



Spring 5-7-2021

## Streaming Down the Stern-Brocot Tree: Finding and Expressing Solutions to Pell's Equation in $SL(2, \mathbb{Z})$

Marcus L. Shell

Jacksonville State University, [mshell2@stu.jsu.edu](mailto:mshell2@stu.jsu.edu)

Follow this and additional works at: [https://digitalcommons.jsu.edu/etds\\_theses](https://digitalcommons.jsu.edu/etds_theses)



Part of the [Algebra Commons](#), and the [Number Theory Commons](#)

---

### Recommended Citation

Shell, Marcus L., "Streaming Down the Stern-Brocot Tree: Finding and Expressing Solutions to Pell's Equation in  $SL(2, \mathbb{Z})$ " (2021). *Theses*. 13.

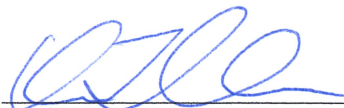
[https://digitalcommons.jsu.edu/etds\\_theses/13](https://digitalcommons.jsu.edu/etds_theses/13)

This Thesis is brought to you for free and open access by the Theses, Dissertations & Graduate Projects at JSU Digital Commons. It has been accepted for inclusion in Theses by an authorized administrator of JSU Digital Commons. For more information, please contact [digitalcommons@jsu.edu](mailto:digitalcommons@jsu.edu).

THESIS APPROVAL

Candidate: Marcus Lee Shell  
Major: Mathematics  
Thesis Title: Streaming Down the Stern-Brocot  
Tree: Finding and Expressing  
Solutions to Pell's Equation in  $SL_2(\mathbb{Z})$

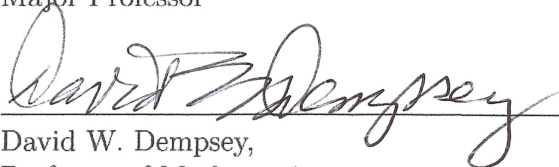
Approval:



4/19/21

Jaedeok Kim,  
Professor of Mathematics  
Major Professor

Date



4/19/21

David W. Dempsey,  
Professor of Mathematics

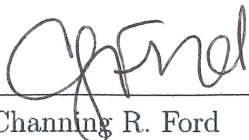
Date



4/19/21

Thomas E. Leathrum,  
Professor of Mathematics

Date



4/20/2021

Channing R. Ford  
Director of Graduate Studies

Date

STREAMING DOWN THE STERN-BROCOT TREE:  
FINDING AND EXPRESSING SOLUTIONS  
TO PELL'S EQUATION IN  $SL_2(\mathbb{Z})$

A Thesis Submitted to the  
Graduate Faculty  
of Jacksonville State University  
in Partial Fulfillment of the  
Requirements for the Degree of  
Master of Science  
with a Major in Mathematics

By

MARCUS LEE SHELL

Jacksonville, Alabama

May 7, 2021

copyright 2021  
All Rights Reserved

Marcus Lee Shell      4/27/2021  
Marcus Lee Shell      Date

## ABSTRACT

This paper explores and elaborates on a method of solving Pell's equation as introduced by Norman Wildberger. In the first chapters of the paper, foundational topics are introduced in expository style including an explanation of Pell's equation. An explanation of continued fractions and their ability to express quadratic irrationals is provided as well as a connection to the Stern-Brocot tree and a convenient means of representation for each in terms of  $2 \times 2$  matrices with integer elements. This representation will provide a useful way of navigating the Stern-Brocot tree computationally and permit us a means of computing continued fractions without the tedium of unraveling nested denominators. The paper also introduces simple unary operations for describing select permutations on continued fractions and, more importantly, their matrix-product counterparts. In the latter chapters of the paper, interesting symmetries appear as a result of using the Wildberger Algorithm. Quadratic forms and the subset of balanced quadratic forms will be shown to act as  $SL_2(\mathbb{Z})$ -sets. Using this language we explore solutions to the generalized Pell equation and demonstrate a generalization for Norman Wildberger's algorithm.

viii., 87 pages

## ACKNOWLEDGMENTS

I would like to thank the MCIS Department at JSU for always being positive and supportive.

I would like to thank Dr. Jaedeok Kim especially for his patience and his willingness to work with me. He has always encouraged and believed in me, and for that he has my gratitude.

I would like to thank both Dr. David Dempsey and Dr. Thomas Leathrum for their service on my thesis committee and for providing constructive criticism when necessary.

Also, I would like to thank Dr. Heidi Dempsey for providing her insight on some unfamiliar editorial territory.

Most of all, I would like to express my deepest thanks to my wife Fallon. She has carried more than her share of the load during my education. I could not have come this far without her care and support, for which I am most grateful.

To all of you, thank you for seeing me through my journey as mentors, family, and friends.

Marcus Lee Shell

# TABLE OF CONTENTS

	PAGE
<b>TABLE OF CONTENTS</b>	<b>vi</b>
<b>LIST OF FIGURES</b>	<b>viii</b>
<b>1 INTRODUCTION</b>	<b>1</b>
<b>2 PELL'S EQUATION</b>	<b>3</b>
<b>3 CONTINUED FRACTIONS</b>	<b>7</b>
3.1 Continued fractions for positive rational numbers . . . . .	9
3.2 Convergents and their defining recurrence relations . . . . .	10
3.3 Accessory unary operations . . . . .	13
3.4 Infinite solutions per infinite fractions . . . . .	14
<b>4 THE STERN-BROCOT TREE</b>	<b>17</b>
4.1 Farey fractions . . . . .	17
4.2 An infinite tree with only two roots . . . . .	18
4.3 Navigation: boxing it all into $SL_2(\mathbb{Z})$ . . . . .	19
<b>5 THE MAGIC OF <math>SL_2(\mathbb{Z})</math></b>	<b>27</b>
5.1 Quadratic forms as a vehicle . . . . .	29
5.2 A wild algorithm . . . . .	31
<b>6 REFLECTIONS ON SYMMETRY</b>	<b>35</b>
6.1 Unary operation revisited . . . . .	35
6.2 Types of symmetry . . . . .	39
<b>7 BUILDING OUR VEHICLE</b>	<b>42</b>
7.1 Rebuilding the engine . . . . .	42
7.2 Preparing to drive . . . . .	48
7.3 Driving in circles . . . . .	54
<b>8 CLOSING THOUGHTS</b>	<b>58</b>

<b>REFERENCES</b>	<b>61</b>
<b>APPENDICES</b>	<b>62</b>
<b>A ACCESSORY PROOFS AND DEFINITIONS</b>	<b>62</b>
A.1 Omitted proofs . . . . .	62
A.2 Unary operations and permutations . . . . .	63
<b>B AMBIGUITY OF FORM</b>	<b>66</b>
<b>C EXAMPLES</b>	<b>72</b>
C.1 Examples for appendices . . . . .	77
<b>D PROVIDED CODE</b>	<b>79</b>
D.1 Coefficients of $\sigma$ - $n$ . . . . .	79
D.2 Farey set generator . . . . .	80
D.3 Generating solutions for $-ab - b^2 = D$ . . . . .	82
D.4 The Wildberger algorithm . . . . .	84



## LIST OF FIGURES

	Page
1. An illustration of the Stern-Brocot tree .....	19
2. Left and right children for node $\frac{5}{2}$ .....	20
3. Parent nodes for $\frac{3}{8}$ in the Stern-Brocot tree .....	22
4. Cayley table for defined unary operations .....	39

# 1. INTRODUCTION

Choose a positive nonsquare integer  $d$ . We will update this language in a later chapter, but for now, let us refer to  $(a, b, c) \in \mathbb{Z}^3$  as a *step*. At each step we may go to the next step by choosing right or left. To decide, calculate the *total*:  $a + 2b + c$ . If the total is less than zero, we choose right; if the total is positive, we choose left. If our total is zero, we retrace our steps as this should not happen. Now, if we choose a right step, then we set the next step  $(a', b', c') = (a, a + b, \text{total})$ . If we choose a left step, we set the next step  $(a', b', c') = (\text{total}, b + c, c)$ . Now, we write the first step  $(a, b, c)$  and as we take steps, we record them by writing  $N(a', b', c')$  where  $N$  is the chosen direction. We will start with  $a = 1, b = 0$ , and  $c = -d$  as our first step. Then, if at any point we reach a step that is equal to the original  $(a, b, c)$ , stop. Let us see an example; the symbol  $\mathbb{W}\mathbb{A}$  will be explained later on. Let  $d = 14$ ; then

$\mathbb{W}\mathbb{A}$ :	$(1, 0, -14)$	
$R$	$(1, 1, -13)$	
$R$	$(1, 2, -10)$	
$R$	$(1, 3, -5)$	
$L$	$(2, -2, -5)$	
$R$	$(2, 0, -7)$	
$R$	$(2, 2, -5)$	
$L$	$(1, -3, -5)$	
$R$	$(1, -2, -10)$	
$R$	$(1, -1, -13)$	
$R$	$(1, 0, -14)$	<i>stop.</i>

So we have taken three right steps, one left step, then two right, one left, and finally three right steps. If the reader has followed along in the process with his or her own choice of  $d$ , then set up the obtained counts of sequential left and right decisions in a similar fashion to the product just below. Finally, compute the matrix multiplication for the resulting string of matrices.

$$\begin{bmatrix} 1 & \mathbf{3} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \mathbf{1} & 1 \end{bmatrix} \begin{bmatrix} 1 & \mathbf{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \mathbf{1} & 1 \end{bmatrix} \begin{bmatrix} 1 & \mathbf{3} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 15 & 56 \\ 4 & 15 \end{bmatrix}.$$

The reader may find it interesting that dividing the top right corner entry by the bottom left corner entry will yield 14. Even more impressive however, is the fact that the left column of the resulting matrix forms a solution to the equation

$$x^2 - 14y^2 = 1.$$

This equation is known as ‘‘Pell’s equation’’ and the fact that our little game has produced a solution to it is no trivial matter. This famous equation has been studied for centuries and is still under examination today. If one can find nonnegative integral values for  $x$  and  $y$  that satisfy the equation, then  $\frac{x}{y}$  will be a reasonable approximation to the irrational number  $\sqrt{d}$ . The accuracy of such an approximation is measurable without knowing the precise value of  $\sqrt{d}$ .

If one is compelled to try and find solutions to Pell’s equation, it becomes rather useful to develop tools with which solutions may be reliably found and evaluated. During this exposition the reader will find some elegant (and convenient!) relationships among different representations of rational numbers.

## 2. PELL'S EQUATION

Let  $d$  be any nonsquare integer. Then define the set

$$G_d = \{x + y\sqrt{d} \mid x, y \in \mathbb{Z}, x^2 - dy^2 = 1\}.$$

Let us denote solutions to the equation  $x^2 - dy^2 = 1$  as  $(x, y)$ .

**THEOREM 1.**  *$G_d$  is a group under multiplication.*

*Proof.* Let  $d$  be a positive nonsquare integer. Since elements  $x + y\sqrt{d}$  of  $G_d$  are real-valued, we know that multiplication in the ordinary sense is associative (and commutative). The only things remaining to be shown in order to demonstrate that  $G_d$  is a group are: closure of  $G_d$  under its operation, the existence of an identity element in  $G_d$ , and for every element  $G_d$  contains,  $G_d$  contains the multiplicative inverse of that element also.

Let  $(x_1, y_1)$  and  $(x_2, y_2)$  be integer solutions to the equation  $x^2 - dy^2 = 1$ . Then,  $x_1 + y_1\sqrt{d}, x_2 + y_2\sqrt{d} \in G_d$ . Let us see if the product of these two arbitrary elements is also in  $G_d$ :

$$(x_1 + y_1\sqrt{d})(x_2 + y_2\sqrt{d}) = x_1x_2 + y_1y_2d + (x_1y_2 + y_1x_2)\sqrt{d},$$

$$x_1x_2 + y_1y_2d, x_1y_2 + y_1x_2 \in \mathbb{Z} .$$

Taking these integers as potential solutions, we insert them into of  $x^2 - dy^2 = 1$  and see that

$$\begin{aligned}
& (x_1x_2 + y_1y_2d)^2 - d(x_1y_2 + y_1x_2)^2 \\
&= (x_1^2x_2^2 + 2x_1x_2y_1y_2d + y_1^2y_2^2d^2) - d(x_1^2y_2^2 + 2x_1x_2y_1y_2 + y_1^2x_2^2) \\
&= x_1^2x_2^2 + y_1^2y_2^2d^2 - dx_1^2y_2^2 - dy_1^2x_2^2 \\
&= x_1^2(x_2^2 - dy_2^2) - dy_1^2(-dy_2^2 + x_2^2) \\
&= x_1^2(1) - dy_1^2(1) = 1;
\end{aligned}$$

therefore,  $G_d$  is closed under multiplication.

Notice:

$$1^2 - d(0)^2 = 1.$$

So,  $1 + 0\sqrt{d} = 1 \in G_d$ . Therefore, since the elements of  $G_d$  are real numbers, the identity 1 of the real numbers is inherited. The multiplicative identity element  $1 + 0\sqrt{d}$  is known as the **trivial solution**:  $(1, 0)$ . To show the existence of multiplicative inverses, let us return to the original equation with a solution  $(a, b)$  substituted in for variables  $x$  and  $y$ . If  $a^2 - db^2 = 1$ , then

$$(1) \quad a^2 - db^2 = (a + b\sqrt{d})(a - b\sqrt{d}) = 1.$$

We know that if  $(a, b)$  is a solution, that  $a + b\sqrt{d} \in G_d$ , but now we should show that  $(a, -b)$  is a solution as well.

$$a^2 - b^2\sqrt{d} = a^2 - (-b)^2\sqrt{d} = 1.$$

So, for all  $a + b\sqrt{d} \in G_d$ ,  $(a + b\sqrt{d})^{-1} = a - b\sqrt{d} \in G_d$  and  $G_d$  is a group. □

Every  $x, y \in \mathbb{Z}$  such that  $x + y\sqrt{d} \in G_d$  forms a solution  $(x, y)$  to Pell's equation for  $d$ . Since  $G_d$  is a group,  $(x + y\sqrt{d})^n$  will represent a solution for Pell's equation for any  $n \in \mathbb{Z}$ . In fact, if  $(a, b)$  is the **fundamental solution**, the solution for which  $a$  and  $b$  are the smallest positive values they can be as part of a solution, then  $G_d = \langle a + b\sqrt{d} \rangle$ . Less succinctly that is,  $G_d$  is the group *generated by*  $a + b\sqrt{d}$ . Take note that for all eligible selections of  $d$ ,  $x$  may not be zero since otherwise we reach a contradiction:

$$[x^2 - dy^2 = 1 \text{ and } x = 0] \implies -dy^2 = 1 \implies d = -1/y^2 < 0.$$

Furthermore,  $x = \pm 1$  if and only if  $y = 0$ :

$$(-1)^2 - dy^2 = (1)^2 - dy^2 = 1 \iff -dy^2 = 0 \iff d = 0 \text{ or } y = 0.$$

Since  $d$  was assumed to be positive<sup>1</sup>,  $y$  must be 0.

As mentioned before, the special case of  $x = 1$  and  $y = 0$  is known as the trivial solution and is valid for all  $d$ . Also, as noted earlier,  $1 + 0\sqrt{d} = 1$  is the identity for  $G_d$ . Clearly,  $1^n = 1$  for all  $n$ , but consider a nontrivial solution  $(a, b)$ , then  $a > 1$  and  $b > 0$ . Let  $(a, b)$  be a solution in  $G_d$ . Let  $(a_n, b_n)$  be the precise element in  $G_d$ ,  $(a + b\sqrt{d})^n$ . Then,  $a_1 + b_1\sqrt{d} = (a + b\sqrt{d})^1$  and  $a_2 + b_2\sqrt{d} = (a + b\sqrt{d})^2 = (a + b\sqrt{d})(a_1 + b_1\sqrt{d}) = aa_1 + ab_1\sqrt{d} + ba_1\sqrt{d} + bb_1d = (aa_1 + bb_1d) + (ab_1 + ba_1)\sqrt{d}$ .

So,  $a_2 = aa_1 + bb_1d$  and  $b_2 = ab_1 + ba_1$ . To prove these relationships will hold for all  $n$ , assume the relationship holds for some integer  $n > 1$  and consider the  $n + 1$  case. We have  $a_{n+1} + b_{n+1}\sqrt{d} = (a + b\sqrt{d})^{n+1} = (a + b\sqrt{d})(a + b\sqrt{d})^n = (a + b\sqrt{d})(a_n + b_n\sqrt{d}) = aa_n + ab_n\sqrt{d} + ba_n\sqrt{d} + bb_nd = (aa_n + bb_nd) + (ab_n + ba_n)\sqrt{d}$ .

---

<sup>1</sup>If  $d$  is negative, the result is unchanged. We will not allow  $d = 0$  since 0 is among the perfect squares which have only the trivial solutions  $(1, 0)$  and  $(-1, 0)$ .

The inductive step complete, we have:

$$(2) \quad a_{n+1} = aa_n + bb_nd$$

and

$$(3) \quad b_{n+1} = ab_n + ba_n.$$

LEMMA 1. *Let  $a, b, d \in \mathbb{Z}^+$  and  $a^2 - db^2 = 1$ , with  $d$  nonsquare,  $a > 1$ . If  $a_n + b_n\sqrt{d} = (a + b\sqrt{d})^n$ , then  $\{b_n\}$  is a strictly increasing sequence.*

*Proof.* See Appendix A.1

Note, since  $\{b_n\}$  is strictly increasing and  $b_n$  is an integer for all  $n$ , it follows that as  $n \rightarrow \infty$ ,  $b_n \rightarrow \infty$ . So, if  $(a, b)$  is indeed a nontrivial solution, then  $(a + b\sqrt{d})^n$  will generate an infinite number of solutions  $(a_n, b_n)$ . Moreover, as  $n$  increases,  $\frac{a_n}{b_n}$  will improve as an approximation for  $\sqrt{d}$ .

In fact, David Burton in [1] provides a corollary that in our context says, if  $(a, b)$  is the fundamental solution then

$$(4) \quad \left| \frac{a_n}{b_n} - \sqrt{d} \right| \leq \frac{1}{b_n^2}.$$

We now shift our focus to the notion of “continued fractions.” For the reader interested in a more detailed introduction to Pell’s equation, [2] and [1] are recommended readings.

### 3. CONTINUED FRACTIONS

Consider the set of nonnegative rational numbers (denoted  $\mathbb{Q}_0^+$ ). No matter what numerical base one is using, at least some of these numbers will have nonterminating representations. However, there is a notation that will allow us to express any number in  $\mathbb{Q}_0^+$  as a finite array of integers. Take the number  $\frac{9}{7}$  as an example. The idea is that we would like to write  $\frac{9}{7}$  as the sum of some whole number  $a_0$  and a fraction whose denominator recursively contains a similar sum<sup>2</sup> until we reach a denominator whose value is simply a whole number. We will require each level of nested fraction to have a numerator of precisely 1. This will be clear after we perform the algorithm on our example.

$$\frac{9}{7} = 1 + \frac{2}{7} = 1 + \frac{1}{\frac{7}{2}} = 1 + \frac{1}{3 + \frac{1}{2}}.$$

Once we have reached that last 2, there is no nontrivial mixed number representation, so we stop here and write either  $1 + \frac{1}{3 + \frac{1}{2}}$  or  $1 + \frac{1}{3 + \frac{1}{1 + \frac{1}{1}}}$ . We will stick to the previous convention for the most part<sup>3</sup>. Now, let us represent our example fraction by recording the whole number part  $a_0 = 1$  as well as the whole number part of each denominator nested within the larger fraction as a list of coefficients. Then, we can write

$$\frac{9}{7} = 1 + \frac{1}{3 + \frac{1}{2}} \equiv [1; 3, 2].$$

In general, this type of fraction is known as a *simple continued fraction*.

---

<sup>2</sup>of some whole number  $a_1$  and a fraction whose denominator recursively contains a similar sum...

<sup>3</sup>See Appendix B.



DEFINITION 1. A continued fraction  $c$  is a real number constructed by a sequence of real numbers  $\{a_i\}$  where  $a_i > 0$  for  $i > 0$  in the following way:

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots + \frac{1}{a_n}}}}$$

For each  $i > 0$ ,  $a_i$  is called a “partial denominator.” If the sequence used to construct  $c$  is finite, we say  $c$  is a **finite** continued fraction. If the sequence used to construct  $c$  contains only integers, we say  $c$  is **simple**. We will refer to this sequence as the **coefficients** of  $c$ .

We will use the following conventions of notation. We may use a boldface letter to refer to a continued fraction. If  $c$  is a nonnegative rational number, and  $\mathbf{c}$  is the list of coefficients of a continued fraction of equal value to  $c$  then we may write  $c \equiv \mathbf{c}$  since the value may be equal, but there may be more than one continued fraction with the same value but different coefficients. We also wish to think of coefficients as a notation that can be processed to retrieve a value without thinking of them as a value in their own right. We will denote the coefficients of a continued fraction  $\mathbf{c}$  by  $\mathbf{c} = [a_0; a_1, a_2, \dots]$  if  $\mathbf{c}$  has infinitely many associated coefficients and  $\mathbf{c} = [a_0; a_1, \dots, a_n]$  if  $\mathbf{c}$  has a finite sequence of partial denominators.

**3.1. Continued fractions for positive rational numbers.** Let us begin with a couple of examples of continued fractions for rational numbers.

*Example 1:*

$$\frac{4}{19} = 0 + \frac{4}{19} = 0 + \frac{1}{\frac{19}{4}} = 0 + \frac{1}{4 + \frac{3}{4}} = 0 + \frac{1}{4 + \frac{1}{\frac{4}{3}}} = 0 + \frac{1}{4 + \frac{1}{1 + \frac{1}{3}}} \equiv [0; 4, 1, 3]$$

*Example 2:*

$$\frac{56}{15} = 3 + \frac{11}{15} = 3 + \frac{1}{1 + \frac{4}{11}} = 3 + \frac{1}{1 + \frac{1}{2 + \frac{3}{4}}} = 3 + \frac{1}{1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{3}}}} \equiv [3; 1, 2, 1, 3]$$

Note that this last example has a certain symmetry about it. We may find later that there is a special property related to this type of symmetry.

**THEOREM 2.** *Every rational number may be expressed as a finite simple continued fraction by using the quotients of Euclid's Algorithm.*

Rather than attempting to reinvent the wheel, we will point out the application of David Burton's proof of Theorem 15.1 in [1].

Let  $\frac{p}{q}$  be a rational number. Then, by an application of the Euclidean algorithm, we may produce a sequence of integers (which *will* terminate) in the following way:

$$\begin{aligned}
p &= qa_0 + r_1, & 0 \leq r_1 < q \\
q &= r_1a_1 + r_2, & 0 \leq r_2 < r_1 \\
r_1 &= r_2a_2 + r_3 \\
&\vdots \\
r_{i-1} &= r_ia_i + r_{i+1} \\
&\vdots \\
r_{n-1} &= r_n a_n + 0 \\
&\text{stop.}
\end{aligned}
\tag{5}$$

By taking the quotient  $a_0$  from the first expression as the whole number part and the quotients  $a_i$  of each of the remaining expressions as the partial denominators,  $c$  is expressed as a finite simple continued fraction.

Also, if a continued fraction is finite, then starting at the last partial denominator and “folding” it up into a typical mixed number, we will only find elements of  $\mathbb{Q}$ . It is interesting to note that the value of any infinite continued fraction must then be an irrational number (refer again to [1] for further reading), unlike the case of infinite decimal representation (in any base) of real numbers (e.g.  $0.\bar{3}$  in base-10).

**3.2. Convergents and their defining recurrence relations.** Now, unpacking a continued fraction can become quite tedious, and in fact, it does take a lot of paper and caution to ensure that no error is made. Fortunately, we have a way to generate  $c$  from its coefficients (if they are available). This method will again rely upon a pair of recurrence relations. Now, this pair of relations is defined and proven reliable in a similar fashion in [1], but we will provide the demonstration here for

convenience. For any rational number  $c$ , we may write  $c = \frac{p}{q} \equiv [a_0; a_1, \dots, a_k, \dots, a_n]$ . We say that  $c_k \equiv [a_0; a_1, \dots, a_k]$  is the “ $k$ th convergent” of  $c$ . Since  $c_k$  is a rational number, let  $c_k = \frac{p_k}{q_k}$  for relatively prime integers  $p_k$  and  $q_k$ .

Notice that  $c_0 = a_0 = \frac{a_0}{1} = \frac{p_0}{q_0}$ . So naturally, choose  $p_0 = a_0$  and  $q_0 = 1$ .

Next, notice that

$$c_1 = a_0 + \frac{1}{a_1} = \frac{a_0 a_1 + 1}{a_1} = \frac{p_1}{q_1}.$$

So, let  $p_1 = a_0 a_1 + 1$  and  $q_1 = a_1$ . To find hypothetical definitions for  $p_k$  and  $q_k$ ,  $k > 1$ , let us examine the  $c_2$  case.

$$\begin{aligned} c_2 &= a_0 + \frac{1}{a_1 + \frac{1}{a_2}} = a_0 + \frac{1}{\frac{a_1 a_2 + 1}{a_2}} \\ &= a_0 + \frac{a_2}{a_1 a_2 + 1} = \frac{a_0(a_1 a_2 + 1) + a_2}{a_1 a_2 + 1} \\ &= \frac{a_0 a_1 a_2 + a_0 + a_2}{a_2 a_1 + 1} = \frac{a_0 a_1 a_2 + a_2 + a_0}{a_2 a_1 + 1} \\ &= \frac{a_2(a_0 a_1 + 1) + a_0}{a_2(a_1) + 1} = \frac{a_2(p_1) + p_0}{a_2(q_1) + q_0} = \frac{p_2}{q_2}. \end{aligned}$$

Since  $p_2 = a_2 p_1 + p_0$  and  $q_2 = a_2 q_1 + q_0$ , it would be convenient if it were true in general that  $p_k = a_k p_{k-1} + p_{k-2}$  and  $q_k = a_k q_{k-1} + q_{k-2}$ , and in fact, this is the case.

Assume the relationships for  $p_k$  and  $q_k$  hold for all  $k > 1$ . Then, we know for  $c_k \equiv [a_0; a_1, \dots, a_k]$ ,

$$(6) \quad c_k = \frac{p_k}{q_k} = \frac{a_k p_{k-1} + p_{k-2}}{a_k q_{k-1} + q_{k-2}}.$$

Now consider  $c_{k+1}$ . Look at just the tail end of the fraction:

$$(7) \quad a_0 + \frac{1}{a_1 + \frac{1}{\ddots + \frac{1}{a_{k-1} + \frac{1}{a_k + \frac{1}{r}}}}}$$

where  $r = a_{k+1}$ . Then we can say  $c_{k+1}$  is equivalently the continued fraction  $c_{k+1} \equiv [a_0, a_1, \dots, a_{k-1}, (a_k + \frac{1}{r})]$ . Granted, this fraction is not simple, but this only leaves our demonstration more generalized. So consider the following:

$$\begin{aligned} c_{k+1} &= \frac{p_{k+1}}{q_{k+1}} = \frac{(a_k + \frac{1}{r})p_{k-1} + p_{k-2}}{(a_k + \frac{1}{r})q_{k-1} + q_{k-2}} \\ &= \frac{a_k p_{k-1} + \frac{p_{k-1}}{r} + p_{k-2}}{a_k q_{k-1} + \frac{q_{k-1}}{r} + q_{k-2}} = \frac{\frac{r a_k p_{k-1} + p_{k-1} + r p_{k-2}}{r}}{\frac{r a_k q_{k-1} + q_{k-1} + r q_{k-2}}{r}} \\ &= \frac{r a_k p_{k-1} + p_{k-1} + r p_{k-2}}{r a_k q_{k-1} + q_{k-1} + r q_{k-2}} = \frac{r a_k p_{k-1} + r p_{k-2} + p_{k-1}}{r a_k q_{k-1} + r q_{k-2} + q_{k-1}} \\ &= \frac{r(a_k p_{k-1} + p_{k-2}) + p_{k-1}}{r(a_k q_{k-1} + q_{k-2}) + p_{k-1}} = \frac{a_{k+1}(p_k) + p_{k-1}}{a_{k+1}(q_k) + q_{k-1}} = \frac{p_{k+1}}{q_{k+1}}, \end{aligned}$$

so by induction, we see that our relations for  $p_k$  and  $q_k$  hold for all  $k$ . Bear in mind that calculating the value of a continued fraction in this way is tedious still, but it is also much simpler to think about—especially when a computer may be used.

**3.3. Accessory unary operations.** Here we will define some unary operations that may help us in our discussion.

DEFINITION 2. Let  $\mathbf{c} \equiv c \in \mathbb{Q}$ . Then

$$\sigma(\mathbf{c}) = \sum a_i .$$

That is,  $\sigma$  will yield the sum of the coefficients for a given simple finite continued fraction.

DEFINITION 3. Let  $\mathbf{c} \equiv c \in \mathbb{Q}$ . Then  $\lambda(\mathbf{c})$  is the number of nonzero coefficients appearing in  $\mathbf{c}$ . That is,  $\lambda(\mathbf{c})$  is the number of partial denominators in the continued fraction given by the coefficients of  $\mathbf{c}$ —plus one if the whole number part is nonzero.

Example 3: Let  $\mathbf{c} = [0; 3, 2, 5]$ . Then

$$\sigma(\mathbf{c}) = 0 + 3 + 2 + 5 = 10$$

and

$$\lambda(\mathbf{c}) = 3.$$

DEFINITION 4. Let  $\mathbf{c} = [a_0; a_1, \dots, a_n]$ . If  $\mathbf{c} \equiv c \in \mathbb{Q}^+$ , then

$$\frac{1}{c} \equiv \rho(\mathbf{c}) = \begin{cases} [0; a_0, a_1, \dots, a_n] & a_0 \neq 0 \\ [a_1; \dots, a_n] & a_0 = 0. \end{cases}$$

$\rho$  may be referred to as the reciprocal of a set of coefficients, but we might also say that to perform  $\rho$  on  $\mathbf{c}$  is to “flip”  $\mathbf{c}$ .

DEFINITION 5. If  $\mathbf{c} = [a_0; a_1, \dots, a_n]$  is the list of coefficients for a finite simple continued fraction, we say that  $\bar{\mathbf{c}}$  is the **conjugate** of  $\mathbf{c}$ . Let  $rev(\mathbf{b})$  denote the reversal of the nonzero coefficients of some finite simple continued fraction  $\mathbf{b}$ . To perform conjugation on  $\mathbf{c}$ , first note of the value  $\lambda(\mathbf{c})$ , whether it be even or odd. Then,

$$\bar{\mathbf{c}} = \begin{cases} rev(\mathbf{c}) & \text{if } \lambda(\mathbf{c}) \text{ is odd} \\ rev(\rho(\mathbf{c})) & \text{if } \lambda(\mathbf{c}) \text{ is even} \end{cases}$$

Simply put, under conjugation, the nonzero coefficients of a finite simple continued fraction with an odd value under  $\lambda$  are reversed. The same is true for conjugation of similar fractions with even values under  $\lambda$  except that in this case the leading zero is “toggled” with the  $\rho$  function also.

*Example 4:* Consider  $\mathbf{b} = [4; 3, 2, 1]$ , and  $\mathbf{c} = [2; 4, 5]$ .

It is easy to see  $\rho(\bar{\mathbf{c}}) = \overline{\rho(\mathbf{c})}$  and  $\rho(\bar{\mathbf{b}}) = \overline{\rho(\mathbf{b})}$ .

$$\lambda(\mathbf{c}) = 3$$

$$\lambda(\mathbf{b}) = 4$$

$$\rho(\mathbf{c}) = [0; 2, 4, 5]$$

$$\rho(\mathbf{b}) = [0; 4, 3, 2, 1]$$

$$\bar{\mathbf{c}} = [5; 4, 2]$$

$$\bar{\mathbf{b}} = rev(\rho(\mathbf{b})) = [0; 1, 2, 3, 4]$$

$$\overline{\rho(\mathbf{c})} = [0; 5, 4, 2]$$

$$\overline{\rho(\mathbf{b})} = \overline{[0; 4, 3, 2, 1]} = rev([4; 3, 2, 1])$$

$$\rho(\bar{\mathbf{c}}) = [0; 5, 4, 2]$$

$$\rho(\bar{\mathbf{b}}) = [1; 2, 3, 4] = rev([4; 3, 2, 1])$$

**3.4. Infinite solutions per infinite fractions.** Our definition for  $\lambda$  is not defined for infinite fractions, but there is a special case of infinite fractions that do have a sort of length as is the case with repeating decimal expansions and their period. As in the case of  $0.\overline{234}$ , we can also have continued fractions of the form:

$$[a_0; \overline{a_1, a_2, \dots, a_n}].$$

In fact, the coefficients of an irrational number  $x$  will become periodic if and only if  $x$  is a quadratic irrational number<sup>4</sup>. However, for now we need only examine a special case of fractions demonstrating periodicity. It is a commonly demonstrated fact that if  $d$  is a nonsquare integer, then the coefficients generated by  $\sqrt{d}$  will have the special form:

$$\sqrt{d} \equiv [a_0; \overline{a_1, a_2, \dots, a_2, a_1, 2a_0}]$$

The allure of algebraically valued numbers with this nearlypalindromic nature is that they are proven to generate solutions to Pell's equation. Since we have seen that one solution expressed in the language of  $G_d$  can be exponentiated to yield other solutions, it would seem only natural that we see something of that nature here and indeed we do. If the coefficients of  $\sqrt{d}$  have a period length of  $n$ , then the convergents  $\frac{p_{kn}}{q_{kn}}$  will form solutions  $(p_{kn-1}, q_{kn-1})$  to the equation  $x^2 - dy^2 = \pm(-1)^{kn}$  for all  $k \in \mathbb{N}$ .

*Example 5:* Consider the cases of  $\sqrt{77}$  and  $\sqrt{58}$ .

For  $\sqrt{77} \equiv [8; \overline{1, 3, 2, 3, 1, 16}]$ , we see a period length of 6 and we have

$$\frac{p_5}{q_5} = 8 + \frac{1}{1 + \frac{1}{3 + \frac{1}{2 + \frac{1}{3 + \frac{1}{1}}}}} = \frac{351}{40}. \text{ Notice that } 351^2 - 77(40)^2 = 1.$$

There is no solution to the negative Pell equation  $x^2 - 77y^2 = -1$ .

Now, for  $\sqrt{58} \equiv [7; \overline{1, 1, 1, 1, 1, 1, 14}]$  we see a period length of 7 and we have

$$\frac{p_6}{q_6} = 7 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1}}}}} = \frac{99}{13}. \text{ Notice that } 99^2 - 58(13)^2 = -1. \text{ While this calculation can}$$

be verified, we will omit the continued fraction expansion for:  $\frac{p_{13}}{q_{13}} = \frac{19603}{2574}$ . Assuming its correctness, one can easily verify that  $19603^2 - 58(2574)^2 = 1$  by using a calculator.

In this manner, solutions to Pell's equations have been deduced, but the process of generating the coefficients and proceeding to calculate their rational value can be lengthy. However, we will soon explore a method of solving Pell's equation

---

<sup>4</sup>I.e.  $x$  is irrational and the root of an irreducible degree two polynomial with rational coefficients.



that skirts the tedium of using Euclid's algorithm, the floor function, or the need to manually generate coefficients for  $\sqrt{d}$ . However, the periodicity of coefficients will be integral to our approach (pardon the pun).

## 4. THE STERN-BROCOT TREE

The Stern-Brocot tree is a dressed up binary tree that allows one to navigate to any rational number in a binary search style. The most beautiful feature of the tree is that it is quite simple to generate algorithmically and employs a special set of numbers to do so. These fractions are known as “Farey fractions.”

### 4.1. Farey fractions.

DEFINITION 6. Let  $a, b, c$ , and  $d$  be positive integers. For fractions  $\frac{a}{b}$  and  $\frac{c}{d}$ , we call the fraction  $\frac{a+c}{b+d}$  their **mediant**. It can be shown that the mediant of two fractions falls between them on the real number line.

LEMMA 2. Let  $a, b, c, d \in \mathbb{Z}^+$ . If  $\frac{a}{b} < \frac{c}{d}$ , then  $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$ .

*Proof.* Let  $\frac{a}{b} < \frac{c}{d}$ . Then

$$\frac{a}{b} < \frac{c}{d} \Leftrightarrow ad < bc \Leftrightarrow ad + dc < bc + dc \Leftrightarrow d(a+c) < c(b+d) \Leftrightarrow \frac{a+c}{b+d} < \frac{c}{d}.$$

Also,

$$\frac{a}{b} < \frac{c}{d} \Leftrightarrow ad < bc \Leftrightarrow ad + ab < bc + ab \Leftrightarrow a(d+b) < b(a+c) \Leftrightarrow \frac{a}{b} < \frac{a+c}{b+d}.$$

Therefore,  $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$ . □

Now consider the ordered set of rational numbers  $F_1 = \{\frac{0}{1}, \frac{1}{1}\}$ . Define  $F'_i$  to be the collection of mediants of each adjacent pair of fractions in  $F_i$ . Then, for all  $i \geq 1$ , let  $F_{i+1} = F_i \cup F'_i$ , whose elements are ordered by their position on the real number

line. For each  $n \geq 1$ ,  $F_n$  is called a *Farey sequence*. The elements of these sequences known as **Farey fractions** have some interesting properties, not the least of which is that for any two adjacent Farey fractions  $\frac{a}{b}$  and  $\frac{c}{d}$  in some sequence  $F_n$  with  $\frac{a}{b} < \frac{c}{d}$ ,  $bc - ad$  will be equal to 1. Furthermore, each Farey fraction (ignoring  $\frac{0}{1}$  and  $\frac{1}{1}$  which are not proper fractions), will be in least terms.

*Example 6:*

$$F_1 = \left\{ \frac{0}{1}, \frac{1}{1} \right\}; \quad F'_1 = \left\{ \frac{1}{2} \right\}$$

$$F_2 = F_1 \cup F'_1 = \left\{ \frac{0}{1}, \frac{1}{2}, \frac{1}{1} \right\}; \quad F'_2 = \left\{ \frac{1}{3}, \frac{2}{3} \right\}$$

$$F_3 = F_2 \cup F'_2 = \left\{ \frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1} \right\}; \quad F'_3 = \left\{ \frac{1}{4}, \frac{2}{5}, \frac{3}{5}, \frac{3}{4} \right\}$$

Some Python code has been provided in Appendix D.2 that will allow for the generation of Farey sequences.

**4.2. An infinite tree with only two roots.** Now that we have a canonical understanding of Farey fractions, let us extend their definitions to include all nonnegative rational numbers. Rather than filling the interval  $[0, 1]$ , we desire to cover the interval  $[0, \infty)$  or, adopting the idea of the extended real numbers,  $[\frac{0}{1}, \frac{1}{0}]$ . Now, the endpoints here must not be viewed as numbers in the usual sense, as this creates a singular problem. Indeed, we could select  $\frac{2}{2}$  or  $\frac{7}{0}$  to respectively achieve the same values for our endpoints, but the mediants of these ordered pairs would not be the same. So we choose the fractions as noted just now with precision. Define  $F_1 = \{\frac{0}{1}, \frac{1}{0}\}$ . Then retain the original definitions of  $F'_i$  and  $F_n$ ,  $1 \leq i$  and  $1 < n$ .

Arranging the elements of  $F_n$  in rows, starting at the top with  $n = 2$ , each fraction of any given row will contribute to creating precisely 2 mediant fractions in the following row. By connecting each fraction to each of its two children, we will

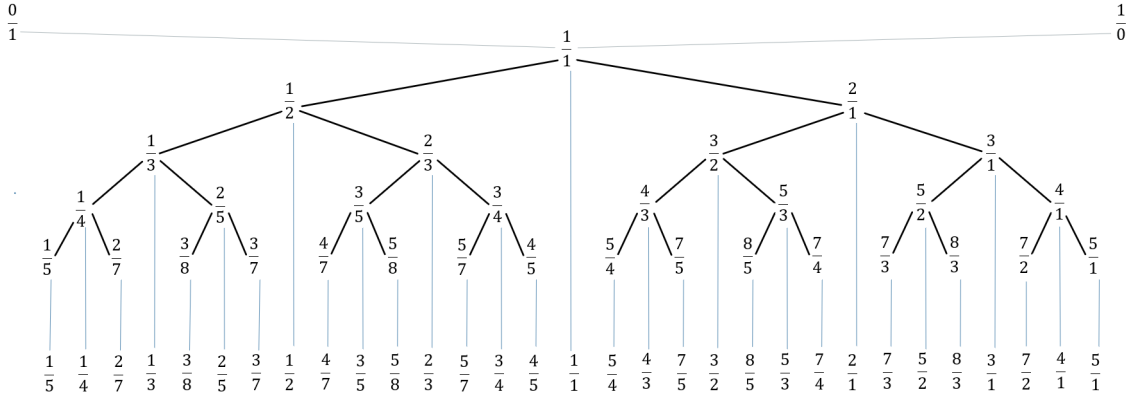


FIG. 1. An illustration of the Stern-Brocot tree

construct the **Stern-Brocot tree**.

Some interesting notes:

- If we use our extended definition for Farey fractions ,  $c \in (0, \infty), c \equiv \mathbf{c}$  if and only if  $c \in F'_{\sigma(\mathbf{c})}$ .
- For any positive integer  $n$  there are precisely  $2^n$  many  $c \in \mathbb{Q}^+$  with  $\sigma(\mathbf{c}) = n$ .
- Each row of the Stern-Brocot tree of index  $n$  is constructed with precisely the elements of  $F'_{n-1}$ ; therefore, each row is the collection of fractions with a sigma value of  $n - 1$ .

This idea can be demonstrated using a Python script provided in Appendix D.

**4.3. Navigation: boxing it all into  $SL_2(\mathbf{Z})$ .** We would like a language in which we can discuss path traversal in the tree. It is a pleasure to introduce the framework in which we will work for the remainder of the paper. The reader may recall the observation that each  $q \in \mathbb{Q}$  has two different representations as a continued fraction. We will see that each of these representations points to one of two nodes on the Stern-Brocot tree, and they share a parent  $q$ . These children are the mediants of  $q$  and its neighbors in the Farey sequence in which  $q$  first appears. Adding one to

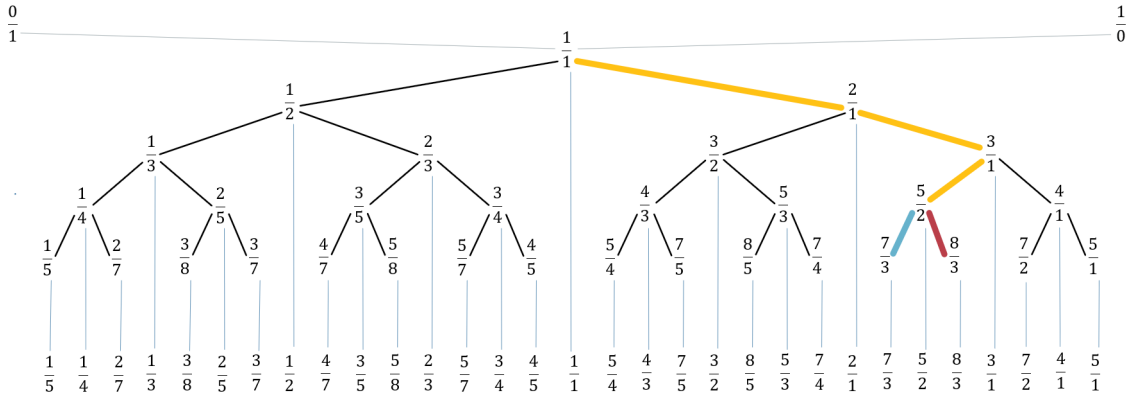


FIG. 2. Left and right children for node  $\frac{5}{2}$

$a_n$  would continue in the last direction traveled along the tree, while starting a new coefficient  $a_{n+1} = 1$  would equivocate changing direction.

*Example 7:* Navigate to  $\frac{5}{2}$  on the tree. Notice that  $\frac{5}{2}$  has two forms as a continued fraction:  $[2; 2]$  and  $[2; 1, 1]$ . The coefficients can read as “Go right 2 and then left 2,” or taking the latter set of coefficients, “Go right 2, left 1, and then right 1.” In either case, we would travel through  $\frac{5}{2}$  to get to one of its two child nodes in the succeeding Farey Sequence.

If this is still confusing, the reader may find it useful to ask: “Why does each node in a particular level of the tree have the same value under  $\sigma$ ?” This is an enlightening premise once understood and it will become easier to see as we introduce a most elegant notation using  $2 \times 2$  matrices. However, we shall point out (though it is easily guessed) that we can navigate to *any* positive rational number on the tree by starting at 1 and going left or right at any node based on whether our current position is too great or too small. If our position is greater than the desired number, we choose left. Otherwise, we choose right. So, if one wanted to navigate to  $\frac{4}{7}$ , one would begin at  $\frac{1}{1}$  and choose left, right, left, left.

We would like to express the rational numbers on the Stern-Brocot tree as  $2 \times 2$  matrices. We know that the mediant of a pair of neighbors in the Farey sequence is uniquely expressed. So, we can uniquely represent a fraction with our matrices by placing the left parent in the the right column vector position and the right parent in the left column vector position. So, for the parents  $0/1$  and  $1/0$ , we would have the *identity matrix!* For parents  $\frac{2}{3}$  and  $\frac{3}{4}$  we would have  $\begin{bmatrix} 3 & 2 \\ 4 & 3 \end{bmatrix}$ . Let us look at some other examples by referring to Figure 2. Every node has two parents. One is the previous node with a line drawn upward. The other parent will be the node indicated by the faint vertical line descending as that is the “nearest” (as appearing on the number line) of the nodes traversed so far, that is opposite the connected parent.

*Example 8:* Since  $\frac{2}{1}$  has parents  $\frac{1}{1}$  and  $\frac{1}{0}$ , we will make  $(1, 0)^T$  the left column and  $(1, 1)^T$  the right column. To the right of  $\frac{1}{1}$ , we have:

$$\frac{2}{1} \equiv \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

On the left of  $\frac{1}{1}$ , we have  $\frac{1}{2}$  :

$$\frac{1}{2} \equiv \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Now, before dismissing this example as trivial, please make note that these two matrices are very dear to us—so much so that they bear a pair of special names:

$$R = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

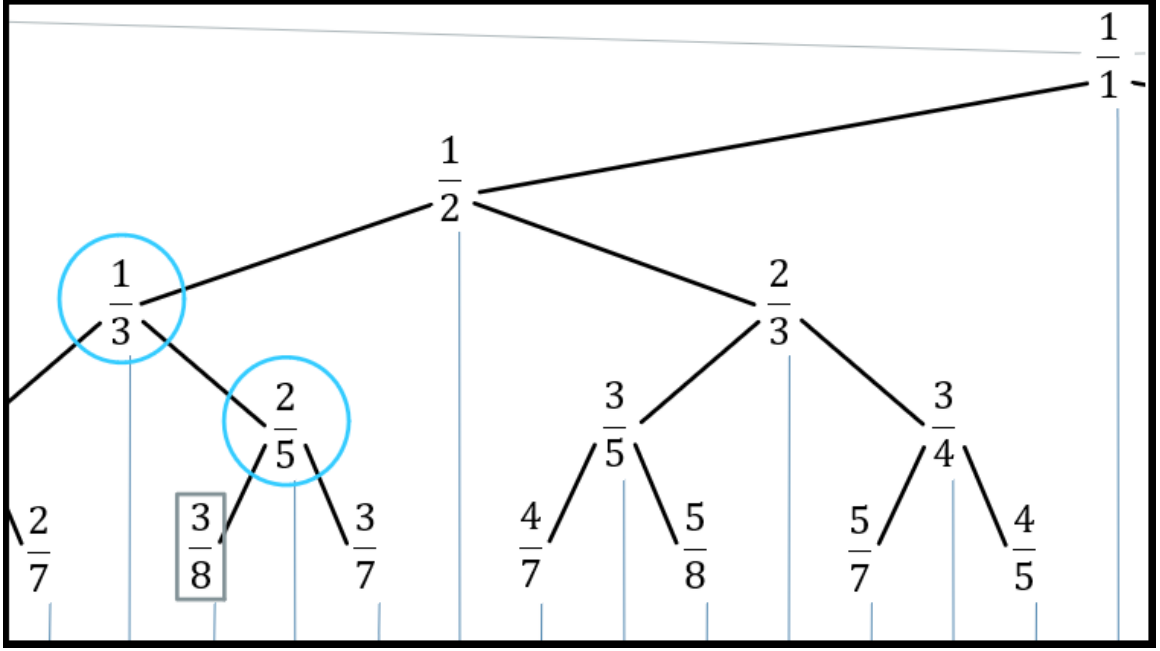


FIG. 3.  $\frac{3}{8}$  has parents  $\frac{1}{3}$  and  $\frac{2}{5}$

$$L = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

The names  $L$  and  $R$  were adopted after their use in [4] within the section entitled *Relative Primality*.

*Example 9:* Since  $\frac{3}{8}$  has the connected parent  $\frac{2}{5}$  to the right side, place  $(2, 5)^T$  in the *left* column, and then, since the nearest left-hand-side fraction is  $\frac{1}{3}$ , we put  $(1, 3)^T$  on the right.

$$\frac{3}{8} \equiv \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$$

In both examples, our matrices have determinant 1, and in fact, this will be the case throughout. By swapping the order of the parents in our placing them as column

vectors in matrices, we guarantee this lovely property since for Farey neighbors  $\frac{a}{b} < \frac{c}{d}$  we have  $bc - ad = 1$ . The most splendid part in all this is each matrix we have mentioned is in  $SL_2(\mathbb{Z})$  which, as it turns out, is generated by the special matrices  $L$  and  $R$  mentioned in Example 8. We employ the adjective *splendid* because of the elegant and simple properties belonging to these generators. Why, the fact that each is the transpose to the other is enough to make one sigh, but nicer still are the following facts, which are left as an easy exercise for the reader:

$$R^n = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$$

$$L^m = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^m = \begin{bmatrix} 1 & 0 \\ m & 1 \end{bmatrix}$$

It cannot be overstated how convenient this is for our purposes.

Let  $\mathbf{c} = [a_0; a_1, \dots, a_n]$ , and let  $c_k = \frac{p_k}{q_k}$  be the  $k$ th convergent for simple continued fraction  $\mathbf{c}$ . Consider the following:

$$(8) \quad R^{a_0} L^{a_1} = \begin{bmatrix} 1 & a_0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a_1 & 1 \end{bmatrix} = \begin{bmatrix} a_0 a_1 + 1 & a_0 + 0 \\ a_1 + 0 & 1 + 0 \end{bmatrix} = \begin{bmatrix} p_1 & p_0 \\ q_1 & q_0 \end{bmatrix}$$

So, the column vectors of  $R^{a_0} L^{a_1}$  are precisely the  $0_{th}$  and  $1_{st}$  convergents of  $c$ .

$$(9) \quad (R^{a_0} L^{a_1}) R^{a_2} = \begin{bmatrix} p_1 & p_0 \\ q_1 & q_0 \end{bmatrix} \begin{bmatrix} 1 & a_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} p_1 & a_2 p_1 + p_0 \\ q_1 & a_2 q_1 + q_0 \end{bmatrix} = \begin{bmatrix} p_2 & p_1 \\ q_2 & q_1 \end{bmatrix}$$



We will soon prove that the relationship demonstrated above will hold, but first we need another definition.

DEFINITION 7. Let  $\mathbf{C} := \{[a_0; a_1, \dots, a_n] \mid n \geq 0; \forall i > 0, a_i \in \mathbb{Z}^+\}$ . Then, let  $h : \mathbf{C} \rightarrow SL_2(\mathbb{Z})$  such that if  $\mathbf{c} \in \mathbf{C}$ , then

$$h(\mathbf{c}) = \prod_{i=0}^n M_i^{a_i}, \text{ where } M_i = \begin{cases} R & \text{if } i \text{ is even} \\ L & \text{if } i \text{ is odd.} \end{cases}$$

We may refer to such products as **strings** or **words** in the alphabet  $\{\mathbf{L}, \mathbf{R}\}$ .

If we refer to such a product as a word, we will assume that the product has been simplified such that all adjacent factors of the same base have been combined by adding their exponents. For example,  $R^0 L R R L L^2 R^5$  would become  $LR^2 L^3 R^5$ . If we need to specify, this simplification will be called the **abbreviated form** of the word in question. We will assume that all words have only nonnegative exponents for our purposes as indicated by the restriction of coefficients in the preimage of  $h$  to having only nonnegative values. This implies that all matrices generated in this way will have only positive entries and will be an important feature later on.

*Example 10:* Let  $\mathbf{a} = [2; 3, 5]$ ,  $\mathbf{b} = [0; 3, 4]$ , and  $\mathbf{c} = [2; 2, 5, 1]$ .

$$h(\mathbf{a}) = R^2 L^3 R^5$$

$$h(\mathbf{b}) = R^0 L^3 R^4 = L^3 R^4$$

$$h(\mathbf{c}) = R^2 L^2 R^5 L$$

We map from the set of appropriate coefficients for finite simple continued fractions rather than from  $\mathbb{Q}^+$  because it is more meaningful to do so. From here

on, we will shift our attention to the convergents generated by quadratic irrationals; furthermore, we need not worry about the interpretation of the matrices as rational numbers. That is, we will use the convergents of quadratic irrationals to generate coefficients, but after doing so, we will be less concerned with their origin and more concerned with their interpretation as exponents for producing products in  $SL_2(\mathbb{Z})$ ; therefore make note of the following convenient result.

**THEOREM 3.** *Let  $\mathbf{c} = [a_0; a_1, \dots, a_n]$ , and let  $c_k = \frac{p_k}{q_k}$  be the  $k$ th convergent for simple continued fraction  $\mathbf{c}$ . Then, for all  $n > 0$ , if  $n$  is odd, then*

$$h(\mathbf{c}) = \begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix},$$

and if  $n$  is even, then

$$h(\mathbf{c}) = \begin{bmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{bmatrix}.$$

*Proof.* We have already shown in equations (8) and (9) that this theorem holds when  $n = 1$  and  $n = 2$  respectively. So, we may prove the theorem inductively if we assume it to be true for all  $n$  and prove it for the even and odd cases for  $n + 1$ .

Let  $n$  be even and naturally  $n + 1$  will be odd.

Then,  $h([a_0; a_1, \dots, a_{n+1}]) = R^{a_0} L^{a_1} \dots L^{a_{n-1}} R^{a_n} L^{a_{n+1}} = h([a_0; a_1, \dots, a_n]) L^{a_{n+1}}$

$$= \begin{bmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a_{n+1} & 1 \end{bmatrix} = \begin{bmatrix} a_{n+1}p_n + p_{n-1} & p_n \\ a_{n+1}q_n + q_{n-1} & q_n \end{bmatrix} = \begin{bmatrix} p_{n+1} & p_n \\ q_{n+1} & q_n \end{bmatrix}.$$

Next, let  $n$  be odd, so that  $n + 1$  will be even.

Then,  $h([a_0; a_1, \dots, a_{n+1}]) = R^{a_0} L^{a_1} \dots R^{a_{n-1}} L^{a_n} R^{a_{n+1}} = h([a_0; a_1, \dots, a_n]) R^{a_{n+1}}$

$$= \begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix} \begin{bmatrix} 1 & a_{n+1} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} p_n & a_{n+1}p_n + p_{n-1} \\ q_n & a_{n+1}q_n + q_{n-1} \end{bmatrix} = \begin{bmatrix} p_n & p_{n+1} \\ q_n & q_{n+1} \end{bmatrix}.$$

Thus, by induction, the theorem holds for all  $n > 0$ . □

One immediate thing to ponder is that incrementing the sigma value of a continued fraction by 1 is essentially to move down the Stern-Brocot tree one level by choosing to go left or right one step. By incrementing the lambda value, we indicate that our current choice *differs* from the previous one; by choosing the same direction twice in a row, we increase the final coefficient rather than adding a new coefficient to the end. So, the larger  $\sigma(\mathbf{c})$  is, the further down the tree we will go. The larger  $\lambda(\mathbf{c})$  is, the more turns we will take. For this reason, one might refer to  $\lambda(\mathbf{c})/\sigma(\mathbf{c})$  as the “indecisiveness” of  $c$ . The other unary operations mentioned before also have meaning in our  $SL_2(\mathbb{Z})$  representation. We will explain these in a later chapter.

## 5. THE MAGIC OF $SL_2(\mathbb{Z})$

$SL_2(\mathbb{Z})$  is a famous group that can be generated by the matrices  $L$  and  $R$ , as we mentioned above. The fact that  $SL_2(\mathbb{Z})$  is a group whose elements each have determinant equal to one is precisely what we need in order to do meaningful work regarding Pell's equation. Contemplate this: If we have a matrix in  $SL_2(\mathbb{Z})$  of the form

$$(10) \quad \begin{bmatrix} a & bd \\ b & a \end{bmatrix},$$

then the determinant of this matrix will give us the equation

$$(11) \quad \begin{vmatrix} a & bd \\ b & a \end{vmatrix} = a^2 - db^2 = 1.$$

And thus,  $(a, b)$  forms a solution to Pell's equation for  $d$ .

**DEFINITION 8.** *Let  $d$  be a positive nonsquare integer. Define  $H_d$  to be the subset of  $SL_2(\mathbb{Z})$  whose elements are of the form:  $\begin{bmatrix} a & bd \\ b & a \end{bmatrix}$ . That is,*

$$\left\{ \begin{bmatrix} a & bd \\ b & a \end{bmatrix} \mid a^2 - db^2 = 1 \right\}.$$

THEOREM 4. ( $G_d \cong H_d$ ) Let  $d$  be a nonsquare positive integer. Then, the set  $H_d$  equipped with the ordinary matrix multiplication is isomorphic to the multiplicative group  $G_d$ .

*Proof.* Let  $\psi : G_d \rightarrow H_d$  be given by  $\psi(a+b\sqrt{d}) = \begin{bmatrix} a & bd \\ b & a \end{bmatrix}$ . First, we show  $\psi$  to be injective and surjective. To show injectivity, suppose  $\psi(a+b\sqrt{d}) = \psi(u+v\sqrt{d})$  and therefore,  $\begin{bmatrix} a & bd \\ b & a \end{bmatrix} = \begin{bmatrix} u & vd \\ v & u \end{bmatrix}$ . This can only be if  $a = u$  and  $b = v$ , and so,  $a + b\sqrt{d} = u + v\sqrt{d}$ . Thus,  $\psi$  is injective. Surjectivity comes naturally as  $\begin{bmatrix} a & bd \\ b & a \end{bmatrix} \in H_d$  if and only if  $a^2 - db^2 = 1$ , which is precisely the only trait required to show  $a + b\sqrt{d} \in G_d$ . So,  $\psi$  is surjective also.

Lastly, we show that  $\psi$  is a homomorphism.

$$\begin{aligned} \psi(a + b\sqrt{d})\psi(u + v\sqrt{d}) &= \begin{bmatrix} a & bd \\ b & a \end{bmatrix} \begin{bmatrix} u & vd \\ v & u \end{bmatrix} = \begin{bmatrix} au + bvd & (bu + av)d \\ bu + av & au + bvd \end{bmatrix} = \\ &= \psi((au + bvd) + (bu + av)\sqrt{d}) = \psi(au + bu\sqrt{d} + av\sqrt{d} + bvd) = \\ &= \psi((a + b\sqrt{d})u + (a + b\sqrt{d})v\sqrt{d}) = \psi((a + b\sqrt{d})(u + v\sqrt{d})). \end{aligned}$$

So,  $\psi$  is a homomorphism as well. So,  $\psi$  is an isomorphism for  $G_d$  with  $H_d$ .  $\square$

Now, since finding any solution represented in one of these groups means that we can generate infinitely many of them, we just need a way to find the first one outside of knowing it is in the set already. What we would like is to find a way to generate a solution in a finite number of steps guaranteed. That is precisely what we will do next. As mentioned before, we can navigate the Stern-Brocot tree in pursuit of a particular rational number and arrive after a finite number of correct steps if at

each step we choose left if our target is less than our current node and choose right if our target is greater. We will do a similar thing using  $R$  and  $L$ ; since  $H_d \subset SL_2(\mathbb{Z})$ , we can generate elements of  $H_d$  using only  $L$  and  $R$ .

### 5.1. Quadratic forms as a vehicle.

DEFINITION 9. Let  $Q : \mathbb{R}^2 \rightarrow \mathbb{R}$  be given by  $Q(x, y) = ax^2 + 2bxy + cy^2$  where  $a, b, c \in \mathbb{Z}$ . We say that  $Q$  is a quadratic form. We will sometimes refer to a quadratic form by its coefficients (in the canonical polynomial sense of the term) in this manner:

$$Q \equiv (a, b, c)$$

We can also write quadratic forms in a more convenient way using matrices.

Let  $Q \equiv (a, b, c)$ . Let  $(x, y)^T = \mathbf{v}$ .

$$\begin{aligned} Q(x, y) &= ax^2 + 2bxy + cy^2 = ax^2 + bxy + bxy + cy^2 = \\ (12) \quad &= (ax + by, bx + cy)(x, y)^T = \mathbf{v}^T \begin{bmatrix} a & b \\ b & c \end{bmatrix} \mathbf{v} = Q(\mathbf{v}). \end{aligned}$$

If  $Q \equiv (a, b, c)$ , then we will call  $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$  the matrix of  $Q$ , and  $\begin{vmatrix} a & b \\ b & c \end{vmatrix}$  the determinant of  $Q$ . We will refer to  $Q$  as *balanced* whenever  $a > 0$  and  $c < 0$ .

DEFINITION 10. Let  $M$  be an invertible  $2 \times 2$  matrix with integer entries. Then,  $M^{-1}$  will also have integer entries, so that the determinant of each must be  $\pm 1$ . If  $Q(\mathbf{v}) = \mathbf{v}^T A \mathbf{v}$  and  $A' = M^T A M$ , then we say  $Q'(\mathbf{v}) = \mathbf{v}^T A' \mathbf{v}$  and  $Q$  are equivalent forms.

And why not, since  $M\mathbf{v}$  is just a vector in the domain of  $Q$  and

$$Q'(\mathbf{v}) = \mathbf{v}^T A' \mathbf{v} = \mathbf{v}^T (M^T A M) \mathbf{v} = (\mathbf{v}^T M^T) A (M \mathbf{v}) = (M \mathbf{v})^T A (M \mathbf{v}) = Q(M \mathbf{v}).$$

Also of note, since  $\det M = \pm 1$ , we must have  $|A'| = |M^T A M| = |M| |A| |M| = |A|$ . This is actually a particularly important note. The matrix of any quadratic form will be symmetric, so let  $A'$  be some symmetric matrix  $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ . Since  $L$  and  $R$  both fit the requirements for the matrix  $M$ , we know that  $Q(L\mathbf{v})$  and  $Q(R\mathbf{v})$  will be equivalent forms for  $Q$ . Casually speaking, this is to set the matrix of some form  $Q'$  to be one of  $A' = R^T A R = L A R$  (called a **Right Step**) or  $B' = L^T A L = R A L$  (called a **Left Step**). Moreover, we can iteratively devise equivalent forms by taking  $Q^{(j)}(M_j \mathbf{v})$  at each step where  $M_j$  is one of  $L$  or  $R$  and  $Q^{(j)}$  is the current form.

Because  $R$  and  $L$  are generators for  $SL_2(\mathbb{Z})$ , we can create a quadratic form  $Q'(\mathbf{v}) = Q(N\mathbf{v})$  for any  $N$  in  $SL_2(\mathbb{Z})$ . And, we can do so constructively if we know how to deduce the sequence of left and right steps to get there, although it is likely that what we are after is not a known matrix. Indeed, what we desire is to use this framework to find  $N \in SL_2(\mathbb{Z})$  for which  $Q(N\mathbf{v}) = 1$  because in doing so, we find that the left column of  $N$  forms a solution to Pell's equation! Thanks to the work of Norman Wildberger, we can do just this. However, first we need to decide what is an appropriate choice of  $A$  to seed the algorithm we wish to introduce; we would like any and every step of the way to allow us to evaluate potential solutions to Pell's equation, so why not start there.

**DEFINITION 11.** *Let  $Q \equiv (1, 0, -d)$  where  $d$  is a nonsquare integer. Then,  $Q(x, y) = x^2 - dy^2$ . This special quadratic form denoted  $Q_d$  will be referred to as the Pell quadratic form.*

The strategy we will employ is actually quite clever. Suppose we search for and find  $N$  such that  $N^T AN = A$ ? In this case, we already know that if  $A$  is the matrix of  $Q$ , then using the trivial solution  $e = (1, 0)^T$ , we see that  $Q(1, 0)$  will be 1. But also if  $A' = N^T AN$  is the matrix of  $Q'$ , then  $Q'(e) = Q(Ne) = 1$  as well.

Let  $N = \begin{bmatrix} q & r \\ s & t \end{bmatrix}$  be such a matrix so that  $Ne = (q, s)^T$  and  $Q(Ne) = Q(q, s) = 1$ . Then, we see that the left column of  $N$  will in fact be a solution to Pell's equation. Then, since  $N^T AN = A$ , it is also true that  $(N^2)^T AN^2 = A$ . This means we may repeat the process starting with our latest  $Q'$  each time and exponentiate  $N$  in doing so. In this manner we will continue to generate these forms. So, we will search for such  $N$ , and in doing so, we will see that there is a way to guarantee that we will not search amiss.

**5.2. A wild algorithm.** We will now summarize the algorithm introduced by Norman Wildberger in his article *Pell's Equation without Irrational Numbers* [5] without proof; however, we will formally define, prove and elaborate on some of these statements in a later chapter.

We have chosen to start with Pell's quadratic form  $Q_d \equiv (1, 0, -d)$ , which is balanced. We will take *left* and *right steps*, creating equivalent forms for  $Q_d$  all along the way. To do this, we will ensure that each step yields a balanced form, which is not automatic. For any given quadratic form  $Q \equiv (a, b, c)$ , left and right steps follow these patterns:

*Left Step:*

$$(13) \quad L^T AL = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} a + 2b + c & b + c \\ b + c & c \end{bmatrix}$$



*Right Step:*

$$(14) \quad L^T AL = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & a+b \\ a+b & a+2b+c \end{bmatrix}$$

We will denote these steps by listing the previous form, an “L” or an “R” for whichever step we took, and then the new form obtained like this:

$$(a, b, c)L(a + 2b + c, b + c, c)$$

for a left step and

$$(a, b, c)R(a, a + b, a + 2b + c)$$

for a right step.

We will call the sum of all entries of a form  $Q$ 's matrix the **total** and denote it  $T$ . At each step,  $T$  will either be greater or less than 0.  $T$  will never equal zero because the determinant of  $(a, b, c)$  must be that of  $(1, 0, -d)$  which is precisely  $-d$ . If  $a + 2b + c$  were to be zero, then we would reach a contradiction since

$$-d = ac - b^2 = ac + (2bc + c^2) - b^2 - (2bc + c^2) = (a + 2b + c)c - (b^2 + 2bc + c^2) = -(b + c)^2$$

(recall that  $d$  is a nonsquare integer).

After taking a step, let  $(a', b', c')$  be the newly obtained equivalent form. At each step, we want a balanced form; therefore, we would like  $a'$  to be positive and  $c'$  to be negative. So, if  $T = a + 2b + c < 0$ , notice that a right step would set  $c' = a + 2b + c < 0$  and  $a' = a > 0$  ( $a > 0$  since we started with a balanced form). Then, similarly, if  $T > 0$ , a left step would put  $a' = a + 2b + c > 0$  and  $c' = c < 0$ . But now that we have our rule for taking steps, we would like to know why this is helpful.

At each step, we must have  $\det N^T AN = ac - b^2 = -d$  where  $N$  is our cumulative word built from the alphabet  $\{L, R\}$ . Since  $ac < 0$ , we will have only finitely many solutions to the equation  $ac - b^2 = -d$ <sup>5</sup>

With there being only a finite number of forms for us to traverse, we will certainly loop back to one we have encountered previously. Once this happens, the state will provide identical criteria with which we decided to go left or right at that step before. Once we reach such a familiar place, this will be our signal to stop; otherwise, the algorithm will loop continually through the same decision sequence. The first of these familiar forms will indeed be  $(1, 0, -d)$ . At any step, we can check to see what type of step was required to arrive there in a similar fashion to how we decided what step to take next. That is, we may calculate  $a' - 2b' + c'$  and check if it be positive or negative. Since  $a' - 2b' + c' = (a + 2b + c) - 2(b + c) + c = a > 0$  if a left step was taken, and  $a' - 2b' + c' = a - 2(a + b) + a + 2b + c = c < 0$  in the event of a right step, we may simply determine if  $a' - 2b' + c'$  is greater or less than 0.

Since we have both a finite sequence of possible steps we can take, and since at each step we can deduce what step was taken to get there, we can also trace the path back to the starting point. Once we have found  $N$ , we know we can iterate through an infinite number of solutions since products of  $R$  and  $L$  have only positive entries. This is easily seen if one refers back to the section on navigation through the Stern-Brocot tree using  $R$  and  $L$ . And this makes sense as our resulting  $N$  will be an element of  $H_d$  and a second iteration will result in an identical string of  $L$ s and  $R$ s. The entries of our matrices will only continue to grow and, therefore, we will never return to the same matrix  $N$  in all our iterations.<sup>6</sup>

---

<sup>5</sup>An algorithm, written in Python, that returns all the solutions to  $ac - b^2 = -d$  as quadratic forms can be found in Appendix D.

<sup>6</sup>See Appendix D.4 for a Python program that will perform this algorithm for any given balanced quadratic form.

Returning to the example provided in the introduction, we can see an interesting pattern appear. If the reader will only take the steps as indicated in the second column, he or she will find that the computation of each form is done by the matrix multiplication and the original shortcut we introduced is equivalent.

*Example 11:* Let  $d = 14$ .

$$\mathbb{W}\mathbb{A} : (1, 0, -14) \quad Q(\mathbf{v})$$

$$R(1, 1, -13) \quad Q^{(1)}(\mathbf{v}) = Q(R\mathbf{v})$$

$$R(1, 2, -10) \quad Q^{(2)}(\mathbf{v}) = Q^{(1)}(R\mathbf{v})$$

$$R(1, 3, -5) \quad Q^{(3)}(\mathbf{v}) = Q^{(2)}(R\mathbf{v})$$

$$L(2, -2, -5) \quad Q^{(4)}(\mathbf{v}) = Q^{(3)}(L\mathbf{v})$$

$$R(2, 0, -7) \quad Q^{(5)}(\mathbf{v}) = Q^{(4)}(R\mathbf{v})$$

$$R(2, 2, -5) \quad Q^{(6)}(\mathbf{v}) = Q^{(5)}(R\mathbf{v})$$

$$L(1, -3, -5) \quad Q^{(7)}(\mathbf{v}) = Q^{(6)}(L\mathbf{v})$$

$$R(1, -2, -10) \quad Q^{(8)}(\mathbf{v}) = Q^{(7)}(R\mathbf{v})$$

$$R(1, -1, -13) \quad Q^{(9)}(\mathbf{v}) = Q^{(8)}(R\mathbf{v})$$

$$R(1, 0, -14) \quad Q^{(10)}(\mathbf{v}) = Q^{(9)}(R\mathbf{v}) = Q(N\mathbf{v})$$

Where  $N = R^3LR^2LR^3$ .

Notice anything intriguing about the exponents of the factorization of  $N$ ? Recall that the period of a simple continued fraction  $\sqrt{d}$  for some positive nonsquare integer  $d$  has a palindromic flavor to it. This pattern arises directly from this feature of periodic continued fractions, but we did not directly employ any knowledge of continued fractions to perform this algorithm. Naturally this pattern will continue.

## 6. REFLECTIONS ON SYMMETRY

As we look a little closer at this algorithm, several questions arise. It is quite interesting that since each step guarantees we have a balanced form, we could start at any step along the path and generate all the same stepping stones in a loop. Before we get into that however, let us revisit our unary operations from before.

**6.1. Unary operations revisited.** Clearly, increasing the value of  $\lambda(\mathbf{c})$  would be to add some  $a_{n+1}$  to the end of the coefficients of  $\mathbf{c}$ , which would translate as appending to the end of product  $h(\mathbf{c})$  an  $L^{a_{n+1}}$  if  $n$  is even and an  $R^{a_{n+1}}$  if  $n$  is odd.

**DEFINITION 12.** *If  $W$  is a word in the alphabet  $\{L, R\}$ , then  $\lambda : SL_2(\mathbb{Z}) \rightarrow \mathbb{Z}^+ \cup \{0\}$  be the **length** of  $W$  so that  $\lambda(W)$  is the number of factors in the abbreviated product of  $W$ .*

Now  $\sigma$  is quite simple to revisit, simply take the sum of the exponents in the abbreviated form of a word and be sure to add 1 for each letter whose exponent is omitted so that if  $W = R^2LR$ , then  $\sigma(W) = 4$ .

And now we arrive at the more interesting operations and will finally be able to introduce a well known matrix operation in terms of our own operations here.

**DEFINITION 13.** *If  $W$  is a word in the alphabet  $\{L, R\}$ , then the **conjugate** of  $W$ , denoted  $W^C$ , is the reversal of the order of its letters such that if  $W = \prod_{i=0}^n M_i$  where  $M_i$  is the  $i^{\text{th}}$  letter of  $W$  from the left, then  $W^C = \prod_{i=0}^n M_{n-i}$ .*

Let us see an example:

*Example 12:* Let  $U = RL^2R^2$  and let  $V = R^2L^3RL^4$ .

Reversing the order of the letters we have

$$U^C = (RLLRR)^C = RLLR = R^2L^2R$$

and also

$$V^C = (RRLLLRLLLL)^C = LLLLRLLLR = L^4RL^3R^2.$$

DEFINITION 14. If  $W$  is a word in the alphabet  $\{L, R\}$  such that  $W = \prod_{i=0}^n M_i$  where  $M_i$  is the  $i^{\text{th}}$  letter of  $W$  from the left, then  $W^F = \prod_{i=0}^n M_i^T$ , is called the **flip** of  $W$ .

This is analogous to the operation  $\rho$  and in fact, if  $h(\mathbf{c}) = W$ , then  $h(\rho(\mathbf{c})) = W^F$ . To see this, recall that  $\rho(\mathbf{c})$  is the same as adding or removing a 0 in the  $a_0$  position. If  $W = h(\mathbf{c})$ , then by Definition 7 of  $h$ ,  $W = M_0^{a_0} M_1^{a_1} \dots M_n^{a_n}$  then  $h(\rho(\mathbf{c})) = M_0^0 M_1^{a_0} M_2^{a_1} \dots M_{n+1}^{a_n} = (M_0^{a_0})^T (M_1^{a_1})^T \dots (M_n^{a_n})^T = W^F$ . Note, the case where  $a_0 = 0$  is covered here since this just means that  $M_i^{a_0} = M_i^0 = I$ . More importantly, flipping a matrix  $W$  is an intermediate step of transposing it since we have  $(A_1 A_2 \dots A_{n-1} A_n)^T = A_n^T A_{n-1}^T \dots A_2^T A_1^T$  as a known theorem about transposition of matrices—notice that if  $A_i \in \{L, R\}$ , this is by definition the conjugation of a flipped word (or the flip of a conjugated word). Let us see some example to demonstrate the relationship between transposition, flipping, and conjugation.

*Example 13:* Let  $W = LRRLRLL$ . Recall that  $R^T = L$

$$W^F = L^T R^T R^T L^T R^T L^T L^T = RLLRRLRR$$

$$W^T = (LR^2LRL^2)^T = (L^T)^2 R^T L^T (R^2)^T L^T = (W^C)^F = R^2LRL^2R$$

also note:

$$W^F = RLLRRLRR = L^T (R^2)^T L^T R^T (L^2)^T = (L^2RLR^2L)^T = (W^C)^T$$

The fact that transposition can be written as both a conjugation and flip (or reciprocation if thinking of convergents) in either order is quite nice. But one may equally think of a flip as both a conjugation and a transpose (in either order). We may now use these to describe the symmetry that we began to see at the end of chapter 5 more generally. The reader may find it an interesting preliminary exercise to refer to definitions 4 and 7 and see that if one computes the product of factors in  $L, R$  for both some  $h(\mathbf{c})$  and  $h(\rho(\mathbf{c}))$ , the result of one will be a rotation of each of the elements of the other. However, if a proof of this concept is desired, one may refer to Appendix A.2.

**THEOREM 5.** *Let  $\mathcal{P} = \{\emptyset, F, C, T\}$  where  $\emptyset$  is the identity function mapping a matrix  $A$  onto itself.  $\mathcal{P}$  forms an abelian group under function composition. This group, which preserves the determinant of any matrix it permutes, is isomorphic to the Klein four-Group  $K_4 = \{(1), (14), (23), (14)(23)\}$ , a subgroup of  $S_4$ .*

*Proof.* If  $\mathcal{P}$  is a group, it is abelian as every group of order less than six is abelian. There are only 2 group structures for sets containing only four elements and the Klein four-group is one. Therefore, to show isomorphism we need only demonstrate that  $\mathcal{P}$  is a group and that it shares a singular property with  $K_4$  which is each element of  $\mathcal{P}$  is an involution. We will, however, indicate the correspondence of

elements to those of  $S_4$  for simplicity. Let  $A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$  be a word in the alphabet  $\{L, R\}$ . The property of associativity is inherited from function composition in the general sense.

The identity function  $\emptyset$  is given by  $\emptyset(A) = A$ . As such,  $\emptyset$  is analogous to (1) which is the identity element of  $S_4$ .

So, there exists an identity element for  $\mathcal{P}$ .

We know that  $\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}^T = \begin{bmatrix} a_1 & a_3 \\ a_2 & a_4 \end{bmatrix}$  and that transposing the result again will return the matrix to its original state. In this way,  $T$  is analogous to (13).

We know that  $\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}^C = \begin{bmatrix} a_4 & a_2 \\ a_3 & a_1 \end{bmatrix}$  by Corollary 4 in Appendix A.2. Clearly, repeating this operation a second time will result in a return to  $A$ .  $C$  represents the element (14) from  $S_4$ .

Finally, we know  $\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}^F = \begin{bmatrix} a_4 & a_3 \\ a_2 & a_1 \end{bmatrix}$  by Corollary 3 in Appendix A.2. This action of flipping is equivalent to performing both  $C$  and  $T$  to the matrix  $A$  once in either order. As with the other cases, performing another operation of  $F$  is to return to  $A$ .

So, each element is its own inverse; this last property is the exact one required to show  $\mathcal{P} \cong K_4$ .

□

It is known that  $\langle \mathbb{Z}_2 \times \mathbb{Z}_2, + \rangle$  is isomorphic to  $K_4$  as well. To see this, compare the Cayley Table below to those known for the Klein four-group and  $\mathbb{Z}_2 \times \mathbb{Z}_2$ .

There is much more that could be said about this—namely that by introducing some specific other functions one could generate other groups effecting even permutations on  $A$ .

$\mathcal{P}$	$\emptyset$	$C$	$F$	$T$
$\emptyset$	$\emptyset$	$C$	$F$	$T$
$C$	$C$	$\emptyset$	$T$	$F$
$F$	$F$	$T$	$\emptyset$	$C$
$T$	$T$	$F$	$C$	$\emptyset$

FIG. 4. Cayley Table for  $\mathcal{P}$

**6.2. Types of symmetry.** When examining the words produced by Norman Wildberger's Algorithm, if we start with a Pell's quadratic form, we will end up with a word that forms a palindrome. However, there are at times some sub-symmetries that we might wish to examine. To describe these, we will employ the following terms.

DEFINITION 15. A word  $W$  in the alphabet  $\{L, R\}$  is called **chiral** if there exists a word  $K$  in the alphabet  $\{L, R\}$  such that  $W = KK^T$ .

Note that if  $W$  is chiral, then  $W = W^T$ ; however, if  $W = W^T$ , then  $W$  does not necessarily factor into the form  $W'(W')^T$ .

DEFINITION 16. A word  $W$  in the alphabet  $\{L, R\}$  is called **palindromic** if  $W = W^C$ .

If  $W$  is either palindromic or chiral, we will say that  $W$  is **symmetrical**. Naturally, the matrix resulting from performing the product indicated by  $W$  in  $SL_2(\mathbb{Z})$

will result in either a proper symmetric matrix  $\begin{bmatrix} e & f \\ f & g \end{bmatrix}$  whenever  $W$  is chiral or a

*persymmetric* matrix  $\begin{bmatrix} a & b \\ c & a \end{bmatrix}$  if  $W$  is palindromic.



*Example 14:* Let  $U = RL^3R^2LR$  and  $V = RL^4RL$ .

If  $S = RL^3R^2LRLRL^2R^3L = (RL^3R^2LR)(LRL^2R^3L) = UU^T$ , then  $S$  is chiral.

Similarly, if  $S' = LR^4LRLRL^4R = (LR^4LR)(LRL^4R) = V^TV$ , then  $S'$  is chiral. If  $P = R^2L^4RL^4R^2 = (R^2L^4RL^4R^2)^C = P^C$ , then  $P$  is palindromic. Also, if  $P' = RL^4RL^2RL^4R = (RL^4RL)(LRL^4R) = VV^C$ , then  $P'$  is palindromic.

In either case of palindromic or chiral matrices, notice that the list of exponents for the letters of the words will always form a palindrome. Bearing this in mind, assume that the exponents of the word  $W$  in the alphabet  $\{L, R\}$  form a palindrome. Then there is a way to classify these symmetries using  $\lambda$  and  $\sigma$ . If  $\lambda(W)$  is even, we know that the word  $W$  will start with  $R$  and end with  $L$  or it will begin with  $L$  and end with  $R$  and therefore cannot be a palindrome. So, assuming that the exponents of  $W$  are a palindrome,  $\lambda(W)$  is even if and only if  $W$  is chiral. Under the same assumption,  $\lambda(W)$  is odd if and only if  $W$  is palindromic. If  $\lambda(W)$  is odd and  $\sigma(W)$  is even, then  $W$  is the product of some subword and its conjugate, each in the alphabet  $\{L, R\}$ . While this may seem of no consequence, it might provide some interesting insight useful to those who wish to run reports programmatically. Especially, since we sometimes have words  $N = KK^C$  constructed from  $\{L, R\}$  for which  $K$  is chiral,  $N(1, 0)^T$  forms a solution to the Pell's equation  $x^2 - dy^2 = 1$ , and  $K(0, 1)^T$  forms a solution to the negative Pell's equation  $x^2 - dy^2 = -1$ . If  $K$  is chiral, then  $\sigma(K) = 2k$  for some  $k \in \mathbb{Z}^+$ . If this is the case though, we also know that  $\sigma(N) = 2\sigma(K) = 2(2k) = 4k$  for some  $k \in \mathbb{Z}^+$ . Furthermore, this can only be the case when  $\sigma(N)$  is even in the first place.

One point of interest (which is demonstrated in Example 16 found in Appendix C) is that whenever these chiral sub-words appear, the middle quadratic form will look like  $(d, 0, -1)$ . Moreover, when this occurs, there is a tendency for forms between the  $(1, 0, -d)$  and  $(d, 0, -1)$  to be of the forms  $(a, b, -a)$  and  $(a, -b, -a)$  where  $d = a^2 + b^2$ .

Upon review of Theorem 15.14 of [1] and its corollary, the fact that these negative solutions arise when they do will seem only natural as in these cases we also find that  $\sigma(N)$  is precisely twice the period length of  $\sqrt{d}$ . And since  $K$  is chiral,  $N$  can be factored further into  $N = KK^C = BB^T(BB^T)^C = B B^T(B^T)^C B^C$  or equivalently,  $B B^T B^F B^C$ . In either of these forms, it is nice to see that they are just permutations on the same sub-word.

## 7. BUILDING OUR VEHICLE

Wildberger's Algorithm is nice indeed and in his article (see [5]), he take things a step further by offering a means of speeding up the algorithm. However, our interest is exploratory, so it is valuable that we should explore the mechanics of quadratic forms a bit more.

### 7.1. Rebuilding the engine.

DEFINITION 17. *Let  $\mathcal{F}$  denote the set of all quadratic forms whose matrices have only integer entries. Let  $*$  :  $\mathcal{F} \times SL_2(\mathbb{Z}) \rightarrow \mathcal{F}$  such that if  $N \in SL_2(\mathbb{Z})$ ,  $f_A \in \mathcal{F}$  with matrix  $A$ , then  $f_A(\mathbf{v}) * N = f_A(N\mathbf{v}) = (N\mathbf{v})^T A (N\mathbf{v}) = \mathbf{v}^T (N^T A N) \mathbf{v}$ , an element of  $\mathcal{F}$  with matrix  $N^T A N$ . We will use a capital  $\Delta f$  to indicate the determinant of a quadratic form  $f$ .*

Note that for any balanced form  $f_A \equiv (a, b, c) \equiv \begin{bmatrix} a & b \\ b & c \end{bmatrix} = A$ ,  $\Delta f_A = ac - b^2 < 0$ . Also, if  $f_A(\mathbf{v}) * N = \mathbf{v}^T (N^T A N) \mathbf{v}$  then as shorthand, let us denote the function  $f_A$  without the evaluation of vector  $\mathbf{v}$  so that  $f_A(\mathbf{v}) * N = f_A(N\mathbf{v})$  is equivalently recorded as  $f_A * N \equiv N^T A N$ . We will also drop the subscript and simply comment what matrix belongs to a quadratic form  $f$ .

THEOREM 6.  $\mathcal{F}$  is an  $SL_2(\mathbb{Z})$ -set where  $*$  is a **right group action** of  $SL_2(\mathbb{Z})$  on  $\mathcal{F}$ .

*Proof.* There are two conditions:

1.  $f * I = f$  for all  $f \in \mathcal{F}$ .
2.  $f * (MN) = (f * M) * N$  for all  $f \in \mathcal{F}, M, N \in SL_2(\mathbb{Z})$

let  $f \equiv A$  be any quadratic form in  $\mathcal{F}$ . For the first condition,  $f * I \equiv I^T A I = A \equiv f$ . For the second condition, let  $M, N \in SL_2(\mathbb{Z})$ . Then, consider  $f * (NM) \equiv (NM)^T A (NM) = (M^T N^T) A (NM) = M^T (N^T A N) M = M^T (f * N) M \equiv (f * N) * M$ .  
□

Note that for  $f_A, f_B \in \mathcal{F}$ , if there exists invertible  $M$  such that  $f_A(M\mathbf{v}) = f_B(\mathbf{v})$ , then  $f_A$  and  $f_B$  are equivalent forms by definition. Therefore, for all  $N \in SL_2(\mathbb{Z})$ , if  $f \in \mathcal{F}$ , then  $f * N$  is an equivalent form of  $f$ .

We cannot say that  $f$  is equivalent to  $g = f * N$  if and only if  $N \in SL_2(\mathbb{Z})$  since this equivalence could also use matrices from  $GL_2(\mathbb{Z})$  which contains matrices of determinant  $-1$ . So, we will now introduce a more restrictive equivalence relation concerned only with matrices of  $SL_2(\mathbb{Z})$  acting on elements of  $\mathcal{F}$ .

THEOREM 7. Let  $f, g \in \mathcal{F}$ . Then, define the relation  $\sim$  such that  $f \sim g$  if and only if  $\exists N \in SL_2(\mathbb{Z})$  such that  $f(\mathbf{v}) * N = f(N\mathbf{v}) = g(\mathbf{v})$ .  $\sim$  is an equivalence relation.

According to Theorem 16.14 of [3], this property is established solely by the fact  $\mathcal{F}$  is an  $SL_2(\mathbb{Z})$ -set. Note that while Fraleigh makes no mention of “right group actions” specifically, it is established elsewhere that right group actions are equivalent to the generic or “left” actions described in [3]. With this in mind, we will talk about our right action as if it were any other group action without issue.

Now, to say that two forms are equal is to say that their matrices are equal as well. So if we say that  $f_A * N = f_B$  it is because  $N^T A N = B$ . This makes it quite easy to show that the determinant of any two equivalent forms in  $\mathcal{F}$  are equal.

LEMMA 3. *Let  $f, g \in \mathcal{F}$ . Then, if  $f \sim g$ , then  $\Delta f = \Delta g$ .*

*Proof.* This follows directly from the multiplicative property of determinants for matrices. Recall that all elements of  $SL_2(\mathbb{Z})$  have determinant 1.  $\square$

So, if  $f, g \in \mathcal{F}, f \sim g$  with  $f \equiv (a, b, c)$  and  $g \equiv (a', b', c')$ , it follows that  $ac - b^2 = a'c' - (b')^2$ .

Clearly, under our restriction for  $\sim$  equivalence, two equivalent forms will have the same determinant. The size of this equivalence class is infinite however. This is easy to demonstrate once we have finished rebuilding our vehicle, but for now, we would like to define a finite set of forms to examine. According to [3], each equivalence class  $[f]$  for representative  $f \in \mathcal{F}$  is an **orbit in  $\mathcal{F}$  under  $SL_2(\mathbb{Z})$** . The following two definitions are respectively more specific versions of Definitions 16.13 and 16.15 taken from [3].

DEFINITION 18. *For each  $f \in \mathcal{F}$ , the set  $SL_2(\mathbb{Z})_f = \{N \in SL_2(\mathbb{Z}) \mid f * N = f\}$  is known as the **isotropy subgroup of  $f$** .*

Fraleigh states that  $SL_2(\mathbb{Z})_f$  is indeed a subgroup of  $SL_2(\mathbb{Z})$  in [3]. We will see later that this is not such a surprise in some contexts.

DEFINITION 19. *Let  $f \in \mathcal{F}$ . Then, the equivalence class  $[f]$ , called the **orbit of  $f$** , is given by  $[f] = \{f * N \mid N \in SL_2(\mathbb{Z})\}$ .*

Since there are infinitely many elements of  $SL_2(\mathbb{Z})$ , it may be that each  $[f]$  has infinite elements also. We *can*, however, prove that the set of *balanced* forms in  $\mathcal{F}$  equivalent to some  $f$  is finite.

DEFINITION 20. For  $f \in \mathcal{F}$ , let  $[f]_b$  denote the subset of  $fSL_2(\mathbb{Z})$  containing only balanced forms.

Note that if  $f$  is not balanced,  $[f]_b \subseteq [f]$  but  $f \notin [f]_b$ .

THEOREM 8. For all balanced forms  $f \in \mathcal{F}$ ,  $[f]_b$  is finite.

*Proof.* Let  $f \equiv (a, b, c)$  with  $a > 0, c < 0$ . Then,  $ac - b^2 = \Delta f$ . Note that  $ac < 0$ . Then,  $a'c' - (b')^2 = \Delta f \iff a'c' = (b')^2 + \Delta f < 0$ . Since  $(b')^2 > 0$ , we know that there are a finite number of integer solutions to  $0 \leq (b')^2 < -\Delta f$  (clearly,  $\Delta f$  must be negative). Since  $(b')^2$  can only take on finitely many values, and there are a finite number of factorizations  $a'c'$  of  $(b')^2 + \Delta f$ , we know that there are only a finite number of values for  $a', b', c'$  for which  $ac - b^2 = a'c' - (b')^2$  and thus, only finitely many forms  $(a', b', c')$  for which  $(a, b, c) \sim (a', b', c')$ .  $\square$

Now that we know there are finitely many balanced quadratic forms in  $[f]_b$ , we can easily recognize that for any balanced form used as a seed, the Wildberger Algorithm will lead to a step reached previously. Each step taken produces yet another balanced form and there are finitely many stepping stones available as seen in Theorem 8. However, we can do better; we know that we will return to the very form we began with if it is able to continue far enough. When the determinant of the starting matrix is of the form  $-m^2$  for some integer  $m$ , we may reach some point where the total is zero in which case our algorithm will stop. As noted earlier, this is never an issue when we begin with a Pell's quadratic form because we demanded a nonsquare determinant; however, we wish to generalize the Wildberger algorithm to process other forms than the Pell's quadratic form. So before we go on to show that the algorithm will return to the first form, let us refine the definition of the Wildberger Algorithm.

LEMMA 4. If  $a + 2b + c = 0$ , then  $\begin{vmatrix} a & b \\ b & c \end{vmatrix} = -m^2$ .

**Proof:** If  $a + 2b + c = 0$ , it follows that  $a = -2b - c$ . Now let  $ac - b^2 = D$ . Then, we also have  $(-2b - c)c - b^2 = -c^2 - 2bc - b^2 = D$ . Therefore,  $D = -(b^2 + 2bc + c^2) = -(b + c)^2$ .  $\square$

DEFINITION 21. Let  $\mathcal{F}^- = \{(a, b, c) \in \mathcal{F} \mid a > 0, c < 0; \nexists m \in \mathbb{Z} \text{ s.t. } \Delta(a, b, c) = -m^2\}$ ; that is, the subset of  $\mathcal{F}$  of balanced forms whose determinant is not a negative perfect square.

We now formally introduce the generalized Wildberger Algorithm.

DEFINITION 22 (The Wildberger Algorithm). Taking any starting form  $f_0 \in \mathcal{F}^-$  as a **seed**, we may define a sequence of balanced forms equivalent to  $f_0$  using  $L, R$  by computing the value  $T = a + 2b + c$  (called the **total**) and applying the following rule:

$$f_i = f_{i-1} * M_i \text{ where } M_i = \begin{cases} L & \text{if } T > 0 \\ R & \text{if } T < 0 \end{cases} \text{ for all } i \in \mathbb{N}.$$

The **Wildberger Algorithm**, denoted  $\mathbb{W}\mathbb{A}$  refers to the process of computing a sequence of such steps and recording the actions performed until a chosen criterion is met. By default, the criterion for terminating will be reaching a quadratic form  $f$  satisfying  $f = f_0$ . Repeating the process until the termination criterion is met, a single cycle will be referred to as an **iteration**.

The sequence of steps generated may be expressed either as a word in the alphabet  $\{L, R\}$  or as a list of actions without asterisks such as:  $\mathbb{W}\mathbb{A} : (a, b, c)M_i(a', b', c')$  where  $(a, b, c) * M_i = (a', b', c')$  and  $M_i$  is an appropriate choice of  $L$  or  $R$ . A quadratic form equivalent to  $f_0$  reached from taking appropriate left and right steps may be referred to as a **stone** in the **path** of  $f_0$  (taken from the term “stepping-stones”).

Note  $\Delta f_i = \Delta f_0$  for all  $i > 0$ , under the rule of the  $\mathbb{W}\mathbb{A}$ . By the contrapositive of Lemma 4, the  $\mathbb{W}\mathbb{A}$  will never hang on a total of 0 since we will not take any seed whose determinant is additive inverse of a perfect square. We will also never encounter any stone in the path of  $f_0$  that is unbalanced since at each step we appropriately place  $T$  into one of the slots  $a'$  or  $c'$  to ensure they retain the sign of the former stone and allow the other to retain the previous value from the previous stone.

One might find it convenient that whenever  $f_0 \in \mathbb{F}^-$ ,  $[f_0] \cap \mathcal{F}^- = [f_0]_b$ . Whether or not the stones generated by  $\mathbb{W}\mathbb{A}$  seeded with  $f_0$  ever make up the entire set  $[f_0]_b$  has yet to be seen. Later, we will demonstrate that it is certainly not always the case. For now we will ponder the cyclic nature of  $\mathbb{W}\mathbb{A}$ . Since at each step  $M_i$  is unambiguous, if two stones in the path of  $f_0$  are equal, then they will have the same next step  $M_{i+1}$ .

LEMMA 5. *Let  $f, g \in \mathcal{F}$ ,  $M \in SL_2(\mathbb{Z})$ . Then,  $f(\mathbf{v}) * M = g(\mathbf{v})$  if and only if  $g(\mathbf{v}) * M^{-1} = f(\mathbf{v})$ .*

*Proof.* To show the forward direction, let  $f$  have matrix  $A$  and let  $f(\mathbf{v}) * M = g(\mathbf{v})$ .

Then,

$$f(\mathbf{v}) * M = f(M\mathbf{v}) = (M\mathbf{v})^T A (M\mathbf{v}) = \mathbf{v}^T (M^T A M) \mathbf{v} = g(\mathbf{v}).$$

Now notice

$$\begin{aligned} g(\mathbf{v}) * M^{-1} &= g(M^{-1}\mathbf{v}) = (M^{-1}\mathbf{v})^T (M^T A M) (M^{-1}\mathbf{v}) = \\ &= \mathbf{v}^T (M M^{-1})^T A (M M^{-1}) \mathbf{v} = \mathbf{v}^T A \mathbf{v} = f(\mathbf{v}) \end{aligned}$$

To show the reverse direction, simply swap  $f$  and  $M$  with  $g$  and  $M^{-1}$  respectively.  $\square$

THEOREM 9. *Let  $f_0 \in \mathcal{F}^-$  be a seed for a sequence generated by  $\mathbb{W}\mathbb{A}$ . Then,  $\mathbb{W}\mathbb{A}$  must return to  $f_0$  after a sufficient number of steps. That is,  $\exists n > 0$  such that  $f_n = f_0$ .*



*Proof.* By Theorem 8 and Lemma 4, we know that  $\exists k \in \mathbb{N}$  such that  $f_k = f_n$  for some  $0 \leq k < n$ . We can show that the smallest  $n$  for which this is true will yield  $f_n = f_0$ .

Suppose  $0 \leq k < n$  where  $n$  is the smallest possible selection allowing  $f_k = f_n$ . Since the choices of  $M_k$  and  $M_n$  are unambiguous decisions for  $f_{k-1} * M_k = f_k$  and  $f_{n-1} * M_n = f_n$ , it follows from Lemma 5 that  $M_k = M_n$ . Therefore, by another application of Lemma 5,  $f_n * M_n^{-1} = f_{n-1} = f_{k-1} = f_k * M_k^{-1}$ . Now, simply replace  $n$  and  $k$  with  $n - i$  and  $k - i$  respectively and let  $0 \leq i \leq k$  since  $n$  and  $k$  are just nonnegative integers and we know that for each increment down from  $n$  and  $k$ , we will still have  $f_{n-i} = f_{k-i}$  by the same logic as for the  $f_{n-1} = f_{k-1}$  case. Then, if we let  $i = k$  we will find that  $f_{n-k} = f_0$  where  $n - k > 0$ . Notice then that since we said  $n$  should be the smallest possible selection allowing  $f_k$  to equal  $f_n$  and we must have started with  $n = n - k$  and  $k = 0$ . So  $f_n = f_0$   $\square$

**7.2. Preparing to drive.** Before we start the engine, we should be sure we know how to navigate properly once we are no longer on a rail and then gas up with some motivations for proceeding to drive.

**THEOREM 10.** *Let  $c < 0 < a$  for  $a, c \in \mathbb{Z}$  and let  $A$  be the matrix of  $(a, 0, c)$ . If,  $N = \begin{bmatrix} p & q \\ r & s \end{bmatrix} \in SL_2(\mathbb{Z})$  and  $N^T AN = A$ , then  $N$  is persymmetric.*

*Proof.* Let  $N^T AN = A$ . It follows that  $AN = (N^T)^{-1}A$ .

$$\begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix} \begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} p & r \\ q & s \end{bmatrix}^{-1} \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix} = \begin{bmatrix} s & -r \\ -q & p \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix}$$

$$\begin{bmatrix} ap & aq \\ cr & cs \end{bmatrix} = \begin{bmatrix} as & -rc \\ -qa & pc \end{bmatrix} \implies ap = as, cs = cp.$$

It follows that  $p = s$  meaning that  $N$  is persymmetric.  $\square$

We also have

$$(15) \quad aq = -rc.$$

Notice also that if  $N$  is persymmetric, it is in the right shape for elements of  $H_d$ .

**THEOREM 11.** *Let  $a, b, c \in \mathbb{Z}$  with  $a + c = 0, 0 < a$  and let  $A$  be the matrix of  $(a, b, c)$ . If,  $N = \begin{bmatrix} p & q \\ r & s \end{bmatrix} \in SL_2(\mathbb{Z})$  and  $N^T AN = A$ , then  $N$  is symmetric.*

*Proof.* Let  $N^T AN = A$ . It follows that  $N^T A = AN^{-1}$ . Note that  $c = -a$ .

$$\begin{bmatrix} p & r \\ q & s \end{bmatrix} \begin{bmatrix} a & b \\ b & -a \end{bmatrix} = \begin{bmatrix} a & b \\ b & -a \end{bmatrix} \begin{bmatrix} s & -q \\ -r & p \end{bmatrix}$$

$$\begin{bmatrix} pa + rb & pb - ra \\ qa + sb & qb - sa \end{bmatrix} = \begin{bmatrix} as - br & bp - aq \\ bs + ar & -bq - ap \end{bmatrix}$$

From the resulting equations, we can derive  $q = r$ ; therefore,  $N = N^T$  meaning that  $N$  must be symmetric.  $\square$

$$(16) \quad \text{Also, } 2rb = a(s - p)$$

Note that this last equation is derived without the assumption that  $a + c = 0$ .

It is clear to us that we cannot have a nontrivial word that is both chiral and a palindrome. It is left to the reader as a simple exercise to show that  $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ,  $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$  are the only matrices  $N$  in  $SL_2(\mathbb{Z})$  that satisfy  $N^T A N = A$  when  $A = \begin{bmatrix} a & 0 \\ 0 & -a \end{bmatrix}$ . This can be done similarly to the proofs of the last two Theorems. The reader may refer again to Example 16 of Appendix C and note the different symmetries that arise. If one chooses to do so, look back to the stones generated for  $\sqrt{13}$  as you go to see how these matrices correspond to them.

Our next theorem will illustrate the motivation for our efforts more clearly.  $\mathbb{W}\mathbb{A}$  was introduced as a way to solve Pell's equation and Pell's equation was introduced as a means of approximating square roots for positive nonsquare integers. However, using our generalization of  $\mathbb{W}\mathbb{A}$ , we will show that we can use it to generate approximations of any quadratic irrational of the form  $\sqrt{\frac{c}{a}}$  whenever  $a, c$  are relatively prime square-free positive integers.

**THEOREM 12** (The infinite road to  $\sqrt{\frac{c}{a}}$ ). *Let  $a, c \in \mathbb{Z}^+$  be integers with  $\gcd(a, c) = 1$  where  $a$  and  $c$  are not both square. Suppose that  $n \in SL_2(\mathbb{Z})$  is a word generated by  $\mathbb{W}\mathbb{A}$  seeded with  $f \equiv (a, 0, -c)$  such that  $(a, 0, -c)N(a, 0, -c)$ . Then,  $N = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$  satisfies the following criteria:*

1.  $N$  is persymmetric (in fact, a palindrome).
2.  $|\frac{p}{r} - \sqrt{\frac{c}{a}}| < \frac{1}{r^2}$ .
3.  $\frac{q}{r} = \frac{c}{a}$

*Proof.* 1 is an immediate result from Theorem 10; therefore,

$$\begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} p & q \\ r & p \end{bmatrix}$$

To show 2 note that since  $f * N = f$ ,  $f(\mathbf{v}) = f(N\mathbf{v})$ . Now, we know that  $f(x, y) = ax^2 - cy^2$  and it follows that  $f(1, 0) = a$ . With this we can assume that  $f(N\mathbf{v}) = a$  where  $\mathbf{v} = (1, 0)^T$ .

$$\begin{bmatrix} p & q \\ r & p \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} p \\ r \end{bmatrix}$$

So,  $f(p, r) = ap^2 - cr^2 = a$ . Now the algebra.

$$ap^2 - cr^2 = a$$

$$p^2 - \frac{c}{a}r^2 = 1$$

$$\frac{p^2}{r^2} - \frac{c}{a} = \frac{1}{r^2}$$

$$\frac{p^2}{r^2} = \frac{1}{r^2} + \frac{c}{a}$$

$$\frac{p}{r} = \sqrt{\frac{1}{r^2} + \frac{c}{a}}$$

$$0 < \frac{p}{r} < \frac{1}{r} + \sqrt{\frac{c}{a}}$$

$$\left| \frac{p}{r} - \sqrt{\frac{c}{a}} \right| < \frac{1}{r}$$

To show 3, we notice that  $N^T A = A(N^T)^{-1}$  and then borrowing from 1, we see

$$\begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} p & q \\ r & p \end{bmatrix} \text{ also. Therefore by}$$

$$\begin{bmatrix} p & r \\ q & p \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & -c \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & -c \end{bmatrix} \begin{bmatrix} p & -q \\ -r & p \end{bmatrix} \iff \begin{bmatrix} ap & -cr \\ qa & -pc \end{bmatrix} = \begin{bmatrix} ap & -aq \\ cr & -cp \end{bmatrix},$$

we obtain  $cr = aq$  and finally,  $\frac{q}{r} = \frac{c}{a}$ .  $\square$

Note that, while we do not prove it,  $\mathbb{WA}$  traces the path to  $\sqrt{\frac{c}{a}}$  in the Stern-Brocot tree. Therefore, because  $N$  is the resulting matrix product of the  $L$  and  $R$ -steps traced, the column vectors of  $N$  will be the convergents of  $\sqrt{\frac{c}{a}}$ . Since  $\frac{p}{r}$  is one of these convergents, the above inequality can be improved to  $|\frac{p}{r} - \sqrt{\frac{c}{a}}| < \frac{1}{r^2}$  as commented on above equation 4 on page 6. Moreover, if one finds some  $N$  satisfying Theorem 12, since the left column of  $N$  is a convergent of  $\sqrt{\frac{c}{a}}$ , Lemma 1 says that  $r$  will increase in size as we raise  $N$  to some power  $m$ . And,  $N^m$  will also satisfy Theorem 12. So, we can find as good an approximation as we want in this way.

Also, since 1 is a positive integer, selecting 1 as  $a$ , we generalize the algorithm to include calculating ratios of quadratic irrationals in addition to its original integral application. Now, consider Equation (7) from section 3.2. If we remove the  $r$ -term, we get a rational number  $q$ ; however, if we replace  $r$ , we get a new value  $x$ . But can't we write this as  $x = q + r'$ ? With a little effort we can see that by adding the  $r$ -term, we are adding or removing (if  $r$  is a coefficient with odd index) some value  $r'$ . If  $r'$  is rational this is obvious. If  $r'$  is a quadratic irrational, then  $r$  is irrational and the sequence of denominators from there on will become periodic. Looking at this in reverse, approximating  $r'$  is the tricky bit (that we now can do easily), but any extra rational value added to  $r'$  can be thought of as appending some finite sequence of coefficients to the beginning of the fraction.

*Example 15:* Consider  $\frac{61}{16}$ . While we will not display the full results here, running the code found in Appendix D.4, you can verify the result yourself. According to our generalized algorithm, we will let  $a = 16$  and  $c = 61$  and then evaluate  $\text{WA}((16, 0, -61))$ . The abbreviated word following is the result:

$$N = RLR^{20}L^{12}R^2L^4R^3L^{15}R^3L^4R^2L^{12}R^{20}LR.$$

$$\sigma(N) = 101, \lambda(N) = 15.$$

Since  $\lambda(N) = 15$ , if we abbreviate the path and note only the stones between changes from a left to right (or right to left) decision, we can summarize the path in 15 stones. However, it is more interesting to check the accuracy of our obtained approximation.

The resulting matrix is:

$$N = \begin{bmatrix} 1,766,319,049 & 3,448,848,195 \\ 904,615,920 & 1,766,319,049 \end{bmatrix}.$$

Also,  $\frac{3,448,848,195}{904,615,920} = 3.8125 = \frac{61}{16}$  and  $\sqrt{\frac{61}{16}}$  is approximately:

$$1.9525624189766635985324306839398.$$

Using our obtained left hand column,  $\frac{1,766,319,049}{904,615,920}$  is exactly:

$$1.9525624189766635988453530643149.$$

The resulting error is around  $0.3129223803751 \times 10^{-18}$  and  $\frac{1}{r^2}$  is  $1.2220009599... \times 10^{-18}$ .

Now, suppose that we let  $a = 1, c = d$ . Then, Theorem 12 provides us with the following fact:  $N$  is persymmetric where  $(1, 0, -d)N(1, 0, -d)$  and  $\frac{a}{r} = d$ . With a bit of effort, one might realize that  $N \in H_d$ . Remember the isotropy groups mentioned before?  $H_d$  is the isotropy group for  $(1, 0, d)$ .

**7.3. Driving in circles.** Concerning orbits, if one were to try seeding  $\mathbb{W}\mathbb{A}$  with an unbalanced form, the algorithm is likely to run infinitely generating unbalanced forms along the way. There is a special condition on the value of  $b$  that allows  $\mathbb{W}\mathbb{A}$  to return to balanced forms and work correctly, but this is not the intended use. Regardless, since the algorithm is potentially capable of producing an infinite sequence of steps with only unbalanced forms, the orbits of a given seed form  $f_0$  may be many. However, since we know that  $[f]_b$  is a finite subset of the orbit, for any two forms  $f, g \in \mathcal{F}$  showing  $[f] \cap [g] = \emptyset$  is sufficient to show  $[f]_b$  and  $[g]_b$  are disjoint. What is more, since  $\Delta g = \Delta f$  for all  $g \in [f]$ , we can look at a collection of orbits classified by their determinant. If we restrict ourselves to the balanced subset of the orbits we have:

$$\bigcup_{\Delta f = -d} [f]_b \subseteq \{(a, b, c) \mid a > 0, c < 0, ac - b^2 = -d\}.$$

This set is finite we know, but it may also have some intriguing features. Namely, the stones generated by  $\mathbb{W}\mathbb{A}$  for some  $f$  with  $\Delta f = -d$  will all be in this set. If we find that two seeds  $f, g$  both have determinant  $-d$  but the stones generated by  $\mathbb{W}\mathbb{A}$  for each are disjoint (and they must be either disjoint or equal if  $f, g$  are both balanced). this begs the question: “Are the orbits of  $f$  and  $g$  disjoint?”

THEOREM 13. If  $\exists N \in SL_2(\mathbb{Z})$  such that  $(1, 0, -d)N(d, 0, -1)$ , then  $N(0, 1)$  is a solution to  $x^2 - dy^2 = -1$ .

*Proof.* Suppose there is some  $N \in SL_2(\mathbb{Z})$  such that  $(1, 0, -d) \sim (d, 0, -1)$ .

Let  $N = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$ . Then,

$$\begin{bmatrix} p & r \\ q & s \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -d \end{bmatrix} \begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} d & 0 \\ 0 & -1 \end{bmatrix}$$

$$\begin{bmatrix} p & r \\ q & s \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -d \end{bmatrix} = \begin{bmatrix} d & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} s & -q \\ -r & p \end{bmatrix}$$

$$\begin{bmatrix} p & -rd \\ q & -ds \end{bmatrix} = \begin{bmatrix} ds & -dq \\ r & -p \end{bmatrix}$$

$$p = ds$$

$$q = r.$$

So,  $N = \begin{bmatrix} ds & q \\ q & s \end{bmatrix}$  and  $\Delta N = 1 = ds^2 - q^2$ . But this also means that  $q^2 - ds^2 = -1$  so  $N(0, 1)$  forms a solution for the negative Pell equation for  $d$ .  $\square$

It is noteworthy that the constraints on  $N$  required symmetry. In fact, this is precisely what we see in the example for  $d = 13$ . Any time the form  $(d, 0, -1)$  shows up as a stone in the path of  $(1, 0, -d)$ , we are guaranteed to have the symmetric matrix  $N$  for which  $N(0, 1)$  is a negative Pell solution for  $d$ .



COROLLARY 1. *If there is no solution for the negative Pell equation for a non-square positive integer  $d$ , then the orbits of  $(1, 0, -d)$  and  $(d, 0, -1)$  are disjoint.*

*Proof.* This follows directly from Theorem 13. If there can be no equivalence between any two quadratic forms, then their equivalence classes must be disjoint due to the transitivity of the relation.  $\square$

Here we find an interesting thought: if we can show that there is an equivalence  $f \sim g$  for *any*  $f \in [(1, 0, -d)]$  and  $g \in [(d, 0, -1)]$ , then we will show the orbits of  $(1, 0, -d)$  and  $(d, 0, -1)$  are equal and that there is a negative Pell solution available for  $d$ . However, it would be nice to have a stronger statement.

COROLLARY 2. *The equation  $x^2 - dy^2 = -1$  has an integral solution if and only if  $(1, 0, -d) \sim (d, 0, -1)$ .*

*Proof.* By Theorem 13, the back direction of this corollary is shown. We will now show the forward direction is also true with ease. Suppose there is a solution  $(q, s)$  to the equation  $x^2 - dy^2 = -1$ . Then,  $ds^2 - q^2 = 1$ , so let us examine the matrix  $N = \begin{bmatrix} ds & q \\ q & s \end{bmatrix} \in SL_2(\mathbb{Z})$ . Note that  $N^T = N$

$$\begin{aligned} & \begin{bmatrix} ds & q \\ q & s \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -d \end{bmatrix} \begin{bmatrix} ds & q \\ q & s \end{bmatrix} = \begin{bmatrix} ds & -qd \\ q & -sd \end{bmatrix} \begin{bmatrix} ds & q \\ q & s \end{bmatrix} \\ & = \begin{bmatrix} d^2s^2 - dq^2 & dsq - dsq \\ dsq - dsq & q^2 - ds^2 \end{bmatrix} = \begin{bmatrix} d(ds^2 - q^2) & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} d & 0 \\ 0 & -1 \end{bmatrix} \end{aligned}$$

Therefore, anytime there is a negative Pell solution  $(q, s)$  for  $d$ , we will know that the matrix  $N = \begin{bmatrix} ds & q \\ q & s \end{bmatrix}$  satisfies  $(1, 0, -d)N(d, 0, -1)$ .  $\square$

The trouble that remains for us is just this: if  $\mathbb{W}\mathbb{A}$  produces two distinct paths for forms  $f, g$  this doesn't mean they are disjoint necessarily. Although the equivalence would have intermediate steps into forms that are unbalanced, there still may be some path of equivalence between them. So we cannot say that there is no negative solution based solely on the fact  $(d, 0, -1)$  is not a stone in the path of  $(1, 0, -d)$ .

## 8. CLOSING THOUGHTS

In mathematics we often discuss the cardinality of sets and the poetic thing about topics such as these are that asking questions often leads to exponentially more questions. The Stern-Brocot tree being a binary tree, it has countably many nodes, but uncountably many infinite paths. It should then be only natural that there will always be more to learn and know. When beginning this research I had an interest in algebra and my major professor was currently interested in number theoretic problems. To his amusement, I may have made it an algebra problem anyway. This is easily done—algebra is everywhere.

More work needs to be done to classify, and even restrict, the group and orbits used to evaluate the forms we have examined. We have only mentioned in passing the forms that produce solutions to the sum of squares problem but they are present nonetheless.

Concerning the Stern-Brocot tree and its representation in  $SL_2(\mathbb{Z})$ , it is interesting: it seems quite possible to express everything as discussed using the inverse matrices  $L^{-1}$  and  $R^{-1}$  and one might suspect that this language can be expanded into the negative rational numbers. Furthermore, if one were to take inverse steps amidst the usual left and right steps on the tree, one would find the resulting matrix isn't in the tree—at least not in the usual sense. If it is true that the Stern-Brocot tree can be described using other combinations of  $L, L^{-1}, R, R^{-1}$ , then the question would be, can these alternate trees be accessed using these "incorrect" intermediate steps? Think about it. If we make a series of left and right decisions on the Stern-Brocot tree

represented through matrix multiplication, we could easily backtrack by multiplying by the inverses of steps taken in reverse order. However, if we were to take steps such as  $LR^{-1}L$ , this enclosed inverse right decision would find us somewhere we do not recognize.

There is also the question of other notations. There is potential for representation of continued fractions as strings in a more robust sense than that of the matrix products alone that utilizes concatenation. The study of strings under concatenation is useful in its own right, so to think that continued fractions might be expressible in this way is nice. It is natural though, since we know that adding more terms to the end of a fraction is the same as adding more factors to the right side of a matrix product in  $SL_2(\mathbb{Z})$ .

Concerning the generalized Wildberger algorithm, we have placed a restriction on the numbers  $a$  and  $c$  that they not both be square, but in fact, we may let both be squares and we will get the path down the Stern-Brocot tree for the rational number  $\sqrt{\frac{a}{c}}$ . In the exploration of these quadratic forms used as a seed for the algorithm, it was devised that we may take any number  $\frac{u+\sqrt{v}}{w}$  where  $u^2 < v$  and first propose  $a = 1, b = \frac{2u}{w}$ , and  $c = \frac{u^2-v}{w}$  and then simply multiply each by the least integral scalar  $m$  so that  $am, bm$ , and  $cm$  are integers and in this way we acquire a seed for  $\mathbb{W}\mathbb{A}$  allowing for approximations to  $\frac{u+\sqrt{v}}{w}$ . If  $v$  is a perfect square so that  $v' = \sqrt{v}$  and  $u = 0$  then we have a seed form of  $(w^2, 0, (v')^2) = (w^2, 0, v)$  and will find not an approximation, but an exact expression of  $\frac{v'}{w}$  when we take the mediant of the resulting matrix produced from the left and right steps taken.

This particular topic needs more exploration, especially since we do not have an mechanism for quadratic irrationals of the mentioned form when  $u^2 > v$ . Allowing  $u^2 > v$  causes unbalanced forms, so the algorithm we have examined is not sufficient in its current state. Alas, there are so many questions that can be asked, but having

finite time and resource, one must be selective concerning which topics should be pursued. It is here we conclude the primary work, but not without hope for the continuation thereof.

## REFERENCES

- [1] D. M. BURTON, *Elementary number theory*, 7th ed., McGraw-Hill, Boston, MA, 2010.
- [2] K. CONRAD, *Pell's Equation, I*,  
<https://kconrad.math.uconn.edu/blurbs/ugradnumthy/pelleqn1.pdf> .
- [3] J. B. FRALEIGH, *A first course in abstract algebra*, 7th ed., Addison-Wesley, Boston, MA, 2003.
- [4] R. L. GRAHAM, D. E. KNUTH, AND O. PATASHNIK, *Concrete Mathematics: A Foundation for Computer Science*, 2nd ed., Addison-Wesley, Reading, MA, 1991.
- [5] N. J. WILDBERGER, *Pell's equation without irrational numbers*, *J. Integer. Seq.*, 13 (2010), 10.4.3.

# APPENDIX A

## ACCESSORY PROOFS AND DEFINITIONS

If mathematics is a miracle, John 21:25 waxes poignant.

### A.1. Omitted proofs.

LEMMA 1. *Let  $a, b, d \in \mathbb{Z}^+$  and  $a^2 - db^2 = 1$ , with  $d$  nonsquare,  $a > 1$ . If  $a_n + b_n\sqrt{d} = (a + b\sqrt{d})^n$ , then  $\{b_n\}$  is a strictly increasing sequence.*

*Proof.* Since  $b_1 = b$  and, notice that

$$b_1 > 0$$

$$b_1 + b_1 > 0 + b_1$$

$$ab_1 + a_1b > b_1$$

and by equation (3) on page 6

$$b_2 > b_1.$$

We also have  $a_1a > a_1$  and  $b > 0$  (since  $a > 1$ ); therefore,  $b_1b > 0$  as well. Combining  $a_1a > a_1$  with  $b_1b > 0$  we obtain  $a_1a + b_1b > a_1 + 0$  and so by equation (2) on page 6, we have  $a_2 > a_1$  also. Now, assume  $b_n > b_{n-1}$  and  $a_n > a_{n-1}$  for all  $n > 1$ . Then  $b > 0 \implies a_nb > 0$ . Also  $a > 1 \implies b_na > b_n$  and thus  $a_nb + b_na > 0 + b_n$ .

Therefore,  $b_{n+1} > b_n$  □

## A.2. Unary operations and permutations.

**THEOREM 14** (Reciprocal Representation of Coefficients). *If  $c > 1$  and  $c \equiv \mathbf{c} = [a_0; a_1, a_2, \dots, a_n]$  has convergents  $\frac{p_k}{q_k}$  defined in the usual way, then  $\frac{1}{c} \equiv \rho(\mathbf{c})$  has convergents  $\frac{p'_k}{q'_k}$  where  $p'_0 = 0, q'_0 = 1$ , and for all  $k > 0, p'_k = q_{k-1}$  and  $q'_k = p_{k-1}$ .*

*Proof.* Let for rational number  $c$  greater than 1,  $c \equiv \mathbf{c} = [a_0; a_1, a_2, \dots, a_n]$ . Then  $\rho(\mathbf{c}) = [b_0; b_1, b_2, \dots, b_{n+1}]$  where  $b_0 = 0$  and  $b_{i+1} = a_i$  for all  $i > 0$ . We know that the convergents of  $\mathbf{c}$  will be given by  $\frac{p_k}{q_k}$  where  $p_k, q_k$  are defined by

$$\begin{aligned} p_0 &= a_0 & q_0 &= 1 \\ p_1 &= a_1 a_0 + 1 & q_1 &= a_1 \\ p_k &= a_k p_{k-1} + p_{k-2} & q_k &= a_k q_{k-1} + q_{k-2}. \end{aligned}$$

Similarly, the convergents  $\frac{p'_k}{q'_k}$  of  $\rho(\mathbf{c})$  will be given by

$$\begin{aligned} p'_0 &= b_0 & q'_0 &= 1 \\ p'_1 &= b_1 b_0 + 1 & q'_1 &= b_1 \\ p'_k &= b_k p'_{k-1} + p'_{k-2} & q'_k &= b_k q'_{k-1} + q'_{k-2}. \end{aligned}$$

By the definitions for the convergents of  $\mathbf{c}$  and  $b_i$ , we obtain

$$\begin{aligned} p'_0 &= 0 & q'_0 &= 1 \\ p'_1 &= a_0(0) + 1 = q_0 & q'_1 &= a_0 = p_0 \end{aligned}$$

by mere substitution. Having satisfied  $p'_0 = 0, q'_0 = 1$ , we now wish to demonstrate

that  $\frac{p'_k}{q'_k} = \frac{q_{k-1}}{p_{k-1}}$  for all  $k > 0$ . This is satisfied for  $k = 0$  as seen in the last block of equations, so to demonstrate the general case, we can use induction on  $k$ . Assume



the hypothesis holds for all  $k > 0$ . Then,

$$p'_{k+1} = b_{k+1}p'_k + p'_{k-1} = a_k q_{k-1} + q_{k-2} = q_k$$

and

$$q'_{k+1} = b_{k+1}q'_k + q'_{k-1} = a_k p_{k-1} + p_{k-2} = p_k.$$

This completes the induction and the proof.  $\square$

CORLLARY 3. If  $W = \begin{bmatrix} u & v \\ x & w \end{bmatrix}$  is the resulting matrix product of a word in the alphabet  $\{L, R\}$ , then  $\begin{bmatrix} u & v \\ x & w \end{bmatrix}^F = \begin{bmatrix} w & x \\ v & u \end{bmatrix}$

*Proof.* If  $W$  is a word in the alphabet  $\{L, R\}$ , then there exists  $\mathbf{c} \in \mathbf{C}$  (namely the exponents of  $W$ 's abbreviated form) such that  $h(\mathbf{c}) = W$ . Without loss of generality, Let  $c \equiv \mathbf{c}$  be greater than 1. Let  $\frac{p_{n-1}}{q_{n-1}}$  and  $\frac{p_n}{q_n}$  be the final two convergents of  $\mathbf{c}$ .

Then we have  $W = \begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix}$  when  $n$  is odd and  $W = \begin{bmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{bmatrix}$  when  $n$  is even. Since  $h^{-1}(W^f) = \rho(\mathbf{c})$ , as noted immediately following Definition 14, if we let  $\frac{p'_n}{q'_n}$  and  $\frac{p'_{n+1}}{q'_{n+1}}$  be the final two convergents of  $\rho(\mathbf{c})$ , we can infer that  $W^F = h(\rho(\mathbf{c}))$  is equal to

$$\begin{bmatrix} p'_{n+1} & p'_n \\ q'_{n+1} & q'_n \end{bmatrix}$$

when  $n$  is even and

$$\begin{bmatrix} p'_n & p'_{n+1} \\ q'_n & q'_{n+1} \end{bmatrix}$$

when  $n$  is odd. It follows from Theorem 14 that  $\begin{bmatrix} p'_{n+1} & p'_n \\ q'_{n+1} & q'_n \end{bmatrix} = \begin{bmatrix} q_n & q_{n-1} \\ p_n & p_{n-1} \end{bmatrix}$  and

$$\begin{bmatrix} p'_n & p'_{n+1} \\ q'_n & q'_{n+1} \end{bmatrix} = \begin{bmatrix} q_{n-1} & q_n \\ p_{n-1} & p_n \end{bmatrix}$$

These are the only cases, and each satisfies the requirement desired to show  $\begin{bmatrix} u & v \\ x & w \end{bmatrix}^F = \begin{bmatrix} w & x \\ v & u \end{bmatrix}$ . □

Thus emphasizing once more the term “flip.”

COROLLARY 4. Let  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  be the resulting matrix product from a word in the alphabet  $\{L, R\}$ . Then  $A^C = \begin{bmatrix} d & b \\ c & a \end{bmatrix}$

*Proof.* By Corollary 3; we know that  $A^F = \begin{bmatrix} d & c \\ b & a \end{bmatrix}$ . Since we also know that for word  $W$ ,  $W^T = (W^F)^T$ , we can infer that  $(A^F)^T = ((A^F)^F)^C = A^C$  meaning  $A^C = (A^F)^T = \begin{bmatrix} d & c \\ b & a \end{bmatrix}^T = \begin{bmatrix} d & b \\ c & a \end{bmatrix}$ . □

DEFINITION 23 (Continued Fractions). A (positive) simple continued fraction  $c$  is a real number constructed by a sequence of integers  $\{a_i\}$  where  $a_i > 0$  for  $i > 0$  in the following way:  $c = a_0 + \frac{1}{b_1}$  where for all  $i \geq 1$ ,  $b_i = \begin{cases} a_i + \frac{1}{b_{i+1}} & a_{i+1} \text{ exists} \\ a_i & \text{otherwise.} \end{cases}$  Note that  $b_i \neq 0$  for any  $i$  since every  $a_i$  after the first must be positive.

## APPENDIX B

### AMBIGUITY OF FORM

There are some subtleties desiring attention regarding the representation of rational numbers as continued fractions. One such subtlety is the handling of intermediate zeros. We might wish to construct new fractions from joining two sets of coefficients together, but what if one has a value between 0 and 1? In this case we would have a zero in the middle of our coefficients. But never fear: We may handle intermediate zeros with ease given that the right criteria are met.

**THEOREM 15** (Handling intermediate zeros in coefficients). *If continued fraction  $c \equiv [a_0; a_1, \dots, a_{k-1}, a_k, a_{k+1}, \dots, a_n]$ ,  $a_k = 0$ , and for all  $i, a_i \geq 0$ , then  $c \equiv [a_0; a_1, \dots, a_{k-1} + a_{k+1}, \dots, a_n]$ , so long as there exists  $n > k$  for which  $a_n > 0$ .*

*Proof.* Let  $c \equiv [a_0; a_1, \dots, a_n]$  where  $a_n \neq 0$ . Suppose  $k$  is the largest integer less than  $n$  for which  $a_k = 0$ . Then, suppose  $r \equiv [a_{k+1}; \dots, a_n]$  and no coefficient for the expansion of  $r$  will be equal to zero. It follows that

$$c = a_0 + \frac{1}{a_1 + \frac{1}{\dots + \frac{1}{a_{k-1} + \frac{1}{0 + \frac{1}{r}}}}} = a_0 + \frac{1}{a_1 + \frac{1}{\dots + \frac{1}{a_{k-1} + \frac{1}{r}}}}$$

$$= a_0 + \frac{1}{a_1 + \frac{1}{\dots + \frac{1}{a_{k-1} + r}}} = a_0 + \frac{1}{a_1 + \frac{1}{\dots + \frac{1}{a_{k-1} + a_{k+1} + \frac{1}{\dots + \frac{1}{a_n}}}}}$$

Now,  $a_{k-1} + a_{k+1}$  is just an integer and  $a_{k+1}$  is positive, so  $a_{k-1} + a_{k+1} > 0$  and  $c \equiv [a_0; a_1, \dots, a_{k-1} + a_{k+1}, \dots, a_n]$   $\square$

This theorem is sufficient to claim that we can remove all intermediate zeros. If we have some continued fraction with multiple intermediate zeros, we may perform the process described in the proof for the last zero, relabel the coefficients with their new indices, and finally repeat the process for the new last zero whose index will be less than that of the first. In this way, we will eliminate each zero in less than  $n$  steps.

Now, if  $\mathbf{c}$  represents an infinite fraction, so long as there are only finitely many zeros in  $\mathbf{c}$ , the theorem will still hold. The reader may wish to explore this idea on their own.

Now that we have handled intermediate zeros, we will create a rule for trailing zeros.

**DEFINITION 24.** *If  $c \equiv [a_0; a_1, \dots, a_n, \dots]$  and  $a_i = 0$  for all  $i > n$ , we will say instead that  $c \equiv [a_0; a_1, \dots, a_n]$ .*

For an example of handling zeros, see Example 18 in Appendix C.1

Now, the intention in addressing ambiguous forms is that if we want to define a clear mapping of the rational numbers (or at least the positive rational numbers) onto the set of continued fraction coefficients and then onto the Stern-Brocot tree, we would prefer if our mapping were one-to-one and therefore invertible.

For reference, here are the recurrence relations for  $p_k, q_k$  as introduced by [1].

$$\begin{aligned}
 (17) \quad p_0 &= a_0 & q_0 &= 1 \\
 p_1 &= a_1 a_0 + 1 & q_1 &= a_1 \\
 p_k &= a_k p_{k-1} + p_{k-2} & q_k &= a_k q_{k-1} + q_{k-2}
 \end{aligned}$$

Now, consider the scenario where we have some rational number  $c_k = \frac{p_k}{q_k}$  and let  $a_{k+1} = 1$

$$(18) \quad [a_0; a_1, \dots, a_k] \equiv \frac{p_k}{q_k} = \frac{a_k p_{k-1} + p_{k-2}}{a_k q_{k-1} + q_{k-2}}$$

$$\begin{aligned}
 (19) \quad [a_0; a_1, \dots, a_k + 1] &\equiv \frac{(a_k + 1)p_{k-1} + p_{k-2}}{(a_k + 1)q_{k-1} + q_{k-2}} = \frac{a_k p_{k-1} + p_{k-1} + p_{k-2}}{a_k q_{k-1} + q_{k-1} + q_{k-2}} \\
 &= \frac{(a_k p_{k-1} + p_{k-2}) + p_{k-1}}{(a_k q_{k-1} + q_{k-2}) + q_{k-1}} = \frac{p_k + p_{k-1}}{q_k + q_{k-1}}
 \end{aligned}$$

$$(20) \quad [a_0; a_1, \dots, a_k, a_{k+1}] \equiv \frac{a_{k+1} p_k + p_{k-1}}{a_{k+1} q_k + q_{k-1}} = \frac{1(p_k) + p_{k-1}}{1(p_k) + p_{k-1}} = \frac{p_k + p_{k-1}}{p_k + p_{k-1}}$$

It is of no real surprise that we find  $[a_0; a_1, \dots, a_k + 1] \equiv [a_0; , a_1, \dots, a_k, 1]$  since we did comment in passing earlier that every fraction may be written in an alternate form whose last partial denominator is 1 so long as we subtract 1 from the next to last partial denominator. However, for those concerned with well-defined notation, this becomes slightly troublesome. Our coefficient notation does not uniquely define the

rational numbers (except for 0 which has no trailing 1 to be placed), so an alternate interpretation is proposed. Notice that each of the two forms for any *positive* rational number  $c$  will have the same value for  $a_{k-1}$ . Because of this, we can easily *and uniquely* represent any positive rational number with the usual coefficients if we simply subtract 1 from  $a_k$  so, let us now define an alternative to the normal convention.

DEFINITION 25. Let  $\mathbf{C} := \{[a_0; a_1, \dots, a_n] \mid n \geq 0; \forall i > 0, a_i \in \mathbb{Z}^+\}$ . Also, let  $c = \frac{p}{q} \in \mathbb{Q}^+$  and let  $a_i$  be quotients obtained by the euclidean algorithm for  $\frac{p}{q}$  (as seen in equation (5) of page 10). Then, let  $g : \mathbb{Q}^+ \rightarrow \mathbf{C}$  such that  $g(c) = [b_0; b_1, \dots, b_m]$  where  $m = \begin{cases} n-1 & a_n = 1 \\ n & \text{otherwise,} \end{cases} \quad \forall i < m, b_i = a_i, \text{ and } b_m = \begin{cases} a_{n-1} & a_n = 1 \\ a_n - 1 & \text{otherwise.} \end{cases}$

Note that the coefficients of  $g(c)$  are still coefficients in the same right as the original version—they still operate in the same way. That is, we treat them no differently regarding our unary operations. The only difference is that if  $\mathbf{c} = g(c) = [b_0; b_1, \dots, b_m]_+$ , then the convergents for  $c_k$  remain the same when the recurrence relations for  $p_k$  and  $q_k$  are defined for  $b_k$  and  $k < m$ . However,  $c_m$  has some nuance to it.

THEOREM 16. Let  $a_i$  be the quotients obtained by performing the euclidean algorithm on  $\frac{p}{q}$ . If  $\mathbf{c} = [a_0; a_1, \dots, a_n]$  and  $\mathbf{b} = g(c) = [b_0; b_1, \dots, b_m]$ , then  $c = \frac{(b_m(p_{n-1}) + p_{n-2}) + p_{n-1}}{(b_m(q_{n-1}) + q_{n-2}) + q_{n-1}}$ . That is,  $c$  is the mediant of convergents  $c_{m-1}$  and  $c_{m-2}$  as defined for  $b_i$ .

*Proof.* Recall that  $c = c_n = \frac{p_n}{q_n}$ . Then,

$$\begin{aligned} \frac{(b_m(p_{n-1}) + p_{n-2}) + p_{n-1}}{(b_m(q_{n-1}) + q_{n-2}) + q_{n-1}} &= \frac{(a_n - 1)p_{n-1} + p_{n-2} + p_{n-1}}{(a_n - 1)q_{n-1} + q_{n-2} + q_{n-1}} \\ &= \frac{a_n p_{n-1} - p_{n-1} + p_{n-2} + p_{n-1}}{a_n q_{n-1} - q_{n-1} + q_{n-2} + q_{n-1}} = \frac{a_n p_{n-1} + p_{n-2}}{a_n q_{n-1} + q_{n-2}} = \frac{p_n}{q_n} = c \end{aligned}$$

To see the clarification that  $c$  is the mediant of the last two convergents of the coefficients  $\mathbf{b}$ , simply use the fact that  $[b_0; b_1, \dots, b_m] = [a_0; a_1, \dots, a_n - 1]$ . By an application of equation (19) we can deduce that  $c_n \equiv [a_0; a_1, \dots, a_n] = [b_0; b_1, \dots, b_m + 1] \equiv \frac{p_m + p_{m-1}}{q_m + q_{m-1}}$  where  $p_k, q_k$  are defined on the coefficients  $\mathbf{b}$ .  $\square$

While this mapping under  $g$  may only be defined for  $\mathbb{Q}^+$ , this is enough. The Stern-Brocot tree only contains positive rationals. That is, we don't really consider  $F_1$  to be in the tree. What we wanted was a unique representation for some classification of rational numbers and what we have obtained is a representation that will cover the whole Stern-Brocot tree which makes our work with matrices a bit cleaner. Notice that we have not done anything that should negatively impact our use of unary operations. The unary operations were defined on the set of coefficients, and the codomain of  $g$  is still in that set. The interpretation for convergents does not change. We only make the change  $c$  does not equal  $c_m$  but rather  $c$  is the mediant of  $c_m$  and  $c_{m-1}$ . Since our coefficients are still working correctly under  $g$ , we may now build further.

DEFINITION 26. Let  $\phi : \mathbb{Q}^+ \rightarrow SL_2(\mathbb{Z})$  such that  $\phi(c) = h(g(c))$ .

THEOREM 17. Let  $c \in \mathbb{Q}^+$ . If  $\phi(c) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ , then  $c = \frac{a+b}{c+d}$ . That is,

$$\phi^{-1} \left( \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = \frac{a+b}{c+d}.$$

*Proof.* We know that the convergents of  $c$  (as defined under  $g$ ) will make up the column vectors of  $\phi(c)$  so that  $\phi(c)$  is one of  $\begin{bmatrix} p_m & p_{m-1} \\ q_m & q_{m-1} \end{bmatrix}$  and  $\begin{bmatrix} p_{m-1} & p_m \\ q_{m-1} & q_m \end{bmatrix}$ . In either case,  $\frac{p_m + p_{m-1}}{q_m + q_{m-1}} = c$   $\square$

What we now have is a function that lets us travel between the positive rational numbers and Stern-Brocot tree in both directions with unique representations

(because the Stern-Brocot tree has every positive rational number precisely once). To see a demonstration, refer to Example 20 in Appendix C.1.

We leave the reader with something to explore.

DEFINITION 27. Let  $C_n := \{c \in \mathbb{Q}^+ \mid \sigma(g(c)) = n - 1\}$ .  $\phi[C_n]$  is the set of matrix representations of nodes on the  $n_{th}$  level of the Stern-Brocot tree.



# APPENDIX C

## EXAMPLES

*Example 16* (Let  $d = 13$  and note  $3^2 + 2^2 = 13$ ):

$\mathbb{W}\mathbb{A}$  :  $(1, 0, -13)$   
 $R$   $(1, 1, -12)$   
 $R$   $(1, 2, -9)$   
 $R$   $(1, 3, -4)$   
 $L$   $(3, -1, -4)$   
 $R$   **$(3, 2, -3)$**   
 $L$   $(4, -1, -3)$   
 $R$   $(4, 3, -1)$   
 $L$   $(9, 2, -1)$   
 $L$   $(12, 1, -1)$   
 $L$   **$(13, 0, -1)$**   
 $L$   $(12, -1, -1)$   
 $L$   $(9, -2, -1)$   
 $L$   $(4, -3, -1)$   
 $R$   $(4, 1, -3)$   
 $L$   **$(3, -2, -3)$**   
 $R$   $(3, 1, -4)$   
 $L$   $(1, -3, -4)$   
 $R$   $(1, -2, -9)$   
 $R$   $(1, -1, -12)$   
 $R$   $(1, 0, -13)$

The resulting matrix  $N$  after one iteration of  $\mathbb{W}\mathbb{A}$  is  $R^3LRLRL^6RLRLR^3$ . The exponents form a palindrome, but also,  $\sigma(N) = 20$  which is even. Notice that the middle factor has an even exponent.

$$R^3LRLRL^6RLRLR^3 = (R^3LRLRL^3)(L^3RLRLR^3)$$

Now while the resulting two factors on the left of this last expression are not palindromic, they *are* chiral.

$$(R^3LRLRL^3)(L^3RLRLR^3) = [(R^3LR)(LRL^3)] [(L^3RL)(RLR^3)]$$

So, if  $A = R^3LR$  then,  $A^C = RLR^3$ ,  $A^F = L^3RL$ , and  $A^T = (A^F)^C = (L^3RL)^C = LRL^3$ . With these identities, we obtain

$$N = AA^T A^F A^C.$$

Since we know the permutation interpretation of these matrices by the work in Appendix A.2, we can generate a simplified product for  $N$  quite easily. First, find  $A$  :

$$\begin{aligned} A = R^3LR &= \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 7 \\ 1 & 2 \end{bmatrix} \end{aligned}$$

Then we obtain

$$N = \begin{bmatrix} 4 & 7 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 7 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 7 & 4 \end{bmatrix} \begin{bmatrix} 2 & 7 \\ 1 & 4 \end{bmatrix}$$

But, we also know that  $N = (AA^T)(AA^T)^C$  therefore, we can simplify further by finding  $AA^T$ .

$$AA^T = \begin{bmatrix} 4 & 7 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 7 & 2 \end{bmatrix} = \begin{bmatrix} 16 + 49 & 4 + 14 \\ 4 + 14 & 1 + 4 \end{bmatrix} = \begin{bmatrix} 65 & 18 \\ 18 & 5 \end{bmatrix}$$

So,

$$N = \begin{bmatrix} 65 & 18 \\ 18 & 5 \end{bmatrix} \begin{bmatrix} 5 & 18 \\ 18 & 65 \end{bmatrix} = \begin{bmatrix} 649 & 2340 \\ 180 & 649 \end{bmatrix}.$$

Also, notice that from the right hand column of  $AA^T$  we obtain

$$18^2 - 13(5^2) = 324 - 13(25) = 324 - 325 = -1.$$

And from the left hand column of  $N$  we obtain

$$649^2 - 13(180)^2 = 421,201 - 421,200 = 1.$$

*Example 17:* The possible solutions to the equation  $ab - b^2 = -30$  where  $a > 0, c < 0$  are given below. There are 40 solutions. These are generated using the algorithm in Appendix D.3.

(1, 0, -30)	(30, 0, -1)	(2, 0, -15)	(15, 0, -2)
(3, 0, -10)	(10, 0, -3)	(5, 0, -6)	(6, 0, -5)
(1, 1, -29)	(29, 1, -1)	(1, -1, -29)	(29, -1, -1)
(1, 2, -26)	(26, 2, -1)	(1, -2, -26)	(26, -2, -1)
(2, 2, -13)	(13, 2, -2)	(2, -2, -13)	(13, -2, -2)
(1, 3, -21)	(21, 3, -1)	(1, -3, -21)	(21, -3, -1)
(3, 3, -7)	(7, 3, -3)	(3, -3, -7)	(7, -3, -3)
(1, 4, -14)	(14, 4, -1)	(1, -4, -14)	(14, -4, -1)
(2, 4, -7)	(7, 4, -2)	(2, -4, -7)	(7, -4, -2)
(1, 5, -5)	(5, 5, -1)	(1, -5, -5)	(5, -5, -1)

Then, using  $\mathbb{W}\mathbb{A}$  by running the code in Appendix D.4 we will list the different paths generated by  $\mathbb{W}\mathbb{A}$  that contain these solutions. The bottom row will contain the resulting matrices composed of left and right steps. Each path is arranged as a column for easy comparison so they read from top to bottom, not left to right. Notice that in this case, there is no indicated equivalence between (1, 0, -30) and (30, 0, -1). Instead, the orbit containing (30, 0, -1) has been absorbed into the equivalence for (5, 0, -6).

$\mathbb{W}\mathbb{A}(1, 0, -30)$	$\mathbb{W}\mathbb{A}(2, 0, -15)$	$\mathbb{W}\mathbb{A}(3, 0, -10)$	$\mathbb{W}\mathbb{A}(5, 0, -6)$	$\mathbb{W}\mathbb{A}(30, 0, -1)$
$R(1, 1, -29)$	$R(2, 2, -13)$	$R(3, 3, -7)$	$R(5, 5, -1)$	$L(29, -1, -1)$
$R(1, 2, -26)$	$R(2, 4, -7)$	$L(2, -4, -7)$	$L(14, 4, -1)$	$L(26, -2, -1)$
$R(1, 3, -21)$	$L(3, -3, -7)$	$R(2, -2, -13)$	$L(21, 3, -1)$	$L(21, -3, -1)$
$R(1, 4, -14)$	$R(3, 0, -10)$	$R(2, 0, -15)$	$L(26, 2, -1)$	$L(14, -4, -1)$
$R(1, 5, -5)$	$R(3, 3, -7)$	$R(2, 2, -13)$	$L(29, 1, -1)$	$L(5, -5, -1)$
$L(6, 0, -5)$	$L(2, -4, -7)$	$R(2, 4, -7)$	$L(30, 0, -1)$	$R(5, 0, -6)$
$L(1, -5, -5)$	$R(2, -2, -13)$	$L(3, -3, -7)$	$L(29, -1, -1)$	$R(5, 5, -1)$
$R(1, -4, -14)$	$R(2, 0, -15)$	$R(3, 0, -10)$	$L(26, -2, -1)$	$L(14, 4, -1)$
$R(1, -3, -21)$			$L(21, -3, -1)$	$L(21, 3, -1)$
$R(1, -2, -26)$			$L(14, -4, -1)$	$L(26, 2, -1)$
$R(1, -1, -29)$			$L(5, -5, -1)$	$L(29, 1, -1)$
$R(1, 0, -30)$			$R(5, 0, -6)$	$L(30, 0, -1)$
$\begin{bmatrix} 11 & 60 \\ 2 & 11 \end{bmatrix}$	$\begin{bmatrix} 11 & 30 \\ 4 & 11 \end{bmatrix}$	$\begin{bmatrix} 11 & 20 \\ 6 & 11 \end{bmatrix}$	$\begin{bmatrix} 11 & 12 \\ 10 & 11 \end{bmatrix}$	$\begin{bmatrix} 11 & 2 \\ 60 & 11 \end{bmatrix}$

### C.1. Examples for appendices.

*Example 18:* Notice how we remove zeros: We first remove the trailing zeros, then we add coefficients to the left and right of an intermediate zero, starting from the rightmost instance.

$$[0; 0, 2, 1, 0, 4, 1, 0, 0]$$

$$[0; 0, 2, 1, 0, 4, 1]$$

$$[0; 0, 2, 1 + 4, 1]$$

$$[2; 0, 2, 5, 1]$$

$$[2; 0 + 2, 5, 1]$$

$$[2; 5, 1]$$

*Example 19:* The reader may recall that in Example 7 of page 20, there was some ambiguity when we tried to map to  $\frac{5}{2} \equiv [2; 2] \equiv [2; 1, 1]$ . However, under the rule of  $\phi$  we have  $\phi\left(\frac{5}{2}\right) = h(g(c)) = h([2; 1]) = R^2L$  which is precisely what we would like in order to navigate to precisely  $\frac{5}{2}$  in the Stern-Brocot tree.

*Example 20:* Consider  $\frac{17}{47}$ .

$$17 = 47(0) + 17$$

$$47 = 17(2) + 13$$

$$17 = 13(1) + 4$$

$$13 = 4(3) + 1$$

$$4 = 1(4) + 0$$

By taking each of these quotients, we can deduce that  $g\left(\frac{17}{47}\right) = [0; 2, 1, 3, 3]$  and therefore  $\phi\left(\frac{17}{47}\right) = h\left(g\left(\frac{17}{47}\right)\right) = R^0L^2RL^3R^3 = L^2RL^3L^3$ .

$$L^2RL^3R^3 = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}.$$

With a bit of calculator magic, we arrive at

$$\phi\left(\frac{17}{47}\right) = \begin{bmatrix} 4 & 13 \\ 11 & 36 \end{bmatrix}$$

and of course

$$\phi^{-1}\left(\begin{bmatrix} 4 & 13 \\ 11 & 36 \end{bmatrix}\right) = \frac{4 + 13}{11 + 36} = \frac{17}{47}.$$

## APPENDIX D

### PROVIDED CODE

Herein the reader will find some humble scripts for computing various things related to this paper.

**D.1. Coefficients of  $\sigma$ - $n$ .** For a given positive integer  $m$  we may generate all coefficients  $\mathbf{c}$  such that  $\sigma(\mathbf{c}) = m$  using the code below.

```
#Recursively generates all possible  
#collections of coefficients  
#summing to n.  
def permutations(n):  
    #base case  
    if (n==1):  
        return [[1]]  
    #smaller-caller  
    else :  
        collection = []  
        for i in range(n):  
            sub = permutations(i)  
            for s in sub:  
                collection.append([n-i] + s)
```



```

    return collection

#Set m to be the desired value for sigma.
m = 5
collection = permutations(m)
"""

#Enabling this block will add the reciprocal expressions
#to the list.
reciprocals = []
for p in collection:
    reciprocals.append([0] + p)
collection = collection + reciprocals
"""

for p in collection:
    print(p)

```

**D.2. Farey set generator.** The following program will generate a directory of Farey sets  $F_n$ .

```

#Generates mediants of each tier
#of Farey fractions and returns
#the mth set.

def farey(m):
    #Define the first level of Farey fractions.

    f = [(0, 1), (1, 1)]

```

```

#Enable the following line to generate the extended
#Farey fractions covering the interval [0, infinity]:

#f = [(0, 1), (1, 0)]

#Define a map for each Farey set starting at F1.
farey = {1:f}
for i in range(1, m):
    g = []
    #Add the mediant of each pair of the current
    #Farey set to the list, 'g.'
    for j in range(len(f)-1):
        g.append((f[2*j][0] + f[2*j+1][0],
                 f[2*j][1] + f[2*j+1][1]))
        #For each mediant generated, insert it where
        #it belongs in the cumulative farey set 'f.'
        f.insert(2*j+1, (f[2*j][0] + f[2*j+1][0],
                        f[2*j][1] + f[2*j+1][1]))
    farey[i+1] = g
return farey[m]

#Change m to set how many levels of fractions to generate.
m = 5
for f in farey(m):
    print(f)

```

**D.3. Generating solutions for  $-ac - b^2 = -D$ .** For a given positive integer  $D$ , the following program will generate a list of solutions to the equation  $ac - b^2 = -D$  where  $a > 0, c < 0$ . The algorithm does this by the equivalent logic for solving  $ac + b^2 = D$  where  $a, c$  are positive.

```

from math import ceil
from math import floor
from math import sqrt
def abc_solver(D):
    solutions = []
    #b must be less than or equal to the square root of D.
    root_D = ceil(sqrt(D))
    #We iterate through the finite possible values of b.
    for b in range(root_D):
        d_i = D - b**2

        #Rearranging the equation to ac = D - b^2, we may
        #find solutions by factoring D-b^2 = d_i into
        #integers a and c.

        #Checking up to the floor of the square root of d_i
        #is sufficient.
        root_d_i = floor(sqrt(d_i))
        a = 1

        while(a <= root_d_i):

```

```

if(d_i % a == 0):
    c = d_i//a
    solutions.append((a, b, -c))
    #This step is skipped if a = c to
    #avoid redundancy.
    if(c != a):
        #For each value of c = d_i/a, we can get
        #another solution from a = d_i/c.
        solutions.append((c, b, -a))
    #This step is skipped if b is zero
    #to avoid redundancy.
    if(b != 0):
        #We obtain another solution for a, c
        #when b is not zero by
        #replacing b with its additive inverse
        #since (-b)^2 = b^2.
        solutions.append((a, -b, -c))
        if(c != a):
            solutions.append((c, -b, -a))
    a +=1
return solutions
#Set D to the desired value.
D = 14
for s in solutions:
    print(s)

```

**D.4. The Wildberger algorithm.** Below is the code implementation for the algorithm introduced by Norman Wildberger. We may run any balanced form as a seed and get the list of equivalent forms traversed as well as the summary of left and right steps taken. The reader might find it of interest to run each of the solutions found by the algorithm above for a given  $D$ . Doing so, one may find there are some collections grouped by sequence of steps in a loop. To do this, we must be sure we have access to both functions in a file. Then, iterate through the resulting list from `abc_solver()` and run each tuple as a seed to the function below.

```
#Q should be of the form Q = (a, b, c)
#If Q is not balanced, the algorithm may not terminate;
#therefore, be sure to select a>0 and c<0. Or, explore the
#cases when the algorithm will terminate
#for an unbalanced seed.

#Set n according to how many iterations of N you would like
#to compute.
def WBA(Q, n = 1):

    #Some accessory and initialization code.

    #Q recorded as starting point.
    start = Q
    stones = [Q]

    #An empty string to record left and right decisions.
```

```

#Useful for feeding into other algorithms.
path = ""

#First line of console output which logs the algorithm.
#If computing long paths, it may speed up processing to
#disable print statements.
print("\\n>>start:\\n" + " _ _ %s" % str(Q) )

#Counter for tracking iterations.
count = 0

#Store the starting quadratic form in a map.
Q_p = {'a': Q[0], 'b': Q[1], 'c': Q[2]}
form = Q
while(count < n):
    #Compute the total.
    T = Q_p['a'] + (2 * Q_p['b']) + Q_p['c']

    #If T>0, take a left step.
    if(T > 0):
        #Take a left step and store the resulting
        #form in the map.
        Q_p = {'a':Q_p['a'] + 2*Q_p['b'] + Q_p['c'],
              'b':Q_p['b'] + Q_p['c'], 'c':Q_p['c']}

        #Record the step taken in the path string.

```

```

path = path + "L"

#Express the new form as a tuple and print
#it to the console with the step taken.
form = (Q_p['a'], Q_p['b'], Q_p['c'])
print("L", form)

#If T<0, take a right step.
elif(T < 0):
    #Same logic at as for a left step.

    Q_p = {'a': Q_p['a'], 'b':Q_p['a'] + Q_p['b'],
           'c':Q_p['a'] + 2*Q_p['b'] + Q_p['c']}

    path = path + "R"

    #Console output
    form = (Q_p['a'], Q_p['b'], Q_p['c'])
    print("R", form)

else:
    #This should only happen if the seed form has a
    #determinant that is a perfect square.
    print("T=_0_or_T_is_not_a_number.")

#Record the tuple representing the current

```

```

    #quadratic form
    stones.append(form)

    if(start == form):
        count = count + 1

    #Enable this line to pause at the end of
    #each iteration of the while loop:
    #input("Press return to step.")

#Returns a string of characters 'L' and 'R' to show the
#path taken.
    return path

#Set the seed quadratic form and print the results by calling
#the function. Below are three examples using balanced forms.
Q = (1, 0, -14)
print(WBA(Q))

Q = (5, 2, -5)
print(WBA(Q))

Q = (3, 0, -7)
print(WBA(Q))

```