PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 00, Number 0, Pages 000–000
S 0002-9939(XX)0000-0

ON BENFORD'S LAW FOR MULTIPLICATIVE FUNCTIONS

VORRAPAN CHANDEE, XIANNAN LI, PAUL POLLACK, AND AKASH SINGHA ROY

(Communicated by Amanda Folsom)

Abstract. We provide a criterion to determine whether a real multiplicative function is a strong Benford sequence. The criterion implies that the k-divisor functions, where $k \neq 10^j$, and Hecke eigenvalues of newforms, such as Ramanujan tau function, are strong Benford. In contrast to some earlier work, our approach is based on Halász's Theorem.

1. Introduction

Benford's law is a phenomenon about the first digits of the numbers in data sets. In particular, the leading digits do not exhibit uniform distribution as might be naively expected, but rather, the digit 1 appears the most, followed by $2, 3$, and so on until 9. More precisely, we define Benford's law for a sequence of numbers below.

Definition 1.1. Let $\{a_k\}$ be a sequence of positive real numbers. Suppose that

$$a_k = 10^{M_k} \sum_{j=0}^{\infty} d_j^{(k)} 10^{-j},$$

where $M_k \in \mathbb{Z}$, $1 \leq d_0^{(k)} \leq 9$, and $0 \leq d_j^{(k)} \leq 9$ for all $j \geq 1$. $\{a_k\}$ satisfies strong Benford's law and is called a *strong Benford sequence* or *strong Benford*, if for all strings $S = d_0 d_1 ... d_{K-1}$, where $d_0 \neq 0$, and for all positive integers K,

$$\begin{aligned}\lim_{N\to\infty} \frac{\#\Big\{k \leq N \mid d_0^{(k)} d_1^{(k)} d_2^{(k)} .... d_{K-1}^{(k)} = d_0 d_1 d_2 ... d_{K-1}\Big\}}{N} &= \log_{10}\big(r + 10^{-K+1}\big) - \log_{10} r \\ &= \log_{10}\left(1 + \frac{1}{S}\right),\end{aligned}$$

where $r = \sum_{j=0}^{K-1} d_j 10^{-j}$.

Definition 1.2. Let $\{b_k\}$ be a sequence of non-negative real numbers. Let $\{b_j'\}$ be the subsequence of $\{b_k\}$ obtained by removing all zero terms. We say that $\{b_k\}$ is strong Benford if $\{b_j'\}$ is strong Benford.

2020 *Mathematics Subject Classification.* Primary 11N60, 11A41, 11B99 .
Simons Foundation Collaboration Grant for Mathematicians and NSF grant DMS-2101806.
Simons Foundation Collaboration Grant for Mathematicians.
NSF grant DMS-2001581.

Remark 1.3. When $K = 1$, the condition above becomes that for $d = 1, 2, ..., 9$,

$$\lim_{N\to\infty} \frac{\#\Big\{k \le N \mid d_0^{(k)} = d\Big\}}{N} = \log_{10}(d+1) - \log_{10} d = \log_{10}\bigg(1 + \frac{1}{d}\bigg).$$

From now on, we will refer to *strong Benford's law* simply as "Benford's law."

If a sequence of numbers is Benford, then the probability that its leading digit is 1 is $\log 2 \approx 30.1\%$ while that of 9 as a leading digit is only $\log(1+1/9) \approx 4.6\%$. The Benford's law phenomena was first observed by the astronomer Simon Newcomb in 1881 [15] when he noticed that the first pages of a book of logarithm tables were the most worn. Later, in 1938, Frank Benford [4] discovered similar patterns and found numerical evidence from a multitude of data sets, e.g. population numbers, areas of rivers, and physical constants.

A number of familiar sequences in mathematics such as Fibonacci sequences, exponential functions, and factorial functions have been proven to follow Benford's law. Sequences in number theory appear as well. For instance, Kontorovich and Miller [11] showed that the distribution of values of L-functions and certain statistics concerning the iterates of the $3x + 1$ problem follow Benford's law. A refinement of Lagarias and Soundararajan [12] proved that iterates of $3x + 1$ problem follow Benford's law for most initial seeds. Recently, the first author and Aursukaree [2] proved that the divisor function, which counts the number of positive divisors of n, is Benford.

The divisor function is an example of the wider class of real multiplicative functions. We say that h is a real multiplicative function if $h\colon \mathbb{N} \to \mathbb{R}$ satisfies that for all natural numbers m and n with $(m, n) = 1$,

$$h(mn) = h(m)h(n).$$

In this article, we will give a criterion for when real multiplicative functions follow Benford's law. Then we will apply the criterion to a number of interesting examples, including the k-divisor functions, Euler's Phi function and Hecke eigenvalues of new forms.

The proof of our criterion is a nice application of Halász's theorem on the sum of multiplicative functions. Roughly speaking, we find that $\{h(n)\}$ is a strong Benford sequence if and only if $e^{2\pi i\ell \log_{10}|h(n)|}$ is not "close to" $n^{i\alpha}$ for all nonzero integers ℓ and all real α. Another important feature of the criterion is that we can determine if $\{h(n)\}$ is Benford as soon as information at prime numbers is known.

After completing this work, we noticed that our criterion stated in Theorems 2.3 and 2.2 is essentially equivalent to the result of Delange in [5], where a criterion is proven for values of a real additive function to be uniformly distributed mod 1. The connection here is that for positive valued multiplicative functions $f(n)$, $\log f(n)$ is a real additive function. We have chosen to preserve our presentation and proof of the criterion for completeness. This is contained in Section 2, where we provide the statements of Halász's theorem, Weyl's criterion and then explicitly state and prove our main criterion (Theorem 2.3). Before that, we illustrate our criterion with some applications below.

1.1. **Applications of the criterion - Theorem 2.3.** We start with the sequence $\{n^a\}$ where a is fixed. Previously, this sequence has been proven to be not strong Benford by other methods, e.g. Fejér's Theorem (see [21]). Applying Theorem 2.3, we have an alternative simple proof of the same result.

Corollary 1.4. Let a be a fixed real constant. Then $\{n^a\}$ is not a strong Benford sequence.

Next, we will consider the k-divisor function, which counts the number of ways to write n as the product of k natural numbers, i.e.

$$d_k(n) = \sum_{n_1 n_2 \ldots n_k = n} 1.$$

As previously mentioned, $\{d_2(n)\} = \{d(n)\}$ is a strong Benford sequence. The first author and Aursukaree [2] used Selberg-Delange's method [13, Chapter 7], which involves analysis with contour integrals. Theorem 2.3 applies to $d_k(n)$ for general k to give an alternate simpler proof.

Corollary 1.5. Let $k \geq 2$, and $d_k(n)$ be the k-divisor function. Then $\{d_k(n)\}$ is a strong Benford sequence if and only if $k \neq 10^j$ for all positive integers j.

Note that when k is a power of 10, so is $d_k(p)$ for all primes p, and hence $d_k(n)$ is also a power of 10 for all squarefree integers n. Since the number of squarefree integers up to x is about $\frac{6}{\pi^2}x$, the probability of the first digit of $d_k(n)$ being 1 exceeds $60.8\% > 30.1\%$.

Another important multiplicative function is Euler's phi function $\varphi(n)$, which counts the number of positive integers up to n that are relatively prime to n. The formula is

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right).$$

Corollary 1.6. Let $\varphi(n)$ be Euler's phi function. Then $\{\varphi(n)\}$ is not a strong Benford sequence.

Finally, let $f(z) = \sum_{n=1}^{\infty} \lambda_f(n) q^n \in S_k^{new}(\Gamma_0(N))$ where $q = e^{2\pi i z}$, be a newform (i.e., a holomorphic cuspidal normalized Hecke eigenform) of even weight k and trivial nebentypus on $\Gamma_0(N)$ that does not have complex multiplication. We will investigate the sequence of Fourier coefficients $\{\lambda_f(n)\}$. A classic example of this type of sequence is the Ramanujan tau function $\tau(n)$. Previously, Jameson, Thorner and Ye [10] showed that the sequence $\{\lambda_f(p)\}$, where p is prime, does not satisfy Benford's law, but it does follow Benford's law with logarithmic density, which is [1]

$$\lim_{n\to\infty} \frac{\sum_{\substack{p \leq n \\ \text{the first } K\text{-digits of } \lambda_f(p) = S}} \frac{1}{p}}{\sum_{p \leq n} \frac{1}{p}} = \log_{10}\left(1 + \frac{1}{S}\right).$$

Our Theorem 2.3 allows us to consider the sequence over natural numbers n, not restricted to primes.

Corollary 1.7. Let $f(z) = \sum_{n=1}^{\infty} \lambda_f(n) q^n \in S_k^{new}(\Gamma_0(N))$ where $q = e^{2\pi i z}$, be a newform of even weight k and trivial nebentypus on $\Gamma_0(N)$ that does not have complex multiplication. Then $\{\lambda_f(n)\}$ is a strong Benford sequence.

We remark that previously Anderson, Rolen and Stoehr [1] proved that the non-zero coefficients of a special family of weakly holomorphic modular forms and certain partition functions are Benford. We refer the reader there for more precise statements.

[1] We state only a special case of [10], where the base is 10.

As an application of Theorem 2.2 below, it is not difficult to show that the collection of multiplicative functions $f\colon \mathbb{N} \to \mathbb{R} - \{0\}$ for which $\{f(n)\}$ is not a strong Benford sequence forms a group under pointwise multiplication. This observation is equivalent to a result established in p. 51 of [5]. As an illustration, we observe that since $\{d(n)\}$ is a strong Benford sequence while $\{\phi(n)\}$ is not, $\{d(n)\phi(n)\}$ is a strong Benford sequence.

Finally, we remark that while we prefer to work in base 10, the definitions and results can all be ported easily to an arbitrary base $g \geq 2$. No significant change is needed to either the results or the proofs, but we must be somewhat careful when formulating the correct analogue of Corollary 1.4: We find that $d_k(n)$ is strong Benford in base g precisely when $\log_g k \notin \mathbb{Q}$.

2. The Main Theorem

Our criterion for real multiplicative functions to satisfy Benford's law is stated in terms of the "distance" between two multiplicative functions. We start by defining this distance.

Definition 2.1. Let f and g be multiplicative functions taking values in the unit disc $\{z \in \mathbb{C} : |z| \leq 1\}$. The distance up to x between f and g is defined to be

$$\mathbb{D}(f, g; x) := \left(\sum_{p \leq x} \frac{1 - \mathrm{Re}(f(p)\overline{g(p)})}{p} \right)^{1/2}.$$

This notion of distance comes from the work of Granville and Soundararajan [9]. It is known (see [9, p. 364]) that the distance function satisfies the triangle inequality, i.e. for any multiplicative function f, f', g, and g',

$$\mathbb{D}(ff', gg'; x) \ \leq \ \mathbb{D}(f, g; x) + \mathbb{D}(f', g'; x).$$

If $\mathbb{D}(f, g; \infty) < \infty$, then we say that f *pretends* to be g. Let $h\colon \mathbb{N} \to \mathbb{R} - \{0\}$ be a multiplicative function. It is obvious that $f_\ell(n) := e^{2\pi i \ell \log_{10} |h(n)|}$ is also multiplicative.

We will start with the criterion for nonzero multiplicative functions. The criterion for $\{h(n)\}$ to be a strong Benford sequence is that f_ℓ does not pretend to be $n^{i\alpha}$ for all $\alpha \in \mathbb{R}$ and all nonzero integers ℓ. More precisely, we have the following.

Theorem 2.2. *Let $h\colon \mathbb{N} \to \mathbb{R} - \{0\}$ be a multiplicative function. Let $f_\ell(n) = e^{2\pi i \ell \log_{10} |h(n)|}$. $\{h(n)\}$ is a strong Benford sequence if and only if $\mathbb{D}(f_\ell, n^{i\alpha}; \infty) = \infty$ for all $\alpha \in \mathbb{R}$ and all $\ell \neq 0$.*

For multiplicative functions which are possibly zero for some n, we need an additional condition on the frequency of nonzero terms.

Theorem 2.3. *Let $h\colon \mathbb{N} \to \mathbb{R}$ be a multiplicative function. Define*

$$\mathcal{T}_N = |\{n \leq N \mid h(n) \neq 0\}|.$$

Suppose $\frac{N}{\mathcal{T}_N} \ll 1$. Let $f_\ell(n)$ be defined as in Theorem 2.2, and $g_\ell(n) = f_\ell(n)$ if $h(n) \neq 0$ and is 0 otherwise. Then $\{h(n)\}$ is a strong Benford sequence if and only if $\mathbb{D}(g_\ell, n^{i\alpha}; \infty) = \infty$ for all $\alpha \in \mathbb{R}$ and all $\ell \neq 0$.

Remark 2.4. $g_\ell(n)$ is also multiplicative. For $(m,n)=1$, if both $h(m)$ and $h(n)$ are not zero, we have

$$g_\ell(mn) = f_\ell(mn) = f_\ell(m)f_\ell(n) = g_\ell(m)g_\ell(n).$$

Otherwise

$$g_\ell(m)g_\ell(n) = 0 = g_\ell(mn).$$

Remark 2.5. Theorem 2.2 is a corollary of Theorem 2.3, where $\mathcal{T}_N = N$ and $g_\ell(n) = f_\ell(n)$. We state Theorem 2.2 for clarity.

The notion of being strong Benford is closely connected to the notion of being uniformly distributed modulo 1, which we define formally below.

Definition 2.6. A sequence $\{a_k\}$ is uniformly distributed modulo 1 if and only if the fractional parts of all numbers in the sequence distribute uniformly on the interval $[0,1]$, i.e.,

$$\lim_{n\to\infty} \frac{|\{k \le n : \ a_k \pmod 1 \in (a,b)\}|}{n} = b-a,$$

where $(a,b) \subset [0,1]$.

Diaconis [7] showed that being strong Benford is equivalent to having a base 10 logarithm that is uniformly distributed mod 1; the following result is Theorem 1 of [7].

Lemma 2.7. *A nonzero sequence $\{a_k\}$ is a strong Benford sequence if and only if the sequence $\{\log_{10}|a_k|\}$ is uniformly distributed modulo 1.*

A classic result of Weyl gives a necessary and sufficient condition for a sequence to be uniformly distributed mod 1.

Theorem 2.8 (Weyl's criterion)**.** *The sequence $\{a_k\}$ is uniformly distributed modulo 1 if and only if*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} e^{2\pi i \ell a_k} = 0$$

for all integers $\ell \neq 0$.

We recommend Travaglini's text [19] (see in particular Chapter 6) for a very readable exposition of Weyl's criterion as well as an account of its connection with Benford's law.

Finally, Halász's Theorem enables us to understand the averages of multiplicative functions by comparing them to $n^{i\alpha}$.

Theorem 2.9 (Halász's Theorem)**.** *Let f be a multiplicative function with $|f(n)| \le 1$ for all integers n. If $\mathbb{D}(f, n^{i\alpha}; \infty) = \infty$ for all $\alpha \in \mathbb{R}$, then f has mean value zero, in the sense that*

$$\lim_{x\to\infty} \frac{1}{x}\sum_{n\le x} f(n) = 0.$$

Otherwise there is a unique $\alpha \in \mathbb{R}$ with $\mathbb{D}(f, n^{i\alpha}; \infty) < \infty$. In that case, f has mean value 0 if and only if

$$f(2^k) = -2^{ik\alpha} \quad \textit{for every positive integer } k. \tag{2.1}$$

Theorem 2.9 is essentially contained in Theorem 6.3 on pp. 226–227 of Elliott's monograph [8]. The uniqueness of α with $\mathbb{D}(f, n^{i\alpha}; \infty) < \infty$ is not explicit in that statement but is proved on p. 248 of that reference.

By Weyl's criterion, it suffices to apply Halász's Theorem to $g_\ell(n)$. As mentioned earlier, Theorem 2.2 is a simple corollary of 2.3. Hence, we focus on Theorem 2.3.

2.1. **Proof of Theorem 2.3.** Let $\{h^*(k)\}$ be the subsequence of $\{h(n)\}$ where all zero terms are removed. By Weyl's criterion, $\{h^*(k)\}$ is a strong Benford sequence if and only if

$$\lim_{\mathcal{T}_N \to \infty} \frac{1}{\mathcal{T}_N} \sum_{k=1}^{\mathcal{T}_N} e^{2\pi i \ell \log_{10} |h^*(k)|} = 0$$

for all nonzero integers ℓ. However, the function $e^{2\pi i \ell \log_{10} |h^*(k)|}$ is not necessarily multiplicative so we cannot apply Halász's Theorem. Therefore we add back some zero terms, expressed in terms of the function g_ℓ from Theorem 2.3. Recall that $g_\ell(n) = e^{2\pi i \ell \log_{10} |h(n)|}$ if $h(n) \neq 0$ and $g_\ell(n) = 0$ otherwise. Thus,

$$\frac{1}{\mathcal{T}_N} \sum_{k=1}^{\mathcal{T}_N} e^{2\pi i \ell \log_{10} |h^*(k)|} = \frac{N}{\mathcal{T}_N} \frac{1}{N} \sum_{n=1}^{N} g_\ell(n).$$

Since $\frac{N}{\mathcal{T}_N} \ll 1$, $\{h^*(k)\}$ is a strong Benford sequence if and only if

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} g_\ell(n) = 0$$

for all integers $\ell \neq 0$. The "if" direction of Theorem 2.3 now follows immediately from applying Halász's Theorem (Theorem 2.9) to the functions g_ℓ.

For the "only if" direction, suppose that $\{h^*(k)\}$ is strong Benford, so that each g_ℓ, with $\ell \neq 0$, has mean value 0. By Theorem 2.9, it is enough to show that there is no nonzero integer ℓ and real number α for which $\mathbb{D}(g_\ell, n^{i\alpha}; \infty) < \infty$ and $g_\ell(2^k) = -2^{ik\alpha}$ for all positive integers k. If such ℓ and α exist, consider $g_{2\ell} = g_\ell^2$. By the triangle inequality,

$$\mathbb{D}(g_{2\ell}, n^{2i\alpha}; \infty) \leq 2 \cdot \mathbb{D}(g_\ell, n^{i\alpha}; \infty) < \infty.$$

Also, $g_{2\ell}(2) = (g_\ell(2))^2 = 2^{i(2\alpha)} \neq -2^{i(2\alpha)}$. Invoking Halász's theorem again, we see that $g_{2\ell}$ does not have mean value 0, a contradiction.

3. Proof of Corollary 1.4 – the function n^a

Here, $f_\ell(p) = e^{2\pi i \ell \log_{10} p^a} = p^{\frac{2\pi i \ell a}{\ln 10}}$. Let $\alpha_\ell = \frac{2\pi \ell a}{\ln 10}$. It is easy to see that

$$\mathbb{D}(f_\ell, n^{i\alpha_\ell}; \infty)^2 = \sum_p \frac{1 - \mathrm{Re}(f_\ell(p) p^{-i\alpha_\ell})}{p} = 0.$$

By Theorem 2.2, $\{n^a\}$ is not Benford.

4. Proof of Corollary 1.5 – k-divisor functions

For this case, $f_\ell(n) = e^{2\pi i \ell \log_{10} d_k(n)}$ so $f_\ell(p) = e^{2\pi i \ell \log_{10} k}$. The corollary will follow if we show the following.

(1) For $k \neq 10^j$,

$$\mathbb{D}(f_\ell, n^{i\alpha}; \infty)^2 = \sum_p \frac{1 - \mathrm{Re}(e^{2\pi i \ell \log_{10} k} p^{-i\alpha})}{p} = \infty \tag{4.1}$$

for all $\alpha \in \mathbb{R}$.

(2) For $k = 10^j$, $\mathbb{D}(f_\ell, 1; \infty) < \infty$ for some $\ell \neq 0$.

First, we consider the case $k = 10^j$, and we chose $\ell = 1$.

$$\mathbb{D}(f_1, 1; \infty)^2 = \sum_p \frac{1 - \mathrm{Re}(e^{2\pi i \log_{10} 10^j})}{p} = \sum_p \frac{1-1}{p} = 0.$$

Thus by Theorem 2.3, $\{d_k(n)\}$ is not a Benford sequence.

Now we consider the case $k \neq 10^j$. Here, we will use some classical results. The proof of the Lemma below can be found in [13, Theorem 2.7].

Lemma 4.1 (Mertens' Theorem)**.** *For $x \geq 2$, we have*

$$\sum_{p \leq x} \frac{1}{p} = \ln\ln x + A + \mathfrak{b}(x), \qquad \textit{where} \quad \mathfrak{b}(x) = O\left(\frac{1}{\ln x}\right),$$

and

$$\sum_{p \leq x} \frac{\ln p}{p} = \ln x + O(1).$$

We also state some classical bounds on $\zeta(s)$ near the line $\mathrm{Re}(s) = 1$. We refer the reader to (3.5.1) and (3.11.8) of Titchmarsh's book [18].

Lemma 4.2. *Let $s = \sigma + it$. There exists some constant $c > 0$ such that for $1 - \frac{c}{\ln(|t|+4)} < \sigma$ and $|t| > 2$, we have*

$$\frac{1}{\ln|t|} \ll \zeta(s) \ll \ln|t|.$$

On the other hand, if $|t| \leq 2$,

$$\zeta(s) \ll \frac{1}{|t|} + O(1).$$

4.1. **Proof of** (4.1) **when** $k \neq 10^j$**.** Note that

$$\mathrm{Re}(e^{2\pi i \ell \log_{10} k} p^{-i\alpha}) = \cos(2\pi \ell \log_{10} k)\cos(\alpha \ln p) + \sin(2\pi \ell \log_{10} k)\sin(\alpha \ln p).$$

When $\alpha = 0$,

$$\mathbb{D}(f_\ell, 1; \infty)^2 = \sum_p \frac{1 - \cos(2\pi \ell \log_{10} k)}{p}.$$

Since $\cos(2\pi \ell \log_{10} k) < 1$ when $k \neq 10^j$, $\mathbb{D}(f_\ell, 1; \infty) = \infty$ by Lemma 4.1.

From now on, we focus on $\alpha \neq 0$. Since $\sum_p \frac{1}{p} = \infty$, to prove (4.1), it suffices to show that

$$\sum_{p \leq x} \frac{\cos(\alpha \ln p)}{p} = O(1), \quad \text{and} \quad \sum_{p \leq x} \frac{\sin(\alpha \ln p)}{p} = O(1)$$

uniformly in x. Since $\mathrm{Re}(p^{i\alpha}) = \cos(\alpha \ln p)$ and $\mathrm{Im}(p^{i\alpha}) = \sin(\alpha \ln p)$, it suffices to prove the following Lemma.

Lemma 4.3. *Let α be a fixed nonzero real number. Then*

$$\sum_{p\le x}\frac{1}{p^{1+i\alpha}} = O_\alpha(1),$$

where the implied constant depends on α.

Proof. We claim that for any $y \ge 2$,

$$\sum_{p\le y}\left(\frac{1}{p^{1+i\alpha}} - \frac{1}{p^{1+\frac{1}{\ln y}+i\alpha}}\right) = O(1), \tag{4.2}$$

and

$$\sum_{p\le y}\frac{1}{p^{1+\frac{1}{\ln y}+i\alpha}} = \ln\zeta\left(1+\frac{1}{\ln y}+i\alpha\right) + O(1). \tag{4.3}$$

Here (4.2) follows from

$$\left|\sum_{p\le y}\left(\frac{1}{p^{1+i\alpha}} - \frac{1}{p^{1+\frac{1}{\ln y}+i\alpha}}\right)\right| \le \sum_{p\le y}\frac{1-e^{-\frac{\ln p}{\ln y}}}{p} \ll \frac{1}{\ln y}\sum_{p\le y}\frac{\ln p}{p} \ll 1$$

by Lemma 4.1. For (4.3), we start from the fact that

$$\zeta(s) = \prod_p\left(1-\frac{1}{p^s}\right)^{-1}$$

for $\mathrm{Re}(s) > 1$. Taking logarithms and making straightforward estimates,

$$\ln\zeta\left(1+\frac{1}{\ln y}+i\alpha\right) - \sum_{p\le y}\frac{1}{p^{1+\frac{1}{\ln y}+i\alpha}} = \sum_{p>y}\frac{1}{p^{1+\frac{1}{\ln y}+i\alpha}} + O(1).$$

Keeping in mind the first assertion of Lemma 4.1,

$$\sum_{p>y}\frac{1}{p^{1+\frac{1}{\ln y}+i\alpha}} \ll \sum_{p>y}\frac{1}{p^{1+\frac{1}{\ln y}}} \ll \sum_{j\ge 0} e^{-2^j}\sum_{y^{2^j}<p\le y^{2^{j+1}}}\frac{1}{p} \ll \sum_{j\ge 0} e^{-2^j} \ll 1,$$

and (4.3) follows.

If $x \le 2$, the estimate of Lemma 4.3 is trivial. Otherwise, by (4.2) and (4.3),

$$\begin{aligned}\sum_{p\le x}\frac{1}{p^{1+i\alpha}} &= \ln\zeta\left(1+\frac{1}{\ln x}+i\alpha\right) + O(1)\\ &\ll_\alpha 1,\end{aligned} \tag{4.4}$$

by Lemma 4.2. □

5. Proof of Corollary 1.6 – Euler's phi function

For Euler's phi function, $f_\ell(p) = e^{2\pi\ell i\log_{10}(p-1)} = (p-1)^{\frac{2\pi i\ell}{\ln 10}}$. Let $\alpha = \frac{2\pi}{\ln 10}$. The Corollary follows from

$$\mathbb{D}(f_1, n^{i\alpha};\infty)^2 < \infty.$$

Now,

$$\mathbb{D}(f_1, n^{i\alpha};\infty)^2 = \sum_p\frac{1-\mathrm{Re}((p-1)^{i\alpha}p^{-i\alpha})}{p} = \sum_p\frac{1-\cos\left(\alpha\ln\left(1-\frac{1}{p}\right)\right)}{p}.$$

For large prime p such that $|\frac{\alpha}{p}| < \frac{1}{2}$,

$$\cos\left(\alpha \ln\left(1-\frac{1}{p}\right)\right) = \cos\left(\alpha\left(-\frac{1}{p}+O\left(\frac{1}{p^2}\right)\right)\right) = 1+O\left(\frac{\alpha^2}{p^2}\right).$$

Let $J = \max\{10, 2|\alpha|\}$. Thus

$$\sum_p \frac{1-\cos\left(\alpha\ln\left(1-\frac{1}{p}\right)\right)}{p} \ll \sum_{p\le J}\frac{1}{p}+\sum_{p>J}\frac{\alpha^2}{p^3} \ll \ln\ln J + \frac{\alpha^2}{J^2} = O_\alpha(1).$$

Thus, we conclude that $\{\varphi(n)\}$ is not a strong Benford sequence.

6. Proof of Corollary 1.7 – Hecke eigenvalues of newforms

Contrary to the previously considered multiplicative functions, $\lambda_f(n)$ might be 0 for some n. Our Corollary follows from Theorem 2.3 and the following lemmas.

Lemma 6.1. *Let $\mathcal{T}_N$ be defined as in Theorem 2.3. Then*

$$\frac{N}{\mathcal{T}_N} \ll 1.$$

Lemma 6.2. *Let $g_\ell(n) = e^{2\pi i \ell \log_{10}|\lambda_f(n)|}$ if $\lambda_f(n)\neq 0$ and 0 otherwise. Then*

$$\mathbb{D}(g_\ell, n^{i\alpha}; \infty) = \infty$$

for all $\alpha\in\mathbb{R}$ and $\ell\neq 0$.

Lemma 6.1 follows immediately from the work of Serre stated in Theorem 6.4 below. We start by stating some now well known properties of $\lambda_f(p)$.

6.1. Properties of $\lambda_f(p)$.

Weil conjectured that for all primes p,

$$|\lambda_f(p)| \le 2p^{\frac{k-1}{2}},$$

and this is proven by Deligne [6]. Therefore for each p, there exists unique $\theta_p\in[0,\pi]$ such that

$$\lambda_f(p) = 2p^{\frac{k-1}{2}}\cos\theta_p.$$

Sato and Tate studied the distribution of $\cos\theta_p$, varying through p for newforms f associated with elliptic curves. They conjectured that $\{\cos\theta_p\}$ is equidistributed in $[-1,1]$ with respect to a certain measure. Later Barnet-Lamb, Geraghty, Harris, and Taylor [3] proved the conjecture, and in fact they generalized it for the larger class of Hecke newforms, which we state below.

Theorem 6.3 (The Sato-Tate Conjecture)**.** *Let $f(z)=\sum_{n=1}^\infty \lambda_f(n)q^n \in S_k^{new}(\Gamma_0(N))$ be a newform of even weight $k\ge 2$ without complex multiplication. Let $F\colon[-1,1]\to\mathbb{C}$ be a Riemann-integrable function. Then*

$$\lim_{x\to\infty}\frac{1}{\pi(x)}\sum_{p\le x}F(\cos(\theta_p)) = \int_{-1}^1 F(t)\,\mathrm{d}\mu_{ST},$$

where $\pi(x)=\#\{p\le x\}$ is the prime counting function, and

$$\mathrm{d}\mu_{ST} = \frac{2}{\pi}\sqrt{1-t^2}\,\mathrm{d}t.$$

The next result was proven by Serre in [16] (see Theorem 15, the accompanying Corollary 2 and Theorem 16). Interested readers may also peruse the later works of Wan [20] for an improvement, and of Murty [14] and Thorner and Zaman [17] for further refinements.

Theorem 6.4. *Let $f(z)$ be defined as in Theorem 6.3. Then for any $\delta < 1/2$,*

$$\#\{p \le x : \lambda_f(p) = 0\} \ll_{f,\delta} \frac{x}{\log^{1+\delta} x}.$$

Hence

$$\sum_{\substack{p \\ \lambda_f(p)=0}} \frac{1}{p} < \infty,$$

and

$$\#\{n \le x : \lambda_f(n) \neq 0\} \gg x.$$

Note that Serre's result immediately implies Lemma 6.1. We prove Lemma 6.2 below.

6.2. **Proof of Lemma 6.2.** From the definition of $g_\ell(n)$,

$$\mathbb{D}(g_\ell, n^{i\alpha}; \infty)^2 = \sum_{\substack{p \\ \lambda_f(p)\neq 0}} \frac{1 - \mathrm{Re}(e^{2\pi i \ell \log_{10} |\lambda_f(p)|} p^{-i\alpha})}{p} + \sum_{\substack{p \\ \lambda_f(p)=0}} \frac{1}{p}.$$

We will deduce Lemma 6.2 from the following Lemma. We remind the reader that $\lambda_f(p) = 2p^{(k-1)/2}\cos\theta_p$.

Lemma 6.5. *Let $\beta = \frac{2\pi\ell}{\ln 10}$ for $\ell \neq 0$, and $\gamma = \frac{k-1}{2}\beta - \alpha$. Then*

$$\left| \sum_{\substack{p \le x \\ \lambda_f(p)\neq 0}} \frac{\mathrm{Re}\left(|2\cos\theta_p|^{i\beta} p^{i\gamma}\right)}{p} \right| = (\mathcal{K}_\gamma + o(1)) \ln\ln x$$

where

$$\mathcal{K}_\gamma = \begin{cases} \frac{4}{\pi}\int_0^1 \cos(\beta\ln(2y))\sqrt{1-y^2}\,\mathrm{d}y & \text{if } \gamma = 0 \\ 0 & \text{otherwise.} \end{cases}$$

There is an interval of positive measure inside $[0,1]$ on which $\cos(\beta\ln(2y)) < 1$, so by continuity,

$$\begin{aligned} &\frac{4}{\pi}\int_0^1 \cos(\beta\ln(2y))\sqrt{1-y^2}\,\mathrm{d}y \\ &< \frac{4}{\pi}\int_0^1 \sqrt{1-y^2}\,\mathrm{d}y = 1, \end{aligned} \tag{6.1}$$

so $\mathcal{K}_\gamma < 1$. Thus, by Lemma 6.5,

$$\mathbb{D}(g_\ell, n^{i\alpha}; x)^2 = \sum_{\substack{p \le x \\ \lambda_f(p)\neq 0}} \frac{1 - \mathrm{Re}\left(|2\cos\theta_p|^{i\beta} p^{i\gamma}\right)}{p} + \sum_{\substack{p \le x \\ \lambda_f(p)=0}} \frac{1}{p} \gg \ln\ln x$$

since $\sum_{p\le x}\frac{1}{p}\sim \ln\ln x$ (Lemma 4.1), and so

$$\mathbb{D}(g_\ell, n^{i\alpha};\infty)=\infty,$$

as required for Lemma 6.2. We now prove Lemma 6.5.

Proof. Note that

$$\operatorname{Re}\big(|2\cos\theta_p|^{i\beta}p^{i\gamma}\big)=\cos(\beta\ln(2|\cos\theta_p|))\cos(\gamma\ln p)-\sin(\beta\ln(2|\cos\theta_p|))\sin(\gamma\ln p).$$

Thus, it suffices to show that

$$\sum_{\substack{p\le x\\ \lambda_f(p)\neq 0}}\frac{\cos(\beta\ln(2|\cos\theta_p|))\cos(\gamma\ln p)}{p}=(\mathcal{K}_\gamma+o(1))\ln\ln x \tag{6.2}$$

and

$$\sum_{\substack{p\le x\\ \lambda_f(p)\neq 0}}\frac{\sin(\beta\ln(2|\cos\theta_p|))\sin(\gamma\ln p)}{p}=o(\ln\ln x). \tag{6.3}$$

We proceed to prove (6.2), noting that (6.3) follows similarly.

By partial summation, the left-hand side of (6.2) is

$$\left(\sum_{\substack{p\le x\\ \lambda_f(p)\neq 0}}\cos(\beta\ln(2|\cos\theta_p|))\right)\frac{\cos(\gamma\ln x)}{x}-\int_2^x\left(\sum_{\substack{p\le t\\ |\cos\theta_p|>0}}\cos(\beta\ln(2|\cos\theta_p|))\right)\mathrm{d}\frac{\cos(\gamma\ln t)}{t} \tag{6.4}$$

Let $I=[-1,0)\cup(0,1]$ and $\chi_I(x)$ denote the characteristic function of I, which is 1 if $x\in I$, and 0 otherwise. Note that $\cos(\beta\ln(2|y|))$ is Riemann integrable over $[-1,1]$, the singularity at $y=0$ being benign.

Thus, by the Sato-Tate conjecture (Theorem 6.3),

$$\begin{aligned}\lim_{x\to\infty}\frac{1}{\pi(x)}\sum_{\substack{p\le x\\ \lambda_f(p)\neq 0}}\cos(\beta\ln(2|\cos\theta_p|))&=\lim_{x\to\infty}\frac{1}{\pi(x)}\sum_{p\le x}\cos(\beta\ln(2|\cos\theta_p|))\chi_I(\cos\theta_p)\\&=\frac{4}{\pi}\int_0^1\cos(\beta\ln(2y))\sqrt{1-y^2}\,\mathrm{d}y=\mathcal{K}_0,\end{aligned}$$

since the integrand is even. By the prime number theorem, $\pi(x)\sim\frac{x}{\ln x}$, we obtain that (6.4) is

$$\begin{aligned}&(\mathcal{K}_0\pi(x)+o(\pi(x)))\frac{\cos(\gamma\ln x)}{x}-\int_2^x(\mathcal{K}_0\pi(t)+o(\pi(t)))\,\mathrm{d}\frac{\cos(\gamma\ln t)}{t}\\&=\mathcal{K}_0\left(\pi(x)\frac{\cos(\gamma\ln x)}{x}-\int_2^x\pi(t)\,\mathrm{d}\frac{\cos(\gamma\ln t)}{t}\right)+o\left(\frac{1}{\ln x}+\int_2^x\frac{1}{t\ln t}\,\mathrm{d}t\right)\\&=\mathcal{K}_0\sum_{p\le x}\frac{\cos(\gamma\ln p)}{p}+o(\ln\ln x).\end{aligned} \tag{6.5}$$

We consider two cases.

Case 1: $\gamma \neq 0$. Here $\mathcal{K}_\gamma = 0$. Applying Lemma 4.3 with $\alpha = -\gamma$ implies that

$$\sum_{p \leq x} \frac{\cos(\gamma \ln p)}{p} = O_\gamma(1).$$

Thus (6.5) is $o(\ln \ln x)$, so we derive the desired result when $\gamma \neq 0$.

Case 2: $\gamma = 0$. Here, $\mathcal{K}_\gamma = \mathcal{K}_0$, and we have

(6.6)
$$\sum_{\substack{p \leq x \\ \lambda_f(p) \neq 0}} \frac{\cos(\beta \ln(2|\cos\theta_p|))\cos(\gamma \ln p)}{p} = \mathcal{K}_0 \sum_{p \leq x} \frac{1}{p} + o(\ln \ln x) = \mathcal{K}_0 \ln \ln x + o(\ln \ln x),$$

as desired. □

Acknowledgement

The authors would also like to thank the reviewer for his/her careful reading and thoughtful suggestions.

References

1. T. Anderson, L. Rolen, and R. Stoehr, *Benford's law for coefficients of modular forms and partition functions.* Proc. Amer. Math. Soc. 139 (2011), no. 5, 1533–1541.
2. S. Aursukaree and V. Chandee, *Equidistribution of* $\log(d(n))$, Proceedings of Annual Pure and Applied Mathematics Conference, Chulalongkorn University, Thailand, May 2016, 399–410.
3. T. Barnet-Lamb, D. Geraghty, M. Harris, and R. Taylor, *A family of Calabi-Yau varieties and potential automorphy II*, Publ. Res. Inst. Math. Sci. 47 (2011), no. 1, 29–98.
4. F. Benford, *The law of anomalous numbers*, Proc. Amer. Philos. Soc. 78 (1938), no. 4, 551–572.
5. H. Delange, *Quelques résultats nouveaux sur les fonctions additives*, Mémories de la S. M. F., 25 (1971), p. 45-53.
6. P. Deligne, *La conjecture de Weil. I*, Publications Mathématiques de l'IHÉS, 43 (1974): 273–307,
7. P. Diaconis, *The distribution of leading digits and uniform distribution mod 1*, Ann. Probability 5 (1977), no. 1, 72–81.
8. P.D.T.A. Elliott, *Probabilistic number theory I. Mean-value theorems*, Grundlehren der Mathematischen Wissenschaften, vol. 239, Springer-Verlag, New York-Berlin, 1979.
9. A. Granville and K. Soundararajan, *Large character sums: Pretentious characters and the Pólya-Vinogradov theorem*, J. Amer. Math. Soc. 20 (2007), 357–384.
10. M. Jameson, J. Thorner, and L. Ye, *Benford's Law for coefficients of newforms*, Int. J. Number Theory 12 (2016), no. 2, 483–494.
11. A.V. Kontorovich and S.J. Miller, *Benford's law, values of L-functions and the* $3x + 1$ *problem*, Acta Arith. 120 (2005), no. 3, 269–297.
12. J. Lagarias and K. Soundararajan, *Benford's law for the 3x+1 function.* J. London Math. Soc. (2) 74 (2006), no. 2, 289–303.
13. H. Montgomery and R. Vaughan, *Multiplicative Number Theory I. Classical Theory*, Cambridge University Press, 2006.
14. V.K. Murty, *Modular forms and the Chebotarev density theorem. II.* In Analytic number theory (Kyoto, 1996), volume 247 of London Math. Soc. Lecture Note Ser., pages 287-308. Cambridge Univ. Press, Cambridge, 1997.
15. S. Newcomb, *Note on the Frequency of Use of the Different Digits in Natural Numbers*, Amer. J. Math. 4 (1881), no. 1–4, 39–40.
16. J.-P. Serre, *Quelques applications du theoreme de densite de Chebotarev*, Inst. Hautes Etudes Sci. Publ. Math. 54 (1981), 323–401.

17. J. Thorner and A. Zaman, *A Chebotarev variant of the Brun-Titchmarsh theorem and bounds for the Lang-Trotter conjectures.* Int. Math. Res. Not. IMRN 2018, no. 16, 4991–5027.
18. E.C. Titchmarsh, *The theory of the Riemann zeta-function.* Second edition. The Clarendon Press, Oxford University Press, New York, 1986.
19. G. Travaglini, *Number theory, Fourier analysis and geometric discrepancy.* London Mathematical Society Student Texts, 81. Cambridge University Press, Cambridge, 2014.
20. D.Q. Wan, *On the Lang-Trotter conjecture.* J. Number Theory, 35(3):247–268, 1990.
21. S. Zheng, *Necessary and Sufficient Conditions for Benford Sequences*, Pi Mu Epsilon Journal (2013), Vol. 13, No. 9, 1914 CENTENNIAL ISSUE 1 2014, 553–561.

Mathematics Department, Kansas State University, Manhattan, KS 66503
Email address: chandee@ksu.edu

Mathematics Department, Kansas State University, Manhattan, KS 66503
Email address: xiannan@math.ksu.edu

Mathematics Department, University of Georgia, Athens, GA 30602
Email address: pollack@uga.edu

Mathematics Department, University of Georgia, Athens, GA 30602
Email address: akash01s.roy@gmail.com