Benford's Law in Linear Recurrences and Continued Fractions

Michelle He, Brian Huang, Aadeesh Shastry, Mirilla Zhu

August 9, 2017

1 Introduction

Given a sequence of real numbers, it seems reasonable to expect that the first digits of its terms will occur with equal frequencies. However, in reality, this is far from the case: in many naturally occurring sequence of numbers, it is more likely for the leading digit to be smaller. Benford's law provides a precise mathematical formulation for this phenomenon of first digit bias:

Definition 1. Let $\{a_n\}$ be a sequence of real numbers and $a_n = 10^k m$ where $k \in \mathbb{Z}$ and $m \in \mathbb{R}$ such that $1 \leq m < 10$. Furthermore, let $M(a_n) = m$. If

$$\mathbb{P}\{1 \leq M(a_n) \leq d\} = \log_{10}(d)$$

for all $1 \leq d < 10$, then $\{a_n\}$ satisfies Benford's law. If $\{a_n\}$ meets this condition, then we call $\{a_n\}$ a Benford sequence.

In particular, Benford's law states that the frequency of terms in a sequence with leading digit d will be

$$\mathbb{P}\{d \leq M(a_n) \leq d+1\} = \mathbb{P}\{1 \leq M(a_n) \leq d+1\} - \mathbb{P}\{1 \leq M(a_n) \leq d\} = \log_{10}\left(1 + \frac{1}{d}\right).$$

In this paper, we explore whether certain recurrence relations satisfy Benford's law. We prove that the Fibonacci sequence follows Benford's law, generalizing our findings to other families of linear recurrences. We investigate if the terms and convergents of certain simple continued fractions follow Benford's Law. We find an explicit formula for the recurrence relation of the numerators and denominators of the convergents of quadratic irrationals to show these sequences follow Benford's Law. Furthermore, we make use of the Gauss-Kuzmin distribution and the Levy's constant to make statements about all simple continued fractions, except for a set of measure zero, concerning Benford's Law. Lastly, we investigate the simple continued fractions of certain transcendentals to make conjectures related to Benford's Law.

2 Equidistribution and Benford's Law

We begin by establishing some preliminary results on Benford's law that will allow us to prove whether certain sequences are Benford. One particularly useful concept that relates to Benford's law is that of equidistribution, which measures whether a sequence is evenly distributed within a given interval.

Definition 2. A sequence $\{a_n\}$ is equidistributed modulo 1 if for any subinterval $[c, d] \in [0, 1)$,

$$\lim_{N \to \infty} \frac{\#\{a_n \in [c, d] : n \leq N\}}{N} = d - c.$$

Theorem 1 ([KM05])**.** *The sequence $\{\log a_n\}$ is equidistributed modulo 1 if and only if the sequence $\{a_n\}$ follows Benford's law.*

In order to prove that a sequence $\{a_n\}$ follows Benford's law, Proposition 1 tells us that we only need to show that the logs of the terms are equidistributed modulo 1. There exist two useful theorems that tell us when sequences are equidistributed in such a manner, whose proofs can be found in [KN74]:

Theorem 2 (Weyl's Criterion)**.** *The sequence $\{a_n\}$ is equidistributed modulo 1 if and only if for all nonzero integers h,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h a_n} = 0.$$

Theorem 3 (Difference Theorem)**.** *If $\lim_{n\to\infty} a_n - a_{n-1} = \alpha$, where α is irrational, then $\{a_n\}$ is equidistributed modulo 1.*

The difference theorem is particularly useful in form of the following corollaries:

Corollary 1. *For a given sequence $\{a_n\}$, if $\lim_{n\to\infty} \log a_n - \log a_{n-1} = \alpha$, where α is irrational, then $\{\log a_n\}$ is equidistributed modulo 1.*

Proof. The proof is immediate from substituting $\log a_n$ for a_n. ■

Corollary 2. *For any irrational α, the sequence $\{a_n = n\alpha\}$ is equidistributed modulo 1.*

Proof. If $a_n = n\alpha$, then $\lim_{n\to\infty} n\alpha - (n-1)\alpha = \alpha$, so $\{a_n\}$ is equidistributed modulo 1. ■

When given a certain sequence, it is often convenient to consider related sequences that are easier to work with. Using Weyl's criterion, we can determine whether a sequence follows Benford's law given that a related sequence is Benford. We first show that the sequence created by multiplying each term of a Benford sequence by a constant creates another Benford sequence:

Proposition 1. *If the sequence a_n is equidistributed modulo 1, then the sequence $b_n = a_n + C$ is equidisitributed modulo 1.*

Proof. We compute

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h b_n} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h(a_n + C)} = \left(e^{2\pi i h C}\right) \left(\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h a_n}\right).$$

Since a_n is equidistributed modulo 1, we can substitute

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h a_n} = 0$$

to obtain

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h b_n} = \left(e^{2\pi i h C}\right)(0) = 0,$$

as desired. Thus b_n satisfies Weyl's criterion and is equidistributed modulo 1. ■

Proposition 2. *Let $\{a_n\}$ be a sequence following Benford's law. Then if $\{b_n\} = C\{a_n\}$ where $C > 0$, $\{b_n\}$ follows Benford's law.*

Proof. If $\{b_n\}_{n=1}^{\infty} = C\{a_n\}_{n=1}^{\infty}$, then

$$\{\log b_n\}_{n=1}^{\infty} = \log(C\{a_n\}_{n=1}^{\infty}) = \log C + \{\log a_n\}_{n=1}^{\infty}.$$

Since $\{a_n\}_{n=1}^{\infty}$ follows Benford's law, $\{\log a_n\}_{n=1}^{\infty}$ is equidistributed modulo 1. By Proposition 2, $\{\log b_n\}_{n=1}^{\infty}$ is equidistributed modulo 1 and thus $\{b_n\}_{n=1}^{\infty}$ follows Benford's law. ■

In addition to considering sequences whose individual terms are based upon a Benford sequence, we can show that sequences that differ from a Benford sequence by a finite number of terms remain Benford.

Proposition 3. *Consider a sequence $\{b_n\}$ formed by adding or removing a finite number of terms from the sequence $\{a_n\}$. Then $\{a_n\}$ follows Benford's law if and only if $\{b_n\}$ follows Benford's law.*

Proof. Without loss of generality, let $\{c_n\}_{n=1}^{k}$ be removed and let $\{d_n\}_{n=1}^{j}$ be added to $\{a_n\}_{n=1}^{\infty}$ to form $\{b_n\}_{n=1}^{\infty}$. Evaluating the sum with the Weyl criterion for $\{b_n\}_{n=1}^{\infty}$, we obtain

$$\begin{aligned}
\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e^{2\pi ih\log b_n} &= \lim_{N\to\infty}\frac{1}{N}\left(\sum_{n=1}^{N}e^{2\pi ih\log a_n}-\sum_{n=1}^{k}e^{2\pi ih\log c_n}+\sum_{n=1}^{j}e^{2\pi ih\log d_n}\right)\\
&= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e^{2\pi ih\log a_n}-\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{k}e^{2\pi ih\log c_n}+\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{j}e^{2\pi ih\log d_n}\\
&= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e^{2\pi ih\log a_n}.
\end{aligned}$$

From here, we see that $\{\log a_n\}_{n=1}^{\infty}$ being equidistributed modulo 1 implies that $\{\log b_n\}_{n=1}^{\infty}$ is equidistributed modulo 1, and vice versa. Thus, $\{b_n\}_{n=1}^{\infty}$ follows Benford's law if and only if $\{a_n\}_{n=1}^{\infty}$ follows Benford's law. ■

Furthermore, a combination of Benford sequences creates another Benford sequence. This technique will be especially useful when we consider the recurrence relations of continued fractions, which we will partition into subsequences satisfying linear recurrences.

Definition 3. Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of infinite length. The sequences $s_1, s_2, \ldots, s_k$ form a partition of $\{a_n\}$ if each term of $\{a_n\}$ is contained in exactly one of $s_1, s_2, \ldots, s_k$.

Proposition 4. *If $s_1, s_2, \ldots, s_k$ form a partition of $\{a_n\}_{n=1}^{\infty}$ and each of $s_1, s_2, \ldots, s_k$ satisfies Benford's law, then $\{a_n\}_{n=1}^{\infty}$ satisfies Benford's law.*

Proof. To prove our result, we evaluate Weyl's criterion:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e^{2\pi ih\log a_n}=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{k}\sum_{n=1}^{N}e^{2\pi ih\log s_{i,n}}=\sum_{i=1}^{k}\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e^{2\pi ih\log s_{i,n}}$$

Since every s_i satisifies Benford's law, $\log s_i$ is equidistributed modulo 1 for all i from 1 to k, and thus

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e^{2\pi ih\log s_{i,n}}=0$$

for every s_i. As a result,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e^{2\pi ih\log a_n}=0,$$

so $\log a_n$ satisfies Weyl's criterion and is equidistributed modulo 1. Thus $\{a_n\}$ satisfies Benford's law. ■

Lastly, we demonstrate that a sequence whose limit approaches a Benford sequence is also Benford.

Proposition 5. *Let $\{a_n\}_{n=1}^{\infty}$ be a sequence follows Benford's law. Furthermore, let $\{b_n\}_{n=1}^{\infty}$ be a sequence such that $\lim_{n\to\infty} b_n = a_n$. Then $\{a_n\}_{n=1}^{\infty}$ also follows Benford's law.*

Proof. Let $b_n - a_n = \epsilon_n$. We have

$$\lim_{n\to\infty}\epsilon_n = 0, b_n = a_n + \epsilon_n.$$

When n is sufficiently large, $b_n \approx a_n$ and the first digits of the sequences are the same for most n, unless $a_n, b_n \approx 10^x$ for some $x \in \mathbb{N}$. Thus, there are a finite number of terms in b_n which do not have the same leading digit as the corresponding term in a_n. By Proposition 3 these terms and their corresponding terms in a_n can be removed without changing whether or not the sequences follow Benford's law. Thus, $\{a_n\}_{n=1}^{\infty}$ is still a Benford sequence, and every term in the corresponding $\{b_n\}_{n=1}^{\infty}$ sequence has the same first digit, and thus $\{b_n\}_{n=1}^{\infty}$ also follows Benford's law. ■

3 Linear Recurrences

From Proposition 2 and Corollary 2 to the difference theorem, we see that all geometric sequences of the form ar^n satisfy Benford's law when r is irrational. The Benford nature of geometric sequences motivates us to investigate whether linear recurrences of the form $a_n = c_1a_{n-1} + c_2a_{n-2} + ... + c_ka_{n-k}$ for constants c_i follow Benford's law, because their closed form consists of sums of geometric series:

Definition 4. The characteristic polynomial of the recurrence $a_n = c_1a_{n-1} + c_2a_{n-2} + ... + c_ka_{n-k}$ is the polynomial

$$r^k - \sum_{i=1}^{k} c_i r^{k-i}$$

Proposition 6. *Each linear recurrence $\{a_n\}$ of order k can be written in a closed form. If the roots r_i of the characteristic polynomial for a_n are all distinct, then $a_n = C_1r_1^n + C_2r_2^n + ... + C_kr_k^n$ for some constants C_i determined by the initial values of the recurrence. If the characteristic polynomial has repeated roots $r_1 = r_2 = ... = r_\gamma$, then $a_n = r_\gamma^n \sum_{i=1}^{\gamma} C_i n^{\gamma-i} + \sum_{i=\gamma+1}^{k} C_k r_i^n$.*

3.1 Fibonacci Numbers

Proposition 7. *The Fibonacci sequence with terms $F_n = F_{n-1}+F_{n+2}$ and $F_0 = 0, F_1 = 1$ satisfies Benford's law.*

Proof. Let F_n be the nth Fibonacci number. We have

$$\log d \leq \text{frac}(\log F_n) \leq \log(d+1),$$

where d denotes the leading digit of F_n.

It is well known that the closed form formula for the sequence F_n is

$$F_n = \frac{1}{\sqrt{5}}(\phi^n - (\frac{1-\sqrt{5}}{2})^n).$$

From this, we see that

$$\begin{aligned}
\log F_n &= \log\frac{1}{\sqrt{5}} + \log(\phi^n + (\frac{1-\sqrt{5}}{2})^n) \\
&= \log\frac{1}{\sqrt{5}} + \log\phi^n + \log(1 + (\frac{1-\sqrt{5}}{1+\sqrt{5}})^n) \\
&= n\log\phi + c + E,
\end{aligned}$$

where $c = \log\frac{1}{\sqrt{5}}$ and $E = \log(1 + (\frac{1-\sqrt{5}}{1+\sqrt{5}})^n)$, with $\lim_{n\to\infty} E = 0$.

Using Weyl's criterion, we want to verify whether

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi ik\text{frac}(\log F_n)} = 0.$$

As $n \to \infty$, $\text{frac}(\log F_n) \to \text{frac}(n\log\phi)$. Since $\log\phi$ is irrational, it is equidistributed by Corollary 2 to the difference theorem. Then $\sum_{n=1}^{N} e^{2\pi ik\text{frac}(\log F_n)}$ stops growing as $N \to \infty$, but the denominator N continues to grow. Therefore Weyl's criterion is satisfied and the Fibonacci sequence is Benford. ■

3.2 Linear Recurrences with a Maximal Root

We can generalize our findings about Benford's Law for Fibonacci numbers to other linear recurrences with a maximal root.

Proposition 8. *Let* $a_n = c_1 r_1^n + c_2 r_2^n + ... + c_k r_k^n$*, where* $|r_1| > |r_i|$ *for all* $2 \le i \le k$*. If* $\log r_1$ *is irrational then* $\{a_n\}$ *satisfies Benford's law.*

Proof. We have

$$\log d \le \{\log a_n\} \le \log(d+1)$$

We can write that

$$\{\log a_n\} = \{\log(c_1 r_1^n + c_2 r_2^n + ... + c_k r_k^n)\}$$

$$= \{\log(c_1 r_1^n) + \log(1 + \frac{c_2}{c_1}\frac{r_2}{r_1}^n + ... + \frac{c_k}{c_1}\frac{r_k}{r_1}^n)\}$$

$$= \{n \log r_1 + \log c_1 + E\}$$

where $E = \log(1 + \frac{c_2}{c_1}\frac{r_2}{r_1}^n + ... + \frac{c_k}{c_1}\frac{r_k}{r_1}^n)$. We can see that $|\frac{r_i}{r_1}| < 1$ for all i such that $2 \le i \le k$, so

$$\lim_{n\to\infty} \frac{r_i}{r_1}^n = 0 \to \lim_{n\to\infty} E = 0$$

Using Weyl's criterion, we want to verify whether

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i k \{\log a_n\}}$$

As $n \to \infty$, $\{\log a_n\} = \{n \log r_1 + \log c_1 + E\} \to \{n \log r_1 + \log(c_1)\}$ which is known to be equidistributed (the constant $\log c_1$ does not affect equidistribution) because $logr_1$ is irrational. Then $\sum_{n=1}^{N} e^{2\pi i k\{\log a_n\}}$ stops growing as $N \to \infty$, but the denominator N continues to grow. Therefore Weyl's criterion is satisfied, and $\{a_n\}$ satisfies Benford's law. ■

Proposition 9. *Let* $a_n = \sum_{i=1}^{k} \beta_i \alpha_i^n$ *where* $|\alpha_1| \le |\alpha_2| \le |\alpha_3| \le \ldots \le |\alpha_{k-1}| < |\alpha_k|$ *and* $\log \alpha_k$ *is irrational. Furthermore, let* $P(n) = \sum_{i=1}^{j} \gamma_i n^i$ *where* $\gamma_j \ne 0$*. Then* $b_n = P(n)a_n$ *follows Benford's law.*

Proof. We proceed by showing $\log b_n$ is equidistributed modulo 1:

$$\lim_{n\to\infty} \log b_n - \log b_{n-1} = \log \frac{b_n}{b_{n-1}} = \lim_{n\to\infty} \log \frac{a_n P(n)}{a_{n-1}P(n-1)} = \lim_{n\to\infty} \log \frac{a_n}{a_{n-1}} = \log \alpha_k.$$

Since $\log \alpha_k$ is irrational, the sequence $\log b_n$ is equidistributed modulo 1, so b_n follows Benford's law. ■

3.3 Positive Second Order Linear Recurrences

Even when the closed form of the recurrence does not have a maximal root, the recurrence can still satisfy Benford's law in some cases. We proceed to show that any positive second-order linear recurrence follows Benford's law, whether its roots are distinct or not.

Proposition 10. *Let $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ where $c_1, c_2 \in \mathbb{R}$ and $a_n \in \mathbb{R}, a_n > 0$. Then a_n follows Benford's law.*

Proof. Consider the roots, r_1, r_2 of the characteristic polynomial of a_n. We will show that $r_1, r_2 \in \mathbb{R}$ through proof by contradiction. Assume that $r_1, r_2 \notin \mathbb{R}$. By the complex conjugate root theorem, $r_2 = \overline{r_1}$. Let $r_1 = re^{i\theta}$ and $r_2 = re^{-i\theta}$ where $r = |r_1| = |r_2|$ and $\theta \neq 0$. Then $a_n = \alpha_1 r_1^n + \alpha_2 r_2^n$. Since $a_n \in \mathbb{R}$, $\alpha_1 r^n e^{in\theta} + \alpha_2 r^n e^{-in\theta} \in \mathbb{R}$, which implies $\mathrm{Im}(\alpha_1 r^n e^{in\theta} + \alpha_2 r^n e^{-in\theta}) = 0$. Since this means $r^n(\alpha_1 - \alpha_2) \sin n\theta = 0$, we have $\alpha_1 = \alpha_2$. As a result, a_n can be written as $a_n = 2\alpha_1 r^n \cos n\theta$.

But since $a_n > 0$, $\cos n\theta > 0$ for all natural numbers n. Thus $2\pi k - \frac{\pi}{2} < n\theta < 2\pi k + \frac{\pi}{2}$ for some $k \in \mathbb{Z}$. Since $\cos n\theta > 0$ for all natural numbers n, $\cos \theta > 0$, which implies that $\frac{-\pi}{2} < \theta < \frac{\pi}{2}$. Suppose $0 < \theta < \frac{\pi}{2}$. Let j be the largest natural number such that $0 < j\theta < \frac{\pi}{2}$ Then $\frac{\pi}{2} \leq (j+1)\theta < \pi$ which implies $-1 < \cos(j+1)\theta \leq 0$ implying $\cos n\theta \not> 0$ for all natural numbers n, a contradiction. A similar argument holds if $\frac{-\pi}{2} < \theta < 0$. Thus $\cos n\theta$ cannot be positive for all natural numbers n and $\theta \neq 0$ and we have a contradiction. Thus $r_1, r_2 \in \mathbb{R}$.

If $r_1 = r_2$, then $a_n = (\alpha_1 + \alpha_2) r_1^n$ which follows Benford's law by Proposition 3.2. If $r_1 = -r_2$ then $a_n = (\alpha_1 + \alpha_2) r_1^n$ for even n and $a_n = (\alpha_1 - \alpha_2) r_1^n$ for odd n. Since $a_n > 0, a_n \neq 0$ and thus $\alpha_1 + \alpha_2 \neq 0, \alpha_1 - \alpha_2 \neq 0$ and a_n is a combination of two Benford sequences. By Proposition 4, a_n follows Benford's law. All other cases are covered in Proposition 8. Thus a_n follows Benford's law if $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ where $c_1, c_2 \in \mathbb{R}$ and $a_n \in \mathbb{R}, a_n > 0$. ■

4 Continued Fractions

An interesting variation on linear recurrences is the sequence of terms and convergents of a continued fraction $\alpha \in \mathbb{R}$. The continued fraction of $\alpha = [a_1, a_2, a_3...]$ takes the form

$$\alpha = a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\cdots}}}$$

where each natural number a_i represents the greatest integer within its corresponding fraction. Thus, we see that the sequence of terms $\{a_n\}$ of the continued fraction is given by the recurrence

$$a_n = \left\lfloor \frac{1}{\mathrm{frac}(a_n)} \right\rfloor.$$

Closely related to the terms $a_1, a_2, ..., a_n$ are the values P_n and Q_n, which respectively represent the numerator and denominator of the nth convergent $[a_1, a_2, a_3, ..., a_n]$ of α. They, too, can be calculated through well-known recursions:

$$P_n = a_n P_{n-1} + P_{n-2} \text{ where } P_0 = 1 \text{ and } P_1 = a_1$$

$$Q_n = a_n Q_{n-1} + Q_{n-2} \text{ where } Q_0 = 0 \text{ and } Q_1 = 1$$

While the recursions for $\{P_n\}$ and $\{Q_n\}$ are not linear since the terms a_i can vary, their general structure resembles the second-order linear recurrences discussed in the previous section. In fact, as we will see later, we can use these recurrences to create formulas for subsequences of $\{P_n\}$ and $\{Q_n\}$ that *are* second-order linear recurrences. For now, however, we will use them to establish another recurrence that relates the terms of a continued fraction to its convergents:

Proposition 11. *Let $\alpha = [a_1, a_2, \ldots, a_n]$ where n can be ∞. Then*

$$\prod_{i=1}^{k} \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} P_k & P_{k-1} \\ Q_k & Q_{k-1}. \end{pmatrix}$$

Proof. We will prove this using the principle of induction. For the base case, $k = 1$,

$$\prod_{i=1}^{1} \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} P_1 & P_0 \\ Q_1 & Q_0 \end{pmatrix}$$

Thus,

$$\prod_{i=1}^{k} \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} P_k & P_{k-1} \\ Q_k & Q_{k-1} \end{pmatrix}$$

is true for the base case. Assume it is true for $k = j$.We will show it is also true for $k = j + 1$.

$$\prod_{i=1}^{j+1} \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} = \left(\prod_{i=1}^{j} \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} \right) \begin{pmatrix} a_{j+1} & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} P_j & P_{j-1} \\ Q_j & Q_{j-1} \end{pmatrix} \begin{pmatrix} a_{j+1} & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} P_j a_{j+1} + P_{j-1} & P_j \\ Q_j a_{j+1} + Q_{j-1} & Q_j \end{pmatrix}$$

Using the recursions for P_k and Q_k, this matrix is simply

$$\begin{pmatrix} P_{j+1} & P_j \\ Q_{j+1} & Q_j \end{pmatrix}$$

Thus

$$\prod_{i=1}^{k} \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} P_k & P_{k-1} \\ Q_k & Q_{k-1} \end{pmatrix}$$

holds for $k = j + 1$ and by the principle of induction holds for all $k \in \mathbb{N}$. ■

By considering these recursions, we proceed to investigate whether the terms and convergents of different classes of continued fractions satisfy Benford's law.

4.1 Continued Fractions of Rationals

The continued fractions of rational numbers are among the simplest to work with, since they have a finite number of terms. Here, we see how this finiteness affects whether their terms and convergents satisfy Benford's law.

Proposition 12. *If α is rational, then its continued fraction expansion is finite.*

Proof. Let $\alpha = \frac{P}{Q}$ for relatively prime $P, Q \in \mathbb{N}$. To construct the terms $a_1, a_2, ..., a_n$ of α, we apply Euclid's algorithm:

$$\begin{aligned} P &= a_1 Q + r_1 \\ Q &= a_2 r_1 + r_2 \\ &\vdots \\ r_n &= a_n r_{n+1} + 0. \end{aligned}$$

Because there are a finite number of natural numbers less than P by well-ordering, Euclid's algorithm will terminate, resulting in a finite number of terms a_i. ■

Proposition 13. *If α is rational, then the terms of its continued fraction $[a_1, a_2, ..., a_n]$ do not satisfy Benford's law.*

Proof. Let α be rational so that it has a finite number of terms in its continued fraction expansion by Proposition 12. Let there be n terms in the continued fraction expansion of α, and of these n terms, let j have a leading digit of 1. Then the probability of 1 being the leading digit of a term in the sequence $\{a_i\}_{i=1}^n$ is $\frac{j}{n}$ which is rational. But by Benford's law, the probability that 1 is the leading digit of a term in a Benford sequence is $\log 2$, which is irrational. Thus $\{a_i\}_{i=1}^n$ does not follow Benford's law. ■

Since rational α have a finite number of terms in their continued fractions, they will also have a finite number of convergents. By similar reasoning to Proposition 13, it follows that the finite sequences $\{P_n\}$ and $\{Q_n\}$ corresponding to such α also fail to satisfy Benford's law. Thus, if P_n and Q_n were to satisfy Benford's Law, α would have to have a continued fraction expansion with an infinite number of terms. Since both P_n and Q_n do not follow Benford's Law for rationals, a natural question would be if a necessary condition for P_n to follow Benford's law would be Q_n following Benford's Law. This statement is in fact true.

Proposition 14. *$\{P_n\}_{n=1}^{\infty}$ follows Benford's Law if and only if $\{Q_n\}_{n=1}^{\infty}$ follows Benford's Law.*

Proof. We first prove the forward direction wanting to show that $\{P_n\}_{n=1}^{\infty}$ if $\{Q_n\}_{n=1}^{\infty}$ satisfies Benford's Law. We have that

$$\lim_{n\to\infty} \frac{P_n}{Q_n} = \alpha$$

which means that

$$\lim_{n\to\infty} P_n = Q_n\alpha$$

But since $\{Q_n\}_{n=1}^{\infty}$ follows Benford's Law, by Proposition 2 $\{Q_n\alpha\}_{n=1}^{\infty}$ follows Benford's law and by Proposition 5 $\{P_n\}_{n=1}^{\infty}$ follows Benford's Law. Similarly for the reverse direction, we have

$$\lim_{n\to\infty} P_n\frac{1}{\alpha} = Q_n$$

and the proof is the same as the forward direction.

■

4.2 Continued Fractions of Quadratic Irrationals

Quadratic irrationals α, which satisfy the equation $a\alpha^2 + b\alpha + c = 0$, are similar to rationals in the sense that both of their continued fractions have a finite number of different terms. Lagrange proved that the continued fractions of all quadratic irrationals are periodic; that is, they can be written in the form $\alpha = [a_1, a_2, \ldots, a_k, \overline{a_{k+1}, a_{k+2}, \ldots, a_{n+k}}]$. As it turns out, this fact is key to determining whether the terms and the convergents of such irrationals follow Benford's law.

Proposition 15. *If α is a quadratic irrational, then the terms of its continued fraction do not satisfy Benford's law.*

Proof. Since α is a quadratic irrational, $\alpha = [a_1, a_2, \ldots, a_k, \overline{a_{k+1}, a_{k+2}, \ldots, a_{n+k}}]$. Consider the sequence $\{a_i\}_{i=k+1}^{n+k}$. Let j of these terms have a leading digit of 1. Thus the probability of 1 being a leading digit of a term in the sequence $\{a_i\}_{i=k+1}^{n+k}$ is $\frac{j}{n}$. Now consider the purely periodic sequence $\{a_i\}_{i=k+1}^{\infty}$. The probability of 1 being a leading digit of a term in the sequence is $\frac{j}{n}$, which is a rational number. If $\{a_i\}_{i=k+1}^{\infty}$ followed Benford's law, the probability of 1 being a leading digit of a term in the sequence is $\log 2$ which is irrational. Thus $\{a_i\}_{i=k+1}^{\infty}$ does not follow Benford's law. The sequence $\{a_i\}_{i=1}^{\infty}$ is simply the sequence formed by adding a finite number of terms (k terms) to the sequence $\{a_i\}_{i=k+1}^{\infty}$. Since $\{a_i\}_{i=k+1}^{\infty}$ does not follow Benford's law, by Proposition 3 $\{a_i\}_{i=1}^{\infty}$ also does not follow Benford's law. ■

While the terms of the continued fraction of a quadratic irrational do not satisfy Benford's law, the corresponding convergents do. In order to prove this, we will first consider the convergents of an α with a purely periodic continued fraction expansion.

Proposition 16. *Let $\alpha = [\overline{a_1, a_2, \ldots, a_n}]$. Then the sequence of convergent numerators $\{P_j\}_{j=1}^{\infty}$ and the sequence of convergent denominators $\{Q_j\}_{j=1}^{\infty}$ both satisfy Benford's Law.*

We provide two separate proofs of this statement. In each proof, we partition $\{P_n\}$ and $\{Q_n\}$ into subsequences $\{P_{nk+i}\}_{n=1}^{\infty}$ and $\{Q_{nk+i}\}_{n=1}^{\infty}$ for each integer $1 \leq i \leq k$ and demonstrate that each subsequence is a linear recurrence satisfying Benford's law. In the first proof, we use an algorithm to show that such a linear recurrence exists, while in the second, we use linear algebra to find an explicit formula for the recurrence.

Proof 1. From Proposition 4, we reduce the problem to showing that each $\{P_{kn+i}\}_{k=1}^{\infty}$ is a second order linear recurrence. First we introduce new notation. Given a sequence $x_0, x_1, x_2, ...$ we write $x_{c_0k+d_0} \to (x_{c_1k+d_1}, x_{c_2k+d_2}, ..., x_{c_jk+d_j})$ if, for some choice of $c_0, c_1, ..., c_j$ and $d_0, d_1, ..., d_j$, $x_{c_0k+d_0}$ is a linear recurrence in terms of $x_{c_1k+d_1}, x_{c_2k+d_2}, ..., x_{c_jk+d_j}$ for all values of k. From the recurrence $P_k = a_kP_{k-1} + P_{k-2}$ and the α being periodic with period n, we have

$$P_{kn+i} \to (P_{kn+i-1}, P_{kn+i-2})$$

$$P_{kn+i} \to (P_{kn+i+1}, P_{kn+i-1})$$

$$P_{kn+i} \to (P_{kn+i+1}, P_{kn+i+2}).$$

We want to show that $P_{kn+i} \to (P_{(k-1)n+i}, P_{(k-2)n+i})$. Using our relations, we find that

$$\begin{aligned} P_{kn+i} &\to (P_{kn+i-1}, P_{kn+i-2}) \\ &\to (P_{kn+i-2}, P_{kn+i-3}, P_{kn+i-2}) \to (P_{kn+i-2}, P_{kn+i-3}) \\ &\to (P_{kn+i-3}, P_{kn+i-4}). \end{aligned}$$

Continuing this algorithm, we obtain

$$P_{kn+i} \to (P_{kn+i-n}, P_{kn+i-(n+1)}) \to (P_{(k-1)n+i}, P_{(k-1)n+i-1}).$$

To complete our algorithm, we show that $P_{kn+i-1} \to (P_{kn+i}, P_{(k-1)n+i})$ by induction. As our base case, we know that

$$P_{kn+i-1} \to (P_{kn+i}, P_{kn+i-2}).$$

Assuming that $P_{kn+i-1} \to (P_{kn+i}, P_{kn+i-r})$, we can perform a series of moves to obtain $P_{kn+i-1} \to (P_{kn+i}, P_{kn+i-(r+1)})$:

$$\begin{aligned} P_{kn+i-1} &\to (P_{kn+i}, P_{kn+i-r}) \\ &\to (P_{kn+i}, P_{kn+i-(r+1)}, P_{kn+i-(r-1)}) \\ &\to (P_{kn+i}, P_{kn+i-(r+1)}, P_{kn+i-(r-2)}, P_{kn+i-(r-3)}). \end{aligned}$$

Repeatedly using $P_k \to (P_{k-1}, P_{k-2})$ gives us

$$\begin{aligned} P_{kn+i-1} &\to (P_{kn+i}, P_{kn+i-(r+1)}, P_{kn+i-(r-(r-1))}, P_{kn+i-(r-r)}) \quad \to (P_{kn+i}, P_{kn+i-(r+1)}, P_{kn+i-1}, P_{kn+i}) \\ &\to (P_{kn+i}, P_{kn+i-(r+1)}). \end{aligned}$$

Having established this inductive step, we see that $P_{kn+i-1} \to (P_{kn+i}, P_{kn+i-r})$ is true for all r. If we let $r = n$, we obtain $P_{kn+i-1} \to (P_{kn+i}, P_{(k-1)n+i})$. Substituting $P_{(k-1)n+i-1} \to (P_{(k-1)n+i}, P_{(k-2)n+i})$ into $P_{kn+i} \to (P_{(k-1)n+i}, P_{(k-1)n+i-1})$, we obtain $P_{kn+i} \to (P_{(k-1)n+i}, P_{(k-2)n+i})$.

Because α has period n, $\{P_{kn+i}\}_{k=1}^{\infty}$ is a second order linear recurrence for each $1 \leq i \leq n$. Furthermore, since each P_i is positive, $\{P_{kn+i}\}_{k=1}^{\infty}$ satisfies Benford's law by Proposition 10. Thus $\{P_k\}_{k=1}^{\infty}$ satisfies Benford's law by Proposition 4.

Proof 2. Once again, we show that $\{P_{ni+k}\}_{k=0}^{\infty}$ and $\{Q_{ni+k}\}_{k=0}^{\infty}$ are second order linear recurrences. By Proposition 11

$$\prod_{i=1}^{k+2n} \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} P_{k+2n} & P_{k+2n-1} \\ Q_{k+2n} & Q_{k+2n-1} \end{pmatrix}$$

Also, by the associativity of matrix multiplication,

$$\prod_{i=1}^{k+2n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = (\prod_{i=1}^{k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix})(\prod_{i=k+1}^{k+2n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix})$$

Since $\{a_i\}_{i=1}^{\infty}$ is periodic with period n,

$$= (\prod_{i=1}^{k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix})(\prod_{i=1}^{n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix})^2$$

Similarly,

$$\prod_{i=1}^{k+n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = (\prod_{i=1}^{k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix})(\prod_{i=1}^{n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix})$$

Let

$$\prod_{i=1}^{k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & \alpha_4\end{pmatrix}$$

and

$$\prod_{i=1}^{n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & \beta_4\end{pmatrix}$$

Thus

$$\prod_{i=1}^{k+n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & \alpha_4\end{pmatrix}\begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & \beta_4\end{pmatrix} = \begin{pmatrix}\alpha_1\beta_1+\alpha_2\beta_3 & \alpha_1\beta_2+\alpha_2\beta_4\\ \alpha_3\beta_1+\alpha_4\beta_3 & \alpha_3\beta_2+\alpha_4\beta_4\end{pmatrix}$$

and

$$\prod_{i=1}^{k+2n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & \alpha_4\end{pmatrix}\begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & \beta_4\end{pmatrix}^2 = \begin{pmatrix}\alpha_1\beta_1^2+\alpha_1\beta_2\beta_3+\alpha_2\beta_1\beta_3+\alpha_2\beta_3\beta_4 & \alpha_1\beta_1\beta_2+\alpha_1\beta_2\beta_4+\alpha_2\beta_2\beta_3+\alpha_2\beta_4^2\\ \alpha_3\beta_1^2+\alpha_3\beta_2\beta_3+\alpha_4\beta_1\beta_3+\alpha_4\beta_3\beta_4 & \alpha_3\beta_1\beta_2+\alpha_3\beta_2\beta_4+\alpha_4\beta_2\beta_3+\alpha_4\beta_4^2\end{pmatrix}$$

Since

$$\prod_{i=1}^{k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & \alpha_4\end{pmatrix}$$

varies with k, $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ vary with k.

$$\prod_{i=1}^{n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & \beta_4\end{pmatrix}$$

is a constant, and thus $\beta_1, \beta_2, \beta_3, \beta_4$ are constants too. Since we want to show $\{P_{ni+k}\}_{i=0}^{\infty}$ and $\{Q_{ni+k}\}_{i=0}^{\infty}$ are linear recurrences we want to find x, y in terms of $\beta_1, \beta_2, \beta_3, \beta_4$ such that

$$\prod_{i=1}^{k+2n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = x\prod_{i=1}^{k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} + y\prod_{i=1}^{k+n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix}$$

which is equivalent to finding x, y in terms of $\beta_1, \beta_2, \beta_3, \beta_4$ such that

$$\begin{pmatrix}\alpha_1\beta_1^2+\alpha_1\beta_2\beta_3+\alpha_2\beta_1\beta_3+\alpha_2\beta_3\beta_4 & \alpha_1\beta_1\beta_2+\alpha_1\beta_2\beta_4+\alpha_2\beta_2\beta_3+\alpha_2\beta_4^2\\ \alpha_3\beta_1^2+\alpha_3\beta_2\beta_3+\alpha_4\beta_1\beta_3+\alpha_4\beta_3\beta_4 & \alpha_3\beta_1\beta_2+\alpha_3\beta_2\beta_4+\alpha_4\beta_2\beta_3+\alpha_4\beta_4^2\end{pmatrix} = x\begin{pmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & \alpha_4\end{pmatrix} + y\begin{pmatrix}\alpha_1\beta_1+\alpha_2\beta_3 & \alpha_1\beta_2+\alpha_2\beta_4\\ \alpha_3\beta_1+\alpha_4\beta_3 & \alpha_3\beta_2+\alpha_4\beta_4\end{pmatrix}$$

We algebraically solve $x = \beta_2\beta_3 - \beta_1\beta_4$ and $y = \beta_1 + \beta_4$. By Proposition 11,

$$\prod_{i=1}^{n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}P_n & P_{n-1}\\ Q_n & Q_{n-1}\end{pmatrix} = \begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & \beta_4\end{pmatrix}$$

and as a result

$$x = \beta_2\beta_3 - \beta_1\beta_4 = P_{n-1}Q_n - P_nQ_{n-1} = (-1)^{n+1}$$
$$y = \beta_1 + \beta_4 = P_n + Q_{n-1}$$
$$\to P_{2n+k} = (P_n + Q_{n-1})P_{n+k} + (-1)^{n+1}P_k$$
$$\to Q_{2n+k} = (P_n + Q_{n-1})Q_{n+k} + (-1)^{n+1}Q_k$$

Thus, both $\{P_{ni+k}\}_{i=0}^{\infty}$ and $\{Q_{ni+k}\}_{i=0}^{\infty}$ are second order linear recurrences. Since they are both positive second order linear recurrences, by Proposition 10, $\{P_{ni+k}\}_{i=0}^{\infty}$ and $\{Q_{ni+k}\}_{i=0}^{\infty}$ follow Benford's law. Thus $\{P_k\}_{k=1}^{\infty}$ and $\{Q_k\}_{k=1}^{\infty}$ satisfy Benford's law by Proposition 4.

We can extend the previous result to quadratic irrationals with any periodic continued fraction expansion. To do so, we show that the non-periodic terms of an eventually periodic continued fraction generate a finite number of convergents that are not Benford, meaning that the overall sequence of convergents remains Benford.

Proposition 17. *Let $\alpha = [a_1, a_2, \ldots, a_k, \overline{a_{k+1}, \ldots, a_{n+k}}]$. Then $\{P_j\}_{j=1}^{\infty}$ and $\{Q_j\}_{j=1}^{\infty}$ satisfy Benford's law.*

Proof. Let A_j and B_j be the numerator and denominator, respectively, of the j^{th} convergent of $\beta = [\overline{a_{k+1}, a_{k+2}, \ldots, a_{n+k}}]$. Furthermore, let P_j and Q_j be the numerator and denominator, respectively, of the j^{th} convergent of $[a_1, a_2, \ldots, a_k, \overline{a_{k+1}, \ldots, a_{n+k}}]$. Using the principle of strong induction, it will be shown that $P_{k+j} = A_jP_k + B_jP_{k-1}$ for $j \in \mathbb{N}$. Consider the base case, $j = 1$.

$$P_{k+1} = a_{k+1}P_k + P_{k-1} = A_1P_k + B_1P_{k-1}$$

Thus $P_{k+j} = A_jP_k + B_jP_{k-1}$ is true for $j = 1$. Assume $P_{k+j} = A_jP_k + B_jP_{k-1}$ is true for $j = 1, 2, \ldots, m$. Then,

$$P_{k+m+1} = a_{k+m+1}P_{k+m} + P_{k+m-1}$$
$$= a_{k+m+1}(A_mP_k + B_mP_{k-1}) + (A_{m-1}P_k + B_{m-1}P_{k-1})$$
$$= (a_{k+m+1}A_j + A_{m-1})P_k + (a_{k+m+1}B_j + B_{m-1})P_{k-1}$$
$$= A_{k+m+1}P_k + B_{k+m+1}P_{k-1}$$

Thus $P_{k+j} = A_jP_k + B_jP_{k-1}$ holds for $j = m + 1$ and by the principle of strong induction, $P_{k+j} = A_jP_k + B_jP_{k-1}$ holds for all $j \in \mathbb{N}$. But

$$\lim_{j\to\infty} \frac{A_j}{B_j} = \beta$$

so

$$\lim_{j\to\infty} P_{k+j} = B_j\beta P_k + B_jP_{k-1} = B_j(\beta P_k + P_{k-1})$$

But by Proposition 16 B_j follows Benford's law, and $0 < \beta, P_k, P_{k-1}$ which means $\beta P_k + P_{k-1} > 0$ and thus $B_j(\beta P_k + P_{k-1})$ follows Benford's law by Proposition 2 and by Proposition 5 $\{P_{k+j}\}_{j=1}^{\infty}$ follows Benford's law. Thus adding $\{P_j\}_{j=1}^{k}$, a finite number of terms, to the sequence $\{P_{k+j}\}_{j=1}^{\infty}$ creates a new sequence $\{P_j\}_{j=1}^{\infty}$ which follows Benford's law by Proposition 3. By Proposition 14 $\{Q_j\}_{j=1}^{\infty}$ also follows Benford's law. ■

To prove that $\{P_j\}_{j=1}^{\infty}$ and $\{Q_j\}_{j=1}^{\infty}$ followed Benford's law for purely periodic continued fractions, we found an explicit formula for the recurrence relations for the convergent numerators and denominators. We can extend this result to any periodic continued fraction.

Proposition 18. *Let $\alpha = [a_1, a_2, \ldots, a_k, \overline{a_{k+1}, \ldots, a_{n+k}}]$. Then*

$$P_{k+j+2n} = (-1^k(P_{n+k}Q_{k-1} - P_{n+k-1}Q_k + P_kQ_{n+k-1} - P_{k-1}Q_{n+k}))P_{k+j+n} + (-1^{n+k+1})P_{k+j}$$
$$Q_{k+j+2n} = (-1^k(P_{n+k}Q_{k-1} - P_{n+k-1}Q_k + P_kQ_{n+k-1} - P_{k-1}Q_{n+k}))Q_{k+j+n} + (-1^{n+k+1})Q_{k+j}$$

Proof. Like we did when finding the recurrence relation for purely periodic continued fractions, we make use of Proposition 11 to get

$$\prod_{i=1}^{k+j+2n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}P_{k+j+2n} & P_{k+j+2n-1}\\ Q_{k+j+2n} & Q_{k+j+2n-1}\end{pmatrix} = (\prod_{i=1}^{k+j}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix})(\prod_{i=k+1}^{n+k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix})^2$$

$$\prod_{i=1}^{k+j+n}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}P_{k+j+n} & P_{k+j+n-1}\\ Q_{k+j+n} & Q_{k+j+n-1}\end{pmatrix} = (\prod_{i=1}^{k+j}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix})(\prod_{i=k+1}^{n+k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix})$$

$$\prod_{i=1}^{k+j}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}P_{k+j} & P_{k+j-1}\\ Q_{k+j} & Q_{k+j-1}\end{pmatrix}$$

If

$$\prod_{i=k+1}^{n+k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & \beta_4\end{pmatrix}$$

then

$$P_{k+j+2n} = (\beta_1+\beta_4)P_{k+j+n} + (\beta_2\beta_3 - \beta_1\beta_4)P_{k+j}$$
$$Q_{k+j+2n} = (\beta_1+\beta_4)Q_{k+j+n} + (\beta_2\beta_3 - \beta_1\beta_4)Q_{k+j}$$

which we showed when finding the explicit form of the recursion for purely periodic continued fractions. Rewriting this product, we can compute $\beta_1, \beta_2, \beta_3, \beta_4$.

$$\begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & \beta_4\end{pmatrix} = \prod_{i=k+1}^{n+k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix} = \frac{\prod_{i=1}^{n+k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix}}{\prod_{i=1}^{k}\begin{pmatrix}a_i & 1\\ 1 & 0\end{pmatrix}} = \frac{\begin{pmatrix}P_{n+k} & P_{n+k-1}\\ Q_{n+k} & Q_{n+k-1}\end{pmatrix}}{\begin{pmatrix}P_k & P_{k-1}\\ Q_k & Q_{k-1}\end{pmatrix}}$$

Since the determinant of the matrix in the denominator is not zero, it has an inverse. Let

$$\begin{pmatrix}P_k & P_{k-1}\\ Q_k & Q_{k-1}\end{pmatrix}^{-1} = \begin{pmatrix}\gamma_1 & \gamma_2\\ \gamma_3 & \gamma_4\end{pmatrix}$$

and by the definition of the inverse of a matrix,

$$\begin{pmatrix}P_k & P_{k-1}\\ Q_k & Q_{k-1}\end{pmatrix}\begin{pmatrix}\gamma_1 & \gamma_2\\ \gamma_3 & \gamma_4\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

which results in the following set of four equations in four variables $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$.

$$P_k\gamma_1 + P_{k-1}\gamma_3 = 1$$

$$P_k\gamma_2 + P_{k-1}\gamma_4 = 0$$
$$Q_k\gamma_1 + Q_{k-1}\gamma_3 = 0$$
$$Q_k\gamma_2 + Q_{k-1}\gamma_4 = 1$$

which results in

$$\gamma_1 = -1^k Q_{k-1}$$
$$\gamma_2 = -1^{k+1} P_{k-1}$$
$$\gamma_3 = -1^{k+1} Q_k$$
$$\gamma_4 = -1^k P_k$$

Thus,

$$\begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & \beta_4\end{pmatrix} = \begin{pmatrix}P_{n+k} & P_{n+k-1}\\ Q_{n+k} & Q_{n+k-1}\end{pmatrix}\begin{pmatrix}-1^k Q_{k-1} & -1^{k+1} P_{k-1}\\ -1^{k+1} Q_k & -1^k P_k\end{pmatrix}$$

We are only concerned with calculating $\beta_2\beta_3 - \beta_1\beta_4$ and $\beta_1 + \beta_4$ which we compute to be

$$\beta_1 + \beta_4 = (-1^k(P_{n+k}Q_{k-1} - P_{n+k-1}Q_k + P_kQ_{n+k-1} - P_{k-1}Q_{n+k}))$$

$$\beta_2\beta_3 - \beta_1\beta_4 = -1^{n+k+1}$$

and thus

$$P_{k+j+2n} = (-1^k(P_{n+k}Q_{k-1} - P_{n+k-1}Q_k + P_kQ_{n+k-1} - P_{k-1}Q_{n+k}))P_{k+j+n} + (-1^{n+k+1})P_{k+j}$$

$$Q_{k+j+2n} = (-1^k(P_{n+k}Q_{k-1} - P_{n+k-1}Q_k + P_kQ_{n+k-1} - P_{k-1}Q_{n+k}))Q_{k+j+n} + (-1^{n+k+1})Q_{k+j}$$

∎

If we let $k = 0$ then α becomes purely periodic. Letting $P_{-1} = 0, Q_{-1} = 1$, the recursive formulas derived for periodic continued fractions become the ones derived for purely periodic fractions. Setting P_{-1} and Q_{-1} to these values coupled with $P_0 = 1, Q_0 = 0$ results in $P_1 = a_1, Q_1 = 1$. With these explicit formulas for the recurrence relations involving the convergent numerators and denominators, we can prove that $\{P_j\}_{j=1}^{\infty}$ and $\{Q_j\}_{j=1}^{\infty}$ satisfy Benford's Law for any periodic continued fraction using a similar method used for purely periodic continued fraction. However, this result would be stronger because it is a generalization of purely periodic continued fractions. Now that we have found a large class of numbers (the quadratic irrationals) whose $\{P_j\}_{j=1}^{\infty}$ and $\{Q_j\}_{j=1}^{\infty}$ follow Benford's Law, we consider arbitrarily chosen numbers.

4.3 The Gauss-Kuzmin Distribution

Having shown that the terms of the simple continued fractions of rationals and quadratic irrationals do not satisfy Benford's Law, a natural question to ask would be if these results can be extended for some arbitrary α.

Although real numbers come in many different varieties, their continued fractions share a common behavior. Gauss discovered that the terms in the continued fraction expansions of almost all $x \in [0, 1)$ follow the same distribution:

Theorem 4 (Gauss-Kuzmin [Kar13]). *Let $[a_1, a_2, a_3 \ldots]$ be the continued fraction expansion of a randomly chosen $x \in [0, 1)$. Then as $n \to \infty$, the probability that $a_n = k$ for any $k \in \mathbb{N}$ is*

$$\lim_{n\to\infty} P(a_n = k) = -\log_2(1 - \frac{1}{(k+1)^2}).$$

Using this distribution, we can determine whether almost all α are Benford.

Proposition 19. *Let α have a simple continued fraction expansion of $[a_1, a_2, \ldots]$. Then for all α except a set of measure zero, the sequence $\{a_n\}_{n=1}^{\infty}$ does not follow Benford's Law.*

Proof. Let x be a number chosen randomly from the interval (0,1). Furthermore, let the simple continued fraction expansion of x be $[a_1, a_2, \ldots]$. Then by the Gauss-Kuzmin theorem, for all x except for a set of measure zero,

$$\lim_{n\to\infty} P(a_n = k) = \frac{-\log(1 - \frac{1}{(k+1)^2})}{\log 2}$$

. Consider the probability that a_n has a leading digit of 1. This probability is

$$\sum_{i=1}^{\infty} \frac{-\log(1-\frac{1}{(b_i+1)^2})}{\log 2} = \frac{1}{\log 2}\sum_{i=1}^{\infty} -\log(1 - \frac{1}{(b_i+1)^2}) = \frac{1}{\log 2}\sum_{i=1}^{\infty} \log\frac{(b_i+1)^2}{b_i^2+2b_i} = (\frac{1}{\log 2})(\log\prod_{i=1}^{\infty}\frac{(b_i+1)^2}{b_i^2+2b_i})$$

where b_i is the i^{th} natural number with a leading digit of 1. Note that $\frac{(b_i+1)^2}{b_i^2+2b_i} = 1 + \frac{1}{b_i^2+2b_i} > 1$ for $b_i \in \mathbb{N}$. Thus $\prod_{i=1}^{k}\frac{(b_i+1)^2}{b_i^2+2b_i} < \prod_{i=1}^{k+1}\frac{(b_i+1)^2}{b_i^2+2b_i}$ and as a result

$$(\frac{1}{\log 2})(\log\prod_{i=1}^{1}\frac{(b_i+1)^2}{b_i^2+2b_i}) = \frac{\log\frac{4}{3}}{\log 2} < (\frac{1}{\log 2})(\log\prod_{i=1}^{\infty}\frac{(b_i+1)^2}{b_i^2+2b_i})$$

If a_n followed Benford's law, then the probability that the leading digit is 1 would be $\log 2$. But

$$\log 2 < \frac{\log \frac{4}{3}}{\log 2} < (\frac{1}{\log 2})(\log \prod_{i=1}^{\infty} \frac{(b_i+1)^2}{b_i^2+2b_i} \to \log 2 \neq (\frac{1}{\log 2})(\log \prod_{i=1}^{\infty} \frac{(b_i+1)^2}{b_i^2+2b_i})$$

Thus the sequence a_n does not follow Benford's law. Consider a real number $y = \lfloor y \rfloor + \{y\}$ chosen from the interval $(1, \infty)$. The continued fraction expansion of y is thus $[\lfloor y \rfloor, b_2, b_3, \ldots]$ where the continued fraction expansion of $\{y\}$ is $[0, b_2, b_3, \ldots]$. Removing and adding a finite number of terms from a sequence that does not follow Benford's law results in another sequence that does not follow Benford's law, by Proposition 3. Thus, the sequence b_n does not follow Benford's law. Thus, for all α in $((0, \infty)$, except for a set of measure zero, $\{a_n\}_{n=1}^{\infty}$ follows Benford's Law. ■

4.4 The Levy Constant

Just as the terms of almost all continued fractions follow the Gauss-Kuzmin distribution, the convergents of almost all continued fractions also follow a certain distribution. This phenomenon is characterized by the following theorem, which was proved by Lévy:

Theorem 5 ([Bax09]). *Let $\frac{P_n}{Q_n}$ be the nth convergent of α. Then for all α but a set of measure zero,*

$$\lim_{n\to\infty} Q_n^{1/n} = e^{\beta},$$

where $\beta = \frac{\pi^2}{12 \log 2}$.

Using Levy's theorem, we can conjecture that the sequences of convergents of almost all α satisfy Benford's law.

Conjecture. $\log_{10} e^{\beta}$ *is irrational.*

If the conjecture is true, then we have the following proposition:

Proposition 20. *Let $\frac{P_n}{Q_n}$ be the n^{th} convergent of α. For all α but a set of measure zero, the sequences $\{Q_n\}_{n=1}^{\infty}$ and $\{P_n\}_{n=1}^{\infty}$ follow Benford's Law.*

Proof. By Theorem 5, for all α but a set of measure zero, $\lim_{n\to\infty} Q_n^{1/n} = e^{\beta}$, which implies $\lim_{n\to\infty} Q_n = e^{\beta n}$. Thus

$$\lim_{n\to\infty} \log Q_n - \log Q_{n-1} = \lim_{n\to\infty} \log \frac{Q_n}{Q_{n-1}} = \log \frac{e^{\beta n}}{e^{\beta(n-1)}} = \log e^{\beta}.$$

If $\log e^{\beta} = \frac{\pi^2}{12ln2}$ is irrational, then $\{\log Q_n\}_{n=1}^{\infty}$ is equidistributed modulo 1 by Proposition 3, which means that $\{Q_n\}_{n=1}^{\infty}$ follows Benford's law by Proposition 1. By Proposition 14 $\{P_n\}_{n=1}^{\infty}$ also follows Benford's Law. ■

What follows from this conjecture is a very strong result: for almost all α, $\{P_n\}_{n=1}^{\infty}$ and $\{Q_n\}_{n=1}^{\infty}$ follow Benford's Law. However, this relies on $\log_{10} e^{\beta}$ being irrational. One way to gather numerical evidence to support this conjecture is to see if $\{P_n\}_{n=1}^{\infty}$ and $\{Q_n\}_{n=1}^{\infty}$ follow Benford's Law for α that obey Levy's constant. If they do follow Benford's Law, then we have reasons to believe that our conjecture is true.

4.5 Continued Fractions of Transcendentals

While Gauss and Levy's conditions hold true for almost all irrationals, it is not known how to determine whether a given transcendental number such as e and π satisfies them. Thus, we decided to separately investigate whether the terms and convergents of these two transcendentals satisfies Benford's law.

The continued fraction expansion for e has an elegant structure: $e = [2; 1, 2, 1, 1, 4, 1, 1, 6, ...]$. Since the terms of this continued fraction have a first digit of 1 more than half of the time, it is clear that the terms of e do not satisfy Benford's law. On the other hand, the continued fraction of $\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2...]$

follows no recognizable pattern. Yet when we perform numerical analysis, we discover that the convergents of e and π behave similarly. To gather data, we wrote a Sage program that calculated the successive convergents with a linear recursion and computed the frequency of their first digits. We then compared the frequencies with the those predicted by Benford's law by taking the average distance between them, measured by the absolute value of their difference. As the number of terms increased, we saw that their average distances quickly tended towards zero:

# Terms	Average Distance (e)	Average Distance (π)
100	$2.86 \cdot 10^{-2}$	$3.00 \cdot 10^{-2}$
1000	$9.55 \cdot 10^{-3}$	$5.74 \cdot 10^{-3}$
10,000	$3.74 \cdot 10^{-3}$	$2.62 \cdot 10^{-3}$

Given that the predicted first digit frequency ranges from 0.3 for a first digit of 1 to 0.05 for a first digit of 9, the size of the disparity is less than the range by an order of 10^2. Thus we have sufficient evidence to conjecture that the convergents of e and π both satisfy Benford's law.

5 Conclusion

In this paper we have applied the method of proving that Benford's law arises in the Fibonacci sequence to prove that all linear recurrences satisfy Benford's law. We have also derived results for simple continued fraction expansions of irrationals, including that the convergents of periodic continued fractions follow Benford's law. By using the Gauss-Kuzmin distribution, we showed that the terms of almost all simple continued fractions do not follow Benford's Law. Furthermore, we conjectured that the the convergents of all simple continued fractions but a set of measure of zero follow Benford's Law by conjecturing theat the common log of the Levy constant is irrational.

Along the line of continued fractions, the most immediate future direction includes proving the satisfaction of Benford's law for the continued fraction expansions of e and π, which are neither periodic nor confirmed to follow the Gauss-Kuzmin distribution. Another unexplored problem would involve investigating Benford's law for non-simple continued fractions, which have values other than 1 in the numerators of the expansion. Through numerical calculations, we found that the convergents of the simple continued fraction expansion of π converged to a Benford sequence faster than e with both converging slower than quadratic irrationals. Thus it would be natural to explore which numbers' continued fraction expansions converge to a Benford sequence fastest.

6 Acknowledgements

The authors would like to thank Dr. Jayadev Athreya for proposing this problem. We would also like to thank our PROMYS research lab counselor Adam Block for his support and guidance throughout our project. Furthermore, we would like to thank Dan Jerison, David Fried, Glenn Stevens, the PROMYS Foundation, and the Clay Mathematics Institute for making this mathematical research possible.

References

[Bax09] Chrostoph Baxa. Levy Constants of Transcendental Numbers. *Proceedings of the American Mathematical Society* 137(7):2243-2249, 2009.

[Kar13] Oleg Karpenkov. *Geometry of Continued Fractions*, Springer, Berlin, 2013.

[KM05] Alex Kontorovich, Steven Miller. Benford's Law, Values of L-Functions and the 3x + 1 Problem. *Acta Arithmetica* 120(3), 269–297, 2005.

[KN74] L Kuipers, H. Niederreiter. *Uniform Distribution of Sequences*, Wiley, New York, 1974.