ON BENFORD'S LAW AND THE COEFFICIENTS OF THE RIEMANN MAPPING FUNCTION FOR THE EXTERIOR OF THE MANDELBROT SET

FILIPPO BERETTA, JESSE DIMINO, WEIKE FANG, THOMAS C. MARTINEZ, STEVEN J. MILLER, AND DANIEL STOLL

Abstract. We investigate Benford's law in relation to fractal geometry. Basic fractals like the Cantor set and Sierpinski triangle are obtained as the limiting iterations of sets, the unique measures of whose components follow a geometric distribution which is Benford in most bases. Building on this intuition, we study this distribution in more complicated fractals such as the Mandelbrot set. We examine the Laurent and Taylor coefficients of a Riemann mapping and its reciprocal function from the exterior of the Mandelbrot set to the complement of the unit disk. We utilize statistical testing to show that the coefficients are a good fit for Benford's law, and we offer additional conjectures and observations about these coefficients. In particular, we highlight certain arithmetic subsequences related to the coefficients' denominators, provide an estimate for their slope, and describe efficient methods to compute some of them.

1. Introduction

Analyzing the family of functions $f_c(z) = z^2 + c$, Douady and Hubbard introduced the Mandelbrot set $\mathcal{M}$ as the set of parameters c for which the orbit of 0 under f_c remains bounded. We study Benford's law in relation to the Mandelbrot set to both investigate the distribution's extension to fractal geometry and search for patterns in the Mandelbrot set.

In 1980, Douady and Hubbard were able to prove the connectedness of $\mathcal{M}$ by constructing a conformal isomorphism

$$\Phi : \mathbb{C} \setminus \mathcal{M} \longrightarrow \mathbb{C} \setminus \overline{\mathbb{D}}$$

between the complement of the Mandelbrot set and the complement of the closed unit disk [DouH1]. Using the Douady-Hubbard map Φ, we can define related conformal

Date: February 4, 2022.

2020 *Mathematics Subject Classification.* primary: 30B10, 30C20, 62P99; secondary: 62-08, 68Q25.

Key words and phrases. Mandelbrot Set, Complex Dynamics, Benford's Law, Power Series.

This work was partially supported by NSF Grant DMS2113535. We thank our colleagues from the 2021 Polymath Jr REU program for helpful comments on earlier drafts.

isomorphisms,

$$\Psi : \mathbb{C} \setminus \overline{\mathbb{D}} \longrightarrow \mathbb{C} \setminus \mathcal{M},$$
$$\Theta : \mathbb{D} \longrightarrow \mathbb{C} \setminus \mathcal{M}^{-1},$$

where $\mathcal{M}^{-1} = \{1/c : c \in \mathcal{M}\}$, by setting $\Psi = \Phi^{-1}$ and $\Theta(c) = 1/\Psi(1/c)$. By a theorem of Caratheodory [LuoY1], these two maps can be extended continuously to the unit circle if and only if the Mandelbrot set is locally connected. One of the most heavily studied questions in complex dynamics is whether or not $\mathcal{M}$ is locally connected (MLC). As such, we focus on studying these maps and their respective Laurent and Taylor expansions:

$$\Psi(z) = z + \sum_{m=0}^{\infty} b_m z^{-m}, \tag{1.0.1}$$

$$\Theta(z) = z + \sum_{m=2}^{\infty} a_m z^{m}. \tag{1.0.2}$$

In Section 3, we outline the methods [Shi2, Shi3] we used to compute the coefficients a_m and b_m. The computation time grows exponentially, so methods of improving computation are explored. With our design of data structures and utilization of high-performance computing resources, we were able to compute the numerators, denominators, and decimal expansions of the first 1024 coefficients.

In Section 4, we discuss Benford's law along with statistical testing to determine whether the coefficients obey a Benford distribution. Given a base $b \geq 2$, a data set is Benford base b if the probability of observing a given leading digit, d, is $\log_b\left(\frac{d+1}{d}\right)$ (see [Benf1, Mil1]).[1] We consider the standard χ^2 distribution and the sequence of the data's logarithms modulo 1 for our statistical testing, and a standard goodness of fit test shows that the coefficients are a good fit for Benford's law.

In Section 5, we present new results and conjectures on the a_m and b_m coefficients. In Theorem 5.0.1, taken from [Shi1], we recall that they are 2-adic rational numbers; in other words, they are of the form $p/2^{-\nu}$, where p is an odd integer. The integer ν is, by definition, the 2-adic valuation ν_2 of a_m or b_m. Therefore we focus on the denominator's exponents $-\nu(a_m)$, $-\nu(b_m)$. Setting $m = 2^n m_0$, with m_0 odd and n fixed; the subsequences $\{-\nu(a_m)\}$, $\{-\nu(b_m)\}$ appear to be near-arithmetic progressions. We present the results observed in the following conjecture.

Conjecture 1.0.1. *Let m be written as $2^n m_0$ as above, with $n = \overline{n}$ fixed. Then the sequence $\{-\nu(a_m)\}_{n=\overline{n}}$ is asymptotically linear, with slope*

$$2/(2^{\overline{n}+1} - 1). \tag{1.0.3}$$

[1]We can write any positive x as $S_b(x)b^{k_b(x)}$, where $S_b(x) \in [1, b)$ is the significand and $k_b(x)$ is an integer. If the probability of observing a significand of $s \in [1, b)$ is $\log_b(s)$ we say the set is strongly Benford (or frequently just Benford).

We also present an efficient way to compute the denominator's exponents for the cases $n = 0, 1, 2$.

2. Preliminaries in Complex Dynamics

We give a brief introduction to complex dynamics. For more detailed proofs see [Miln1] and [Sto1].

Let $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ be a rational map. The *Julia set* J_f associated to the map f may be defined as the closure of the set of repelling periodic points of f. For a rational map of degree 2 or higher, the Julia set J_f is non-empty. We now restrict our attention to polynomial maps of degree $d \geq 2$, which have a superattracting fixed point at infinity. We can thus define the *filled Julia set* K_f as the complement of the basin of attraction of infinity:

$$K_f \; = \; \left\{ z \in \mathbb{C} : \{f^n(z)\}_{n \geq 0} \text{ is bounded} \right\}.$$

Making use of the above, it is possible to redefine J_f as the boundary of the filled Julia set.

Lemma 2.0.1. *Let f be a polynomial of degree $d \geq 2$. The filled Julia set $K_f \subset \mathbb{C}$ is compact, with boundary $\partial K_f = J_f$ equal to the Julia set. The complement $\hat{\mathbb{C}} \setminus K_f$ is connected and equal to the basin of attraction $A(\infty)$ of the point ∞.*

It follows that the Julia set J_f of a polynomial f is precisely the boundary of the basin of attraction $\mathcal{A}_f(\infty)$.

We may now characterize the connectedness of J_f. This is determined entirely by the activity of the critical points of f.

Theorem 2.0.2. *The Julia set J_f for a polynomial of degree $d \geq 2$ is connected if and only if the filled Julia set K_f contains every critical point of f.*

2.1. **The Mandelbrot set.**

We now focus primarily on the family of quadratic functions of the form $\{f_c(z) = z^2 + c\}_{c \in \mathbb{C}}$. Since f_c has a single critical point 0, it follows from Theorem 2.0.2 that J_{f_c} is connected if and only if the orbit $\{f_c^n(0) \mid n \in \mathbb{N}\}$ is bounded. This motivates the following definition of the Mandelbrot set, $\mathcal{M}$.

Definition 2.1.1. $\mathcal{M} \subset \mathbb{C}$ is the set of all the parameters $c \in \mathbb{C}$ such that the Julia set J_{f_c} is connected. Equivalently, $\mathcal{M}$ is the set of all the c such that the orbit of 0 under f_c remains bounded:

$$\mathcal{M} \; := \; \{c \in \mathbb{C} : \text{there exists } R > 0 \text{ such that for all } n, |f_c^n(0)| < R\}.$$

Remark 2.1.2. It is possible to generalize this definition and most of the following results to the family of unicritical degree d polynomials $f_{c,d}(z) = z^d + c$, where d is an integer $d \geq 2$. In this case, $\mathcal{M}_d$ is called the *multibrot set* of degree d. For simplicity, we focus only on $\mathcal{M} = \mathcal{M}_2$, which has historically been the object of greatest interest. For more information on $\mathcal{M}_d$ see [Shi3, Shi4].

It is possible to show that the interior of $\mathcal{M}$ is nonempty. We can use computer graphics obtain visualizations of $\mathcal{M}$.

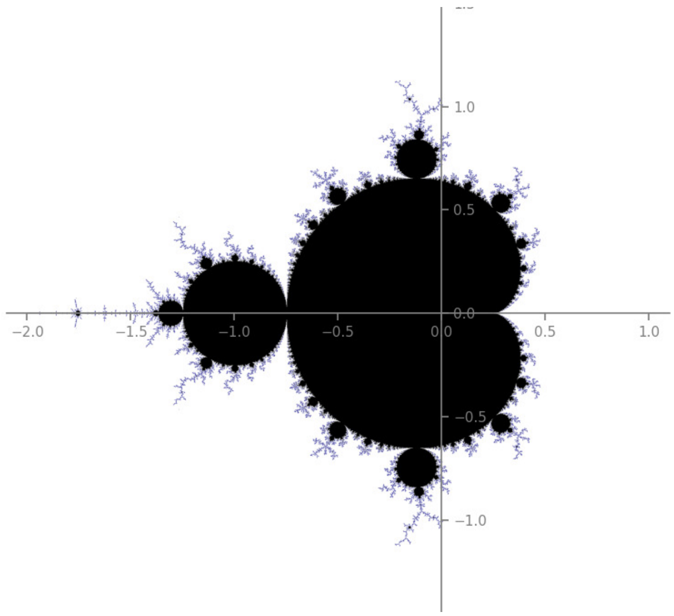


FIGURE 2.1.1. The Mandelbrot set $\mathcal{M}$ in the complex plane.

When the first computer images of $\mathcal{M}$ were generated, Benoit Mandelbrot observed small regions that appeared to be separate from the main cardioid, and conjectured that $\mathcal{M}$ was disconnected, which was later disproved.

Theorem 2.1.3 (Douady, Hubbard)**.** *The Mandelbrot set is connected.*

This result was first proved by Douady and Hubbard [DouH1] by explicitly constructing a conformal isomorphism $\Phi : \hat{\mathbb{C}} \setminus \mathcal{M} \to \hat{\mathbb{C}} \setminus \overline{\mathbb{D}}$. Douady and Hubbard's proof is significant not only for the result, but also since it provides an explicit formula for the uniformization of the complement of the Mandelbrot set.

A large amount of research has been devoted to the local connectivity of the Mandelbrot set, which is generally regarded as one of the most important open problems in complex dynamics. We recall that a set A in a topological space X is locally connected at $p \in A$ if for every open set $V \subset X$ containing p, there is an open subset U with $p \in U \subset V$ such that $U \cap X$ is connected. The set A is said to be *locally connected* if it is locally connected at p for all $p \in A$.

As above, let $\Phi : \hat{\mathbb{C}} \setminus \mathcal{M} \to \hat{\mathbb{C}} \setminus \overline{\mathbb{D}}$ be the conformal isomorphism constructed by Douady and Hubbard; notice that the map $\Psi : \hat{\mathbb{C}} \setminus \overline{\mathbb{D}} \to \hat{\mathbb{C}} \setminus \mathcal{M}$ is the Riemann mapping function of $\hat{\mathbb{C}} \setminus \mathcal{M}$. We consider its the Laurent expansion at ∞

$$\Psi(z) \;=\; z + \sum_{m=0}^{\infty} b_m z^{-m}.$$

Another possibility is to consider $\Theta(z) := 1/\Psi(1/z)$, which is the Riemann mapping of the bounded domain $\mathbb{C} \setminus \{1/z : z \in \mathcal{M}\}$. We have the corresponding Taylor expansion for Θ at the origin:

$$\Theta(z) \;=\; z + \sum_{m=2}^{\infty} a_m z^{m}.$$

To underline the importance of these maps, we reference a lemma from Caratheodory [LuoY1].

Theorem 2.1.4. ***(Caratheodory's continuity lemma)*** *Let $G \subset \hat{\mathbb{C}}$ be a simply connected domain and a function f maps $\mathbb{D}$ onto G via a conformal isomorphism. Then f has a continuous extension to $\overline{\mathbb{D}}$ if and only if the boundary of G is locally connected.*

Therefore, if the map Ψ or the map f can be extended continuously to the unit circle $\partial\mathbb{D}$, then $\mathcal{M}$ is locally connected. To show this extension, it would be sufficient to prove that one of the two series converges uniformly on $\mathbb{D}$.

There have been numerous attempts to prove this result. For example, Ewing and Schober showed that the inequality $0 < |b_m| < 1/m$ holds for every $m < 240000$. A bound of the type $|b_m| < K/m^{(1+\epsilon)}$ would lead to the result desired; however, this would imply that the extension of Ψ is Holder continuous, which does not happen as expressed in [BieFH1]. Proving that $|b_m| < K/(m\ \log^2(m))$ would prove that the series converges absolutely, however modern computations suggest that such a bound does not exist [BieFH1]. Therefore, the MLC conjecture and its consequences remain object of study.

2.2. **Consequences of local connectedness.**

This paper tackles problems related behaviour of the coefficients $\{a_m\}$ and $\{b_m\}$, with the ultimate aim of improving the state-of-art on local connectedness of $\mathcal{M}$. In this section we highlight the importance of this property of $\mathcal{M}$. In particular, it implies a conjecture known as Density of Hyperbolicity. A *hyperbolic component* is an interior connected component of $\mathcal{M}$ in which the sequences $\{f_c^n(0)\}$ have an attracting periodic cycle of period p. The *Density of Hyperbolicity* conjecture states that these are the only interior regions of $\mathcal{M}$. Another consequence of MLC is related to a topologically-equivalent description of $\partial\mathcal{M}$. In particular, the boundary of the Mandelbrot set can be identified with the unit circle S_1 under a specific relation $\sim$, known as the *abstract Mandelbrot set* [BanK1]. More information and other implications of MLC and Density of Hyperbolicity may be found in [Beni1] and [DouH1].

3. Algorithms and Complexity

There are algorithms to compute both the coefficients a_m and b_m. While these algorithms work for a generic degree $d \geq 2$, we focus on $d = 2$, which is the most interesting case historically, since it is the one associated with $\mathcal{M}_2 = \mathcal{M}$. The behaviour for other values is similar. Derivations for the explicit form of the b_m coefficients may be found in [Shi3]. Derivations for the explicit form of the a_m coefficients may be found in [Shi2]. Recursive algorithms and the formula to switch between the coefficients are outlined in [Shi1].

Theorem 3.0.1. *Let $n \in \mathbb{N}$ and $1 \leq m \leq d^{n+1} - 3$. Then*

$$\begin{aligned} b_{d,m} &= -\frac{1}{m} \sum C_{j_1}\left(\frac{m}{d^n}\right) C_{j_2}\left(\frac{m}{d^{n-1}} - dj_1\right) C_{j_3}\left(\frac{m}{d^{n-2}} - d^2 j_1 - dj_2\right) \cdots \\ & \qquad C_{j_n}\left(\frac{m}{d} - d^{n-1} j_1 - \cdots - dj_{n-1}\right), \end{aligned}$$

where the sum is over all non-negative indices $j_1, \ldots, j_n$ such that

$$\left(d^n - 1\right) j_1 + \left(d^{n-1} - 1\right) j_2 + \left(d^{n-2} - 1\right) j_3 + \cdots + (d-1)\, j_n \;=\; m + 1 \qquad (3.0.1)$$

and $C_j\left(a\right)$ is the generalized binomial coefficient

$$C_j\left(a\right) \;=\; \frac{a\left(a-1\right)\cdots\left(a-j+1\right)}{j\left(j-1\right)\cdots 2 \cdot 1}.$$

The coefficient a_m can be obtained from the b_m using the formula

$$a_{d,m} \;=\; -b_{d,m-2} - \sum_{j=2}^{m-1} a_{d,j} b_{d,m-1-j}$$

or can be directly calculated as in the following theorem.

Theorem 3.0.2. *Let $n \in \mathbb{N}$ and $2 \leq m \leq d^{n+1} - 1$. Then*

$$\begin{aligned} a_{d,m} &= \frac{1}{m} \sum C_{j_1}\left(\frac{m}{d^n}\right) C_{j_2}\left(\frac{m}{d^{n-1}} - dj_1\right) C_{j_3}\left(\frac{m}{d^{n-2}} - d^2 j_1 - dj_2\right) \cdots \\ & \qquad C_{j_n}\left(\frac{m}{d} - d^{n-1} j_1 - \cdots - dj_{n-1}\right), \end{aligned}$$

where the sum is over all non-negative indices $j_1, \ldots, j_n$ such that

$$(d^n - 1)j_1 + (d^{n-1} - 1)j_2 + (d^{n-2} - 1)j_3 + \cdots + (d-1)j_n \;=\; m - 1 \qquad (3.0.2)$$

and $C_j(a)$ is the generalized binomial coefficient

$$C_j(a) \;=\; \frac{a(a-1)\cdots(a-j+1)}{j(j-1)\cdots 2 \cdot 1}.$$

3.1. **Computation.**

We wrote a program in Python to compute $b_{2,m}$ and $a_{2,m}$ based on the formulas given in Theorem 3.0.1 and Theorem 3.0.2, and we obtained the first 1024 exact values of both coefficients' sequences.

Our methodology for computing the m^{th} coefficient was to first set an upper limit on the degree for which to obtain coefficients, generate the solutions to the Diophantine equations (3.0.1) and (3.0.2) for the highest order coefficient, and create data structures for dynamic storage. We stored each individual solution as a tuple of length n where the k^{th} entry denoted the value of the coefficient j_k. Every solution for the upper bound was then given a reference in a linked list. The next step is to iterate through the solutions stored in the linked list to compute the highest order coefficient with the algorithms outlined in Theorem 3.0.1 and Theorem 3.0.2. The solution stored for the upper bound can then be modified through by decrementing the value for j_n for each tuple and then deleting the reference to the tuple and deallocating the memory in the linked list when the value for j_n reaches zero. This will obtain all of the solutions for the next lowest-order coefficient, and the process may be repeated to obtain the solutions to all of the lower order coefficients.

There is a recursive method to generate these coefficients as outlined by Shiamuchi in [Shi1] and Bielefeld, Fisher, and v. Haeseler in [BieFH1]. Recursive methods are more practical for this computation because they are able to reuse information to compute the next coefficient, but we focus on computing the coefficients directly in order to get a better understanding of their basic structure. The main computational issue in generating these coefficients directly is the exponentially growing amount of computations needed due to the generalized binomial coefficients. Since the main computational sink is calculating the binomial coefficients, any way to reuse previously calculated information from the binomials would lead to a much more efficient program. A feasible solution could be to create a look-up table in parallel, but we were unable to find an efficient way to implement one. There is also an issue of memory as the coefficients are large rational numbers, so the numerators and denominators must be computed with arbitrary precision and stored separately to avoid truncation error. Arbitrary precision integers must be stored in the heap, so they have much slower access as well.

3.2. **Utilizing high-performance computing resources.**

To deal with the time sink in generating the binomial coefficients, we utilized multi-core parallel computing. Each coefficient can be computed independently after we obtain the solutions to the Diophantine equations. Therefore, we structure our code for concurrent computation and use generator expressions so that we can use multiple cores where our code is executing simultaneously.

Computing on local machines, we utilize 12 cores to compute the coefficients, with each core dealing with one coefficient at a time. When one coefficient calculation is done, the core takes the next awaiting task that is not being taken by other cores. To improve our calculations, we switched to machine with higher performance. Specifically, we chose a server machine with two 12 core Intel Haswell processors with 256 GB RAM, and we ran our code in a parallel environment with 72 valid cores. We obtained the first 1024 coefficients with a CPU time of 166 hours and a total run time of 7 days.

4. Benford's Law

Frank Benford's 1938 paper, *The Law of Anomalous Numbers* [Benf1], illustrated a profound result, the first digits of numbers in a given data set are not uniformly distributed in general. Benford applied statistical analysis to a variety of well-behaved but uncorrelated data sets such as the areas of rivers, financial returns, and lists of physical constants; in an overwhelming amount of the data, 1 appeared as the leading digit around 30% of the time, and each higher digit was successively less likely [Benf1, Mil1]. He then outlined the derivation of a statistical distribution which maintained that the probability of observing a leading digit, d, for a given base, b, is $\log_b\left(\frac{d+1}{d}\right)$ for such data sets [Benf1, Mil1]. This logarithmic relation is referred to as Benford's Law, and its resultant probability measure for base 10 is outlined in Figure 4.0.1. Benford's Law has been the subject of intensive research over the past several decades, arising in numerous fields; see [Mil1] for an introduction to the general theory and numerous applications.

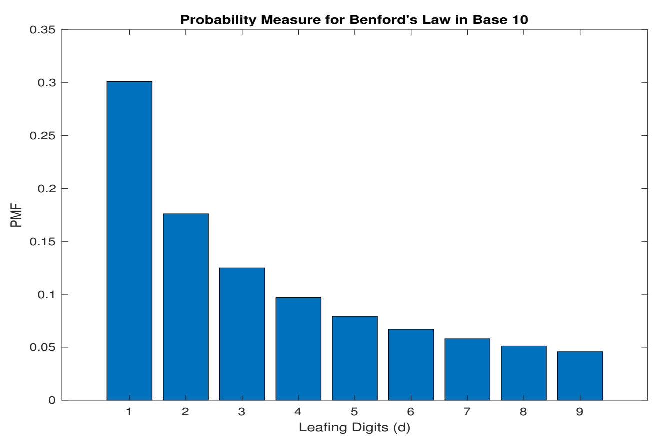


Figure 4.0.1. Probability Measure for Benford's Law in Base 10.

Benford's law appears throughout purely mathematical constructions as well as geometric series, recurrence relations, and geometric Brownian motion. Its ubiquity makes it one of the most interesting objects in modern mathematics, as it arises in many disciplines. Therefore, it worthwhile to consider non-traditional data sets where it may appear, such as fractal sets. Basic fractals like the Cantor set and Sierpinski triangle are obtained as the limiting iterations of sets, and as a result their unique component measures follow a geometric distribution which are Benford in most bases. Building on these results, it is plausible that more complicated fractals obey this distribution as well. We studied the Riemann mapping of the inverse of the Mandelbrot set to the exterior of the unit disk, along with its reciprocal. These mappings are given by a Taylor and Laurent series respectively, and we studied their coefficients to determine their fit for Benford's saw. These coefficients are of interest as their asymptotic convergence is intimately related to the conjectured local connectivity of the Mandelbrot set, which is an important open problem in complex dynamics.

4.1. **Statistical testing for Benford's law.**

A common statistical methodology for evaluating whether a data set is distributed according to Benford's law is to utilize the standard χ^2 goodness of fit test. As we are investigating Benford's law in base 10 we utilize 8 degrees of freedom for our χ^2 testing.[2] If there are N observations, letting $p_d = \log_{10}\left(\frac{d+1}{d}\right)$ be the Benford probability of having a first digit of d, we expect $p_d N$ values to have a first digit of d. If we let $\mathcal{O}_d$ be the observed number with a first digit of d, the χ^2 value is

$$\chi^2 \ := \ \sum_{d=1}^{9} \frac{(\mathcal{O}_d - p_d N)^2}{p_d N}.$$

With 8 degrees of freedom, if the data is Benford then 95% of the time the χ^2 value will be at most 15.51; thus we used 15.51 as our threshold for statistical significance. We considered the distribution of the χ^2 values as a function of the sample size on a logarithmic scale to account for random fluctuations and periodic behaviors. The distribution's limiting behavior in relation to our chosen threshold value provided our benchmark for the fit to Benford's law. An equivalent test is to consider the distribution of the base 10 logarithm of the data set modulo 1, as a necessary and sufficient condition for a Benford distribution is that this sequence converges to the uniform distribution ([MilTB1]).

The coefficients we studied are rational numbers, so we considered the distributions of the numerators, denominators, and decimal expansions separately. We considered the non-zero coefficients since zero is not defined for our probability measure, and since certain theorems and conjectures outlined by Shiamuchi in [Shi2] already describe the distribution of the zeroes in the coefficients. These tools will not explicitly prove that the distribution of the coefficients is Benford, and an analytical proof of such a result is outside the scope of this paper. Still, utilizing statistical testing yields insight to the distribution of the coefficients, and it is hoped that this research will prove useful for eventually devising a formal proof.

Above is an example of the coefficients computed. When the coefficient is 0, numerators are set to 0 and denominators to -. We then use them to compute the exact values in decimal expansion for a_m and b_m.

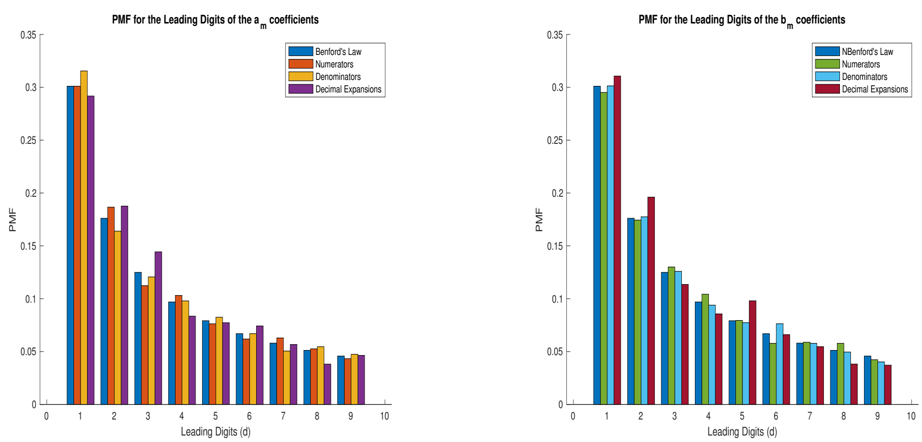
[2]There are nine possible first digits, but once we know the proportion that are digits 1 through 8 the percentage that start with a 9 is forced, and thus we lose one degree of freedom.

m	a_m num	a_m denom	b_m num	b_m denom
1	0	-	1	8
2	1	2	-1	4
3	1	8	15	128
4	1	4	0	-
5	15	128	-47	1024
6	0	-	-1	16
7	81	1024	987	32768
8	1	8	0	-
9	1499	32768	-3673	262144
10	1	32	1	32

TABLE 4.1.1. The table of first 10 coefficients computed.

4.2. **Benfordness of the Taylor and Laurent coefficients.**

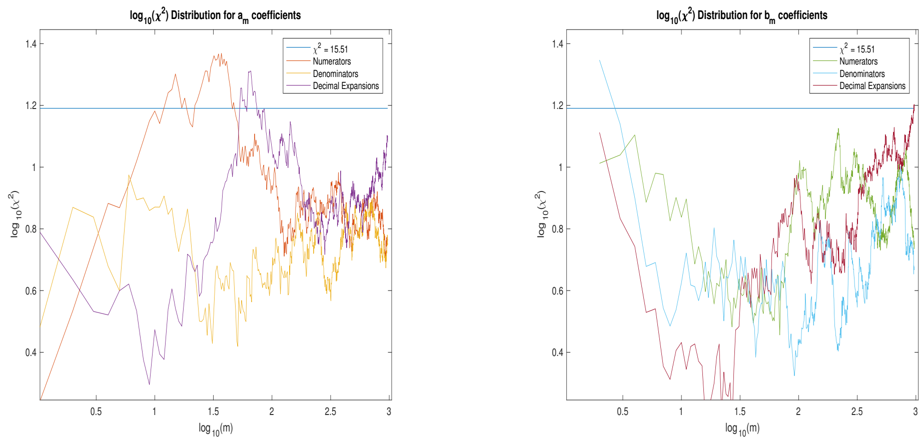
We examine the distribution of the first digits of the a_m and b_m coefficients. As mentioned earlier, we restrict our discussion to the non-zero coefficients. We examined the distribution of the leading digits for all of the data collected, which are shown in Figure 4.2.1. The data is good fit empirically for Benford's law, which motivates conducting more explicit statistical tests on the data set.



(A) a_m (B) b_m

FIGURE 4.2.1. Probability mass functions for the leading digits of the a_m and b_m coefficients.

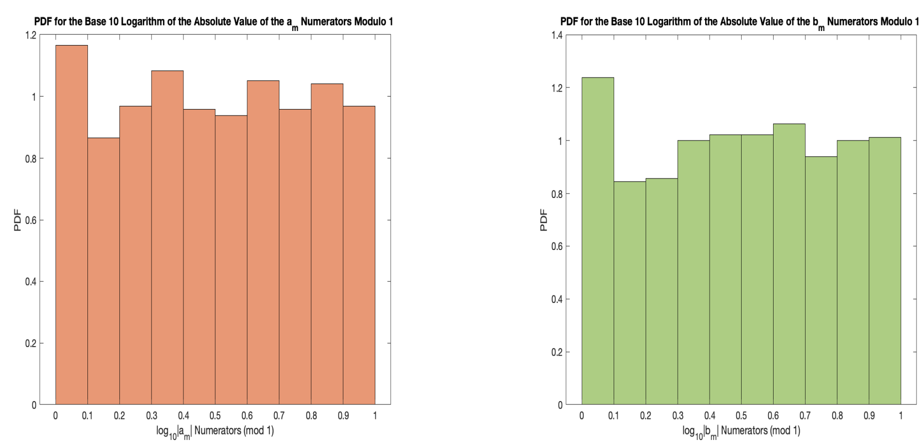
We conduct the χ^2 test and distribution of the base 10 logarithms modulo 1, as discussed above to evaluate the data. The plots of the χ^2 values are shown in Figure 4.2.2, and the plots of the logarithms modulo 1 are provided in Figures 4.2.3 - 4.2.5



(A) a_m (B) b_m

FIGURE 4.2.2. χ^2 distribution for the a_m and b_m coefficients.

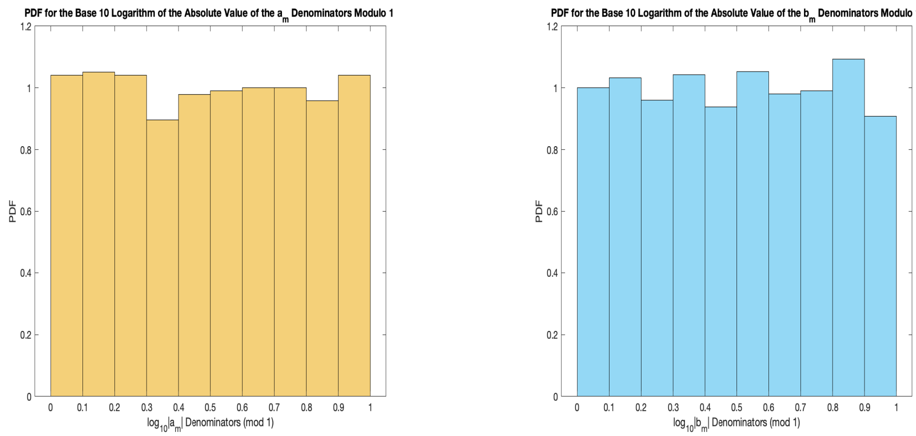
The data provides a relatively good fit for Benford's Law, as it tends to stay below a χ^2 value of 15.51. That being said, there is still noise in the data set, with many fluctuations that can push it above the threshold. More analysis is needed to determine exactly what is causing these fluctuations and how they might be obstructing the Benfordness of the coefficients.



(A) a_m (B) b_m

FIGURE 4.2.3. a_m and b_m numerators.

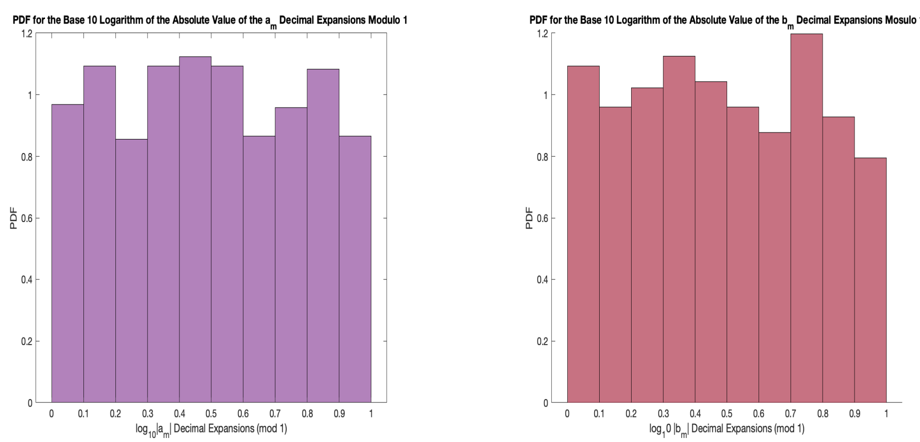
The numerators have a skew towards zero in the distribution of their logarithms. There is regularity when the b_m numerators are equal to 1 as discussed by Bielefeld, Fisher, and v. Haeseler in [BieFH1]. This helps to explain the skew for the b_m numerators, and it is plausible that a similar relation holds for the a_m numerators as well.



(A) a_m (B) b_m

Figure 4.2.4. a_m and b_m denominators.

The denominators consist of integer powers of 2. Since $\log_{10}(2)$ is irrational, the sequence $x_n = 2^n$ is Benford in base 10, and $\log_{10}(2^n) \pmod 1$ converges to the uniform distribution [MilTB1]. Since the denominators span many orders of magnitude, it is expected that they will similarly converge in distribution. There is noise in the se-



(A) a_m (B) b_m

Figure 4.2.5. a_m and b_m decimal expansions.

quence for the decimal expansions which may be attributed to the regularities of the numerators and the random fluctuations in the data set. It may still be reasonable to expect convergence to a uniform distribution, but it would likely converge slower due to interference from the numerators and denominators.

In general the data suggests a convergence to the uniform distribution, which is another indicator that the data is a good fit for Benford's law. We may also investigate how many orders of magnitude a data set spans on average by computing the arithmetic mean and standard deviation of $\log_{10}|x_n|$ to tell where the data set is centered and how spread out it is. It is typical, but not necessary, for a data set to be Benford if it spans many orders of magnitude (see Chapter 2 of [Mil1] and [BerH1] for an analysis that a sufficiently large spread is not enough to ensure Benfordness). Our findings for the Taylor and Laurent coefficients are summarized in Table 4.2.1.

data set	μ	σ
a_m numerators	192.20	178.42
a_m denominators	195.59	178.69
a_m decimals	-3.39	0.68
b_m numerators	192.00	178.40
b_m denominators	195.75	178.67
b_m decimals	-3.75	0.70

Table 4.2.1. Parameters for the $\log_{10}$ distribution of the data sets.

The numerators and denominators span many orders of magnitude, but the decimal expansions cluster. The mean for the decimal expansions being negative indicates that the denominators are larger than the numerators on average. The ratio between the growth rates of the numerators and denominators likely has some form of regularity as well to account for the small standard deviation, but more analysis would be needed to determine the exact relationship. These observations are consistent with the previously discussed conjecture that $0 < |b_m| < 1/m$ for all m. Ultimately, this provides insight into the growth of the coefficients and the shape of the data, and this motivates our conclusion that the numerators and denominators appear to converge to a Benford distribution more quickly than their respective decimal representations.

5. On the Taylor and Laurent coefficients

Theorem 2.1.4 highlights the relevance of the study of the Riemann mappings Ψ and Φ. Much effort has been put into the understanding the behaviour of both series. We now refer to a few important results and introduce new conjectures on the behaviour of the coefficients. Proofs may be found in Shimauchi's work [Shi1, Shi2, Shi3, Shi4].

Theorem 5.0.1. *The coefficients $a_{d,m}$ and $b_{d,m}$ are d-adic rational numbers. In other words, $a_{d,m}$ and $b_{d,m}$ are of the form*

$$\frac{p}{d^{-v}},$$

where p is an integer indivisible by d. The integer ν is, by definition, the d-adic valuation ν_d of $a_{d,m}$ or $b_{d,m}$.

In the following, we consider only the case $d = 2$. The majority of the results holds also for a general integer $d \geq 2$, under simple modifications.

5.1. **Results for the Taylor coefficients.**

Remark 5.1.1. For any integers k and ν satisfying $k \geq 1$ and $2^\nu \geq k+1$, let $m = (2k+1)2^\nu$. Then $a_{2,m} = 0$.

It is unknown whether the converse is true. The proof may be found in [EwiS1]. The authors have reported that their computation of 1000 terms of $a_{2,m}$ has not produced a zero-coefficient besides those indicated by the theorem, which is consistent with our observations. The result may also be expressed as following corollary:

Corollary 5.1.2. *Let $m = m_0 2^n$ with $n \geq 0$, and m_0 odd. If $3 \leq m_0 \leq 2^{n+1}$, then $a_{2,m} = 0$.*

Making use of the 2-adic valuation ([Shi2]), it is possible to obtain the following theorem.

Theorem 5.1.3. *We have $-\nu_2(a_m) \leq \nu_2((2m-2)!)$ for all m, with equality attained exactly when m is odd.*

In the following, since our interest is only for the 2-adic valuation, we will make use of the notation $\nu(x) := \nu_2(x)$.

Remark 5.1.4. In the case that m is odd, we may obtain an efficient algorithm to compute $-\nu(a_m)$ by making use of $\nu((2m-2)!)$. From the properties of the d-adic evaluation outlined in [Shi3], we have:

$$\nu((2m-2)!) = \sum_{l=1}^{\infty} \left\lfloor \frac{2m-2}{2^l} \right\rfloor = \sum_{l=0}^{\infty} \left\lfloor \frac{m-1}{2^l} \right\rfloor. \tag{5.1.1}$$

Therefore, if we set a value, N, the denominator's exponent for every odd number $m < 2^N$ is given by

$$-\nu(a_m) = \nu((2m-2)!) = \sum_{l=0}^{N} \left\lfloor \frac{m-1}{2^l} \right\rfloor. \tag{5.1.2}$$

In the following it is reported another relevant result, which is a summary of Theorem 3.1 and Corollary 3.5 from [Shi4] for the case that $d = 2$.

Theorem 5.1.5. *Given $m \in \mathbb{N} \setminus \{1\}$, we have that $-\nu(a_m) \leq x(m)$, where*

$$x(m) \;=\; \nu((m-1)!) + m - 1.$$

Under the same assumptions, the result is also true with

$$x(m) \;=\; \lceil \nu(m-1) + m - 1 \rceil .$$

5.2. **Results for the Laurent coefficients.**

Similar results hold for the b_m coefficients. We use the notation $m = 2^n m_0$, where m_0 is odd.

Remark 5.2.1. If $n \geq 2$ and $m_0 \leq 2^{n+1} - 5$, then $b_m = 0$.

It is still unknown whether the converse of this theorem is true. In [EwiS2], the only coefficients which have been observed to be zero are those mentioned in this theorem. The following result, from [BieFH1], underlines that a result similar to the one for the a_m holds.

Theorem 5.2.2. *We have $-\nu(b_m) \leq \nu((2m+2)!)$ for all m, and equality is attained exactly when m is odd.*

Don Zagier has made several observations and conjectures about the exponents of b_m. Our objective to extend them to the a_m coefficients. The original conjectures are outlined in [BieFH1].

Conjecture 5.2.3. *For the b_m coefficients, we have the following.*

- *If $n \;=\; 0$, then $-\nu(b_m) \;=\; \nu((2m+2)!)$.*
- *If $n \;=\; 1$, then $-\nu(b_m) \;=\; \nu(((2m+2)/3)!) + \epsilon(m_0)$, where $\epsilon(m_0) \;=\; 0$ if $m_0 \;\equiv\; 11$ mod 12 and 1 otherwise.*
- *If $n \;=\; 2$, then $-\nu(b_m) \;=\; \nu(((2m-25)/7)!) + \epsilon(m_0)$, where $\epsilon(m_0)$ moves with periodicity of 28, as expressed in* [BieFH1].

Remark 5.2.4. Theorem 5.2.2 and the results on $\nu(x!)$ reported in Remark 5.1.4 imply that the denominator's coefficients for b_m have regularity in term of consecutive differences for m odd. Zagier's conjectures posit similar patterns for m even. This gives costant discrete derivative in the denominator's exponents of b_m for $m \equiv 2 \bmod 4$ ($n = 1$), $m \equiv 4 \bmod 8$ ($n = 2$), $m \equiv 8 \bmod 16$ ($n = 3$); $\ldots$ $m \equiv 2^N \bmod 2^{N+1} (n = N)$. Note that for each of these subsequences, there are some starting 0-values. These correspond exactly with the values predicted by Corollaries 5.1.2 and 5.2.1 respectively; when $a_m = 0$, the algorithm gives 0 for the denominator's exponent. In general, for each n there is a partial periodicity with period $2(2^{n+1} - 1)$ in m_0, or of $2^{n+1}(2^{n+1} - 1)$ in m.

The previous remark is an essential part of current studies related to the behaviour of the coefficients. The correction term $\epsilon(m_0)$, appears to depend on n. Another possible solution, building on this periodicity, has been presented in the form of another conjecture, stated in [BieFH1].

Conjecture 5.2.5. *Let $m = 2^n m_0$ and m_0 be uniquely written as $2(2^{n+1}-1)k+l$, where l is odd, $k \geq 0$, and $1 \leq l \leq 2^{n+2}-3$. Then*

$$-\nu(b_m) = \nu((2^{n+2}k)!) + \tau(k,l) \tag{5.2.1}$$

where

- $\tau(k,l) = l-1$ *if* $l = 2^{n+2}-3$, *k odd,*
- $\tau(k,l) = l$ *if* $l = 2^{n+2}-3$, *k even, and*
- $\tau(k,l) = l+1$ *if* $2^{n+1}-3 \leq l \leq 2^{n+2}-5$.

Note that this result holds for every $n \geq 0$, but at fixed n it does not cover all the associated values of l except for the case $n = 1$. In any case, for a fixed n, it covers more than half of the values of m_0 of the form $2(2^{n+1}-1)k+l$, varying k and l.

One last possibility to tackle the problem is calculate the slope of each of the subsequences, since they seem to grow linearly. The observations have led to the Conjecture 1.0.1.

The numerators tend to follow a similar behaviour as the denominators. In particular, the modulus of the numerators in the subsequences tend to organize as follows: $\{a_m\}_{n=0} > \{a_m\}_{n=1} > \cdots \{a_m\}_{n=N}$. The possibility of bounding the numerators by making use of its associated denominator is a subject of further study.

5.3. **New remarks on the coefficients.**

We now extend the previous results to the a_m coefficients, following the same approach of Section 5.2.

Conjecture 5.3.1. *For the a_m coefficients we have:*

- *if* $n = 0$, $-\nu(a_m) = \nu((2m-2)!)$, *and*
- *if* $n = 1$, $-\nu(a_m) = \nu(((2m-2)/3)!) + \epsilon(m_0)$, *where* $\epsilon(m_0) = 1$ *if* $m_0 \not\equiv 3$ *mod 12. Otherwise, it follows the pattern in Table 5.3.1.*

This suggests a partial periodicity with period $2^4(2^{n+1}-1)$ in m_0, or of $2^{n+4}(2^{n+1}-1)$ in m. As before, it is possible to write m_0 as $2(2^{n+1}-1)k+l$, but it is more difficult to identify a general pattern in this case. The observations made are summarized in Conjecture 5.3.2. In particular, this suggests a dependence on n, which was not present in the previous result for b_m.

Conjecture 5.3.2. *Let m_0 be written in the same conditions of the Conjecture 5.2.5. Then*

- *if* $l = 2^{n+2}-3$, $-\nu(a_m) = \nu((2^{n+2}k)!) + l$, *and*
- *if* $l = 1$, $-\nu(a_m) = \nu((2^{n+2}k)!) + l + n - 1$.

Note that also in this case, as for Conjecture 5.2.5, for a fixed n, only specific values of l (and therefore of $m = 2^n m_0$) are covered. One of the objectives of further studies in this field may be to extend these results for every l.

$m_0 \pmod{192}$	$\epsilon(m_0)$
3	-2
15	-2
27	-4
39	-3
51	-2
63	-2
75	-2
87	-3
99	-4
111	-2
123	-2
135	-6
147	-3
159	-2
171	-2
183	-3

Table 5.3.1. The distribution of $\epsilon(m_0)$ when $m_0 \equiv_{12} 3$ has periodicity $16 * 12 = 192$. From $m_0 = 195$, it repeats itself, and will have the same $\epsilon(m_0)$ of $m_0 = 3$ and following.

6. Future Work

The most natural extension of our work would be utilizing an improved algorithm, for example the recursive algorithm, to generate more coefficients, allowing us to do more thorough statistical testing. Given more data, we could also go the route of looking at the coefficients over a certain zoom or averaging the coefficients over certain subsets. It would be particularly interesting if certain subsets of the coefficients also converge to Benford distribution.

The algorithms for computing the coefficients of the Mandelbrot can also be easily generalized to obtain other abstract Multibrot sets which could be analyzed using the same methods. We could look at the data in different bases to see if it Benfordness holds there. It would be interesting to see if the numerators and decimal expansions of the coefficients for the Multibrot set of degree d follow a Benford distribution in base d as it is trivial to see that the denominators will not.

The results of Section 5 also present interesting extensions for future work. In particular, Remark 5.2.4 suggests that dividing the coefficients in subsequences to be bounded separately may be the best approach to study the convergence of the Laurent series of the coefficients. This approach, which has not been followed in the past, to the best of our knowledge, could provide valuable results in the study of the local connectedness of $\mathcal{M}$.

References

[BanK1] C. Bandt and K. Keller, *Symbolic dynamics for angle-doubling on the circle. II. Symbolic description of the abstract Mandelbrot set*, Nonlinearity (1993), 377–392

[Benf1] F. Benford, *The Law of Anomalous Numbers*, Proceedings of the American Philosophical Society, **78** (1938), 551-572

[Beni1] A. M. Benini, *A survey on MLC, Rigidity and related topics*, arXiv, (2018)

[BerH1] A. Berger and T. P. Hill, *Benford's law strikes back: no simple explanation in sight for mathematical gem*, Math. Intelligencer **33** (2011), 21–30

[BieFH1] B. Bielefeld, Y. Fisher, and F. v. Haeseler, *Computing the Laurent Series of the Map* $\Psi : \mathbb{C}\backslash\overline{\mathbb{D}} \to \mathbb{C}\backslash\mathcal{M}$, Adv. Appl. Math., **14** (1993), 25–38

[DouH1] A. Douady and J. Hubbard, *Exploring the Mandelbrot set*, The Orsay Notes, (1985)

[EwiS1] J. H. Ewing and G. Schober, *Coefficients associated with the reciprocal of the Mandelbrot Set*, Journal of Mathematical Analysis and Applications, (1992), 104–114

[EwiS2] J. H. Ewing and G. Schober, *The area of the Mandelbrot set*, Numerische Mathematik, (1992), 59–72

[LuoY1] J. Luo and X. Yao, *To Generalize Carathéodory's Continuity Theorem*, arXiv, (2020)

[MilTB1] S. J. Miller and R. Takloo-Bighash, *An Invitation to Modern Number Theory*, Princeton University Press, Princeton, 2006

[Mil1] S. J. Miller, *Theory and Applications of Benford's Lawn*, Princeton University Press, Princeton, 2015

[Miln1] J. Milnor, *Dynamics in one complex variable: Third edition*, Princeton University Press, Princeton, 2011

[Shi1] H. Shiamuchi, *On the coefficients of the Riemann mapping function for the complement of the Mandelbrot set*, RIMS Kôkyûroku Bessatsu, (2011), 109–113

[Shi2] H. Shiamuchi, *On the coefficients of the Riemann mapping function for the exterior of the Mandelbrot set*, RIMS Kôkyûroku Bessatsu, (2012), 21–30

[Shi3] H. Shiamuchi, *A remark on Zagier's observation of the Mandelbrot set*, Osaka Journal of Mathematics, **52** (2015), 737–747

[Shi4] H. Shiamuchi, *Taylor Coefficients of the Conformal Map for the Exterior of the Reciprocal of the Multibrot Set*, Interdisciplinary Information Sciences, (2019)

[Shis1] M. Shishikura, *The Hausdorff Dimension of the Boundary of the Mandelbrot Set and Julia Sets*, Annals of Mathematics, **147** (1998), 225–267

[Sto1] D. Stoll, *A Brief Introduction to Complex Dynamics*, University of Chicago REU, (2015), 1–23

(Filippo Beretta) University of Milan

Email address: filo0820@gmail.com

(Jesse Dimino) Department of Mathematics, CUNY College of Staten Island, Staten Island, NY 10314

Email address: jessedimino@gmail.com, jesse.dimino@macaulay.cuny.edu

(Weike Fang) Department of Mathematics, University of Notre Dame, IN 46556

Email address: wfang@nd.edu

(Thomas C. Martinez) Department of Mathematics, Harvey Mudd College, Claremont, CA 90701

Email address: tmartinez@hmc.edu

(Steven J. Miller) Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267

Email address: sjm1@williams.edu, Steven.Miller.MC.96@aya.yale.edu

(Daniel Stoll) University of Michigan, Ann Arbor, MI 48109

Email address: dastoll@umich.edu