathbb{Z})$ having finite order (i.e., having bounded growth for P_n) has order 1, 2, 3, 4 or 6. The order 1 case corresponds to $P_n = I$. The proof can be found in Kuzmanovich and Pavlichenkov [47], for example, where it is shown that for the prime decomposition $m = p_1^{e_1} p_2^{e_2} \cdots p_t^{e_t}$ with $p_1 < p_2 < \cdots < p_t$, an element in $\mathrm{GL}(\mathrm{n},\mathbb{Z})$ has order m if and only if one of the following holds:

$$\sum_{i=1}^{t} (p_i - 1)p_i^{e_i - 1} - 1 \leq n \text{ for } p_1^{e_1} = 2,$$
$$\sum_{i=1}^{t} (p_i - 1)p_i^{e_i - 1} \leq n \text{ otherwise.}$$

In the case of $n = 2$, we can deduce that $M \neq \pm I$ in $\mathrm{GL}(2,\mathbb{Z})$ has order 2, 3, 4 or 6. These are therefore the only possible orders for our product matrices $P_n \in G \leq \mathrm{GL}(2,\mathbb{Z})$.

Also, Weinstein [73, p. 94] states that $[M] \in \mathrm{PSL}(2,F)$ has order 4 if and only if $\mathrm{tr}(M) = \pm\sqrt{2}$ and $[M] \in \mathrm{PSL}(2,F)$ has order 6 if and only if $\mathrm{tr}(M) = \pm\sqrt{3}$. Since our group G contains only integer matrices, $[M] \in \mathrm{PG}$ cannot have orders 4 or 6. This does not prevent M from having order 6, however, because we may still have the case $\mathrm{ord}([M]) = 3$.

McGuire [55] gives a proof of Corollary 3.7 using eigenvalues in polar coordinates. In his proof he uses the facts that if $\lambda_1, \lambda_2 = \pm 1$ then $\mathrm{ord}(P_n) = 2$ (in which case $\det(P_n) = -1$ and $\mathrm{tr}(P_n) = 0$), and if $\mathrm{tr}(P_n) = \pm 1$ for P_n with finite order, then $\mathrm{ord}(P_n) = 6, 3$. He also gives the following necessary condition for sequences to be periodic (i.e., bounded by Corollary 2.2). In our terminology, given a coefficient cycle of length n determining a bounded periodic coefficient sequence, we must have

$$\mathrm{ord}(P_n) \cdot n \equiv 0 \pmod 3.$$

He uses the fact that if the initial values of our sequence are x and y, the subsequent terms will contain a Fibonacci number of x's and y's. The result follows from some parity arguments and the idea that for the sequence to be periodic, we must eventually return to terms x and y.

McGuire's result is implicit in our previous results on order and trace. In Corollary 3.2 we have shown that for n odd, a bounded periodic coefficient sequence implies $n \equiv 0 \pmod 3$. Further, for n even, i.e., $\det(P_n) = 1$, Corollary 3.6 tells us that our periodic coefficient sequence is bounded if and only if $\mathrm{ord}(P_n) = 3$ or 6. Therefore, combining these corollaries we can conclude that for a bounded sequence, either n or $\mathrm{ord}(P_n)$ is divisible by 3. The converse of this statement is not true, however. If $n = 6$ then $\mathrm{ord}(P_n)n \equiv 0 \pmod 3$, but Table 2.1 gives several examples of length 6 coefficient cycles which do not give bounded growth.

Interestingly, McGuire [56] generalized this result to random m-nacci (m^{th} order) sequences. Instead of generating periodic coefficient sequences using coefficient cycles of $\pm$ signs, he uses size-n sets of $(m-1)$-tuples of $\pm$ signs. Using products of $m \times m$ matrices he shows that for a bounded periodic coefficient sequence, we must have

$$\mathrm{ord}(P_n) \cdot n \equiv 0 \pmod{m+1}.$$

We have been working with powers of matrices above and we obtain the following neat trace result when we consider the square of a matrix M.

Theorem 3.10. *Let* $M \in \mathrm{S}^*\mathrm{L}(2,\mathbb{R})$ *with* $\det(M) = \pm 1$. *Then* $\mathrm{tr}(M^2) = (\mathrm{tr}(M))^2 \mp 2$.

Proof: Suppose we have $M = \left(\begin{smallmatrix} M_1 & M_2 \\ M_3 & M_4 \end{smallmatrix}\right) \in \mathrm{S}^*\mathrm{L}(2,\mathbb{R})$. The trace is given by $\mathrm{tr}(M) = M_1 + M_4$ and so $\mathrm{tr}(M)^2 = M_1^2 + M_4^2 + 2M_1M_4$. Squaring M gives

$$M^2 = \begin{pmatrix} M_1^2 + M_2M_3 & M_2(M_1+M_4) \\ M_3(M_1+M_4) & M_2M_3 + M_4^2 \end{pmatrix},$$

with $\mathrm{tr}(M^2) = M_1^2 + M_4^2 + 2M_2M_3$. Because we know $\det(M) = M_1M_4 - M_2M_3 = \pm 1$, we can rewrite $\mathrm{tr}(M^2)$ as

$$\begin{aligned} \mathrm{tr}(M^2) &= M_1^2 + M_4^2 + 2M_1M_4 \mp 2, \\ &= \mathrm{tr}(M)^2 \mp 2, \end{aligned}$$

completing the proof. $\diamond$

It would also be interesting to consider some of the results in this section using product matrices formed from $A, \hat{B}$. We have already noted that $\hat{B}^3 = -I$, and so $\text{ord}([\hat{B}]) = 3$. Another interesting product is $(AA\hat{B}\hat{B})^3 = I$. Further, we could look at orders of products of $A, \hat{B}$ and B in the non-linear case. Those results which apply to matrices in the general groups $\mathrm{SL}(2, \mathbb{Z})$ and $\mathrm{S^*L}(2, \mathbb{Z})$ will carry over.

3.3 Approximating the Almost Sure Growth Rate Using Averages

In approximating the almost sure growth rate we will only need the absolute value of the traces of product matrices. Recall from Proposition 2.3 that given a product matrix P_n, the growth rate of the associated sequence is given by $|\lambda_1|^{\frac{1}{n}}$, i.e., the n^{th} root of the absolute value of the dominant eigenvalue of P_n. Also recall that the corresponding characteristic equation is given by $x^2 - ux - v$, so that the eigenvalues are $\lambda_1, \lambda_2 = \frac{u \pm \sqrt{u^2+4v}}{2}$. Using the facts that $u = \text{tr}(P_n)$ and $v = \pm 1$ for n odd and even respectively, the absolute value of the dominant eigenvalue is thus

$$|\lambda_1| = \frac{\left||\text{tr}(P_n)| + \sqrt{\text{tr}(P_n)^2 \pm 4}\right|}{2}. \tag{3.14}$$

We have in fact used this method of eigenvalues to find the growth rate of a random Fibonacci sequence corresponding to a given product matrix, in Chapter 1. Recall that Table 1.4 gave the averages of 20 growth rates of random Fibonacci sequences with length 40002, using the program in Figure A.3 to compute 20 product matrices P_n at random, with $n = 40000$, and find the average of the n^{th} roots of their dominant eigenvalues. This is in contrast to computing a length-n random Fibonacci sequence and approximating its growth rate using $|t_n|^{\frac{1}{n}}$. Note that this eigenvalue approximation of Viswanath's constant is independent of the initial values of the sequence.

Removing the randomness, we now consider the set of all possible length-n coefficient cycles σ_n. Note that this generates a corresponding set of periodic coefficient sequences of infinite length which may approximate random Fibonacci sequences as n gets large. By taking the average of the growth rates for all corresponding product matrices of length n, and letting n get very large, we can effectively estimate Viswanath's constant. We know that with probability 1, any random Fibonacci sequence has growth rate e^{γ_f} (recall this notation for Viswanath's constant from Chapter 1), and so the average of all possible growth rates must also be e^{γ_f}, because we can combine terms to rewrite this average (in the limit) as the expected value $1(e^{\gamma_f}) + 0(c) = e^{\gamma_f}$, where c is any other growth rate. We can write our approximation as

$$e^{\gamma_f} \approx \frac{1}{2^n} \sum_{\text{all } P_n} \left(\frac{||\mathrm{tr}(P_n)| + \sqrt{\mathrm{tr}(P_n)^2 \pm 4}|}{2} \right)^{\frac{1}{n}}, \tag{3.15}$$

where we choose the $+4$ term if n is odd and -4 if n is even. Table 3.2 gives approximate growth rates using averages, for $n \leq 19$. The Maple program used here is given in Figure A.4 of Appendix A, for $n = 9$. Notice that for $n = 19$ we are averaging

n	average of g.r.
1	1.618033988
2	1.309016994
3	1.154508497
4	1.145259161
5	1.186117799
6	1.083345517
7	1.184039573
8	1.157659225
9	1.142422955
10	1.140061077
11	1.151816913
12	1.119038560
13	1.152772740
14	1.143234055
15	1.138393665
16	1.137198656
17	1.142272071
18	1.127914590
19	1.142873068

Table 3.2: Average values of growth rates.

the growth rates of 524288 sequences, and still we have only achieved accuracy to one decimal place. This gives us an indication of how difficult the computation of Viswanath's constant really is.

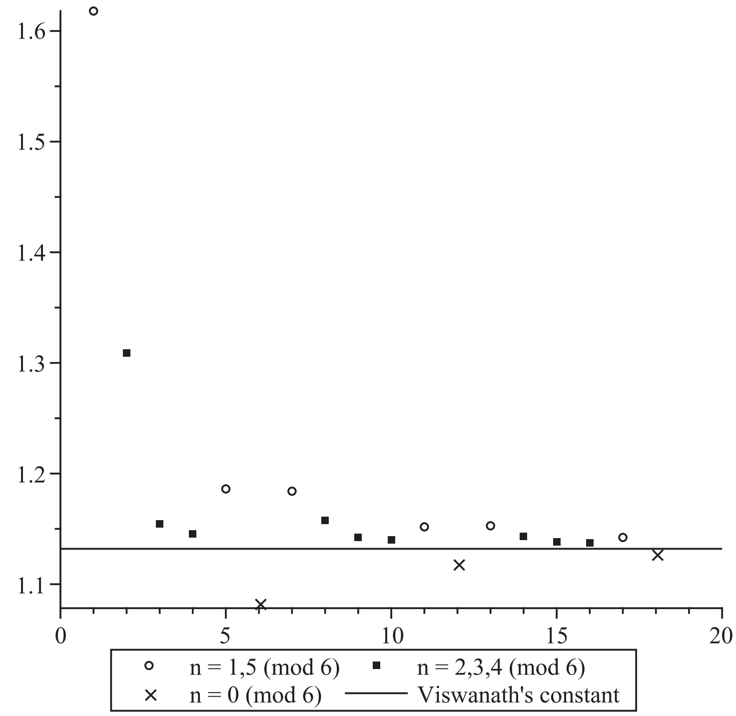


Figure 3.1: Average values of growth rates.

Plotting the values in Table 3.2 to obtain Figure 3.1 gives us some new insights. It is easily recognizable from Figure 3.1 that the growth rate follows a pattern for values of $n \pmod 6$. We can partially explain this pattern by considering some of the previously gathered results on growth type. First consider the points marked with an x. These occur when $n \equiv 0 \pmod 6$. Corollary 3.1 tells us that if growth is linear then $6|n$. Also, for linear growth we have $|\mathrm{tr}(P_n)| = 2$ (by Theorem 3.3) and growth rate

$$\left(\frac{2+\sqrt{4-4}}{2}\right)^{\frac{1}{n}} = 1^{\frac{1}{n}} = 1,$$

and so we might expect that when we take the average, the growth rate for $n \equiv 0 \pmod 6$ will be lower than the average growth rate for all n. Similarly, for n odd, Corollary 3.2 tells us that if $3 \nmid n$, growth is exponential. Therefore, for $n \equiv 1, 5 \pmod 6$, the growth rates must be strictly greater than 1 and we might expect the average of growth rates for such n to be higher than the average growth rate for all n, as evidenced by the circles in Figure 3.1. If $n \equiv 2, 3, 4 \pmod 6$ we have a mixture of bounded and exponential growth and so we may expect the average of the growth rates to be lower than the exponential only case but higher than the linear case, as evidenced by the squares in Figure 3.1.

We can simplify our average growth rate calculation by noticing that in certain cases, for large values of n, we have the approximation $|\text{tr}(P_n)| = |u| \approx |\lambda_1|$ as follows. For large n, if $|u|$ is also large, the ± 4 term becomes negligible and we can write Equation (3.14) as

$$|\lambda_1| = \frac{||u| + \sqrt{u^2 \pm 4}|}{2} \approx \frac{||u| + \sqrt{u^2}|}{2} = |u|. \tag{3.16}$$

If n is even and $|u| = 1$ then growth is bounded (by Theorem 3.3) and we have $|\lambda_1| = \left|\frac{1+\sqrt{-3}}{2}\right| = 1$. Here $|u| = |\lambda_1|$ and our simplification is exact, and it is is the only case where we can obtain a complex eigenvalue. Also, notice that in general, when we calculate the average growth rate, we must take the n^{th} root of each λ_1, so that for any non-zero $|u|$ and n large enough,

$$|u|^{\frac{1}{n}} \approx |\lambda_1|^{\frac{1}{n}},$$

making the approximation even closer. Assuming we can always make the simplification to $|u|$, Equation (3.15) for the average growth rate can now be approximated as

$$e^{\gamma_f} \approx \frac{1}{2^n} \sum_{\text{all } P_n} |u|^{\frac{1}{n}}. \tag{3.17}$$

Table 3.3 shows the average values of such *simplified growth rates* for $n \leq 20$. The Maple program used here is similar to that given in Figure A.4 of Appendix A. If it were possible to characterize the occurrence of trace values $|u|$ for a given n, Equation 3.17 would become deterministic (instead of relying on calculation) and we may be able to find an exact expression for e^{γ_f}.

In Figure 3.2 we plot the values in Table 3.3 (circles) along with those in Figure 3.1 (asterisks). We can see that for each n, the points alternate between being greater than and less than the other, and both sets seem to tend to Viswanath's constant. The errors are plotted in Figure 3.3. We can attribute this discrepancy to the occurrence of small values of $|u|$ in the average. We have seen that in the linear case $|u| = 2$ and $\lambda_1 = 1$ and so the growth rate is raised slightly. Also when n is odd, $u = 0$

n	average of simplified g.r.
1	1
2	1.366025404
3	0.396850262
4	1.157340573
5	1.115253060
6	1.180264329
7	1.152601279
8	1.164282290
9	0.7066641139
10	1.143815487
11	1.134226076
12	1.148636161
13	1.14323635
14	1.145579314
15	0.8681194271
16	1.138800815
17	1.135386984
18	1.140510038
19	1.138864119
20	1.139510204

Table 3.3: Average values of simplified growth rates.

implies bounded growth by Theorem 3.3. Here, the approximation of $|\lambda_1|^{\frac{1}{n}} = 1$ is off by 1 from $|u|^{\frac{1}{n}}$, resulting in a much lower approximate average growth rate. We know from Theorem 3.2 that when $u = 0$ we must have $n \equiv 3 \pmod 6$. This explains the low growth rates for $n = 3, 9$ and 15 in Figure 3.2. We could remove some of the poorer approximations by considering only $n \equiv 1, 5 \pmod 6$ (the circles in Figure 3.1), which correspond to exponential growth, and still appear converge to Viswanath's constant. Also for n even we are ignoring the -4 term resulting in a slightly higher approximate average growth rate, and for n odd we are ignoring the $+4$ term, which slightly lowers the growth rate, which may explain the alternating larger value of circles and asterisks in Figure 3.2.

It may be a good idea to consider *balanced* coefficient cycles, i.e., cycles with an equal number of $+$ and $-$ signs. This may be a better approximation to a random Fibonacci sequence, because we know $+$ and $-$ occur with equal probability. Table 3.4

gives the average growth rate of balanced coefficient cycles, along with the previous values for all coefficient cycles given in Table 3.2. When n is odd we considered the *almost balanced* coefficient cycles, namely those where the number of $+$ and $-$ signs differs by one. The balanced growth rates appear to be a bit smaller, but do not seem to give a better estimate of Viswanath's constant.

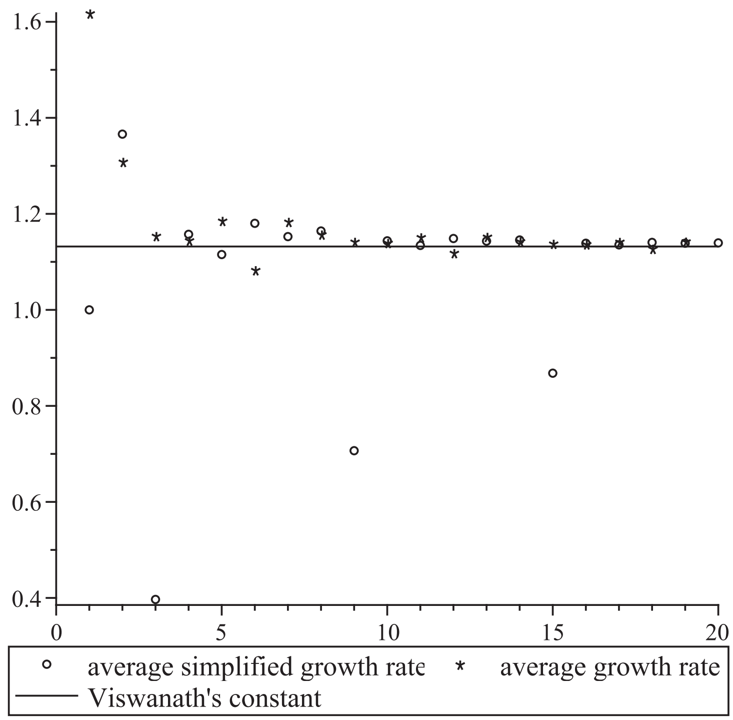


Figure 3.2: Comparing average values of growth rates and simplified growth rates.

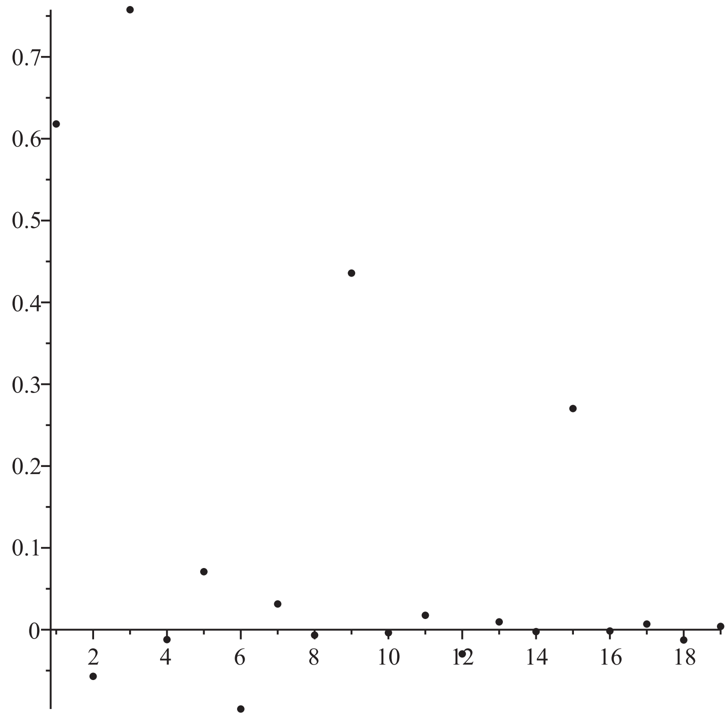


Figure 3.3: Error between average values of growth rates and simplified growth rates.

n	average growth rate	average for balanced cases
1	1.618033988	1.618033988
2	1.309016994	1
3	1.154508497	1
4	1.145259161	1.181346427
5	1.186117799	1.185472140
6	1.083345517	1
7	1.184039573	1.155972089
8	1.157659225	1.148606476
9	1.142422955	1.142604829
10	1.140061077	1.111498520
11	1.151816913	1.158370490
12	1.119038560	1.103502154

Table 3.4: Approximate average growth rate for all, and balanced coefficient cycles.

Chapter 4

Equivalence Classes

4.1 Continuant Polynomials

A useful tool in the study of periodic coefficient sequences is the *continuant polynomial* or simply *continuant*. It is closely connected to the continued fraction, hence the name. The information in this section can be found in Graham et al. [33, p. 301–305, p. 318]. Proofs of results are included to give a deeper understanding of this polynomial.

Definition 4.1. The *continuant*, denoted $K_n(x_1, \ldots, x_n)$, in the n variables $x_1, \ldots, x_n$ is defined recursively as

$$K_n(x_1, \ldots, x_n) = K_{n-1}(x_1, \ldots, x_{n-1})x_n + K_{n-2}(x_1, \ldots, x_{n-2}),$$

for $n \geq 2$, where $K_0() = 1$ and $K_1(x_1) = x_1$.

The next few polynomials are given by

$$\begin{aligned}
K_2(x_1, x_2) &= K_1(x_1)x_2 + K_0() = x_1x_2 + 1, \\
K_3(x_1, x_2, x_3) &= K_2(x_1, x_2)x_3 + K_1(x_1) = (x_1x_2 + 1)x_3 + x_1 = x_1x_2x_3 + x_1 + x_3, \\
K_4(x_1, x_2, x_3, x_4) &= K_3(x_1, x_2, x_3)x_4 + K_2(x_1, x_2) = (x_1x_2x_3 + x_1 + x_3)x_4 + x_1x_2 + 1 \\
&= x_1x_2x_3x_4 + x_1x_2 + x_1x_4 + x_3x_4 + 1.
\end{aligned}$$

We will soon see a method to evaluate continuants non-recursively. The following theorem gives the connection between continuant polynomials and our periodic coefficient sequences.

Theorem 4.1. *For any product matrix P_n, with $n \geq 2$, we have*

$$P_n = \begin{pmatrix} 0 & 1 \\ 1 & x_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & x_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & x_n \end{pmatrix} = \begin{pmatrix} K_{n-2}(x_2, \ldots, x_{n-1}) & K_{n-1}(x_2, \ldots, x_n) \\ K_{n-1}(x_1, \ldots, x_{n-1}) & K_n(x_1, \ldots, x_n) \end{pmatrix},$$

where $x_i \in \{\pm 1\}$ *and* K_i *is a continuant.*

Proof: The initial case, for $n = 2$, gives the product matrix

$$P_2 = \begin{pmatrix} 0 & 1 \\ 1 & x_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & x_2 \end{pmatrix} = \begin{pmatrix} 1 & x_2 \\ x_1 & x_1x_2 + 1 \end{pmatrix} = \begin{pmatrix} K_0() & K_1(x_2) \\ K_1(x_1) & K_2(x_1, x_2) \end{pmatrix},$$

which is the required matrix of continuants. Now suppose the theorem holds for n. For the case of $n+1$ we get the product matrix

$$\begin{aligned} P_{n+1} &= \begin{pmatrix} 0 & 1 \\ 1 & x_1 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & x_n \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & x_{n+1} \end{pmatrix} \\ &= \begin{pmatrix} K_{n-2}(x_2, \ldots, x_{n-1}) & K_{n-1}(x_2, \ldots, x_n) \\ K_{n-1}(x_1, \ldots, x_{n-1}) & K_n(x_1, \ldots, x_n) \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & x_{n+1} \end{pmatrix} \\ &= \begin{pmatrix} K_{n-1}(x_2, \ldots, x_n) & K_{n-2}(x_2, \ldots, x_{n-1}) + K_{n-1}(x_2, \ldots, x_n)x_{n+1} \\ K_n(x_1, \ldots, x_n) & K_{n-1}(x_1, \ldots, x_{n-1}) + K_n(x_1, \ldots, x_n)x_{n+1} \end{pmatrix} \\ &= \begin{pmatrix} K_{n-1}(x_2, \ldots, x_n) & K_n(x_2, \ldots, x_{n+1}) \\ K_n(x_1, \ldots, x_n) & K_{n+1}(x_1, \ldots, x_{n+1}) \end{pmatrix}, \end{aligned}$$

by Definition 4.1, as required. ◇

Note that this theorem is true for general values of x_i and not just $x_i \in \{1, -1\}$. Euler, who studied these polynomials deeply, found a way to write down any continuant without using the recursive definition. He noticed that given any product $x_1 \cdots x_n$, one can "strike out" all disjoint adjacent pairs $x_i x_{i+1}$ in all possible ways to obtain $K_n(x_1, \ldots, x_n)$. So, given $x_1 \cdots x_n$ we can strike out zero pairs, leaving $x_1 \cdots x_n$, we can strike out one adjacent pair $x_i x_{i+1}$, for $1 \le i \le n-1$, leaving $x_1 \cdots x_{i-1} x_{i+2} \cdots x_n$, We can strike out two adjacent pairs $x_i x_{i+1}$ and $x_j x_{j+1}$ for $1 \le i \le n-1$, $1 \le j \le n-1$, $x_j \ne x_{i-1}, x_i, x_{i+1}$, leaving $x_1 \cdots x_{i-1} x_{i+2} \cdots x_{j-1} x_{j+2} \cdots x_n$, and so on. We then sum these products of remaining terms to obtain a polynomial.

Example 4.1. Given $x_1x_2x_3x_4$, we could strike out zero pairs leaving the product unchanged, we could strike out one pair among x_1x_2, x_2x_3 and x_3x_4 leaving behind x_3x_4, x_1x_4 or x_1x_2 or we could strike out the two pairs x_1x_2 and x_3x_4 leaving behind a 1. Summing the possibilities gives us

$$K_4(x_1, x_2, x_3, x_4) = x_1x_2x_3x_4 + x_1x_2 + x_1x_4 + x_3x_4 + 1,$$

matching our previous definition of K_4.

We now need a proof of the general result.

Theorem 4.2. *The polynomial produced by striking out disjoint adjacent pairs in the product $x_1 \cdots x_n$ is the continuant $K_n(x_1, \ldots, x_n)$, for $n \geq 0$.*

Proof: Given the product $x_1 \cdots x_n$ we can strike out any combination of adjacent pairs, leaving us with a new product which either contains x_n or does not contain x_n. In the latter case we know that $x_{n-1}x_n$ has been struck out. The possible remaining combinations are then by induction given by $K_{n-2}(x_1, \ldots, x_{n-2})$. In the former case, x_n remains in the product and the possible remaining combinations excluding x_n are given by $K_{n-1}(x_1, \ldots, x_{n-1})$. Therefore the remaining combinations including x_n are given by $K_{n-1}(x_1, \ldots, x_{n-1})x_n$. Altogether we get that striking out adjacent pairs in all possible combinations gives $K_{n-1}(x_1, \ldots, x_{n-1})x_n + K_{n-2}(x_1, \ldots, x_{n-2})$, for $n \geq 2$ as required. For $n = 0$ or $n = 1$, we have either the empty product 1, or the single term x_1. In either case there are no pairs to strike out and we are left with $1 = K_0()$ or $x_1 = K_1(x_1)$ respectively. $\diamond$

One further thing we can note about continuants is their size.

Theorem 4.3. *The number of terms in the continuant $K_n(x_1, \ldots, x_n)$ is the Fibonacci number F_{n+1}, for $n \geq 0$. In other words, $K_n(1, \ldots, 1) = F_{n+1}$.*

Proof: This is a simple proof by strong induction. We know that $K_0() = 1$, which contains 1 term and $K_1(1) = 1 = F_2$. Suppose that $K_{n-2}(1, \ldots, 1) = F_{n-1}$ and $K_{n-1}(1, \ldots, 1) = F_n$. For n we have

$$\begin{aligned} K_n(1, \ldots, 1) &= K_{n-1}(1, \ldots, 1) \times 1 + K_{n-2}(1, \ldots, 1) \\ &= F_n + F_{n-1} = F_{n+1}, \end{aligned}$$

completing the proof. $\diamond$

4.2 Equivalence Classes

We know how to classify the growth of a given periodic coefficient sequence by looking at the trace and determinant (i.e., u and v values) of its associated product matrix, and we can determine the exact growth rate using the dominant eigenvalue. There are several things we can do to reduce the number of sequences we have to test in this way. According to certain properties of the sequences, we can group them into equivalence classes, where the elements in a class all have the same growth rate.

We have seen in Theorems 2.8 and 2.9 that reversing or rotating entries in a coefficient cycle does not change the characteristic equation. We alternately (and more simply) can prove these results using continuants.

Theorem 4.4. *Given a coefficient cycle $\sigma_n = (s_1, \ldots, s_n)$, with $n \geq 1$, reversing the terms gives a new cycle $\sigma'_n = (s_n, \ldots, s_1)$ with corresponding product matrix having the same characteristic equation as the original product matrix.*

Proof: From Theorem 4.1 we know that

$$u = K_n(x_1, \ldots, x_n) + K_{n-2}(x_2, \ldots, x_{n-1}),$$

for $n \geq 2$. If we reverse the terms in the product matrix, the u value becomes

$$u = K_n(x_n, \ldots, x_1) + K_{n-2}(x_{n-1}, \ldots, x_2).$$

But these expressions for u are equivalent because striking out adjacent pairs from the product $x_1 \cdots x_n$ leaves us with the same polynomial as striking pairs from $x_n \cdots x_1$, and similarly for $x_2 \cdots x_{n-1}$. Therefore the value of u remains the same, and because the cycle length n is unchanged, v also remains the same. In the $n = 1$ case, reversing the terms in the cycle $\sigma_1 = (s_1)$ leaves the cycle and hence the characteristic polynomial unchanged. ◇

Similarly we can prove the rotation theorem using continuants.

Theorem 4.5. *Given a coefficient cycle $\sigma_n = (s_1, \ldots, s_n)$, with $n \geq 1$, rotating the entries (to the right) by k, where $0 \leq k \leq n-1$, gives a new cycle $\sigma'_n = (s_{n-k+1}, \ldots, s_n, s_1, \ldots, s_{n-k})$ with corresponding product matrix having the same characteristic equation as the original product matrix.*

Proof: We know that $K_n(x_1, \ldots, x_n)$ gives all possible products that remain when we strike out adjacent pairs. Now consider $K_{n-2}(x_2, \ldots, x_{n-1})$, the other term in the expression for u, where $n \geq 2$. We can think of this as containing the sum of all possible products remaining after the non-adjacent pair $x_1 x_n$ has been removed from $x_1 \cdots x_n$. Therefore, if we think of the product $x_1 \cdots x_n$ as a loop (i.e., x_n is adjacent to x_1), u gives us the sum of all possible products remaining when we strike adjacent pairs from the loop. If we then rotate the entries by k terms, u remains unchanged. Again since the cycle length n is fixed, the characteristic equation is unchanged. In the case of $n = 1$, rotating the cycle $\sigma_1 = (s_1)$ has no effect, as in the previous theorem. Also, a rotation of 0 terms, which is equivalent to a rotation of n terms leaves the cycle unchanged. $\diamond$

In Theorems 4.4 and 4.5, we have shown that altering the coefficient cycle in a particular way does not change the characteristic equation. By Theorem 3.3, the values of u and v uniquely determine the growth type, and so it remains unchanged by altering the coefficient cycle. Also, Definition 2.3 tells us that the growth rate is dependent on the eigenvalues (which in turn depend on u and v), and the cycle length n. This tells us that the specific growth rate is also unchanged when we reverse or rotate the terms in our coefficient cycle.

There is another operation we can perform on our cycles without changing the growth rate. We will prove this using continuants, although it can be done (somewhat tediously) using induction on product matrices.

Theorem 4.6. *Given a coefficient cycle $\sigma_n = (s_1, \ldots, s_n)$, with $n \geq 1$, switching each term s_n from $+$ to $-$ or vice versa gives us a new cycle σ_n' with corresponding product matrix having the same growth rate as the original product matrix.*

Proof: Notice that in the expression $u = K_n(x_1, \ldots, x_n) + K_{n-2}(x_2, \ldots, x_{n-1})$, for $n \geq 2$, the indices of the continuants are of the same parity. Also notice that because we are striking out pairs, the number of x_i terms in the remaining products will have the same parity as n. If n is even, u will contain products of even numbers of x_i terms as well as two 1's, which are the result of striking out all pairs in each product. Switching the sign of each x_i therefore does not affect the value of u. If n is odd, u will contain products of odd numbers of x_i terms. This holds also for $n = 1$. Switching the sign of each x_i therefore switches the sign of u.

We know from Proposition 2.1 that for n odd or even, $v = 1$ or -1, respectively, and so switching signs does not change the value of v because the period length remains the same. Therefore when n is odd our eigenvalues, $\lambda_1, \lambda_2 = u \pm \frac{\sqrt{u^2+4v}}{2}$, switch sign, whereas for n even they are unchanged. In any case, the absolute value of the dominant eigenvalue, hence the growth rate, stays the same when we switch the $\pm$ signs. $\diamond$

There is one further thing we can do to group coefficient cycles with the same growth rate. We can simply look for repeating patterns of s_i terms and reduce our cycles to their primitive cycles, which are defined as follows.

Definition 4.2. Given a coefficient cycle $\sigma_n = (s_1, \ldots, s_n)$, where $(s_1, \ldots, s_n) = (s_1, \ldots, s_p, s_1, \ldots, s_p, \ldots, s_1, \ldots, s_p)$, with p the smallest possible such number, we say that σ_n has a *primitive coefficient cycle* or simply *primitive cycle* of length p, denoted by $\sigma_p = (s_1, \ldots, s_p)$.

The cycle $\sigma_6 = (+-+-+-)$ has primitive cycle $\sigma_2 = (+-)$ and both define the sequence $+-+-+-+-\ldots$. It is clear that σ_n and σ_p must give the same growth rate because they define the same sequence.

Example 4.2. Consider Example 2.6, in which we looked at the coefficient cycle $\sigma_4 = (++--)$ with product matrix $P_n = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 3 \end{smallmatrix}\right)$. We have characteristic equation $x^2 - 3x + 1 = 0$ and eigenvalues $\frac{3\pm\sqrt{5}}{2}$. The growth rate here is

$$\left(\frac{3+\sqrt{5}}{2}\right)^{\frac{1}{4}} = 1.272019\ldots.$$

If we consider the cycle $\sigma_8 = (++--++--)$, we have $P_8 = P_4^2 = \left(\begin{smallmatrix} -1 & 3 \\ -3 & 8 \end{smallmatrix}\right)$, which has characteristic equation $x^2 - 7x + 1$ and eigenvalues $\frac{7\pm\sqrt{45}}{2}$. In both cases the value of $|u|$ tells us we have exponential growth for n even. The growth rate here, which should be the same as that for P_4, is

$$\left(\frac{7+\sqrt{45}}{2}\right)^{\frac{1}{8}} = 1.272019\ldots.$$

There is no contradiction of growth rates!

We can now define a proper equivalence relation using the operations on cycles described in Theorems 4.4, 4.5 and 4.6, and Definition 4.2.

Definition 4.3. We write $\sigma_n \sim \tau_d$ if the cycle $\sigma_n = (s_1, \ldots, s_n)$, where $n \geq 1$, can be transformed into the cycle $\tau_d = (t_1, \ldots, t_d)$ by applying any finite combination of the following operations:

1. *reversal*: $(s_1, \ldots, s_n) \to (s_n, \ldots, s_1)$;
2. *rotation*: $(s_1, \ldots, s_n) \to (s_{k+1}, \ldots, s_n, s_1, \ldots, s_k)$;
3. *negation*: $(s_1, \ldots, s_n) \to (-s_1, \ldots, -s_n)$;
4. *period reduction* or *extension*: $(s_1, \ldots, s_n) \to (s_1, \ldots, s_d)$, where $p|d$, i.e., σ_n is reduced or extended by a multiple of its primitive cycle σ_p.

Note that only one of these operations involves changing the length of the coefficient cycle.

Theorem 4.7. *The relation defined in Definition 4.3 is in fact an equivalence relation.*

Proof: We must show that this relation is reflexive, symmetric and transitive for all cycles $\sigma_n = (s_1, \ldots, s_n)$, where $n \geq 1$. There are several ways we can show that our relation is reflexive, i.e., $\sigma_n \sim \sigma_n$. One is simply the trivial rotation, where we rotate σ_n by zero terms, or equivalently by n terms. In order for our relation to be symmetric we must show that if $\sigma_n \sim \tau_d$, then $\tau_d \sim \sigma_n$. If we list all of the operations required to transform σ_n to τ_d, we can transform τ_d back to σ_n by applying the inverses of these operations (finitely many) in reverse order. Reversal and negation are self-inverses and the inverse of rotation by k terms is a rotation of $n-k$ terms. Lastly, the inverse of reduction by a multiple of σ_p is extension by a multiple of σ_p and vice versa. Note that the inverses of the four operations again belong to the list of operations and so it is valid to use them. To show our relation is transitive, we first assume that $\sigma_n \sim \tau_d$ and $\tau_d \sim \rho_m$. It easily follows that $\sigma_n \sim \rho_m$ because starting with σ_n we can apply all operations, i.e., those used to take σ_n to τ_d and to take τ_d to ρ_m, in series to obtain ρ_m. $\diamond$

Note that in Theorems 2.9 and 4.5, rotation of a coefficient cycle σ_n took on the form $\sigma_n' = (s_{n-k+1}, \ldots, s_n, s_1, \ldots, s_{n-k})$, where we have rotated k terms to the right, whereas in Definition 4.3, we rotated k terms to the left (or equivalently $n-k$ terms to the right), to give $\sigma_n' = (s_1, \ldots, s_n) \to (s_{k+1}, \ldots, s_n, s_1, \ldots, s_k)$. The set of rotations is equivalent in either case, although the former suited the proof of Theorem 2.9 better and the latter has simpler notation. When calculating subsequences as in Example 2.6, however, we must rotate the coefficient cycle to the left, implying the product matrix (of the form $M_n \cdots M_1$) is rotated to the right.

The following useful fact about equivalence classes is one of our main results.

Corollary 4.1. *All elements in an equivalence class have the same growth rate.*

Proof: We have already seen that each of our four operations does not affect the growth rate, hence all elements in a particular equivalence class must have the same rate of growth. $\diamond$

4.3 Necklaces

There is an alternate way in which we can view our coefficient cycles σ_n. Instead of using a repeated coefficient cycle to generate a periodic coefficient sequence, we can picture it as a loop which is continually traversed. The following definition can be found in Graham et al. [33, p. 139].

Definition 4.4. A *necklace* is a string of n characters of q different types which is unchanged by rotation, i.e., two necklaces are equivalent if one can be transformed into the other by a rotation of k characters.

It is most common to speak of the characters as beads and the different types as colors. In our case, the $\pm$ signs in the coefficient cycles σ_n are represented by two colors. We can now generalize our cycles σ_n so that they contain elements s_i which can take on q different colors. We can view our new combinatorial object in light of the equivalence classes introduced in Definition 4.3. One natural question is — how many elements belong to each equivalence class for a given n? We will break this down in steps, starting with necklaces, which are equivalent under rotation, and then consider the other three operations given in Definition 4.3.

The key tool in enumerating our necklaces is the following combinatorial theorem due to Pólya [62].

Theorem 4.8 (Pólya Enumeration Theorem)**.** *Suppose H is a finite group of order h, of transformations t which act on a finite set of objects. Further, suppose that two objects are equivalent if one can be transformed into the other by some $t \in H$. Then the number of inequivalent objects in the set is given by*

$$T := \frac{1}{h} \sum_{t} I(t),$$

where $I(t)$ is the number of objects which are left invariant by transformation $t \in H$ and the sum is taken over all t.

The following theorem due to MacMahon (1892) enumerates our necklaces and can be found in Graham et al. [33, p. 140], for example. Recall that Euler's totient function, $\phi(n)$, denotes the number of positive integers $k \leq n$ such that $\gcd(k, n) = 1$. The proof of this theorem, as well as those of the remaining theorems in this chapter, are given for completion.

Theorem 4.9. *The number of distinct necklaces with n beads and q colours is given by*

$$N(n, q) := \frac{1}{q} \sum_{d|n} q^{\frac{n}{d}} \phi(d),$$

where $\phi(d)$ is Euler's totient function.

Proof: The set of rotations $(s_1, \ldots, s_n) \to (s_{n-k+1}, \ldots, s_n, s_1, \ldots, s_{n-k})$ for $0 \leq k \leq n-1$ (where a rotation of 0 terms is the same as a rotation of n terms) can be represented as the cyclic group C_n. We will denote an element of this group by $R^k \in C_n$, where R stands for rotation to the right. In order to apply Pólya's Enumeration Theorem, we must find $I(R^k)$, i.e., the number of cycles which are left invariant by a general element $R^k \in C_n$. Let us instead consider a general element R^{n-k}. Suppose the cycle $\sigma_n = (s_1, \ldots, s_n)$ is left invariant by this element, i.e.,

$$(s_1, \ldots, s_n) = (s_{k+1}, \ldots, s_n, s_1, \ldots, s_k),$$

This requires the following equalities:

$$s_1 = s_{k+1},\ \ s_2 = s_{k+2},\ \ \ldots,\ \ s_k = s_{2k},\ \ s_{k+1} = s_{2k+1},\ \ \ldots,\ \ s_n = s_k.$$

Notice that we have $s_1 = s_{k+1} = s_{2k+1} = \cdots = s_{(lk+1) \bmod n}$ for $l \in \mathbb{N}$, and in general $s_j = s_{(lk+j) \bmod n}$ for $j = 1, \ldots, n$. We are using the convention that $n \pmod n \equiv n$.

It is known that given any $k \in \mathbb{N}$, $\{lk \pmod n \mid l \in \mathbb{N}\} \equiv \{0, d, 2d, \ldots, n-d \mid d = \gcd(k,n)\}$ (not using the mentioned convention). Therefore we must have

$$s_j = s_{j+ld},\ \ ld \in \{0, d, 2d, \ldots, n-d\}, \tag{4.1}$$

for $j = 1, \ldots, d$. In other words, if we think about the group C_n as a subgroup of the symmetric group of permutations, S_n, every d^{th} term in σ_n belongs to the same cycle in the permutation, so that we have d cycles of length $\frac{n}{d}$.

What we can now infer is that the colours of $s_1, s_2, \ldots, s_d$ are each chosen independently from a set of size q, and the colors of $s_{d+1}, \ldots, s_n$ are determined by Equation (4.1). In other words, since we want our coefficient cycle σ_n to be invariant under rotation, each element in a given cycle of the rotation permutation must have the same color. Therefore we have q^d ways to choose d colors, i.e., q^d coefficient cycles which are equal to themselves when rotated by k and thus $I(R^{n-k}) = q^d$. Note also that we must have d as the length of our primitive cycle. Applying Pólya's Theorem, and using the fact that $|C_n| = n$, we conclude that

$$N(n,q) = \frac{1}{n} \sum_{R^{n-k}} q^d = \frac{1}{n} \sum_{R^k} q^d. \tag{4.2}$$

Now because $d = \gcd(k,n)$, we have that $d|n$. We can split the above sum according to the value of d for each $0 \le k < n$ to give

$$N(n,q) = \frac{1}{n} \sum_{d|n} q^d \sum_{\substack{0 \le k < n \\ \gcd(k,n)=d}} 1.$$

The first sum corresponds to each possible value of d, while the second corresponds to the number of times each d-value arises. By noting that $d = \gcd(k,n)$ if and only if $\gcd(\frac{k}{d}, \frac{n}{d}) = 1$, we can rewrite the previous sum as

$$N(n,q) = \frac{1}{n} \sum_{d|n} q^d \sum_{\substack{0 \le k < n \\ \gcd(\frac{k}{d}, \frac{n}{d})=1}} 1.$$

Now writing $k = jd$ (we know $d|k$), our expression becomes

$$N(n,q) = \frac{1}{n} \sum_{d|n} q^d \sum_{\substack{0 \le j < \frac{n}{d} \\ \gcd(j, \frac{n}{d})=1}} 1.$$

The second sum is by definition $\phi\left(\frac{n}{d}\right)$, so we obtain

$$N(n,q) = \frac{1}{n} \sum_{d|n} q^d \phi\left(\frac{n}{d}\right) = \frac{1}{n} \sum_{d|n} q^{\frac{n}{d}} \phi(d), \tag{4.3}$$

since d and $\frac{n}{d}$ give the same list of divisors of n, completing the proof. $\diamond$

When $q = 2$, the following sequence (see [67, A000031]) gives the number of necklaces of two colors, for increasing $n \geq 0$:

$$1, 2, 3, 4, 6, 8, 14, 20, 36, 60, 108, 188, 352, \ldots.$$

Example 4.3. Figure 4.1 lists all necklaces with 4 beads and 2 colors. From the above sequence we see that $N(4,2) = 6$. Notice that the first two pairs of necklaces

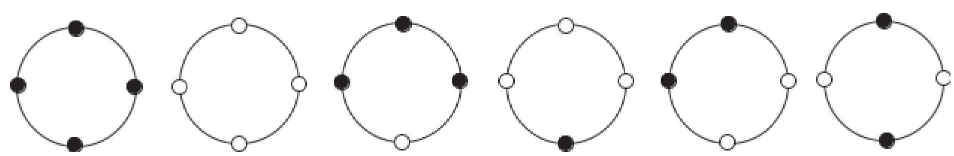


Figure 4.1: Necklaces with $n = 4$, $q = 2$.

have opposite colors. If we reversed the colors on either of the last two necklaces, they would remain unchanged under rotation. For example, Figure 4.2 gives two equivalent necklaces.

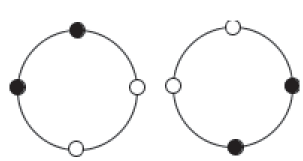


Figure 4.2: Equivalent necklaces.

Next we will apply the operation of negation, which is color swapping in this case, to our necklaces. This means that swapping colors will produce a new necklace which is equivalent to the old one. We will denote the number of such necklaces by $N_s(n, q)$. Also, we will need to make use of the symmetric group on q elements, S_q. The following theorem is found in Gilbert and Riordan [32].

Theorem 4.10. *The number of distinct necklaces with n beads and q colors, where color swapping is allowed, is given by*

$$N_s(n, q) := \frac{1}{q!n} \sum_{d,P} \phi(d) N(c_1, \ldots, c_q)(m(d))^{\frac{n}{d}},$$

where the sum is taken over all divisors d of n and all partitions P of q. Here $m(d) := \sum_{j|d} jc_j$, and $N(c_1, \ldots, c_q)$ is the number of permutations of $\pi \in S_q$ which have c_j cycles of length j, where $j = 1, \ldots, q$.

Proof: As we have seen in the proof of Theorem 4.9, the set of rotations can be represented as the cyclic group C_n. Given q colors, the set of permutations (possible swappings of colors) can be represented by the symmetric group S_q. As given in Gilbert and Riordan [32], the set of transformations of cycles which may arise from rotation, color swapping or both may be represented by the direct product $C_n \times S_q$. We will denote an element of this group by $R^k\pi$, where $R^k \in C_n$ and $\pi \in S_q$. We have defined a group action, where an element $R^k\pi$ acts on a cycle $\sigma_n = (s_1, \ldots, s_n)$ by permuting the colors, π, and rotating the terms, R^k. The former is a permutation of the set of q colors, where the latter is a permutation of the n elements of the cycle σ_n. These operations are independent of order because they are applied on two different sets.

In order to apply Pólya's Theorem, we must find $I(R^k\pi)$, i.e., the number of cycles which are left invariant by a general element $R^k\pi \in C_n \times S_q$. As in Theorem 4.9, suppose the cycle $\sigma_n = (s_1, \ldots, s_n)$ is left invariant by $R^{n-k}\pi$. Then for all i,

$$s_i = \pi s_{i+k} = \pi^2 s_{i+2k} = \pi^3 s_{i+3k} = \cdots \tag{4.4}$$

Again the colors of terms s_i are independently chosen for $i = 1, \ldots, d$, where $d = \gcd(n, k)$ and are determined by Equation (4.4) for $i > d$, and where $\{lk \pmod n) | l \in$

$\mathbb{Z}\} \equiv \{0, d, 2d, \ldots, n-d\}$. However, in the case of $C_n \times S_q$ we have a further restriction imposed by S_q, which we will now explain.

We can write the above set as $\{\, ld \,|\, 0 \leq l \leq \frac{n}{d} - 1\}$ and use it to replace the lk terms in Equation (4.4). Continuing the terms we get

$$s_i = \pi s_{i+d} = \pi^2 s_{i+2d} = \cdots = \pi^{\frac{n}{d}-1} s_{i+(\frac{n}{d}-1)d} = \pi^{\frac{n}{d}} s_{i+\frac{nd}{d}},$$

where we have included the next l-value, $\frac{n}{d}$. (We did not include n in the set $\{ld\}$ because it is equivalent to zero modulo n.) We then have

$$s_i = \pi^{\frac{n}{d}} s_{i+\frac{nd}{d}} = \pi^{\frac{n}{d}} s_i \pmod{n}.$$

Therefore $\pi^{\frac{n}{d}}$ preserves s_i for $i = 1, \ldots, n$. Recall that the permutation R^{n-k} has d cycles of length $\frac{n}{d}$, so repeating the rotation $\frac{n}{d}$ times is equivalent to applying the identity element in our group C_n. We can deduce that since $\pi^{\frac{n}{d}} = 1$, the value of s_i, i.e., its color, has order dividing $\frac{n}{d}$ for $i = 1, \ldots, n$. In other words, the value of s_i belongs to a cycle of π which has length dividing $\frac{n}{d}$, because the order of an element is equal to the length of its cycle in the decomposition of $\pi \in S_q$. If π decomposes into c_j cycles of length j for $j = 1, \ldots, q$, each $s_i \in \sigma_n$ is restricted to the colors (values $1, \ldots, q$) which belong to cycles having lengths j, dividing $\frac{n}{d}$, so that $\pi^j s_i = s_i$. Therefore s_i can take on

$$m\left(\frac{n}{d}\right) = \sum_{j|\frac{n}{d}} j c_j \tag{4.5}$$

different colors, since these colors are unchanged by the permutation π. We have seen that we have choice of color for the first d terms $(s_1, \ldots, s_d)$ in σ_n, i.e., each of the d cycles in R^{n-k} contains elements of only one color. Therefore, combining our information, we see that $I(R^{n-k}\pi)$, i.e., the number of coefficient cycles σ_n that remain invariant under the permutation $R^{n-k}\pi$, is given by

$$I(R^{n-k}\pi) = (m(\tfrac{n}{d}))^d.$$

Note that $I(R^{n-k}\pi) = I(R^k\pi)$ because $\gcd(n, k) = \gcd(n, n-k) = d$.

We can now apply Pólya's Theorem, which tells us that the number of equivalence classes of sequences under rotation and color swapping is given by

$$N_s(n,q) = \frac{1}{q!n}\sum_{R^k\pi} I(R^k\pi) = \frac{1}{q!n}\sum_{R^k\pi}\left(m\left(\frac{n}{d}\right)\right)^d. \tag{4.6}$$

Here $q!n$ is the order of the direct product $C_n \times S_q$ and the sums are taken over all elements in the group. Summing over $C_n \times S_q$ means our terms are dependent on k and π. We can simplify this formula by combining terms which have the same d-value. We can see from Equation (4.2) and Equation (4.3) in the proof of Theorem 4.9 that

$$\sum_{0\le k\le n} q^d = \sum_{d|n} q^d \phi\left(\frac{n}{d}\right). \tag{4.7}$$

We can apply the same argument to Equation (4.6), summing over d instead of k and multiplying our sum by a factor of $\phi\left(\frac{n}{d}\right)$.

Now, of the $q!$ elements $\pi \in S_q$, we can split them into sets according to partitions P of q given by $c_1 + 2c_2 + \cdots + qc_q = q$, where each partition corresponds to elements π with c_j cycles of length j. It is a well known fact in group theory that two elements in the symmetric group S_q are conjugate (i.e., for $a, b \in S_q$ there exists $x \in S_q$ such that $xax^{-1} = b$) if and only if they consist of the same number of disjoint cycles of the same length, i.e., they are distinct permutations of the same partition of q. Recall that conjugacy classes (here, partitions) are equivalence classes induced by conjugacy of elements. We will denote the size of the conjugacy class associated with a given partition by $N(c_1, \ldots, c_q)$. To obtain an expression for this number (as shown in Bóna [10, p. 80]), start by considering all $q!$ possible permutations of q elements. Now choose a partition, where the cycles are arranged in some fixed order (for example, length 2, length 3, length 3) and think about placing this on top of each of the $q!$ permutations. Each cycle of length j can be rotated j different ways to give j equivalent cycles. Since there are c_j cycles of length j, we can divide $q!$ by j^{c_j} for each value of j, $j = 1, \ldots, q$, to remove these repetitions. Also, the c_j cycles of length j can be arranged in $c_j!$ different ways, without changing the fixed structure of the cycle. Therefore we can also divide $q!$ by $c_j!$ for each value of j giving the number of distinct permutations of a partition (by distinct we mean different elements in S_q), or in other words, the size of the conjugacy class, as

$$N(c_1, \ldots, c_q) = \frac{q!}{c_1! \cdot 2^{c_2} c_2! \cdots q^{c_q} c_q!}. \tag{4.8}$$

Now instead of summing over permutations π, we will sum over partitions P and multiply by a factor of $N(c_1, \ldots, c_q)$. Equation (4.6) can now be written as

$$N_s(n, q) = \frac{1}{q!n} \sum_{d,P} \phi\left(\frac{n}{d}\right) N(c_1, \ldots, c_q) \left(m\left(\frac{n}{d}\right)\right)^d.$$

Finally, because $\{d\}$ and $\{\frac{n}{d}\}$ give the same set of divisors of n, we can rewrite the above as

$$N_s(n, q) = \frac{1}{q!n} \sum_{d,P} \phi(d) N(c_1, \ldots, c_q)(m(d))^{\frac{n}{d}},$$

completing the proof. ◇

We can now consider the special case of $q = 2$, which is the case we are interested in. This can be found in Fine [27].

Corollary 4.2. *The number of distinct necklaces with n beads and two colors, which may be swapped, is given by*

$$N_s(n, 2) = \sum_{d|n} \frac{\phi(2d) 2^{\frac{n}{d}}}{2n}.$$

Proof: Theorem 4.10 tells us that

$$N_s(n, 2) = \frac{1}{2n} \sum_{d,P} \phi(d) N(c_1, c_2)(m(d))^{\frac{n}{d}}.$$

Now, how many partitions $c_1 + 2c_2 = 2$ are there in S_2 with c_j cycles of length j? We can have one cycle of length 2 so that $c_1 = 0$ and $c_2 = 1$, e.g., $\pi_1 = (12)$, or two cycles of length 1 so that $c_1 = 2$ and $c_2 = 0$, e.g., $\pi_2 = (1)(2)$. Now we need to find $N(c_1, c_2)$, i.e., the number of permutations of a given partition. For π_1 we have

$$N(0, 1) = \frac{2!}{0!1!2^1} = 1,$$

and for π_2 we have

$$N(2, 0) = \frac{2!}{2!0!2^0} = 1.$$

It is clear from the above example that there is only one permutation of each partition. We can now split up $N_s(n, 2)$ according to our two partitions as

$$N_s(n, 2) = \frac{1}{2n} \left(\sum_{d,\ \pi_1} \phi(d)(m(d))^{\frac{n}{d}} + \sum_{d,\ \pi_2} \phi(d)(m(d))^{\frac{n}{d}} \right), \tag{4.9}$$

where the first and second sums correspond to π_1 and π_2 respectively, and each sum is over all divisors d of n. Now consider the term $m(d) = \sum_{j|d} jc_j$. This places the restriction that the color of each $s_i \in \sigma_n$ belongs to a cycle of π which has length dividing d, for a given permutation π. Therefore in the case of π_1 we must have $2|d$ and in π_2 we must have $1|d$. This tells us that the first sum in Equation (4.9) must have d even and the second can have d even or odd. For π_1 we must have $m(d) = 1(0) + 2(1) = 2$ because j can take on the value 1 or 2. For π_2, if d is even, j can again take on the value 1 or 2 and $m(d) = 1(2) + 2(0) = 2$. If d is odd, j can only take on the value 1 and so $m(d) = 1(2) = 2$.

Rewriting the sums in Equation (4.9) according to the parity of d gives

$$N_s(n, 2) = \frac{1}{2n} \left(\sum_{d \text{ even}} 2\phi(d) 2^{\frac{n}{d}} + \sum_{d \text{ odd}} \phi(d) 2^{\frac{n}{d}} \right).$$

We can now use the facts that for d even we have $\phi(2d) = 2\phi(d)$ and for d odd, $\phi(2d) = \phi(d)$, to give

$$N_s(n, 2) = \frac{1}{2n} \left(\sum_{d \text{ even}} \phi(2d) 2^{\frac{n}{d}} + \sum_{d \text{ odd}} \phi(2d) 2^{\frac{n}{d}} \right).$$

Combining these terms gives

$$N_s(n, 2) = \frac{1}{2n} \sum_{d|n} \phi(2d) 2^{\frac{n}{d}},$$

which completes the proof. $\diamond$

The sequence $N_s(n, 2)$, and corresponding formula, can be found in [67, A000013]. The first few terms for $n \geq 0$ are

$$1, 1, 2, 2, 4, 4, 8, 10, 20, 30, 56, 94, 180, \ldots . \tag{4.10}$$

Example 4.4. If we introduce equivalence under the operation of color swapping to the necklaces given in Example 4.3, our set of necklaces is reduced to that in Figure 4.3. From the above sequence we see that $N_s(4, 2) = 4$.

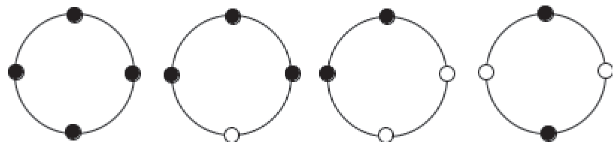


Figure 4.3: Necklaces with color swapping for $n = 4$, $q = 2$.

4.4 Bracelets

We can extend the definition of the necklace to form another combinatorial object, the bracelet.

Definition 4.5. A *bracelet* is a string of n characters of q different types which is unchanged by rotation and reversal, i.e., two bracelets are equivalent if one can be transformed into the other by a combination of rotation of k characters and reversal of character order.

This term is more recent, and the papers referenced earlier, namely Gilbert and Riordan [32] and Fine [27], do not use it. Gilbert and Riordan refer to a "mirror image necklace". Note that a bracelet is simply a necklace with the additional property that we can pick it up and flip it over in the plane, which is equivalent to reversing the order of the characters. The combined set of rotations $(s_1, \ldots, s_n) \to (s_{n-k+1}, \ldots, s_n, s_1, \ldots, s_{n-k})$ and reversals $(s_1, \ldots, s_n) \to (s_n, \ldots, s_1)$ for $k = 0, \ldots, n-1$ can be represented by the dihedral group D_{2n}. If we rotate by k terms and then reverse the cycle, our operation is

$$(s_1, \ldots, s_n) \to (s_{n-k}, \ldots, s_1, s_n, \ldots, s_{n-k+1}). \tag{4.11}$$

For each of the n rotations, we can either reverse the cycle or leave it unchanged; this gives us a total of $2n$ different permutations. Let $F^t R^k$ be a typical element of D_{2n} where $t \in \{0, 1\}$ (we reduce its value modulo 2). The rightmost operation, the rotation, is applied first and our group operation is concatenation of elements. Note that the permutations F^t (F for flip) and R^k are not commutative. If we reverse the cycle $(s_1, \ldots, s_n)$ and then rotate by k terms we obtain

$$(s_1, \ldots, s_n) \to (s_k, \ldots, s_1, s_n, \ldots, s_{k+1}),$$

which is not the same cycle as that given in the expression in (4.11). We do have that $F^sF^t = F^{s+t}$ and $R^jR^k = R^{j+k}$. We will use the convention that $F^0 = 1$, so that $F^0R^k = R^k$.

It will be useful to describe the conjugacy classes of elements in the group D_{2n}. Before we do this, a few notes on conjugacy class and cycle type will be useful. Note that $D_{2n} \leq S_n$ for $n \geq 3$. Every element of D_{2n} can be viewed as a permutation of n elements, and these permutations can be represented as cycles, in S_n, as we will see. We mentioned in Theorem 4.10 that two elements in S_n are conjugate if and only if they consist of the same number of disjoint cycles of the same length, i.e., they have the same cycle type. If elements a, b have the property that $a \sim b$ in D_{2n}, then $a \sim b$ in S_n as well. Therefore a and b have the same cycle type. But the converse is not necessarily true, i.e., if a and b have the same cycle type, then it is not necessarily true that $a \sim b$ in D_{2n}. The following Lemma can be found in Riordan [63, p. 149].

Lemma 4.1. *The conjugacy classes of D_{2n} for n even are given by*

$$\{1\}, \{R^{\frac{n}{2}}\}, \{R, R^{n-1}\}, \{R^2, R^{n-2}\}, \ldots, \{R^{\frac{n-2}{2}}, R^{\frac{n+2}{2}}\},$$
$$\left\{FR^{2k+1} \mid k = 0, \ldots, \frac{n-2}{2}\right\}, \left\{FR^{2k} \mid k = 0, \ldots, \frac{n-2}{2}\right\}.$$

For n odd, the conjugacy classes are

$$\{1\}, \{R, R^{n-1}\}, \{R^2, R^{n-2}\}, \ldots, \{R^{\frac{n-1}{2}}, R^{\frac{n+1}{2}}\}, \{FR^k \mid k = 0, \ldots, n-1\}.$$

Proof: First suppose we have a cycle $\sigma_n = (s_1, \ldots, s_n)$ to which we apply the permutation FRF, i.e., we flip, rotate, then flip again. This gives

$$(s_1, \ldots, s_n) \to (s_n, \ldots, s_1) \to (s_1, s_n, \ldots, s_2) \to (s_2, \ldots, s_n, s_1). \tag{4.12}$$

The permutation (as an element of S_n) can be represented as $(1 \;\; n \;\; n-1 \;\; \ldots \;\; 2)$ because s_1 moves from position 1 to position n, s_n moves from position n to position $n-1$ and so on. This permutation is in fact R^{n-1}. We have therefore proven the relation $FRF = R^{n-1}$. Because $F^2 = 1$, we know that $F = F^{-1}$, and thus R and R^{n-1} are conjugate elements in D_{2n}. By iterating, it is easy to see that $FR^kF = R^{n-k}$ for any rotation of size k. This tells us that R^k and R^{n-k} are also conjugate in D_{2n}, and in fact they are inverses of each other, since if we apply one after the other we obtain $R^n = 1$. We can use this result to prove another conjugacy relation.

Consider the permutation $R^{n-k}(FR^s)R^k$ for some integer s, and use the just-mentioned conjugacy in the form $R^{n-k}F = FR^k$ to give

$$R^{n-k}(FR^s)R^k = (R^{n-k}F)R^{s+k} = (FR^k)R^{s+k} = FR^{s+2k}.$$

This tells us that FR^s and FR^{s+2k} are conjugates for any integral values of s and k. Using this we will be able to deduce the conjugacy classes for elements of the form $FR^k \in D_{2n}$.

Consider the case of n even. For $s = 0$ we have that F is conjugate to FR^{2k} for all $k = 0, \ldots, \frac{n-2}{2}$ (note that $FR^n = F$). We have that the permutation F can be represented by

$$F = (1 \;\; n)(2 \;\; n-1)\cdots\left(\frac{n}{2} \;\; \frac{n}{2}+1\right). \tag{4.13}$$

We can therefore deduce that all elements of the form FR^{2k} for n even have $\frac{n}{2}$ cycles of length 2. For $s = 1$ we have that FR is conjugate to FR^{2k+1} for all $k = 0, \ldots, \frac{n-2}{2}$. Applying the permutation FR to the cycle $\sigma_n(s_1, \ldots, s_n)$ (starting with the rightmost operation) gives

$$(s_1, \ldots, s_n) \to (s_n, s_1, \ldots, s_{n-1}) \to (s_{n-1}, \ldots, s_1, s_n). \tag{4.14}$$

This permutation can be represented as

$$FR = (n)\left(\frac{n}{2}\right)(1 \;\; n-1)(2 \;\; n-2)\cdots\left(\frac{n}{2}-1 \;\; \frac{n}{2}+1\right). \tag{4.15}$$

Here we have, for example, that s_1 moves from position 1 to position $n-1$, and s_{n-1} moves from position $n-1$ to position 1. We can therefore deduce that all elements of the form FR^{2k+1} for n even have $\frac{n-2}{2}$ cycles of length 2 and two cycles of length 1. Note that because these two conjugacy classes have different cycle types, they remain distinct. (By the contrapositive of the fact that conjugacy implies same cycle type.)

Now consider the case for n odd. For $s = 0$ we again have that F is conjugate to FR^{2k}, except that now this holds for all $k = 0, \ldots, \frac{n-1}{2}$. We have that

$$F = \left(\frac{n+1}{2}\right)(1 \;\; n)(2 \;\; n-1)\cdots\left(\frac{n-1}{2} \;\; \frac{n+3}{2}\right). \tag{4.16}$$

For $s = 1$ we have that FR is conjugate to FR^{2k+1} for all $k = 0, \ldots, \frac{n-3}{2}$. Applying the permutation FR to the cycle $(s_1, \ldots, s_n)$ again gives the expression in (4.14), which can be represented as

$$FR = (n)(1 \quad n-1)(2 \quad n-2)\cdots\begin{pmatrix}\frac{n-1}{2} & \frac{n+1}{2}\end{pmatrix}. \tag{4.17}$$

Note that regardless of the parity of s, when n is odd we have $\frac{n-1}{2}$ cycles of length 2, and one cycle of length 1. We can show that these two sets of elements $\{FR^{2k}\}$ and $\{FR^{2k+1}\}$ in fact belong to the same conjugacy class, but cannot assume this based on the fact that they have the same cycle type. Consider the element $F \in D_{2n}$ and multiply R^k and its inverse on either side as follows:

$$(R^kF)R^{n-k} = (FR^{n-k})R^{n-k} = FR^{2n-2k} = FR^{n-2k},$$

which tells us that $F \sim FR^{n-2k}$. For n odd, $n - 2k$ is also odd and so $F \sim FR^{2k+1}$. We know that $F \sim FR^{2k}$, and by transitivity of conjugation we can conclude that all elements of the form FR^k are conjugate for $k = 0, \ldots, n-1$.

This takes care of the conjugacy classes of elements of D_{2n} which are of the form R^kF^t where $t = 1$. What about the case of $t = 0$? Here we want the conjugacy classes of the group C_n. We have seen above that $R^k \sim R^{n-k}$. From this we can deduce that for n odd we have the sets of conjugate elements

$$\{1\}, \{R, R^{n-1}\}, \{R^2, R^{n-2}\}, \ldots, \{R^{\frac{n-1}{2}}, R^{\frac{n+1}{2}}\},$$

and for n even the sets are

$$\{1\}, \{R^{\frac{n}{2}}\}, \{R, R^{n-1}\}, \{R^2, R^{n-2}\}, \ldots, \{R^{\frac{n-2}{2}}, R^{\frac{n+2}{2}}\}.$$

We cannot conclude that these sets represent conjugacy classes in D_{2n} until we show that no other conjugacies exist among the sets. Consider the element R^k, and suppose we take a general element $F^tR^j \in D_{2n}$ to test for further elements conjugate to R^k. Using the fact that $(F^tR^j)^{-1} = R^{n-j}F^t$ we get

$$(F^tR^j)R^k(R^{n-j}F^t) = F^tR^{n+k}F^t = F^tR^kF^t.$$

If $t = 0$ we get no conjugacy, and if $t = 1$, a previous relation tells us that $FR^kF = R^{n-k}$. This tells us that $R^k \sim R^{n-k}$. We can now conclude that there are no other elements conjugate to R^k and therefore the sets listed above do in fact represent conjugacy classes in D_{2n}. $\diamond$

We can now generalize Theorem 4.10 to count bracelets rather than necklaces. A general form of the following theorem can be found in Gilbert and Riordan [32].

Theorem 4.11. *The number of distinct bracelets with n beads and q colors, where color swapping is allowed is given by*

$$B_s(n,q) := \frac{1}{2nq!} \sum_{P,P'} M(e_1,\dots,e_n) N(c_1,\dots,c_q) \prod_{l=1}^{n} (m(l))^{e_l},$$

where the sum is taken over all partitions P and P' which are given as $c_1+2c_2+\cdots+qc_q = q$ and $e_1+2e_2+\cdots+ne_n = n$, respectively. Here $m(l) = \sum_{j|l} jc_j$, $N(c_1,\dots,c_q)$ is the number of permutations $\pi \in S_q$ having c_j cycles of length j, and $M(e_1,\dots,e_n)$ is the number of permutations $F^tR^k \in D_{2n}$ having e_l cycles of length l.

Proof: Recall that D_{2n} represents the set of permutations of cycles arising from rotation and reversal. The set of permutations of q colors is represented by S_q. The set of transformations of cycles which may arise from any combination of rotation, reversal or color swapping may be represented by the direct product $D_{2n} \times S_q$. Let $F^tR^k\pi$ be a typical element of $D_{2n} \times S_q$, where $t \in \{0,1\}$, $k \in \{0,\dots,n-1\}$ and $\pi \in S_q$. Keep in mind that the flips F^t and rotations R^k do not commute, but the color swaps π do commute with elements F^tR^k.

To apply Pólya's Theorem we must first find $I(F^tR^k\pi)$, i.e., the number of cycles σ_n (where we are thinking of cycles as loops) which are left invariant by a general element $F^tR^k\pi \in D_{2n} \times S_q$. Suppose the permutation F^tR^k, considered as an element of S_n, decomposes into e_l cycles of length l. As before, in order for a cycle σ_n to be invariant under such a permutation, we will need each cycle in the decomposition to contain only one color. Now, as in Theorem 4.10 we will investigate the additional restraint imposed by color swapping.

Consider an element $s_i \in \sigma_n$ in a length-l cycle of some permutation F^tR^k. For example, when n is even, the permutation F is represented by Equation (4.13). The element $s_1 \in \sigma_n$ is in position 1 and so belongs to the cycle $(1\, n)$, which is a length-2 cycle of the permutation F. Applying the permutation F^tR^k to the element s_i, l times, gives $(F^tR^k)^l(s_i) = s_i$, because l is the order of the cycle to which s_i belongs. This tells us that s_i is invariant under $(F^tR^k)^l$. If we attach color swapping to our permutation, we also want to have $\pi^l(s_i) = s_i$. (Note that $(F^tR^k\pi)^l = (F^tR^k)^l\pi^l$ because color swapping commutes with the other operations.) As in the proof of Theorem 4.10, this tells us that the color of s_i belongs to one of the c_j cycles of length j, of $\pi \in S_q$ such that $j|l$. Therefore, the number of possible colors the element s_i can take on is given by Equation (4.5):

$$m(l) = \sum_{j|l} jc_j.$$

Since there are e_l cycles (containing elements of the same color) of length l in D_{2n} we must choose one of $m(l)$ colors e_l times and so taking the product over all $l = 1, \dots, n$ gives

$$I(F^tR^k\pi) = (m(1))^{e_1}(m(2))^{e_2} \cdots (m(n))^{e_n} = \prod_l (m(l))^{e_l}.$$

Here we make the convention that if $e_l = 0$ then $m(l)^{e_l} = 1$ for all values of $m(l)$, including zero. By Pólya's Theorem, we have that the number of inequivalent bracelets where swapping color is allowed is given by

$$B_s(n,q) = \frac{1}{2nq!} \sum_{F^tR^k\pi} \prod_l (m(l))^{e_l}, \tag{4.18}$$

where we have used the fact that $|D_n \times S_q| = 2nq!$.

Now, instead of summing over elements $F^tR^k\pi$, i.e., permutations $\pi \in S_q$ and $F^tR^k \in D_{2n}$, we can sum over partitions of these permutations. This means we can split up our sum according to the partitions of q and n. A partition P of q given by $c_1 + 2c_2 + \cdots + qc_q = q$ corresponds to elements π with c_j cycles of length j, and the number of permutations it has (i.e., the size of its conjugacy class) is given in Equation (4.8) as

$$N(c_1, \dots, c_q) = \frac{q!}{c_1! \cdot 2^{c_2}c_2! \cdots q^{c_q}c_q!}. \tag{4.19}$$

Similarly, a partition P' of n given by $e_1 + 2e_2 + \cdots + ne_n = n$ corresponds to the elements F^tR^k with e_l cycles of length l, and the number of permutations it has in D_{2n}

is denoted by $M(e_1, \ldots, e_n)$. We can again make the connection to sizes of conjugacy classes.

We have already noted that if elements a, b have the property that $a \sim b$ in D_{2n}, then a and b have the same cycle structure but the converse is not necessarily true. This means that it is not necessary for the conjugacy classes in D_{2n} to include all elements with a given cycle structure. It is possible that two elements have the same cycle structure and are conjugate in S_n but not in D_{2n}. Therefore we must check if any of the conjugacy classes in D_{2n} can be combined to give a larger group of elements all having the same cycle structure.

We have seen in Lemma 4.1 that elements of the form $R^kF \in D_{2n}$ belong in one or two conjugacy classes depending on the parity of n. We have already determined the cycle types, which are given by Equations (4.13), (4.15) and (4.16). Therefore we cannot group these any further. What about the cycle structures of elements of the type $R^k \in D_{2n}$? Suppose we apply a rotation R^k to a cycle σ_n. If we let $d = \gcd(n, k)$, we have that the permutation R^k can be represented as d cycles of length $\frac{n}{d}$. We can now group elements R^k according to their associated value of d. The number of permutations R^k for which $\gcd(n, k) = d$ is precisely $\phi\left(\frac{n}{d}\right)$. (This was shown in Equations (4.2) and (4.3) and the explanation in between.) We can conclude that for each $d|n$, each element in the set

$$\{R^k \mid d = \gcd(n, k)\},$$

has the same cycle structure, and this structure differs for each set. The size of each set is $\phi\left(\frac{n}{d}\right)$. We know from Lemma 4.1 that R^k and R^{n-k} are conjugate. This fits with what we have just proven because $\gcd(n, k) = \gcd(n, n-k)$.

We can write the sum in Equation (4.18) as

$$B_s(n, q) = \frac{1}{2nq!} \sum_{P,P'} M(e_1, \ldots, e_n) N(c_1, \ldots, c_q) \prod_l (m(l))^{e_l},$$

where, combining what we have just proven with Lemma 4.1, we have

$$M(e_1, \ldots, e_n) = \begin{cases} \phi\left(\frac{n}{d}\right), & \text{for elements } R^k; \\ \frac{n}{2}, & \text{for elements } FR^k,\ n \text{ even}; \\ n, & \text{for elements } FR^k,\ n \text{ odd}, \end{cases}$$

and $N(c_1, \ldots, c_q)$ is given by Equation (4.19). $\diamond$

Corollary 4.3. *The number of distinct bracelets with* n *beads and* 2 *colors, where color swapping is allowed, is given by*

$$B_s(n,2) = \frac{2^{\lfloor \frac{n}{2} \rfloor} + N_s(n,2)}{2}.$$

Proof: We want to write $B_s(n,2)$ in terms of $N_s(n,2)$, so it becomes easier to use Equation (4.18) rather than the equation in the statement of Theorem 4.11. The reason is that we want to first separate elements F^tR^k based on the value of t instead of looking at the partitions P, P'. From Equation (4.18) we have that for $q = 2$,

$$B_s(n,2) = \frac{1}{4n} \sum_{F^tR^k\pi} \prod_l (m(l))^{e_l}, \tag{4.20}$$

where $F^tR^k\pi \in D_{2n} \times S_2$, $m(l)$ is given by Equation (4.5), and e_l is the number of cycles of length l in a permutation $F^tR^k \in D_{2n}$. In order to rewrite this sum, we will need to split it up in a number of steps.

First, notice that we can divide $D_{2n} \times S_2$ into two sets of elements $F^tR^k\pi$, depending on the value of t. If $t = 0$, our bracelet is unflipped and we are left with the elements $R^k\pi$, i.e., the group $C_n \times S_2$. If $t = 1$, our bracelet is flipped and we have the elements $FR^k\pi$. The sum in Equation (4.20) can be rewritten as

$$B_s(n,2) = \frac{1}{4n} \left(\sum_{R^k\pi} \prod_l (m(l))^{e_l} + \sum_{FR^k\pi} \prod_l (m(l))^{e_l} \right). \tag{4.21}$$

Recall from the proof of Theorem 4.11 that $\prod_l (m(l))^{e_l}$ counts the number of elements of $D_{2n} \times S_2$ that are invariant under some $F^tR^k\pi$. Now, because we are summing over $R^k\pi$, we are instead counting the number of elements of $C_n \times S_2$ that are invariant under some $R^k\pi$, and e_l still represents the number of cycles of length l in an element of D_{2n} although we are only concerned with elements in the subgroup $C_n \times S_2$. If we sum and divide by $2n$ (the order of $C_n \times S_2$), we are, by Pólya's Theorem, counting the number of distinct necklaces with n beads and 2 colors where color swapping is allowed, i.e., we have

$$N_s(n,2) = \frac{1}{2n} \sum_{R^k\pi} \prod_l (m(l))^{e_l}.$$

Equation (4.21) can now be written as

$$B_s(n,2) = \frac{1}{2}N_s(n,2) + \frac{1}{4n}\sum_{FR^k\pi}\prod_{l}(m(l))^{e_l}. \tag{4.22}$$

It will soon be useful to consider separate cases for n even and odd. We saw in the proof of Lemma 4.1 that for elements $FR^k \in D_{2n}$, there is one conjugacy class for n odd and two for n even. In the proof of Theorem 4.11 we verified that for n odd, all elements in $\{FR^k\}$ have cycle structure consisting of one cycle of length 1 and $\frac{n-1}{2}$ cycles of length 2, i.e., we have the partition $n = 1 + 2e_2$ where $e_1 = 1$ and $e_2 = \frac{n-1}{2}$. For n even, elements in $\{FR^{2k}\}$ have cycle structure consisting of $\frac{n}{2}$ cycles of length 2, i.e., the partition $n = 2e_2$ with $e_1 = 0$ and $e_2 = \frac{n}{2}$. Elements in $\{FR^{2k+1}\}$ have cycle structure consisting of $\frac{n-2}{2}$ cycles of length 2 and two cycles of length 1, i.e., the partition $n = 2 + 2e_2$, where $e_1 = 2$ and $e_2 = \frac{n-2}{2}$. To incorporate this information we must split the sum in Equation (4.22) even further, according to the parities of n and k, and the fact that $e_l = 0$ for $l \geq 3$. This gives

$$\begin{aligned} B_s(n,2) &= \frac{1}{2}N_s(n,2) + \frac{1}{4n}\sum_{FR^k\pi} m(1)(m(2))^{\frac{n-1}{2}} \qquad (n \text{ odd}),\\ B_s(n,2) &= \frac{1}{2}N_s(n,2) + \frac{1}{4n}\sum_{\substack{FR^k\pi\\ k \text{ odd}}} (m(1))^2(m(2))^{\frac{n-2}{2}}\\ &\quad + \frac{1}{4n}\sum_{\substack{FR^k\pi\\ k \text{ even}}} (m(1))^0(m(2))^{\frac{n}{2}} \qquad (n \text{ even}). \end{aligned}$$

Now, in order to evaluate the $m(1)$ and $m(2)$ terms we must have values for c_1 and c_2. These are dependent on the partition of $q = 2$. For $\pi_1 = (12)$ we have $m(1) = c_1 = 0$ and $m(2) = c_1 + 2c_2 = 0 + 2(1) = 2$. For $\pi_2 = (1)(2)$ we have $m(1) = c_1 = 2$ and $m(2) = c_1 + 2c_2 = 2 + 2(0) = 2$. To incorporate these values into the above equations we must again split each sum into two, this time depending on the partition of S_2. In the case of π_1, our general element can be written as $F^tR^k\pi_1$, whereas for π_2, our general element is F^tR^k since this is the identity permutation in S_2. We can now write

$$B_s(n,2) = \frac{1}{2}N_s(n,2) + \frac{1}{4n}\sum_{FR^k} m(1)(m(2))^{\frac{n-1}{2}} + \frac{1}{4n}\sum_{FR^k\pi_1} m(1)(m(2))^{\frac{n-1}{2}} \qquad (n \text{ odd}),$$

$$B_s(n,2) = \frac{1}{2}N_s(n,2) + \frac{1}{4n}\sum_{\substack{FR^k \\ k \text{ odd}}}(m(1))^2(m(2))^{\frac{n-2}{2}} + \frac{1}{4n}\sum_{\substack{FR^k\pi_1 \\ k \text{ odd}}}(m(1))^2(m(2))^{\frac{n-2}{2}}$$
$$+\frac{1}{4n}\sum_{\substack{FR^k \\ k \text{ even}}}(m(2))^{\frac{n}{2}} + \frac{1}{4n}\sum_{\substack{FR^k\pi_1 \\ k \text{ even}}}(m(2))^{\frac{n}{2}} \quad (n \text{ even}).$$

We can explain the fact that $m(1) = 0$ for permutations with $\pi_1 = (12)$ in the following way. The term $m(1)$ counts the number of colors an element s_i in a length one cycle can take on, whilst maintaining invariance under some permutation. Since the permutation π_1 switches the color of the element in a length-one cycle, it cannot possibly remain invariant and $m(1) = 0$. Therefore the entire cycle σ_n cannot remain invariant under this permutation, and the entire term drops out of the sum. Note that we are using the convention $0^0 = 1$ in the final term of the n even case, since we have $m(1)^{e_1} = 0^0$.

Now, substituting the values of $m(1)$ and $m(2)$ gives

$$B_s(n,2) = \frac{1}{2}N_s(n,2) + \frac{1}{4n}\sum_{FR^k} 2 \cdot 2^{\frac{n-1}{2}} \quad (n \text{ odd}),$$
$$B_s(n,2) = \frac{1}{2}N_s(n,2) + \frac{1}{4n}\sum_{\substack{FR^k \\ k \text{ odd}}} 2^2 2^{\frac{n-2}{2}}$$
$$+\frac{1}{4n}\sum_{\substack{FR^k \\ k \text{ even}}} 2^{\frac{n}{2}} + \frac{1}{4n}\sum_{\substack{FR^k\pi_1 \\ k \text{ even}}} 2^{\frac{n}{2}} \quad (n \text{ even}).$$

We can take each sum over the number of elements in the given set, noting that for n odd, $|\{FR^k\}| = n$, and for n even, $|\{FR^k | k \text{ odd}\}| = \frac{n}{2}$, $|\{FR^k | k \text{ even}\}| = \frac{n}{2}$ and $|\{FR^k\pi | k \text{ even}\}| = \frac{n}{2}$. This gives

$$B_s(n,2) = \frac{1}{2}N_s(n,2) + \frac{1}{4n}(n)2 \cdot 2^{\frac{n-1}{2}} \quad (n \text{ odd}),$$
$$B_s(n,2) = \frac{1}{2}N_s(n,2) + \frac{1}{4n}\left(\frac{n}{2}\right) 2^2 2^{\frac{n-2}{2}}$$
$$+\frac{1}{4n}\left(\frac{n}{2}\right)2^{\frac{n}{2}} + \frac{1}{4n}\left(\frac{n}{2}\right)2^{\frac{n}{2}} \quad (n \text{ even}).$$

Simplifying gives

$$B_s(n,2) = \frac{1}{2}N_s(n,2) + \frac{1}{2}2^{\frac{n-1}{2}} \quad (n \text{ odd}),$$
$$B_s(n,2) = \frac{1}{2}N_s(n,2) + \frac{1}{2}2^{\frac{n-2}{2}} + \frac{1}{8}2^{\frac{n}{2}} + \frac{1}{8}2^{\frac{n}{2}}$$
$$= \frac{1}{2}N_2(n,2) + \frac{1}{2}2^{\frac{n}{2}} \quad (n \text{ even}).$$

We can now combine these into one equation as

$$B_s(n,2) = \frac{N_s(n,2) + 2^{\lfloor \frac{n}{2} \rfloor}}{2},$$

completing the proof. ◇

The sequence $B_s(n,2)$, and corresponding formula, can be found in [67, A000011]. The first few terms, for $n \geq 0$, are

$$1, 1, 2, 2, 4, 4, 8, 9, 18, 23, 44, 63, 122, \ldots.$$

Example 4.5. Notice that the above sequence for bracelets with color swapping matches the sequence given in (4.10) for necklaces with color swapping, up to $n = 6$. For $n = 7$ we have $N_s(7,2) = 10$ and $B_s(7,2) = 9$. This difference can be seen in Figure 4.4. Both objects belong to the set of necklaces with color swapping, but are

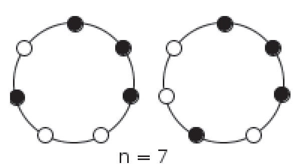


Figure 4.4: Equivalent bracelets.

equivalent in the set of bracelets with color swapping, because they are mirror images of each other. This is the first instance where taking the mirror image of a necklace does not simply give the same necklace back under rotation.

4.5 Enumerating the Equivalence Classes

Next we want to count the number of primitive bracelets (primitive cycles with rotation and reversal) where color swapping is allowed. This means that we are now allowed to reduce a cycle σ_n to its primitive cycle σ_k, where $k|n$. We have now incorporated all four operations in our equivalence relation and so we are counting the number of equivalence classes.

Fine [27] explains how we can obtain the number of primitive classes from the number of classes which are not necessarily primitive using Möbius inversion. Recall that the Möbius function, $\mu(n)$, is defined as follows:

$$\mu(n) := \begin{cases} 0, & \text{if } n \text{ has at least one repeated prime factor;} \\ 1, & \text{if } n = 1; \\ (-1)^k, & \text{if } n \text{ is a product of } k \text{ distinct primes.} \end{cases}$$

Theorem 4.12. *The number of distinct primitive bracelets with n beads and q colors, where color swapping is allowed, i.e., the number of equivalence classes of cycles σ_n, is given by*

$$B_{sp}(n, q) := \sum_{d|n} \mu(d) B_s\left(\frac{n}{d}, q\right),$$

where $\mu(d)$ is the Möbius function.

Proof: First denote the number of primitive bracelets with n beads and q colors, where color swapping is allowed, by $B_{sp}(n, q)$. Now consider $B_s(n, q)$, which by Theorem 4.11 is the total number of bracelets with n beads and q colors, where color swapping is allowed, whether or not the bracelets are primitive. If we break this number down according to primitive cycle, we need to count the number of primitive cycles of length d where $d|n$. Summing over all d we obtain

$$B_s(n, q) = \sum_{d|n} B_{sp}(d, q).$$

To find an expression for $B_{sp}(n, q)$ we simply use Möbius inversion. This tells us that given the expression $g(n) = \sum_{d|n} f(d)$, we can write the function f as $f(n) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right)$, where $\mu(d)$ is the Möbius function. In our case, applying Möbius inversion gives

$$B_{sp}(n, q) = \sum_{d|n} \mu(d) B_s\left(\frac{n}{d}, q\right).$$

◇

For the case we are concerned with, $q = 2$, the following corollary is immediately obtained:

Corollary 4.4. *The number of distinct primitive bracelets with* n *beads and* 2 *colors, where color swapping is allowed, i.e., the number of equivalence classes of coefficient cycles* σ_n*, is given by*

$$B_{sp}(n, 2) = \sum_{d|n} \mu(d) B_s\left(\frac{n}{d}, 2\right).$$

The sequence $B_{sp}(n, 2)$, and corresponding formula, can be found in [67, A000046]. The first few terms, for $n \geq 0$, are

$$1, 1, 1, 1, 2, 3, 5, 8, 14, 21, 39, 62, 112, \ldots.$$

Notice that prime-indexed terms in this sequence are one less than those in $B_s(n, 2)$ ([67, A000011]). This is due to the fact that when n is prime, we only have one non-primitive coefficient cycle, namely $((+)^n)$.

Example 4.6. In Figures 4.5 and 4.6 we have listed the set of equivalence classes for $n \leq 7$.

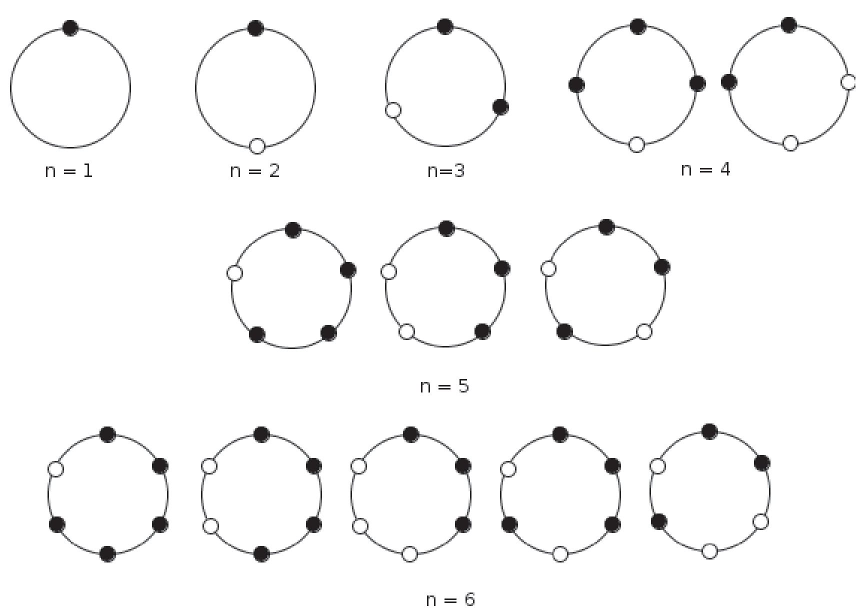


Figure 4.5: Equivalence classes for $n \leq 6$.

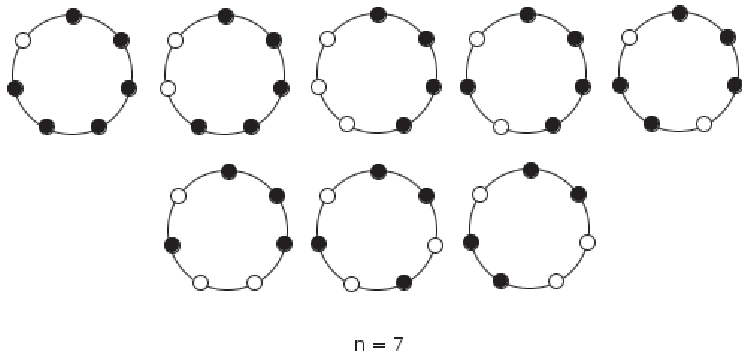


Figure 4.6: Equivalence classes for $n = 7$.

Here we have equivalence among bracelets with the same primitive cycle. For example, the bracelets in Figure 4.7 are now equivalent.

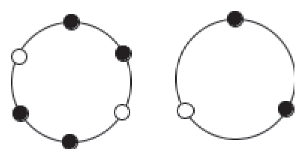


Figure 4.7: Members of the same equivalence class.

The following result relates the value of n to the growth of the set of equivalence classes of coefficient cycles (or equivalently, periodic coefficient sequences).

Corollary 4.5. *If at least one equivalence class of periodic coefficient sequences is bounded then n is divisible by 2 or 3. Also if n is not divisible by 2 or 3 then all equivalence classes of sequences grow exponentially.*

Proof: Suppose at least one equivalence class of periodic coefficient sequences is bounded and consider the cases where n is odd. Corollary 3.2 tells us that for the bounded equivalence classes of sequences, we have $3 \mid n$. If n is even, we have $2 \mid n$. Note that for $n = 0$, we have the single equivalence class containing the corresponding product matrix $\pm I$. In this case growth is bounded and $2, 3 \mid n$. The second, more powerful statement follows from the fact that if $2 \nmid n$ then n is odd, and by Corollary 3.2 again, if $3 \nmid n$ then growth is exponential. $\diamond$

We can now consider the number of equivalence classes of each growth type. Table 4.1 gives these numbers as well as the number of coefficient cycles that fall into these categories. Recall that there are 2^n different length-n coefficient cycles that are separated into equivalence classes. Corollary 4.5, which said that for n not divisible

n	coefficient cycles	E	B	L	equivalence classes	E	B	L
1	2	2	0	0	1	1	0	0
2	4	2	2	0	1	0	1	0
3	8	2	6	0	1	0	1	0
4	16	6	10	0	2	1	1	0
5	32	32	0	0	3	3	0	0
6	64	14	8	42	5	1	0	4
7	128	128	0	0	8	8	0	0
8	256	182	74	0	14	11	3	0
9	512	290	222	0	21	14	7	0
10	1024	672	352	0	39	27	12	0
11	2048	2048	0	0	62	62	0	0
12	4096	2082	16	1998	112	57	0	55

Table 4.1: Number of coefficient cycles and equivalence classes of each growth type.

by 2 and 3, all equivalence classes have exponential growth, is apparent here. We can also consider equivalence classes of balanced coefficient cycles, and their breakdown into growth types. The number of equivalence classes which are balanced, for even $n \geq 0$, is given by the sequence

$$1, 1, 1, 2, 5, 12, 31, 84, 250, 762, 2504, 8358, 28928, \ldots,$$

which can be can found in [67, AA045633].

Example 4.7. For $n = 8$ we have 14 equivalence classes of coefficient cycles, and 5 of these are balanced. The balanced cases are $(++++----)$, $(+++-+---)$, $(+++--+--)$, $(++-+-+--)$ and $(++-+--+-)$, and they have growth types E, E, E, B and E, respectively.

Our original intent was to try to shed some light on Viswanath's constant by determining the number of equivalence classes, the size of each equivalence class, and the growth type/rate of each equivalence class. It may then have been possible to combine this information into a formula for the average of the growth rates of periodic coefficient sequences with period length n. By letting $n \to \infty$, we would be in theory be computing Viswanath's constant. This approach turned out to be quite difficult. We have found the number of equivalence classes for period length n, and a way to determine the growth rate of each, but finding the size of each class and constructing a formula are complex problems.

4.6 Some Applications of Necklaces and Bracelets

Interestingly, the concept of necklaces and bracelets appears in a paper by Grünbaum and Shephard [35] on the geometry of fabrics, in particular, those which arise from weaving. The mathematical treatment of this subject has a short history, despite mankind's long history of weaving. We can think of the weaved fabric as an infinite grid where horizontal strips covering vertical strips are represented by black squares, and vertical strips covering horizontal strips are represented by white squares, to form repeating patterns. A simple class of fabrics known as twills is formed from a periodic length n pattern, occurring in consecutive rows, but shifted by one square to the right as we move down. Therefore the entire pattern can be represented by a portion of a single row. How many distinct twill patterns are there? We can think of each row as being generated by a size n necklace made up of two colors. Furthermore we will consider the twill equivalent to itself if we swap colors (which is equivalent to interchanging the horizontal and vertical sides of the fabric, i.e., flipping the fabric over the line $x + y = 0$) or if we flip the fabric over in the plane (which in this case is equivalent to a 90° rotation). Therefore we are enumerating bracelets of two colors which are equivalent under color swapping, the number of which we have seen in Corollary 4.3 to be

$$B_s(n, 2) = \frac{2^{\lfloor \frac{n}{2} \rfloor} + N_s(n, 2)}{2}.$$

Grünbaum and Shephard give a sketch of this proof and enumerate the possible patterns for $n \leq 8$.

In [39], Hoskins and Penfold Street carry on the work of Grünbaum and Shephard [35] on twills. They devise an algorithm for computing the list of twills with period

n, and also derive a formula for evaluating the number of period n twills with a given number of breaks (i.e., changes in color, and similar to the sign-flips mentioned in Chapter 1). Interestingly, they also define an equivalence relation on period n binary sequences such that sequences S and T are equivalent if and only if one can be transformed into the other by a shift, reversal, complementation or a finite number of these operations. In other words, equivalence classes are determined by the action of the group $D_{2n} \times S_2$. The equivalence classes defined in Definition 4.3 have the additional condition that two sequences (not necessarily of the same length) are equivalent if one can be transformed into the other by reducing or extending it by a multiple of the primitive cycle. The idea of a "balanced twill" is also introduced, where the number of black squares is equal to the number of white squares. Recall that balanced coefficient cycles were discussed in Section 3.3.

There has also been some interesting work (e.g., Chen [17]) relating the combinatorics of binary necklaces with DNA sequences. A DNA molecule can be seen as a sequence over four bases: A, C, G and T, whose double helix form is composed of two complementary strands. Furthermore, the repetition of patterns within the sequence can be viewed cyclically. Classification of length n DNA patterns can be viewed in terms of equivalence classes of length n words on four letters under rotation and complementation. Furthermore, there exists a bijection between length n words on four letters and length $2n$ words on 2 letters (our binary necklaces), which allows for enumeration of the equivalence classes of length n DNA sequences.

Chapter 5

The Tree $\boldsymbol{R}$ and the Growth Rate of the Expected Value of a Random Fibonacci Sequence

5.1 The Reduced Tree $\boldsymbol{R}$

Recall the tree $\widetilde{T}$ given in Figure 1.1(c), and further generated in Figure 5.1. From here on we will refer to $\widetilde{T}$ simply as T, as it is the complete binary tree we will focus on. In [64], Rittaud develops a variation of the tree T called R, which will prove very

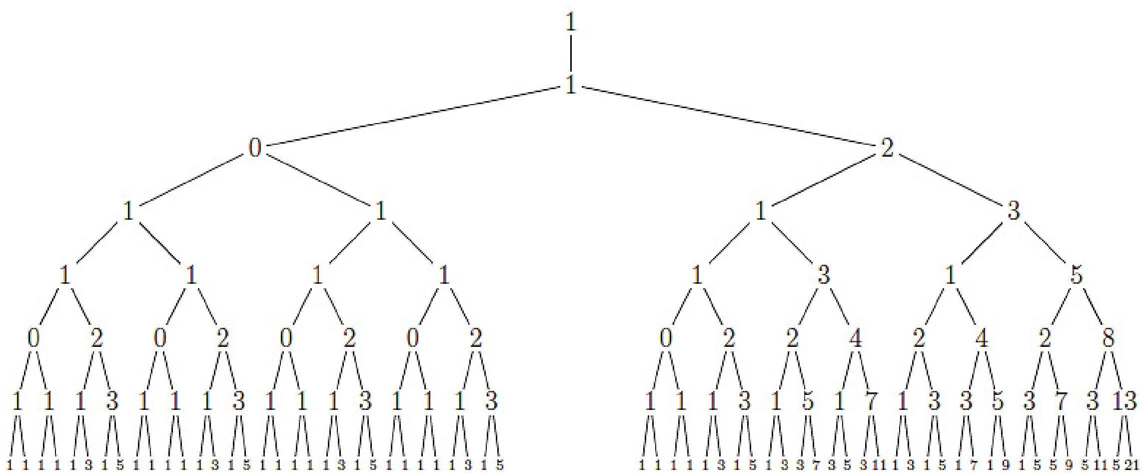


Figure 5.1: The tree T (Rittaud [64]).

useful in determining results about random Fibonacci sequences. Before considering it, we will look at the following interesting properties of T, as found in [64]. The properties of T and R in this section are given with proof in order to help illuminate the structure and connection between the trees.

Proposition 5.1. *If a is a node in the tree T with child b, then a and b are relatively prime.*

Proof: Let $d = \gcd(a, b)$, and let z be a's parent. The node b must then be of the form $b = z + a$ or $b = |z - a|$. We must also have that $\gcd(z, a) = d$. Tracing backwards we see that z and its parent must also have greatest common divisor d, and by induction this pattern continues until we reach the initial nodes $1, 1$. Clearly $\gcd(1, 1) = 1$ and so $d = 1$. $\diamond$

The converse, which states that any relatively prime pair a, b can be found as an edge in the tree T, is also true.

Recall from Chapter 1 that $T^{(a,b)}$ is the tree formed from all possible random Fibonacci sequences with $t_1 = a$ and $t_2 = b$. We then have that $T = T^{(1,1)}$. We can assume that a and b are relatively prime; otherwise if $\gcd(a, b) = d$, every node in $T^{(a,b)}$ is a multiple of d.

Proposition 5.2. *Given a relatively prime pair (a, b), the tree $T^{(a,b)}$ appears infinitely many times in the tree T.*

Proof: The proof uses the converse of Proposition 5.1, which says that $T^{(a,b)}$ appears in T. Therefore, if we can show that T in turn appears in $T^{(a,b)}$, then the proposition is proved. This requires the pair $(1, 1)$ to appear in $T^{(a,b)}$. This must happen because given a parent-child pair (a, b), the left child is $|b-a|$, and we must have that $\max(b, |b - a|) \leq \max(a, b)$. Continuing to take left children, we obtain a sequence of positive numbers which must eventually reach $(1, 1)$ because because by Proposition 5.1 we cannot have two successive equal nodes t in a branch, where $t > 1$, and so the above property of maxima cannot give equality for two successive pairs. $\diamond$

The following is a "characterization of shortest walks" in the tree T.

Proposition 5.3. *If (a, b) is a pair of successive nodes in T, the shortest path to reach it from the root is made up of successive nodes (c, d) where the parent of c is $|c - d|$.*

Rittaud [64] goes on to generate the reduced tree R, which will play a vital role in results related to Viswanath's constant. The tree R, as seen in Figure 5.2, is generated from T by pruning any edge (and subsequent branch) which already occurs at an earlier level in the tree, i.e., it is composed of the shortest walks in T. There is an exception, however. Node 0 in row 3 of T belongs to the edge $(1, 0)$ as well as both edges $(0, 1)$, which do not appear at an earlier level of the tree, but are not included in R. Therefore, the entire left half of the tree T does not appear in R. This pruning process removes infinitely many subtrees of R from T. The tree T is

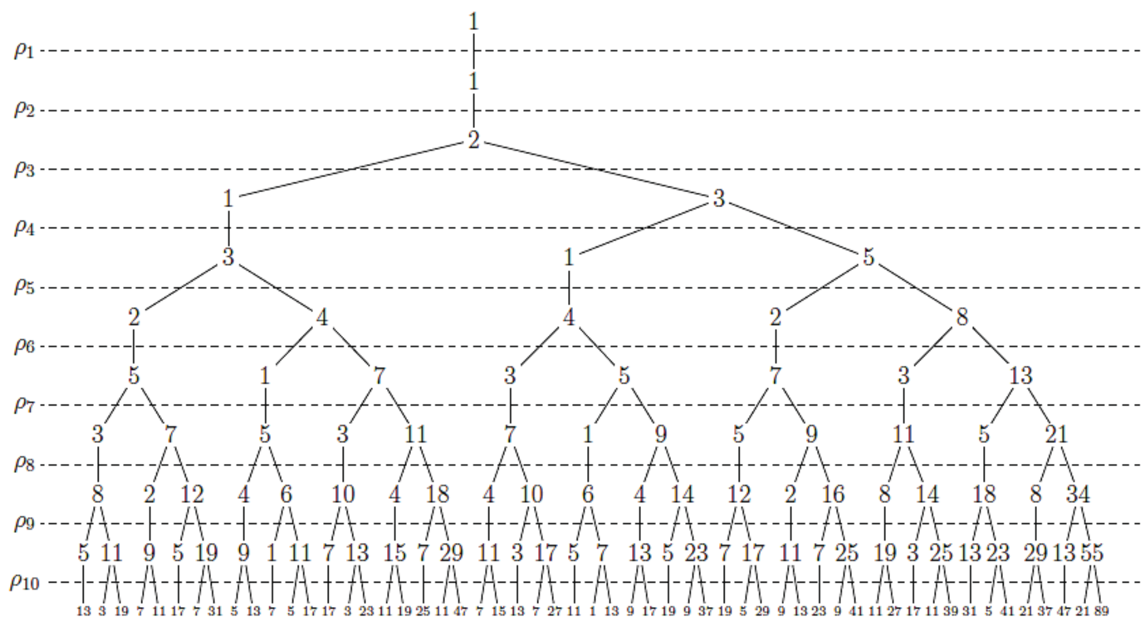


Figure 5.2: The reduced tree R (Rittaud [64]).

therefore composed of R and pieces of R. The growth of T must then follow from the growth pattern in R.

Let ρ_n and τ_n denote the n^{th} rows of entries in the trees R and T respectively. Note that here we use ρ_n to list nodes, whereas Rittaud uses it to list edges. The labels in Figure 5.2 therefore refer to the nodes above the dashed line. We will also use the notation $c_R(\rho_n)$ to denote children in the tree R, so that $c_R(\rho_n) = \rho_{n+1}$.

Example 5.1. Let us look at the shortest walk characterization given in Proposition 5.3. We know that any pair of successive nodes (a, b) in a branch of R can be traced back through the parents $|a - b|$ to reach the root. Consider the pair $(9, 2)$ with nodes in ρ_8 and ρ_9. If we take the sequence of absolute values of the differences in consecutive pairs, we get $2, 9, 7, 2, 5, 3, 2, 1, 1$. This is exactly the path which takes us back to the root of R.

The following propositions from Rittaud [64] contains some vital facts about R.

Proposition 5.4. *Any edge (a, b) in the tree R occurs only once.*

Proof: We know from the definition of R that a given edge can only appear at one level of the tree R. It remains to show that an edge cannot appear more than once at a level of R. By Proposition 5.3, any node a with child b has parent $|a - b|$, and so if two such nodes appeared in a given row, we would be able to trace both back to the root through the same set of nodes. But we can see that the initial levels of R contain no repeated edges. Therefore any edge (a, b) appears only once in the tree R. $\diamond$

Proposition 5.5. *The following are important properties about the trees R.*

1. *Right children in R are greater than their parents and left children are smaller, with the exception of the initial nodes* $1, 1$.
2. *Right children in R have two children and left children only have a right child.*
3. *The left child of a left or right child in R is equal to its great grandparent.*

Proof: Suppose a is a left child in row $n - 1$ of tree R. Also suppose a's parent is b, and b is the right child of c as seen in Figure 5.3. The left child of b is $a = |b - c|$ and so may be either $b - c$ or $c - b$. By Proposition 5.3, if the child of b is $b - c$, the parent of b must be $|b - (b - c)| = c$, which is true, and so $b - c$ is in R. Therefore $b > c$, and $a = b - c < b$, i.e., the left child of a right child is smaller than its parent. The fact that $b > c$ also tells us that right children are greater than their parents, although this is clear because we are dealing with positive integers.

We have that $a = b - c$, and so $c = b - a$. In this case the left child of a is equal to its great grandparent. In fact, the left child of any (left or right) node a with parent b is equal to its great grandparent because by definition, the left child of a is $|b - a|$, and by Proposition 5.3 the parent of b is $|b - a|$, proving Statement 3. Using this, we have that a's grandparent must have parent a. Putting this information together we see that the edge $(a, c = b - a)$ is repeated and so a's left child in row n does not exist

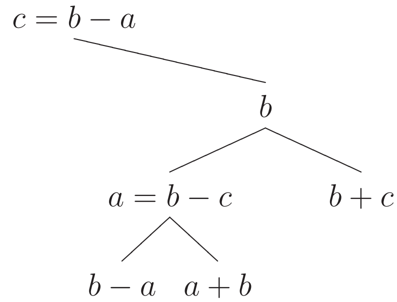


Figure 5.3: Left children of left children equal great grandparents.

in the tree R. This argument tells us that the left child of a left child of a right node cannot exist in R. Further, we have already shown that left children of right nodes are less than their parents, and this is now equivalent to stating that all left nodes in R are less than their parents. Statement 1 has now been proven.

Next we can show that all nodes have right children. The right child of a with parent b is $a + b$, but how do we know the edge $(a, a + b)$ has not already occurred in the tree? By Proposition 5.3, the shortest walk to a must have $|a - (a + b)| = b$ as a parent, which we know is true. We can conclude that left children have a right child only. Lastly we must show that right nodes have left children. Consider the right node b with parent c in Figure 5.3. We have seen above that the left child of b is $b - c$, which exists in R, and so we can conclude that right nodes have two children and left nodes only have a right child, proving Statement 2. $\diamond$

Note that this shortest walk argument does not hold for the case of a left child of a left child. For example, in Figure 5.3 we have that the left child of left node a is $b - a$. By Proposition 5.3, the parent of a in the shortest walk is then $|a - (b - a)| = |2a - b| \neq b$, since $a \neq 0$. This tells us that $b - a$ does not belong in R.

5.2 Variations of the Tree R

Let us first consider what happens when we generalize the initial values of our tree. If we start with the tree $T^{(a,b)}$, the corresponding tree $R^{(a,b)}$ may be quite different from R, and the subsequent material in this chapter may not carry over. If $a < b$,

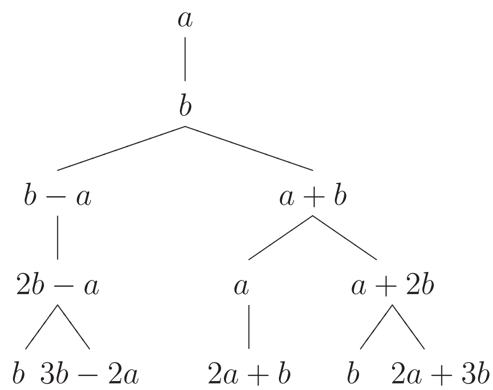
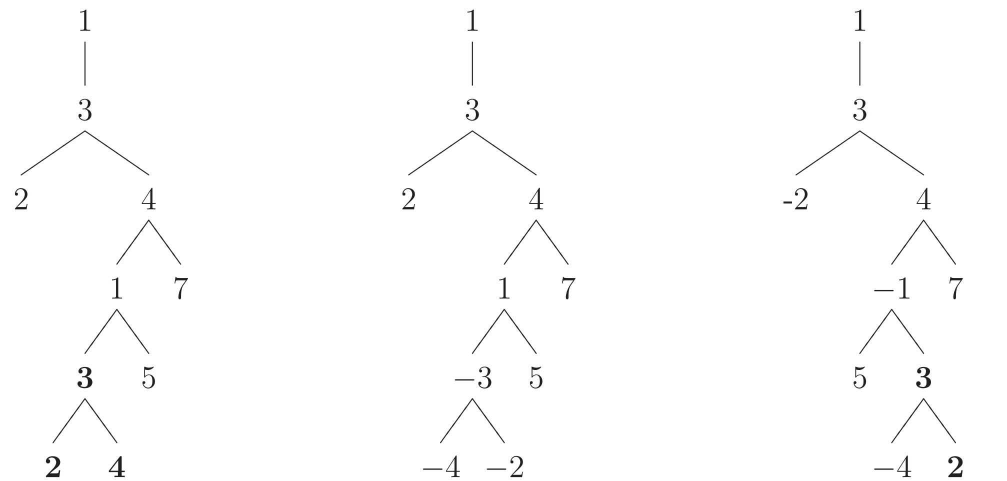


Figure 5.4: The tree $R^{(a,b)}$ for $a < b$.

Proposition 5.5 holds. We have b as the only child of a, which we can then think of as a right child which is bigger than its parent. Further, left children of left children cause repeated edges and are therefore removed from $R^{(a,b)}$, as can be seen in Figure 5.4. If $b = a$, we have that $b - a = 0$ in row 3, which leads to the left side of the tree being repetitive and hence removed, as is the case for $R^{(1,1)}$. Note that when $a < b$, we have non-repetitive nodes in the left half of the tree, although the overall structure of $R^{(a,b)}$ is the same as that of $R^{(1,1)}$ if we think of the initial values as shifted to $\rho_2 = a$, $\rho_3 = b$.

If $a > b$, Proposition 5.5 does not hold. The orientation of nodes in $T^{(a,b)}$ differs from that in the $a < b$ case and as a result the repetition in the tree does not always occur at the left child of a left node.

We can now consider another variation of the tree R. Recall that Rittaud [64] used Equation (1.8), $\widetilde{f}_n = |\widetilde{f}_{n-1} \pm \widetilde{f}_{n-2}|$, to generate the tree T from which R was obtained. What happens if we instead use Equation (1.6), $t_n = \pm t_{n-1} + t_{n-2}$, or Equation (1.7) $f_n = f_{n-1} \pm f_{n-2}$? We looked at the corresponding matrix representations of each of these recurrences in Section 2.1. Also, in Section 1.3 we analyzed the trees $\widetilde{T} = T$, T_1 and T_2 resulting from the three recurrences and saw in Theorem 1.2 that the trees T, $|T_1|$ and $|T_2|$ are all comprised of the same set of sequences. Figure 5.5 demonstrates the behaviour of sequences in each of the three trees. We will denote the reduced tree



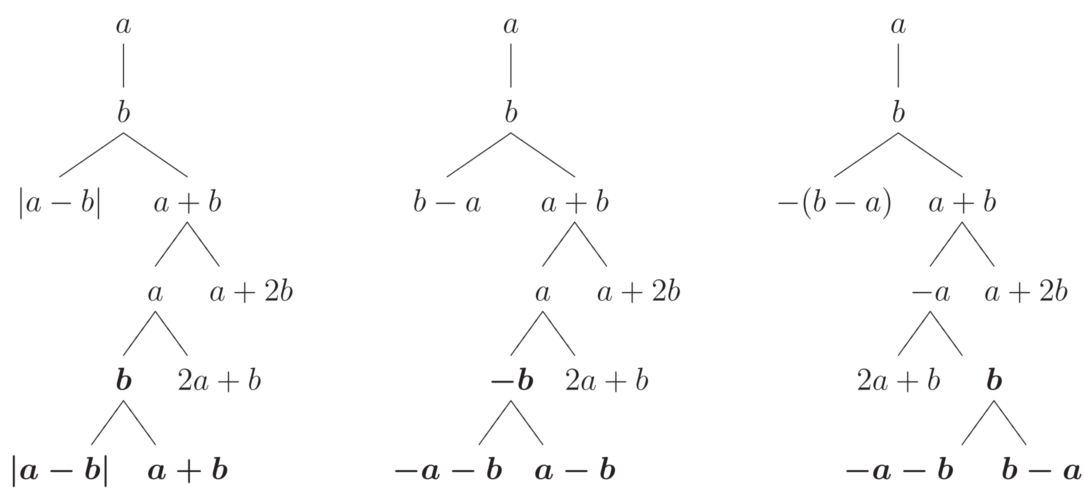
(a) Branch of R. (b) Branch of R_2. (c) Branch of $R_1^{(2,1)}$.

Figure 5.5: Example of sequence behaviour in R and its variations.

of T_1 by R_1 and the reduced tree of T_2 by R_2.

Figure 5.5(a) contains a branch from the tree R originating in ρ_4, with the exception of the bolded terms, which belong to T only. As was seen in Proposition 5.5, the left child of a left child is removed when creating the tree R, and the left child (3 in this case) is equal to its great grandparent. Because $(1, 3)$ is a repeated edge, the children of 3 are also repeated earlier in the tree. Now consider the branch of R_2 given in Figure 5.5(b). This branch, along with the bolded terms from T_2, was generated using the recurrence $f_n = f_{n-1} \pm f_{n-1}$, and so contains negative terms. The removal of repeated edges applies to the absolute value of the tree, since we are ultimately interested in the absolute value of our sequences. The branch has the same shape as that in Figure 5.5(a), and we can see that -3, which is the left child of the left child of 4, is removed. This -3 has the negative value of its great grandparent, and its children are negated and switched, compared to the children of the great grandparent. The branch in Figure 5.5(c) was generated using the recurrence $t_n = \pm t_{n-1} + t_{n-2}$ and so also contains negative terms. When we remove repeated edges (in absolute value) we see that the tree takes on a different shape than the previous two. The node 3 is removed; however, it is a right child. It is equal in value to its great grandparent, and again its children are negated and switched compared to the children of the great grandparent. The node 5, which is a left child of a left child, is not removed from the tree. In contrast to the restriction on two consecutive left nodes, it seems in the tree R_2 we cannot have the pattern right, left, right. Note that the exact branch in Figure 5.5(c) is not found in tree R_1. The shape of R_1 differs from R and so the shape previous to this branch differs also, creating a permutation of nodes. In the fourth row of R_1, we have node -1 with left child 3. The branch in Figure 5.5(c) is found, however, in the tree $R_1^{(2,1)}$, where an initial value differs.

We can prove that the above observations are true in general using the analogous branches in Figure 5.6, which start with values a and b, where b is a right child of a and $b > a$ by Proposition 5.5. The following are useful Propositions on R and its variations.



(a) Branch of R. (b) Branch of R_2. (c) Branch of R_1.

Figure 5.6: Comparing sequence behaviour in the general case of R and its variations.

Proposition 5.6. *The trees R, R_1 and R_2 all contain the same set of sequences in absolute value.*

Proof: From Theorem 1.2 we know that the trees T, $|T_1|$ and $|T_2|$ all contain the same set of sequences, except they are permuted. To create R, we removed edges from the tree T that were repeated at an earlier level. We know that $|T_1|$ and $|T_2|$ are simply rearrangements of T, and so those same repeated edges will be removed in creating R_1 and R_2. We are then left with trees R, $|R_1|$ and $|R_2|$, which still all contain the same set of sequences. $\diamond$

Proposition 5.7. *The trees R and R_2 are equal and positive and the tree R_1 contains negative nodes.*

Proof: By definition, the tree R is positive. We can use induction to show that the tree R_2 is positive and equal to R. The first right child occurs in row 3 and has value 2, which is greater than the value of its parent. Calculating the next few rows, we see that right children of both left and right nodes are positive and greater than their parents, and equal to the corresponding rows in R. Now suppose that any right node b in row n is greater than its parent a, where a and b are positive, as shown in Figure 5.6(b). By Equation (1.7), the left child of b is $b - a > 0$. This is equal to the corresponding left child in R. The right child is simply $a + b$, which is positive and also equal to the corresponding right child in R. If we now consider the left child $b - a$ in Figure 5.6(b), we see that its right child is $2b - a$, which is positive and again equal to the right child in R. We conclude that all nodes in R_2 must be positive, and since R and R_2 have the same left and right children, they are equal.

In R_1, however, Equation (1.8) tells us that the left child of b with parent a is $a - b$, which is negative if a and b are positive with $b > a$, as seen in Figure 5.6(c). By calculating the first few rows we see that row 4 contains -1, and so this case does indeed occur. $\diamond$

Rittaud [64] develops what he calls the $\mathrm{SL}(2, \mathbb{N})$ tree. The matrix tree has root I, with child A. Right children of any node M after the root are given by AM, whereas left children are given by $B'M$. The matrix tree is based on the tree $R^{(a,b)}$ and so follows the rule of no consecutive left children. These product matrices, formed from $A, B' = \left(\begin{smallmatrix} 0 & 1 \\ \pm 1 & 1 \end{smallmatrix}\right)$, model the sequence $f_n = f_{n-1} \pm f_{n-2}$, as given in Equation (2.6), and are all positive like the nodes of R_2. This tree will be further discussed in Chapter 6.

In [40], Janvresse et al. use the random Fibonacci sequence generated by $f_n = f_{n-1} \pm f_{n-2}$. We have seen in Equation (2.7) that this recurrence can be modeled using matrices $\left(\begin{smallmatrix} 0 & \pm 1 \\ 1 & 1 \end{smallmatrix}\right)$, which we have denoted by A and $\hat{B}$. (Recall that moving left or right down the tree corresponds to multiplying a product matrix by $\hat{B}$ or A respectively.) Their characterization of the tree R_2 involves a reduction process, where $A\hat{B}\hat{B}A$ is replaced by $-\hat{B}$ and $A\hat{B}\hat{B}\hat{B}$ is replaced by $-A$. This matrix representation implies the fact that in Figure 5.6(b), traveling right, left, left, from the first node b, is equivalent to switching and negating the following children. This is shown below, where the second terms in the vectors are the children in question. The top (bottom) pair of equations represents right (left) children of $-b$ and b respectively.

$$(a,b)A\hat{B}\hat{B}A = (a,b)\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} = (-b, a-b), \qquad (a,b)A = (b, a+b),$$

$$(a,b)A\hat{B}\hat{B}\hat{B} = (a,b)\begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix} = (-b, -a-b), \qquad (a,b)\hat{B} = (b, b-a).$$

This process is even simpler in the positive case R, given in Figure 5.6(a) and generated by the non-linear recurrence $\widetilde{f}_n = |\widetilde{f}_{n-1} \pm \widetilde{f}_{n-2}|$. Here, traveling right, left, left, is equivalent to not moving at all. In both the linear and non-linear cases we may remove the last left which gets us to this point of repetition, i.e., left children do not have left children, as we have seen in Proposition 5.5.

What happens in the case of Figure 5.6(c) in terms of matrices $A, B = \left(\begin{smallmatrix} 0 & 1 \\ 1 & \pm 1 \end{smallmatrix}\right)$? We have that $AABA \neq -B$ and $BABA \neq -A$, contrary to what might be expected, because we are again switching and negating children. (Remember we must take the reverse product matrix because we are using left multiplication, and our matrices represent moving right, left, right in the tree, which is the combination that leads to cancelation in R_1.) Taking a closer look at what is happening, we have

$$AABA\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ b-a \end{pmatrix}, \quad A\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ a+b \end{pmatrix},$$

$$BABA\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ -a-b \end{pmatrix}, \qquad B\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ a-b \end{pmatrix}.$$

We see that the children are switched and negated as required, but that the first entry in the vectors is not negated. This is the reason for our matrix equations not holding in this case. Further, it can be shown that traveling left, right, left in the tree R_1 produces repetition as well, and in this case the children are switched but not negated. So in contrast to the trees R and R_2, in which we cannot have left, left, the tree R_1 gives a new characterization, where we cannot have right, left, right or left, right left. In terms of coefficient cycles, this means that we can not have the patterns $(+-+)$ or $(-+-)$. Recall, from Proposition 5.6 that in absolute value, all three trees are comprised of the same set of sequences and so there are no differences in growth rates. In [1], Alperin uses a similar construction of words on two letters using the group $\mathrm{PSL}(2,\mathbb{Z})$. He views it as a free product of the cyclic group of order 2 generated by the image of S, and the cyclic group of order 3 generated by the image of $ST = R$. Recall that S and T are the generators of $\mathrm{SL}(2,\mathbb{Z})$ defined in Theorem 2.1, and they satisfy the given relations $S^2 = (ST)^3 = I$ in $\mathrm{PSL}(2,\mathbb{Z})$. Non-identity elements can then be described as unique strings of S's and R's, which cannot contain S^2 or R^3.

5.3 The Decomposition of Tree T

Before proceeding, the following definition will be useful (see McCliment [54]).

Definition 5.1. A *multiset* $\mathcal{M}$ is a generalization of a set, where elements may appear more than once. The *multiset sum* of two multisets $\mathcal{M}$ and $\mathcal{N}$, denoted $\mathcal{M} \uplus \mathcal{N}$, is defined as the multiset which contains all elements of $\mathcal{M}$ and all elements of $\mathcal{N}$. We therefore have that $|\mathcal{M} \uplus \mathcal{N}| = |\mathcal{M}| + |\mathcal{N}|$.

We have seen that τ_n and ρ_n are multisets because a given row in T or R may contain an element multiple times. Keep in mind that the elements in a multiset are unordered, unlike the elements in a string. For convenience we will write the elements in the order they appear in the tree. This becomes useful when taking multiset sums, as we are not required to simply concatenate the sets as with strings. We can also take a multiset sum of a multiset with itself. We define $2\mathcal{M} = \mathcal{M} \uplus \mathcal{M}$ to be the multiset containing two copies of every element in $\mathcal{M}$, including the repetitions. In general we can write $n\mathcal{M} = \mathcal{M} \uplus \mathcal{M} \uplus \cdots \uplus \mathcal{M}$, where we have taken the multiset sum of n copies of $\mathcal{M}$.

In T, we have that the children of row ρ_n are given by ρ_{n+1} as well as the elements in τ_{n+1} which are left children of the left children in ρ_n, as these are the elements that are removed when forming ρ_{n+1}. Denote by $c_T(\rho_n)$, the multiset of children in T of row ρ_n in R. Also, the following notation from Rittaud will be needed. Denote by ρ_n^- and ρ_n^+ the multisubsets consisting of left and right nodes, respectively, of the multiset ρ_n. Let $c_R^+(\rho_n)$ and $c_R^-(\rho_n)$ be the multisets of left and right children, respectively, of ρ_n in R.

Our aim is to write each row τ_n in the tree T as a multiset sum of rows ρ_j in tree R, where $j \leq n$. We will then be able to write the growth rate of the expected value of nodes in τ_n in terms of the growth rate of the expected value of nodes in ρ_n, the latter of which is easy to deduce. We will give a different proof of this than Rittaud [64] does, and consider both methods later in the chapter. Given row ρ_n, we want to deduce what elements would need to be added in order to form τ_n. In particular we look for missing descendants of parents, grandparents and so on. Some new notation will be useful here. Let $m(\rho_n)$ denote the multiset of elements of τ_n which are descendants of ρ_{n-1} (the parents) and are missing from ρ_n. Similarly let $m^2(\rho_n)$ denote the multiset of elements of τ_n which are descendants of ρ_{n-2} (the grandparents) and are missing from ρ_n. In general, let $m^i(\rho_n)$ denote the multiset of elements of τ_n which are descendants of ρ_{n-i}, i.e., the (great)$^{i-2}$ grandparents, missing from ρ_n. We can also write $m^0(\rho_n) = \rho_n$. Using this new notation, we can write the children of ρ_n in T, for $n \geq 1$, as

$$c_T(\rho_n) = \rho_{n+1} \uplus m(\rho_{n+1}). \tag{5.1}$$

Example 5.2. Consider the portion of the tree T given in Figure 5.7. For $n = 6$, we

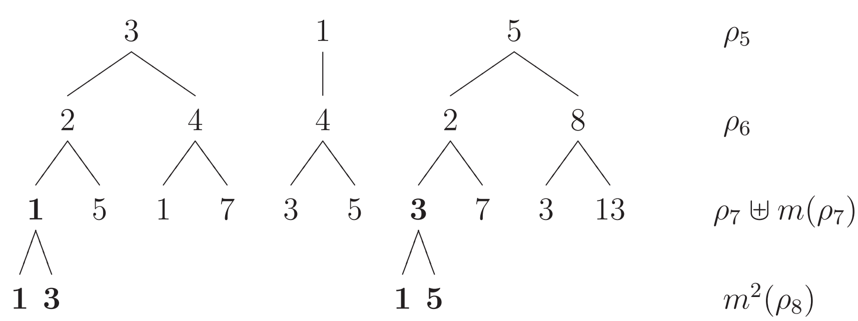


Figure 5.7: Illustration of $m(\rho_7)$ and $m^2(\rho_8)$.

have that by Equation (5.1), $c_T(\rho_6) = \rho_7 \uplus m(\rho_7)$. The children of ρ_6 are comprised of ρ_7 as well as the elements missing from τ_7 which are left children of left children in ρ_6. This multiset of missing elements, $m(\rho_7) = \{1, 3\}$, is bolded in Figure 5.7. Also, $m^2(\rho_8)$ is the multiset of elements missing from τ_8 which are the missing grandchildren of left children in ρ_6. This is the multiset of bolded elements $\{1, 3, 1, 5\}$.

For $n = 3, \ldots, 10$ explicit calculation of trees gives us Table 5.1. Note that for $n = 1$ and 2, the trees T and R both have the same entries. Each row τ_n in T is the multiset sum of ρ_n and all multisets of missing elements, i.e., we take the multiset sum of all the multisets corresponding to n in Table 5.1. Also, we have bolded those multisets ρ_j in $m^i(\rho_n)$ which come from the term $m^{i-1}(\rho_{n-3})$. We will see in Equation (5.4) why this term is of importance. The table contains some easily observable patterns, which we will prove for the general case. The following important characterization of repetition in T is found in Rittaud [64].

Corollary 5.1. *The elements missing from ρ_n, which come from the parents, form precisely the multiset ρ_{n-3}, i.e., $m(\rho_n) = \rho_{n-3}$, for $n \geq 5$. Further, we can write*

$$c_T(\rho_n) = \rho_{n+1} \uplus \rho_{n-2}, \tag{5.2}$$

for $n \geq 4$.

Proof: We know from Proposition 5.5 that the left child of a left node a is missing in R, and this child is precisely the grandparent of a. Therefore the missing children in ρ_n which come from left node parents belong to the multiset ρ_{n-3}. Now, any node c in ρ_{n-3} has eight great grandchildren in T but only one of those has form $m(\rho_n)$. The reason is that two of these great grandchildren are left children of left nodes, but one of them is also the grandchild of a left node, so has been previously removed. Therefore the missing left children in row ρ_n of the elements in row ρ_{n-1} are precisely the elements in ρ_{n-3}. Note that for $n = 4$ there are no missing elements from the parent row. Finally, Equation (5.2) is a direct consequence of Equation (5.1). ◇

Table 5.1: Breakdown of rows τ_n into ρ_n and missing elements, $3 \le n \le 10$.

n	$m^i(\rho_n)$	ρ_n and missing elements
3	ρ_3	2
	$m(\rho_3)$	0
4	ρ_4	1,3
	$m(\rho_4)$	-
	$m^2(\rho_4)$	1,1
5	ρ_5	3,1,5
	$m(\rho_5)$	1
	$m^2(\rho_5)$	-
	$m^3(\rho_5)$	1,1,1,1
6	ρ_6	2,4,4,2,8
	$m(\rho_6)$	2
	$m^2(\rho_6)$	0,2
	$m^3(\rho_6)$	-
	$m^4(\rho_6)$	{0,2} (×4)
7	ρ_7	5,1,7,3,5,7,3,13
	$m(\rho_7)$	1,3
	$m^2(\rho_7)$	1,3
	$m^3(\rho_7)$	**1,1**,1,3
	$m^4(\rho_7)$	-
	$m^5(\rho_7)$	{1,1,1,3} (×4)
8	ρ_8	3,7,5,3,11,7,1,9,5,9,11,5,21
	$m(\rho_8)$	3,1,5
	$m^2(\rho_8)$	**1**,3,1,5
	$m^3(\rho_8)$	1,3,1,5
	$m^4(\rho_8)$	**1,1,1,1**,1,3,1,5
	$m^5(\rho_8)$	-
	$m^6(\rho_8)$	{1,1,1,1,1,3,1,5} (×4)

Continued on next page–

Table 5.1 – Continued from previous page		
n	$m^i(\rho_n)$	ρ_n and missing elements
9	ρ_9	8,2,12,4,6,10,4,18,4,10,6,4,14,12,2,16,8,14,18,8,34
	$m(\rho_9)$	2,4,4,2,8
	$m^2(\rho_9)$	2,4,**2**,4,2,8
	$m^3(\rho_9)$	**0,2**,2,4,2,4,2,8
	$m^4(\rho_9)$	0,2,2,4,2,4,2,8
	$m^5(\rho_9)$	{**0,2**} (×4) 0,2,2,4,2,4,2,8
	$m^6(\rho_9)$	-
	$m^7(\rho_9)$	{{0,2} (×4) 0,2,2,4,2,4,2,8} (×4)
10	ρ_{10}	5,11,9,5,19,9,1,11,7,13,15,7,29,11,3,17,5,7,13,5,23, 7,17,11,7,25,19,3,25,13,23,29,13,55
	$m(\rho_{10})$	5,1,7,3,5,7,3,13
	$m^2(\rho_{10})$	**1**,5,1,7,3,5,**3**,7,3,13
	$m^3(\rho_{10})$	1,5,1,7,**1,3**,3,5,3,7,3,13
	$m^4(\rho_{10})$	**1,1,1,3**,1,5,1,7,1,3,3,5,3,7,3,13
	$m^5(\rho_{10})$	1,1,1,3,1,5,1,7,1,3,3,5,3,7,3,13
	$m^6(\rho_{10})$	{**1,1,1,3**} (×4)1,1,1,3,1,5,1,7,1,3,3,5,3,7,3,13
	$m^7(\rho_{10})$	-
	$m^8(\rho_{10})$	{{1,1,1,3} (×4) 1,1,1,3,1,5,1,7,1,3,3,5,3,7,3,13} (×4)

Example 5.3. We can see from tree R that $\rho_6 = \{2, 4, 4, 2, 8\}$. Using Equation (5.2) we have that

$$\begin{aligned} c_T(\rho_6) &= \rho_7 \uplus \rho_4 \\ &= \{5, 1, 7, 3, 5, 7, 3, 13\} \uplus \{1, 3\}. \end{aligned}$$

Proposition 5.8. *Any row τ_n in T, for $n \geq 1$, is made up of a multiset sum of rows ρ_j in R, where $j \leq n$.*

Proof: The first few rows in the tree R are $\{1\}, \{1\}, \{2\}, \{1, 3\}, \{1, 3, 1, 5\}$ and the first few in T are $\{1\}, \{1\}, \{0, 2\}, \{1, 1, 1, 3\}, \{1, 1, 1, 1, 1, 3, 1, 5\}$. We will consider the rows $\{2\}$ and $\{0, 2\}$ to be equivalent because ultimately we will be taking the average value of the row entries, which is unaffected by extra zero terms. We can prove this result using induction, the initial cases being clear from the rows listed above. Suppose row τ_n is comprised of rows ρ_j where $j \leq n$. We have that for each ρ_j in τ_n,

$$c_T(\rho_j) = \rho_{j+1} \uplus \rho_{j-2},$$

by Equation (5.2). Therefore, row τ_{n+1} of T is also comprised of rows of R, completing the proof. ◇

Example 5.4. We can see from Table 5.1 that τ_7 is composed of ρ_7 together with 7 copies of $\rho_4 = \{1, 3\}$ and 10 copies of $\rho_1 = \{1\}$.

Note that the multiset of children in T of elements missing from ρ_n from ancestors i generations back is also the multiset of missing elements in ρ_{n+1} from ancestors $i+1$ generations back, i.e.,

$$c_T(m^i(\rho_n)) = m^{i+1}(\rho_{n+1}), \tag{5.3}$$

for $1 \leq i \leq n-2$ and $n \geq 3$.

Example 5.5. For $n = 7$ and $i = 3$, Equation (5.3) says that

$$c_T(m^3(\rho_7)) = m^4(\rho_8).$$

From Table 5.1 we see that this gives us

$$c_T(\{1, 1, 1, 3\}) = \{1, 1, 1, 1, 1, 3, 1, 5\},$$

as can be verified from the tree T.

We know from Corollary 5.1 that $m(\rho_n) = \rho_{n-3}$. It follows that the children of each of these multisets are also equal.

Corollary 5.2. *For any row ρ_n in tree R with $n \geq 5$, we have that*

$$c_T(m(\rho_n)) = c_T(\rho_{n-3}).$$

Proof: A node a in $m(\rho_n)$ is missing from the tree T because the edge (b, a) was repeated three rows back. Specifically a appeared in ρ_{n-3}, for $n \geq 5$ by Corollary 5.1. Since the parent of a is b in both cases, the children of a (in T) must be the same in both cases, and so $c_T(m(\rho_n)) = c_T(\rho_{n-3})$. $\diamond$

Proposition 5.9. *Let $n \geq 6$. The set of elements missing from ρ_n which are descendants in T of elements in row ρ_{n-i}, i.e., which come from the* (great)$^{i-2}$ *grandparents, can be written as the multiset sum*

$$m^i(\rho_n) = m^{i-1}(\rho_n) \uplus m^{i-1}(\rho_{n-3}), \tag{5.4}$$

for $2 \leq i \leq n-4$.

Proof: We will prove this result using induction on i. The initial case, for $i = 2$, can be written as

$$m^2(\rho_n) = m(\rho_n) \uplus m(\rho_{n-3}).$$

In other words, the elements missing from ρ_n from grandparents are precisely the elements missing from ρ_n from the parents along with the elements missing from ρ_{n-3} from the parents. By Equation (5.3), the elements missing from ρ_n from grandparents can be thought of as the multiset of children of elements missing from ρ_{n-1} from parents, i.e., the multiset of children of ρ_{n-4}, namely

$$m^2(\rho_n) = c_T m(\rho_{n-1})) = c_T(\rho_{n-4}),$$

for $n \geq 6$ by Corollary 5.2. The children in T of ρ_{n-4} are given by the next row in R, i.e., ρ_{n-3} (which by Corollary 5.1 is $m(\rho_n)$), as well as what is missing from ρ_{n-3} from the parents. This can be written as $m(\rho_n) \uplus m(\rho_{n-3})$, completing the proof for the initial case for $n \geq 6$.

Now suppose the statement holds for $i - 1$. By Equation (5.3) we have

$$m^i(\rho_n) = c_T(m^{i-1}(\rho_{n-1})),$$

for $2 \leq i \leq n-2$ and $n \geq 4$, and by our assumption,

$$\begin{aligned} m^i(\rho_n) &= c_T(m^{i-2}(\rho_{n-1}) \uplus m^{i-2}(\rho_{n-4})) \\ &= c_T(m^{i-2}(\rho_{n-1})) \uplus c_T(m^{i-2}(\rho_{n-4})). \end{aligned} \tag{5.5}$$

Now using Equation (5.3) again we can write

$$m^i(\rho_n) = m^{i-1}(\rho_n) \uplus m^{i-1}(\rho_{n-3}),$$

where conditions for the first term on the right-hand side are $3 \leq i \leq n-1$, $n \geq 4$, and conditions for the second term are $3 \leq i \leq n-4$, $n \geq 7$. This completes the induction for $3 \leq i \leq n-4$, with the following consideration. Note that when applying Equation (5.3) to the second term in the right-hand side of Equation (5.5), we must have $n \geq 7$, because the equation is true starting with row 3 in tree T, as this is when missing elements first appear. When $n = 6$, we must have $i = 2$, and this was already proven as the initial case. ◇

Example 5.6. For $n = 9$ and $i = 3$ we get

$$m^3(\rho_9) = m^2(\rho_9) \uplus m^2(\rho_6),$$
$$\{0, 2, 2, 4, 2, 4, 2, 8\} = \{2, 4, 2, 4, 2, 8\} \uplus \{0, 2\}.$$

Note that the bolded entries in $m^3(\rho_9)$ in Table 5.1 are those coming from the second term, $m^2(\rho_6)$.

We now consider what happens to values of i outside the range $2 \leq i \leq n-4$. We know from Corollary 5.1 that $m(\rho_n) = \rho_{n-3}$ for $i = 1$. We also have the following result.

Proposition 5.10. *For $i = n-3$ and $n \geq 4$ we have $m^i(\rho_n) = \{\}$, and for $i = n-2$ and $n \geq 5$ we have $m^i(\rho_n) = 4\tau_{n-3}$. Also, for $i = n-4$ and $n \geq 5$ we have $m^i(\rho_n) = \tau_{n-3}$. Furthermore, for $i = n-2$ and $n \geq 5$ we also have $m^i(\rho_n) = 4m^{i-2}(\rho_n)$ for $i = n-2$.*

Proof: In the case of $i = n-3$, we are looking for elements missing from row n which come from row 3. In the tree R we have that $\rho_3 = \{2\}$, which has left and right children 1 and 3 respectively, and hence no missing children. Therefore there are no children missing from ρ_n which are a result of ρ_3. We can observe this for rows past row 3, i.e., $n \geq 4$.

For $i = n - 2$, we are looking for elements missing from ρ_n which come from $\rho_2 = \{1\}$. In R, this node 1 has right child 2, but no left child. In T we see that this deleted left child was 0. The descendants of this 0 node form the entire left half of the tree T. It is easy to see that the left half of row τ_n is in fact $4\tau_{n-3}$ because the tree T is contained four times in the left-hand side of T itself, starting in row 4 with $\{1, 1\}$. It is only in row 5 where we have nodes $\{1, 1, 1, 1\}$, i.e., τ_2 is contained four times. Therefore, $m^i(\rho_n) = 4\tau_{n-3}$ for $i = n - 2$ for $n \geq 5$.

For $i = n - 4$, we are looking for elements missing from ρ_n which come from $\rho_4 = \{1, 3\}$. The node 1 is a left node, with children 1 and 3. The child 1 (τ_2) in row 5 is removed, and with it, so is a (almost complete) copy of T as can be seen from tree T. We therefore have $m^i(\rho_n) = \tau_{n-3}$ for $n \geq 5$. We have this relation in addition to the one given in Proposition 5.9 for $i = n - 4$. Note that combining statements of the theorem gives $m^i(\rho_n) = 4m^{i-2}(\rho_n)$ for $i = n - 2$ and $n \geq 5$, which explains the repetition in the last three rows for each value of $n \geq 5$ in Table 5.1. $\diamond$

We have now proven the patterns needed to extend Table 5.1 to any value of n.

Example 5.7. Consider τ_7. Proposition 5.10 tells us that $m^4(\rho_7) = \{\}$, which can be seen from Table 5.1. According to Proposition 5.9, $m^4(\rho_7)$ reappears in the equation

$$m^5(\rho_{10}) = m^4(\rho_{10}) \uplus m^4(\rho_7).$$

Because $m^4(\rho_7)$ is the empty set, we have that

$$m^5(\rho_{10}) = m^4(\rho_{10}) = \{1, 1, 1, 3, 1, 5, 1, 7, 1, 3, 3, 5, 3, 7, 3, 13\}.$$

Corollary 5.3. *For $n \geq 7$ we have*

$$m^{n-5}(\rho_n) = m^{n-6}(\rho_n)$$

which is evident as a repeated row in Table 5.1.

Proof: Let $i = n - 5$. From Proposition 5.9 we have that

$$m^i(\rho_n) = m^{i-1}(\rho_n) \uplus m^{i-1}(\rho_{n-3}),$$

for $n \geq 7$. Also, from Proposition 5.10 we have that $m^{n-3}(\rho_n) = \{\}$ for $n \geq 4$, which means that $m^{i-1}(\rho_{n-3}) = m^{n-6}(\rho_{n-3}) = \{\}$ for $n \geq 7$. Therefore, we can rewrite the above Equation as

$$m^i(\rho_n) = m^{i-1}(\rho_n),$$
$$m^{n-5}(\rho_n) = m^{n-6}(\rho_n),$$

completing the proof. $\diamond$

We can make a couple of remarks about the above results. Proposition 5.8 tells us that a row τ_n is made up of rows ρ_j where $j \leq n$. Furthermore, we can add that we must have $j \equiv n \pmod 3$ because Proposition 5.9 and Corollary 5.1 tell us that elements missing from row ρ_n depend on elements missing from rows ρ_n and ρ_{n-3} for $1 \leq i \leq n-4$. In turn, elements missing from row ρ_{n-3} depend on elements missing from rows ρ_{n-3} and ρ_{n-6} and so on. We can conclude that τ_n is a multiset union of some combination of rows $\rho_n, \rho_{n-3}, \rho_{n-6}, \ldots$. This also holds for $i = n-2$ with $n \geq 5$, as we have seen in Proposition 5.10 that $m^i(\rho_n) = 4m^{i-2}(\rho_n)$. Recall that Example 5.4 showed that τ_7 is composed of copies of ρ_7, ρ_4 and ρ_1.

Secondly, we can remark that since $m(\rho_n) = \rho_{n-3}$ by Corollary 5.1, and each multiset $m^i(\rho_n)$ contains the multiset $m^{i-1}(\rho_n)$ by Proposition 5.9, every multiset $m^i(\rho_n)$ for $1 \leq i \leq n-4$ and $n \geq 5$ must contain ρ_{n-3}. This is clearly observed in Table 5.1. Again this holds also for the case $i = n-2$, but not for the empty row $i = n-3$. Also, by induction, it is easy to see that not only every row τ_n, but every multiset $m^i(\rho_n)$ inside τ_n (except the empty row), is made up of multisets ρ_j where $j \leq n$. For example, in τ_{10},

$$m^4(\rho_{10}) = \rho_7 \uplus 3\rho_4 \uplus 2\rho_1.$$

The following theorem gives us the critical relation in our determination of the growth rate of the expected value of the n^{th} term in a random Fibonacci sequence.

Theorem 5.1. *Let $t(n,k)$ denote the number of copies of ρ_{n-3k} in row τ_n of T, where $n \geq 1$ and $0 \leq k \leq \lfloor \frac{n-1}{3} \rfloor$. Then we have the relation*

$$t(n,k) = t(n-1,k) + t(n-1,k-1), \tag{5.6}$$

where $n \geq 5$, $k \geq 1$, *with the exception of* $t(3k+1,k)$ *and* $t(3k+2,k)$ *(in which case* $k = \frac{n-1}{3}$ *and* $k = \frac{n-2}{3}$ *respectively). If* $k = 0$ *we have* $t(n,0) = 1$ *for* $n \geq 1$, *and if* $k = 1$ *we have* $t(n,1) = n$, *for* $n \geq 5$. *In the exceptional cases we have* $t(3k+1,k) = 2t(3k-1,k-1)$ *and* $t(3k+2,k) = 2t(3k+1,k) + t(3k+1,k-1)$ *for* $k \geq 1$.

Proof: Suppose we want to count the number of copies of $\rho_j = \rho_{n-3k}$ in τ_n, i.e., $t(n,k)$. The fact that $j \geq 1$ implies that $k \leq \left\lfloor \frac{n-1}{3} \right\rfloor$. By Equation (5.2), there are two ways a multiset ρ_j can arise in τ_n. We can write

$$c_T(\rho_{j-1}) = \rho_j \uplus \rho_{j-3},$$

for $j \geq 5$, which tells us that each ρ_{j-1} in τ_{n-1} gives rise to a ρ_j in τ_n. More generally, by Equation (5.1) we have that for $j \geq 2$ the children of ρ_{j-1} in τ_{n-1} are contained in τ_n. We also have

$$c_T(\rho_{j+2}) = \rho_{j+3} \uplus \rho_j, \tag{5.7}$$

for $j \geq 2$, which tells us that each ρ_{j+2} in τ_{n-1} gives rise to a ρ_j in τ_n. Comparing indices, the number of copies of ρ_{j-1} in τ_{n-1} and the number of copies of ρ_{j+2} in τ_{n-1} are the terms $t(n-1,k)$ and $t(n-1,k-1)$ respectively. Note that for $j = n-3k \geq 2$ and $k \geq 1$ we must have $n \geq 5$. When $k = 0$, $t(n,0)$ is counting the number of copies of ρ_n in τ_n. We know that τ_n is made up of exactly one copy of ρ_n plus missing elements, so $t(n,0) = 1$ for all values of n.

The exceptions can be explained in terms of the tree T. The expression $t(3k+1,k) = 2t(3k-1,k-1)$ can be restated as: the number of multisets ρ_1 in τ_n is equal to twice the number of multisets ρ_2 in τ_{n-2}. (Note that for $n = 3k+1$, we have $3k-1 = n-2$.) Recall that $\rho_1 = \rho_2 = \{1\}$. Any time a ρ_2 appears in the tree it must have $\rho_3 = \{2\}$ as a child and therefore have parent 1. (See Figure 5.8.) The node 0, which is also a child of ρ_2 in T, then has $\{1,1\}$ as children. Each of these 1's has children $\{1,1\}$ and is therefore itself equal to a copy of ρ_1. We can conclude that any ρ_2 appearing in the tree has two multisets ρ_1 as grandchildren. Similarly any ρ_1 appearing in τ_n for $n \geq 4$ must belong to a pair with grandparent ρ_2. This gives us the desired relation for $n \geq 4$, i.e., $k \geq 1$. This is an exception to Equation (5.6)

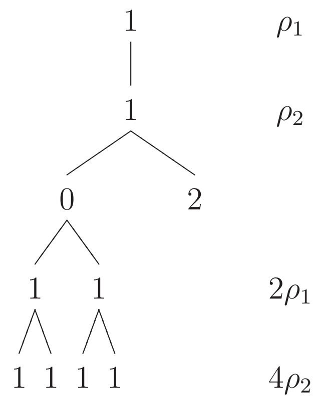


Figure 5.8: Occurrence of ρ_1 and ρ_2 in T.

because when $j = 1$ we cannot count the number of multisets ρ_j in τ_n by counting multisets ρ_{j-1} in τ_{n-1}.

The second exception is given by $t(3k + 2, k) = 2t(3k + 1, k) + t(3k + 1, k - 1)$, and can be restated as: the number of multisets ρ_2 in τ_n is equal to twice the number of multisets ρ_1 in τ_{n-1} plus the number of multisets $\rho_4 = \{1, 3\}$ in τ_{n-1}. According to Equation (5.6) there are two ways ρ_j can occur; for ρ_2, the first is that it follows ρ_1. Secondly, by Equation (5.7) with $j = 2$, we could have ρ_2 appear as a child of ρ_4. The number of times the latter occurs, i.e., the number of times ρ_4 occurs in τ_{n-1}, is $t(3k + 1, k - 1)$, which we could write as $t(n - 1, k - 1)$ for $n = 3k + 2$. We can see from the tree T that we have $n \geq 5$, implying $k \geq 1$. The number of multisets ρ_2 appearing as a consequence of ρ_1, however, is not simply the term $t(n - 1, k) = t(3k + 1, k)$ given in Equation (5.6). We have seen above that each ρ_1 appearing in τ_4 or higher has children $\{1, 1\}$, which are each multisets ρ_2. Therefore, the number of sets ρ_2 is doubled to give $2t(3k + 1, k)$, and altogether we obtain $t(3k + 2, k) = 2t(3k + 1, k) + t(3k + 1, k - 1)$ for $n \geq 5$, i.e., $k \geq 1$. This case is an exception because if we tried to follow our tree backwards to the initial nodes, we would see that before the first node 1, we should have a 0, implying that the children of 1 should be $\{1, 1\}$. This is not the case because we have fixed the second row, ρ_2, to be a single 1.

Lastly, we must show that $t(n, 1) = n$ for $n \geq 5$. Letting $k = 1$ in $t(3k + 1, k) = 2t(3k - 1, k - 1)$ gives $t(4, 1) = 2t(2, 0) = 2$, using the fact that $t(n, 0) = 1$. Now letting

$k = 1$ in $t(3k+2, k) = 2t(3k+1, k) + t(3k+1, k-1)$ gives $t(5, 1) = 2t(4, 1) + t(4, 0) = 2(2) + 1 = 5$. We can now induct on $t(n, k) = t(n-1, k) + t(n-1, k-1)$ for $k = 1$. Suppose $t(n, 1) = n$ holds. We then have $t(n+1, 1) = t(n, 1) + t(n, 0) = n + 1$, as required. $\diamond$

Example 5.8. As an example of the general case, consider $t(7, 1)$, i.e., the number of multisets $\rho_4 = \{1, 3\}$ appearing in τ_7. The above corollary tells us that $t(7, 1) = t(6, 1) + t(6, 0)$. In other words $t(7, 1)$ equals the number of multisets $\rho_3 = \{2\}$ in τ_6 plus the number of multisets $\rho_6 = \{2, 4, 4, 2, 8\}$ in τ_6. Counting in Table 5.1 we get $7 = 6 + 1$.

Example 5.9. As an example of the second exception, consider $t(8, 2)$, i.e., the number of multisets $\rho_2 = \{1\}$ appearing in τ_8. We have the relation $t(8, 2) = 2t(7, 2) + t(7, 1)$ from Theorem 5.1. In other words, $t(8, 2)$ equals twice the number of multisets $\rho_1 = \{1\}$ in τ_7 plus the number of multisets $\rho_4 = \{1, 3\}$ in τ_7. Counting again in Table 5.1 gives $27 = 2(10) + 7$.

Note that when counting multisets $\rho_3 = \{2\}$ in T, we ignore the 0 term that is often paired with the 2. We do not need to take the 0 terms into consideration when writing τ_n as multiset union of the ρ_j because as mentioned earlier, our focus will be on finding sums of rows.

The relations in Theorem 5.1 allow us to generate Table 5.2, of $t(n, k)$ values. In Table 5.2 we see that k increases by 1 for every third increase in n. Corner numbers occur when $n \equiv 1 \pmod 3$, in which case $k = \frac{n-1}{3}$. We will use the notation $t(k) = t\left(n, \frac{n-1}{3}\right) = t(3k+1, k)$ because of their special nature. The two entries below the corners, $t\left(n+1, \frac{n-2}{3}\right)$ and $t\left(n+2, \frac{n-3}{3}\right)$ are the final entries in their respective rows. We have that $t(0) = 1$ according to the table (i.e., one copy of ρ_n appears in τ_n), but the patterns suggest that this value should in fact be $\frac{1}{2}$. If we had a column of 0's previous to the column of 1's, the expression $t(3k+2, k) = 2t(3k+1, k) + t(3k+1, k-1)$ in Theorem 5.1 would require $t(0) = \frac{1}{2}$. For future calculations, we will need this new value of $t(0)$.

Note that Table 5.2 bears some resemblance to Pascal's triangle, particularly the relation among entries, which in our case is given by Equation (5.6). There are many

n \ k	0	1	2	3	4	5	6	7
1	1							
2	1							
3	1							
4	1	2						
5	1	5						
6	1	6						
7	1	7	10					
8	1	8	27					
9	1	9	35					
10	1	10	44	54				
11	1	11	54	152				
12	1	12	65	206				
13	1	13	77	271	304			
14	1	14	90	348	879			
15	1	15	104	438	1227			
16	1	16	119	542	1665	1758		
17	1	17	135	661	2207	5181		
18	1	18	152	796	2868	7388		
19	1	19	170	948	3664	10256	10362	
20	1	20	189	1118	4612	13920	30980	
21	1	21	209	1307	5730	18532	44900	
22	1	22	230	1516	7037	24262	63432	61960
23	1	23	252	1746	8553	31299	87694	187352
24	1	24	275	1998	10299	39852	118993	275046

Table 5.2: Values of $t(n,k)$ for $n \leq 24$, $k \leq 7$.

triangles with similar properties, for example Catalan's triangle [67, A009766], the entries of which are given by the expression

$$c(n,k) := \frac{(n+k)!(n-k+1)}{k!(n+1)!}.$$

We will soon see that an expression for $t(n,k)$ is far more complicated.

5.4 Sums of Rows of R

We will denote the sums of the elements of the multisets ρ_n and τ_n by $S(\rho_n)$ and $S(\tau_n)$ respectively. The results in this section are due to Rittaud [64]. As with the properties of the tree R, we have included proofs of these results to give a deeper understanding of the tree and the important role the it plays in this thesis. The following important property counts the number of nodes per row of R.

Proposition 5.11. *The number of nodes in row n of the tree R is given by $|\rho_n| = F_{n-1}$ for $n \geq 2$, and the numbers of left and right nodes are given by $|\rho_n^-| = F_{n-3}$ and $|\rho_n^+| = F_{n-2}$ respectively, for $n \geq 3$, where F_n is the n^{th} Fibonacci number.*

Proof: We can see by looking at Figure 5.2 that $|\rho_1| = |\rho_2| = |\rho_3| = 1$ and $|\rho_4| = 2$. Now suppose that $|\rho_n^-| = F_{n-3}$ and and $|\rho_n^+| = F_{n-2}$ (implying $|\rho_n| = F_{n-1}$). We want to show that this property holds for $n+1$. We know from Proposition 5.5 that every element of ρ_n has a right child, so the number of right children in row $n+1$ is

$$|c_R^+(\rho_n)| = |\rho_{n+1}^+| = |\rho_n| = F_{n-1}.$$

Also, we know that only right children (and all right children) have left children and so the number of left children in row $n+1$ is

$$|c_R^-(\rho_n)| = |\rho_{n+1}^-| = |\rho_n^+| = F_{n-2}.$$

Therefore the total number of nodes in row $n+1$ is

$$|\rho_{n+1}^+| + |\rho_{n+1}^-| = F_{n-1} + F_{n-2} = F_n,$$

completing the induction ◇

The following relation among sums of rows is the key to their growth rate. The proof differs from that given by Rittaud.

Lemma 5.1. *The sum of the nodes in row n of tree R is given by $S(\rho_1) = S(\rho_2) = 1$, $S(\rho_3) = 2$, $S(\rho_4) = 4$, and for $n \geq 5$:*

$$S(\rho_n) = 2S(\rho_{n-1}) + S(\rho_{n-3}).$$

Proof: The first few values of $S(\rho_n)$ are easily verifiable from Figure 5.2. We can write ρ_n as the multiset of children of elements in ρ_{n-1}, and summing gives

$$S(\rho_n) = S(c_R(\rho_{n-1})).$$

Now splitting $c_R(\rho_{n-1})$ into children of left nodes and children of right nodes gives

$$S(\rho_n) = S(c_R(\rho_{n-1}^+)) + S(c_R(\rho_{n-1}^-)).$$

Notice that $S(c_R(\rho_n^+)) = 2S(\rho_n^+)$, i.e., the sum of the children of a multiset of right nodes is equal to two times the sum of the right nodes themselves. This is clear from Figure 5.3, where the children of right child b are $b+c$ and $b-c$. Summing we get $2b$. We can therefore write the above equation as

$$S(\rho_n) = 2S(\rho_{n-1}^+) + S(c_R(\rho_{n-1}^-)). \tag{5.8}$$

We must now deal with the term $S(c_R(\rho_{n-1}^-))$, i.e., the sum of the children of left nodes in ρ_{n-1}. In Figure 5.3, a is a general left node with right child $a+b$. Note that a's grandparent is $b-a$ and so $a+b$ can be written as $a+b = 2(a)+(b-a)$, i.e., twice its parent plus its great grandparent. This is due to the fact that $a+b$ is a right child and so can be written as the sum of its parent a and grandparent b. The node b must be a right node, because its child a is a left node, and we know from Proposition 5.5 that left nodes cannot have left children. Therefore the grandparent is the sum of its two predecessors, namely $b-a$ and its parent a (recall from Proposition 5.3 that a must be the absolute value of the difference between $b-a$ and and b). Therefore $a+b = (a)+(b) = (a)+((b-a)+a)$. We can write the fact that the child of a left node is written as twice its parent plus its great grandparent, as

$$S(c_R(\rho_n^-)) = 2S(\rho_n^-) + S(\rho_{n-2}). \tag{5.9}$$

The multiset of grandparents of ρ_n^- is ρ_{n-2}, since each element in ρ_{n-2} has one left grandchild, the parent of which is the element's right child.

We can substitute Equation (5.9) into Equation (5.8) to give

$$\begin{aligned} S(\rho_n) &= 2S(\rho_{n-1}^+) + 2S(\rho_{n-1}^-) + S(\rho_{n-3}) \\ &= 2S(\rho_{n-1}) + S(\rho_{n-3}), \end{aligned}$$

completing the proof. $\diamond$

The sequence $\{S(\rho_n)\}$ for $n \geq 2$ can be found in [67, A008998] and its first few terms are

$$1, 1, 2, 4, 9, 20, 44, 97, 214, \ldots.$$

We will now consider its growth rate.

Theorem 5.2. *The rate of growth of the sum of elements in the n^{th} row of the tree R is given by*

$$\lim_{n\to\infty} \frac{S(\rho_n)}{S(\rho_{n-1})} = \alpha,$$

where α is the real root of $\alpha^3 - 2\alpha^2 - 1 = 0$.

Proof: We have seen in Lemma 5.1 that the sum of elements in row n of the tree R can be written as

$$S(\rho_n) = 2S(\rho_{n-1}) + S(\rho_{n-3}),$$

for $n \geq 5$. By the theory of linear recurrences, we deduce that the growth rate of terms $S(\rho_n)$ is given by the dominant root of the equation

$$x^3 - 2x^2 - 1 = 0.$$

Solving, we see that these roots are approximately 2.205569431 and $-0.1027847152 \pm 0.6654569515i$. The complex roots have moduli .6733480912, and so the dominant root, $\alpha = 2.2055\ldots$, is our growth rate. $\diamond$

Corollary 5.4. *The growth rate of the expected value of an element in the n^{th} row of R is given by α/ϕ, where ϕ is the golden ratio,* $1.618033\ldots$.

Proof: Instead of looking at the sum of entries in row n of R, we want to look at the average value. The sum is denoted by $S(\rho_n)$, and by Proposition 5.11 we know that there are F_{n-1} terms in this row. Therefore the average value of an entry in row n is $S(\rho_n)/F_{n-1}$. To find the growth rate of this average value, we can take the limit of the ratio:

$$\lim_{n\to\infty} \frac{S(\rho_n)/F_{n-1}}{S(\rho_{n-1})/F_{n-2}} = \lim_{n\to\infty} \frac{S(\rho_n)}{S(\rho_{n-1})} \cdot \lim_{n\to\infty} \frac{1}{F_{n-1}/F_{n-2}} = \alpha \cdot \frac{1}{\phi} = 1.363116873\ldots,$$

which by Theorem 5.2 gives us the value required. $\diamond$

5.5 Sums of Rows of T

We will now consider the sums of the rows τ_n in the tree T and use the fact that we can write this expression as a linear combination of sums of rows ρ_j in R for $j \leq n$ to find the growth rate of the expected value of $S(\tau_n)$. The first few values of $S(\tau_n)$, for $n \geq 1$, are

$$1, 1, 2, 6, 14, 32, 82, 196, 464, 1142, \ldots,$$

as can be found in [67, A083404]. No formula or recurrence for this sequence is given.

Before continuing, we will need some information about a useful type of mean called a Cesàro mean. The following definition and proposition are found in Hardy [36, p. 96–102], in greater generality.

Definition 5.2. Given a sequence $\{a_n\}$, its *Cesàro means* are given by the sequence $\{c_n\}$, where

$$c_n := \frac{1}{n}\sum_{i=1}^{n} a_i.$$

Proposition 5.12. *Given a sequence $\{a_n\}$, and its Cesàro means $\{c_n\}$, if*

$$\lim_{n\to\infty} a_n = A$$

then

$$\lim_{n\to\infty} c_n = A.$$

Theorem 5.3. *The limit of the ratios of the sums $S(\tau_n)$ can be written as*

$$\lim_{n\to\infty}\frac{S(\tau_n)}{S(\tau_{n-1})} = \begin{cases} \lim_{n\to\infty} \dfrac{\sum_{k=0}^{\frac{n-3}{3}} t(n,k)\frac{1}{\alpha^{3k-1}}}{\sum_{k=0}^{\frac{n-3}{3}} t(n-1,k)\frac{1}{\alpha^{3k}}}, & \text{if } n \equiv 0 \pmod 3; \\ \lim_{n\to\infty} \dfrac{\sum_{k=0}^{\frac{n-4}{3}} t(n,k)\frac{1}{\alpha^{3k-1}} + t\left(\frac{n-1}{3}\right)\frac{1}{\alpha^{n-2}}}{\sum_{k=0}^{\frac{n-4}{3}} t(n-1,k)\frac{1}{\alpha^{3k}}}, & \text{if } n \equiv 1 \pmod 3; \\ \lim_{n\to\infty} \dfrac{\sum_{k=0}^{\frac{n-2}{3}} t(n,k)\frac{1}{\alpha^{3k-1}}}{\sum_{k=0}^{\frac{n-5}{3}} t(n-1,k)\frac{1}{\alpha^{3k}} + t\left(\frac{n-2}{3}\right)\frac{1}{\alpha^{n-2}}}, & \text{if } n \equiv 2 \pmod 3, \end{cases}$$

where $t(n,k)$ is the entry in the n^{th} row and k^{th} column of Table 5.2 and $t(k)$ is the k^{th} corner entry, where $t(k) = t\left(n, \frac{n-1}{3}\right)$.

Proof: We want an expression for $S(\tau_n)$, the sum of all elements in row n of the tree T. We have seen in Proposition 5.8 and Theorem 5.1 that $S(\tau_n)$ can be written as a linear combination of terms $S(\rho_j)$, the sums of rows of tree R with $j \leq n$, where coefficients are given by $t(n,k)$. Here $j = n - 3k$, $n \geq 1$ and $0 \leq k \leq \left\lfloor \frac{n-1}{3} \right\rfloor$. We are interested in writing $S(\tau_n)$ in terms of $S(\rho_j)$ because we have seen in Theorem 5.2 that the sums $S(\rho_j)$ grow at the rate α, where α is the positive root of the third degree polynomial $\alpha^3 - 2\alpha^2 - 1 = 0$.

We can write the linear combination for a given row τ_n as

$$\begin{aligned} S(\tau_n) =& t(n,0)S(\rho_n) + t(n,1)S(\rho_{n-3}) + t(n,2)S(\rho_{n-6}) + \cdots + t(n,k)S(\rho_{n-3k}) + \\ & \cdots + t\left(n, \left\lfloor \tfrac{n-1}{3} \right\rfloor\right) S\left(\rho_{n-3\left\lfloor \frac{n-1}{3} \right\rfloor}\right). \end{aligned}$$

For specific values of $n \pmod 3$ we have

$$S(\tau_n) = \begin{cases} t(n,0)S(\rho_n) + \cdots + t(n,k)S(\rho_{n-3k}) + \cdots + t\left(n, \frac{n-3}{3}\right) S(\rho_3), & \text{if } n \equiv 0; \\ t(n,0)S(\rho_n) + \cdots + t(n,k)S(\rho_{n-3k}) + \cdots + t\left(n, \frac{n-1}{3}\right) S(\rho_1), & \text{if } n \equiv 1; \\ t(n,0)S(\rho_n) + \cdots + t(n,k)S(\rho_{n-3k}) + \cdots + t\left(n, \frac{n-2}{3}\right) S(\rho_2), & \text{if } n \equiv 2. \end{cases} \tag{5.10}$$

We are interested in the ratio of sums. For instance, for $n = 12$ we have

$$\frac{S(\tau_{12})}{S(\tau_{11})} = \frac{t(12,0)S(\rho_{12}) + t(12,1)S(\rho_9) + t(12,2)S(\rho_6) + t(12,3)S(\rho_3)}{t(11,0)S(\rho_{11}) + t(11,1)S(\rho_8) + t(11,2)S(\rho_5) + t(11,3)S(\rho_2)}.$$

In general we have the ratio $S(\tau_n)/S(\tau_{n-1})$, for values of $n \pmod 3$. For $n \equiv 0 \pmod 3$ we have the same number of terms in each sum and no corner terms, giving

$$\begin{aligned} S(\tau_n) &= t(n,0)S(\rho_n) + \cdots + t(n,k)S(\rho_{n-3k}) + \cdots + t\left(n, \tfrac{n-3}{3}\right) S(\rho_3), \\ S(\tau_{n-1}) &= t(n-1,0)S(\rho_{n-1}) + \cdots + t(n-1,k)S(\rho_{n-3k-1}) + \\ & \quad \cdots + t\left(n-1, \tfrac{n-3}{3}\right) S(\rho_2). \end{aligned}$$

For $n \equiv 1 \pmod 3$, the last term in the first sum is a corner term, and the first sum has one more term than the second sum because the corner number starts a new column in Table 5.2. We have

$$\begin{aligned} S(\tau_n) &= t(n,0)S(\rho_n) + \cdots + t(n,k)S(\rho_{n-3k}) + \cdots + t\left(n, \tfrac{n-4}{3}\right) S(\rho_4) + t\left(\tfrac{n-1}{3}\right) \\ S(\tau_{n-1}) &= t(n-1,0)S(\rho_{n-1}) + \cdots + t(n-1,k)S(\rho_{n-3k-1}) + \\ & \quad \cdots + t\left(n-1, \tfrac{n-4}{3}\right) S(\rho_3). \end{aligned}$$

Lastly, for $n \equiv 2 \pmod 3$, the last term in the second sum is a corner number and each sum has the same number of terms. We have

$$\begin{aligned} S(\tau_n) &= t(n,0)S(\rho_n) + \cdots + t(n,k)S(\rho_{n-3k}) + \cdots + t\left(n, \tfrac{n-2}{3}\right) S(\rho_2), \\ S(\tau_{n-1}) &= t(n-1,0)S(\rho_{n-1}) + \cdots + t(n-1,k)S(\rho_{n-3k-1}) + \\ &\quad \cdots + t\left(n-1, \tfrac{n-5}{3}\right) S(\rho_4) + t\left(\tfrac{n-2}{3}\right) S(\rho_1). \end{aligned}$$

Our goal is to write the growth rate of $S(\tau_n)$ in terms of the growth rate of $S(\rho_n)$. We can start by dividing both $S(\tau_n)$ and $S(\tau_{n-1})$ by $S(\rho_{n-1})$, for each value of $n \pmod 3$. For $n \equiv 0 \pmod 3$ this gives

$$\begin{aligned} S(\tau_n) &= t(n,0)\frac{S(\rho_n)}{S(\rho_{n-1})} + t(n,1)\frac{S(\rho_{n-3})}{S(\rho_{n-1})} + \cdots + t(n,k)\frac{S(\rho_{n-3k})}{S(\rho_{n-1})} + \\ &\quad \cdots + t\left(n, \tfrac{n-3}{3}\right) \frac{S(\rho_3)}{S(\rho_{n-1})}, \\ S(\tau_{n-1}) &= t(n-1,0) + t(n-1,1)\frac{S(\rho_{n-4})}{S(\rho_{n-1})} + \cdots + t(n-1,k)\frac{S(\rho_{n-3k-1})}{S(\rho_{n-1})} + \\ &\quad \cdots + t\left(n-1, \tfrac{n-3}{3}\right) \frac{S(\rho_2)}{S(\rho_{n-1})}. \end{aligned} \tag{5.11}$$

Now taking the limit of $S(\tau_n)/S(\tau_{n-1})$ as $n \to \infty$ and using the fact from Theorem 5.2 that

$$\lim_{n\to\infty} \frac{S(\rho_n)}{S(\rho_{n-1})} = \alpha,$$

we will be able to rewrite the above expressions.

Now consider the limit of $S(\rho_{n-3})/S(\rho_{n-1})$ as $n \to \infty$. We can write this as

$$\lim_{n\to\infty} \frac{S(\rho_{n-3})}{S(\rho_{n-1})} = \lim_{n\to\infty} \frac{1}{\frac{S(\rho_{n-1})}{S(\rho_{n-3})} \cdot \frac{S(\rho_{n-2})}{S(\rho_{n-2})}} = \frac{1}{\lim_{n\to\infty} \frac{S(\rho_{n-1})}{S(\rho_{n-2})} \cdot \lim_{n\to\infty} \frac{S(\rho_{n-2})}{S(\rho_{n-3})}} = \frac{1}{\alpha^2}. \tag{5.12}$$

In general we can show that for $k \geq 0$

$$\lim_{n\to\infty} \frac{S(\rho_{n-3k})}{S(\rho_{n-1})} = \frac{1}{\alpha^{3k-1}}. \tag{5.13}$$

For $k = 0$, the above limit is equal to α. Following Equation (5.12) we can write

$$\lim_{n\to\infty} \frac{S(\rho_{n-3k})}{S(\rho_{n-1})} = \lim_{n\to\infty} \frac{1}{\frac{S(\rho_{n-1})}{S(\rho_{n-3k})}} = \lim_{n\to\infty} \frac{1}{\frac{S(\rho_{n-1})}{S(\rho_{n-2})} \cdot \frac{S(\rho_{n-2})}{S(\rho_{n-3})} \cdots \frac{S(\rho_{n-3k+1})}{S(\rho_{n-3k})}}. \tag{5.14}$$

Note that for large values of k, the ratios in the denominator do not approximate α well. We need to use a different method to prove our result. Now, we can write the logarithm of the denominator as

$$\log\left(\frac{S(\rho_{n-1})}{S(\rho_{n-3k})}\right) = \log\left(\frac{S(\rho_{n-1})}{S(\rho_{n-2})}\right) + \log\left(\frac{S(\rho_{n-2})}{S(\rho_{n-3})}\right) + \cdots + \log\left(\frac{S(\rho_{n-3k+1})}{S(\rho_{n-3k})}\right).$$

Notice that when $k \geq 1$, the subscripts of the ρ terms in the numerators increase from $n-3k+1$ to $n-1$. There are $3k-1$ terms in the sum so we will divide both sides of the equation by this number to give

$$\frac{1}{3k-1}\log\left(\frac{S(\rho_{n-1})}{S(\rho_{n-3k})}\right) = \frac{1}{3k-1}\left(\log\left(\frac{S(\rho_{n-1})}{S(\rho_{n-2})}\right) + \cdots + \log\left(\frac{S(\rho_{n-3k+1})}{S(\rho_{n-3k})}\right)\right). \tag{5.15}$$

Since we know that $\lim_{n\to\infty}\log\left(\frac{S(\rho_{n-1})}{S(\rho_{n-2})}\right) = \log(\alpha)$, we have by Proposition 5.12 for Cesàro means that the limit of the right-hand side of Equation (5.15) is also equal to $\log(\alpha)$, and so

$$\lim_{n\to\infty}\frac{1}{3k-1}\log\left(\frac{S(\rho_{n-1})}{S(\rho_{n-3k})}\right) = \log(\alpha).$$

Rewriting gives

$$\lim_{n\to\infty}\left(\frac{S(\rho_{n-1})}{S(\rho_{n-3k})}\right) = \alpha^{3k-1},$$

and substituting into Equation (5.14) gives Equation (5.13) as required. Similarly, for terms in Equation (5.11) we have

$$\lim_{n\to\infty}\frac{S(\rho_{n-3k-1})}{S(\rho_{n-1})} = \frac{1}{\alpha^{3k}}.$$

We want to show that the limit of the ratio of sums of rows in T can be written as

$$\lim_{n\to\infty}\frac{S(\tau_n)}{S(\tau_{n-1})} = \lim_{n\to\infty}\frac{\widetilde{S}(\tau_n)}{\widetilde{S}(\tau_{n-1})}, \tag{5.16}$$

where

$$\begin{aligned}\widetilde{S}(\tau_n) &= t(n,0)\alpha + t(n,1)\frac{1}{\alpha^2} + t(n,2)\frac{1}{\alpha^5} + \\ &\qquad \cdots + t(n,k)\frac{1}{\alpha^{3k-1}} + \cdots + t\left(n, \frac{n-3}{3}\right)\frac{1}{\alpha^{n-4}}, \\ \widetilde{S}(\tau_{n-1}) &= t(n-1,0) + t(n-1,1)\frac{1}{\alpha^3} + t(n-1,2)\frac{1}{\alpha^6} + \cdots + t(n-1,k)\frac{1}{\alpha^{3k}} + \\ &\qquad \cdots + t\left(n-1, \frac{n-3}{3}\right)\frac{1}{\alpha^{n-3}}.\end{aligned}$$

We will first need to "normalize" the $t(n,k)$ terms in $\widetilde{S}(\tau_n)$ and $\widetilde{S}(\tau_{n-1})$ so that they are all less than 1. We can do this by dividing by $(n+1, \frac{n-4}{2})$, which is the term directly under $t\left(n, \frac{n-3}{3}\right)$ in Table 5.2, and is greater than all $t(n,k)$ terms in the expressions, by Equation (5.6). We will denote the normalized terms by $t_N(n,k)$ and sums containing these normalized terms by $S_N(\tau_n)$. Equation (5.16) can now be rewritten as

$$\lim_{n\to\infty} \frac{S_N(\tau_n)}{S_N(\tau_{n-1})} = \lim_{n\to\infty} \frac{\widetilde{S}_N(\tau_n)}{\widetilde{S}_N(\tau_{n-1})}. \tag{5.17}$$

If we consider the right-hand side of the above equation, we can split the limit into numerator and denominator for the following reason. We can write the numerator as

$$t_N(n,0)\alpha + t_N(n,1)\frac{1}{\alpha^2} + t_N(n,2)\frac{1}{\alpha^5} + \cdots + t_N(n,k)\frac{1}{\alpha^{3k-1}} + \cdots + t_N\left(n, \frac{n-3}{3}\right)\frac{1}{\alpha^{n-4}},$$

which converges as $n \to \infty$ by the comparison test because the sum $\sum_{n=0}^{\infty} \frac{1}{\alpha^{3k-1}}$ converges by the ratio test, and each $t_N(n,k)$ is less than 1. The same can be said for the denominator. We now want to show that the respective numerators and denominators in Equation (5.17) are equal in the limit. We can do this by showing

$$\lim_{n\to\infty} |S_N(\tau_n) - \widetilde{S}_N(\tau_n)| = 0. \tag{5.18}$$

Recall that $S_N(\tau_n)$ contains ratios of sums of the form $S(\rho_{n-3k})/S(\rho_{n-1})$. For $n \geq 2$, using the theory of linear recurrences, we can show that the sequence $\{S(\rho_n)\}$ satisfies the Binet-type formula

$$S(\rho_n) = a\alpha^n + b\beta^n + \overline{b\beta}^n,$$

where $\alpha \approx 2.20556943$, $\beta \approx 0.10278471 + 0.66545695i$, $a \approx 0.3821595$,
$b \approx -0.19107976 + 0.0885410i$. (These numbers are algebraic and we could also give them in terms of radicals, but this is not necessary.) Using some straightforward calculations involving geometric series, we can show that there exists a constant C, which is independent of k, such that for all n the following inequalities hold:

$$\left| \frac{S(\rho_n)}{S(\rho_{n-1})} - \alpha \right| \leq C \left(\frac{|\beta|}{\alpha} \right)^n,$$

$$\left|\frac{S(\rho_{n-3k})}{S(\rho_{n-1})} - \frac{1}{\alpha^{3k-1}}\right| \leq C\left(\frac{|\beta|}{\alpha}\right)^n,$$
$$\left|\frac{S(\rho_{n-3k-1})}{S(\rho_{n-1})} - \frac{1}{\alpha^{3k}}\right| \leq C\left(\frac{|\beta|}{\alpha}\right)^n.$$

Note that $|\beta|/\alpha \approx 0.30529444 < 1/3$. Using the above inequalities as well as the triangle inequality we can now write

$$\begin{aligned}
|S_N(\tau_n) - \widetilde{S}_N(\tau_n)| &\leq t_N(n,0)\left|\frac{S(\rho_n)}{S(\rho_{n-1})} - \alpha\right| + t_N(n,1)\left|\frac{S(\rho_{n-3})}{S(\rho_{n-1})} - \frac{1}{\alpha^2}\right| + \\
&\quad \cdots + t_N\left(n, \frac{n-3}{3}\right)\left|\frac{S(\rho_3)}{S(\rho_{n-1})} - \frac{1}{\alpha^{n-4}}\right| \\
&\leq t_N(n,0)\cdot C\left(\frac{|\beta|}{\alpha}\right)^n + t_N(n,1)\cdot C\left(\frac{|\beta|}{\alpha}\right)^n + \cdots + t_N\left(n, \frac{n-3}{3}\right)\cdot C\left(\frac{|\beta|}{\alpha}\right)^n \\
&\leq C\left(\frac{1}{3}\right)^n\left(t_N(n,0) + t_N(n,1) + \cdots + t_N\left(n, \frac{n-3}{3}\right)\right) \\
&\leq C\left(\frac{1}{3}\right)^n n.
\end{aligned}$$

The last step is true because there are $n/3$ terms in the sum, each of which is less than 1. Now, taking the limit gives Equation (5.18). The same can be said for the difference of denominators in Equation (5.17). We have therefore proven that Equation (5.17) holds. Multiplying back through by $(n+1, \frac{n-4}{2})$, i.e., "unnormalizing", gives us Equation (5.16), as required.

Similarly, for $n \equiv 1, 2 \pmod 3$, taking the limit of $S(\tau_n)/S(\tau_{n-1})$ gives

$$\lim_{n\to\infty}\frac{S(\tau_n)}{S(\tau_{n-1})} = \lim_{n\to\infty}\frac{\widetilde{S}(\tau_n)}{\widetilde{S}(\tau_{n-1})},$$

where for $n \equiv 1 \pmod 3$ we have

$$\begin{aligned}
\widetilde{S}(\tau_n) &= t(n,0)\alpha + t(n,1)\frac{1}{\alpha^2} + t(n,2)\frac{1}{\alpha^5} + \cdots + t(n,k)\frac{1}{\alpha^{3k-1}} + \\
&\qquad \cdots + t\left(n, \frac{n-4}{3}\right)\frac{1}{\alpha^{n-5}} + t\left(\frac{n-1}{3}\right)\frac{1}{\alpha^{n-2}}, \\
\widetilde{S}(\tau_{n-1}) &= t(n-1,0) + t(n-1,1)\frac{1}{\alpha^3} + t(n-1,2))\frac{1}{\alpha^6} + \cdots + t(n-1,k)\frac{1}{\alpha^{3k}} + \\
&\qquad \cdots + t\left(n-1, \frac{n-4}{3}\right)\frac{1}{\alpha^{n-4}},
\end{aligned}$$

and for $n \equiv 2 \pmod 3$,

$$\begin{aligned}
\widetilde{S}(\tau_n) &= t(n,0)\alpha + t(n,1)\frac{1}{\alpha^2} + t(n,2)\frac{1}{\alpha^5} + \\
&\quad \cdots + t(n,k)\frac{1}{\alpha^{3k-1}} + \cdots + t\left(n, \frac{n-2}{3}\right)\frac{1}{\alpha^{n-3}}, \\
\widetilde{S}(\tau_{n-1}) &= t(n-1,0) + t(n-1,1)\frac{1}{\alpha^3} + t(n-1,2)\frac{1}{\alpha^6} + \cdots + t(n-1,k)\frac{1}{\alpha^{3k}} + \\
&\quad \cdots + t\left(n-1, \frac{n-5}{3}\right)\frac{1}{\alpha^{n-5}} + t\left(\frac{n-2}{3}\right)\frac{1}{\alpha^{n-2}}.
\end{aligned}$$

This completes the proof. ◇

The following result is due to Rittaud [64]. His proof, which will be discussed in the next section, differs from the one presented below. We feel this, in conjunction with Theorem 5.3, is one of our main contributions.

Theorem 5.4. *The growth rate of the sum of entries in row τ_n of the tree T is given by*

$$\lim_{n\to\infty} \frac{S(\tau_n)}{S(\tau_{n-1})} = 2(\alpha - 1),$$

where α is the real root of $x^3 - 2x^2 - 1 = 0$.

Proof: We must consider three cases, depending on the value of $n \pmod 3$, as given in Theorem 5.3. First suppose that $n \equiv 0 \pmod 3$, which corresponds to the expression

$$\lim_{n\to\infty} \frac{S(\tau_n)}{S(\tau_{n-1})} = \lim_{n\to\infty} \frac{\sum_{k=0}^{\frac{n-3}{3}} t(n,k)\frac{1}{\alpha^{3k-1}}}{\sum_{k=0}^{\frac{n-3}{3}} t(n-1,k)\frac{1}{\alpha^{3k}}}. \tag{5.19}$$

We can now write this ratio as a single polynomial in α by setting it equal to

$$a_n(0)\alpha + a_n(1) + a_n(2)\frac{1}{\alpha} + a_n(3)\frac{1}{\alpha^2} + \cdots + a_n(j)\frac{1}{\alpha^{j-1}} + \cdots,$$

and cross-multiplying to solve for the coefficients $a_n(j)$. We obtain

$$\begin{aligned}
&t(n,0)\alpha + t(n,1)\frac{1}{\alpha^2} + t(n,2)\frac{1}{\alpha^5} + \cdots + t(n,i)\frac{1}{\alpha^{3i-1}} + \cdots \\
&= a_n(0)t(n-1,0)\alpha + a_n(0)t(n-1,1)\frac{1}{\alpha^2} + \cdots + a_n(0)t(n-1,i)\frac{1}{\alpha^{3i-1}} + \cdots \\
&+ a_n(1)t(n-1,0) + a_n(1)t(n-1,1)\frac{1}{\alpha^3} + \cdots + a_n(1)t(n-1,i)\frac{1}{\alpha^{3i}} + \cdots \\
&+ a_n(2)t(n-1,0)\frac{1}{\alpha} + a_n(2)t(n-1,1)\frac{1}{\alpha^4} + \cdots + a_n(2)t(n-1,i)\frac{1}{\alpha^{3i+1}} + \cdots \\
&+ a_n(3)t(n-1,0)\frac{1}{\alpha^2} + a_n(3)t(n-1,1)\frac{1}{\alpha^5} + \cdots + a_n(3)t(n-1,i)\frac{1}{\alpha^{3i+2}} + \cdots \\
&+ a_n(j)t(n-1,0)\frac{1}{\alpha^{j-1}} + a_n(j)t(n-1,1)\frac{1}{\alpha^{j+2}} + \cdots + a_n(j)t(n-1,i)\frac{1}{\alpha^{3i+j-1}} + \cdots \\
&+ \cdots
\end{aligned}$$

Rearranging terms gives

$$\begin{aligned}
&t(n,0)\alpha + t(n,1)\frac{1}{\alpha^2} + t(n,2)\frac{1}{\alpha^5} + \cdots + t(n,i)\frac{1}{\alpha^{3i-1}} + \cdots \\
&= a_n(0)t(n-1,0)\alpha + a_n(1)t(n-1,0) + a_n(2)t(n-1,0)\frac{1}{\alpha} \\
&+ [a_n(0)t(n-1,1) + a_n(3)t(n-1,0)]\frac{1}{\alpha^2} + \cdots \\
&+ [a_n(0)t(n-1,i) + a_n(3)t(n-1,i-1) + \cdots + a_n(3i)t(n-1,0)]\,\frac{1}{\alpha^{3i-1}} \\
&+ [a_n(1)t(n-1,i) + a_n(4)t(n-1,i-1) + \cdots + a_n(3i+1)t(n-1,0)]\,\frac{1}{\alpha^{3i}} \\
&+ [a_n(2)t(n-1,i) + a_n(5)t(n-1,i-1) + \cdots + a_n(3i+2)t(n-1,0)]\,\frac{1}{\alpha^{3i+1}} \\
&+ \cdots
\end{aligned}$$

Equating coefficients we see that $t(n,0) = a_n(0)t(n-1,0)$, which implies $a_n(0) = 1$ because we know from Theorem 5.1 that $t(n,0) = 1$ for all $n \geq 1$. Similarly we obtain

$$a_n(0)t(n-1,1) + a_n(3)t(n-1,0) = t(n,1).$$

Here we can use the fact from Theorem 5.1 that $t(n,1) = n$ for $n \geq 5$ to obtain $(n-1) + a_n(3) = n$, which implies $a_n(3) = 1$. Also we have $a_n(1)t(n-1,0) = 0$ and $a_n(2)t(n-1,0) = 0$ which imply that $a_n(1) = a_n(2) = 0$ because $t(n-1,0)$ is non-zero.

We can now show that $a_n(j) = 0$ for all $j \geq 4$. The proof of this fact can be broken into two cases, namely $j \equiv 0 \pmod 3$ and $j \equiv 1, 2 \pmod 3$. In the latter case, our result will follow from the fact that the numerator in Equation (5.19) contains only terms with $\frac{1}{\alpha^{3i-1}}$. Continuing to equate coefficients gives, for coefficients of $\frac{1}{\alpha^{3i}}$ and $\frac{1}{\alpha^{3i+1}}$ respectively,

$$a_n(1)t(n-1,i) + a_n(4)t(n-1,i-1) + \cdots + a_n(3i+1)t(n-1,0) = 0$$
$$a_n(2)t(n-1,1) + a_n(5)t(n-1,i-1) + \cdots + a_n(3i+2)t(n-1,0) = 0.$$

We have already seen that $a_n(1) = a_n(2) = 0$ and strong induction will show that $a_n(j) = 0$ for all $j \equiv 1, 2 \pmod 3$.

Consider the case $j \equiv 1 \pmod 3$ and suppose $a_n(1) = a_n(4) = \cdots = a_n(3i+1) = 0$. The coefficient for $\frac{1}{\alpha^{3(i+1)}}$ is given by

$$a_n(1)t(n-1,i+1)+a_n(4)t(n-1,i)+\cdots+a_n(3i+1)t(n-1,1)+a_n(3i+4)t(n-1,0) = 0.$$

This reduces to $a_n(3i+4)t(n-1,0) = 0$, implying $a_n(3i+4) = 0$ because $t(n-1,0) = 1$. The same argument holds for $j \equiv 2 \pmod 3$, in which case we are dealing with coefficients of $\frac{1}{\alpha^{3i+1}}$. Now we must show that $a_n(j) = 0$ for $j \geq 6$, where $j \equiv 0 \pmod 3$. Equating coefficients of $\frac{1}{\alpha^{3i-1}}$ gives

$$t(n,i) = a_n(0)t(n-1,i)+a_n(3)t(n-1,i-1)+a_n(6)t(n-1,i-2)+\cdots+a_n(3i)t(n-1,0),$$

where we have shown that $a_n(0) = a_n(3) = 1$. We again use strong induction. For the initial case we must show that $a_n(6) = 0$. Equating coefficients gives

$$a_n(0)t(n-1,2) + a_n(3)t(n-1,1) + a_n(6)t(n-1,0) = t(n,2).$$

We have seen that $a_n(0) = a_n(3) = 1$, and using the facts from Theorem 5.1 that $t(n-1,0) = 1$ and $t(n-1,2) + t(n-1,1) = t(n,2)$, we can rewrite the above as

$$t(n,2) + a_n(6) = t(n,2),$$

implying $a_n(6) = 0$. Now suppose that $a_n(9) = \cdots = a_n(3i) = 0$. Then we have that

$$\begin{aligned} t(n,i+1) = a_n(0)t(n-1,i+1) + a_n(3)t(n-1,i) + \qquad\qquad (5.20)\\ \cdots + a_n(3i)t(n-1,1) + a_n(3i+1)t(n-1,0) \\ = t(n-1,i+1) + t(n-1,i) + a_n(3i+3). \qquad\qquad (5.21)\end{aligned}$$

From Theorem 5.1, we know that our entries $t(n,k)$ follow the equation

$$t(n,k+1) = t(n-1,k+1) + t(n-1,k), \tag{5.22}$$

for $n \equiv 0 \pmod 3$ (which give non-corner elements). Therefore Equation (5.21) can be written as

$$t(n,i+1) = t(n,i+1) + a_n(3i+3),$$

implying $a_n(3i+3) = 0$, and completing the induction.

We may now conclude that

$$a_n(0)\alpha + a_n(1) + a_n(2)\frac{1}{\alpha} + a_n(3)\frac{1}{\alpha^2} + \cdots + a_n(j)\frac{1}{\alpha^{j-1}} + \cdots = \alpha + \frac{1}{\alpha^2}.$$

We know that α is the real root of $x^3 - 2x^2 - 1 = 0$, which can be rearranged to give $x + \frac{1}{x^2} = 2x - 2$. Therefore, $\alpha + \frac{1}{\alpha^2} = 2\alpha - 2$ and we have by Equation (5.19) that

$$\lim_{n\to\infty} \frac{S(\tau_n)}{S(\tau_{n-1})} = 2(\alpha - 1), \tag{5.23}$$

as required, for $n \equiv 0 \pmod 3$.

We must now prove that this limiting value is the same when $n \equiv 1, 2 \pmod 3$. The expression for $n \equiv 1 \pmod 3$ in Theorem 5.3 is given by

$$\begin{aligned}\lim_{n\to\infty} \frac{S(\tau_n)}{S(\tau_{n-1})} &= \lim_{n\to\infty} \frac{\sum_{k=0}^{\frac{n-4}{3}} t(n,k)\frac{1}{\alpha^{3k-1}} + t\left(\frac{n-1}{3}\right)\frac{1}{\alpha^{n-2}}}{\sum_{k=0}^{\frac{n-4}{3}} t(n-1,k)\frac{1}{\alpha^{3k}}} \\ &= \lim_{n\to\infty} \frac{\sum_{k=0}^{\frac{n-4}{3}} t(n,k)\frac{1}{\alpha^{3k-1}}}{\sum_{k=0}^{\frac{n-4}{3}} t(n-1,k)\frac{1}{\alpha^{3k}}} + \lim_{n\to\infty} \frac{t\left(\frac{n-1}{3}\right)\frac{1}{\alpha^{n-2}}}{\sum_{k=0}^{\frac{n-4}{3}} t(n-1,k)\frac{1}{\alpha^{3k}}},\end{aligned} \tag{5.24}$$

where we will show that each of the limits in the sum is defined. We have already seen that considering infinite sums, the first term in Equation (5.24) is $2(\alpha - 1)$ for $n \equiv 0 \pmod 3$. We must now show that this result holds for $n \equiv 1 \pmod 3$. We can follow through the cross multiplication and inductions as in the $n \equiv 0 \pmod 3$ case. The only difference is that we must verify Equation (5.22) for $n \equiv 1 \pmod 3$. We have already removed the corner term (which is an exception to the relation) from the sum, and so the remaining entries do indeed adhere to the relation. We must also show that the second term in the right-hand side of Equation (5.24) is 0.

We will begin by proving that

$$\lim_{n\to\infty} t\left(\frac{n-1}{3}\right)\frac{1}{\alpha^{n-2}} = 0,$$

which can be accomplished by showing that the series

$$\sum_{n=0}^{\infty} t\left(\frac{n-1}{3}\right)\frac{1}{\alpha^{n-2}}$$

converges. To do this, we can use the ratio test to show that for $n \equiv 0 \pmod 3$,

$$\begin{aligned}\lim_{n\to\infty} \frac{t\left(\frac{n+2}{3}\right)}{t\left(\frac{n-1}{3}\right)}\frac{1}{\alpha^3} &= L < 1,\\ \lim_{n\to\infty} \frac{t\left(\frac{n+2}{3}\right)}{t\left(\frac{n-1}{3}\right)} &= \alpha^3 L < \alpha^3 = 10.729\ldots.\end{aligned} \tag{5.25}$$

In order to show this, we can use the relations among the $t(n,k)$ terms in Table 5.2 (given in Theorem 5.1 in slightly different notation). We are looking at the ratios of consecutive corner terms $t(k)$. We can write

$$t\left(\tfrac{n+2}{3}\right) = 2t\left(n+1, \tfrac{n-1}{3}\right) = 2\left(t\left(n, \tfrac{n-4}{3}\right) + 2t\left(\tfrac{n-1}{3}\right)\right),$$

so that our ratio becomes

$$\frac{2t\left(n, \frac{n-4}{3}\right) + 4t\left(\frac{n-1}{3}\right)}{t\left(\frac{n-1}{3}\right)} = 4 + \frac{2t(n, \frac{n-4}{3})}{t\left(\frac{n-1}{3}\right)}.$$

Now we can write the denominator as

$$t\left(\tfrac{n-1}{3}\right) = 2t\left(n-2, \tfrac{n-4}{3}\right),$$

and the term in the numerator as

$$\begin{aligned} t\left(n, \tfrac{n-4}{3}\right) &= t\left(n-1, \tfrac{n-7}{3}\right) + t\left(n-1, \tfrac{n-4}{3}\right)\\ &= t\left(n-2, \tfrac{n-10}{3}\right) + 2t\left(n-2, \tfrac{n-7}{3}\right) + t\left(n-2, \tfrac{n-4}{3}\right)\\ &\le 4t\left(n-2, \tfrac{n-4}{3}\right),\end{aligned}$$

because the entries in the Table 5.2 are increasing for each row with $n-2 \equiv 2 \pmod 3$. (Recall we have $n \equiv 1 \pmod 3$.) The ratio in Equation (5.25) can now be written as the inequality

$$\lim_{n\to\infty} \frac{t\left(\frac{n+2}{3}\right)}{t\left(\frac{n-1}{3}\right)} \le 4 + \frac{2\left(4t\left(n-2, \frac{n-4}{3}\right)\right)}{2t\left(n-2, \frac{n-4}{3}\right)} = 4 + 4 = 8 < \alpha^3.$$

We now know that the numerator of the second term in the right-hand side of Equation (5.24),

$$\lim_{n\to\infty} \frac{t\left(\frac{n-1}{3}\right)\frac{1}{\alpha^{n-2}}}{\sum_{k=0}^{\infty} t(n-1,k)\frac{1}{\alpha^{3k}}}, \tag{5.26}$$

tends to zero. Consider the denominator. Each term in the sum is positive and the first term is $t(n-1,0)=1$. The sum is therefore greater than 1, and so the terms in the above limit are bounded above by $\lim_{n\to\infty} t\left(\frac{n-1}{3}\right)\frac{1}{\alpha^{n-2}}$, which goes to zero. Therefore, the term in (5.26) goes to zero also. We can conclude that for $n \equiv 1 \pmod 3$, Equation (5.23) holds.

Finally, in a similar fashion, we can prove this result for $n \equiv 2 \pmod 3$. We have from Theorem 5.3 that

$$\lim_{n\to\infty} \frac{S(\tau_n)}{S(\tau_{n-1})} \tag{5.27}$$

$$= \lim_{n\to\infty} \frac{\sum_{k=0}^{\frac{n-2}{3}} t(n,k)\frac{1}{\alpha^{3k-1}}}{\sum_{k=0}^{\frac{n-5}{3}} t(n-1,k)\frac{1}{\alpha^{3k}} + t\left(\frac{n-2}{3}\right)\frac{1}{\alpha^{n-2}}} = \lim_{n\to\infty} \frac{\sum_{k=0}^{\frac{n-5}{3}} c_n(k)\frac{1}{\alpha^{3k-1}} + t\left(n, \frac{n-2}{3}\right)\frac{1}{\alpha^{n-3}}}{\sum_{k=0}^{\frac{n-5}{3}} t(n-1,k)\frac{1}{\alpha^{3k}} + t\left(\frac{n-2}{3}\right)\frac{1}{\alpha^{n-2}}}$$

$$= \lim_{n\to\infty} \frac{\sum_{k=0}^{\frac{n-5}{3}} t(n,k)\frac{1}{\alpha^{3k-1}}}{\sum_{k=0}^{\frac{n-5}{3}} t(n-1,k)\frac{1}{\alpha^{3k}} + t\left(\frac{n-2}{3}\right)\frac{1}{\alpha^{n-2}}} + \lim_{n\to\infty} \frac{t\left(n, \frac{n-2}{3}\right)\frac{1}{\alpha^{n-3}}}{\sum_{k=0}^{\frac{n-5}{3}} t(n-1,k)\frac{1}{\alpha^{3k}} + t\left(\frac{n-2}{3}\right)\frac{1}{\alpha^{n-2}}}, \tag{5.28}$$

where we will show that each of the limits in the sum is defined. The first term can be rewritten as

$$\frac{1}{\lim_{n\to\infty} \frac{\sum_{k=0}^{\frac{n-5}{3}} t(n-1,k)\frac{1}{\alpha^{3k}} + t\left(\frac{n-2}{3}\right)\frac{1}{\alpha^{n-2}}}{\sum_{k=0}^{\frac{n-5}{3}} t(n,k)\frac{1}{\alpha^{3k-1}}}},$$

and splitting up the denominator (where it will be shown that the limits in the sum are defined) gives

$$= \frac{1}{\lim_{n\to\infty} \frac{\sum_{k=0}^{\frac{n-5}{3}} t(n-1,k)\frac{1}{\alpha^{3k}}}{\sum_{k=0}^{\frac{n-5}{3}} t(n,k)\frac{1}{\alpha^{3k-1}}} + \lim_{n\to\infty} \frac{t\left(\frac{n-2}{3}\right)\frac{1}{\alpha^{n-2}}}{\sum_{k=0}^{\frac{n-5}{3}} t(n,k)\frac{1}{\alpha^{3k-1}}}}.$$

We have that

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{\frac{n-5}{3}} t(n,k)\frac{1}{\alpha^{3k-1}}}{\sum_{k=0}^{\frac{n-5}{3}} t(n-1,k)\frac{1}{\alpha^{3k}}} = 2(\alpha-1),$$

by our previous argument - the fact that this expression contains no corner terms - and so its reciprocal tends to $\frac{1}{2(\alpha-1)}$. Now consider the term $t\left(\frac{n-2}{3}\right)$. We have already seen that $\lim_{n\to\infty} t\left(\frac{n-1}{3}\right)\frac{1}{\alpha^{n-2}} = 0$ when $n \equiv 1 \pmod 3$ because corner numbers $t\left(\frac{n-1}{3}\right)$ grow at a rate less than α^3. The same argument holds in the $n \equiv 2 \pmod 3$ because we are dealing with the same set of corner numbers. We can again conclude that

$$\lim_{n\to\infty} \frac{t\left(\frac{n-2}{3}\right)\frac{1}{\alpha^{n-2}}}{\sum_{k=0}^{\frac{n-5}{3}} t(n,k)\frac{1}{\alpha^{3k-1}}} = 0,$$

because the denominator is always greater than 1. We then have that the first term in Equation (5.28) is $\frac{1}{\frac{1}{2(\alpha-1)}} = 2(\alpha-1)$.

Now we must show that the second term goes to zero. The numerator contains $t\left(n, \frac{n-2}{3}\right)$, which for $n \equiv 2 \pmod 3$ is the sequence given by the numbers directly below the corners in our Table 5.2. We can write these numbers as

$$t\left(n, \frac{n-2}{3}\right) = \frac{1}{2}t\left(\frac{n+1}{3}\right).$$

The rate of growth of this sequence is therefore the same as that of the corner numbers, $t\left(\frac{n+1}{3}\right)$, which we have shown is less than α^3. We can then conclude that

$$\lim_{n\to\infty} t\left(n, \frac{n-2}{3}\right)\frac{1}{\alpha^{n-3}} = 0.$$

The second term in Equation (5.28) must then go to zero because the denominator is always greater than one and we obtain Equation (5.23), completing the proof for the case $n \equiv 2 \pmod 3$. $\diamond$

The following corollary (Rittaud [64]) gives us the growth rate of the expected value of the n^{th} term in a given row of T. In other words, we are taking the growth rate of an average. Rittaud calls this value the “average growth rate” but we choose to reserve this term for the average value of a list of growth rates of random Fibonacci sequences.

Corollary 5.5. *The growth rate of the expected value of the n^{th} term in a random Fibonacci sequence is given by*

$$\lim_{n\to\infty} \frac{\mathbb{E}(|t_n|)}{\mathbb{E}(|t_{n-1}|)} = \alpha - 1 = 1.205569431\ldots.$$

Proof: We know from Theorem 5.4 that

$$\lim_{n\to\infty} \frac{S(\tau_n)}{S(\tau_{n-1})} = 2(\alpha - 1).$$

We are dealing with the sum of entries in τ_n of the tree T. Since this tree is a representation of all possible random Fibonacci sequences, $S(\tau_n)$ also represents the sum of all possible n^{th} terms in a random Fibonacci sequence, taken in absolute value. We wish to consider the expected value of such terms, which is simply the average value. We must divide the sum by the total number of entries in row τ_n.

Tree T is a full binary tree with initial rows having sizes $|\tau_1| = |\tau_2| = 1$, $|\tau_3| = 2$, $|\tau_4| = 4$, and the n^{th} row having size $|\tau_n| = 2^{n-2}$ for $n \geq 2$. Therefore the expected value of the absolute value of the n^{th} term in a random Fibonacci sequence is given by

$$\mathbb{E}(|t_n|) = \frac{S(\tau_n)}{2^{n-2}}.$$

The growth rate can then be written as

$$\begin{aligned}\lim_{n\to\infty} \frac{\mathbb{E}(|t_n|)}{\mathbb{E}(|t_{n-1}|)} = \lim_{n\to\infty} \frac{S(\tau_n)/2^{n-2}}{S(\tau_{n-1})/2^{n-3}} &= \lim_{n\to\infty} \frac{S(\tau_n)}{S(\tau_{n-1})} \cdot \frac{2^{n-3}}{2^{n-2}} \\ &= 2(\alpha - 1) \cdot \frac{1}{2} = \alpha - 1,\end{aligned}$$

completing the proof. $\diamond$

This value is in agreement with the generalized mean inequality (Hardy et al. [36, p. 26]). The following definition can be found in [36, p. 12].

Definition 5.3. We define the *generalized mean* of a list of positive real numbers a_i by

$$\mathcal{M}_r(a) := \left(\frac{1}{n}\sum_{i=1}^{n} a_i^r\right)^{\frac{1}{r}},$$

where r is a real non-zero number.

The Generalized Mean Inequality says that if $r < s$ then $\mathcal{M}_r(a) \leq \mathcal{M}_s(a)$, with equality if and only if all of the a_i are equal. If we sum from $i = 1$ to 2^n (the number of length-$(n+2)$ random Fibonacci sequences), and let $a_i = t_i$, $r = \frac{1}{n}$ and $s = 1$, we have $r < s$ for $n > 1$ and our inequality becomes

$$\left(\frac{1}{2^n}\sum_{i=1}^{2^n}(|t_i|)^{\frac{1}{n}}\right)^n < \frac{1}{2^n}\sum_{i=1}^{2^n}|t_i|, \tag{5.29}$$

$$\frac{1}{2^n}\sum_{i=1}^{2^n}|t_i|^{\frac{1}{n}} < \left(\frac{1}{2^n}\sum_{i=1}^{2^n}|t_i|\right)^{\frac{1}{n}},$$

$$1.13198824\ldots < 1.205569431\ldots. \tag{5.30}$$

The term on the right calculates the growth rate of the average value of the n^{th} term in a random Fibonacci sequence (in absolute value), which we have seen in Corollary 5.5 is $1.205569431\ldots$ as $n \to \infty$. The term on the left calculates the average value of the approximated growth rates of all length-n random Fibonacci sequences (in absolute value), which in the limit gives Viswanath's constant, $1.13198824\ldots$.

Recall that Equation (3.17) gave us a way to approximate Viswanath's constant by taking the average of the n^{th} roots of traces $|u|$ for all length-n product matrices P_n (corresponding to the full binary tree T). We can perform a similar calculation by first taking the average of the $|u|$ values, and then finding the approximate growth rate (n^{th} root) of this average. This in fact seems to give another approximation to Rittaud's number $\alpha - 1 = 1.205569431\ldots$, and again the Generalized Mean Inequality holds. The trace values $|u|$ approximate the dominant eigenvalues and so we are essentially calculating the growth rate of the expected value of the dominant eigenvalue, as opposed to the growth rate of the expected value of the n^{th} term in a random Fibonacci sequence, as in Corollary 5.5.

The sums of the traces $|u|$ for product matrices P_n give us the sequence

$$2, 8, 8, 36, 72, 208, 464, 1204, 2528, 6768, \ldots$$

for increasing values of n. Since the number of product matrices for a given n is 2^n, we can simply find the growth rate of the above sequence of sums, and divide by two to find the growth rate of the average value, as in Corollary 5.5. This case seems to lack a simple recurrence for the sums. Table 5.3 gives the growth rate of the average trace for $6 \leq n \leq 16$.

n	sum of traces	approx. g.r. of average trace
6	208	1.217065337
7	464	1.201990802
8	1204	1.213523848
9	2528	1.194143022
10	6768	1.207858443
11	15688	1.203331371
12	39040	1.206690300
13	92744	1.205227170
14	226836	1.206482036
15	532128	1.204216052
16	1310484	1.205894980

Table 5.3: Approximate growth rates of average trace values.

Example 5.10. For $n = 16$, we have that the sum of traces is 1310484, giving a growth rate approximation of $1310484^{\frac{1}{16}} = 2.411789959$, using the root form of Definition 2.4. Therefore the growth rate of the average trace is approximated by $\frac{2.411789959}{2} = 1.205894980$, which is accurate to three decimal places to Rittaud's calculation.

Let $u_{n,i}$ and $\lambda_{n,i}$ be the trace and dominant eigenvalue respectively of the i^{th} product matrix $P_{n,i}$.

Conjecture 5.1. *We have the following alternate formula for the growth rate of the expected value of n^{th} terms in a random Fibonacci sequence,*

$$\lim_{n\to\infty} \frac{(|u_{n,1}| + |u_{n,2}| + \cdots + |u_{n,i}| + \cdots + |u_{n,2^n}|)^{\frac{1}{n}}}{2} = \alpha - 1 = 1.20556943\ldots, \quad (5.31)$$

where $u_{n,i}$ is the trace of the i^{th} product matrix $P_{n,i}$.

Based on Equation (3.16), which says that for large n and $|u|$ we have $|u| \approx |\lambda|$, we suspect that we may be able to make a similar calculation using the absolute value of the dominant eigenvalue instead of the trace value. This would give

$$\lim_{n\to\infty} \frac{(|\lambda_{n,1}| + |\lambda_{n,2}| + \cdots + |\lambda_{n,i}| + \cdots + |\lambda_{n,2^n}|)^{\frac{1}{n}}}{2} = \alpha - 1 = 1.20556943\ldots. \quad (5.32)$$

5.6 Rittaud's Proof

The main result of Rittaud's paper [64] is Corollary 5.5, which gives the growth rate of the expected value of the n^{th} term in a random Fibonacci sequence. Further, he claims that this result is true for any initial values $a, b \geq 0$, with $ab \neq 0$. An error in his proof is pointed out in Janvresse et al. [41], and corrected in a more general setting. Rittaud breaks his proof into four steps. We will need to make use of the notation $\rho_n^{(a,b)}$ and $\tau_n^{(a,b)}$ to denote the n^{th} rows in trees $R^{(a,b)}$ and $T^{(a,b)}$ respectively, and so $\rho_n^{(1,1)} = \rho_n$.

The first step of the proof is to show that the growth rate of $\{S(\rho_n^{(a,b)})\}$ is independent of the initial values a, b, where $ab \neq 0$. By Theorem 5.2, we know that this growth rate, for $a = b = 1$, is $\alpha = 2.20556943\ldots$, where α is the real root of $\alpha^3 - 2\alpha^2 - 1 = 0$. The corresponding recurrence is $S(\rho_n) = 2S(\rho_{n-1}) + S(\rho_{n-3})$. Recall in the case of the Fibonacci numbers, Equation (1.4) gave G_n (the Fibonacci sequence with initial values a, b) as $G_n = aF_{n-1} + bF_{n-2}$, and the sequence $\{G_n\}$ also had growth rate ϕ. By the theory of linear recurrences, the same is true for $\{S(\rho_n^{(a,b)})\}$. We can write

$$S(\rho_n^{(a,b)}) = au_n + bv_n, \tag{5.33}$$

where $u_n = 2u_{n-1} + u_{n-3}$ for $u \geq 5$ with $u_1 = 1, u_2 = 0, u_3 = 1, u_4 = 2$ and $v_n = 2v_{n-1} + v_{n-3}$ for $u \geq 4$ with $v_1 = 0, v_2 = 1, v_3 = 1, v_4 = 2$. The first few terms of $S(\rho_n^{(a,b)})$ are therefore

$$a, b, (a + b), (2a + 2b), (4a + 5b), (9a + 11b), \ldots,$$

which adheres to our recurrence for $n \geq 5$. The growth rate of $\{S(\rho_n^{(a,b)})\}$ is therefore also α. Recall the tree $R^{(a,b)}$ for $a < b$ given in Figure 5.4. The sums of rows are given by

$$a, b, 2b, 4b + a, 9b + 2a, \ldots,$$

which also adhere to the recurrence, for $n \geq 4$. The former sequence of sums corresponds to the tree $R^{(a,b)}$ with $\rho_1 = a$, $\rho_2 = b$, $\rho_3 = a + b$, which has the same shape as R.

The second step in Rittaud's proof focuses on finding an expression for $\tau_n^{(a,b)}$ in terms of $\rho_j^{(|b-a|,a)}$ with $j \leq n$, for the specific cases $(a, b) = (1, \phi)$ and $(a, b) = (1, \phi^{-1})$.

Recall from Proposition 5.3 that for any node a in R with child b, a's parent is $|b-a|$. We will denote $z := |b-a|$, so that $z = \phi^{-1}$ and $z = \phi^{-2}$ in our cases. Rittaud shows that for the specific choice of initial values we have

$$\tau_{n+2}^{(a,b)} = \biguplus_{m=0}^{\lfloor \frac{n}{3} \rfloor} \left(\binom{n}{m} - 2\binom{n}{m-1} \right) \nu_{n+2-3m},$$

where

$$\nu_m := \biguplus_{i=0}^{\lfloor \frac{m}{2} \rfloor} z^i \rho_{m-2i}^{(z,1)}. \tag{5.34}$$

(Recall that we are taking the multiset sum of nodes here.) These relations do not necessarily hold for any a, b. The proof uses the fact that for

$$c_{n,m} := \left(\binom{n}{m} - 2\binom{n}{m-1} \right),$$

Pascal's formula holds, i.e., $c_{n+1,m+1} = c_{n,m+1} + c_{n,m}$.

The third step in Rittaud's proof is to find the explicit growth rate of the sequence $\{S(\tau_n^{(a,b)})\}$ for the specific values of (a, b). He uses the fact that from the theory of linear recurrences we can write the solution to the recurrence as

$$S(\rho_n^{(a,b)}) = c_1\alpha^n + c_2\beta^n + \overline{c_2}\overline{\beta}^n,$$

for any a, b with $ab \neq 0$, where constants $c_1 \in \mathbb{R}, c_2 \in \mathbb{C}$ depend on the initial values of $\{S(\tau_n^{(a,b)})\}$. Also, α is the real root of $x^3 = 2x^2 + 1$ and the complex roots are $\beta = -0.1027847\ldots + i0.6654569\ldots$. By Equation (5.34), we can write $S(\nu_m)$ in terms of $S(\rho_j^{(z,1)})$, and therefore in terms of α, β and $\overline{\beta}$. Rittaud shows that the latter two terms are negligible and that the sequence $\{S(\tau_n^{(a,b)})\}$ grows at the rate $2(\alpha - 1)$. This is the result we obtained in Theorem 5.4 for $(a, b) = (1, 1)$.

The last step in Rittaud's proof is to show that this growth rate applies to random Fibonacci sequences with any initial values $a, b \geq 0$ with $ab \neq 0$. (Since we are talking about the positive tree $T^{(a,b)}$, we can assume the initial values are positive.) He starts by noting that we can write any pair (a, b) as a linear combination of our special initial values as

$$(a, b) = u(1, \phi) + v(1, \phi^{-1}),$$

for constants $uv \neq 0$. Rittaud then deduces that by linearity, we can write the row $\tau_n^{(a,b)}$ as a linear combination as

$$\tau_n^{(a,b)} = u\tau_n^{(1,\phi)} \uplus v\tau_n^{(1,\phi^{-1})}, \tag{5.35}$$

and taking sums over these multisets gives

$$S(\tau_n^{(a,b)}) = uS(\tau_n^{(1,\phi)}) + vS(\tau_n^{(1,\phi^{-1})}). \tag{5.36}$$

Since we know that both sequences $\{S(\tau_n^{(1,\phi)})\}$ and $\{S(\tau_n^{(1,\phi^{-1})})\}$ grow exponentially at rate $2(\alpha - 1)$, we can deduce, using the same logic as in the first step, that $\{S(\tau_n^{(a,b)})\}$ does also, for any initial values a, b with $ab \neq 0$. By Corollary 5.5, the growth rate of the expected value of the n^{th} term in a random Fibonacci sequence is given by $\alpha - 1$. The proof of this is independent of initial values, so we can assume that this growth rate holds also for random Fibonacci sequences with initial values a, b, and in particular the sequence with $a = b = 1$.

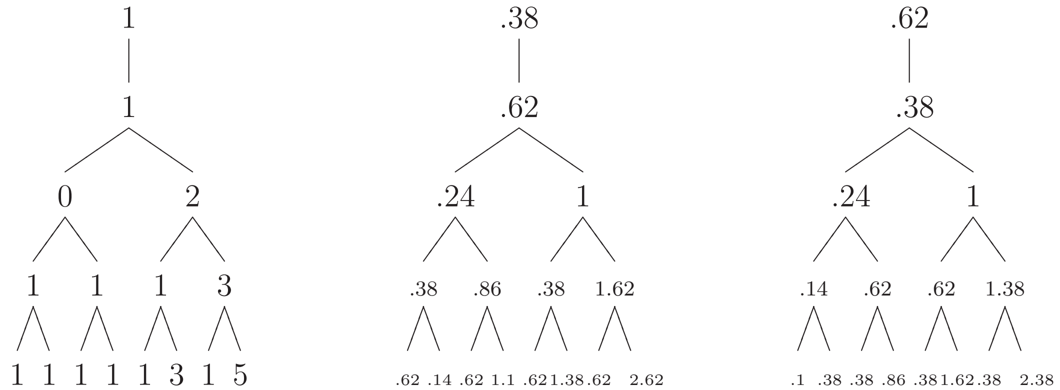
There is a mistake in Rittaud's proof, however, and it occurs in Equation (5.35). The assumption that we can write the n^{th} row of $T^{(a,b)}$ as a linear combination is false. This is pointed out by Janvresse et al. in [41], and as an example they consider writing $(a, b) = (1, 1)$ as a linear combination of the specific initial values. Solving for u and v gives

$$(1, 1) = \frac{1}{\phi^2}(1, \phi) + \frac{1}{\phi}(1, \phi^{-1}) = (\phi^{-2}, \phi^{-1}) + (\phi^{-1}, \phi^{-2}).$$

Figure 5.9 gives the three trees $T^{(a,b)}$ corresponding to the initial values above, in approximated decimal form. Recall that multiplying initial values by a constant is equivalent to multiplying all nodes in a tree by that constant.

The left grandchild of the root node in Figure 5.9(a) is 0, whereas the sum of the corresponding left grandchildren in Figures 5.9(b) and 5.9(c) is approximately 0.48. Therefore the tree in Figure 5.9(a) is not simply the sum of the trees in Figures 5.9(b) and 5.9(c). Comparing further in the trees, we see that some nodes do sum correctly and others do not. Rittaud's error can be explained by the fact that his tree $T^{(a,b)}$ is generated using Equation (1.8), namely,

$$\widetilde{f}_n = |\widetilde{f}_{n-1} \pm \widetilde{f}_{n-2}|.$$



(a) The tree $T^{(1,1)}$ (b) The tree $T^{(\phi^{-2},\phi^{-1})}$ (c) The tree $T_{(\phi^{-1},\phi^{-2})}$

Figure 5.9: $T^{(a,b)}$ for different initial values.

This equation is non-linear. The fact that we are taking the absolute value at each step means that some nodes will be permuted and negated, compared to the linear Equation (1.7), $f_n = f_{n-1} \pm f_{n-2}$. Had we generated the trees in Figure 5.9 using this equation, the summation of trees would hold and there would be no flaw in Rittaud's proof. We know from Corollary 1.1 that the set of n^{th} terms is the same for the sequences $\{|f_n|\}$ and $\{\widetilde{f}_n\}$, so we can repeat Rittaud's argument in step four by generating trees using Equation (1.7), and taking the absolute values afterwards. The growth rate of the sum of n^{th} terms of the sequence $\{|f_n|\}$ with initial values a, b is unchanged and can be written as the linear combination in Equation (5.36). Therefore the growth rate of the expected value of a random Fibonacci sequence with initial values a, b is in fact $\alpha - 1$.

Notice that the nodes in Figures 5.9(a) which belong to tree R, are the sum of those in Figures 5.9(b) and 5.9(c). This is because Equations (1.7) and (1.8) generate the same tree $R^{(a,b)}$, by Proposition 5.7. This also explains why the growth rate of $S(\rho_n^{(a,b)})$ in Equation (5.33) carries through the linear combination.

As pointed out earlier, Janvresse et al. [41] found Rittaud's error, and also provide a solution in the more general setting of (p,α)-random Fibonacci sequences for the non-linear case. As mentioned in Chapter 1, sums of rows of $T^{(a,b)}(p,\lambda_k)$ are approximated by splitting this tree into infinitely many copies of the tree $R^{(l_{s+1},l_{s+2})}(p,\lambda_k)$. Here the initial nodes l_{s+1} and l_{s+2} are consecutive nodes in the leftmost branch of

$T^{(a,b)}(p, \lambda_k)$. The expected value of a node in the n^{th} row in the tree $T^{(a,b)}(p, \lambda_k)$ is given by

$$\mathbb{E}(\tau_n^{(a,b)}) = \sum_{m=0}^{\lfloor \frac{n}{k} \rfloor} \sum_{s=0}^{n-km} c_{n,m} (pq^{k-1})^m q^s \mathbb{E}(\rho_{n+2-s-km}^{(l_{s+1},l_{s+2})}(p, \lambda_k)), \tag{5.37}$$

where $q = 1-p$, $\rho_{n+2-s-km}^{(l_{s+1},l_{s+2})}(p, \lambda_k)$ denotes row $n+2-s-km$ of the tree $R^{(l_{s+1},l_{s+2})}(p, \lambda_k)$ and

$$c_{n,m} = \binom{n}{m} - (k-1)\binom{n}{m-1}. \tag{5.38}$$

Recall from Chapter 1 that $\lambda_k = 2\cos(\pi/k)$, so that $\lambda_3 = 1$. If we also choose $p = \frac{1}{2}$, we have the case discussed by Rittaud in [64]. We can rewrite Equation (5.37) using these values to give

$$\mathbb{E}(\tau_n^{(a,b)}) = \sum_{m=0}^{\lfloor \frac{n}{3} \rfloor} \sum_{s=0}^{n-3m} c_{n,m} \left(\frac{1}{2}\right)^{3m+s} \mathbb{E}(\rho_{n+2-s-3m}^{(l_{s+1},l_{s+2})}),$$

where

$$c_{n,m} := \binom{n}{m} - 2\binom{n}{m-1}.$$

Pascal's formula holds for the general coefficient $c_{n,m}$ given in Equation (5.38). Table 5.4 gives the coefficients for $k = 3$, and Pascal's formula can easily be observed.

n \ m	0	1	2	3
1	1			
2	1			
3	1			
4	1	1		
5	1	2		
6	1	3		
7	1	4	3	
8	1	5	7	
9	1	6	12	
10	1	7	18	12
11	1	8	25	30
12	1	9	33	55

Table 5.4: Values of $c_{n,m}$ for $n \leq 12$.

Notice that our proof of Theorem 5.4 (and consequently Corollary 5.5), as well as the proof by Rittaud and the generalized proof by Janvresse et al. have all used the same technique of partitioning the nodes in some form of the tree T into nodes in its subtree R. Growth rates are easier to find for R and then can be translated back to the tree T. Our method used Equation (5.10), which is a linear combination of sums of nodes in R, namely $S(\rho_{n-3k})$. The coefficients are quite complicated, as will be shown in the following sections. The methods of Rittaud and Janvresse et al. contrast ours in that their coefficients are fairly straightforward, consisting primarily of $c_{n,m}$, while the partition into nodes in R is much more complicated.

The proof we have given of Theorem 5.4, for the growth rate of the sum $S(\tau_n)$, for a random Fibonacci sequence with initial values $1, 1$, does not easily generalize to initial values a, b. The breakdown of rows τ_n into rows ρ_j for $j \leq n$ as given in Table 5.1 no longer holds. Furthermore, given initial values a, b, the tree $R^{(a,b)}$ may not be constructed in the same way as $R^{(1,1)}$ because different edge repetitions will occur. This was discussed in Section 5.1. There we stated that for $a < b$, the overall structure of $R^{(a,b)}$ is the same as that of $R^{(1,1)}$ if we think of the initial values as shifted to $a = \rho_2$ and $b = \rho_3$. In this case we would then have $\rho_1 = b - a$, and our tree is in fact $R^{(b-a,a)}$, as was used by Rittaud to partition $T^{(a,b)}$. In this case, if we try to partition rows $\tau_n^{(a,b)}$ into rows $\rho_j^{(b-a,a)}$, we still do not obtain the breakdown given in Table 5.1.

5.7 An Explicit Formula for Corner Terms $t(k)$

Our goal is to find a recurrence for the corner terms $t(k)$ for $k \geq 0$, which form the sequence

$$\{t(k)\} = \frac{1}{2}, 2, 10, 54, 304, 1758, \ldots.$$

By observing Table 5.2, we can represent such a recurrence geometrically and derive a combinatorial formula. If we try to write the first few values of $\{t(k)\}$ in terms of

$t(j)$ for $j \leq k-1$, we obtain the following recurrences:

$$\begin{aligned} 2 &= 4t(0), \\ 10 &= 4t(0) + 4t(1), \\ 54 &= 12t(0) + 4t(1) + 4t(2), \\ 304 &= 48t(0) + 12t(1) + 4t(2) + 4t(3), \\ 1758 &= 220t(0) + 48t(1) + 12t(2) + 4t(3) + 4t(4). \end{aligned}$$

The coefficients, which we will call $A(m)$ appear to form a sequence given by

$$\{A(m)\} = 4, 4, 12, 48, 220, \ldots$$

for $m \geq 0$. The sequence $\{A(m)/4\}$ is identified in [67, A001764] as having form $\frac{1}{2m+1}\binom{3m}{m}$, and so

$$A(m) = \frac{4}{2m+1}\binom{3m}{m}. \tag{5.39}$$

Graham et al. [33, p. 361] identify this as a generalized Catalan number, called a *Fuss-Catalan number*, given by

$$C_m^{(k)} := \binom{km}{m}\frac{1}{(k-1)m+1}.$$

For $k = 2$, we obtain the regular Catalan numbers $C_m = C_m^{(2)}$, and for $k = 3$ we obtain $A(m)/4$. We must prove that the coefficients we obtain actually adhere to this pattern.

Example 5.11. Let us derive the corner recurrence for $t(4) = 304$. We can write, by Theorem 5.1,

$$\begin{aligned} 304 &= 2(152) = 2(44 + 2 \cdot \mathbf{54}) \\ &= 2(9 + 35) + 4 \cdot \mathbf{54} \\ &= 2(1 + 8 + 8 + 27) + 4 \cdot \mathbf{54} \\ &= 2(1 + 2(1 + 7) + 7 + 2 \cdot \mathbf{10}) + 4 \cdot \mathbf{54} \\ &= 2(3(1) + 3(1 + 6)) + 4 \cdot \mathbf{10} + 4 \cdot \mathbf{54} \end{aligned}$$

$$
\begin{aligned}
&= 2(6(1) + 3(1+5)) + 4 \cdot \mathbf{10} + 4 \cdot \mathbf{54} \\
&= 2(9(1) + 3(1 + 2 \cdot \mathbf{2}) + 4 \cdot \mathbf{10} + 4 \cdot \mathbf{54} \\
&= 2(12(1) + 6 \cdot \mathbf{2}) + 4 \cdot \mathbf{10} + 4 \cdot \mathbf{54} \\
&= 48 \cdot \tfrac{\mathbf{1}}{\mathbf{2}} + 12 \cdot \mathbf{2} + 4 \cdot \mathbf{10} + 4 \cdot \mathbf{54},
\end{aligned}
$$

where corner terms are bolded.

We can proceed to represent this process geometrically. Notice in Table 5.2 that we can write any non-corner term $t(n,k)$ not only according to Theorem 5.1, but as twice the corner term $t(k)$ plus entries in column $k-1$. For example, $t(14,3) = 348$ can be written as $2 \cdot \mathbf{54} + 44 + 54 + 65 + 77$. This is a simple consequence of repeating Equation (5.6) until we arrive at the corner entry. Visually we have Figure 5.10. The structure of the dots in this and subsequent figures is taken from Table 5.2. Numbers

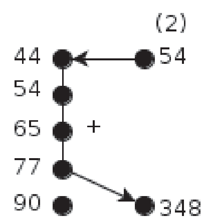


Figure 5.10: Breakdown of 348 into smaller terms.

occurring in parentheses represent nodes that are counted multiple times. We can use this idea to write any term $t(n,k)$ as a linear combination of corner terms.

Let us first consider how to write the corner terms themselves using this method. We have that

$$
\begin{aligned}
t(k) = t(n,k) &= 2t(n-2, k-1) \\
&= 2(t(n-3, k-2) + 2t(k-1)) \\
&= 4t(k-1) + 2t(n-3, k-2). \qquad (5.40)
\end{aligned}
$$

Example 5.12. Consider the breakdown $t(4) = 304 = 4 \cdot \mathbf{54} + 2(44)$ from Example 5.11. Now we must write 44 in terms of $t(2) = 10$, and we can view this process geometrically in Figure 5.11. We can write $44 = 2 \cdot \mathbf{10} + 7 + 8 + 9$ and we are left to

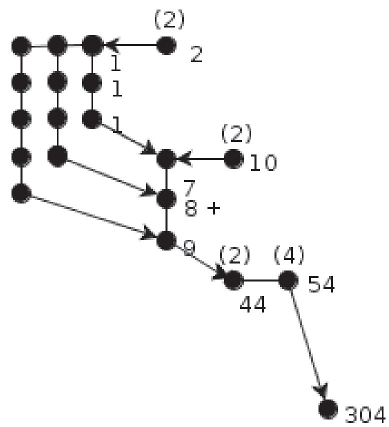


Figure 5.11: Breakdown of 304 into corner terms.

write $7, 8, 9$ in terms of $t(1) = 2$. We have

$$7 = 2 \cdot \mathbf{2} + 1 + 1 + 1,$$
$$8 = 2 \cdot \mathbf{2} + 1 + 1 + 1 + 1,$$
$$9 = 2 \cdot \mathbf{2} + 1 + 1 + 1 + 1 + 1.$$

Each term 1 can then be written as $2t(0) = 2 \cdot \mathbf{\frac{1}{2}}$.

Notice that in Equation (5.40) the corner term $t(k-1)$ (54 in Example 5.12) is multiplied by 4 and the term $t(n-3, k-2)$ (44 in Example 5.12) is doubled. Also, each time we come upon a new corner number in the recurrence, it is doubled. Therefore, the number of times a corner number will appear is a multiple of 4. For example, each term $7, 8, 9$ appears once in each term 44 and contains two copies of $t(1) = 2$. So the number of $t(1)$ terms appearing in 304 is $4(3) = 12$. This term 3 is visualized as the line in Figure 5.12. We can continue to count "dots" in our geometric diagram to aid us in counting corners. Next, to count the number of times $t(0) = \frac{1}{2}$ occurs, we must count the number of 1 terms, each of which contributes 2 terms $\frac{1}{2}$. For example, consider Figure 5.13. We have a 2-dimensional shape containing 12 dots, each of which counts the term $t(0) = \frac{1}{2}$ four times in 304; twice for each term 1, and again doubled because $7, 8, 9$ all occur once in each of the two terms 44. It is representative

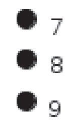


Figure 5.12: Line of size 3.

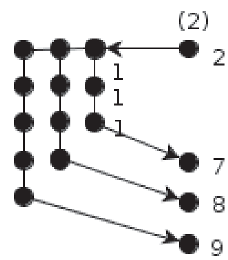


Figure 5.13: Counting terms $t(0)$ in $7, 8, 9$.

of the set of equations after Figure 5.11. Each dot in the shape on the right generates a vertical line, the length of which is determined by the distance from the dot to the top of the figure.

This same pattern holds for the general case, by the structure of Table 5.2. If we are dealing with a corner larger than 304, the next stage in the progression of geometric figures would look like Figure 5.14. By the previous logic, the number of

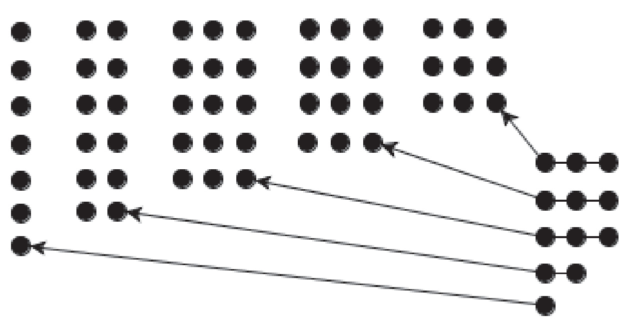


Figure 5.14: Counting the next level of corner terms.

corners would be four times the number of dots, which is $4(55) = 220$. We can now redraw Figure 5.14 as the group of 2-dimensional shapes in Figure 5.15, and think of it as a shape in 3 dimensions. We can continue creating a new line for each dot $t(n, k)$

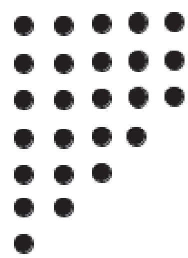


Figure 5.15: Shape of dots in three dimensions.

in the same manner — by moving to the point $t(n-1, k-1)$ and tracing upwards until the next corner is reached (three rows above the top of the previous shape).

We now require a formula for counting dots. We can write the recurrence for the corner term $t(k)$ as

$$t(k) = 4s_0t(k-1) + 4s_1t(k-2) + 4s_2t(k-3) + \cdots + 4s_mt(k-m-1) + \cdots + 4s_{k-1}t(0), \tag{5.41}$$

where the coefficient s_m, for $0 \le m \le k-1$, is therefore counting the dots in the sequence of shapes given in Figure 5.16.

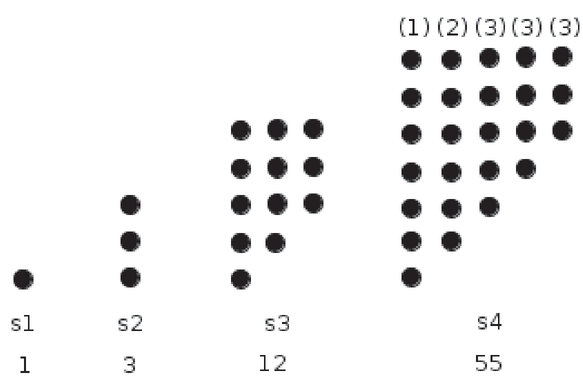


Figure 5.16: Sequence of coefficients s_m for the corner terms.

We will set $s_0 = s_1 = 1$. Here s_2 counts the number of dots in the single line, s_3 counts the number of dots in the 2-dimensional shape, s_4 counts the number of dots in the group of 2-dimensional shapes (i.e., the 3-dimensional shape), and so on. Notice how these shapes grow in groups of three. (See Figure 5.15 for s_4.) More generally, we will let $s(i,0) = s(i,1) = 1$ for all $i \geq 1$ and define $s(i,m)$, for $i \geq 1$ and $m \geq 1$, to be the total number of dots in the m^{th} shape of the sequence, where the sequence starts with $m = 0$. Here i is the number of dots in the 1-dimensional column at $m = 2$, i.e., $s(i,2) = i$. Note that the m^{th} shape of the sequence, for $i \geq 2$, has dimension $m-1$. The above s_m terms then correspond to the case $i = 3$. We can observe the following relations for $i = 3$:

$$\begin{aligned}
s(3,3) &= s(3,2) + s(4,2) + s(5,2) \\
&= 3 + 4 + 5 = 12, \\
s(3,4) &= s(3,3) + s(4,3) + s(5,3) \\
&= 12 + 18 + 25 = 55.
\end{aligned}$$

These equations are summing the number of dots in columns and 2-dimensional shapes respectively. We also have the trivial case

$$\begin{aligned}
s(3,2) &= s(3,1) + s(4,1) + s(5,1) \\
&= 1 + 1 + 1 = 3.
\end{aligned}$$

The $i = 3$ case is important because it gives us the coefficients for our corner recurrence. Using the new $s(i,m)$ notation, Equation (5.41) becomes

$$\begin{aligned}
t(k) = 4s(3,0)t(k-1) &+ 4s(3,1)t(k-2) + 4s(3,2)t(k-3) + \cdots \\
&+ 4s(3,m)t(k-m-1) + \cdots + 4s(3,k-1)t(0). \qquad (5.42)
\end{aligned}$$

In order to prove an explicit formula for $s(3,m)$ we need to consider the general case $s(i,m)$.

Lemma 5.2. *A shape at level m (having $s(i,m)$ dots) has $m-1$ dimensions and is mapped to a shape at level $m+1$ (having $s(i,m+1)$ dots) with m dimensions, for $i \geq 2$. If $i = 1$, a shape at level m has $m-2$ dimensions and is mapped to a shape with $m-1$ dimensions.*

Proof: Consider what happens when we move up a level in our sequence of geometric shapes. We know that since each dot turns into a line with i dots where $i \geq 3$, each line turns into a 2-dimensional shape with i columns, as represented by $s(i,2) \mapsto s(i,3)$. Now consider $s(i,m)$ for $i \geq 2$, i.e., the number of dots in a shape of $m-1$ dimensions. Each line in this shape turns into a 2-dimensional shape, and so by induction the shape itself will turn into one with m dimensions, and containing i shapes of $m-1$ dimensions. In the case of $i = 1$, we have a line at $m = 2$ containing one dot, which turns into another line with 3 dots, and so on. These shapes still increase in dimension, but instead we have an $m-2$ dimensional shape at level m. ◇

Now we can prove our summation relation for the general case.

Theorem 5.5. *For $m \geq 2$ and $i \geq 1$ we have that*

$$s(i,m) = s(3,m-1) + s(4,m-1) + \cdots + s(i+2,m-1). \tag{5.43}$$

Proof: For the initial case we have $s(1,2) = 1$ because we have seen that $s(i,2) = i$. Also, we have defined $s(i,1) = 1$ and so $s(3,1) = 1$, implying

$$s(1,2) = s(3,1).$$

Now we must induct over the two variables i and m. First assume Equation (5.43) holds for $s(i,2)$ and show it is true for $s(i+1,2)$. We are dealing with the trivial case for m and so we can simply write that

$$\begin{aligned} s(i+1,2) &= i+1 = 1 + \cdots + 1 \\ &= s(3,1) + s(4,1) + \cdots + s(i+3,1), \end{aligned}$$

proving the relation without needing to induct.

Now inducting over m, we can assume Equation (5.43) for $s(i,m)$, for all i. The term $s(i,m+1)$ gives us the number of dots in the shape one dimension higher. As shown in Lemma 5.2, each of the terms $s(j,m-1)$ for $3 \leq j \leq i+2$ is transformed into $s(j,m)$ so that we obtain

$$s(i,m+1) = s(3,m) + s(4,m) + \cdots + s(i+2,m), \tag{5.44}$$

completing the induction. $\diamond$

We are now ready to use this summation formula to find an explicit formula for $s(i,m)$. Let us first consider a couple of initial cases as examples.

Example 5.13. For $m = 3$, Theorem 5.5 gives

$$\begin{aligned} s(i,3) &= s(3,2) + s(4,2) + \cdots + s(i+2,2) \\ &= 3 + 4 + \cdots + (i+2) \\ &= \frac{(i+2)(i+3)}{2} - 2 - 1 \\ &= \frac{i(i+5)}{2}. \end{aligned}$$

So when $i = 3$, we obtain $s(3,3) = 12 = \frac{1}{4}A(3)$. Similarly, when $m = 4$ we have

$$\begin{aligned} s(i,4) &= s(3,3) + s(4,3) + \cdots + s(i+2,3) \\ &= \sum_{j=3}^{i+2} \frac{j(j+5)}{2} = \frac{1}{2}\left(\sum_{j=3}^{i+2} j^2 + 5\sum_{j=3}^{i+2} j\right) \\ &= \left(\frac{(i+2)(i+3)(2i+5)}{6} - 5 + 5\left(\frac{i(i+5)}{2}\right)\right) \\ &= \frac{i(i+7)(i+8)}{6}. \end{aligned}$$

Continuing in this fashion, we come up with the following explicit formula.

Theorem 5.6. *Let $s(i,1) = 1$. For $i \geq 1$ and $m \geq 2$ the term $s(i,m)$ can be written as follows:*

$$s(i,m) = \frac{i(i+2m-1)(i+2m)\cdots(i+3m-4)}{(m-1)!} = \binom{i+3m-4}{i+2m-2}\frac{i}{m-1}. \tag{5.45}$$

Proof: To start, we have that for the initial case, $s(1,2) = \binom{3}{3}\frac{1}{1} = 1$, as required. We must now induct on both variables i and m. We can start by showing that if Equation (5.45) holds for $s(i,2)$, then it holds for $s(i+1,2)$, and then show that if it holds for $s(i,m)$, it must also hold for $s(i,m+1)$. We can now assume $s(i,2) = \binom{i+2}{i+2}\frac{i}{1} = i$. It is true then, that $s(i+1,2) = \binom{i+3}{i+3}\frac{i+1}{1} = i+1$ because we have already seen that $s(i,2) = i$ for all i.

Now we must assume Equation (5.45) holds for m and all values of i, and show it also holds for $m+1$. We know that from Theorem 5.5 that

$$s(i, m+1) = s(3, m) + s(4, m) + \cdots + s(i+2, m),$$

which by our assumption can be written as

$$s(i, m+1) = \sum_{j=3}^{i+2} \binom{j+3m-4}{j+2m-2} \frac{j}{m-1}.$$

Using software such as Maple, we can see that this expression in fact evaluates as

$$\binom{i+3m-1}{i+2m} \frac{i}{m} = \binom{i+3(m+1)-4}{i+2(m+1)-2} \frac{i}{(m+1)-1},$$

completing the induction. $\diamond$

Corollary 5.6. *In the recurrence of corner numbers $t(k)$ given in Equation* (5.42)*, the term $t(k-m-1)$ has coefficient $A(m)$, for $k \geq 1, m \leq k-1$.*

Proof: We know that the coefficient of $t(k-m-1)$ in our recurrence is given by $4s(3, m)$, which by Theorem 5.6 can be written as

$$4s(3, m) = \frac{12}{m-1} \binom{3m-1}{2m+1}.$$

We can now rewrite this binomial term. We have that

$$\begin{aligned}
\frac{12}{m-1} \binom{3m-1}{2m+1} &= \frac{12}{m-1} \left(\frac{(3m-1)(3m-2)\cdots(2m+2)}{(m-2)!} \right) \\
&= \frac{4m}{m} \left(\frac{3(3m-1)(3m-2)\cdots(2m+2)}{(m-1)!} \right) \\
&= \frac{4}{2m+1} \left(\frac{(3m)(3m-1)\cdots(2m+2)(2m+1)}{m!} \right) \\
&= \frac{4}{2m+1} \binom{3m}{m} = A(m),
\end{aligned}$$

by Equation (5.39) for $A(m)$, completing the proof. $\diamond$

Now we can rewrite Equation (5.42), for our corner terms, as

$$\begin{aligned}
t(k) &= A(0)t(k-1) + A(1)t(k-2) + A(2)t(k-3) + \cdots + A(k-1)t(0) \\
&= \sum_{j=0}^{k-1} A(k-j-1)t(j). \qquad (5.46)
\end{aligned}$$

We wish to use this corner recurrence to find a closed form expression for $t(k)$. We will need to make use of the generating functions with coefficients $A(k)$ and $t(k)$, which are defined respectively as

$$\begin{aligned} G_A(x) &:= \sum_{k=0}^{\infty} A(k)x^k = \sum_{k=0}^{\infty} \frac{4}{2k+1}\binom{3k}{k}x^k, \\ G_t(x) &:= \sum_{k=0}^{\infty} t(k)x^k. \end{aligned} \tag{5.47}$$

Multiplying these two generating functions gives the following Cauchy product:

$$\begin{aligned} G_A(x)G_t(x) &= \left(\sum_{k=0}^{\infty} A(k)x^k\right)\left(\sum_{k=0}^{\infty} t(k)x^k\right) \\ &= \sum_{k=0}^{\infty}\left(\sum_{j=0}^{k} A(k-j)t(j)\right)x^k \\ &= \sum_{k=0}^{\infty} t(k+1)x^k, \end{aligned} \tag{5.48}$$

where we have used Equation (5.46) to obtain the last line.

This new generating function can be rewritten as

$$\begin{aligned} \sum_{k=0}^{\infty} t(k+1)x^k &= \frac{1}{x}\sum_{k=0}^{\infty} t(k+1)x^{k+1} = \frac{1}{x}\sum_{k=1}^{\infty} t(k)x^k \\ &= \frac{1}{x}\left(\sum_{k=0}^{\infty} t(k)x^k - t(0)\right) = \frac{1}{x}\left(G_t(x) - \frac{1}{2}\right). \end{aligned} \tag{5.49}$$

By combining Equations (5.48) and (5.49) we are able to solve for $G_t(x)$ in terms of $G_A(x)$ as follows:

$$\begin{aligned} G_A(x)G_t(x) &= \frac{1}{x}\left(G_t(x) - \frac{1}{2}\right), \\ G_t(x)\left(G_A(x) - \frac{1}{x}\right) &= -\frac{1}{2x}, \\ G_t(x) &= \frac{1}{2 - 2xG_A(x)}, \\ 2G_t(x) &= \frac{1}{1 - xG_A(x)}. \end{aligned}$$

Notice that $2G_t(x)$ is in fact a generating function, which contains the generating function $G_A(x)$. By rewriting $2G_t(x)$ as a power series within a power series, we can use Faà di Bruno's formula to obtain a single power series for this expression. We start by manipulating $G_A(x)$ in the denominator to obtain

$$xG_A(x) = x\sum_{k=0}^{\infty} A(k)x^k = \sum_{k=0}^{\infty} A(k)x^{k+1} = \sum_{k=0}^{\infty} \overline{A}(k+1)x^{k+1} = \sum_{k=0}^{\infty} \overline{A}(k)x^k,$$

where we have defined $\overline{A}(k+1) = A(k)$ for $k \geq 0$ and $\overline{A}(0) = 0$. Now if we define $xG_A(x)$ by $G_{\overline{A}}(x)$, we can write

$$2G_t(x) = \frac{1}{1 - G_{\overline{A}}(x)} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{\infty} \overline{A}(k)x^k \right)^n, \tag{5.50}$$

using a geometric series, where Stirling's formula is used to show that $G_{\overline{A}}(x)$ has a positive radius of convergence. Before applying Faà di Bruno's formula we will need the following definition (Comtet [20, p. 134]).

Definition 5.4. The *exponential partial Bell polynomial* $B_{k,n}$ is given by

$$B_{k,n}(x_1, x_2, \ldots, x_{k-n+1}) := \sum_{\substack{c_1+2c_2+\cdots=k \\ c_1+c_2+\cdots=n}} \frac{n!}{c_1!c_2!\cdots(1!)^{c_1}(2!)^{c_2}\cdots} x_1^{c_1}x_2^{c_2}\cdots\,,$$

where $c_i, k, n \geq 0$.

Here we are partitioning the integer k into the sum of n positive integers, where c_i counts the number of integers i in the partition. By the pigeonhole principle, none of these integers can be larger than $k - n + 1$, and so we need only consider terms c_i for $1 \leq i \leq k - n + 1$. Note that the coefficients of the polynomial are similar to the terms $N(c_1, \ldots, c_q)$ given in Equation (4.8). The following theorem can be found in Comtet ([20, p. 137]).

Theorem 5.7 (*Faà di Bruno*)**.** *Let f and g be two formal power series, given by*

$$f := \sum_{n=0}^{\infty} f_n \frac{u^n}{n!}, \qquad g := \sum_{m=0}^{\infty} g_m \frac{x^m}{m!},$$

where $g_0 = 0$, and let h be the formal power series which is the composition of f and g,

$$h := f \circ g = \sum_{k=0}^{\infty} h_k \frac{x^k}{k!}. \tag{5.51}$$

Then the coefficients h_k are given by

$$h_0 = f_0, \qquad h_k = \sum_{n=1}^{k} f_k B_{k,n}(g_1, g_2, \ldots, g_{k-n+1}),$$

where the $B_{k,n}$ are the exponential partial Bell polynomials.

We can now apply this theorem to rewrite $2G_t(x)$ as a single formal power series.

Theorem 5.8. *The corner terms $t(k)$ in Table 5.2 can be written explicitly as*

$$t(k) = \frac{1}{2} \sum_{j_1+2j_2+\cdots+kj_k=k} \binom{n}{j_1, j_2, \ldots, j_k} A(0)^{j_1} A(1)^{j_2} \cdots A(k-1)^{j_k},$$

for $k \geq 0$ and for $n = j_1 + j_2 + \cdots + j_k$.

Proof: Applying Theorem 5.7 to Equation (5.50) gives

$$f = \sum_{n=0}^{\infty} f_n \frac{u^n}{n!} = \sum_{n=0}^{\infty} u^n,$$

where $f_n = n!$ and $u = \sum_{k=0}^{\infty} \overline{A}(k)x^k$, and

$$g = u = \sum_{m=0}^{\infty} g_m \frac{x^m}{m!} = \sum_{m=0}^{\infty} \overline{A}(m)x^m,$$

where $g_m = m!\overline{A}(m)$. Now we can conclude that $h = \sum_{k=0}^{\infty} h_k \frac{x^k}{k!}$, where

$$\begin{aligned}
h_k &= \sum_{n=1}^{k} f_k B_{k,n}(g_1, g_2, \ldots, g_{k-n+1}) \\
&= \sum_{n=1}^{k} k! B_{k,n}(1!\overline{A}(1), 2!\overline{A}(2), \ldots, (k-n+1)!\overline{A}(k-n+1)) \\
&= \sum_{n=1}^{k} k! \sum_{\substack{j_1+2j_2+\cdots=k \\ j_1+j_2+\cdots=n}} \frac{n!}{j_1!j_2!\cdots(1!)^{j_1}(2!)^{j_2}\cdots} (1!\overline{A}(1))^{j_1}(2!\overline{A}(2))^{j_2}\cdots \\
&\quad \cdots(\overline{A}(k-n+1))^{j_{k-n+1}} \\
&= \sum_{n=1}^{k} \sum_{\substack{j_1+2j_2+\cdots=k \\ j_1+j_2+\cdots=n}} \frac{n!k!}{j_1!j_2!\cdots} \overline{A}(1)^{j_1}\overline{A}(2)^{j_2}\cdots\overline{A}(k-n+1)^{j_{k-n+1}}.
\end{aligned}$$

Notice that by Equation (5.50), $2G_t(x) = h$, and comparing Equations (5.51) and (5.47), we see that

$$2t(k) = \frac{h_k}{k!}.$$

Also, converting the function $\overline{A}$ back to A, we can write

$$t(k) = \frac{1}{2}\sum_{n=1}^{k} \sum_{\substack{j_1+2j_2+\cdots=k \\ j_1+j_2+\cdots=n}} \frac{n!}{j_1!j_2!\cdots} A(0)^{j_1}A(1)^{j_2}\cdots A(k-n)^{j_{k-n+1}}.$$

We can now replace the factorial term with a multinomial coefficient, and simplify by removing the summation over n and simply summing over all possible partitions of k, regardless of the number of terms n in the partition. Since $1 \leq n \leq k$ we must have $1 \leq i \leq k$ in the term j_i. In other words, since n is not fixed, it does not restrict i. This gives

$$t(k) = \frac{1}{2} \sum_{j_1+2j_2+\cdots+kj_k=k} \binom{n}{j_1, j_2, \ldots} A(0)^{j_1}A(1)^{j_2}\cdots A(k-1)^{j_k},$$

where we know $j_1 + j_2 + \cdots = n$, completing the proof. $\diamond$

5.8 An Explicit Formula for Terms $t(n,k)$

We can now extend these ideas to derive a recursive formula for the non-corner terms, $t(n,k)$, in terms of the corner numbers. We will again use a geometric argument to find the coefficients of the recursion.

Start by considering the numbers in Table 5.2 directly below the corner numbers. This gives us the sequence

$$\{t(3k+2,k)\} = 1, 5, 27, 152, 879, \ldots$$

for $k \geq 0$. Note by Theorem 5.1, we can write

$$t(3k+2,k) = \frac{1}{2}t(3k+4,k+1) = \frac{1}{2}t(3(k+1)+1,k+1) = \frac{1}{2}t(k+1).$$

By Equation (5.46), we can write

$$t(3k+2,k) = \frac{1}{2}\sum_{j=0}^{k} A(k-j)t(j),$$

and so the coefficients in the recursion which gives $t(3k+2,k)$ in terms of $t(j)$ are $\frac{1}{2}A(m)$. We can also think about this by counting dots.

Example 5.14. Consider the term $t(14,4) = 879$. We can write this as $879 = 2\cdot\mathbf{304}+271$. By observing Table 5.2 we can write $271 = 2\cdot\mathbf{54}+44+54+65$. Further we have

$$44 = 2\cdot\mathbf{10}+7+8+9,$$
$$54 = 2\cdot\mathbf{10}+7+8+9+10,$$
$$65 = 2\cdot\mathbf{10}+7+8+9+10+11.$$

Visually, we can represent this as Figure 5.17, which gives the sequence of dots in Figure 5.18.

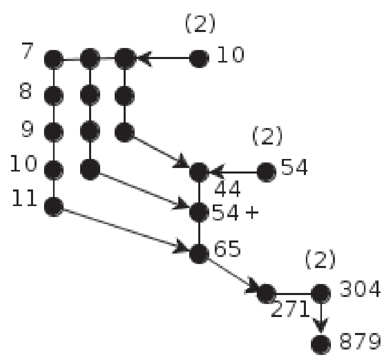


Figure 5.17: Breakdown of 879 into corner terms.

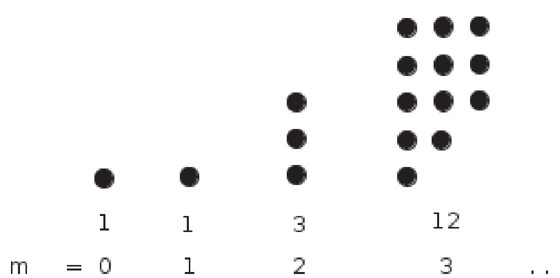


Figure 5.18: Sequence of coefficients $s(3,m)$ for below-corner terms.

This is the same sequence as given in Figure 5.16, and is represented by $s(3, m)$. Note that each corner number is multiplied by 2 and so we must also multiply the number of dots by 2 to obtain the correct coefficient. This gives us $2s(3, m)$, as opposed to $A(m) = 4s(3, m)$. We will instead think about the sequence of dots having its initial column at $m = 1$ and containing one dot, as opposed to the initial size-three column at $m = 2$. Define $\hat{s}(i, m)$ to be the number of dots in the m^{th} shape in the sequence, where i is the number of dots in the column at level $m = 1$. For our example we have

$$\begin{aligned}
\hat{s}(1, 1) &= s(3, 1) = 1, \\
\hat{s}(1, 2) &= s(3, 2) = 3, \\
\hat{s}(1, 3) &= s(3, 3) = 12, \\
\hat{s}(1, m) &= s(3, m) = \frac{1}{4}A(m).
\end{aligned}$$

This is an exceptional case because we can view the initial column as having either size $i = 1$ or 3.

Now, let us continue down the columns of Table 5.2 to those entries which are two below a corner number.

Example 5.15. For $t(15, 4)$, we can write

$$\begin{aligned}
t(15, 4) = 1227 &= 2 \cdot \mathbf{304} + 271 + 348 \\
&= 2 \cdot \mathbf{304} + (2 \cdot \mathbf{54} + 44 + 54 + 65) + (2 \cdot \mathbf{54} + 44 + 54 + 65 + 77).
\end{aligned}$$

Visually, we can represent this as Figure 5.19, from which we obtain the sequence of dots in Figure 5.20, for $m \geq 0$.

This gives us the sequence

$$\{\hat{s}(2, m)\} = 1, 2, 7, 30, \ldots.$$

The $\hat{s}$ function simply shifts the initial column from $m = 2$ in the s function to $m = 1$. We can therefore write

$$\hat{s}(i, m) = s(i, m + 1).$$

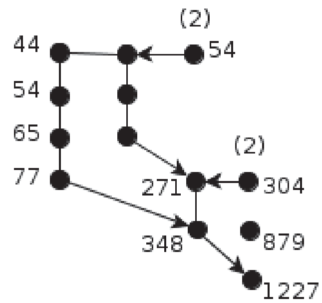


Figure 5.19: Breakdown of 1227 into corner terms.

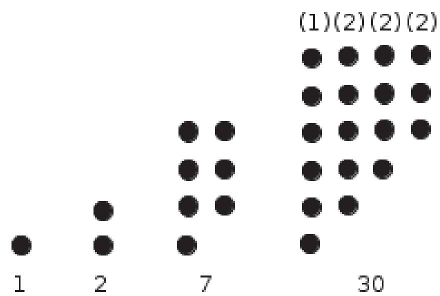


Figure 5.20: Sequence of coefficients for two-below-corner terms, $m \geq 0$.

Note that for numbers one row beneath the corner terms, the initial column for $m = 1$ has one dot, and for numbers two rows beneath the corner term it has two dots. This pattern does in fact continue and it is easy to see by observing Table 5.2, that for numbers i terms below the corner, the initial column has i dots, as in Figure 5.21.

We have defined terms $A(m)$ to be coefficients in the recurrence for writing corner numbers in terms of themselves, and shown them to be equal to $4s(3, m)$. More generally we can define terms $A(i, m)$ to be the coefficients in the recurrence for writing non-corner numbers, lying i rows beneath the corner, in terms of corner numbers, where $m = 0, 1, 2, \ldots$ number the corners in descending order. We have explained above that in terms of counting dots we have

$$A(i, m) = 2\hat{s}(i, m),$$

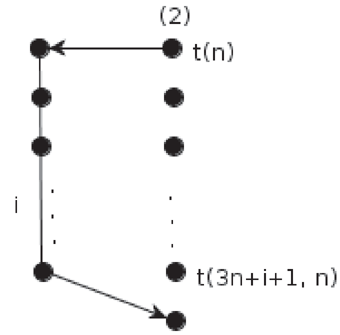


Figure 5.21: Initial column for the case of i terms below the corner.

because corners are multiplied by two. We can now give an expression for $A(i, m)$.

Theorem 5.9. *The coefficients in the recurrence for the non-corner numbers $t(n, k)$ are given by*

$$A(i, m) = \frac{2i}{3m+i}\binom{3m+i}{m},$$

for $m \geq 0$, $i \geq 1$.

Proof: We have by Theorem 5.6 that

$$A(i, m) = 2\hat{s}(i, m) = 2s(i, m+1) = \frac{2i}{(m+1)-1}\binom{i+3(m+1)-4}{i+2(m+1)-2},$$

for $i, m \geq 1$. We can now rewrite this expression as

$$A(i, m) = \frac{2i}{m}\binom{3m+i-1}{2m+i} = \frac{2i}{m}\binom{3m+i-1}{m-1} = \frac{2i}{3m+i}\binom{3m+i}{m},$$

where we have used the fact that $\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$. When $m = 0$, we have $A(i, 0) = 2s(i, 1) = 2$ by definition, which agrees with $\frac{2i}{i}\binom{i}{0}$ as stated in the theorem. $\diamond$

We can now compile our previous results into a theorem concerning non-corner numbers.

Theorem 5.10. *The non-corner numbers $t(n, k)$ in Table 5.2 can be written in terms of the corner numbers $t(k)$ as*

$$t(n, k) = \sum_{j=0}^{k} A(n-3k-1, k-j)t(j),$$

for $n \geq 1$*, where* $t(k)$ *is given explicitly in Theorem 5.8, and* $A(i, m)$ *is given in Theorem 5.9.*

Proof: A term in column k of Table 5.2 will be dependent on corner terms $t(j)$ for $j = 0, \ldots, k$. Each corner term has a coefficient $A(n, m)$, where $A(n, 0)$ corresponds to the largest corner $t(k)$, $A(n, 1)$ corresponds to corner $t(k-1)$ and so on. Also, since n is the row number in Table 5.2, $n - 3k - 1$ is the number of rows from n to the corner term. $\diamond$

Chapter 6

Using the Tree R to Derive Viswanath's Constant

6.1 Further Properties of the Tree R

Rittaud [64] studied a variety of properties of the tree R. Although it would be ideal to be able to explicitly derive the values of the nodes in a given row of tree R, this has proved extremely difficult. Rittaud has a few interesting results here, including the following lemma.

Lemma 6.1. *The nodes in the tree R with value i, where $i \geq 1$, are included in a union of $2\phi(i)$ different walks in R.*

This relies on the fact that parent-child pairs (i, j) in R have i and j relatively prime (recall Proposition 5.1) and the fact that the coefficient cycle $\sigma_3 = (+ - +)$ produces the sequence

$$i,\ j,\ i+j,\ i,\ 2i+j,\ 3i+j,\ i,\ 4i+j,\ 5i+j,\ i,\ 6i+j,\ 7i+j,\ i, \ldots.$$

Rittaud goes even further to say that for any $i \geq 1$ there exists an integer $n(i)$ such that for any $n > n(i)$, there are $2\phi(i)$ nodes with value i amongst rows n, $n+1$ and $n+2$ of R. Rittaud further studies the exact positions of the 0 nodes in the tree T. He also uses the generalized tree $R^{(a,b)}$ to construct the $\mathrm{SL}(2, \mathbb{N})$ tree mentioned in Section 5.2. Here, the matrix entries correspond to the coefficients of a and b in the nodes of $R^{(a,b)}$. The following is an interesting property of the group $\langle A, \hat{B} \rangle$. (Recall from Section 5.1 that multiplying a matrix in the $\mathrm{SL}(2, \mathbb{N})$ tree by $\hat{B}$ and A gives its left and right children respectively.)

Proposition 6.1. *The* $\mathrm{SL}(2, \mathbb{N})$ *tree contains all elements of* $SL(2, \mathbb{N})$ *exactly once.*

Recall from Theorem 2.1 that product matrices $P_n \in K$ (where $K = G \cap \mathrm{SL}(2, \mathbb{Z})$) form a subgroup of index two of $\mathrm{SL}(2, \mathbb{Z})$. Rittaud further studies properties of the determinant and trace trees obtained from the matrix tree $\mathrm{SL}(2, \mathbb{N})$. Lastly, he uses continued fractions to give an expression for the size of a walk in R, i.e., the value of n such that the parent-child pair (i, j) appears with i in the n^{th} row.

We can now compile a list of relations among products of entries in the tree R. Recall that in Chapter 5 we used the notation $S(\tau_n)$ and $S(\rho_n)$ to represent the sums of the elements in the n^{th} rows of T and R respectively. We can similarly use $P(\tau_n)$ and $P(\rho_n)$ to represent the products of the non-zero elements in the n^{th} rows. We will also consider the products of non-zero left and right children in the n^{th} row, which we can denote by $P(\rho_n^-)$ and $P(\rho_n^+)$ respectively. For simplicity, however, we will denote $P(\rho_n)$, $P(\rho_n^-)$ and $P(\rho_n^+)$ by A_n, L_n (not to be confused with the Lucas numbers) and R_n, respectively. Table 6.1 lists these products for $n \leq 10$.

n	L_n	R_n	A_n
1	1	1	1
2	1	1	1
3	1	2	2
4	1	3	3
5	1	15	15
6	4	128	512
7	9	15925	143325
8	225	50426145	11345882625
9	65536	143269901107200	9389336238961459200
10	2282450625	33576172121209656047484375	76635955043162555201615 741396484375

Table 6.1: Products of rows in R for left, right and all nodes.

We can assume that the empty sets of children give product 1. This is the case for L_n with $n \leq 3$ and R_n with $n \leq 2$. A breakdown of the products in Table 6.1 into their prime factorizations did not reveal any patterns.

Theorem 6.1. *The following properties are true for nodes in the tree R:*

1. $A_n = L_n R_n$;

2. $L_{n+3} = L_n R_n^2$;

3. $L_{n+3} = A_n R_n$;

4. $A_n = \sqrt{L_n L_{n+3}}$*;*

5. $L_{n+3} = (R_n R_{n-3} R_{n-6} \cdots)^2$*;*

6. $A_n = R_n (R_{n-3})^2 (R_{n-6})^2 \cdots$*;*

7. $L_{n+3} = \left(\frac{A_n}{A_{n-3}} \cdot fracA_{n-6}A_{n-9} \cdot \cdots \right)^2$*;*

8. $\frac{A_{n+3}}{A_n} = R_n R_{n+3}$.

Proof: Statement 1 is clearly true, as the product of left nodes with right nodes equals the product of all nodes. Statement 2 can easily be derived from the above propositions. We have from Proposition 5.5 that left nodes are equal to their great grandparents and so the left nodes in ρ_{n+3} are derived from nodes in ρ_n. But by Proposition 5.5, all left nodes have one left great grandchild and all right nodes have two. Therefore, each element in ρ_n^- produces one great grandchild (equal to itself) in ρ_{n+3}^- and each element in ρ_n^+ produces two great grandchildren (equal to itself) in ρ_{n+3}^-. Taking products, we have that $L_{n+3} = L_n R_n^2$. Statement 3 is simply a result of substituting Statement 1 into Statement 2. (Rittaud [64] also points out that the n^{th} row of R is included in the $(n+3)^{\text{rd}}$ row of R.)

Rearranging Statement 2 and substituting it into Statement 1 gives,

$$A_n = L_n R_n = L_n \sqrt{\frac{L_{n+3}}{L_n}} = \sqrt{L_n L_{n+3}},$$

which is precisely Statement 4. Statement 5 can be proved by repeatedly substituting Statement 2 into itself to get

$$L_{n+3} = L_n R_n^2 = L_{n-3}(R_{n-3})^2 R_n^2 = L_{n-6}(R_{n-6})^2 (R_{n-3})^2 R_n^2 = (R_n R_{n-3} R_{n-6} \cdots)^2.$$

Here we have $L_1 = L_2 = L_3 = 1$, so this term disappears from the product. Statement 6 follows directly from Statement 5 for L_n by using Statement 3 to rewrite L_{n+3} as $A_n R_n$. We can obtain Statement 7 from Statement 5 by considering the term $R_n R_{n-3} R_{n-6} \cdots$. Substituting Statements 1 and 3 in an alternating fashion we get

$$R_n R_{n-3} R_{n-6} \cdots = \frac{A_n}{L_n} \cdot \frac{L_n}{A_{n-3}} \cdot \frac{A_{n-6}}{L_{n-6}} \cdot \frac{L_{n-6}}{A_{n-9}} \cdot \cdots = \frac{A_n}{A_{n-3}} \cdot \frac{A_{n-6}}{A_{n-9}} \cdot \cdots ,$$

and substituting this expression into Statement 5 gives Statement 7. Lastly, Statement 8 comes directly from Statement 6 by taking a ratio. ◇

Notice that Statements 1 and 2 are independent of each other and the rest are derived from these two. It would be ideal if we could find a third independent relation for the products of nodes in R. This would allow us to solve for an exact expression for the product A_n, which we will soon see would lead us to an expression for Viswanath's constant! Using what we know about products of nodes in R we can attempt to reconstruct A_n. Statement 2 of Theorem 6.1 says that $L_{n+3} = L_n R_n^2$, and so we have a recursion for the product of left nodes in a given row. What can we say about the right nodes? We know from Proposition 5.5 that left nodes have three great grandchildren, one of which we know is equal to itself because it is a left child. We can introduce a constant c and write the product of the great grandchildren of left nodes as $c(L_{n-3})^3$. Similarly, right nodes have five great grandchildren, two of which are equal to itself, so we can write the product of these great grandchildren as $d(R_{n-3})^5$, using a constant d. Combining we have that

$$A_n = \alpha(L_{n-3})^3(R_{n-3})^5,$$

where $\alpha = cd$. Using values in the extension of Table 6.1 (for $n \leq 29$), we are able to determine that $\alpha \approx 1.5836413\ldots$. If this number could be identified, we would have an additional relation among the products of nodes in R and could therefore find an expression for A_n.

Another interesting pattern to note is the following.

Conjecture 6.1. *The last F_n entries of ρ_{n+1} plus the last F_n entries of ρ_{n+2} respectively, give the last F_n entries of ρ_{n+3}. This holds also for subsets of R with the same shape as R.*

Example 6.1. Consider the seventh row of R, which is $\rho_7 = \{5, 1, 7, 3, 5, 7, 3, 13\}$. We know from Proposition 5.11 that $|\rho_7| = F_6 = 8$. The multiset given by the last 8 entries of ρ_8 is $\{7, 1, 9, 5, 9, 11, 5, 21\}$ and for the last 8 entries of ρ_9 we have $\{12, 2, 16, 8, 14, 18, 8, 34\}$. It is easy to see that adding respective members of the first two multisets gives the third, as conjectured.

In Section 4.5 we described the equivalence classes of length-n coefficient cycles which generate periodic coefficient sequences. These sequences approximate the sequences in T when n gets large. We can similarly describe the equivalence classes of coefficient cycles corresponding to sequences in R. We know that we cannot have two consecutive lefts in R, and this is equivalent to prohibiting $(--)$ in our coefficient cycles. The operation of color swapping does not come into play in this case because with the exception of $((+-)^k)$, switching $+$ and $-$ signs will result in the coefficient cycle containing $(--)$, for $n \geq 2$. Again, we will think of our coefficient cycles as necklaces.

The number of necklaces of two colors with no $(--)$ is given by the following sequence (see [67, A000358]) for $n \geq 1$:

$$1, 2, 2, 3, 3, 5, 5, 8, 10, 15, 19, 31, 41, 64, \ldots,$$

and if we restrict these necklaces to be primitive we get the sequence (see [67, A006206])

$$1, 1, 1, 1, 2, 2, 4, 5, 8, 11, 18, 25, 40, 58, \ldots.$$

Notice that prime-indexed terms in A006206 are one less than those in A000358. This is due to the fact that when n is prime, we only have one non-primitive coefficient cycle, namely $((+)^n)$, as was the case for sequences $B_s(n,2)$ and $B_{sp}(n,2)$ in Chapter 4. The former sequence has formula

$$a_n = \frac{1}{n} \sum_{d|n} \phi\left(\frac{n}{d}\right)(F_{d-1} + F_{d+1}),$$

as given in [67, A000358].

Example 6.2. For $n = 8$, the eight necklaces with no $(--)$ (in A000358) are $(++++++++)$, $(+++++++-)$, $(+++++-+-)$, $(+++-+-+-)$, $(++++-++-)$, $(++-++-+-)$, $(+++-+++-)$ and $(+-+-+-+-)$. If we restrict this list to primitive necklaces (in A006206), the first and the last two necklaces are removed, and we are left with five.

We can further restrict the number of coefficient cycles by considering them as bracelets rather than necklaces. The number of bracelets of two colors with no $(--)$

is given by the following sequence (see [67, A129526]) for $n \geq 1$:

$$1, 2, 2, 3, 3, 5, 5, 8, 9, 14, 16, 26, 31, 49, \ldots.$$

If we restrict these bracelets to be primitive, we obtain the number of equivalence classes of coefficient sequences corresponding to sequences in R. The first few terms are

$$1, 1, 1, 1, 2, 2, 4, 5, 7, \ldots.$$

Notice that the this sequence is identical to A006206 until $n = 9$, and similarly A129526 is identical to A000358 until this point. This is the first instance where reversing a coefficient cycle (equivalently, flipping a bracelet over in the plane) produces a different coefficient cycle under rotation. Specifically, for $n = 9$ we have the equivalent bracelets $(+++-++-+-)$ and $(+++-+-++-)$. This distinction occurs whenever we have three different sized groups of $+$ signs in our coefficient cycle. Recall the case in Example 4.5, where for $n = 7$ we had equivalent bracelets $(+++-+--)$ and $(+++--+-)$, which were distinct necklaces. This does not occur at $n = 7$ for coefficient cycles in R because $(--)$ is prohibited.

The growth types and rates of our equivalence classes can be found by considering the trace or order of the corresponding product matrices, as was done in Chapter 3. Recall that length n coefficient cycles are used only to approximate sequences in R. The tree R does not contain repeated edges, and so it cannot possibly contain a bounded (periodic) sequence. It again proves difficult, however, to determine the sizes of the equivalence classes and construct a formula for the average of growth rates.

We can also look at the equivalence classes of coefficient cycles of R_1, the variation of R discussed in Section 5.1. We saw that coefficient cycles which approximate sequences in R_1 cannot contain $(+-+)$ or $(-+-)$. Also recall that in R and R_1, the first direction we travel is right, and so coefficient cycles must begin with $+$. Notice that we cannot talk about necklaces in this case because we must take into account the fact that some or all rotations of a given coefficient cycle may be restricted because they contain $(+-+)$ or $(-+-)$. By definition, a necklace is equivalent to itself under rotation. We may still speak of equivalence classes of allowable coefficient cycles, where equivalence under rotation applies only to cycles with no $(+-+)$ or $(-+-)$.

Let us first consider the number of coefficient cycles with equivalence under rotation, where $(+-+)$ and $(-+-)$ are not allowed. We can split our cycles into two types - those of the form $((+)^i(-)^j)$, and those which do not have this form. Coefficient cycles of this simple form are easy to count. If the cycle is of length n, it may begin with i $+$ signs, where $1 \leq i \leq n$. Therefore we have n distinct coefficient cycles of this type. The number of coefficient cycles not of this type is 0 for $n \leq 5$ due to the restrictions in place. For $6 \leq n \leq 9$ we have $1, 4, 10, 18$. The first instance of a non-simple coefficient cycle is $(+--++-)$. Note that none of its rotations are allowable. Combining both types we have that the number of coefficient cycles with equivalence under rotation for $n \geq 1$ is given by the sequence

$$1, 2, 3, 4, 5, 7, 11, 18, 27, \ldots.$$

Let us now remove equivalence under rotation and simply count the number of coefficient cycles appearing in R_1. We know that there are F_{n-1} terms in row ρ_n of R, but to reach ρ_n, it requires $n-2$ choices of $+$ or $-$ (i.e., left or right in the tree). Therefore, if we have n terms in our coefficient cycle, there are F_{n+1} different coefficient cycles of length n. This also holds for R_1 because it contains the same set of sequences as R by Proposition 5.6. Now let us count the total number of coefficient cycles of each type. For cycles with form $((+)^i(-)^j)$, we need to count the number of allowable rotations. If we have $i \leq n-2$ $+$ signs for $n \geq 3$, all rotations which start with $+$ are allowable (recall they must start with $+$), so there are i coefficient sequences of this form. If $i = n-1$, our cycle is of the form $(+\cdots+-)$ and any rotation starting with $+$ would contain $(+-+)$. Also, if $i = n$, all rotations give us the same cycle back. The total number of coefficient cycles of this simple type is therefore given by

$$2 + \sum_{i=1}^{n-2} i,$$

for $n \geq 3$. For $n = 1$ we have the coefficient cycle $(+)$ and for $n = 2$ we have $(++)$ and $(+-)$. Therefore for $n \geq 1$, the total number of coefficient cycles of this type, is given by the sequence

$$1, 2, 3, 5, 8, 12, 17, 23, 30, \ldots.$$

Since we know that there are F_{n+1} coefficient cycles of length n, the number of coefficient cycles not of the form $((+)^i(-)^j)$ is

$$F_{n+1} - \left(2 + \sum_{i=1}^{n-2} i\right), \tag{6.1}$$

for $n \geq 3$. For $n = 1, 2$, this value is 0, and in fact it is 0 for $n \leq 5$. Interestingly, the sequence defined by Equation (6.1), namely

$$0, 0, 0, 0, 0, 1, 4, 11, 25, 51, 97, 176, 309, \ldots,$$

can be found in [67, A014162] for $n \geq 6$, and its formula is given as

$$F_{n+1} - \frac{1}{2}(n^2 - 3n + 6).$$

It describes the partial sum operator applied three times to Fibonacci numbers, and also "the number of 132-avoiding two-stack sortable permutations which contain exactly one subsequence of type 51234, with offset 4." This connection to permutations with forbidden subsequences is something which could be further examined.

Example 6.3. For $n = 7$, one of our coefficient cycles is $(+ - - - + + -)$, and we also have the cycle obtained by switching signs and reversing, $(+ - - + + + -)$. The coefficient cycles obtained by switching signs in the former cycle or reversing it do not meet our criteria because those cycles and all possible rotations either begin with $-$, or contain $(+ - +)$ or $(- + -)$. For $n = 8$, the coefficient cycle $(+ + + - - + + -)$, its sign-switched counterpart $(+ - - - + + - -)$, as well as the reversals of both of these, $(+ + - - + + + -)$ and $(+ - - + + - - -)$ all occur. This example highlights the importance of the initial term in a coefficient cycle. Previously, when dealing with necklaces, we could always rotate our cycle so that it began with $+$.

If we now restrict ourselves to the number of equivalence classes (primitive coefficient cycles with equivalence under rotation, reversal and color swapping), we have the following sequence for $n \geq 1$:

$$1, 1, 1, 2, 2, 4, 5, 8, 12, \ldots.$$

Curiously, this sequence matches that of the number of primitive necklaces with no $(--)$ ([67, A006206]) term up to $n = 8$.

Example 6.4. For $n = 8$, the 8 equivalences classes of coefficient cycles with $(+-+)$ and $(-+-)$ prohibited are $(+++++++-)$, $(++++++--)$, $(+++++---)$, $(++++----)$, $(+++--++-)$, $(+--++++-)$, $(+--+++--)$, $(+---+++-)$. The equivalence class of $(+++--++-)$ contains the rotation $(-+++--++)$, but none of the other six. The equivalence class of $(+--++++-)$ does not contain any of its rotations. For $n = 9$, the equivalence class of $(+++--+++--)$ contains all of its rotations that begin with $+$ because neither $(+-+)$ nor $(-+-)$ occur.

6.2 A New Computation of Viswanath's Constant

Recall from Equation (3.15) that we can approximate Viswanath's constant by calculating the growth rate of every periodic coefficient sequence of length n, taking the average, and then letting n go to infinity. Each growth rate was given by $|\lambda_1|^{\frac{1}{n}}$, where λ_1 was the dominant eigenvalue of the product matrix P_n. Recall from Chapter 5 that we have also used $\lambda_{n,i}$ to denote the dominant eigenvalue of $P_{n,i}$, the i^{th} product matrix in row n. Viswanath's constant, which it will soon be convenient to denote as τ rather than e^{γ_f}, can be rewritten as

$$\tau = \lim_{n\to\infty} \frac{|\lambda_{n,1}|^{\frac{1}{n}} + |\lambda_{n,2}|^{\frac{1}{n}} + \cdots + |\lambda_{n,i}|^{\frac{1}{n}} + \cdots + |\lambda_{n,2^n}|^{\frac{1}{n}}}{2^n},$$

There are 2^n periodic coefficient sequences of length n, and so 2^n product matrices. Further, we also approximated Viswanath's constant using $|\lambda_1| \approx |\mathrm{tr}(P_n)|$, as in Equation (3.17). Both approximations to τ converged very slowly.

We can now give a formulation for Viswanath's constant that does not take matrices into consideration. In [44], Kalmár-Nagy notes that we can do this using the geometric mean of the n^{th} terms in our periodic coefficient sequences, but does not provide a proof.

Theorem 6.2. *Viswanath's constant can be written as the limit*

$$\tau = \lim_{n\to\infty} \left(\prod_{\text{all } t_n \neq 0} |t_n| \right)^{\frac{1}{n2^{n-2}}}, \tag{6.2}$$

where t_n is the n^{th} term in a random Fibonacci sequence, and the product is taken over all possible non-zero n^{th} terms in absolute value.

Proof: Let $0 < \varepsilon < \frac{1}{2}$. From Theorem 1.1 we have that

$$\tau = \lim_{n \to \infty} |t_n|^{\frac{1}{n}},$$

almost surely. This means that there is a positive integer N such that, with the exception of a proportion ε of the 2^{N-2} terms at level N, we have

$$\tau - \varepsilon \leq |t_{N,i}|^{\frac{1}{N}} \leq \tau + \varepsilon, \tag{6.3}$$

where $t_{n,i}$ denotes the n^{th} term in the i^{th} random Fibonacci sequence. If there were no such N, then a positive proportion of terms would always lie outside of that interval, which would be a contradiction to Theorem 1.1.

Now, relabel the terms $t_{N,i}$ at level N, where $i = 1, 2, \cdots, 2^{N-2}$, such that those in the exceptional set occur first. We then have that the $t_{N,i}$ with

$$i = 1, 2, \ldots, \lfloor \varepsilon 2^{N-2} \rfloor,$$

are in the exceptional set. Note that for any term $|t_{n,i}|^{\frac{1}{n}}$ we have

$$1 \leq |t_{n,i}|^{\frac{1}{n}} \leq \phi, \tag{6.4}$$

because any term $|t_n|$ lies between 1 and F_n in τ_n. Set $i_0 := \lfloor \varepsilon 2^{N-2} \rfloor$, and consider the geometric mean

$$P(N) := \left(\prod_{i=1}^{2^{N-2}} |t_{N,i}|^{\frac{1}{N}} \right)^{\frac{1}{2^{N-2}}}, \tag{6.5}$$

where zero-terms have been removed from the product. Now, split this product into two parts; those terms which belong in the exceptional set, and those which do not. This gives

$$\begin{aligned}
P(N) &= \left(\prod_{i=1}^{i_0} |t_{N,i}|^{\frac{1}{N}} \prod_{i=i_0+1}^{2^{N-2}} |t_{N,i}|^{\frac{1}{N}} \right)^{\frac{1}{2^{N-2}}} \\
&= \left(\prod_{i=1}^{i_0} \frac{|t_{N,i}|^{\frac{1}{N}}}{\tau} \right)^{\frac{1}{2^{N-2}}} \left(\tau^{i_0} \prod_{i=i_0+1}^{2^{N-2}} |t_{N,i}|^{\frac{1}{N}} \right)^{\frac{1}{2^{N-2}}} &(6.6) \\
&= P_1(N) P_2(N). &(6.7)
\end{aligned}$$

Here we have added i_0 terms τ to the product on the right-hand side so that we have a geometric mean, $P_2(N)$, of 2^{N-2} terms. Since a geometric mean lies between the smallest and the largest elements, we have by the inequality in (6.3) (which includes the term τ),

$$\tau - \varepsilon \le P_2(N) \le \tau + \varepsilon. \tag{6.8}$$

Next, in order to estimate $P_1(N)$, we note that

$$\frac{i_0}{2^{N-2}} \le \frac{\varepsilon 2^{N-2}}{2^{N-2}} = \varepsilon. \tag{6.9}$$

Also, we will use the following lemma from Brown, Dilcher and Manna [12]: For $0 < x < \frac{1}{3}$ we have $e^x < 1 + \frac{6}{5}x$. This, together with the form for $P_1(N)$ given in Equation (6.6) and the inequalities in (6.4) and (6.9), we have

$$P_1(N) \le \left(\frac{\phi}{\tau}\right)^{i_0/2^{N-2}} \le \left(\frac{\phi}{\tau}\right)^{\varepsilon} = e^{\varepsilon \log(\phi/\tau)} < 1 + \frac{6}{5}\log\left(\frac{\phi}{\tau}\right)\varepsilon < 1 + \frac{\varepsilon}{2}. \tag{6.10}$$

(Note that $0 < \varepsilon \log(\phi/\tau) < 0.178618\ldots$ and $\frac{6}{5}\log(\phi/\tau) = 0.428683\ldots$.) On the other hand, by the inequalities in (6.4) and (6.9) again,

$$P_1(N) \ge \left(\frac{1}{\tau}\right)^{i_0/2^{N-2}} \ge \left(\frac{1}{\tau}\right)^{\varepsilon} = e^{-\varepsilon \log(\tau)} > 1 - \log(\tau)\varepsilon > 1 - \frac{\varepsilon}{8}. \tag{6.11}$$

(Note that we have used the facts that $e^x > 1 + x$ for all real x, and $\log(\tau) = 0.123975\ldots$.)

Combining Equation (6.7) with the inequalities in (6.8), (6.10) and (6.11), we get

$$\left(1 - \frac{\varepsilon}{8}\right)(\tau - \varepsilon) \le P(N) \le \left(1 + \frac{\varepsilon}{2}\right)(\tau + \varepsilon). \tag{6.12}$$

To simplify these inequalities, we note that

$$\left(1 - \frac{\varepsilon}{8}\right)(\tau - \varepsilon) = \tau - \left(1 + \frac{\tau}{8}\right)\varepsilon + \frac{1}{8}\varepsilon^2 > \tau - 1.15\varepsilon$$

and, since $\varepsilon < \frac{1}{2}$,

$$\left(1 + \frac{\varepsilon}{2}\right)(\tau + \varepsilon) = \tau + \left(\frac{\tau}{2} + 1 + \frac{\varepsilon}{2}\right)\varepsilon < \tau + 2\varepsilon.$$

(Note here that $\frac{\tau}{8} = 0.1414985\ldots$ and $\frac{\tau}{2} + \frac{1}{4} + 1 = 1.8159941\ldots$.) Combining this with the inequality in (6.12), we have that

$$|P(N) - \tau| < 2\varepsilon,$$

which proves the theorem since ε was arbitrary. $\diamond$

We can also give a rough sketch of the proof of the above theorem by multiplying the growth rate formulas for each random Fibonacci sequence $t_{n,i}$ as follows:

$$\lim_{n\to\infty} |t_{n,1}|^{\frac{1}{n}} \lim_{n\to\infty} |t_{n,2}|^{\frac{1}{n}} \cdots \lim_{n\to\infty} |t_{n,i}|^{\frac{1}{n}} \cdots \lim_{n\to\infty} |t_{n,2^{n-2}}|^{\frac{1}{n}} = \tau^{2^{n-2}}.$$

We can rewrite this as the limit of a product to give

$$\lim_{n\to\infty} |t_{n,1}t_{n,2}\cdots t_{n,i}\cdots t_{n,2^{n-2}}|^{\frac{1}{n}} = \tau^{2^{n-2}},$$

where any zero term is removed without consequence. Now, taking the root of order 2^{n-2} gives

$$\lim_{n\to\infty} |t_{n,1}t_{n,2}\cdots t_{n,i}\cdots t_{n,2^{n-2}}|^{\frac{1}{n2^{n-2}}} = \tau,$$

as required.

By Corollary 1.1, the multiset of all possible n^{th} terms $|t_n|$ is the n^{th} row, τ_n, of T. Therefore, we may think of the tree T as being comprised of the set of random Fibonacci sequences $\{|t_n|\}$ in absolute value, or equivalently, the set of sequences $\{\widetilde{f_n}\}$, as originally defined. Using our product notation, Equation (6.2) can now be written as

$$\tau = \lim_{n\to\infty} P(\tau_n)^{\frac{1}{n2^{n-2}}},$$

where again the product is taken over non-zero terms. It is interesting to note that the result of Corollary 5.5, namely

$$\lim_{n\to\infty} \frac{\mathbb{E}|t_n|}{\mathbb{E}|t_{n-1}|} = \alpha - 1 = 1.20556943\ldots,$$

gives us the growth rate of the arithmetic mean of terms in row τ_n (using the ratio form of Definition 2.4). This is in contrast to τ, which gives us the growth rate of the geometric mean of terms in τ_n. The theorem of arithmetic and geometric means (see Hardy et al. [36, p. 17]), which states that the arithmetic mean of a list of non-negative real numbers is always greater than or equal to the geometric mean of those numbers, is satisfied here. Note that we have already verified the inequality between these two numbers using the generalized mean inequality in Chapter 5.

It is also interesting that Theorem 6.2 for the growth rate does not depend on individual sequences, only products of terms at each level. We could therefore rearrange the nodes at each level in T, retaining the same multisets τ_n but creating a new set of sequences with the same a.s. growth rate, τ. So do the individual sequences really matter, or can we concern ourselves with the multisets τ_n only? Suppose we generate a sequence by selecting one element at random from τ_n for each $n \geq 1$. What is its growth rate? The value of the randomly selected element is the expected value (arithmetic mean) of elements in τ_n, and so the sequence will have growth rate $1.20556943\ldots$. This is in contrast to τ, which would occur if the expected value used the geometric mean instead. Perhaps thinking about this problem in a different way could instead lead us to the geometric mean.

We can apply the above ideas about the geometric mean to the reduced tree R defined in Section 5.1. We must first assume that the following limit exists for sequences $\{r_n\}$ in R:

$$\lim_{n\to\infty} {r_n}^{\frac{1}{n}} = \rho,$$

where the constant ρ is the a.s. growth rate of sequences in R and is analogous to τ in T. Since terms r_n are nodes in R, they are positive. In [64], Rittaud gives a heuristic argument for the following result; as an important contribution, we have made it rigorous.

Theorem 6.3. *The following link exists between growth rates $\tau = 1.13198824\ldots$ and $\rho = 1.33683692\ldots$:*

$$\tau = \rho^{1-\frac{3}{2\phi^2}} = \rho^{\frac{3\sqrt{5}-5}{4}}. \tag{6.13}$$

Proof: By definition, we have that $\tau = \lim_{n\to\infty} |t_n|^{\frac{1}{n}}$ and $\rho = \lim_{n\to\infty} {r_n}^{\frac{1}{n}}$, almost surely. We want to project sequences $\{|t_n|\}$ in T into R. This is possible because we know that R is a reduction of the tree T and so contains all edges which appear in T. We also know that R contains only the first appearance of every edge in T. Let $\theta(n)$ be the row number in R where we find $|t_n|$ in the edge $(|t_{n-1}|, |t_n|)$, so that $r_{\theta(n)} = |t_n|$. Note that $\theta(n) \leq n$. We have that for sequences in T which also occur in R, $\theta(n) = n$, i.e., there are no reductions. In the projection of sequences in T into R, we have

$$\rho = \lim_{n\to\infty} {r_{\theta(n)}}^{\frac{1}{\theta(n)}}, \tag{6.14}$$

almost surely. This is because projecting a sequence in T into R should give the same growth rate as that sequence in R. Recalling that sequences $|t_n|$ are positive in T, we can write, almost surely,

$$\tau = \lim_{n\to\infty} |t_n|^{\frac{1}{n}} = \lim_{n\to\infty} r_{\theta(n)}{}^{\frac{1}{n}} = \lim_{n\to\infty} r_{\theta(n)}{}^{\frac{1}{n}\cdot\frac{\theta(n)}{\theta(n)}} = \lim_{n\to\infty} r_{\theta(n)}{}^{\frac{1}{n\theta(n)}\cdot\theta(n)}. \tag{6.15}$$

We now want to find an expression for $\theta(n)$, i.e., the row number in which we find $|t_n|$ in R. We will use an expected value to find an expression for $\theta(n+1)$, i.e., the row number in which we find $|t_{n+1}|$ in R, based on the fact that given $\theta(n)$, there are two different values $\theta(n+1)$ can take on. Recall that if $|t_{n+1}|$ is a child of $|t_n|$ in T, it is not necessarily a child in R because in this tree we have removed branches. Corollary 5.1 says that

$$c_T(\rho_n) = \rho_{n+1} \uplus \rho_{n-2},$$

i.e., children in T of nodes belonging to ρ_n may appear in row $n+1$ or $n-2$. Let p_n be the probability of finding the edge $(|t_n|, |t_{n+1}|)$ at a previous level of R, i.e., the probability that $|t_{n+1}|$ is the left child of a left node $|t_n|$, where $|t_n|$ occurs at level $\theta(n)$ in R (which means it is in the set $\rho_{\theta(n)}$). The expected value can be written as

$$\mathbb{E}(\theta(n+1)) = (1-p_n)(\theta(n)+1) + p_n(\theta(n)-2).$$

We can also think of this expression as a random recurrence by writing

$$\theta(n+1) = \theta(n) + 1 - 3\mu_n,$$

where $\mu_n = 1$ with probability p_n and $\mu_n = 0$ with probability $1-p_n$. Solving gives

$$\theta(n) = \sum_{i=0}^{n-1}(1-3\mu_i) = n - 3\sum_{i=0}^{n-1}\mu_i,$$

where we have assumed $\theta(0) = 0$. The second sum tells us the number of times $\mu_i = 1$. In order to evaluate this sum we must use an expected value as follows:

$$\mathbb{E}(\theta(n)) = n - 3\sum_{i=0}^{n-1}\mathbb{E}(\mu_i) = n - 3\sum_{i=0}^{n-1} p_i = n\left(1 - \frac{1}{n}\sum_{i=0}^{n-1} p_i\right) \tag{6.16}$$

We can write p_n as the ratio of the number of missing children of $\rho_{\theta(n)}$ in $\rho_{\theta(n)+1}$ over the total number of children of $\rho_{\theta(n)}$ in T and so

$$p_n = \frac{|\rho_{\theta(n)-2}|}{|\rho_{\theta(n)+1} \uplus \rho_{\theta(n)-2}|}.$$

We have $|\rho_n| \geq 1$ for $n \geq 1$ and so we will assume $|\rho_n| = 0$ for $n \leq 0$, in which case $p_n = 0$ and hence $\mu_n = 0$ with probability 1 also. We know from Proposition 5.11 the sizes of rows in R, and so

$$p_n = \frac{F_{\theta(n)-3}}{F_{\theta(n)} + F_{\theta(n)-3}} = \frac{F_{\theta(n)-3}}{2F_{\theta(n)-1}}.$$

We can now write Equation (6.16) as

$$\mathbb{E}(\theta(n)) = n\left(1 - \frac{1}{n}\sum_{i=0}^{n-1}\frac{F_{\theta(i)-3}}{2F_{\theta(i)-1}}\right) \tag{6.17}$$

We will be interested in taking the limit of the above sum. Note that if we randomly choose a sequence in T, we know with probability 1 that it must grow exponentially. If we then project this sequence into R, it still grows exponentially because it is the same sequence. The difference is that in R we may be removing some instances of repetition from the sequence, implying $\theta(n) \leq n$, but $\theta(n)$ must still grow infinitely large. Otherwise, the sequence must stay within a finite number of rows in R implying it is bounded, which is a contradiction. Therefore, as $n \to \infty$ we must also have $\theta(n) \to \infty$, implying

$$\lim_{n\to\infty}\frac{F_{\theta(n)-3}}{2F_{\theta(n)-1}} = \lim_{n\to\infty}\frac{F_{n-3}}{2F_{n-1}} = \frac{1}{2\phi^2}.$$

By Proposition 5.12 for Cesàro means, we have

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\frac{F_{\theta(i)-3}}{2F_{\theta(i)-1}} = \lim_{n\to\infty}\frac{F_{\theta(n-1)-3}}{2F_{\theta(n-1)-1}} = \frac{1}{2\phi^2}. \tag{6.18}$$

Since we have an expression for the expected value of $\theta(n)$, we can rewrite Equation (6.15) as

$$\tau = \lim_{n\to\infty}\left(r_{\theta(n)}\right)^{\frac{1}{n\mathbb{E}(\theta(n))}\cdot\mathbb{E}(\theta(n))}, \tag{6.19}$$

keeping in mind that limit expressions for τ and ρ hold almost surely. Substituting Equation (6.17) into this expression for τ gives

$$\begin{aligned}\tau &= \lim_{n\to\infty}\left(r_{\theta(n)}\right)^{\frac{1}{n\mathbb{E}(\theta(n))}\left(1-\frac{1}{n}\sum_{i=0}^{n-1}\frac{F_{\theta(i)-3}}{2F_{\theta(i)-1}}\right)n} \\ &= \lim_{n\to\infty}\left(r_{\theta(n)}\right)^{\frac{1}{\mathbb{E}(\theta(n))}\left(1-\frac{1}{n}\sum_{i=0}^{n-1}\frac{F_{\theta(i)-3}}{2F_{\theta(i)-1}}\right)}.\end{aligned}$$

Also, because $\mathbb{E}(\theta(n))$ almost surely goes to infinity at the same rate as $\theta(n)$ does (recall we have seen that the latter goes infinity almost surely), we can modify Equation (6.14) to give

$$\rho = \lim_{n\to\infty} (r_{\theta(n)})^{\frac{1}{\mathbb{E}(\theta(n))}}. \tag{6.20}$$

We can now use the fact that if $\lim_{n\to\infty} f(n)$ and $\lim_{n\to\infty} g(n)$ both exist, we have

$$\lim_{n\to\infty} f(n)^{g(n)} = \lim_{n\to\infty} f(n)^{\lim_{n\to\infty} g(n)}.$$

This is straightforward to see if we take the logarithm and use limit laws. Using Equation (6.18), we can now write

$$\tau = \lim_{n\to\infty} (r_{\theta(n)})^{\frac{1}{\mathbb{E}(\theta(n))}\lim_{n\to\infty}\left(1-\frac{1}{n}\sum_{i=0}^{n-1}\frac{F_{\theta(i)}-3}{2F_{\theta(i)}-1}\right)}, \tag{6.21}$$

almost surely. Substituting Equations (6.20) and (6.18) into the above gives

$$\tau = \rho^{1-\frac{3}{2\phi^2}},$$

completing the proof. $\diamond$

We can now calculate ρ to be approximately 1.33683692. Similarly, finding a good approximation of ρ will yield a good approximation of τ. The following theorem from Gelfond [31, p. 106] could shed light on the nature of Viswanath's constant, if more was known about the constant ρ.

Theorem 6.4 (Gelfond-Schneider)**.** *Given numbers α and β, where $\alpha \neq 0, 1$ is algebraic and β is algebraic and irrational, the number α^β is transcendental.*

Let $\alpha = \rho$ and $\beta = 1 - \frac{3}{2\phi^2} = \frac{3\sqrt{5}-5}{4}$. We know that β is an algebraic irrational, and so if it could be shown that ρ is algebraic, Viswanath's constant τ would necessarily be transcendental.

We can now write the constant ρ using a geometric mean, as in Theorem 6.2. This formula will soon lead us to the most important result of the thesis.

Theorem 6.5. *The almost sure growth rate of a random Fibonacci sequence $\{r_n\}$ in the tree R is given by*

$$\lim_{n\to\infty} P(\rho_n)^{\frac{1}{nF_{n-1}}} = \rho, \tag{6.22}$$

where ρ_n is the n^{th} row in R.

Proof: We will use the argument given in Theorem 6.2, with a few minor changes. Replacing τ with ρ, $|t_n|$ with r_n and 2^{n-1} with F_{n-1} (recall Proposition 5.11, which states that there are F_{n-1} entries in ρ_n), we can follow the proof up to Equation (6.10), where the inequalities remain true. Here we can write

$$P_1(N) \leq \left(\frac{\phi}{\rho}\right)^{i_0/2^{N-2}} \leq \left(\frac{\phi}{\rho}\right)^{\varepsilon} = e^{\varepsilon \log(\phi/\rho)} < 1 + \frac{6}{5}\log\left(\frac{\phi}{\rho}\right)\varepsilon < 1 + \frac{\varepsilon}{2},$$

where $0 \leq \varepsilon \log(\phi/\rho) \leq 0.095452754$, and so we are permitted to use the result of Brown et al. [12]. Also, $\frac{6}{5}\log\left(\frac{\phi}{\rho}\right) = 0.22908661$, so the final inequality holds. The bounds in Equation (6.11) can be adjusted slightly to give

$$P_1(N) \geq \left(\frac{1}{\rho}\right)^{i_0/2^{N-2}} \geq \left(\frac{1}{\rho}\right)^{\varepsilon} = e^{-\varepsilon \log(\rho)} > 1 - \log(\rho)\varepsilon > 1 - \frac{\varepsilon}{3}.$$

Here, $\log \rho = 0.29030631$ so we moved the lower bound from $1 - \frac{\varepsilon}{8}$ to $1 - \frac{\varepsilon}{3}$. Equation (6.12) can now be written as

$$\left(1 - \frac{\varepsilon}{3}\right)(\rho - \varepsilon) \leq P(N) \leq \left(1 + \frac{\varepsilon}{2}\right)(\rho + \varepsilon).$$

The upper bound is again bounded by $\rho + 2\varepsilon$, and the lower bound can be written as

$$\left(1 - \frac{\varepsilon}{3}\right)(\rho - \varepsilon) = \rho - \left(1 + \frac{\rho}{3}\right)\varepsilon + \frac{1}{3}\varepsilon^2 > \rho - 1.5\varepsilon$$

because $\frac{\rho}{3} = 0.4456123067$. Again we can conclude that

$$|P(N) - \rho| < 2\varepsilon,$$

completing the proof. $\diamond$

We need not take absolute values of the product because all terms in the tree R are positive. Note that the growth rate of the arithmetic mean of terms in ρ_k was given in Corollary 5.4 as $2.20556943/\phi = 1.363116873$.

Table 6.2 gives approximations of τ and ρ using the root form of Definition 2.4 as well as an approximation of τ obtained from ρ using Rittaud's formula. The approximation of τ given in column 1 is achieved using the n^{th} root of the geometric mean of nodes in τ_n, and the Maple program for this geometric mean is given in Figure A.2 of Appendix A. Here we are calculating and storing all nodes $f(k, n)$ in the k^{th} position of the n^{th} row of the tree T based on the value of $n \pmod 4$. The approximation is poor, however, so it was not worth the long computing time to continue for higher values of n.

The fact that random Fibonacci sequences in R do not grow at the rate $\tau = 1.13198824\ldots$ is not contradictory. At level n, we have F_{n-1} sequences in R and 2^{n-2} sequences in T. Since

$$\lim_{n\to\infty} \frac{F_{n-1}}{2^{n-2}} = 0, \tag{6.23}$$

a random Fibonacci sequence chosen at random from T will almost surely have growth rate τ.

We have seen in Definition 2.4 that we can equally well define the growth rate of a sequence in terms of a ratio, rather than an n^{th} root. For sequences $P(\tau_n)^{\frac{1}{2^{n-2}}}$ and $P(\rho_n)^{\frac{1}{F_{n-1}}}$, the growth rates can also be determined by the respective limits

$$\tau = \lim_{n\to\infty} \frac{|P(\tau_n)|^{\frac{1}{2^{n-2}}}}{|P(\tau_{n-1})|^{\frac{1}{2^{n-3}}}}, \qquad \rho = \lim_{n\to\infty} \frac{P(\rho_n)^{\frac{1}{F_{n-1}}}}{P(\rho_{n-1})^{\frac{1}{F_{n-2}}}}. \tag{6.24}$$

As can be seen in Table 6.3, this limit actually gives us a much better approximation to the growth rate ρ than Equation (6.22). We have that $\rho = 1.33683692\ldots$ and again we use Rittaud's formula to approximate $\tau = 1.13198824\ldots$ given ρ. This approximation for $n = 39$ is correct to eight decimal places, and we feel this is the most important result of the thesis. As in Table 6.2, the approximation to τ in column 1 using the ratio of geometric means is far worse than that in Column 3, so this method was not pursued. It is important to note that this good approximation of Viswanath's constant is a result of using the reduced tree R. Because there are far fewer sequences at level n in tree R than in tree T, there were less computations needed in the calculation.

A portion of the Maple program used to calculate $P(\rho_n)^{\frac{1}{F_{n-1}}}$ (from which we can calculate the root or ratio definition of ρ) for $n = 24$ is given in Figure A.1 in Appendix A. To speed up the process, we divided the tree R into 13 subtrees and calculated the product of nodes for each. This required finding the product matrices for the 13 length-6 coefficient cycles at ρ_8 in R. Multiplying the product matrix by the vector $[1,1]^T$ gives us the initial values for each of the 13 subtrees. By taking further product matrices, we can then calculate all possible descendants at a given

n	$\lvert P(\tau_n)\rvert^{\frac{1}{n2^{n-2}}}$	$P(\rho_n)^{\frac{1}{nF_{n-1}}}$	τ approximation
12	1.114731452	1.27818405041443	1.11050575538334
13	1.107641902	1.28245477393190	1.11208880100941
14	1.108178454	1.28627159706030	1.11350104361417
15	1.116093473	1.28960968771617	1.11473418605353
16	1.111784082	1.29248979778539	1.11579667451891
17	1.112282714	1.29505783713997	1.11674289496682
18	1.117453610	1.29734878466390	1.11758611120501
19	1.114747778	1.29939403349511	1.11833817400352
20	1.115191136	1.30124111524162	1.11901678573824
21	1.118703154	1.30291522613447	1.11963137191158
22		1.30443763984302	1.12018987554398
23		1.30582984991015	1.12070028661417
24		1.30710747270355	1.12116841350243
25		1.30828374124988	1.12159917199075
26		1.30937058135497	1.12199698391308
27		1.31037774307515	1.12236546260294
28		1.31131361211891	1.12270771284350
29		1.31218556072738	1.12302646126611
30		1.31299990697781	1.12332404306748
31		1.31376216345969	1.12360249420665
32		1.31447718492731	1.12386360633979
33		1.31514922711472	1.12410894908595
34		1.31578204955417	1.12433990820437
35		1.31637899041374	1.12455771341596
36		1.31694301673026	1.12476345714241
37		1.31747677722124	1.12495811410692
38		1.31798264476916	1.12514255710183
39		1.31846274997541	1.12531756944812

Table 6.2: n^{th} root approximations to τ and ρ.

n	$\frac{\lvert P(\tau_n)\rvert^{\frac{1}{2^{n-2}}}}{\lvert P(\tau_{n-1})\rvert^{\frac{1}{2^{n-3}}}}$	$\frac{P(\rho_n)^{\frac{1}{F_{n-1}}}}{P(\rho_{n-1})^{\frac{1}{F_{n-2}}}}$	τ approximation
12	1.265325659	1.33924654727029	1.13285914893344
13	1.026003498	1.33483022958181	1.13126229129270
14	1.115177325	1.33693642638980	1.13202422951668
15	1.233027320	1.33726287123704	1.13214226270383
16	1.049104383	1.33647141663268	1.13185606667525
17	1.120291308	1.33684733797345	1.13199201480700
18	1.209129616	1.33692083992409	1.13201859348026
19	1.067147962	1.33676482459047	1.13196217655715
20	1.123648513	1.33683881628769	1.13198893326862
21	1.191313146	1.33685345437077	1.13199422655883
22		1.33682247027204	1.13198302233215
23		1.33683722479937	1.13198835776711
24		1.33684029202542	1.13198946691295
25		1.33683394604571	1.13198717212873
26		1.33683697902493	1.13198826889205
27		1.33683762000829	1.13198850067949
28		1.33683630073850	1.13198802361530
29		1.33683693377726	1.13198825252994
30		1.33683706976437	1.13198830170455
31		1.33683679142042	1.13198820105195
32		1.33683692544958	1.13198824951855
33		1.33683695453562	1.13198826003642
34		1.33683689516267	1.13198823856643
35		1.33683692384246	1.13198824893739
36		1.33683693011759	1.13198825120656
37		1.33683691733460	1.13198824658407
38		1.33683692352642	1.13198824882311
39		1.33683692488976	1.13198824931611

Table 6.3: Ratio approximations to τ and ρ.

level for each subtree and find their products. When calculating product matrices, the absolute value was taken at each step, which means we are actually using the recurrence $t_n = |\pm t_{n-1} + t_{n-2}|$. This is equivalent to Equation (1.8), $\widetilde{f}_n = |\widetilde{f}_{n-1} \pm \widetilde{f}_{n-2}|$, which produces the positive tree R as required. We take the root of order F_{n-1} for all 13 products of nodes in ρ_n, and then take the ratio of this quantity for consecutive values of n. It is interesting that taking ratios for $n = 39$ we obtain one of two values, namely $\alpha = 1.01630979\ldots$ and $\beta = 1.02652252\ldots$. If the initial values form a left edge (this occurs 5 times), we obtain the former value and if they form a right edge (this occurs 8 times) we obtain the latter. We can combine this information to write

$$\frac{P(\rho_{39})^{\frac{1}{F_{38}}}}{P(\rho_{38})^{\frac{1}{F_{37}}}} = (1.01630979\ldots)^5 (1.02652252\ldots)^8, \tag{6.25}$$

and this value tends to ρ. But can these new constants tell us anything about the value of ρ? We can observe that $\beta = \alpha^{\phi}$. We can therefore rewrite Equation (6.25) (in the limit) as

$$\rho = \alpha^5 \alpha^{8\phi} = \alpha^{5+8\phi} = \alpha^{\phi^6}, \tag{6.26}$$

and so

$$\alpha = \rho^{\frac{1}{\phi^6}}, \quad \beta = \rho^{\frac{1}{\phi^5}}.$$

This investigation gives us no new insight about ρ, but does verify Proposition 1.1, which says that initial values are unimportant when calculating the a.s. growth rate of a random Fibonacci sequence, this time for sequences in R. Each of our 13 subtrees has different initial values, and each showed the same growth behaviour. Right-edge initial values (in the limit) have ϕ times as many descendants than left-edge initial values, which is demonstrated in the relation $\beta = \alpha^{\phi}$. Also, had we taken the root of order F_{n-7} instead of F_{n-1}, we would have obtained ρ instead of α as a growth rate.

Table 6.4 summarizes the different growth rates we have seen so far.

Tree	g.r. of sums of rows	g.r. of expected value	a.s. growth rate
R	$\alpha = 2.205569431\ldots$	$\alpha/\phi = 1.363116873\ldots$	$\rho = 1.33683692\ldots$
T	$2(\alpha - 1) = 2.411138862\ldots$	$\alpha - 1 = 1.205569431\ldots$	$\tau = 1.13198824\ldots$

Table 6.4: Various growth rates in R and T.

In the previous section we considered the products L_n and R_n, of left and right nodes of ρ_n, respectively. Using the ratio formula, like that given in Equations (6.24), computation reveals that the growth rates of the geometric means of L_n and R_n are given by

$$\lim_{n\to\infty} \frac{(L_n)^{\frac{1}{F_{n-3}}}}{(L_{n-1})^{\frac{1}{F_{n-4}}}} = \lim_{n\to\infty} \frac{(R_n)^{\frac{1}{F_{n-2}}}}{(R_{n-1})^{\frac{1}{F_{n-3}}}} = \rho. \tag{6.27}$$

Recall from Proposition 5.11 that $|\rho_n^-| = F_{n-3}$ and $|\rho_n^+| = F_{n-2}$. Further it is interesting to note that if we split up the product A_n, using the equivalent ratio form of the growth rate given in Definition 2.4, we have

$$\begin{aligned}
\lim_{n\to\infty} (A_n)^{\frac{1}{nF_{n-1}}} &= \lim_{n\to\infty} \left((L_n)^{\frac{1}{nF_{n-1}}} \cdot (R_n)^{\frac{1}{nF_{n-1}}} \right) \\
&= \lim_{n\to\infty} (L_n)^{\frac{1}{nF_{n-3}} \cdot \frac{F_{n-3}}{F_{n-1}}} \cdot \lim_{n\to\infty} (R_n)^{\frac{1}{nF_{n-2}} \cdot \frac{F_{n-2}}{F_{n-1}}}.
\end{aligned}$$

We have seen that $\lim_{n\to\infty} f(n)^{g(n)} = \lim_{n\to\infty} f(n)^{\lim_{n\to\infty} g(n)}$, where both limits exist. Using the growth rate from Equation (6.27) we can write

$$\begin{aligned}
\lim_{n\to\infty} (A_n)^{\frac{1}{nF_{n-1}}} &= \lim_{n\to\infty} (L_n)^{\frac{1}{nF_{n-3}} \cdot \lim_{n\to\infty} \frac{F_{n-3}}{F_{n-1}}} \cdot \lim_{n\to\infty} (R_n)^{\frac{1}{nF_{n-2}} \cdot \lim_{n\to\infty} \frac{F_{n-2}}{F_{n-1}}} \\
&= \rho^{\frac{1}{\phi^2}} \cdot \rho^{\frac{1}{\phi}} = \rho^{\frac{1+\phi}{\phi^2}} = \rho,
\end{aligned}$$

as expected.

We can consider a few other types of growth rates concerning the tree R, and also its variation R_1. We can approximate the a.s. growth rate of sequences in R by taking an average over all possible growth rates (i.e., n^{th} roots of dominant eigenvalues $|\lambda|$) of sequences as was done in Equation (3.15), or approximate the growth rate by the n^{th} root of the trace, like Equation (3.17). In R we will take the limit of the average of F_{n+1} growth rates or n^{th} roots of traces derived from product matrices of length n. We suspect that the a.s. growth rate would converge slowly to ρ, as was the case for sequences in T and their a.s. growth rate τ. First, let us consider the arithmetic mean of n^{th} roots of traces in R, i.e.,

$$\frac{\left(|u_{n,1}|^{\frac{1}{n}} + |u_{n,2}|^{\frac{1}{n}} + \cdots + |u_{n,i}|^{\frac{1}{n}} + \cdots + |u_{n,F_{n+1}|}^{\frac{1}{n}}\right)}{F_{n+1}}, \tag{6.28}$$

where $u_{n,i}$ is the trace of the i^{th} product matrix $P_{n,i}$. For $n = 7$ and 8, the above average is approximately equal to 1.118428 and 1.151009 respectively. We can similarly take the average of growth rates giving

$$\frac{(|\lambda_{n,1}|^{\frac{1}{n}} + |\lambda_{n,2}|^{\frac{1}{n}} + \cdots + |\lambda_{n,i}|^{\frac{1}{n}} + \cdots + |\lambda_{n,F_{n+1}}|^{\frac{1}{n}})}{F_{n+1}}.$$

For $7 \leq n \leq 13$ we obtain the values 1.157415, 1.142256, 1.123831, 1.114312, 1.124051, 1.080061, 1.119459. Strangely, these values appear to be hovering around Viswanath's constant, although we do not have enough data to draw any conclusions. We can repeat the calculation in Equation (6.28) for sequences in the tree R_1, i.e., those which do not contain $(+-+)$ or $(-+-)$. For $5 \leq n \leq 9$ we obtain the values 1.230506, 1.264744, 1.27086, 1.276508, 1.106244.

Similarly, we can modify Equation (5.31) for the growth rate of the expected value of the trace $|u|$, for both R and R_1. If we first consider the sums of trace values (in absolute value), we have the following sequences for R and R_1 respectively, for $n \geq 1$:

$$1, 4, 4, 11, 18, 42, 71, 154, 274, 533, 956, 1906, 3541, 6936, \ldots,$$

$$1, 4, 4, 15, 28, 66, 141, 316, 660, \ldots.$$

Taking the n^{th} root of these terms gives an approximation to the growth rate of the sum, and because we have F_{n+1} product matrices for a given n, dividing by ϕ (as would be done in the limit using the ratio definition) gives us an approximation to the growth rate of the expected value of the trace as follows:

$$\frac{(|u_{n,1}| + |u_{n,2}| + \cdots + |u_{n,i}| + \cdots + |u_{n,F_{n+1}}|)^{\frac{1}{n}}}{\phi}. \tag{6.29}$$

For the case of R, this expression gives us the following values for $6 \leq n \leq 13$: 1.152269, 1.136203, 1.159985, 1.153117, 1.157937, 1.153355, 1.159727, 1.158836. Similarly for R_1, the expression gives the following values for $6 \leq n \leq 9$: 1.242423, 1.253267, 1.269033, 1.271437. We could also repeat this calculation for eigenvalues $\lambda_{n,i}$ as was done in Equation (5.32). When comparing the average value of the growth rate of $|u|$ to the growth rate of the average value, (see Equations (6.28) and (6.29)) we obtain roughly 1.15 for R in both cases and 1.26 for R_1. This differs from the case of T, where we obtained differing values, namely $\tau = 1.13198824\ldots$ (see Equation (3.17)) and $\alpha - 1 = 1.20556943\ldots$ (see Equation (5.31)), in accordance with the generalized mean inequality given in Equation (5.30). Further, we might expect that the calculations of growth rates in R and R_1 would give the same values because by Proposition 5.6 these trees are comprised of the same set of sequences in absolute value. Our calculations here are for small values of n, and need to be further investigated before drawing any conclusions.

As mentioned in Section 6.1, Rittaud [64] constructs a matrix tree and subsequently a trace tree using the matrices $A, B' = \left(\begin{smallmatrix} 0 & 1 \\ \pm 1 & 1 \end{smallmatrix}\right)$. Using these traces, Equation (6.28) gives the values $1.354443, 1.35444$ for $n = 6, 7$. This value is close to $\rho = 1.33683692\ldots$, and would nicely parallel the approximation of τ given in Equation (3.17) (see Table 6.4), but again further investigation is required. The sequence of the sums of trace values (in absolute value) for $n \geq 1$ is given by

$$1, 4, 8, 19, 40, 90, 197, 436, 960, 2119, 4672, \ldots.$$

Rittaud gives the recurrence

$$H_n = 2H_{n-1} + H_{n-3} + 2(-1)^n$$

for $n \geq 3$, where $H_0 = 2, H_1 = 1, H_2 = 4$. Note the similarity to the recurrence given in Lemma 5.1. For $n = 6, 12, 18$, Equation (6.29) for the growth rate of the expected value of the trace gives $1.308336, 1.334862, 1.344085$. This value is close to $\alpha/\phi = 1.363116873\ldots$, and would nicely parallel the approximation of $\alpha - 1$ given in Equation (5.31) (again see Table 6.4). Finding a recurrence for the sums of traces or a characterization of the occurrence of trace values in R (as in the above examples) or T, could lead to a deterministic growth rate formula. This would be useful in those cases where the values are not tending to algebraic numbers, i.e., it could lead to an exact expression of τ or ρ.

6.3 A Second Argument for the Link Between τ and ρ.

Theorem 6.3 gives Rittaud's link between τ and ρ, i.e., the almost sure growth rates of random Fibonacci sequences in the trees T and R respectively. We can explain this relationship between τ and ρ using a different method. We will use our new derivation of τ from geometric means found in Kalmár-Nagy [44], as well as the results in Chapter 5 on writing rows of the tree T in terms of rows of the tree R.

We know from Equation (5.10) that the sum of terms in row n of T can be written as

$$S(\tau_n) = \sum_{k=0}^{\frac{n-3}{3}} t(n,k) S(\rho_{n-3k}), \tag{6.30}$$

where we have $n \equiv 0 \pmod 3$ and

$$\begin{aligned}
t(n,k) &= \sum_{j=0}^{k} A(n-3k-1, k-j) t(j), \\
A(i,m) &= \frac{2i}{3m+i}\binom{3m+i}{m}, \\
t(k) &= \frac{1}{2} \sum_{j_1+2j_2+\cdots+kj_k=k} \binom{n}{j_1, j_2, \ldots, j_k} A(0)^{j_1} A(1)^{j_2} \cdots A(k-1)^{j_k}, \\
A(m) &= \frac{4}{2m+1}\binom{3m}{m}.
\end{aligned}$$

If we write the entries in row n of the tree T as a product rather than a sum, Equation (6.30) can be rewritten as

$$P(\tau_n) = \prod_{k=0}^{\frac{n-3}{3}} P(\rho_{n-3k})^{t(n,k)}.$$

Our goal is to introduce the constants τ and ρ into this equation.

We can start by taking the root of order 2^{n-2} of both sides to get

$$P(\tau_n)^{\frac{1}{2^{n-2}}} = \prod_{k=0}^{\frac{n-3}{3}} P(\rho_{n-3k})^{\frac{t(n,k)}{2^{n-2}}}.$$

This gives us the geometric mean of the entries in τ_n on the left. We similarly want the geometric mean of the entries in ρ_{n-3k} on the right-hand side and so must take the $(F_{n-3k-1})^{\text{th}}$ root to get

$$P(\tau_n)^{\frac{1}{2^{n-2}}} = \prod_{k=0}^{\frac{n-3}{3}} P(\rho_{n-3k})^{\frac{1}{F_{n-3k-1}} \frac{t(n,k)F_{n-3k-1}}{2^{n-2}}}.$$

Now, in order for the left-hand side to approach the growth rate τ in the limit, we must take the n^{th} root of both sides, which gives

$$P(\tau_n)^{\frac{1}{n\cdot 2^{n-2}}} = \prod_{k=0}^{\frac{n-3}{3}} P(\rho_{n-3k})^{\frac{1}{F_{n-3k-1}} \frac{t(n,k)F_{n-3k-1}}{n\cdot 2^{n-2}}}.$$

Similarly, we need to take the $(n-3k)^{\text{th}}$ root of the right-hand side, which gives

$$P(\tau_n)^{\frac{1}{n\cdot 2^{n-2}}} = \prod_{k=0}^{\frac{n-3}{3}} P(\rho_{n-3k})^{\frac{1}{F_{n-3k-1}(n-3k)} \frac{t(n,k)F_{n-3k-1}(n-3k)}{n\cdot 2^{n-2}}} \tag{6.31}$$

Now we can take the limit as n goes to infinity of both sides. The left-hand side goes to τ, and $P(\rho_{n-3k})^{\frac{1}{F_{n-3k-1}(n-3k)}}$ goes to ρ for k small enough. It will soon be conjectured that $\sum_{k=0}^{\frac{n-3}{3}} \frac{t(n,k)F_{n-3k-1}(n-3k)}{n\cdot 2^{n-2}}$ approaches a constant, implying these limits exist. Roughly speaking we have that

$$\begin{aligned}
\tau &= \prod_{k=0}^{\frac{n-3}{3}} \rho^{\lim_{n\to\infty} \frac{t(n,k)F_{n-3k-1}(n-3k)}{n\cdot 2^{n-2}}} \\
&= \rho^{\lim_{n\to\infty} \sum_{k=0}^{\frac{n-3}{3}} \frac{t(n,k)F_{n-3k-1}(n-3k)}{n\cdot 2^{n-2}}}.
\end{aligned} \tag{6.32}$$

To prove this, we may be able to use Proposition 5.12 for Cesàro means on logarithms of terms in Equation (6.31). It is true that

$$\lim_{n\to\infty} \left(\prod_{k=0}^{\frac{n-3}{3}} P(\rho_{n-3k})^{\frac{1}{F_{n-3k-1}(n-3k)}} \right)^{\frac{1}{n/3}} = \lim_{n\to\infty} P(\rho_n)^{\frac{1}{F_{n-1}\cdot n}} = \rho,$$

where we have chosen $k = 0$. The problem is that the term $\frac{t(n,k)F_{n-3k-1}(n-3k)}{n\cdot 2^{n-2}}$ is increasing as k increases, whereas the term in the above equation is decreasing.

We have the following conjecture, which if true, could give Rittaud's link, $\tau = \rho^{1-\frac{3}{2\phi^2}}$, using Equation (6.32).

Conjecture 6.2. *For* $3|n$ *we have*

$$\lim_{n\to\infty} \sum_{k=0}^{\frac{n-3}{3}} \frac{t(n,k)F_{n-3k-1}(n-3k)}{n\cdot 2^{n-2}} = 1 - \frac{3}{2\phi^2},$$

where $t(n,k)$ *is given by Equation 5.10 and Table 5.2.*

We can obtain similar expressions by repeating the above for Equation (5.10) with $n = 1$ or $2 \pmod 3$. Let us denote the sum in our conjecture by

$$\mu_n := \sum_{k=0}^{\frac{n-3}{3}} \frac{t(n,k)F_{n-3k-1}(n-3k)}{n \cdot 2^{n-2}}.$$

Table 6.5 gives some values for μ_n for $n \geq 15$ with $3|n$.

n	μ_n
15	0.47324218
18	0.46752929
21	0.46295275
24	0.45922851
27	0.45615220
30	0.45357654

Table 6.5: Approximating μ_n for Conjecture 6.2.

We want these values to tend to

$$1 - \frac{3}{2\phi^2} = 0.4270509834\ldots,$$

and they appear to be converging as desired.

Chapter 7

Trace Reduction

7.1 A Trace Recursion

To better understand how the growth of a random Fibonacci sequence changes as we increase n, we can study the change in the trace of its product matrix. For this purpose we have the following recursive system for the trace. As before, A and B are the matrices first defined in Section 1.4.

Theorem 7.1. *Let P_n be a product matrix of length n. The following expressions therefore correspond to product matrices of length $n+2$:*

$$\left.\begin{aligned} \mathrm{tr}(P_nAA) &= \mathrm{tr}(P_nA) + \mathrm{tr}(P_n), \\ \mathrm{tr}(P_nBB) &= -\mathrm{tr}(P_nB) + \mathrm{tr}(P_n). \end{aligned}\right\} \tag{7.1}$$

The following expressions correspond to product matrices of length $n+3$:

$$\left.\begin{aligned} \mathrm{tr}(P_nAAA) &= \mathrm{tr}(P_nAA) + \mathrm{tr}(P_nA), \\ \mathrm{tr}(P_nAAB) &= \mathrm{tr}(P_nAB) + \mathrm{tr}(P_nB), \\ \mathrm{tr}(P_nABA) &= \mathrm{tr}(P_nAA) - \mathrm{tr}(P_nB) - 2\mathrm{tr}(P_n), \\ \mathrm{tr}(P_nABB) &= -\mathrm{tr}(P_nAB) + \mathrm{tr}(P_nA), \\ \mathrm{tr}(P_nBAA) &= \mathrm{tr}(P_nBA) + \mathrm{tr}(P_nB), \\ \mathrm{tr}(P_nBAB) &= -\mathrm{tr}(P_nBB) - \mathrm{tr}(P_nA) + 2\mathrm{tr}(P_n), \\ \mathrm{tr}(P_nBBA) &= -\mathrm{tr}(P_nBA) + \mathrm{tr}(P_nA), \\ \mathrm{tr}(P_nBBB) &= -\mathrm{tr}(P_nBB) + \mathrm{tr}(P_nB). \end{aligned}\right\} \tag{7.2}$$

Proof: Start by letting $P_n = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ be a product matrix of length n. We simply need to compute all product matrices of length $n+1$, $n+2$ and $n+3$, and verify the trace relations. For length $n+1$ we have:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} = \begin{pmatrix} b & a \pm b \\ d & c \pm d \end{pmatrix},$$

and so

$$\operatorname{tr}(P_n A) = b + c + d,$$
$$\operatorname{tr}(P_n B) = b + c - d.$$

Note that $\operatorname{tr}(P_n) = a + d$.

For length $n+2$ we have the following cases for $(P_n AA)$ and $(P_n AB)$:

$$\begin{pmatrix} b & a+b \\ d & c+d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} = \begin{pmatrix} a+b & b \pm (a+b) \\ c+d & d \pm (c+d) \end{pmatrix},$$

and so

$$\operatorname{tr}(P_n AA) = a + b + d + (c+d) = a + b + c + 2d,$$
$$\operatorname{tr}(P_n AB) = a + b + d - (c+d) = a + b - c.$$

Similarly for $(P_n BA)$ and $(P_n BB)$ we have

$$\begin{pmatrix} b & a-b \\ d & c-d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} = \begin{pmatrix} a-b & b \pm (a-b) \\ c-d & d \pm (c-d) \end{pmatrix},$$

and so the traces are

$$\operatorname{tr}(P_n BA) = a - b + d + (c-d) = a - b + c,$$
$$\operatorname{tr}(P_n BB) = a - b + d - (c-d) = a - b - c + 2d.$$

In terms of previous traces we have the relations

$$\operatorname{tr}(P_n AA) = (b + c + d) + (a + d) = \operatorname{tr}(P_n +) + \operatorname{tr}(P_n),$$
$$\operatorname{tr}(P_n BB) = -(b + c - d) + (a + d) = -\operatorname{tr}(P_n -) + \operatorname{tr}(P_n).$$

For length $n+3$ we have the following cases:

$$P_n AAA,\ P_n AAB = \begin{pmatrix} a+b & b+(a+b) \\ c+d & d+(c+d) \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} = \begin{pmatrix} a+2b & a+b \pm (a+2b) \\ c+2d & c+d \pm (c+2d) \end{pmatrix},$$

$$P_nABA,\ P_nABB = \begin{pmatrix} a+b & b-(a+b) \\ c+d & d-(c+d) \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} = \begin{pmatrix} -a & a+b \pm (-a) \\ -c & c+d \pm (-c) \end{pmatrix},$$

$$P_nBAA,\ P_nBAB = \begin{pmatrix} a-b & b+(a-b) \\ c-d & d+(c-d) \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} = \begin{pmatrix} a & a-b \pm a \\ c & c-d \pm c \end{pmatrix},$$

$$\begin{aligned} P_nBBA,\ P_nBBB &= \begin{pmatrix} a-b & b-(a-b) \\ c-d & d-(c-d) \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} \\ &= \begin{pmatrix} -a+2b & a-b \pm (-a+2b) \\ -c+2d & c-d \pm (-c+2d) \end{pmatrix}. \end{aligned}$$

It follows that the possible traces are

$$\begin{aligned} \mathrm{tr}(P_nAAA) &= a+2b+c+d+(c+2d) = a+2b+2c+3d, \\ \mathrm{tr}(P_nAAB) &= a+2b+c+d-(c+2d) = a+2b-d, \\ \mathrm{tr}(P_nABA) &= -a+c+d+(-c) = -a+d, \\ \mathrm{tr}(P_nABB) &= -a+c+d-(-c) = -a+2c+d, \\ \mathrm{tr}(P_nBAA) &= a+c-d+c = a+2c-d, \\ \mathrm{tr}(P_nBAB) &= a+c-d-c = a-d, \\ \mathrm{tr}(P_nBBA) &= -a+2b+c-d+(-c+2d) = -a+2b+d, \\ \mathrm{tr}(P_nBBB) &= -a+2b+c-d-(-c+2d) = -a+2b+2c-3d. \end{aligned}$$

Writing these traces in terms of the previous traces gives:

$$\begin{aligned} \mathrm{tr}(P_nAAA) &= (a+b+c+2d)+(b+c+d) = \mathrm{tr}(P_nAA)+\mathrm{tr}(P_nA), \\ \mathrm{tr}(P_nAAB) &= (a+b-c)+(b+c-d) = \mathrm{tr}(P_nAB)+\mathrm{tr}(P_nB), \\ \mathrm{tr}(P_nABA) &= (a+b+c+2d)-(b+c-d)-2(a+d) \\ &= \mathrm{tr}(P_nAA)-\mathrm{tr}(P_nB)-2\mathrm{tr}(P_n), \\ \mathrm{tr}(P_nABB) &= -(a+b-c)+(b+c+d) = -\mathrm{tr}(P_nAB)+\mathrm{tr}(P_nA), \\ \mathrm{tr}(P_nBAA) &= (a-b+c)+(b+c-d) = \mathrm{tr}(P_nBA)+\mathrm{tr}(P_nB), \end{aligned}$$

$$\begin{aligned}\mathrm{tr}(P_nBAB) &= -(a-b-c+2d)-(b+c+d)+2(a+d)\\ &= -\mathrm{tr}(P_nBB)-\mathrm{tr}(P_nA)+2\mathrm{tr}(P_n),\\ \mathrm{tr}(P_nBBA) &= -(a-b+c)+(b+c+d) = -\mathrm{tr}(P_nBA)+\mathrm{tr}(P_nA),\\ \mathrm{tr}(P_nBBB) &= -(a-b-c+2d)+(b+c-d) = -\mathrm{tr}(P_nBB)+\mathrm{tr}(P_nB),\end{aligned}$$

completing the proof. ◇

We can go one step further and write the above eight relations so that each trace is in terms of $\mathrm{tr}(P_nAB), \mathrm{tr}(P_nBA), \mathrm{tr}(P_nA), \mathrm{tr}(P_nB)$ or $\mathrm{tr}(P_n)$. We do this by substituting the equations in System (7.1), giving the alternate relations:

$$\begin{aligned}\mathrm{tr}(P_nAAA) &= 2\mathrm{tr}(P_nA)+\mathrm{tr}(P_n),\\ \mathrm{tr}(P_nABA) &= \mathrm{tr}(P_nA)-\mathrm{tr}(P_nB)-\mathrm{tr}(P_n),\\ \mathrm{tr}(P_nBAB) &= \mathrm{tr}(P_nB)-\mathrm{tr}(P_nA)+\mathrm{tr}(P_n),\\ \mathrm{tr}(P_nBBB) &= 2\mathrm{tr}(P_nB)-\mathrm{tr}(P_n).\end{aligned}$$

In fact, the complete recursive system can be represented by the following six equations, rather than the eight given in System (7.2):

$$\begin{aligned}\mathrm{tr}(P_nAA) &= \mathrm{tr}(P_nA)+\mathrm{tr}(P_n),\\ \mathrm{tr}(P_nBB) &= -\mathrm{tr}(P_nB)+\mathrm{tr}(P_n),\\ \mathrm{tr}(P_nAAB) &= \mathrm{tr}(P_nAB)+\mathrm{tr}(P_nB),\\ \mathrm{tr}(P_nABA) &= \mathrm{tr}(P_nA)-\mathrm{tr}(P_nB)-\mathrm{tr}(P_n),\\ \mathrm{tr}(P_nBAB) &= \mathrm{tr}(P_nB)-\mathrm{tr}(P_nA)+\mathrm{tr}(P_n),\\ \mathrm{tr}(P_nBBA) &= -\mathrm{tr}(P_nBA)+\mathrm{tr}(P_nA).\end{aligned}$$

Notice that the relations for $\mathrm{tr}(P_nAAA), \mathrm{tr}(P_nBAA), \mathrm{tr}(P_nABB)$ and $\mathrm{tr}(P_nBBB)$ belong to the $\mathrm{tr}(P_nAA)$ and $\mathrm{tr}(P_nBB)$ cases.

Example 7.1. Let us calculate the trace of the matrix given by the cycle $(++-+-)$

using our recursive relations. Letting $P_2 = AA$, we have

$$\begin{aligned}
\operatorname{tr}(AABAB) = \operatorname{tr}(P_2BAB) &= \operatorname{tr}(P_2B) - \operatorname{tr}(P_2A) + \operatorname{tr}(P_2) \\
&= \operatorname{tr}(AAB) - \operatorname{tr}(AAA) + \operatorname{tr}(AA) \\
&= \operatorname{tr}(AB) + \operatorname{tr}(B) - (2\operatorname{tr}(A) + \operatorname{tr}(I)) + \operatorname{tr}(A) + \operatorname{tr}(I) \\
&= \operatorname{tr}(AB) - \operatorname{tr}(A) + \operatorname{tr}(B).
\end{aligned}$$

Using the facts that $\operatorname{tr}(AB) = \operatorname{tr}(A) = 1$ and $\operatorname{tr}(B) = -1$, we have that

$$\operatorname{tr}(AABAB) = 1 - 1 + (-1) = -1.$$

Notice that for any product matrix, the trace can be written as a linear combination of $\operatorname{tr}(AB), \operatorname{tr}(BA), \operatorname{tr}(A), \operatorname{tr}(B)$ and $\operatorname{tr}(I)$. Furthermore, we have that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ and $\operatorname{tr}(B) = -\operatorname{tr}(A)$, so that any trace can in fact be written as a linear combination of $\operatorname{tr}(AB), \operatorname{tr}(A)$ and $\operatorname{tr}(I)$, which have values $1, 1$ and 2 respectively. We can think of these as the initial values for our recursive system.

We are ultimately concerned with the growth rate of our sequences, which can be derived from the trace of the product matrix. One thing the recursive system allows us to do is examine the average trace value, for length-n product matrices.

Corollary 7.1. *The average trace value over all product matrices of length n, for $n \geq 0$, is given by*

$$\begin{cases} 0, & \textit{if } n \textit{ is odd;} \\ 2, & \textit{if } n \textit{ is even.} \end{cases}$$

Proof: We can easily prove this result using induction. Consider the initial case. For $n = 0$, we have by default $P_0 = I$, and so $\operatorname{tr}(P_0) = 2$. For $n = 1$, we have $P_1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & \pm 1 \end{smallmatrix}\right)$, which have traces ± 1, averaging to 0. Now consider a length n product matrix, P_n. The eight relations in System (7.2) tell us that the average trace of product matrices of the form $P_nM_1M_2M_3$, where $M_i \in \{A, B\}$ is

$$\begin{aligned}
&\frac{1}{8}(2\operatorname{tr}(P_n) + \operatorname{tr}(P_n) + \operatorname{tr}(P_nAB) + \operatorname{tr}(P_nB) + \operatorname{tr}(P_nA) - \operatorname{tr}(P_nB) \\
&- \operatorname{tr}(P_n) + \operatorname{tr}(P_nBA) + \operatorname{tr}(P_nB) - \operatorname{tr}(P_nAB) + \operatorname{tr}(P_nA) + \operatorname{tr}(P_nB) \\
&- \operatorname{tr}(P_nA) + \operatorname{tr}(P_n) - \operatorname{tr}(P_nBA) + \operatorname{tr}(P_nA) + 2\operatorname{tr}(P_nB) - \operatorname{tr}(P_n)) \\
&= \frac{1}{2}(\operatorname{tr}(P_nA) + \operatorname{tr}(P_nB)).
\end{aligned}$$

This tells us that the average trace for matrices of the form $P_nM_1M_2M_3$ is the same as the average trace of matrices of the form P_nM_1, for a given P_n. Averaging over all 2^n product matrices P_n of length n, we can conclude that the average trace of product matrices of length $n+3$ is equal to the average trace of product matrices of length $n+1$.

Now by induction, suppose n is even, and that the average trace of product matrices of odd length $n+1$ is 0. Then, the average trace of product matrices of length $n+3$ is also 0, by the above-mentioned result. Similarly, if n is odd and product matrices of even length $n+1$ have an average trace of 2, then so do products of length $n+3$. This completes the induction. $\diamond$

This result is easy to see for the n odd case. Recall that the proof of Theorem 4.6 says that when we negate the signs in an odd-length coefficient cycle, we also negate the trace. The set of all coefficient cycles must occur in pairs (all cycles and their negations) and so the sum of all traces, and hence the average, must be zero. Our result is not useful in approximating Viswanath's constant, or proving Conjecture 5.1 because the conjecture and Equation (3.17) require us to use the absolute value of the traces, and we cannot determine which terms in the average are positive and which are negative.

7.2 Some Trace Patterns

By looking at some patterns of $\pm$ signs in our coefficient cycles, we can find patterns in the traces of the corresponding product matrices. Many of these patterns will involve Fibonacci numbers. Recall that we have defined the initial values of the Fibonacci numbers to be $F_0 = 0, F_1 = 1, F_2 = 1$ and $F_3 = 2$. We can take the sequence backwards and define Fibonacci numbers F_n for negative values of n. This gives $F_{-1} = 1, F_{-2} = -1, F_{-3} = 2, F_{-4} = -3, F_{-5} = 5, F_{-6} = -8$ and so on. We see that

$$F_{-n} = (-1)^{n+1} F_n.$$

Lemma 7.1. *For* $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 0 & 1 \\ 1 & -1 \end{smallmatrix}\right)$ *and* $C = AB = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, *we have the* *following product matrices for $n \geq 0$:*

$$A^n = \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix}, \qquad B^n = \begin{cases} \begin{pmatrix} F_{n-1} & -F_n \\ -F_n & F_{n+1} \end{pmatrix}, & n \text{ even}; \\ \begin{pmatrix} -F_{n-1} & F_n \\ F_n & -F_{n+1} \end{pmatrix}, & n \text{ odd}, \end{cases}$$

$$C^n = \begin{cases} \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, & n \equiv 1 \pmod 6; \\ \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, & n \equiv 2 \pmod 6; \\ \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, & n \equiv 3 \pmod 6; \end{cases} \qquad \begin{cases} \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}, & n \equiv 4 \pmod 6; \\ \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}, & n \equiv 5 \pmod 6; \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & n \equiv 0 \pmod 6. \end{cases}$$

Proof: We can prove the form of A^n for $n \geq 1$ using a simple induction. For $n = 1$, we have $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} F_0 & F1 \\ F_1 & F_2 \end{smallmatrix}\right)$. Now assume that $A^n = \left(\begin{smallmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{smallmatrix}\right)$. We can write

$$A^{n+1} = \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} F_n & F_{n-1} + F_n \\ F_{n+1} & F_n + F_{n+1} \end{pmatrix} = \begin{pmatrix} F_n & F_{n+1} \\ F_{n+1} & F_{n+2} \end{pmatrix},$$

completing the induction.

Similarly we have $B = \left(\begin{smallmatrix} 0 & 1 \\ 1 & -1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -F_0 & F_1 \\ F_1 & -F_2 \end{smallmatrix}\right)$ and $B^2 = \left(\begin{smallmatrix} 1 & -1 \\ -1 & 2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} F_1 & -F_2 \\ -F_2 & F_3 \end{smallmatrix}\right)$. Now assume that $B^n = \left(\begin{smallmatrix} -F_{n-1} & F_n \\ F_n & -F_{n+1} \end{smallmatrix}\right)$ for n odd. We can write

$$\begin{aligned} B^{n+1} &= \begin{pmatrix} -F_{n-1} & F_n \\ F_n & -F_{n+1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} F_n & -F_{n-1} - F_n \\ -F_{n+1} & F_n + F_{n+1} \end{pmatrix} = \begin{pmatrix} F_n & -F_{n+1} \\ -F_{n+1} & F_{n+2} \end{pmatrix}, \end{aligned}$$

which is the required form for n even. Further, multiplying by B again gives

$$\begin{aligned} B^{n+2} &= \begin{pmatrix} F_n & -F_{n+1} \\ -F_{n+1} & F_{n+2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} -F_{n+1} & F_n + F_{n+1} \\ F_{n+2} & -F_{n+1} - F_{n+2} \end{pmatrix} = \begin{pmatrix} -F_{n+1} & F_{n+2} \\ F_{n+2} & -F_{n+3} \end{pmatrix}, \end{aligned}$$

which is the required form for n odd, completing the induction.

Lastly we consider the matrix C. Taking successive powers gives the matrices stated in the Lemma. We reach $C^6 = I$ and so the matrices will continue to rotate through this set. For each of the three product matrices, letting $n = 0$ gives us the identity matrix. Therefore our lemma is satisfied for $n \geq 0$, using the convention that $M^0 = I$ for any matrix M. ◇

Let us now consider product matrices of coefficient cycles having the form $(++\cdots+--\cdots-)$, where there are n $+$ signs and i $-$ signs, i.e., $((+)^n(-)^i)$.

Theorem 7.2. *The following product matrices have the given traces, for $n, i \geq 0$:*

1. $\mathrm{tr}(A^nB) = F_{n-2}$,
2. $\mathrm{tr}(A^nB^2) = 3F_{n-1}$,
3. $\mathrm{tr}(A^nB^3) = F_n - 4F_{n-1}$,
4. $\mathrm{tr}(A^nB^n) = -F_n^2 + 2(-1)^n$,
5. $\mathrm{tr}(A^nB^i) = (2F_nF_i - F_{n-1}F_{i-1} - F_{n+1}F_{i+1})(-1)^{i+1}$.

Proof: From Lemma 7.1 we can write

$$A^nB = \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} F_n & F_{n-1} - F_n \\ F_{n+1} & F_n - F_{n+1} \end{pmatrix} = \begin{pmatrix} F_n & -F_{n-2} \\ F_{n+1} & -F_{n-1} \end{pmatrix}.$$

We have $\mathrm{tr}(A^nB) = F_n - F_{n-1} = F_{n-2}$, proving Statement 1. Let us now jump to the general case given in Statement 5, which we could have used to prove Statement 1. Combining the even and odd cases for i in Lemma 7.1, we can write

$$\begin{aligned} A^nB^i &= \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix} \begin{pmatrix} -F_{i-1} & F_i \\ F_i & -F_{i+1} \end{pmatrix} (-1)^{i+1} \\ &= \begin{pmatrix} -F_{n-1}F_{i-1} + F_nF_i & F_{n-1}F_i - F_nF_{i+1} \\ -F_nF_{i-1} + F_{n+1}F_i & F_nF_i - F_{n+1}F_{i+1} \end{pmatrix} (-1)^{i+1}, \end{aligned} \tag{7.3}$$

which has

$$\mathrm{tr}(A^nB^i) = (2F_nF_i - F_{n-1}F_{i-1} - F_{n+1}F_{i+1})(-1)^{i+1}. \tag{7.4}$$

We can now use this general trace to prove Statements 2-4. We have that for $i = 2$,

$$\mathrm{tr}(A^nB^2) = (2F_nF_2 - F_{n-1}F_1 - F_{n+1}F_3)(-1) = -2F_n + F_{n-1} + 2F_{n+1} = 3F_{n-1},$$

which proves Statement 2. Similarly, for $i = 3$, Equation (7.4) gives

$$\mathrm{tr}(A^n B^3) = 2F_n F_3 - F_{n-1}F_2 - F_{n+1}F_4 = 4F_n - F_{n-1} - 3F_{n+1} = F_n - 4F_{n-1},$$

which proves Statement 3. Lastly, for $i = n$, Equation (7.4) gives

$$\mathrm{tr}(A^n B^n) = (2F_n^2 - F_{n-1}^2 - F_{n+1}^2)(-1)^{n+1} = -F_n^2 + 2(-1)^n.$$

The last step is straightforward to prove using properties of Fibonacci numbers. $\diamond$

We know from Equation (3.14) that we can use the trace of a product matrix to find the growth rate of the corresponding periodic coefficient sequence. It can be shown that as $n \to \infty$ in each of the cases in Theorem 7.2, the growth rate approaches ϕ. This makes sense, as we can think of the coefficient cycle approaching $(+++\cdots)$, which corresponds to the regular Fibonacci sequence.

Now we will consider product matrices of coefficient cycles that contain $((+-)^n)$.

Theorem 7.3. *The following product matrices have the given traces and growth types, for $n \geq 0$:*

1. $|\mathrm{tr}(C^n)| = \begin{cases} 1 \text{ for } n \equiv 1, 2 \pmod 3; \\ 2 \text{ for } n \equiv 0 \pmod 3, \end{cases}$ *growth* $= \begin{cases} E \iff n \text{ odd}; \\ B \iff n \text{ even}, \end{cases}$

2. $|\mathrm{tr}(AC^n)| = \begin{cases} 0 \text{ for } n \equiv 1 \pmod 3; \\ 1 \text{ for } n \equiv 0, 2 \pmod 3, \end{cases}$ *growth* $= \begin{cases} E \iff n \equiv 3, 5 \pmod 6; \\ B \Leftrightarrow n \equiv 0, 1, 2 \pmod 6, \end{cases}$

3. $|\mathrm{tr}(AC^nB)| = \begin{cases} 3 \text{ for } n \equiv 1 \pmod 3; \\ 2 \text{ for } n \equiv 2 \pmod 3; \\ 1 \text{ for } n \equiv 0 \pmod 3, \end{cases}$ *growth* $= \begin{cases} E \Leftrightarrow n \equiv 1, 3, 4, 5 \pmod 6; \\ L \iff n \equiv 2 \pmod 6; \\ B \iff n \equiv 0 \pmod 6, \end{cases}$

4. $|\mathrm{tr}(A^nC^n)| = \begin{cases} F_{n-1} \text{ for } n \equiv 1 \pmod 3; \\ F_{n+1} \text{ for } n \equiv 2 \pmod 3; \\ L_{n+1} \text{ for } n \equiv 0 \pmod 3, \end{cases}$ *growth* $= \begin{cases} E \iff n = 3, n \geq 5; \\ L \iff n = 2, 4; \\ B \iff n = 0, 1. \end{cases}$

Proof: We will start by proving the trace results. The proof of Statement 1 follows directly from Lemma 7.1. For $n \equiv 1,2,3,4,5,0 \pmod 6$, we have the respective traces $1,-1,-2,-1,1,2$. For Statement 2, we pre-multiply each matrix in Lemma 7.1 by A, to obtain the following six matrices:

$$\begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & -2 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ -2 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ -1 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

which have respective traces $0,-1,-1,0,1,1$. Similarly for AC^nB we obtain the matrices

$$\begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & -3 \end{pmatrix}, \begin{pmatrix} 1 & -2 \\ 2 & -3 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix},$$

which have traces $3,2,-1,-3,-2,1$. Lastly, for A^nC^n, we obtain the matrices

$$\begin{pmatrix} F_{n+1} & -F_{n-1} \\ F_{n+2} & -F_n \end{pmatrix}, \begin{pmatrix} F_n & -F_{n+1} \\ F_{n+1} & -F_{n+2} \end{pmatrix}, \begin{pmatrix} -F_{n-1} & -F_n \\ -F_n & -F_{n+1} \end{pmatrix},$$

$$\begin{pmatrix} -F_{n+1} & F_{n-1} \\ -F_{n+2} & F_n \end{pmatrix}, \begin{pmatrix} -F_n & F_{n+1} \\ -F_{n+1} & F_{n+2} \end{pmatrix}, \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix},$$

which have traces F_{n-1}, $-F_{n+1}$, $-F_{n-1} - F_{n+1} = -L_{n+1}$, $-F_{n-1}$, F_{n+1}, $F_{n-1} + F_{n+1} = L_{n+1}$, respectively. Recall that L_n denotes the Lucas numbers, namely, $2,1,3,4,7,11,18,\ldots$ for $n \geq 1$.

Now we can use the above information along with the connection between trace and growth type found in Theorem 3.3 to prove the remainder of the theorem. For matrices of the form C^n, we need only classify the traces 1 and 2. From Theorem 3.3 we have that for n odd, traces 1 and 2 imply exponential growth. For n even, trace 1, which occurs for $n \equiv 1,2 \pmod 3$, implies bounded growth, and so this occurs for $n \equiv 2,4 \pmod 6$. Similarly for n even, trace 2, which occurs for $n \equiv 0 \pmod 3$, implies linear growth for product matrices $P_n \neq I$. We have seen above that $C^n = I$ for $n = 0 \pmod 6$ and so for these values of n, growth is actually bounded.

For matrices of the form AC^n, we need to classify traces 0 and 1. For n odd, traces 0 and 1 imply bounded and exponential growth respectively, and for n even trace 1 implies bounded growth, where trace 0 does not occur. Combining this information with the fact that trace 0 occurs for $n \equiv 1 \pmod 3$ and trace 1 occurs otherwise we can deduce that exponential growth occurs for traces 3 and 5, and bounded growth occurs for traces 0, 1, and 2.

For matrices of the form AC^nB, we have traces 1, 2 and 3. When n is odd, we have by Theorem 3.3 that all of these traces imply exponential growth. When n is even, traces 1, 2 and 3 occur for bounded, linear and exponential growth, respectively. These traces occur for values of $n \equiv 0, 2, 1 \pmod 3$ respectively, which correspond to values $n \equiv 0, 2, 4 \pmod 6$.

Lastly we consider the matrices of the form A^nC^n. For $n \geq 0$ we have the traces $2, 0, 2, 4, 2, 8, 18, 8, 34, \ldots$. By Theorem 3.3 we have linear growth when n is even and the trace is 2, which occurs here for $n = 0, 2$, and 4. But recall that this is only true for $P_n \neq I$. In the case of $n = 0$ we have $A^0C^0 = I$ and so growth is bounded. Bounded growth also occurs for $n = 1$, and all other values of n correspond to exponential growth. $\diamond$

It is easy to find many more interesting and more complex examples of this nature. The only case in Theorem 7.3 where the trace increases exponentially as $n \to \infty$ is that of the product matrix A^nC^n. The reason is that it is the only case containing A^n. The other cases have bounded (in fact, periodic) trace values, hence bounded eigenvalues and bounded growth rates. The growth rate $|\lambda_1|^{\frac{1}{n}}$ for a given value of $n \pmod 6$ must actually be decreasing to 1 as n increases. The examples in Theorem 7.2 which contain A^n have increasing traces and corresponding growth rates that tend to ϕ. For A^nC^n, computation shows that the growth rate tends to $\phi^{\frac{1}{3}}$ as $n \to \infty$. This is explainable by noticing that in the product matrix A^n has length n and C^n has length $2n$. Therefore only a third of the product matrix is contributing to the exponential growth towards ϕ.

We will consider another type of trace reduction, where we look at the effect on the trace of multiplying a general matrix $P_n = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ by a particular product matrix. The product matrices we will consider are periodic with period 1 or 2. We will also introduce the matrix $D = BA$. Note that $D^n = BC^{n-1}A$, and produces the following set of matrices for $n \geq 1$:

$$\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Table 7.1 is a simple consequence of multiplying a general product matrix P_n by D^n or one of the powers given in Lemma 7.1. We could continue the table for product matrices with periods greater than 2. Note that a table of this sort might be the best we can do to track the changes of the trace of a product matrix. Given traces $\text{tr}(M_1)$ and $\text{tr}(M_2)$ of matrices M_1 and M_2, there is no known way to compute the $\text{tr}(M_1M_2)$. Note that in Table 7.1, by $n(6)$ we mean $n \pmod 6$.

Example 7.2. We can use Table 7.1 to break down the trace of $P_nAAAABAB$ in steps, where $P_n = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. First rewrite this matrix as $\hat{P}_nABAB$, where $\hat{P}_n = P_nAAA$. From the table we have that

$$\hat{P}_n = \begin{pmatrix} aF_2 + bF_3 & aF_3 + bF_4 \\ cF_2 + dF_3 & cF_3 + dF_4 \end{pmatrix} = \begin{pmatrix} a + 2b & 2a + 3b \\ c + 2d & 2c + 3d \end{pmatrix}.$$

Now we can use the fact that $\text{tr}(\hat{P}_nC^2) = \hat{b} - \hat{c} - \hat{d}$, where $\hat{b}, \hat{c}, \hat{d}$ are the entries in $\hat{P}_n$, to give us

$$\text{tr}(P_nA^3C^2) = (2a + 3b) - (c + 2d) - (2c + 3d) = 2a + 3b - 3c - 5d.$$

7.3 Fibonacci Blocks Type I

The growth type and rate of a random Fibonacci sequence are dependent on the trace of the associated product matrix. (Recall Theorem 3.3 gives the connection between growth type and trace, Definition 2.3 gives us the growth rate in terms of the eigenvalues, and Equation (3.14) gives the dependence of the dominant eigenvalue on the trace.) It would therefore be ideal to have an equation for determining the trace of a product matrix given its particular periodic coefficient sequence. In the previous section we have considered the trace for some specific types of coefficient cycles.

Our product matrices are comprised of the matrices A and B, but it is useful to think of our product as being comprised of the *blocks* A^iB^j. We have from Equation

$n(6)$	form	product matrix	trace
any	P_nA^n	$\begin{pmatrix} aF_{n-1}+bF_n & aF_n+bF_{n+1} \\ cF_{n-1}+dF_n & cF_n+dF_{n+1} \end{pmatrix}$	$aF_{n-1}+bF_n+cF_n+dF_{n+1}$
0,2,4	P_nB^n	$\begin{pmatrix} aF_{n-1}-bF_n & -aF_n+bF_{n+1} \\ cF_{n-1}-dF_n & -cF_n+dF_{n+1} \end{pmatrix}$	$aF_{n-1}-bF_n-cF_n+dF_{n+1}$
1,3,5	P_nB^n	$\begin{pmatrix} -aF_{n-1}+bF_n & aF_n-bF_{n+1} \\ -cF_{n-1}+dF_n & cF_n-dF_{n+1} \end{pmatrix}$	$-aF_{n-1}+bF_n+cF_n-dF_{n+1}$
0	P_nC^n	$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$	$a+d$
1	P_nC^n	$\begin{pmatrix} a+b & -a \\ c+d & -c \end{pmatrix}$	$a+b-c$
2	P_nC^n	$\begin{pmatrix} b & -a-b \\ d & -c-d \end{pmatrix}$	$b-c-d$
3	P_nC^n	$\begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$	$-a-d$
4	P_nC^n	$\begin{pmatrix} -a-b & a \\ -c-d & c \end{pmatrix}$	$-a-b+c$
5	P_nC^n	$\begin{pmatrix} -b & a+b \\ -d & c+d \end{pmatrix}$	$-b+c+d$
0	P_nD^n	$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$	$a+d$
1	P_nD^n	$\begin{pmatrix} a-b & a \\ c-d & c \end{pmatrix}$	$a-b+c$
2	P_nD^n	$\begin{pmatrix} -b & a-b \\ -d & c-d \end{pmatrix}$	$-b+c-d$
3	P_nD^n	$\begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$	$-a-d$
4	P_nD^n	$\begin{pmatrix} -a+b & -a \\ -c+d & -c \end{pmatrix}$	$-a+b-c$
5	P_nD^n	$\begin{pmatrix} b & -a+b \\ d & -c+d \end{pmatrix}$	$b-c+d$

Table 7.1: Change in trace upon matrix multiplication.

(7.3) that

$$A^iB^j = \begin{pmatrix} -F_{i-1}F_{j-1} + F_iF_j & F_{i-1}F_j - F_iF_{j+1} \\ -F_iF_{j-1} + F_{i+1}F_j & F_iF_j - F_{i+1}F_{j+1} \end{pmatrix} (-1)^{j+1}.$$

This way, any product matrix P_n can be written as a product of blocks A^iB^j, $A^kB^l, \ldots$, and given in terms of the exponents $i, j, k, l, \ldots$. We will use these exponents to represent the trace. The number of blocks in a product matrix P_n is not fixed, so when using block representation we will use the notation B_q to denote a product matrix with q blocks.

Note that in the case of one block, as Statement 5 of Theorem 7.2, we have

$$\mathrm{tr}(B_1) = \mathrm{tr}(A^iB^j) = -F_{i-1}F_{j-1} + 2F_iF_j - F_{i+1}F_{j+1}, \tag{7.5}$$

where we have assumed that j is odd. For even j, the trace is negated. As the number of blocks increases, it will be convenient to represent each m-termed product in the trace sum by an m-tuple comprised of the terms $-1, 0, 1$. We can do this using a simple mapping. We let $F_{i-1} \mapsto -1$, $F_i \mapsto 0$ and $F_{i+1} \mapsto 1$ and a product of $m = 2q$ Fibonacci terms is mapped to an ordered $2q$-tuple according to the alphabetical order of the indices. For example, the sum in Equation (7.5) can be written as

$$\mathrm{tr}(B_1) \mapsto -(-1,-1) + 2(0,0) - (1,1).$$

We will call these $2q$-tuples *simplified Fibonacci products*, and their original forms, found in the trace equation (7.5), *Fibonacci products*. Let us look at another example before considering the general case.

Example 7.3. For two blocks, Lemma 7.1 gives us

$$B_2 = A^iB^jA^kB^l =$$
$$\begin{pmatrix} F_{i-1} & F_i \\ F_i & F_{i+1} \end{pmatrix} \begin{pmatrix} -F_{j-1} & F_j \\ F_j & -F_{j+1} \end{pmatrix} (-1)^{j+1} \begin{pmatrix} F_{k-1} & F_k \\ F_k & F_{k+1} \end{pmatrix} \begin{pmatrix} -F_{l-1} & F_l \\ F_l & -F_{l+1} \end{pmatrix} (-1)^{l+1}. \tag{7.6}$$

Rather than write out the product matrix, we will list its terms as simplified Fibonacci products, in Table 7.2. Suppose we have $B_2 = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. (We will also use the terms α,

a	b	c	d
$(0,0,0,0)$	$(0,0,-1,0)$	$(0,0,1,0)$	$(0,0,0,0)$
$(0,1,0,-1)$	$(-1,0,0,0)$	$(1,0,0,0)$	$(0,-1,0,1)$
$(-1,0,1,0)$	$(0,1,1,1)$	$(1,1,0,-1)$	$(1,0,-1,0)$
$(-1,-1-1,-1)$	$(-1,-1,0,1)$	$(0,-1,-1,-1)$	$(1,1,1,1)$
$-(0,1,1,0)$	$-(0,0,0,1)$	$-(0,0,0,-1)$	$-(0,0,1,1)$
$-(0,0,-1,-1)$	$-(0,1,0,0)$	$-(0,-1,0,0)$	$-(1,0,0,1)$
$-(-1,0,0,-1)$	$-(-1,0,1,1)$	$-(1,1,1,0)$	$-(1,1,0,0)$
$-(-1,-1,0,0)$	$-(-1,-1,-1,0)$	$-(1,0,-1,-1)$	$-(0,-1,-1,0)$

Table 7.2: Simplified Fibonacci products for B_2.

β, γ and δ as corresponding general position markers in our matrices.) Recall that each term in Table 7.2 is mapped from a product of Fibonacci numbers, for example, $F_{i-1}F_jF_{k+1}F_l \mapsto (-1,0,1,0)$, and the sum of each column gives us an entry of B_2. We have assumed that j and l are odd. If exactly one of them is even, all terms in the table switch sign.

We cannot give an explicit formula for the trace of a product matrix B_q. We are however, able to completely characterize the trace. We must first examine the structure of the matrix entries of B_q.

Lemma 7.2. *Given the block representation B_q of a product matrix, the simplified Fibonacci products belonging to entries b and c do not contain 1 and -1 adjacent to each other. For simplified Fibonacci products in entries a and d, this is true even if we consider our simplified Fibonacci product as a loop.*

Proof: We can prove this result using two simple inductions. First, suppose that a block representation matrix B_q has entries a, b, c and d composed of simplified Fibonacci products, where those in a do not begin or end with 1, those in d do not begin or end with -1, those in b do not begin with 1 or end with -1 and those in c do not begin with -1 or end with 1. For the initial case, B_1, we can rewrite the matrix in Equation (7.3) as

$$B_1 \mapsto \begin{pmatrix} -(-1,-1)+(0,0) & (-1,0)-(0,1) \\ -(0,-1)+(1,0) & (0,0)-(1,1) \end{pmatrix}, \tag{7.7}$$

where we have extended the mapping of Fibonacci products to simplified Fibonacci products, to matrices. It is clear that these simplified Fibonacci products satisfy the required conditions. We must now show that this is true for B_{q+1}.

Assuming for now that j is odd, we have

$$B_{q+1} = B_q A^i B^j = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} -F_{i-1}F_{j-1} + F_iF_j & F_{i-1}F_j - F_iF_{j+1} \\ -F_iF_{j-1} + F_{i+1}F_j & F_iF_j - F_{i+1}F_{j+1} \end{pmatrix} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix},$$

where

$$\begin{aligned} a' &= a(-F_{i-1}F_{j-1} + F_iF_j) + b(-F_iF_{j-1} + F_{i+1}F_j), \\ b' &= a(F_{i-1}F_j - F_iF_{j+1}) + b(F_iF_j - F_{i+1}F_{j+1}), \\ c' &= c(-F_{i-1}F_{j-1} + F_iF_j) + d(-F_iF_{j-1} + F_{i+1}F_j), \\ d' &= c(F_{i-1}F_j - F_iF_{j+1}) + d(F_iF_j - F_{i+1}F_{j+1}). \end{aligned}$$

We can then write $B_{q+1} \mapsto$

$$\begin{pmatrix} -a(-1,-1) + a(0,0) - b(0,-1) + b(1,0) & a(-1,0) - a(0,1) + b(0,0) - b(1,1) \\ -c(-1,-1) + c(0,0) - d(0,-1) + d(1,0) & c(-1,0) - c(0,1) + d(0,0) - d(1,1) \end{pmatrix}. \tag{7.8}$$

To complete the induction, we need to check that the simplified Fibonacci products in this matrix meet the criteria. The products in position α are $a(-1,-1)$, $a(0,0)$, $b(0,-1)$ and $b(1,0)$. None of these terms ends with 1, and we know terms a and b do not begin with 1. Simplified Fibonacci products in position β are $a(-1,0)$, $a(0,1)$, $b(0,0)$ and $b(1,1)$. As before, none of these terms begins with 1, and it is easy to see none of them end in -1. The products in position γ are $(c,-1,-1)$, $(c,0,0)$, $(d,0,-1)$ and $(d,1,0)$. None of these terms end with 1, and as we know terms c and d do not begin with -1. The products in position δ are $c(-1,0)$, $c(0,1)$, $d(0,0)$ and $d(1,1)$. As before, none of these terms begin with -1 and it is clear that none end in -1.

Now, we can prove the main result using another induction. Assume that for simplified Fibonacci products in B_q, we never find 1 and -1 adjacent. The initial case is easy to see, by again looking at the matrix B_1 in Equation (7.7). Similarly we can look at entries from B_2 in Table 7.2. We must now show our result is true for

B_{q+1}. Consider the term in position α of the matrix in Equation (7.8). Since in B_q, terms in position α do not begin or end with 1 and terms in position β do not begin with 1 or end with -1, the terms $a(-1,-1)$, $a(0,0)$, $b(0,-1)$ and $b(1,0)$ cannot have 1 and -1 adjacent, even in a loop. Similarly, since terms in position δ of B_q do not begin or end with -1 and terms in position γ do not begin with -1 or end with 1, the terms $c(-1,0)$, $c(0,1)$, $d(0,0)$, and $d(1,1)$ in position δ of B_{q+1} cannot have 1 and -1 adjacent, even under a loop. The terms in positions β and γ are also prevented from having 1 and -1 adjacent because of the restrictions. Notice that for these positions we make no mention of a loop. It is therefore possible to have the term $c(-1,-1)$ where simplified Fibonacci products in c begin with 1. Also note that we were able to ignore the ± 1 coefficients of the simplified Fibonacci products here because they do not come into play in our proof. $\diamond$

Proposition 7.1. *Given the block representation B_q of a product matrix, the simplified Fibonacci products contain an even number (including zero) of* 0*'s between equal non-*0 *terms, and an odd number of* 0*'s between unequal non-*0 *terms. For terms in positions a and d in B_q, the results extends to simplified Fibonacci products as loops.*

Proof: We will need to use several simple inductions here, and again we will not worry about coefficients of our matrices, i.e., we will assume powers of B are odd. First, we want to show that if terms in positions a or c of B_q end in 0 (excluding the all-0 term), we must have an even number of 0's preceded by -1, or an odd number of 0's preceded by 1. If terms in positions b or d end in 0 (excluding the all-0 term), we must have an odd number of 0's preceded by -1 or an even number of 0's preceded by 1.

As the initial case, for B_1 in Equation (7.7), the terms which end in 0, but are not all-0, are $(-1,0)$ in position β, which has an odd number of 0's after the -1, and $(1,0)$ in position γ, which has an odd number of 0's after the 1, as required. It is also easy to see that the assumptions are also true for B_2, by observing Table 7.2. Suppose now that they are true for B_q. We can again observe the matrix in Equation (7.8) to show that the rules also hold for $q+1$ blocks. Consider the terms in positions

α and γ of B_{q+1}:

$$a(-1,-1),\ a(0,0),\ b(0,-1),\ b(1,0),$$
$$a(-1,-1),\ c(0,0),\ d(0,-1),\ d(1,0).$$

The terms $a(0,0)$ and $c(0,0)$ add two 0's to the end so that the parity of the number of ending 0's is unchanged. The terms $b(1,0)$ and $d(1,0)$ fit the rule of 1 followed by an odd number of 0's. Similarly, the terms in positions β and δ of B_{q+1} have $b(0,0)$ and $d(0,0)$, so that the parity of 0's is unchanged, and $a(-1,0)$ and $c(-1,0)$, where -1 is followed by an odd number of 0's.

Furthermore, if we consider the beginnings of our simplified Fibonacci products, we have quite similar results, except that the parity of the number of 0's at the beginning switches from what it was at the end of the product for terms in positions β and γ in B_q and remains the same for terms in positions α and δ. This is easily observed in Equation (7.7) for B_1 and in Table 7.2 for B_2. Suppose these conditions are true for B_q, i.e., if terms in position a or b begins with a group of 0's, its size must be odd if it is followed by 1 and even if it is followed by -1. If terms in positions c or d begins with a group of 0's, its size must be odd if followed by -1 and even if followed by 1.

We now want to show that this is true for $q+1$ blocks. Again we will look at the terms in Equation (7.8). Consider the terms in positions α and β of B_{q+1}:

$$a(-1,-1),\ a(0,0),\ b(0,-1),\ b(1,0),$$
$$a(-1,0),\ a(0,1),\ b(0,0),\ b(1,1).$$

The only possible product that can affect the number of initial 0's is the all-0 product and this arises only in terms a and d of B_q. If a contains the all-0 product (which must contain an even number of 0's), adding $(-1,-1)$ or $(-1,0)$ to the end retains the even number of 0's. Adding $(0,1)$ gives us an odd number in accordance with our assumptions. Similarly, for terms in positions γ and δ of B_{q+1}, the affected products are $d(0,-1)$, $d(1,0)$, and $d(1,1)$. If d contains the all-0 product, the first term gives us an odd number of initial 0's and the second two products give us an even number, again in accordance with our assumptions.

We can now use another induction to prove our lemma. For the product matrix B_1 in Equation (7.7), the conditions of our lemma hold, but are only applicable to the terms $(1,1)$ and $(-1,-1)$, where we have zero 0's between the 1's and between the -1's. It is also easy to see that the rules hold for the product matrix B_2, by examining Table 7.2. For example, consider the terms $(0,1,0,-1)$ and $(0,1,1,0)$, both of which occur in position α of the matrix. The first term has an odd number of 0's between the non-equal non-0 terms, even under rotation and the latter term has an even number of 0's, again under rotation, between the two 1 terms. The rotation rule does not apply to the term $(1,0,0,0)$ in position β however. Under rotation, we have three 0's between 1 and itself, which is not an even number.

Now, suppose our rules hold for the product matrix B_q. We can again observe the matrix in Equation (7.8) to show that the rules also hold for $q+1$ blocks. The number of 0's in the products $a(-1,-1)$, $c(-1,-1)$, $b(1,1)$ and $d(1,1)$ is unchanged by adding to the end. If a or c ends in zero, we have seen that we must have -1 or 1 followed by an even or odd number of 0's respectively, which meets the requirement of an even number of zeroes between -1 terms or an odd number of zeroes between 1 and -1. Further, the terms $a(-1,-1)$ and $d(1,1)$ appear in the trace of B_{q+1} so we must consider them under rotation. If a starts with 0, we must have an even or odd number of 0's followed by -1 or 1 respectively, and if d starts with 0 we must have an even or odd number of 0's followed by 1 or -1 respectively. In both cases, we have the correct pattern of zeroes under rotation.

Now let us consider the terms $a(0,0)$, $b(0,0)$, $c(0,0)$ and $d(0,0)$. We know that 1 and -1 cannot be adjacent by Lemma 7.2, so if terms a and d begin and end with non-0 entries, these entries must be equal, and so contain two 0's under rotation. If term a or d begins or ends in 0, adding two 0's to the group (under rotation) does not change the parity, and so groups of 0's must still have the desired property. We need not prove anything about rotation for $b(0,0)$ and $c(0,0)$ because these terms appear in positions β and γ of B_{q+1} respectively. Note from B_1 that the term $(0,0)$ occurs only in terms a and d, and by Equation (7.8) the all-0 term must remain in these positions. We are only concerned with rotations of terms in positions α and δ of B_{q+1} and so if a or d is the all-zero term, we need only check its effects on $a(-1,-1)$ and

$d(1,1)$. Adding an even number of zeroes results in them being contained by equal non-zero terms, as required.

Now consider the terms $b(0,-1)$, $d(0,-1)$, $a(0,1)$ and $c(0,1)$. We have shown in our first induction that if a or c ends in 0, the ending must be 1 or -1 followed by an odd or even number of 0's respectively. This implies that we must have an even group of 0's between two 1's or an odd group of 0's between -1 and 1, as required. If a or c does not end in 0, we have seen in the first induction of Lemma 7.2 that it must end in -1, so that our new ending is $-1, 0, 1$, with an odd number of 0's as required. We can argue similarly for $b(0,-1)$ and $d(0,-1)$. Further, the terms $b(0,-1)$ and $c(0,1)$ belong to the trace of B_{q+1} and so we must check them under rotation. If b begins with 0, we must have an even or odd number of zeroes followed by -1 or 1 respectively, and if c begins with 0, we must have an even or odd number of 0's followed by 1 or -1 respectively. In both cases, we have the correct pattern of zeroes under rotation.

Finally, we must consider the terms $a(-1,0)$, $c(-1,0)$, $b(1,0)$ and $d(1,0)$. If a or c ends in zero, we must have an even or odd number of 0's preceded by -1 or 1 respectively, and if b or d ends in 0 we have an even or odd number of 0's preceded by 1 or -1 respectively. We have the correct parity of 0's between non-0 terms in both cases. Further, only the terms $c(-1,0)$ and $b(1,0)$ appear in the trace of B_{q+1} and so are the only ones we must consider under rotation. If c starts with a non-0 term, it must be 1 by the first induction of Lemma 7.2 and so we have an even number of 0's between -1 and 1 under rotation. If c starts with 0, we have seen in the second induction that if the group of 0's is odd or even, it is followed by -1 or 1 respectively. Adding $(-1,0)$ to the end produces, under rotation, an even or odd number of 0's between equal non-0 terms or unequal non-0 terms, respectively. We can argue similarly for $b(1,0)$. Note also that from Lemma 7.2, 1 and -1 are never found adjacent to each other. Therefore adjacent non-0 terms must be equal, which adheres to this theorem, if we think of these terms as containing zero 0's (an even number) between them. $\diamond$

Corollary 7.2. *The total number of* 0*'s in a simplified Fibonacci product occurring* *in the trace of B_q is always even.*

Proof: We know from Proposition 7.1 that there is an odd number of 0's between terms -1 and 1 and an even number of 0's between 1 and 1 or between -1 and -1. This extends to simplified Fibonacci products in positions α and δ forming loops. Consider any simplified Fibonacci product and remove all even groups of 0's. This does not change the parity of the sum total of 0's. Since even groups of 0's occur between equal non-0 terms, we are now left with groups of 1's and groups of -1's separated by odd-numbered groups of 0's. By Lemma 7.2, we cannot have 1 and -1 adjacent, and by removing even-numbered groups of 0's, this still holds true. We must be left with an even number of non-0 groups of terms separated by odd-numbered groups of 0's. If we had an odd number of non-0 groups of terms, our simplified Fibonacci product would have its first and last non-0 terms equal, in which case we could remove the group of 0's between them, which is even for terms a and d occurring in the trace. We are then left with an even number of odd-numbered groups of 0's, giving an even total. $\diamond$

We will now consider the coefficients of our simplified Fibonacci products.

Proposition 7.2. *If we assume that the power of each matrix B is odd, so by Lemma 7.1,*

$$B^j = \begin{pmatrix} -F_{j-1} & F_j \\ F_j & -F_{j+1} \end{pmatrix},$$

then a simplified Fibonacci product belonging to the trace of B_q has

$$\begin{aligned} \textit{coefficient } 1 &\iff \textit{entries sum to } 0 \pmod 4, \\ \textit{coefficient } -1 &\iff \textit{entries sum to } 2 \pmod 4, \\ \textit{coefficient } 2 &\iff \textit{all-0 product.} \end{aligned}$$

Furthermore, all coefficients are negated if our product matrix contains an odd number of even powers of B.

Proof: We can use another induction here. It is easy to see that the above statements are true for the examples we have already considered, namely, $q = 1, 2$. Now suppose that these statements are true for B_q. For the $(q+1)$-block case we can again consider Equation (7.8). For this reason we must know the sum of the entries in a simplified Fibonacci product belonging to each of the terms a, b, c and d of B_q and not simply for the trace terms. Therefore we will include in our induction the facts that a simplified Fibonacci product belonging to term b has

$$\begin{aligned} \text{coefficient } 1 &\iff \text{entries sum to } 3 \pmod 4, \\ \text{coefficient } -1 &\iff \text{entries sum to } 1 \pmod 4, \end{aligned}$$

and a simplified Fibonacci product belonging to term c has

$$\begin{aligned} \text{coefficient } 1 &\iff \text{entries sum to } 1 \pmod 4, \\ \text{coefficient } -1 &\iff \text{entries sum to } 3 \pmod 4. \end{aligned}$$

The term in position α of B_{q+1} is given by

$$-a(-1,-1) + a(0,0) - b(0,-1) + b(1,0).$$

By simply summing the entries in the simplified Fibonacci products, supposing that a and b have a coefficient of 1, we obtain $2, 0, 2, 0 \pmod 4$ for term in the above sum, respectively, as required. If a and b have coefficient -1, we obtain as sums $0, 2, 0, 2 \pmod 4$. These sums also fit our pattern because the fact that each simplified Fibonacci product is a product of Fibonacci numbers implies that the coefficient -1 of a and b can be moved outside the product, changing the sign of the coefficient in the larger simplified Fibonacci product. Therefore the sums $0, 2, 0, 2$ really correspond to products which have as coefficients $1, -1, 1, -1$. The term in position β is given by

$$a(-1,0) - a(0,1) + b(0,0) - b(1,1).$$

If we assume a and b have coefficient 1, our sums are $3, 1, 3, 1 \pmod 4$ as required, and if a and b have coefficient -1, the coefficient of each of the above terms switches and our sums are $1, 3, 1, 3 \pmod 4$. We can argue similarly for the terms in positions γ and δ in B_{q+1}.

The only thing left to consider is the all-0 term, which has coefficient 2 in the trace. The reason for this is that it occurs once in the term a and once in term d and in positions α and δ of B_{q+1} we find $a(0,0)$ and $d(0,0)$ respectively. In the $q = 1$ case we have the length-2 all-0 term $(0,0)$, and so by adding each new block, a and d will both contain an all-0 term with coefficient 1.

Note that these results rely on the assumption that all powers of B in our product matrix are odd. Recall that

$$B^j = \begin{pmatrix} -F_{j-1} & F_j \\ F_j & -F_{j+1} \end{pmatrix} (-1)^{j+1}.$$

If we have an even value of j, our matrix is negated. Furthermore, if this happens for an odd number blocks, the entire product matrix, and hence the trace, is negated. The coefficients of all simplified Fibonacci products in our trace representation are therefore negated. $\diamond$

Combining our results, we can say the following.

Theorem 7.4. *The trace of a product matrix B_q containing q blocks of the form A^iB^j is comprised of a sum of Fibonacci products which we can write in simplified form using $0, 1,$ and -1, and can consider as a loop. The set of such products contains all $2q$-tuples comprised of $0, 1$ and -1 which obey the following rules:*

1. *The terms 1 and -1 cannot be adjacent.*
2. *The simplified Fibonacci products contain an even number (including zero) of 0's between equal non-0 terms, and an odd number of 0's between unequal non-0 terms.*

Furthermore, we have the following rules concerning the coefficients of simplified Fibonacci products:

3. *The coefficient of a simplified Fibonacci product is 2 for the all-0 term, 1 if and only if the sum of terms in the product is $0 \pmod 4$ and -1 if and only if the sum is $2 \pmod 4$.*
4. *All coefficients are negated if our product matrix contains an odd number of even powers of B.*

This gives us a complete characterization of the trace of any product matrix P_n. The following result gives us an easy way to group simplified Fibonacci products that occur in the trace.

Theorem 7.5. *The set of simplified Fibonacci products occurring in the trace of B_q contains the (almost complete) equivalence class of any simplified Fibonacci product contained in it. Furthermore, all simplified Fibonacci products occurring in the same equivalence class have the same coefficient.*

Proof: We are using the equivalence classes given in Definition 4.3 for periodic coefficient sequences, i.e., the combined set of rotations, reversals and negations. Here we do not consider reduction or extension of our products because we want to keep them at a fixed length. If we think about the rules we have already established, it is easy to see that if a particular simplified Fibonacci product behaves according to Theorem 7.4, all elements of its equivalence class must also. Keep in mind that since we are only interested in trace terms, the rules apply to the simplified Fibonacci products considered as loops.

First, the set of simplified Fibonacci products in the trace must not have adjacent unequal non-0 terms. If this is true of any particular product, it must also be true for those products obtained by rotation, reversal and negation. Second, the parity of the number of 0's between non-0 terms also does not change under rotation, reversal or negation. (And hence the total number of 0's remains fixed also.) Finally, the sums of the non-0 coefficients do not change under rotation or reversal. If we swap 1 and -1, our sums modulo 4 are also unchanged. This is because the swap merely changes the sign of our sum, which can be either 0 or 2 for terms in the trace. Negating these numbers does not change them modulo 4. $\diamond$

7.4 Fibonacci Blocks Type II

In the previous section we investigated the traces of product matrices which were broken into blocks of the form A^iB^j. We gave expressions for the trace in terms of sums of Fibonacci products, which contained F_{i-1}, F_i and F_{i+1}, for the various exponents involved. We can try to extend this idea by varying our blocks and the Fibonacci terms involved. For instance we could write our Fibonacci products in terms of F_i and F_{i-1}, so that our simplified Fibonacci products contain only 0 and 1. We could break our product matrices into blocks of the form A^iB. Note that this corresponds to the tree R, where we cannot have two consecutive lefts (i.e., B's), instead of the original tree T. We can also try writing our product matrices so that they begin and end with powers of A. It turns out that this variation, applied to the Fibonacci blocks in the previous section gives another patterned, but slightly more complicated trace characterization. Of the combinations of the above variations, the following proved to be quite interesting.

Consider Fibonacci blocks of the form A^iB, so that we are dealing with product matrices of the form $A^iBA^jBA^kBA^lB\cdots$. We will devise a reduction scheme so that we can write the trace of a particular product matrix in terms of traces of smaller product matrices, as was done in Section 7.1.

Example 7.4. We start by considering product matrices of the form $P_nA^iBA^iB$, where $P_n = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, and is followed by the repeated block A^iB. Trace reductions for $i \leq 7$ are given in Table 7.3.

i	trace reduction
1	$\mathrm{tr}(P_nABAB) = \mathrm{tr}(P_nAB) - \mathrm{tr}(P_n)$
2	$\mathrm{tr}(P_nA^2BA^2B) = \mathrm{tr}(P_n)$
3	$\mathrm{tr}(P_nA^3BA^3B) = \mathrm{tr}(P_nA^3B) - \mathrm{tr}(P_n)$
4	$\mathrm{tr}(P_nA^4BA^4B) = \mathrm{tr}(P_nA^4B) + \mathrm{tr}(P_n)$
5	$\mathrm{tr}(P_nA^5BA^5B) = 2\mathrm{tr}(P_nA^5B) - \mathrm{tr}(P_n)$
6	$\mathrm{tr}(P_nA^6BA^6B) = 3\mathrm{tr}(P_nA^6B) + \mathrm{tr}(P_n)$
7	$\mathrm{tr}(P_nA^7BA^7B) = 5\mathrm{tr}(P_nA^7B) - \mathrm{tr}(P_n)$

Table 7.3: Trace reduction for repeated Fibonacci blocks, $i \leq 7$.

It is easy to see that the first equation in Table 7.3 holds. We have from Table 7.1

$$\mathrm{tr}(P_nABAB) = \mathrm{tr}\begin{pmatrix} b & -a-b \\ d & -c-d \end{pmatrix} = b - c - d,$$

$$\mathrm{tr}(P_nAB) - \mathrm{tr}(P_n) = \mathrm{tr}\begin{pmatrix} a+b & -a \\ c+d & -c \end{pmatrix} - \mathrm{tr}\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
$$= (a+b-c) - (a+d) = b - c - d.$$

The second equation is trivial because $A^2BA^2B = I$.

We will prove the general case of this form of product matrix in the following propostion. Note that we are again tracking the change in trace of a product matrix P_n upon multiplying by a repeated matrix pattern, as was done in Table 7.1.

Proposition 7.3. *Given a general product matrix $P_n = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we have the following trace formula for product matrices ending in equal blocks:*

$$\mathrm{tr}(P_nA^iBA^iB) = F_{i-2}\mathrm{tr}(P_nA^iB) + (-1)^i\mathrm{tr}(P_n).$$

Proof: We can prove this result quite simply using the fact that

$$A^iB = \begin{pmatrix} F_i & -F_{i-2} \\ F_{i+1} & -F_{i-1} \end{pmatrix}.$$

Expanding the product matrices $P_nA^iBA^iB$ and P_nA^iB and taking the traces gives

$$\begin{aligned} \mathrm{tr}(P_nA^iBA^iB) &= aF_i^2 + bF_iF_{i+1} - aF_{i+1}F_{i-2} - bF_{i+1}F_{i-1} \\ &\quad - cF_{i-2}F_i - dF_{i-2}F_{i+1} + cF_{i-1}F_{i-2} + dF_{i-1}^2, \\ F_{i-2}\mathrm{tr}(P_nA^iB) + (-1)^i\mathrm{tr}(P_n) &= F_{i-2}(aF_i + bF_{i+1} - cF_{i-2} - dF_{i-1}) + (-1)^i(a+d). \end{aligned}$$

Equating, and grouping terms according to their coefficients, we obtain the following four equations:

$$a: \qquad F_i^2 - F_{i-2}F_{i+1} = F_{i-2}F_i + (-1)^i, \tag{7.9}$$

$$b: \qquad F_iF_{i+1} - F_{i-1}F_{i+1} = F_{i-2}F_{i+1}, \tag{7.10}$$

$$c: \qquad F_{i-2}F_{i-1} - F_{i-2}F_i = -F_{i-2}^2, \tag{7.11}$$

$$d: \qquad F_{i-1}^2 - F_{i-2}F_{i+1} = -F_{i-2}F_{i-1} + (-1)^i. \tag{7.12}$$

It is easy to show these four equations hold using simple Fibonacci relations. Let us begin with Equation (7.9). Using Cassini's identity, $F_{n-1}F_{n+1} - F_n^2 = (-1)^n$, we can rewrite this equation as

$$\begin{aligned}
F_{i-2}F_i + F_{i-1}F_{i+1} - F_i^2 &= F_i^2 - F_{i-2}F_{i+1},\\
F_{i-2}F_i - F_i^2 + F_{i+1}(F_{i-2} + F_{i-1}) &= F_i^2,\\
F_iF_{i-2} + F_i(F_{i+1} - F_i) &= F_i^2,\\
F_i(F_{i-2} + F_{i-1}) &= F_i^2,\\
F_i^2 &= F_i^2.
\end{aligned}$$

Similarly, for Equation (7.12) we have

$$\begin{aligned}
-F_{i-2}F_{i-1} + F_{i-1}F_{i+1} - F_i^2 &= F_{i-1}^2 - F_{i-2}F_{i+1},\\
-F_{i-2}F_{i-1} + F_{i+1}(F_{i-2} + F_{i-1}) - F_i^2 &= F_{i-1}^2,\\
-F_{i-2}F_{i-1} + F_i(F_{i+1} - F_i) &= F_{i-1}^2,\\
F_{i-1}(F_i - F_{i-2}) &= F_{i-1}^2,\\
F_{i-1}^2 &= F_{i-1}^2.
\end{aligned}$$

For Equations (7.10) and (7.11) we have the respective expressions

$$\begin{aligned}
F_iF_{i+1} - F_{i-1}F_{i+1} &= F_{i+1}(F_i - F_{i-1}) = F_{i-2}F_{i+1},\\
F_{i-2}F_{i-1} - F_{i-2}F_i &= F_{i-2}(F_{i-1} - F_i) = -F_{i-2}^2,
\end{aligned}$$

completing the proof. ◇

With this result we are now able to remove any repeated blocks A^iBA^iB from our product matrix. We have seen in Theorem 2.9 that rotating terms in a product matrix does not affect the characteristic polynomial, and therefore does not affect the trace. If the repeated block occurs in the center of our product matrix, we can simply rotate it to the end and use Proposition 7.3. Consider the following example.

Example 7.5. We can reduce the trace of the matrix $A^3BA^5BA^3B$ using Proposition 7.3 as follows:

$$\begin{aligned}
\operatorname{tr}(A^3BA^5BA^3B) &= \operatorname{tr}(A^5BA^3BA^3B) = \operatorname{tr}(A^5BA^3B) - \operatorname{tr}(A^5B)\\
&= \operatorname{tr}(A^2A^3BA^3B) - \operatorname{tr}(A^5B) = \operatorname{tr}(A^2A^3B) - \operatorname{tr}(A^2) - \operatorname{tr}(A^5B)\\
&= \operatorname{tr}(A^5B) - \operatorname{tr}(A^2) - \operatorname{tr}(A^5B) = -\operatorname{tr}(A^2).
\end{aligned}$$

Here we have a cancelation in the last line. Also, notice that we have used Proposition 7.3 twice; the second time, we used it to calculate $\mathrm{tr}(A^5BA^3B)$. Note that here we did not have two equal blocks. We can now extend our reduction to unequal blocks occurring in our product matrix.

Corollary 7.3. *Given a general product matrix $P_n = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we have the following trace formula for product matrices ending in blocks of length i and j with $i \leq j$:*

$$\mathrm{tr}(P_nA^iBA^jB) = F_{j-2}\mathrm{tr}(P_nA^iB) + (-1)^j\mathrm{tr}(A^{i-j}P_n). \tag{7.13}$$

Proof: The key is to rewrite our product matrix as $P_nA^iBA^jB = P_nA^{i-j}A^jBA^jB$, so that it is now in the proper form to use Proposition 7.3. Letting P_n in the theorem be P_nA^{i-j}, we have

$$\begin{aligned}\mathrm{tr}(P_nA^{i-j}A^jBA^jB) &= F_{j-2}\mathrm{tr}(P_nA^{i-j}A^jB) + (-1)^j\mathrm{tr}(P_nA^{i-j}) \\ &= F_{j-2}\mathrm{tr}(P_nA^iB) + (-1)^j\mathrm{tr}(A^{i-j}P_n),\end{aligned}$$

where we have $\mathrm{tr}(P_nA^{i-j}) = \mathrm{tr}(A^{i-j}P_n)$ by Theorem 2.9. This allows us to group matrices A at the beginning of the product, and end with B. ◇

We now have a recursive formula for the trace of a product matrix comprised of blocks of the form A^iB. We want to know what we are left with if we apply this recursive formula repeatedly. We have seen in Example 7.5 that our trace boils down to a sum of terms of the form $\mathrm{tr}(A^iB)$ and $\mathrm{tr}(A^j)$, i.e., traces of single blocks, which may or may not contain a B term. As shown in Table 7.3, we may also have Fibonacci coefficients.

A bit of new notation will be useful here. We will use $\hat{B}_q$ to represent a product matrix containing q blocks. We will again map the trace of $\hat{B}_q$ to a simplified form, this time using sums of q-tuples comprised of $0, 1$ and -1. (Recall, traces of matrices B_q were mapped to $2q$-tuples comprised of $0, 1$ and -1.) Now, however, the terms $0, 1$ and -1 will carry a different meaning. The 0 will denote a Fibonacci term and -1 and 1 will correspond to powers of A. The positions of these numbers inside the q-tuple indicate which variables they refer to. The exponent of (-1) will come from the -1 terms in our q-tuple, and initial terms in the q-tuple will determine whether or not the matrix B appears. An example will be most useful here.

Example 7.6. Consider the trace expression in Equation (7.13), but with $P_n = I$:

$$\mathrm{tr}(A^iBA^jB) = F_{j-2}\mathrm{tr}(A^iB) + (-1)^j\mathrm{tr}(A^{i-j}). \tag{7.14}$$

This gives us the trace of $\hat{B}_2$ and contains the variables i and j. Each term in the sum is going to give us an ordered pair. In the first term, i appears as a positive power of A, so we let the first entry in our ordered pair be 1. We also have j appearing in the Fibonacci subscript, so we let the second entry in our order pair be 0. There are no negative powers of A, so we have no -1 terms in the pair. Our ordered pair is thus $(1, 0)$. Similarly in the second term, i and j appear as positive and negative coefficients of A, respectively, and so our ordered pair is $(1, -1)$. There is no 0 term here because we have no Fibonacci coefficients. The exponent of the (-1) term is the sum of the negative exponents of A, which is simply j in this case. We therefore have the mapping

$$\mathrm{tr}(A^iBA^jB) \mapsto (1, 0) + (1, -1),$$

and we call each q-tuple a *simplified Fibonacci-trace product*, or just *simplified product* if the context is known. We will also refer to the original form of the product, given in the trace equation (7.14), as the *Fibonacci-trace product.* Notice that each exponent appears in exactly one of the Fibonacci subscripts or the powers of A.

We can now use this simplified Fibonacci-trace product notation to completely characterize the trace of $\hat{B}_q$ in a non-recursive way.

Theorem 7.6. *The trace of a product matrix $\hat{B}_q$ containing q blocks of the form A^iB is comprised of a sum of Fibonacci-trace products which we can map to simplified form using $0, 1,$ and -1. The set of such products is comprised of exactly all q-tuples containing $0, 1$ and -1 which obey the following two rules:*

1. *The simplified Fibonacci-trace products must begin with either 1 or $(1, -1)$.*
2. *After the initial group we may add either 0 or $(1, -1)$ to the simplified Fibonacci-trace products.*

Furthermore,

3. *A Fibonacci-trace product contains the matrix B if and only if its simplified Fibonacci-trace product begins with* $(1,0)$ *or* $(1,1)$*, i.e., initial group* 1 *rather than* $(1,-1)$.

Proof: For the 1-block case, we have $\mathrm{tr}(\hat{B}_1) = \mathrm{tr}(A^iB)$. In simplified Fibonacci-trace form we can write $\mathrm{tr}(\hat{B}_1) \mapsto (1)$. For the 2-block case, we have $\mathrm{tr}(\hat{B}_2) = \mathrm{tr}(A^iBA^jB)$, which we have seen in Example 7.6 gives the simplified form $\mathrm{tr}(A^iBA^jB) \mapsto (1,0) + (1,-1)$. So we have either started with $(1,-1)$ or started with 1 and added 0. Since we can only add terms to the end of the simplified product, this proves the first rule.

We can now use a strong induction to prove the second rule, i.e., that adding 0 or $(1,-1)$ to the initial group to form all possible q-tuples gives the complete trace of $\hat{B}_q$. For $\hat{B}_2$, we have seen that the two simplified Fibonacci-trace products are the only ones possible. Using Corollary 7.3, we have that for $q = 3$

$$\begin{aligned}
\mathrm{tr}(A^iBA^jBA^kB) &= F_{k-2}\mathrm{tr}(A^iBA^jB) + (-1)^k\mathrm{tr}(A^{i+j-k}B) \\
&= F_{k-2}(F_{j-2}\mathrm{tr}(A^iB) + (-1)^j\mathrm{tr}(A^{i-j})) + (-1)^k\mathrm{tr}(A^{i+j-k}B) \\
&= F_{k-2}F_{j-2}\mathrm{tr}(A^iB) + (-1)^jF_{k-2}\mathrm{tr}(A^{i-j}) + (-1)^k\mathrm{tr}(A^{i+j-k}B).
\end{aligned}$$

In terms of simplified Fibonacci-trace products we have

$$\mathrm{tr}(A^iBA^jBA^kB) \mapsto (1,0,0) + (1,-1,0) + (1,1,-1).$$

Again, we can see that these are the only possible simplified products; if we start with 1, we can add $(1,0)$ or $(0,0)$ and if we start with $(1,-1)$ we can only add 0.

Now suppose our inductive hypothesis is true for product matrices containing q or fewer blocks. We want to show it is also true for $q+1$ blocks. We will use Corollary 7.3 and suppose that the product matrix P_n is of the form $\hat{B}_{q-1}$. If we also suppose that the exponents of A in the last two blocks are r and s, we can write

$$\mathrm{tr}(P_nA^rBA^sB) = F_{s-2}\mathrm{tr}(P_nA^rB) + (-1)^s\mathrm{tr}(A^{r-s}P_n). \tag{7.15}$$

First consider the term $F_{s-2}\mathrm{tr}(P_nA^rB)$. The matrix P_nA^rB contains q blocks and so its trace contains all possible q-tuples as described. By multiplying the trace of this matrix by F_{s-2}, we have taken all simplified products in the trace of P_nA^rB and added a 0 to the s position, i.e., the end. Now consider the term $(-1)^s\mathrm{tr}(A^{r-s}P_n)$.

The matrix P_n contains $q-1$ blocks and so its trace contains all possible $(q-1)$-tuples as described. By multiplying P_n by A^{r-s}, each term $\mathrm{tr}(A^iB)$ or $\mathrm{tr}(A^j)$ in the trace of P_n, becomes $\mathrm{tr}(A^{i+r-s}B)$ or $\mathrm{tr}(A^{j+r-s})$ respectively. We are also multiplying each term by $(-1)^s$. In terms of simplified Fibonacci-trace products, this is the equivalent of adding 1 and -1 to the r and s positions. We now have that the set of simplified Fibonacci-trace products in the trace of $\hat{B}_{q+1}$ is comprised of the set of simplified Fibonacci products in the trace of $\hat{B}_q$ with 0 appended to the end, plus the set of simplified Fibonacci products in the trace of $\hat{B}_{q-1}$ with $(1,-1)$ appended to the end. This tells us that adding 0 or $(1,-1)$ to the initial groups 1 and $(1,-1)$ to form all possible q-tuples gives the complete trace of $\hat{B}_q$.

We can prove Rule 3, that matrix B occurs when our simplified Fibonacci-trace product begins with $(1,0)$ or $(1,1)$, using induction. The first few cases are easy to verify from Table 7.4, which breaks down the Fibonacci trace products for $q \leq 5$, and will be discussed in more detail in Example 7.7. Now suppose that our result holds for matrices containing q or fewer blocks. We want to show that it is also true for $\hat{B}_{q+1}$. In the first term of Equation (7.15) we have that the trace for q blocks, $\mathrm{tr}(P_nA^rB)$, contains B when the simplified product begins with $(1,0)$ or $(1,1)$. We simply multiply by F_{s-2}, which does not affect B, and adds 0 to the end. Similarly for the second term, our assumption is true for the $q-1$ block case, $\mathrm{tr}(P_n)$. Here we add to the exponent of A, and multiply by a power of (-1). This adds $(1,-1)$ to the end of the simplified Fibonacci-trace product and does not affect the presence of matrix B or the initial terms of the simplified product. $\diamond$

We have seen that our simplified Fibonacci-trace products are built up recursively, where those in $\hat{B}_{q+1}$ depend on those in $\hat{B}_q$ and $\hat{B}_{q-1}$. The following example makes this particularly clear.

Example 7.7. For $q \leq 5$, we have Table 7.4. For each value of q we have included the set of simplified Fibonacci-trace products, the expanded forms (Fibonacci-trace products) as well as the variables involved in the subscript of the Fibonacci numbers and the exponents of A. The fourth column indicates whether or not the matrix B is included in the Fibonacci-trace product. Recall that we may have $\mathrm{tr}(A^iB)$ or $\mathrm{tr}(A^j)$

in the expansion of our simplified Fibonacci-trace product. The dotted lines in the table show us how each level has been derived from the previous two by adding either 0 or $(1,-1)$ to the end of the simplified product.

q	F sub.	A exponent	B	simp. prod.	expansion
1	-	i	x	(1)	$\mathrm{tr}(A^iB)$
2	j	i	x	$(1,0)$	$F_{j-2}\mathrm{tr}(A^iB)$
	-	$i-j$		$(1,-1)$	$(-1)^j\mathrm{tr}(A^{i-j})$
3	k,j	i	x	$(1,0,0)$	$F_{k-2}F_{j-2}\mathrm{tr}(A^iB)$
	k	$i-j$		$(1,-1,0)$	$(-1)^jF_{k-2}\mathrm{tr}(A^{i-j})$
	-	$i+j-k$	x	$(1,1,-1)$	$(-1)^k\mathrm{tr}(A^{i+j-k}B)$
4	l,k,j	i	x	$(1,0,0,0)$	$F_{l-2}F_{k-2}F_{j-2}\mathrm{tr}(A^iB)$
	l,k	$i-j$		$(1,-1,0,0)$	$(-1)^jF_{l-2}F_{k-2}\mathrm{tr}(A^{i-j})$
	l	$i+j-k$	x	$(1,1,-1,0)$	$(-1)^kF_{l-2}\mathrm{tr}(A^{i+j-k}B)$
	j	$i+k-l$	x	$(1,0,1,-1)$	$(-1)^lF_{j-2}\mathrm{tr}(A^{i+k-l}B)$
	-	$i-j+k-l$		$(1,-1,1,-1)$	$(-1)^{j+l}\mathrm{tr}(A^{i-j+k-l})$
5	m,l,k,j	i	x	$(1,0,0,0,0)$	$F_{m-2}F_{l-2}F_{k-2}F_{j-2}\mathrm{tr}(A^iB)$
	m,l,k	$i-j$		$(1,-1,0,0,0)$	$(-1)^jF_{m-2}F_{l-2}F_{k-2}\mathrm{tr}(A^{i-j})$
	m,l	$i+j-k$	x	$(1,1,-1,0,0)$	$(-1)^kF_{m-2}F_{l-2}\mathrm{tr}(A^{i+j-k}B)$
	m,j	$i+k-l$	x	$(1,0,1,-1,0)$	$(-1)^lF_{m-2}F_{j-2}\mathrm{tr}(A^{i+k-l}B)$
	m	$i-j+k-l$		$(1,-1,1,-1,0)$	$(-1)^{j+l}F_{m-2}\mathrm{tr}(A^{i-j+k-l})$
	k,j	$i+l-m$	x	$(1,0,0,1,-1)$	$(-1)^mF_{k-2}F_{j-2}\mathrm{tr}(A^{i+l-m}B)$
	k	$i-j+l-m$		$(1,-1,0,1,-1)$	$(-1)^{j+m}F_{k-2}\mathrm{tr}(A^{i-j+l-m})$
	-	$i+j-k+l-m$	x	$(1,1,-1,1,-1)$	$(-1)^{k+m}\mathrm{tr}(A^{i+j-k+l-m}B)$

Table 7.4: Fibonacci trace products for $q \leq 5$.

The following corollaries provide us with a bit more insight into this representation of the trace of $\hat{B}_q$.

Corollary 7.4. *Each Fibonacci-trace product in the trace of $\hat{B}_q$ contains q different variables, each occurring exactly once in either the Fibonacci subscript or the exponent of A.*

Proof: This is a simple consequence of the previous theorem, which states that the set of Fibonacci-trace products in the trace of $\hat{B}_q$ can be represented by the set of all q-tuples which begin with 1 or $1,-1$ and are formed by adding 0 or $1,-1$ to the end. We have assigned each entry in a q-tuple to a specific variable $i,j,k,l,\ldots,$

where the value of the entry tells us the location of the variable in the Fibonacci-trace product. Therefore each variable appears exactly once in the Fibonacci subscript or the exponent of A. Note that we are not including the exponent of the (-1) term in the Fibonacci-trace product here. $\diamond$

Corollary 7.5. *The number of simplified Fibonacci-trace products in the trace of $\hat{B}_q$ is F_{q+1}.*

Proof: This is easy to prove with strong induction. From Table 7.4 we can see that for the first few values of q we have $1, 2, 3$ and 5 simplified products. Now suppose that the number of terms in $\mathrm{tr}(\hat{B}_q)$ is F_{q+1}, and likewise for all block sizes less than q. By Theorem 7.6 the set of simplified Fibonacci-trace products in $\mathrm{tr}(\hat{B}_{q+1})$ comes from adding a 0 to the end of the simplified products in $\mathrm{tr}(\hat{B}_q)$, and $(1, -1)$ to the end of the products in $\mathrm{tr}(\hat{B}_{q-1})$. Therefore the size of such a set is simply $F_{q+1} + F_q = F_{q+2}$, as required. $\diamond$

It is interesting to note that the sequence of final matrices in the Fibonacci-trace products, (either A or B) is the rabbit sequence (see [67, A036299]). This is due to the recursive behaviour of our products and the fact that the final matrix does not change as we extend our simplified Fibonacci-trace products. Recall, the Rabbit Sequence is a string of 0's and 1's which is formed by concatenation. We start with initial terms 1 and 10, then form the next term by concatenating the previous two to obtain 101. Continuing we obtain 10110, 10110101, $1011010110110, \ldots$. This is precisely the sequence of A's and B's read down Table 7.4.

Chapter 8

Conclusion

8.1 Summary

We will now summarize the contents of this thesis and review its contributions. In Chapter 1 we introduced the random Fibonacci sequence along with the result of Viswanath that motivated this thesis, namely, that almost all random Fibonacci sequences grow exponentially at the fixed rate $1.13198824\ldots$. We gave some experimental evidence that the convergence to this number occurs very slowly. We looked at some different ways to generate random Fibonacci sequences by changing the location of the $\pm$ sign or using absolute values, and these ideas are considered again in subsequent chapters of the thesis. We showed that for each of these variations on the definition, the set of all possible $(\frac{1}{2}, 1)$-random Fibonacci sequences is the same in absolute value. We used a binary tree to represent the set of all sequences, an idea which plays a major role later in the thesis.

The majority of the Introduction chapter gives an overview of different types of random sequences, and generalizations of Viswanath's random Fibonacci sequence. In particular, we look at work on the cases where $p \neq \frac{1}{2}$ or there is a non-unity coefficient in the recurrence. We also gave an overview of Viswanath's proof and discussed the role of random matrix theory, as well as the importance of computer calculations in studying random Fibonacci sequences and their variations. Further, we made some connections to physics and other fields of study, and noted some open questions in the literature.

The goal in Chapter 2, and the first major contribution of this thesis, was to remove the randomness from the random Fibonacci sequence. This idea was explored independently and at a lesser depth by McGuire [55]. We defined the coefficient cycle σ_n of length n and used it to generate the periodic coefficient sequence, which for large values of n approximates a random Fibonacci sequence. The set of all possible coefficient cycles of length n represents the set of all possible length-n random coefficient sequences. The key tool for studying these new sequences is the matrix representation of the recurrence, and we saw that for any coefficient cycle, we could associate a product matrix P_n. We reviewed the different types of growth (bounded, linear and exponential), as well as the growth rate, of a general second order linear recurrence. The dominant eigenvalue of the companion matrix of such a recurrence determines the growth rate as well as criteria to determine the growth type, and this information was then translated back to our periodic coefficient sequence and corresponding matrix P_n using a set of n subsequences, a new result. Two important theorems — that rotating or reversing the coefficient cycle does not change the growth rate or type — were also given.

In Chapter 3 we removed the need for eigenvalues to characterize the growth type of our sequences, and focused on the trace and order of a product matrix P_n. Some simple connections were made between trace and n, leading up to our main result of the chapter, namely Theorem 3.3, which gives a characterization of growth type based on the absolute value of the trace and the parity of n. A similar classification using continued fractions is speculated upon. Second, we study the connection between the order of a product matrix and the growth type. To do this we introduce the quotient group $\mathrm{PS^*L}(2,\mathbb{Z})$, which equates matrices $\pm P_n = [P_n]$ in $\mathrm{SL}(2,\mathbb{Z})$. We give a characterization of boundedness based on the order of $[P_n]$ in the cases of trace 0 and trace ± 1 matrices, and show that all bounded periodic coefficient sequences must have product matrices with orders 1, 2, 3 or 6. The latter result was given by McGuire [55] and the former extends results found in a slightly different form in [55]. Lastly, we approximated Viswanath's constant by taking an average of all growth rates obtained from the set of length-n coefficient cycles. The approximation did not even yield two decimal places of accuracy, but an interesting pattern in $n \pmod 6$ is observed. Also, we approximate the growth rate instead by the n^{th} root of the absolute value of the trace, and again find an average.

In Chapter 4 we considered our coefficient cycles as loops, specifically, necklaces, which are equivalent under rotation, and bracelets, which are equivalent under reversal. We have already seen that rotating and reversing the coefficient cycle had no effect on growth rate or type, and we considered two further operations on our bracelets. First, we showed that negating terms in our coefficient cycle, i.e., swapping colors in our bracelet, had no effect on the growth, and second, we showed that reducing or extending by a multiple of the primitive cycle also did not affect growth. We used these four operations to define an equivalence class of coefficient cycles and derived (known) combinatorial formulas which count numbers of necklaces, bracelets and equivalence classes. The point of forming such equivalence classes is that we can reduce the number of coefficient cycles whose growth rate/type we need to find. We made an interesting connection between continuant polynomials and product matrices, and used it to construct some of our proofs.

In Chapter 5 we considered some properties of the binary tree T of all random Fibonacci sequences (in absolute value), and introduced the reduced tree R made from pruning all repeated edges from T. This idea and many related results are found in Rittaud [64], which was the most influential paper for this thesis. Many properties of R were considered, the most important of which tells us that left children of left nodes do not occur in R. We also considered some variations of the tree R based on different definitions of the random Fibonacci sequence and generalization of initial values. Rittaud showed that the growth rate of the expected value of the n^{th} term in a random Fibonacci sequence (in absolute value) is the algebraic number $\alpha = 1.205569431\ldots$. He did this by first finding the analogous growth rate in R, and then using the fact that the rows in T can be written as linear combinations of the rows in R, where the initial values are functions of the golden ratio ϕ. (His proof also contains an error which was pointed out and fixed in a more general setting in a subsequent paper, and which we have corrected.) We gave a new and simpler breakdown of rows in T in terms of those in R and gave an explicit formula for the coefficients of the linear combination, which involved an interesting geometric argument. The coefficients were much more complicated in our case, but formed a neat Pascal-like triangle which is easily extended. We used this breakdown to give a simpler proof of the growth rate α, and conjectured that this value may also be found using sums of traces in T.

In Chapter 6 we looked at products of the n^{th} rows of terms in T and R. We found some relations among the products of all nodes, left nodes only and right nodes only of rows in R. We also counted equivalence classes of coefficient cycles corresponding to the tree R and its variation R_1. The main result of this chapter, and probably the major result of the thesis, is a new computation of Viswanath's constant using a geometric mean of nodes in T given by Kalmár-Nagy [44]. This computation alone does not give an accurate approximation of the constant, so we computed the geometric mean in R instead, which converged much more quickly. A heuristic result of Rittaud [64] links Viswanath's constant to this new value using ϕ, and we provided a rigorous proof. With this, we were able to calculate Viswanath's constant to 8 decimal places of accuracy, although we feel this could be improved with a more efficient program. Viswanath's calculation was extremely difficult and required many different components, including random matrix theory, the Stern-Brocot tree, a fractal measure and an intensive computer calculation, so we feel this was a vast simplification of the problem. Further, we give another formulation of the link between the two constants, using the composition of T in terms of R given in Chapter 5.

Chapter 7 is a detailed study of patterns in the traces of product matrices P_n, corresponding to coefficient cycles in both T and R. Everything in this section is original. We start with a recursive system of equations that allows us to write the trace of a product matrix P_{n+3} in terms of P_{n+2}, P_{n+1} and P_n. This also allows us to consider the average trace values over all product matrices of a given length. We consider the traces and corresponding growth types of k^{th} powers of some particular product matrices, as well as some products of such powers. We also look at the change in trace upon multiplying P_n by particular powers of product matrices. Lastly we give complete characterizations of the traces of products of Fibonacci blocks for both of the forms A^iB^j and A^iB, which correspond to products in trees T and R respectively. We do this using mappings from products of Fibonacci terms to q-tuples containing $0, 1$ and -1. Some interesting structure appears in the traces of such products.

8.2 Further Work

The most important future work we feel that could be done is to write a better program for computing the geometric mean of nodes in ρ_n. As seen, our computation gave us 8 correct digits of Viswanath's constant, but the simplicity of this method leads us to believe that we can achieve a better outcome. It would also be interesting to find the rate of convergence using our method as well as Viswanath's. Another important question that remains open is how our method of computing Viswanath's constant using the geometric mean carries over to sequences defined using a coefficient (perhaps a general coefficient) in the recurrence, or with $p \neq \frac{1}{2}$. In the latter case we would need to use a weighted geometric mean. We could also consider increasing the number of terms in our recurrence, as well as some other of the numerous generalizations mentioned in Section 1.8. It would be interesting to move beyond sequences generated by coefficient cycles and look at the growth of sequences generated by non-periodic patterns of $\pm$ signs. Another question is whether or not we can construct a random Fibonacci sequence (of the non-random type) with a given growth rate.

In Chapter 2, we discussed some properties of our group G, and its subgroup with positive determinant, K. We saw that with elements taken modulo 2, all matrices in G are of the form $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, or $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, and K is exactly equal to the set of matrices in G with determinant 1 and this form modulo 2. It would be nice to have some more information about the group G, and in particular, if possible, a complete characterization of its elements. It would be quite interesting to further investigate the properties of the groups formed by taking matrix products of $A, \hat{B} = \left(\begin{smallmatrix} 0 & \pm 1 \\ 1 & 1 \end{smallmatrix}\right)$ or $A, \hat{B}$ and B, as discussed in Chapters 2 and 3. Further we could study the changes to the trace, growth and order results found in these chapters. We could also extend the criteria for growth based on the trace and order of a product matrix to the generalized matrices $\left(\begin{smallmatrix} 0 & 1 \\ 1 & \pm\beta \end{smallmatrix}\right)$ which are the result of adding a coefficient to the recurrence.

It may be of interest to further investigate our product matrices P_n as compositions of Möbius transformations. Multiplication by matrices A and B correspond to the transformations $f(z) = \frac{1}{z \pm 1}$. We could consider a sequence of A's and B's geometrically, as a series of Möbius transformations which map points or sets in the complex plane. Properties such as fixed points, sinks and sources, and limit sets could be studied. We can use a circle pairing map to study the effects of our Möbius transformations $f(z) = \frac{1}{z\pm1}$. The transformation $f(z) = \frac{1}{z}$ is an inversion, i.e., maps the outside (inside) of one disc to the inside (outside) of another, and $f(z) = z \pm 1$ is a translation. Limit points of a Möbius transformation represent infinite sequences of nested disks, which shrink down to a point. Recall the maps given in Section 3.1. For parabolic maps, nested disks shrink linearly, for loxodromic maps, they shrink exponentially, and for elliptic maps, shrinking does not occur. (See Mumford et al. [57, p. 97, 123, 169].)

In Chapter 3, we gave approximations to Viswanath's constant by calculating the average value of the n^{th} roots of absolute values of dominant eigenvalues, and also trace values. A further look at the proportion of dominant eigenvalues or trace values that are small (or give bounded growth, for example), as n increases, might help us to better understand the approximation given, or to give a better approximation. We also saw in Figure 3.1 that a pattern exists in the average values of growth rates for values of $n \pmod 6$. All three groups of values appear to converge to Viswanath's constant, although focusing on one of these groups, for example $n \equiv 1, 5 \pmod 6$, which gives exponential growth only, may simplify the approximation or give us some insight.

Conjecture 3.1, on the connection between the continued fraction obtained from the coefficient cycle and the growth type deserves to be studied further. We feel the explanation for these results is obtainable without too much difficulty. We could further investigate the connection between product matrices and continuants given in Theorem 4.1. We can write an expression for the growth rate in terms of continuant polynomials by simply finding the dominant eigenvalue of the matrix given in the theorem. We would then need to find a way to average all possible such growth rates, to give an expression for Viswanath's constant.

We mentioned that the original intent of using equivalence classes was to try to shed some light on Viswanath's constant by determining the number of equivalence classes, the size of each equivalence class, the growth type/rate of each equivalence class, and then constructing a formula for the constant with this information. The problem of finding the size of the equivalence classes turned out to be difficult, and this approach could be further considered. We could also try grouping all coefficient cycles of length n, not according to equivalence class, but according to growth rate. Since coefficient cycles of the same length have the same determinant, equal growth rate also means equal trace (in absolute value). Since we know that multiple equivalence classes may share the same trace value, we would essentially be combining existing equivalence classes into larger sets. This could be a useful grouping, in terms of enumeration. We saw in the proof of Theorem 2.11 that similar matrices share trace, determinant, eigenvalues and order, and so we could also look at sets of product matrices $P_n \in G$ which are similar, if they happen to differ from the above groupings.

Recall from Chapter 6 that the sequence defined by Equation (6.1), which gives the number of coefficient cycles with $(--)$ forbidden (i.e., in R) and not of the form $((+)^i(-)^j)$, matched the sequence describing the partial sum operator applied three times to Fibonacci numbers, and also the number of 132-avoiding two-stack sortable permutations which contain exactly one subsequence of type 51234, with offset 4. It would be interesting to examine these possible connections.

There were numerous calculations given in Sections 5.3 and 6.2 of different types of growth rates which require further investigation. Conjecture 5.1 claims that the growth rate of the expected value of the traces (in absolute value) for all product matrices P_n corresponding to level n of the tree T gives us the constant $1.20556943\ldots$, which we have seen in Corollary 5.5 is the growth rate of the expected value of the n^{th} terms in a random Fibonacci sequence. This calculation could be repeated using eigenvalues rather than trace values, and we expect a similar result. We can approximate the a.s. growth rate of sequences in R or R_1 by taking the arithmetic mean of n^{th} roots of traces as in Equation (6.28), or approximate the growth rate of the expected value of the trace as in Equation (6.29). We expect these values to approximate $\rho = 1.33683692\ldots$ and $\frac{\alpha}{\phi} = 1.363116873\ldots$ respectively, as was the case for corresponding growth rates in T (see Table 6.4), but some quick calculations showed that in R, both of these values were around 1.15 and in R_1 they were both around 1.26. These calculations for values in Rittaud's trace tree (see section 6.2) give approximations closer to our anticipated values. Further, we can repeat the above calculations using the dominant eigenvalue instead of the trace.

We can also try taking geometric means of traces and dominant eigenvalues, instead of just the arithmetic means used in the calculations above. We had luck with the geometric mean of the n^{th} terms in a random Fibonacci sequence, as given in Theorem 6.2. This calculation led us to very good approximations of ρ and τ. We can also experiment with both the ratio and n^{th} root definitions of the growth rate for exponential growth, as given in Definition 2.4. Sometimes one works much better than the other, as was the case for the approximations to ρ and τ in Tables 6.2 and 6.3. These ideas about growth rates require further investigation.

In Chapters 5 and 6 we saw that the values $1.13198824\ldots$ and $1.205569431\ldots$ fit the generalized mean inequality and the theorem of arithmetic and geometric means (both of which follow from Jensen's inequality on the expected value of convex functions). These values represent the expected value of the growth rate of a random Fibonacci sequence, and the growth rate of the expected value of a random Fibonacci sequence (both in absolute value), respectively. It is interesting that nothing is known about the nature of the former value (although it's most likely irrational), but the latter value is an algebraic number of degree 3. A deeper study of these ideas may lead us to better understand the connection between these two numbers, or perhaps allow us to find a lower bound for $1.205569431\ldots$ (giving information about $1.13198824\ldots$).

Recall Theorem 6.1, which stated a number of (related) relations among products of all nodes, left nodes and right nodes in rows of R. We mentioned that an additional independent relation could lead us to an exact expression for $P(\rho_n)$ and hence to Viswanath's constant itself. The constant $1.5836413\ldots$ could be used for this purpose as discussed in Section 6.1. Although efforts were not successful, we could continue looking for an additional expression or recurrence for one of the above mentioned products. We could also try to further exploit the fractal nature of R (its subtrees have the same shape and growth as the overall tree) to find an equation for $P(\rho_n)$. The subtrees have different initial values, but this does not affect the growth rate of the product.

In Section 5.1, we discussed the general tree $R_{(a,b)}$ for $a < b$. Attempts to find a pattern in the general nodes were unsuccessful; however, if this was possible, we could write an expression for $P(\rho_n)$, and hence the a.s. growth rate of sequences, in terms of a and b. We could then substitute $a = 1$ and $b = 2$ (recall we had shifted down one row to match the general form of R) to find an expression for Viswanath's constant.

Recall from Chapter 1 that Kalmár-Nagy [44] derives a generating function that characterizes the multisets derived by taking the multiset sum of the Minkowski sums and differences of consecutive rows in T_2. He does this by using properties of the Minkowski sum/difference to construct a recurrence for the generating function, which he then explicitly finds. Further, he states that a recurrence relation for the generating function for the entries in row n of T_2 can be constructed. We have not figured out how to do this, but such a generating function would encode the frequency of nodes occurring in any given row, from which we could deduce the geometric mean and hence growth rate, i.e., Viswanath's constant. We could also look for such a generating function for nodes in R. Similarly, finding a pattern in, or characterization of the occurrence of trace values in R or T, could lead to a deterministic growth rate formula.

In Corollary 7.1, we considered the average trace value over all product matrices P_n. This was not particularly useful, as the average is either 0 or 2 depending on the parity of n. The reason is that we need to take the absolute values of the traces in Theorem 7.1 before averaging. If we could somehow determine the signs of the traces, we could derive a (possibly recursive) expression for the average trace, which could prove Conjecture 5.1, or take products or n^{th} roots of traces and possibly find an expression for Viswanath's constant, as given in Equation (3.17). Also, sequences for these sums of absolute values of traces in T and R were given in Chapters 5 and 6, from which we could continue to look for a pattern or recurrence.

In [61], Trefethen is quoted as saying (about random Fibonacci sequences) "Looking for patterns and trends among such sequences of numbers can be a fascinating pastime." I certainly agree. I've been at it for the last five years!

Appendix A

Maple Programs

```
Digits := 15;
printlevel := 0 :
with(linalg) : with(combinat) :
A[0] := Matrix(2, 2, [0, 1, 1, 1]) :
A[1] := Matrix(2, 2, [0, 1, 1, -1]) : V := matrix(2, 1, [1, 1]) : M := Matrix(2, 2, [3, 2, 4, 3]) :
# a1 := 0:
 P := 1 :
 for a1 from 0 to 1 do
 for a2 from 0 to 1 - a1 do
   for a3 from 0 to 1-a2 do
      for a4 from 0 to 1-a3 do
         for a5 from 0 to 1 - a4 do
            for a6 from 0 to 1 - a5 do
               for a7 from 0 to 1 - a6 do
                  for a8 from 0 to 1 - a7 do
                     for a9 from 0 to 1 - a8 do
                        for a10 from 0 to 1 - a9 do
                           for a11 from 0 to 1 - a10 do
                              for a12 from 0 to 1 - a11 do
                                 for a13 from 0 to 1 - a12 do
                                    for a14 from 0 to 1 - a13 do
                                       for a15 from 0 to 1 - a14 do
                                          for a16 from 0 to 1 - a15 do

            B := evalm(((|A[a16].(|A[a15].(|A[a14].(|A[a13].(|A[a12].(|A[a11].(|A[a10].(|A[a9].
((|A[a8].(|A[a7].(|A[a6].(|A[a5].(|A[a4].(| A[a3].(|A[a2].(|A[a1].M|)|)|)|)|)|)|)|)|)|)|)|)|)|)|)|).V):
            P := P·B[2, 1] :
 od od od od od od od od od od od od od od od od ;
print(n = 24, evalf(P^(1/fibonacci(23))))
```

Figure A.1: 1 of 13 programs used to compute the ratio approximation to ρ for $n = 24$.

```
f := proc(k, n);
  option remember :
  f(1, 1) := 1; f(2, 1) := 1 :
   if modp(n, 4) = 0 then  return f(k − 1, n/2) + f(k − 2, n/4) :
   elif modp(n, 4) = 1 then return |f(k − 1, (n + 1)/2) − f(k − 2, (n + 3)/4)|:
   elif modp(n, 4) = 2 then return f(k − 1, n/2) + f(k − 2, (n + 2)/4) :
   elif modp(n, 4) = 3 then return |f(k − 1, (n + 1)/2) − f(k − 2, (n + 1)/4)|:
   end if
 end proc;

 k := 21; P := 1;
 for n from 1 to 2^(k − 2) do
         if f(k, n) > 0 then
           P := P·f(k, n); fi:
 od;


 print(k, evalf(P^(1/2^(k − 2))));
```

Figure A.2: Approximation of Viswanath's constant using the geometric mean.

```
printlevel := 0 : with(linalg) :
with(RandomTools) :
 A := matrix(2, 2, [0, 1, 1, 1]) :
 B := matrix(2, 2, [0, 1, 1,-1]) :
 for j from 1 to 20 do
 for n from 1 to 40000 do M[n] := Generate(choose({A, B}))od:
 P[1] := M[1] :
 for i from 2 to 40000 do P[i] := multiply(P[i-1], M[i]) od:
  eigenvalues(P[40000]) :
  if  evalf((abs(%[1]))^(1/40000)) ≥ evalf((abs(%[2]))^(1/40000))
 then  g[j] := abs(evalf((abs(%[1]))^(1/40000))) :
   else    g[j] := abs(evalf((abs(%[2]))^(1/40000))) fi:
od: print(average = (add(g[j], j = 1..20))/20);
                 average = 1.132329640
```

Figure A.3: The average of j growth rates of random Fibonacci sequences of length n.

```
printlevel := 0 : s := 0 :
with(linalg) :

A[0] := matrix(2, 2, [0, 1, 1, 1]);
A[1] := matrix(2, 2, [0, 1, 1,-1]);

for a1 from 0 to 1 do
for a2 from 0 to 1 do
 for a3 from 0 to 1 do
 for a4 from 0 to 1 do
 for a5 from 0 to 1 do
 for a6 from 0 to 1 do
 for a7 from 0 to 1 do
 for a8 from 0 to 1 do
 for a9 from 0 to 1 do

B := evalm(A[a1]&*A[a2]&*A[a3]&*A[a4]&*A[a5]&*A[a6]
                &*A[a7]&*A[a8]&*A[a9]) : (eigenvalues(B)) :
 if  evalf((abs(%[1]))^(1/9)) ≥ evalf((abs(%[2]))^(1/9)) then
g := (evalf(abs(%[1])^(1/9))) else   g := (evalf(abs(%[2])^(1/9)))
 fi: s := s + g;
 od od od od od od od od od od od od od od od od od od;
print(s, s/2^9);
```

Figure A.4: Approximation of Viswanath's constant using eigenvalues for $n = 9$.

Bibliography

[1] Roger C. Alperin. Rationals and the modular group. *American Mathematical Monthly*, 106:771–773, Oct 1999.

[2] Zai-Qiao Bai. On the cycle expansion for the Lyapunov exponent of a product of random matrices. *Journal of Physics A: Mathematical and Theoretical*, 40(29):8315–8328, Jul 2007.

[3] Zai-Qiao Bai. An infinite transfer matrix approach to the product of random 2×2 positive matrices. *Journal of Physics A: Mathematical and Theoretical*, 42(1):015003, Jan 2009.

[4] Zai-Qiao Bai. A transfer operator approach to random Fibonacci sequences. *Journal of Physics A: Mathematical and Theoretical*, 44(11):115002, Mar 2011.

[5] Prem N. Bajaj. Equality of limits in ratio and root tests. *Mathematics Magazine*, 71(4):299, Oct 1998.

[6] Alan F. Beardon. *The Geometry of Discrete Groups*, volume 91 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1983.

[7] Richard Bellman. Limit theorems for non-commutative operations I. *Duke Mathematical Journal*, 21(3):491–500, 1954.

[8] E. Ben-Naim and P. I. Krapivsky. Growth and structure of stochastic sequences. *Journal of Physics A: Mathematical and General*, 35(41):L557–L563, Oct 2002.

[9] Florin P. Boca. Products of matrices $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$ and the distribution of reduced quadratic irrationals. *Journal für die reine und angewandte Mathematik*, 606:149–165, 2007.

[10] Miklós Bóna. *Combinatorics of Permutations*, volume 72 of *Discrete Mathematics and its Application*. Chapman & Hall/CRC, New York, 2004.

[11] Philippe Bougerol and Jean Lacroix. *Products of Random Matrices with Applications to Schrödinger Operators*, volume 8 of *Progress in Probability and Statistics*. Birkhäuser, Boston, 1985.

[12] Jason I. Brown, Karl Dilcher, and Dante V. Manna. On a sequence of sparse binomial-type polynomials. *Journal of Mathematical Analysis and Applications*, to appear.

[13] T. Richard Carson. Periodic recurrence relations and continued fractions. *The Fibonacci Quarterly*, 45(4):357–361, Nov 2007.

[14] Hei-Chi Chan. The asymptotic growth rate of random Fibonacci type sequences. *Fibonacci Quarterly*, 43(3):243–255, Aug 2005.

[15] Hei-Chi Chan. The asymptotic growth rate of random random Fibonacci type sequences II. *Fibonacci Quarterly*, 44(1):73–84, Feb 2006.

[16] Philippe Chassaing, Gérard Letac, and Marianne Mora. Brocot sequences and random walks in $SL(2, \mathbb{R})$. In H. Heyer, editor, *Probability Measures on Groups VII,Lecture Notes in Mathematics*, volume 1064, pages 36–48. Springer-Verlag, Berlin, 1983.

[17] William Y. C. Chen. Necklaces, MSS sequences, and DNA sequences. *Advances in Applied Mathematics*, 18:18–32, 1997.

[18] Harvey Cohn. Growth types of Fibonacci and Markoff. *Fibonacci Quarterly*, 17(2):178–183, Apr 1979.

[19] E. R. Colman and G. J. Rodgers. The resistance of randomly grown trees. *Journal of Physics A: Mathematical and Theoretical*, 44(50):505001, Dec 2011.

[20] Louis Comtet. *Advanced Combinatorics: The Art of Finite and Infinite Expansions*. D. Reidel Publishing Company, Dordrecht, Holland, revised and enlarged edition, 1974.

[21] Edgardo Cureg and Arunava Mukherjea. Numerical results on some generalized random Fibonacci sequences. *Computers and Mathematics with Applications*, 59:233–246, 2010.

[22] R. Dawson, G. Gabor, R. Nowakowski, and D. Wiens. Random Fibonacci-type sequences. *Fibonacci Quarterly*, 23(2):169–176, May 1985.

[23] Keith Devlin. New mathematical constant discovered: Descendant of two thirteenth century rabbits. *Devlin's Angle*, Mar 1999. `http://www.maa.org/devlin/devlin_3_99.html`.

[24] Mark Embree and Lloyd N. Trefethen. The growth and decay of random Fibonacci sequences. *London Proceedings, Series A, Mathematical, Physical and Engineering Sciences*, 455(1987):2471–2485, Jul 1999.

[25] Graham Everest, Alf van der Poorten, Igor Shparlinski, and Thomas Ward. *Recurrence Sequences*, volume 104 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, 2003.

[26] Helaman Rolfe Pratt Ferguson. The Fibonacci pseudogroup, characteristic polynomials and eigenvalues of tridiagonal matrices, periodic linear recurrence systems and application to quantum mechanics. *The Fibonacci Quarterly*, 16(5):435–447, Oct 1978.

[27] N. J. Fine. Classes of periodic sequences. *Illinois Journal of Mathematics*, 2(2):285–302, 1958.

[28] Alastair Fletcher. Jordan normal form. `http://homepages.warwick.ac.uk/~masias/jnf.pdf`, Jan 2003. class notes.

[29] H. Furstenberg and H. Kesten. Products of random matrices. *Annals of Mathematical Statistics*, 31(2):457–469, 1960.

[30] Harry Furstenberg. Noncommuting random products. *Transactions of the American Mathematics Society*, 108(3):377–428, Sep 1963.

[31] A. O. Gelfond. *Transcendental and Algebraic Numbers*. Dover Publications, New York, 1960.

[32] E. N. Gilbert and John Riordan. Symmetry types of periodic sequences. *Illinois Journal of Mathematics*, 5(4):657–665, 1961.

[33] Ronald Graham, Donald Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, Boston, MA, second edition, 1994.

[34] Gebhard Greiter. A simple proof for a theorem of Kronecker. *American Mathematical Monthly*, 85(9):756–757, 1978.

[35] Branko Grünbaum and Geoffrey C. Shephard. Satins and twills: An introduction to the geometry of fabrics. *Mathematics Magazine*, 53(3):139–161, May 1980.

[36] G. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, Cambridge, 1952.

[37] Brian Hayes. The Vibonacci numbers. *American Scientist*, 87(4):296–301, Jul-Aug 1999.

[38] Peter Hope. Exponential growth of random Fibonacci sequences. *Fibonacci Quarterly*, 33(2):164–168, May 1995.

[39] W. D. Hoskins and Anne Penfold Street. Twills on a given number of harnesses. *Journal of the Australian Mathematical Society Series A*, 3:1–15, 1982.

[40] Élise Janvresse, Benoît Rittaud, and Thierry de la Rue. How do random Fibonacci sequences grow? *Probability Theory and Related Fields*, 142(3-4):619–648, 2008.

[41] Élise Janvresse, Benoît Rittaud, and Thierry de la Rue. Growth rate for the expected value of a generalized random Fibonacci sequence. *Journal of Physics A: Mathematical and Theoretical*, 42(8):1–18, 2009.

[42] Élise Janvresse, Benoît Rittaud, and Thierry de la Rue. Almost-sure growth rate of generalized random Fibonacci sequences. *Annales de l'Institut Henri Poincaré, Probabilits et Statistiques*, 46(1):135–158, 2010.

[43] Tamás Kalmár-Nagy. The random Fibonacci recurrence and the visible points of the plane. *Journal of Physics A: Mathematical and General*, 39(20):L323–L328, 2006.

[44] Tamás Kalmár-Nagy. A multiset-valued Fibonacci-type sequence. *Advances and Applications in Discrete Mathematics*, 2(2):85–95, 2008.

[45] Ron Knott. Fibonacci numbers and the golden section. `http://www.maths.surrey.ac.uk/hosted-sites/R.Knott/Fibonacci/`, 2011.

[46] I. Krasikov, G. J. Rodgers, and C. E. Tripp. Growing random sequences. *Journal of Physics A: Mathematical and General*, 37(6):2365–2370, Feb 2004.

[47] James Kuzmanovich and Andrey Pavlichenkov. Finite groups of matrices whose entries are integers. *American Mathematical Monthly*, 109(2):173–186, Feb 2002.

[48] Yueheng Lan. Novel computation of the growth rate of generalized random Fibonacci sequences. *Journal of Statistical Physics*, 142:849–861, 2011.

[49] Serge Lang. *Introduction to Modular Forms*. Number 222 in Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, 1976.

[50] Ron Larson, Bruce H. Edwards, and David C. Falvo. *Elementary Linear Algebra*. Houghton Mifflin Company, 1999. Ch. 8–10 found at `http://college.cengage.com/mathematics/larson/elementary_linear/4e/shared/downloads/c10s3.pdf`.

[51] Joseph Lehner. *Discontinuous Groups and Automorphic Functions*. Number VIII in Mathematical Surveys. American Mathematical Society, Providence, 1964.

[52] P. Lévy. Sur les lois de probabilité dont dépendent les quotients complets et imcomplets d'une fraction continue. *Bulletin de la S. M. F.*, 57:178–194, 1929.

[53] Eran Makover and Jeffrey McGowan. An elementary proof that random Fibonacci sequences grow exponentially. *arXiv:math.NT/0510159 v2*, Oct 2005.

[54] Michael L. McCliment. A singular contiguity, multiset sums. `http://singularcontiguity.wordpress.com/2008/05/14/multiset-sum/`, 2008.

[55] Trevor McGuire. On the periodicity of certain recursive sequences. *Fibonacci Quarterly*, 46/47(4):350–355, Nov 2008/2009.

[56] Trevor McGuire. Generalizations of the periodicity of certain recursive sequences. *Fibonacci Quarterly*, 48(2):175–181, May 2010.

[57] David Mumford, Caroline Series, and David Wright. *Indra's Pearls: The Vision of Felix Klein*. Cambridge University Press, 2002.

[58] C. D. Olds. *Continued Fractions*. Random House, 1963.

[59] João Batista Oliveira and Luiz Henrique de Figueiredo. Interval computation of Viswanath's constant. *Reliable Computing*, 8:131–138, 2002.

[60] Yuval Peres. Analytic dependence of Lyapunov exponents on transition probabilities. In Ludwig Arnold, Hans Crauel, and Jean-Pierre Eckmann, editors, *Lyapunov Exponents*, volume 1486 of *Lecture Notes in Mathematics*, pages 64–80. Springer, Berlin, 1991.

[61] Ivars Peterson. Fibonacci at random: Uncovering a new mathematical constant. *Science News*, 155(24):376, Jun 1999.

[62] George Pólya. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Mathematica*, 68(1):145–254, 1937.

[63] John Riordan. *Introduction to Combinatorial Analysis*. John Wiley & Sons, New York, 1958.

[64] Benoît Rittaud. On the average growth of random Fibonacci sequences. *Journal of Integer Sequences*, 10(07.2.4), 2007.

[65] Byron Schmuland. Random harmonic series. *American Mathematical Monthly*, 110(5):407–416, May 2003.

[66] Clément Sire and Paul L. Krapivsky. Random Fibonacci sequences. *Journal of Physics A:Mathematical and General*, 34:9065–9083, Oct 2001.

[67] Neil J. A. Sloane. Online encyclopedia of integer sequences. `http://oeis.org/`, 2012.

[68] Terence Tao. *Topics in Random Matrix Theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012.

[69] S. Vajda. *Fibonacci & Lucas Numbers, and the Golden Section*. Ellis Horwood Limited, Chichester, England, 1989.

[70] A. Vince. Period of a linear recurrence. *Acta Arithmetica*, 39:303–311, 1981.

[71] Divakar Viswanath. *Lyapunov Exponents from Random Fibonacci Sequences to the Lorenz Equations*. PhD thesis, Cornell University, Aug 1998.

[72] Divakar Viswanath. Random Fibonacci sequences and the number 1.13198824... *Mathematics of Computation*, 69(231):1131–1155, Jun 1999.

[73] Michael Weinstein. *Examples of Groups*. Polygonal Publishing House, Passaic, N.J., 1977.

[74] Raymond E. Whitney. Composition of recursive formulae. *The Fibonacci Quarterly*, 4(4):363–366, Dec 1966.

[75] Thomas G. Wright and Lloyd N. Trefethen. Computing Lyapunov constants for random recurrences with smooth coefficients. *Journal of Computational and Applied Mathematics*, 132:331–340, 2000.