Analysis and Applications
© World Scientific Publishing Company

The primes perform a Benford dance

Arno Berger* and Ardalan Rahmatidehkordi
Mathematical and Statistical Sciences, University of Alberta
Edmonton, Alberta, Canada.

Received (19 December 2023)
Revised (25 October 2024)
Accepted (26 October 2024)

This article develops a new description of the asymptotics for the empirical distributions of significands and significant digits associated with (p_n), where p_n denotes the n-th prime number. The work utilizes the space of probability measures on the significand, endowed with a suitable Kantorovich metric, as well as finite-dimensional projections thereof. For sequences sufficiently close to (p_n), it is shown that the limit points of the associated empirical distributions form a circle that is made up of all rescalings of a single absolutely continuous distribution, and is centered at a distribution known as Benford's law (BL). The precise rate of convergence to that circle is determined. Moreover, even in the infinite-dimensional setting of significands the convergence is seen to occur along a distinguished low-dimensional object, in fact, along a smooth curve intimately related to BL. By connecting (p_n) and BL in a new way, the results rigorously confirm well-documented experimental observations and complement known facts in the literature.

Keywords: First significant digit; significand; empirical distribution; Kantorovich metric; Benford's law; prime number theorem.

Mathematics Subject Classification 2020: 11A63, 28A33, 60B10; 11A41, 11B39, 28C10.

1. Introduction

Given an integer $N \geq 1$ and a sequence (x_n) in $\mathbb{R}^+ := \{r \in \mathbb{R} : r > 0\}$, denote by $q_N = q_N[x_n] \in \mathbb{R}^9$ the probability (row) vector whose i-th component is the relative frequency of $x_1, x_2, \ldots, x_N$ having first significant (decimal) digit i, that is,

$$q_{N,i} = \frac{\#\{1 \leq n \leq N : \text{first significant digit of } x_n = i\}}{N} \quad \forall N \in \mathbb{N}, i \in \{1, 2, \ldots, 9\}\,.$$

Understanding, as accurately as possible, the asymptotics of q_N as $N \to \infty$ is a main objective of *digit analysis*; see, e.g., [4,14,16,22,25,26,28]. This article provides such an accurate understanding for a class of sequences that includes $(p_n) = (2, 3, 5, 7, 11, \ldots)$, the sequence of primes, as a most prominent example.

*Corresponding author (berger@ualberta.ca).

To preview and contextualize the main results, it is helpful to first consider q_N for another prominent integer sequence, the Fibonaccis $(F_n) = (1, 1, 2, 3, 5, \ldots)$, for which it is well-known [2,4,14,33] that

$$\lim_{N\to\infty} q_{N,i} = \log\frac{i+1}{i} =: b_{0,i} \qquad \forall i \in \{1, 2, \ldots, 9\}\,; \tag{1.1}$$

here and throughout, $\log t$ and $\ln t$ denote the base-10 and the natural logarithm of $t \in \mathbb{R}^+$, respectively. While the precise rate of convergence in (1.1) is unknown, q_N turns out to be extraordinarily close to b_0 even for not-so-large N; see [4,10].

Quite likely, the reader will recognize the probability vector $b_0 \in \mathbb{R}^9$ in (1.1) as a version of *Benford's law* (BL). In fact, the term BL is often used synonymously with b_0. As in [4,19,24] and many other references, however, BL is understood herein as describing a more general phenomenon of which b_0 is but one facet. For notational clarity, therefore, b_0 is henceforth referred to as *Benford's first digit law*. For the purpose of this article, it suffices to note that (1.1) has been shown to hold for many different classes of sequences, and that, more broadly, BL indeed is a multifaceted phenomenon that has long been of interest to analysts; see, e.g., [4,5,8,11,20,24,35] for background, applications, and further references.

For (p_n), the example of paramount interest in this article, it has been known at least since [34] that the sequence (q_N) diverges, so in particular (1.1) fails. Instead, $(q_{N,1})$, for instance, appears to be oscillating between $\frac{1}{9}$ and $\frac{5}{9}$, in a way that is asymptotically periodic on a logarithmic scale; see Fig. 1 and [4,16]. The fact that for the primes (q_N) not only diverges but also seems to lack any natural connection to BL has motivated much work [16,29] aimed at connecting (p_n) and b_0. To mention but one result in this regard, notice that

$$\lim_{N\to\infty} \frac{1}{\ln N} \sum_{n=1}^{N} \frac{q_n}{n} = b_0\,. \tag{1.2}$$

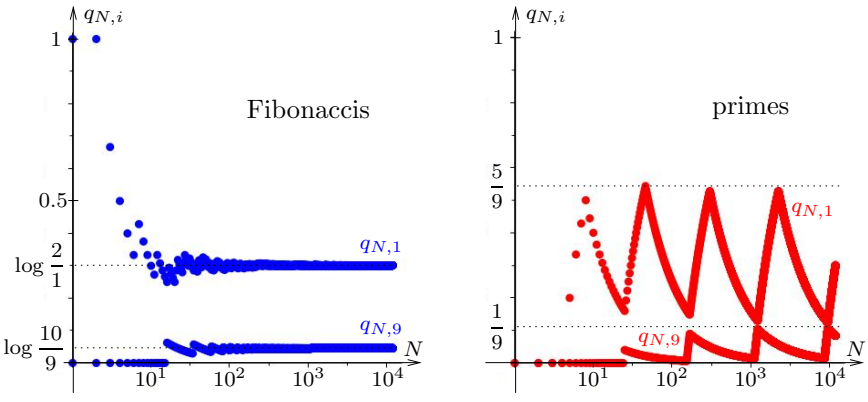


Fig. 1. For the Fibonaccis, each component of (q_N) converges to the corresponding component of Benford's first digit law b_0, that is, $\lim_{N\to\infty} q_{N,i} = \log\big((i+1)/i\big)$ for every $i \in \{1, 2, \ldots, 9\}$. By contrast, for the primes, each component $(q_{N,i})$ oscillates between $1/(9i)$ and $10/(9i+9)$, in a way that appears to be asymptotically periodic on a logarithmic scale.

Thus, while (q_N) does not conform to b_0 in the sense of (1.1), it does conform in the weaker sense of (1.2); see, e.g., [23,28] for many further results in a similar spirit.

This article pursues a novel, altogether different approach towards connecting (p_n) and BL. Rather than utilizing sophisticated but potentially artificial tools to make (q_N) conform to b_0, here q_N, as well as a suitable infinite-dimensional analogue, is studied directly. The simple, natural question, then, is this: Since, in the case of the primes, the sequence (q_N) diverges, what else can be said about its behaviour as $N \to \infty$? The following first main result provides a precise quantitative answer; to state it, let $\|v\|_\infty = \max_{i=1}^9 |v_i|$ for every $v \in \mathbb{R}^9$, and denote $\frac{1}{9}[1,1,\dots,1] \in \mathbb{R}^9$, the discrete uniform distribution on $\{1,2,\dots,9\}$, simply by u.

Theorem 1.1. *There exists a unique continuous, non-constant and* 1*-periodic function* $g : \mathbb{R} \to \mathbb{R}^9$ *with* $g(0) = u$ *so that, with some* $c \in \mathbb{R}^+$*,*

$$\frac{1}{c} < \log N \left\| q_N[p_n] - g(\log p_N) \right\|_\infty < c \qquad \forall N \geq 2\,.$$

Theorem 1.1 is proved in Section 3. To illustrate the result geometrically, let $\Omega[q_N]$ be the set of all limit points of (q_N). Note that $\Omega[q_N] \subset \mathbb{R}^9$ is a non-empty, compact and connected set of probability vectors. By Theorem 1.1, clearly $\Omega[q_N] = g(\mathbb{R})$, the latter set being a closed polygon with precisely 9 vertices, one of which is $u = g(0)$. While $b_0 \notin \Omega[q_N]$, and hence not even a subsequence of (q_N) conforms to b_0 in the sense of (1.1), it turns out that b_0 is the barycenter of $\Omega[q_N]$; see Proposition 2.3 below. In a natural, distributional way, therefore, the primes "dance" around BL. By contrast, for the Fibonaccis simply $\Omega[q_N] = \{b_0\}$, by (1.1); see Fig. 2.

The analysis in this article is motivated in part by experimental observations recorded in [22]. Theorem 1.1 rigorously confirms two of these observations: On the one hand, for the primes and with $N_j = 10^j$ the sequence (q_{N_j}) seems to diverge. As seen in Section 3 below, $\log p_{N_j} + \mathbb{Z}$ is dense in $\mathbb{R}/\mathbb{Z}$, and hence Theorem 1.1 shows that, rather than converge, (q_{N_j}) approaches the closed 9-gon $g(\mathbb{R})$, at the rate $(1/j)$. On the other hand, with $N_j = \pi(10^j)$, i.e., with N_j equal to the total number of primes $\leq 10^j$, (q_{N_j}) seems to converge to u. With all details deferred to Section 3, $\lim_{j\to\infty} \operatorname{dist}(\log p_{N_j}, \mathbb{Z}) = 0$ in this case, and it follows easily from Theorem 1.1 that indeed $\lim_{j\to\infty} q_{N_j} = u$, also at the rate $(1/j)$.

To preview the second main result of this article, let $b_t \in \mathbb{R}^9$ be given by

$$b_{t,i} = \frac{(i+1)^t - i^t}{10^t - 1} \qquad \forall t \in \mathbb{R}^+, i \in \{1,2,\dots,9\}\,. \tag{1.3}$$

Note that $b_1 = u$, whereas $\lim_{t\to 0} b_t = b_0$ and $\lim_{t\to\infty} b_t = [0,\dots,0,1] =: e_9$. Thus the family $(b_t)_{t\in\mathbb{R}^+}$, sometimes referred to as *generalized Benford's* (*first digit*) *law* [22,27], provides a smooth interpolation between b_0, u, and e_9. Now, for the primes and with $N_j = \pi(10^j)$, not only does (q_{N_j}) converge to $b_1 = u$, as seen above, but [22] furthermore observed experimentally that q_{N_j} is strikingly close to $b_{t_j^*}$ for some $t_j^* < 1$. The following result makes this observation rigorous; note that $\liminf_{j\to\infty} \log N_j \, \|q_{N_j}[p_n] - u\|_\infty > 0$, by Theorem 1.1.

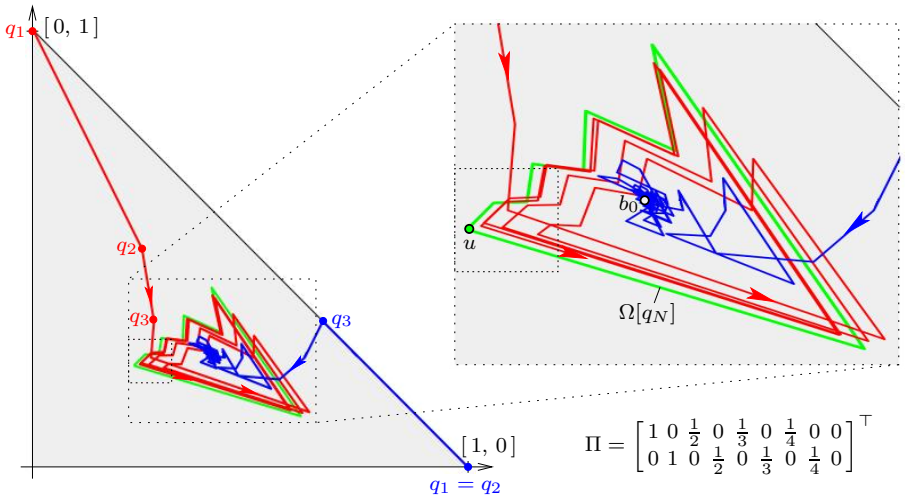


Fig. 2. For the Fibonaccis (blue data), $\lim_{N\to\infty} q_N = b_0$, whereas for the primes (red data), the sequence (q_N) approaches the green closed 9-gon $\Omega[q_N] = g(\mathbb{R})$; see also Fig. 4 and 6 below. Note that this figure only shows a two-dimensional (linear) projection of $b_0, u, q_N \in \mathbb{R}^9$, with the projection matrix Π suitably chosen for graphical depiction. Successive points $q_1, q_2, \ldots$, or rather their projections, are joined by straight line segments to underscore the eventually periodic nature (on a logarithmic scale) of (q_N) for the primes.

Theorem 1.2. *Let (N_j) be an increasing sequence in $\mathbb{N}$, and $t_j^* = 1 - 1/\ln N_j$ for every $j \geq 2$. If $\lim_{j\to\infty} \log N_j \operatorname{dist}(\log p_{N_j}, \mathbb{Z}) = 0$ then*

$$\lim_{j\to\infty} \log N_j \left(q_{N_j}[p_n] - b_{t_j^*}\right) = 0\,.$$

Theorem 1.2 is proved in Section 3 also. Informally put, the result asserts that the convergence $q_{N_j} \to u$ in $\mathbb{R}^9$ takes place asymptotically along the smooth curve $(b_t)_{t\in\mathbb{R}^+}$; see also Fig. 6 below. To illustrate the remarkable closeness of q_{N_j} to $b_{t_j^*}$, take, e.g., $j = 6$ for which $N_6 = 78498$, $t_6^* = 0.9112$,

$$q_{N_6} = \tfrac{1}{9}\big[\,1.098,\, 1.048,\, 1.027,\, 1.002,\, 0.9877,\, 0.9697,\, 0.9670,\, 0.9545,\, 0.9435\,\big]\,,$$
$$b_{t_6^*} = \tfrac{1}{9}\big[\,1.108,\, 1.057,\, 1.026,\, 1.003,\, 0.9858,\, 0.9713,\, 0.9590,\, 0.9484,\, 0.9391\,\big]\,,$$

and correspondingly $\|q_{N_6} - b_{t_6^*}\|_\infty = 1.078 \cdot 10^{-3}$ indeed is much smaller than $\|q_{N_6} - u\|_\infty = 1.099 \cdot 10^{-2}$; here and throughout, all real numbers are displayed to four correct significant digits.

While Theorems 1.1 and 1.2 reveal intriguing aspects of the statistics of (first significant digits of) primes, this article also makes it clear that, contrary to what [22] seems to suggest, neither aspect is in any way characteristic of the primes. As detailled below, both can be observed for a wider class of sequences, and are in fact direct consequences of a simple infinite-dimensional phenomenon.

The remainder of this article is organized as follows. Section 2 reviews the significand function S and space $\mathbb{S}$, as well as the topology and algebra associated with them. The section also reviews the space $\mathcal{P}$ of all probability distributions on $\mathbb{S}$,

and it introduces a finite-dimensional analogue of $\mathcal{P}$, via first significant digits. Section 3 presents the main results about the asymptotics of the empirical significand distributions for sequences sufficiently close to (p_n^a), with $a \in \mathbb{R}^+$; in particular, proofs of Theorem 1.1 and 1.2 are presented, in a somewhat wider context. For the reader's convenience, the results are summarized and put in perspective in Section 4, while the proof of one technical lemma is deferred to Appendix A.

Remark 1.3. By far the most familiar version of Benford's (first digit) law is the one for *decimal* (i.e., base-10) digits. For notational simplicity, only that version is considered throughout this article. Correspondingly, in the law's first digit form, b_0 is the *nine*-dimensional probability vector appearing in (1.1). However, all results carry over *mutatis mutandis* to every other integer base ≥ 2; cf. [4,28].

2. Significands, first significant digits, and their distributions

The main purpose of this section is to assemble the analytical tools to be employed in Section 3 below. As presented here, the key tool is a space of probability measures endowed with a suitable Kantorovich metric. In principle, Theorem 1.1 and 1.2 could be proved directly in $\mathbb{R}^9$, without resorting to the infinite-dimensional approach developed here. This novel approach is documented only indirectly in the literature [29,31]. However, it allows for the underlying mechanism to be presented much more transparently, and may be well-suited for other studies on BL also.

(i) The circular slide rule $(\mathbb{S}, d_{\mathbb{S}})$

Given $t \in \mathbb{R}^+$, let $1 \leq S(t) < 10$ be the unique real number such that $t = S(t) \cdot 10^k$ for some (necessarily also unique) $k \in \mathbb{Z}$. The function $S : \mathbb{R}^+ \to [1, 10[$ thus defined is the decimal *significand* (*function*). Clearly, $S(10t) = S(t)$ for all $t \in \mathbb{R}^+$. The number $S(t)$ incorporates all (decimal) digits of t. In particular, the first significant digit of t simply is $D_1(t) = \lfloor S(t) \rfloor$ where, as usual, $\lfloor a \rfloor$ denotes the largest integer $\leq a$, for any $a \in \mathbb{R}$. The range of S, i.e., the half-open interval $[1, 10[$, with its order inherited from $\mathbb{R}$, is henceforth denoted by $\mathbb{S}$. Letting

$$d_{\mathbb{S}}(r, s) = \min_{k \in \mathbb{Z}} \left| \log \frac{r}{s} + k \right| \qquad \forall r, s \in \mathbb{S}\,,$$

it is readily confirmed that $d_{\mathbb{S}}$ is a metric on $\mathbb{S}$, and endowed with it $\mathbb{S}$ is compact. In geometric terms, $d_{\mathbb{S}}(r, s)$ is nothing other than the minimal (normalized) arclength between r and s on a circular slide rule. Clearly $d_{\mathbb{S}} \leq \frac{1}{2}$, and $d_{\mathbb{S}}(r, s) = \frac{1}{2}$ precisely if $r/s \in \{\sqrt{10}, 1/\sqrt{10}\}$, that is, if r, s are antipodal. Furthermore, letting

$$r * s = S(rs) \qquad \forall r, s \in \mathbb{S}\,,$$

it is readily seen that $(\mathbb{S}, *)$ is a commutative group, with neutral element $s = 1$, and the inverse of $s \neq 1$ is $10/s$. For every $t \in \mathbb{S}$ the rotation by t is an isometry,

$$d_{\mathbb{S}}(t * r, t * s) = d_{\mathbb{S}}(r, s) \qquad \forall r, s, t \in \mathbb{S}\,.$$

Later it will be convenient to consider, for every $t \in \mathbb{R}^+$, the rotation $T_t : \mathbb{S} \to \mathbb{S}$ given by $T_t(\cdot) = S(t) * \cdot = S(t\,\cdot)$; with this, $d_{\mathbb{S}}\big(T_t(r), T_t(s)\big) = d_{\mathbb{S}}(r,s)$ for all $r, s \in \mathbb{S}$ and $t \in \mathbb{R}^+$.

(ii) The space $(\mathcal{P}, d_{\mathbb{S}})$ of all laws on $\mathbb{S}$

The main protagonists of the analysis in this article are the Borel probability measures (or *laws* [13, Sec. 9.1]) on $\mathbb{S}$. Denote by $\mathcal{P}$ the set of all such measures, with generic elements henceforth denoted P, Q, etc. Every $P \in \mathcal{P}$ is uniquely determined by its distribution function $F_P : \mathbb{S} \to [0,1]$, given by

$$F_P(s) = P([1,s]) \qquad \forall s \in \mathbb{S}\,.$$

Notice that F_P is non-decreasing, right-continuous with left limit $F_P(s-) = \lim_{r\uparrow s} F_P(r) \le F_P(s)$, and $F_P(1) \ge 0$ as well as $F_P(10-) = 1$; every function with these properties is the distribution function of one and only one $P \in \mathcal{P}$.

Prominent elements of $\mathcal{P}$ in this article include the Dirac distribution (or unit point mass) δ_s concentrated at $s \in \mathbb{S}$, with $F_{\delta_s} = \mathbf{1}_{[s,10[}$, the indicator function of $[s,10[\subset \mathbb{S}$; the uniform distribution U, with $F_U(s) = \frac{1}{9}(s-1)$; and Benford's law B, with $F_B(s) = \log s$. An important property of B is its invariance under rotations,

$$B \circ T_t^{-1}([1,s]) = \log s = B([1,s]) \qquad \forall t \in \mathbb{R}^+, s \in \mathbb{S}\,.$$

This property of B is traditionally referred to as *scale-invariance* [4,19]; in harmonic analysis parlance [12], it simply asserts that B is the (unique, normalized) Haar measure of the compact commutative group $(\mathbb{S}, *)$.

Recall that $\mathcal{P}$ is a compact metrizable space where $\lim_{N\to\infty} P_N = P$ means that

$$\lim_{N\to\infty} \int_{\mathbb{S}} f \,\mathrm{d}P_N = \int_{\mathbb{S}} f \,\mathrm{d}P \qquad \forall f \in C(\mathbb{S})\,;$$

here, as usual, $C(\mathbb{S})$ denotes the Banach space of all continuous functions $f : \mathbb{S} \to \mathbb{R}$. A natural metric on $\mathcal{P}$ inducing this so-called "weak" topology is defined as follows: Denoting by $\mathrm{Lip}_1(\mathbb{S})$ the set of all 1-Lipschitz functions $\ell : \mathbb{S} \to \mathbb{R}$, let

$$d_{\mathbb{S}}(P,Q) = \sup_{\ell \in \mathrm{Lip}_1(\mathbb{S})} \int_{\mathbb{S}} \ell \,\mathrm{d}(P-Q) \qquad \forall P, Q \in \mathcal{P}\,. \tag{2.1}$$

With this, $(\mathcal{P}, d_{\mathbb{S}})$ is a compact metric space, and $\lim_{N\to\infty} P_N = P$ precisely if $\lim_{N\to\infty} d_{\mathbb{S}}(P_N, P) = 0$. A few comments regarding (2.1) are in order. Firstly, $d_{\mathbb{S}}(\delta_r, \delta_s) = d_{\mathbb{S}}(r,s)$ for all $r, s \in \mathbb{S}$, so $\mathbb{S}$ is isometrically embedded in $\mathcal{P}$; this motivates using the same symbol for the metrics on both $\mathbb{S}$ and $\mathcal{P}$. Also, the definition (2.1) results from a general principle applicable to every compact metric space [13, Ch. 11], with $d_{\mathbb{S}}$ referred to as the *Kantorovich* (or 1-*Wasserstein*) *distance* [6,7,8,30]. Secondly, the supremum in (2.1) is attained, due to the compactness of $\big\{\ell \in \mathrm{Lip}_1(\mathbb{S}) : \ell(1) = 0\big\}$ in $C(\mathbb{S})$. Lastly, though it may seem taunting to compute $d_{\mathbb{S}}(P,Q)$ for specific P, Q using (2.1), it turns out that

$$d_{\mathbb{S}}(P,Q) = \log e \, \min_{a\in\mathbb{R}} \int_1^{10} \big|F_P(s) - F_Q(s) - a\big| \,\frac{\mathrm{d}s}{s} \qquad \forall P, Q \in \mathcal{P}\,, \tag{2.2}$$

which is rather well-suited for computations [6,9]. For example, for every $s \in \mathbb{S}$,

$$d_{\mathbb{S}}(U, \delta_s) = \begin{cases} \log s + \left(1 - \dfrac{2s}{\sqrt{10}+1}\right) \log e & \text{if } 1 \le s < \sqrt{10}\,, \\ 1 - \log s + \left(\dfrac{2s}{10+\sqrt{10}} - 1\right) \log e & \text{if } \sqrt{10} \le s < 10\,, \end{cases}$$

so $d_{\mathbb{S}}(U, \delta_s) + d_{\mathbb{S}}(U, \delta_{s*\sqrt{10}}) = \frac{1}{2} = d_{\mathbb{S}}\big(s, s*\sqrt{10}\big)$, and hence $d_{\mathbb{S}}(U, \delta_s)$ is minmal and maximal precisely for $s = \frac{1}{2}\big(10+\sqrt{10}\big)$ and $s = \frac{1}{2}(\sqrt{10}+1)$, respectively.

Utilizing (2.2), it is easy to see that $d_{\mathbb{S}}(P, B) \le \frac{1}{4}$ for every $P \in \mathcal{P}$, with equality holding precisely if $P = \delta_s$ for some $s \in \mathbb{S}$. As a consequence, $d_{\mathbb{S}}(P, Q) \le \frac{1}{2}$ for all $P, Q \in \mathcal{P}$, with equality holding precisely if $P = \delta_r$, $Q = \delta_s$ for some antipodal $r, s \in \mathbb{S}$. Informally, therefore, one may visualize $\mathcal{P}$ as an (infinite-dimensional) "ball" with "center" B and "radius" $\frac{1}{4}$; see Fig. 3.

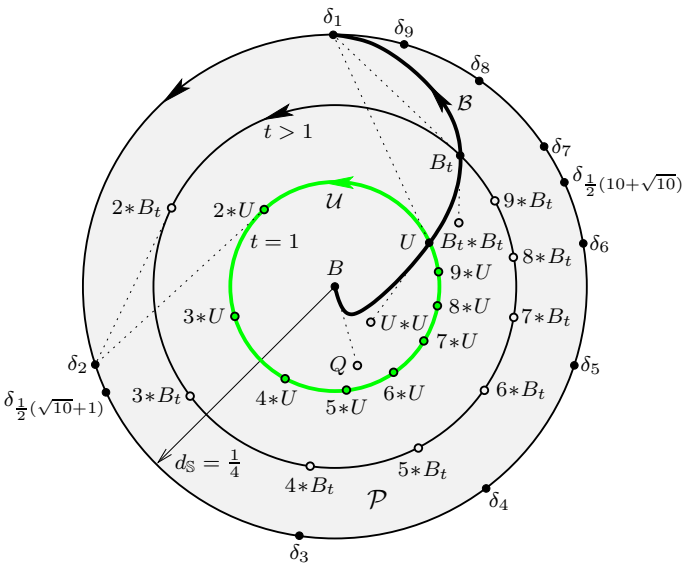


Fig. 3. Visualizing the compact metric space $(\mathcal{P}, d_{\mathbb{S}})$ as a "ball" with "center" B and "radius" $\frac{1}{4}$. Note that the sets $\{s * B_t : s \in \mathbb{S}\}$ with $0 < t \le \infty$ all are homeomorphic to the circle $(\mathbb{S}, d_{\mathbb{S}})$. The family $(B_t)_{t\in\mathbb{R}^+}$ represents a smooth curve $\mathcal{B}$ in $\mathcal{P}$, connecting $B = B_0$, $U = B_1$, and $\delta_1 = B_\infty$.

Finally, define a binary operation $*$ ("convolution") on $\mathcal{P}$, via

$$P * Q(A) = \int_{\mathbb{S}\times\mathbb{S}} \mathbf{1}_A(r * s)\, \mathrm{d}P(r)\, \mathrm{d}Q(s) \qquad \forall P, Q \in \mathcal{P}\,,$$

where $A \subset \mathbb{S}$ is any Borel set. It is readily seen that $*$ is jointly continuous, associative, commutative, and $P * \delta_1 = P$ for every $P \in \mathcal{P}$, but also $\delta_r * \delta_s = \delta_{r*s}$ for all $r, s \in \mathbb{S}$. In probabilistic parlance, if ξ and η are two *independent* positive random variables, and P and Q are the distributions on $\mathbb{S}$ of $S(\xi)$ and $S(\eta)$, respectively,

then $P * Q$ is the distribution of $S(\xi\eta)$. Later it will be convenient to consider the distribution $t * P$, obtained from $P \in \mathcal{P}$ via scaling (or rotation) by any $t \in \mathbb{R}^+$,

$$t * P = \delta_{S(t)} * P = P \circ T_t^{-1} \qquad \forall t \in \mathbb{R}^+, P \in \mathcal{P}\,.$$

For example, $t * \delta_s = \delta_{S(t)} * \delta_s = \delta_{S(ts)}$ for all $t \in \mathbb{R}^+$, $s \in \mathbb{S}$.

Remark 2.1. While $d_{\mathbb{S}}$ is the only metric on $\mathcal{P}$ employed consistently throughout this article, many other metrics could be chosen instead; see, e.g., [17]. One natural choice would be $d_\infty(P,Q) = \|F_P - F_Q\|_\infty$ with $\|f\|_\infty := \sup_{s\in\mathbb{S}} |f(s)|$. An analogue of *star-discrepancy* [1,21], the metric d_∞ has often been used in studies on BL; see, e.g., [3,25,26]. By (2.2), $d_{\mathbb{S}} \le d_\infty \le 1$, and d_∞ induces a finer topology on $\mathcal{P}$ that is not separable, let alone compact. Also, d_∞ is not invariant under scaling or inversion. Not least for these reasons d_∞ may be less convenient to work with than $d_{\mathbb{S}}$. Notice, however, that if $P(\{s\}) = 0$ for every $s \in \mathbb{S}$ then $d_\infty(P_N, P) \to 0$ if and only if $d_{\mathbb{S}}(P_N, P) \to 0$, though the *rates* of convergence may differ greatly [1].

(iii) The generalized BL family (B_t)

For the purpose of this article, one family in $\mathcal{P}$ of particular interest is $(B_t)_{t\in\mathbb{R}^+}$, with distribution functions

$$F_{B_t}(s) = \frac{s^t - 1}{10^t - 1} \qquad \forall t \in \mathbb{R}^+, s \in \mathbb{S}\,.$$

The family $(B_t)_{t\in\mathbb{R}^+}$, or rather its first digit version $\lfloor B_t \rfloor$ to be discussed in Subsection 2(iv) below, is sometimes referred to as *generalized Benford's law* [22,27]. Notice that $t \mapsto B_t$ is continuous and one-to-one, with $B_1 = U$, and (2.2) yields

$$d_{\mathbb{S}}(B_t, B) = \frac{1}{t}\log\frac{10^{t/2}+1}{2} - \frac{1}{4}\,, \quad d_{\mathbb{S}}(B_t, \delta_1) = \frac{\log e}{t}\cdot\frac{10^{t/2}-1}{10^{t/2}+1} \qquad \forall t \in \mathbb{R}^+,$$

which shows that $\lim_{t\to 0} B_t = B$ and $\lim_{t\to\infty} B_t = \delta_1$. With $B_0 := B$ and $B_\infty := \delta_1$, therefore, the set $\mathcal{B} := \{B_t : 0 \le t \le \infty\}$ is homeomorphic to the unit interval, and $(B_t)_{t\in\mathbb{R}^+}$ provides a continuous interpolation between B, U, and δ_1; see Fig. 3.

To appreciate the special role that B plays in $\mathcal{P}$, note that $t * B = B$ for all $t \in \mathbb{R}^+$, and $P * B = B$ for every $P \in \mathcal{P}$. Given $P \ne B$, the set $\{s * P : s \in \mathbb{S}\}$ is homeomorphic to $(\mathbb{S}, d_{\mathbb{S}})$, and hence may be interpreted as a "circle" with "centre" B and "radius" $d_{\mathbb{S}}(P,B) \le \frac{1}{4}$. In support of this interpretation, observe that with every closed path γ in $\mathcal{P}$, i.e., with every continuous function $\gamma : \mathbb{S} \to \mathcal{P}$, one can associate its *barycenter* $\widehat{\gamma} \in \mathcal{P}$, defined uniquely by the property that

$$\int_{\mathbb{S}} f \,\mathrm{d}\widehat{\gamma} = \log e \int_1^{10} \int_{\mathbb{S}} f \,\mathrm{d}\gamma(s)\,\frac{\mathrm{d}s}{s} \qquad \forall f \in C(\mathbb{S})\,.$$

It is readily seen that B is the barycenter of many closed paths in $\mathcal{P}$.

Proposition 2.2. *Given $P \in \mathcal{P}$, let $\gamma(s) = s * P$ for every $s \in \mathbb{S}$. Then $\gamma : \mathbb{S} \to \mathcal{P}$ is continuous, and $\widehat{\gamma} = B$.*

For instance, the set $\{s * B_t : s \in \mathbb{S}\}$, with $0 < t \leq \infty$, is a "circle" in $\mathcal{P}$ with barycenter B and "radius" $d_{\mathbb{S}}(B_t, B) > 0$.

The remainder of this subsection collects a few analytic properties of the family $(s * B_t)_{s\in\mathbb{S},t\in\mathbb{R}^+}$. Specifically, it is straightforward to see that for every $t \in \mathbb{R}^+$,

$$4d_{\mathbb{S}}(B_t, B)\, d_{\mathbb{S}}(r, s) \leq d_{\mathbb{S}}(r * B_t, s * B_t) \leq \frac{t d_{\mathbb{S}}(B_t, \delta_1)}{\log e}\, d_{\mathbb{S}}(r, s) \qquad \forall r, s \in \mathbb{S}\,, \quad (2.3)$$

where the left and the right inequality is sharp for $d_{\mathbb{S}}(r, s) = \frac{1}{2}$ and $d_{\mathbb{S}}(r, s) \to 0$, respectively. Thus, $s \mapsto s * B_t$ is Lipschitz, and in fact is differentiable, with

$$\frac{\mathrm{d}}{\mathrm{d}s} s * B_t = \frac{t}{s}(\delta_s - s * B_t) \qquad \forall s \in \mathbb{S}, t \in \mathbb{R}^+, \quad (2.4)$$

in the sense that

$$\lim_{r\to s} \int_{\mathbb{S}} f \,\mathrm{d}\left(\frac{r * B_t - s * B_t}{r - s}\right) = \frac{t}{s} \int_{\mathbb{S}} f \,\mathrm{d}(\delta_s - s * B_t) \qquad \forall f \in C(\mathbb{S}), s \in \mathbb{S}\,. \quad (2.5)$$

Moreover, locally uniformly on $\mathbb{S} \setminus \{s\}$,

$$\lim_{r\to s} \frac{F_{r*B_t} - F_{s*B_t}}{r - s} = \frac{t}{s}(F_{\delta_s} - F_{s*B_t}) \qquad \forall s \in \mathbb{S}, t \in \mathbb{R}^+, \quad (2.6)$$

which in turn shows that the convergence in (2.5) is uniform in $f \in \mathrm{Lip}_1(\mathbb{S})$. Thus, for instance, the (signed) measure $\delta_1 - B_t$ is a "tangent vector" to the "circle" $\{s * B_t : s \in \mathbb{S}\}$ at B_t. In a similar vein, $t \mapsto B_t$ is differentiable, with

$$\frac{\mathrm{d}}{\mathrm{d}t} B_t = \begin{cases} \dfrac{B_t - B_t * B_t}{t} & \text{if } t > 0\,, \\ \dfrac{Q - B}{2 \log e} & \text{if } t = 0\,, \end{cases} \quad (2.7)$$

where $F_Q(s) = (\log s)^2$ for all $s \in \mathbb{S}$. In analogy to (2.6), uniformly on $\mathbb{S}$,

$$\lim_{s\to t} \frac{F_{B_s} - F_{B_t}}{s - t} = \frac{F_{B_t} - F_{B_t * B_t}}{t} \qquad \forall t \in \mathbb{R}^+\,. \quad (2.8)$$

Thus, for instance, $U - U * U$ is a "tangent vector" to $\mathcal{B}$ at $B_1 = U$; see Fig. 3.

(iv) First significant digit distributions

To set the scene for the proofs of Theorem 1.1 and 1.2 in the next section, note that the first significant digit function $D_1 = \lfloor S \rfloor : \mathbb{R}^+ \to \{1, 2, \ldots, 9\}$ is continuous at t if and only if $S(t) \notin \{1, 2, \ldots, 9\}$. Every $P \in \mathcal{P}$ induces a probability distribution on $\{1, 2, \ldots, 9\}$, namely $P \circ D_1^{-1}$, henceforth denoted $\lfloor P \rfloor$ for convenience; thus,

$$\lfloor P \rfloor_i = P([i, i + 1[) \qquad \forall P \in \mathcal{P}, i \in \{1, 2, \ldots, 9\}\,.$$

In probabilistic parlance, if ξ is a positive random variable and P is the distribution of $S(\xi)$ on $\mathbb{S}$ then $\lfloor P \rfloor$ is the distribution of $D_1(\xi)$ on $\{1, 2, \ldots, 9\}$. Notice that $\lfloor \mathcal{P} \rfloor := \{\lfloor P \rfloor : P \in \mathcal{P}\}$ equals the simplex of all probability vectors in $\mathbb{R}^9$, i.e.,

$$\lfloor \mathcal{P} \rfloor = \left\{ p \in \mathbb{R}^9 : p_i \geq 0 \ \ \forall i \in \{1, 2, \ldots, 9\}\,, \sum\nolimits_{i=1}^{9} p_i = 1 \right\}.$$

Endowing $\lfloor\mathcal{P}\rfloor$ with the usual topology inherited from $\mathbb{R}^9$, observe that $\lfloor\cdot\rfloor$ is continuous at P if and only if $P(\{1,2,\dots,9\})=0$.

It is instructive to observe how the distributions introduced earlier in this section are transported from $\mathcal{P}$ to $\lfloor\mathcal{P}\rfloor$. For instance, applying $\lfloor\cdot\rfloor$ to $s*U$ yields

$$\lfloor s*U\rfloor_i = \frac{1+9\min\{i+1,s\}-9\min\{i,s\}}{9s} \qquad \forall s\in\mathbb{S}, i\in\{1,2,\dots,9\}\,. \tag{2.9}$$

Note that while $s\mapsto s*U$ is differentiable on $\mathbb{S}$ by (2.4) with $t=1$, the continuous function $s\mapsto\lfloor s*U\rfloor$ is differentiable only if $s\notin\{1,2,\dots,9\}$. The lack of differentiability at $s=i\in\{1,2,\dots,9\}$ reflects the fact that $\frac{\mathrm{d}}{\mathrm{d}s}(s*U)\big|_{s=i}=(\delta_i-i*U)/i$, and hence the function $\lfloor\cdot\rfloor$, when extended to all finite (signed) Borel measures on $\mathbb{S}$, is discontinuous at $\frac{\mathrm{d}}{\mathrm{d}s}(s*U)\big|_{s=i}$. From (2.9), it is readily seen that applying $\lfloor\cdot\rfloor$ turns the (infinite-dimensional) "circle" $\mathcal{U}:=\{s*U : s\in\mathbb{S}\}$ into the closed 9-gon $\lfloor\mathcal{U}\rfloor=\{\lfloor s*U\rfloor : s\in\mathbb{S}\}$, with vertices

$$\lfloor i*U\rfloor = \frac{1}{9i}\big[\underbrace{10,\dots,10}_{i-1\text{ entries}},\ \underbrace{1,\dots,1}_{10-i\text{ entries}}\big]\in\mathbb{R}^9 \qquad \forall i\in\{1,2,\dots,9\}\,,$$

so in particular $\lfloor 1*U\rfloor=\lfloor U\rfloor=u$; see Fig. 4.

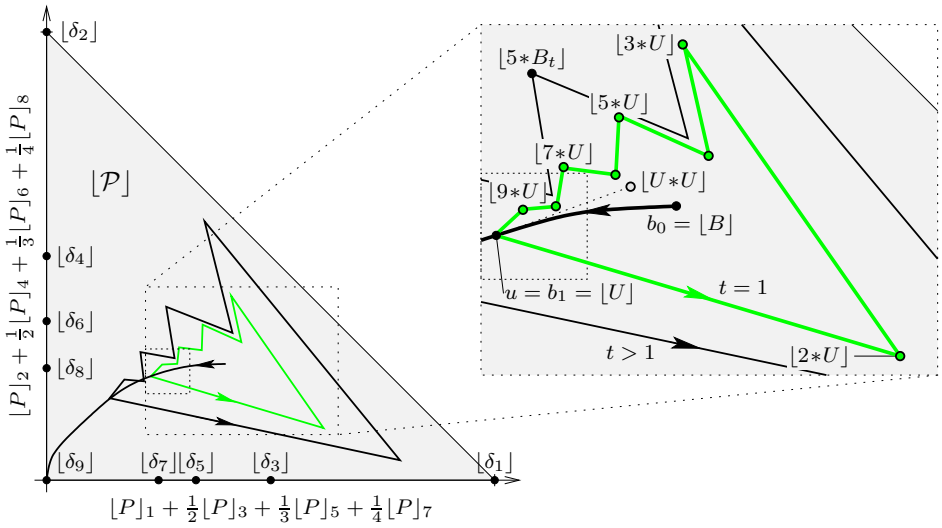


Fig. 4. Considering only the first significant (decimal) digit or, more formally, applying the function $\lfloor\cdot\rfloor$ converts the infinite-dimensional "ball" $\mathcal{P}$ into the nine-dimensional simplex $\lfloor\mathcal{P}\rfloor$. To facilitate comparison with Fig. 2 and 6, the same projection Π has been used in depicting $\lfloor\mathcal{P}\rfloor$.

In a similar vein, applying $\lfloor\cdot\rfloor$ turns the family $(B_t)_{t\in\mathbb{R}^+}$ into the smooth curve $(b_t)_{t\in\mathbb{R}^+}$, with b_t given by (1.3); as observed there already, $\lim_{t\to 0} b_t = b_0 = \lfloor B\rfloor$,

which also follows from the continuity of $\lfloor\cdot\rfloor$ at B. Furthermore, by (2.7),

$$\frac{\mathrm{d}}{\mathrm{d}t}b_t = \begin{cases} \dfrac{b_t - \lfloor B_t * B_t\rfloor}{t} & \text{if } t > 0\,, \\ \dfrac{\lfloor Q\rfloor - b_0}{2\log e} & \text{if } t = 0\,. \end{cases}$$

Thus, for instance $\frac{\mathrm{d}}{\mathrm{d}t}b_t\big|_{t=1} = u - \lfloor U * U\rfloor$. Recall that $\lim_{t\to\infty} B_t = \delta_1$, whereas

$$b_\infty := \lim_{t\to\infty} b_t = \lim_{t\to\infty}\lfloor B_t\rfloor = e_9 \neq [1, 0, \ldots, 0] = \lfloor\delta_1\rfloor\,,$$

reflecting the discontinuity of $\lfloor\cdot\rfloor$ at δ_1. The special role of B as per Proposition 2.2 is preserved by $\lfloor\cdot\rfloor$, in that $b_0 = \lfloor B\rfloor$ is the barycenter of many closed paths in $\lfloor\mathcal{P}\rfloor$.

Proposition 2.3. *Given $P \in \mathcal{P}$ with $P(\{s\}) = 0$ for every $s \in \mathbb{S}$, the function $s \mapsto \lfloor s * P\rfloor$ is continuous, and* $\log e \int_1^{10} \lfloor s * P\rfloor/s \;\mathrm{d}s = b_0$.

3. Empirical distributions of significands and first significant digits

Given any sequence (z_N) in any metric space Z, denote by $\Omega[z_N]$ the set of all limit points of (z_N), that is,

$$\Omega[z_N] = \bigcap\nolimits_{m\in\mathbb{N}} \overline{\{z_n : n \geq m\}} = \big\{z \in Z : \exists N_1 < N_2 < \ldots \text{ s.t. } \lim\nolimits_{j\to\infty} z_{N_j} = z\big\}\,.$$

Plainly, $\Omega[z_N]$ is closed, and indeed is non-empty and compact whenever Z is compact. Also, $\Omega[z_N] = \{z\}$ if and only if $\lim_{N\to\infty} z_N = z$. Whenever (z_N) and Z are clear from the context, write $\Omega[z_N]$ simply as Ω.

The tools assembled in Section 2 are now going to be employed to analyze the significands and first significant digits of arbitrary sequences (x_n) in $\mathbb{R}^+$. To this end, for every $N \in \mathbb{N}$ denote the empirical distribution on $\{S(x_1), S(x_2), \ldots, S(x_N)\}$ by $Q_N[x_n]$, that is,

$$Q_N[x_n] = \frac{1}{N}\sum\nolimits_{n=1}^{N} \delta_{S(x_n)} \in \mathcal{P} \qquad \forall N \in \mathbb{N}\,;$$

whenever (x_n) is clear from the context, write $Q_N[x_n]$ simply as Q_N. Note that (Q_N) is a sequence in $\mathcal{P}$, so $\Omega[Q_N] \neq \varnothing$. The set $\Omega[Q_N] \subset \mathcal{P}$, determined solely by (x_n), is denoted $\Omega\{x_n\}$, rather than $\Omega\big[Q_N[x_n]\big]$, to avoid confusion with $\Omega[x_n] \subset \mathbb{R}$. The basic properties of $\Omega\{x_n\}$ are well-known; see, e.g., [6,31,32].

Proposition 3.1. *Let (x_n) be a sequence in $\mathbb{R}^+$. Then $\Omega\{x_n\}$ is non-empty, compact, and connected. Given any non-empty, compact, and connected set $\mathcal{Q} \subset \mathcal{P}$, there exists an increasing sequence (y_n) in $\mathbb{R}^+$ such that $\Omega\{y_n\} = \mathcal{Q}$.*

By Proposition 3.1, the simplest situation occurs when $\Omega = \{P\}$ for some $P \in \mathcal{P}$, that is, when $\lim_{N\to\infty} Q_N = P$. The case $P = B$ is of particular interest.

Definition 3.2. A sequence (x_n) in $\mathbb{R}^+$ is a *Benford sequence*, or simply *Benford*, if $\Omega\{x_n\} = \{B\}$.

Many basic (deterministic or stochastic) processes have been studied for their potential to generate Benford sequences; see, e.g., [2,4,24] and the many references therein. A basic example of a Benford (integer) sequence encountered in the Introduction is (F_n). Note that if (x_n) is Benford then $\lim_{N\to\infty}\lfloor Q_N\rfloor = b_0$, by the continuity of $\lfloor\cdot\rfloor$ at B. Thus (x_n) conforms to Benford's first digit law in the sense of (1.1). The converse is not true: $(\lfloor Q_N\rfloor)$ may converge to b_0, and yet (Q_N) may diverge or else converge to some $P \neq B$. The latter is the case, e.g., for $(x_n) = \big(D_1(F_n)\big)$ where $\lim_{N\to\infty} Q_N = \sum_{i=1}^{9} b_i\delta_i \neq B$, and yet $\lim_{N\to\infty}\lfloor Q_N\rfloor = b_0$.

As alluded to already in the Introduction, the sequence (p_n) does not satisfy (1.1), and a fortiori is not Benford. Nonetheless, several natural connections between the primes and BL are going to be stablished here. To this end, consider slightly more generally the sequence (p_n^a) with any $a \in \mathbb{R}^+$. The detailled quantitative analysis of $Q_N[p_n^a]$ rests upon the following auxiliary result.

Lemma 3.3. *Let $(x_n) = (p_n^a)$ for some $a \in \mathbb{R}^+$. Then*

$$\lim_{N\to\infty} \big\| \ln N \left(F_{x_N^{-1}*Q_N} - F_{B_{1/a}}\right) - \left(F_{B_{1/a}*B_{1/a}} - F_{B_{1/a}}\right)\big\|_\infty = 0\,. \tag{3.1}$$

For the reader's convenience, the straightforward yet somewhat lengthy proof of Lemma 3.3 is deferred to Appendix A. The argument makes crucial use of an appropriate quantitative refinement of the prime number theorem, but does not rely on the primality of p_n in any other way. As detailled in Appendix A, therefore, (3.1) holds for all sequences (x_n) that are, in some sense, sufficiently close to (p_n^a).

Despite its somewhat technical appearance, Lemma 3.3 yields interesting basic properties of the sequence (Q_N) with very little effort.

Theorem 3.4. *Let $(x_n) = (p_n^a)$ for some $a \in \mathbb{R}^+$. Then*

(i) $\lim_{N\to\infty} \ln N\, d_{\mathbb{S}}(Q_N, x_N * B_{1/a}) = d_{\mathbb{S}}(B_{1/a} * B_{1/a}, B_{1/a})$*, and therefore in particular* $\lim_{N\to\infty} d_{\mathbb{S}}(Q_N, x_N * B_{1/a}) = 0$*;*

(ii) *given an increasing sequence (N_j) in $\mathbb{N}$ and $s \in \mathbb{S}$,* $\lim_{j\to\infty} d_{\mathbb{S}}(Q_{N_j}, s * B_{1/a}) = 0$ *if and only if* $\lim_{j\to\infty} d_{\mathbb{S}}(S(x_{N_j}), s) = 0$*;*

(iii) $\lim_{N\to\infty} d_{\mathbb{S}}(Q_N, B) = d_{\mathbb{S}}(B_{1/a}, B)$*;*

(iv) $\Omega[Q_N] = \{s * B_{1/a} : s \in \mathbb{S}\}$.

Proof. Observe first that for every $t \in \mathbb{R}^+$ and $P \in \mathcal{P}$,

$$F_{t*P}(s) = \begin{cases} F_P\left(\dfrac{10s}{S(t)}\right) - F_P\left(\dfrac{10}{S(t)}-\right) & \text{if } 1 \le s < S(t)\,, \\ F_P\left(\dfrac{s}{S(t)}\right) + 1 - F_P\left(\dfrac{10}{S(t)}-\right) & \text{if } S(t) \le s < 10\,. \end{cases} \tag{3.2}$$

Using this, deduce from Lemma 3.3 that

$$\lim_{N\to\infty} \big\| \ln N \left(F_{Q_N} - F_{x_N*B_{1/a}}\right) - \left(F_{x_N*B_{1/a}*B_{1/a}} - F_{x_N*B_{1/a}}\right)\big\|_\infty = 0\,, \tag{3.3}$$

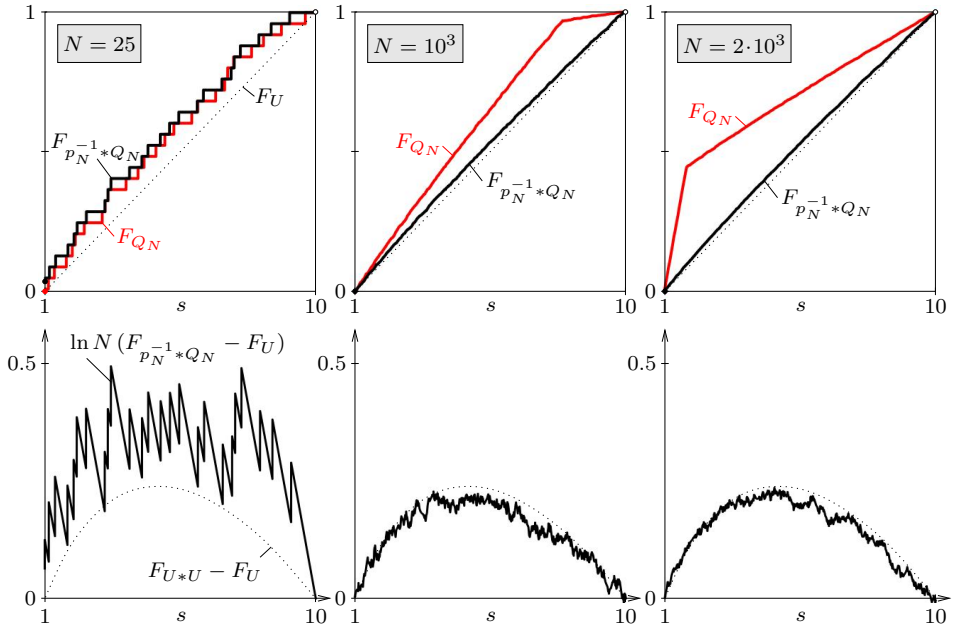


Fig. 5. Illustrating Lemma 3.3 for $a = 1$: Whereas for $(x_n) = (p_n)$ the sequence (Q_N) diverges (red data), its scaled (or rotated) version $(p_N^{-1} * Q_N)$ converges to U (black data), at the precise rate $(1/\ln N)$ and in the "direction" of $U - U * U$.

which in turn implies that, uniformly in $\ell \in \mathrm{Lip}_1(\mathbb{S})$,

$$\ln N \int_{\mathbb{S}} \ell \,\mathrm{d}(Q_N - x_N * B_{1/a}) - \int_{\mathbb{S}} \ell \,\mathrm{d}(x_N * B_{1/a} * B_{1/a} - x_N * B_{1/a}) \overset{N\to\infty}{\longrightarrow} 0\,. \quad (3.4)$$

By utilizing (3.4), all assertions are now going to be established quite easily.

(i) Pick $\ell_0, \ell_N \in \mathrm{Lip}_1(\mathbb{S})$ with $d_{\mathbb{S}}(B_{1/a} * B_{1/a}, B_{1/a}) = \int_{\mathbb{S}} \ell_0 \,\mathrm{d}(B_{1/a} * B_{1/a} - B_{1/a})$ and $d_{\mathbb{S}}(Q_N, x_N * B_{1/a}) = \int_{\mathbb{S}} \ell_N \,\mathrm{d}(Q_N - x_N * B_{1/a})$. Given $\varepsilon > 0$, there exists $N_1 \in \mathbb{N}$ so that

$$\left| \ln N \int_{\mathbb{S}} \ell \,\mathrm{d}(Q_N - x_N * B_{1/a}) - \int_{\mathbb{S}} \ell \,\mathrm{d}(x_N * B_{1/a} * B_{1/a} - x_N * B_{1/a}) \right| < \varepsilon$$

for all $N \geq N_1$ and $\ell \in \mathrm{Lip}_1(\mathbb{S})$. Since T_t is an isometry of $\mathbb{S}$ for every $t \in \mathbb{R}^+$,

$$\begin{aligned} \ln N\, d_{\mathbb{S}}(Q_N, x_N * B_{1/a}) &= \ln N \int_{\mathbb{S}} \ell_N \,\mathrm{d}(Q_N - x_N * B_{1/a}) \\ &< \varepsilon + \int_{\mathbb{S}} \ell_N \,\mathrm{d}(x_N * B_{1/a} * B_{1/a} - x_N * B_{1/a}) \leq \varepsilon + \int_{\mathbb{S}} \ell_0 \,\mathrm{d}(B_{1/a} * B_{1/a} - B_{1/a}) \\ &= \varepsilon + d_{\mathbb{S}}(B_{1/a} * B_{1/a}, B_{1/a}) \qquad \forall N \geq N_1\,, \end{aligned}$$

but also

$$\begin{aligned} d_{\mathbb{S}}(B_{1/a} * B_{1/a}, B_{1/a}) &= \int_{\mathbb{S}} \ell_0 \,\mathrm{d}(B_{1/a} * B_{1/a} - B_{1/a}) \\ &< \varepsilon + \ln N \int_{\mathbb{S}} \ell_0 \circ T_{x_N}^{-1} \,\mathrm{d}(Q_N - x_N * B_{1/a}) \le \varepsilon + \ln N \int_{\mathbb{S}} \ell_N \,\mathrm{d}(Q_N - x_N * B_{1/a}) \\ &= \varepsilon + \ln N \, d_{\mathbb{S}}(Q_N, x_N * B_{1/a}) \qquad \forall N \ge N_1 \,. \end{aligned}$$

Consequently, $\bigl|\ln N \, d_{\mathbb{S}}(Q_N, x_N{*}B_{1/a}) - d_{\mathbb{S}}(B_{1/a}{*}B_{1/a}, B_{1/a})\bigr| < \varepsilon$ for every $N \ge N_1$, and since $\varepsilon > 0$ has been arbitrary,

$$\lim_{N\to\infty} \ln N \, d_{\mathbb{S}}(Q_N, x_N * B_{1/a}) = d_{\mathbb{S}}(B_{1/a} * B_{1/a}, B_{1/a}) \,.$$

(ii) By (i), it is possible to pick $c \in \mathbb{R}^+$ with

$$d_{\mathbb{S}}(Q_N, x_N * B_{1/a}) \le \frac{c}{\ln N} \qquad \forall N \ge 2 \,,$$

and given any increasing sequence (N_j) in $\mathbb{N}$ and any $s \in \mathbb{S}$, note that

$$\bigl|d_{\mathbb{S}}(Q_{N_j}, s * B_{1/a}) - d_{\mathbb{S}}(x_{N_j} * B_{1/a}, s * B_{1/a})\bigr| \le d_{\mathbb{S}}(Q_{N_j}, x_{N_j} * B_{1/a}) \le \frac{c}{\ln N_j}$$

for every $j \ge 2$. Deduce from this, together with (2.3), that

$$d_{\mathbb{S}}(Q_{N_j}, s * B_{1/a}) \le \frac{d_{\mathbb{S}}(B_{1/a}, \delta_1)}{a \log e} \, d_{\mathbb{S}}(S(x_{N_j}), s) + \frac{c}{\ln N_j} \qquad \forall j \ge 2 \,,$$

but also

$$d_{\mathbb{S}}(Q_{N_j}, s * B_{1/a}) \ge 4 d_{\mathbb{S}}(B_{1/a}, B) \, d_{\mathbb{S}}(S(x_{N_j}), s) - \frac{c}{\ln N_j} \qquad \forall j \ge 2 \,.$$

Thus $\lim_{j\to\infty} d_{\mathbb{S}}(Q_{N_j}, s * B_{1/a}) = 0$ if and only if $\lim_{j\to\infty} d_{\mathbb{S}}(S(x_{N_j}), s) = 0$.

(iii) With $c \in \mathbb{R}^+$ as in (ii), and since $P \mapsto x_N * P$ is an isometry of $\mathcal{P}$,

$$\bigl|d_{\mathbb{S}}(Q_N, B) - d_{\mathbb{S}}(B_{1/a}, B)\bigr| \le d_{\mathbb{S}}(Q_N, x_N * B_{1/a}) < \frac{c}{\ln N} \qquad \forall N \ge 2 \,,$$

which shows that $\lim_{N\to\infty} d_{\mathbb{S}}(Q_N, B) = d_{\mathbb{S}}(B_{1/a}, B)$.

(iv) For convenience, let $\mathcal{C}$ be the "circle" $\{s * B_{1/a} : s \in \mathbb{S}\}$. Notice that $\sup_{n\in\mathbb{N}} \ln n \, d_{\mathbb{S}}\bigl(S(x_{n+1}), S(x_n)\bigr) < \infty$ and $x_n \to \infty$, so $\Omega[S(x_n)] = \mathbb{S}$. Together with (ii), this implies that $\Omega[Q_N] \supset \mathcal{C}$. However, $\Omega[Q_N] \subset \mathcal{C}$ also, since every limit point of (Q_N) is of the form $s * B_{1/a}$ for some $s \in \mathbb{S}$, by (i). In summary, $\Omega[Q_N] = \mathcal{C}$. □

It is worthwhile contrasting Theorem 3.4 with its d_∞-counterpart which shows that $\bigl(\ln N \, d_\infty(Q_N, x_N * B_{1/a})\bigr)$ and $\bigl(d_\infty(Q_N, B)\bigr)$ both are bounded but divergent, with $\limsup = 2 \liminf$ for either sequence. This provides yet another reason why, for the purpose of this article, $d_{\mathbb{S}}$ is preferable to d_∞; cf. Remark 2.1. Since the result follows similarly from Lemma 3.3, all details are left to the interested reader.

Proposition 3.5. *Let $(x_n) = (p_n^a)$ for some $a \in \mathbb{R}^+$. Then*

(i) *$\Omega[\ln N \, d_\infty(Q_N, x_N * B_{1/a})] = [\frac{1}{2}c_1, c_1]$ with $c_1 = d_\infty(B_{1/a} * B_{1/a}, B_{1/a})$, and therefore in particular $\lim_{N\to\infty} d_\infty(Q_N, x_N * B_{1/a}) = 0$;*

(ii) *given an increasing sequence* (N_j) *in* $\mathbb{N}$ *and* $s \in \mathbb{S}$, $\lim_{j\to\infty} d_\infty(Q_{N_j}, s*B_{1/a}) = 0$ *if and only if* $\lim_{j\to\infty} d_{\mathbb{S}}(S(x_{N_j}), s) = 0$*;*
(iii) $\Omega[d_\infty(Q_N, B)] = [\frac{1}{2}c_2, c_2]$ *with* $c_2 = d_\infty(B_{1/a}, B)$.

By Theorem 3.4 and Proposition 3.5, for $(x_n) = (p_n^a)$ the set $\Omega\{x_n\}$ simply equals $\{s * B_{1/a} : s \in \mathbb{S}\}$, a "circle" with "center" B and "radius" $d_{\mathbb{S}}(B_{1/a}, B) < \frac{1}{4}$. Moreover, the (discrete) distribution Q_N approaches the (absolutely continuous) distribution $x_N * B_{1/a}$, at the precise rate $(1/\ln N)$, with respect to both $d_{\mathbb{S}}$ and d_∞, and in the "direction" of $x_N * B_{1/a} - x_N * B_{1/a} * B_{1/a}$. (The fact that $\lim_{N\to\infty} d_\infty(Q_N, x_N * B_{1/a}) = 0$ has been observed in [26] also, though somewhat opaquely, and without a precise rate or "direction" of convergence.) To identify an even better absolutely continuous approximation of Q_N, let

$$t_n = \frac{1}{a} - \frac{1}{a \max\{1, \ln n\}} \qquad \forall n \in \mathbb{N}\,. \tag{3.5}$$

Lemma 3.6. *Let* $(x_n) = (p_n^a)$ *for some* $a \in \mathbb{R}^+$*. Then, with* t_N *as in* (3.5)*,*

$$\lim_{N\to\infty} \ln N \, d_\infty(Q_N, x_N * B_{t_N}) = 0\,.$$

Proof. Observe that $a \ln N = -1/(t_N - 1/a)$ for every $N \geq 3$, and

$$\begin{aligned}
\ln N \, d_\infty(Q_N, x_N * B_{t_N}) &\leq \\
&\leq \left\| \ln N \, (F_{Q_N} - F_{x_N * B_{1/a}}) - (F_{x_N * B_{1/a} * B_{1/a}} - F_{x_N * B_{1/a}}) \right\|_\infty \\
&\quad + \frac{1}{a} \left\| a \ln N \, (F_{x_N * B_{t_N}} - F_{x_N * B_{1/a}}) + a(F_{x_N * B_{1/a}} - F_{x_N * B_{1/a} * B_{1/a}}) \right\|_\infty \\
&\leq \left\| \ln N \, (F_{Q_N} - F_{x_N * B_{1/a}}) - (F_{x_N * B_{1/a} * B_{1/a}} - F_{x_N * B_{1/a}}) \right\|_\infty \\
&\quad + \frac{2}{a} \left\| \frac{F_{B_{t_N}} - F_{B_{1/a}}}{t_N - 1/a} - \frac{F_{B_{1/a}} - F_{B_{1/a} * B_{1/a}}}{1/a} \right\|_\infty ,
\end{aligned}$$

where the last inequality follows from (3.2). Both the second-to-last and the last term goes to 0 as $N \to \infty$, by (3.3) and (2.8), respectively. □

Since (Q_N) is divergent for $(x_n) = (p_n^a)$, and also in view of the experimental evidence presented in [22], it is natural to study appropriate subsequences (Q_{N_j}). By Theorem 3.4, (Q_{N_j}) converges in $\mathcal{P}$ if and only if $\big(S(x_{N_j})\big)$ converges in $\mathbb{S}$. Under the slightly stronger assumption that, for some $s \in \mathbb{S}$,

$$\lim_{j\to\infty} \ln N_j \, d_{\mathbb{S}}(S(x_{N_j}), s) = 0\,, \tag{3.6}$$

more can be said about the convergence $Q_{N_j} \to s * B_{1/a}$: It takes place asymptotically along the smooth curve $(s * B_t)_{t\in\mathbb{R}^+}$.

Theorem 3.7. *Let* $(x_n) = (p_n^a)$ *for some* $a \in \mathbb{R}^+$*. Assume* (N_j) *is an increasing sequence in* $\mathbb{N}$ *that satisfies* (3.6) *for some* $s \in \mathbb{S}$*. Then*

$$\lim_{j\to\infty} \left\| \ln N_j \left(F_{Q_{N_j}} - F_{s*B_{1/a}}\right) - \left(F_{s*B_{1/a}*B_{1/a}} - F_{s*B_{1/a}}\right) \right\|_\infty = 0\,,$$

and with $t_j^* = t_{N_j}$ *as in* (3.5) *also* $\lim_{j\to\infty} \ln N_j \, d_\infty(Q_{N_j}, s * B_{t_j^*}) = 0$.

Proof. As a complement to (2.3), a straightforward calculation yields

$$d_\infty(r * B_t, s * B_t) \le \frac{2t^2 10^{2t}}{(\log e)^2(10^t - 1)}\, d_{\mathbb{S}}(r,s) \qquad \forall r,s \in \mathbb{S}, t \in \mathbb{R}^+ . \tag{3.7}$$

With this, the first assertion follows immediately by observing that, for every $j \in \mathbb{N}$,

$$\begin{aligned} \big\| \ln N_j \,(F_{Q_{N_j}} - F_{s*B_{t_j^*}}) - (F_{s*B_{1/a}*B_{1/a}} - F_{s*B_{1/a}}) \big\|_\infty \le \\ \le \big\| \ln N_j \,(F_{Q_{N_j}} - F_{x_{N_j}*B_{t_j^*}}) - (F_{x_{N_j}*B_{1/a}*B_{1/a}} - F_{x_{N_j}*B_{1/a}}) \big\|_\infty \\ + |\ln N_j - 1|\, d_\infty(x_{N_j} * B_{1/a}, s * B_{1/a}) \\ + d_\infty(x_{N_j} * B_{1/a} * B_{1/a}, s * B_{1/a} * B_{1/a}) , \end{aligned}$$

where each of the three terms on the right goes to 0 as $j \to \infty$, by (3.3), by (3.6) and (3.7), and by the continuity of $s \mapsto s * B_{1/a} * B_{1/a}$, respectively. Similarly, the second assertion follows from

$$\ln N_j\, d_\infty(Q_{N_j}, s * B_{t_j^*}) \le \ln N_j\, d_\infty(Q_{N_j}, x_{N_j} * B_{t_j^*}) + \ln N_j\, d_\infty(x_{N_j} * B_{t_j^*}, s * B_{t_j^*}) ,$$

where the first and second term on the right goes to 0 as $j \to \infty$, by Lemma 3.6, and by (3.6) and (3.7), respectively. □

To discuss but one basic example satisfying (3.6), let $N_j = \pi(10^{j/a})$ for all sufficiently large $j \in \mathbb{N}$. Deduce from (A.4), a quantitative refinement of the prime number theorem reviewed in Appendix A, that

$$\log N_j = \frac{j}{a} - \log j + \log(a \log e) + \frac{a(\log e)^2}{j} + \mathcal{O}\left(\frac{1}{j^2}\right) \quad \text{as } j \to \infty ,$$

and consequently, by (A.5),

$$\log(p_{N_j}^a) = j + \mathcal{O}\left(\frac{\ln j}{j^2}\right) \quad \text{as } j \to \infty .$$

Thus $\lim_{j\to\infty} d_{\mathbb{S}}(S(x_{N_j}), 1) = 0$, and $\Omega[Q_{N_j}] = \{B_{1/a}\}$, by Theorem 3.4. Moreover, (3.6) holds, and hence Theorem 3.7 shows that $Q_{N_j} \to B_{1/a}$ along $\mathcal{B}$, in that $B_{t_j^*}$ is an even better approximation of Q_{N_j} than is $B_{1/a}$. By contrast, letting, for instance, $N_j = \lfloor 10^{j/a} \rfloor$ for all sufficiently large $j \in \mathbb{N}$, it is readily seen that $\big(S(x_{N_j})\big)$ and (Q_{N_j}) both diverge, in fact $\Omega[S(x_{N_j})] = \mathbb{S}$ and $\Omega[Q_{N_j}] = \{s * B_{1/a} : s \in \mathbb{S}\}$.

To prove the assertions made in the Introduction regarding the first significant digit distributions (q_N) associated with $(x_n) = (p_n^a)$, apply $\lfloor \cdot \rfloor$ to the earlier results in this section as follows.

Theorem 3.8. *Let $(x_n) = (p_n^a)$ for some $a \in \mathbb{R}^+$. Then*

(i) $\sup_{N\in\mathbb{N}} \ln N \, \|q_N - \lfloor x_N * B_{1/a} \rfloor\|_\infty < \infty$, *and therefore in particular* $\lim_{N\to\infty}(q_N - \lfloor x_N * B_{1/a} \rfloor) = 0$;

(ii) *given an increasing sequence (N_j) in $\mathbb{N}$ and $s \in \mathbb{S}$, $\lim_{j\to\infty} q_{N_j} = \lfloor s * B_{1/a} \rfloor$ if and only if $\lim_{j\to\infty} d_{\mathbb{S}}(S(x_{N_j}), s) = 0$;*

(iii) *given an increasing sequence* (N_j) *in* $\mathbb{N}$ *that satisfies* (3.6) *for some* $s \in \mathbb{S}$, $\lim_{j\to\infty} \ln N_j \, (q_{N_j} - \lfloor s * B_{t_j^*} \rfloor) = 0$, *with* $t_j^* = t_{N_j}$ *as in* (3.5)*;*
(iv) $\inf_{N\in\mathbb{N}} \|q_N - b_0\|_\infty > 0$*;*
(v) $\Omega[q_N] = \{\lfloor s * B_{1/a} \rfloor : s \in \mathbb{S}\}$.

Proof. For convenience, define two functions $g_{1,2} : \mathbb{S} \to \mathbb{R}^9$ by

$$g_1(s) = \lfloor s * B_{1/a} \rfloor \,, \quad g_2(s) = \lfloor s * B_{1/a} * B_{1/a} \rfloor - \lfloor s * B_{1/a} \rfloor \qquad \forall s \in \mathbb{S} \,,$$

so $\lfloor t * B_{1/a} \rfloor = g_1\big(S(t)\big)$ for every $t \in \mathbb{R}^+$. Since $B_{1/a}$ and $B_{1/a} * B_{1/a}$ have no atoms, g_1 and g_2 are continuous. It is readily seen that $g_1(s) \neq b_0$ for every $s \in \mathbb{S}$, so $\min_{s\in\mathbb{S}} \|g_1(s) - b_0\|_\infty > 0$, and $g_{1,2}$ are one-to-one. Applying $\lfloor\cdot\rfloor$ to (3.3) yields

$$\lim_{N\to\infty} \Big(\ln N \, \big(q_N - g_1\big(S(x_N)\big)\big) - g_2\big(S(x_N)\big) \Big) = 0 \,. \tag{3.8}$$

Clearly, (i) is immediate from (3.8). To see (ii), notice that

$$\big| \|q_N - g_1(s)\|_\infty - \big\|g_1(s) - g_1\big(S(x_N)\big)\big\|_\infty \big| \leq \big\|q_N - g_1\big(S(x_N)\big)\big\|_\infty \qquad \forall N \in \mathbb{N}, s \in \mathbb{S} \,.$$

Thus if $S(x_{N_j}) \to s$ then $\|q_{N_j} - g_1(s)\|_\infty \to 0$, and if $\limsup_{j\to\infty} d_{\mathbb{S}}(S(x_{N_j}), s) > 0$ then

$$\limsup_{j\to\infty} \|q_{N_j} - g_1(s)\|_\infty \geq \limsup_{j\to\infty} \big\|g_1(s) - g_1\big(S(x_{N_j})\big)\big\|_\infty > 0 \,,$$

which proves (ii). Applying $\lfloor\cdot\rfloor$ to the second assertion of Theorem 3.7 yields (iii). To see (iv), note that $Nq_N \in \mathbb{Z}^9$ for every $N \in \mathbb{N}$, but $Nb_0 \notin \mathbb{Z}^9$, and so $\|q_N - b_0\|_\infty > 0$. Moreover, by (3.8),

$$\liminf_{N\to\infty} \|q_N - b_0\|_\infty \geq \min_{s\in\mathbb{S}} \|g_1(s) - b_0\|_\infty > 0 \,.$$

Finally, (v) is obvious from (i), together with the fact that $\Omega[S(x_n)] = \mathbb{S}$, as already noticed in the proof of Theorem 3.4. □

At long last, the *Proof of Theorem 1.1* is very short indeed: With $a = 1$ and g_1 as in the proof of Theorem 3.8, the function $g : \mathbb{R} \to \mathbb{R}^9$ with $g(t) = g_1(10^t) = \lfloor 10^t * U \rfloor$ has all the desired properties since, as is readily seen, $q_N[p_n] \notin \{g(t) : t \in \mathbb{R}\}$ for every $N \in \mathbb{N}$. Similarly, for the *Proof of Theorem 1.2* simply apply Theorem 3.8(iii) with $a = 1$, $s = 1$, and $N_j = \pi(10^j)$ for every $j \in \mathbb{N}$. □

4. Conclusion

The sequence (p_n^a) is not Benford for any $a \in \mathbb{R}^+$. Nonetheless it is intimately related to the generalized Benford's law $(B_t)_{t\in\mathbb{R}^+}$ and its first digit version $(b_t)_{t\in\mathbb{R}^+}$, in at least three ways: Firstly, the closed curves $\Omega[Q_N] = \{s * B_{1/a} : s \in \mathbb{S}\}$ and $\Omega[q_N] = \{\lfloor s * B_{1/a} \rfloor : s \in \mathbb{S}\}$ have B and b_0 as their respective barycenters. Thus (Q_N) and (q_N) "dance" around B and b_0, respectively. Secondly, for subsequences (N_j) satisfying (3.6), the convergence $Q_{N_j} \to B_{1/a}$ occurs asymptotically in the "direction" of $B_{1/a} - B_{1/a} * B_{1/a}$, and similarly $q_{N_j} \to b_{1/a}$ in the "direction" $b_{1/a} - \lfloor B_{1/a} * B_{1/a} \rfloor$; see Fig. 6. In this situation, $B_{t_j^*}$ and $b_{t_j^*}$, with $t_j^* = t_{N_j}$ as

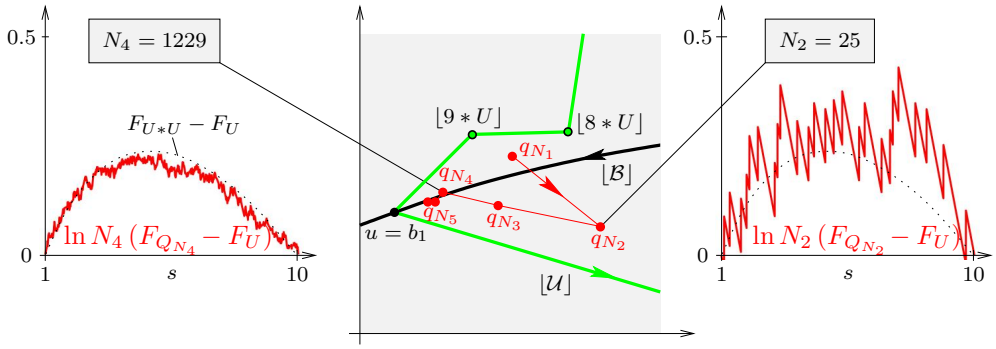


Fig. 6. Illustrating Theorem 3.7 and 3.8 for $a = 1$, $s = 1$: With $N_j = \pi(10^j)$, the distributions (Q_{N_j}) of $(x_n) = (p_n)$ converge to U as $j \to \infty$, at the precise rate $(1/j)$ and in the "direction" of $U - U * U$. The first digit distributions (q_{N_j}) converge to $u = b_1$, asymptotically along the smooth curve $\lfloor \mathcal{B} \rfloor$ of the generalized Benford's first digit law (center). Note that the same projection Π as in Fig. 2 and 4 has been used, but only the portion of $\lfloor \mathcal{P} \rfloor$ indicated in those figures by the smallest rectangles is shown here, as $\|q_{N_j} - u\|_\infty$ is quite small even for $j = 1, 2, \ldots, 6$.

in (3.5), are especially good approximations of Q_{N_j} and q_{N_j}, respectively. Thirdly, notice that $d_{\mathbb{S}}(s * B_{1/a}, B) = d_{\mathbb{S}}(B_{1/a}, B) \to 0$ as $a \to \infty$. Thus the "circle" $\Omega\{p_n^a\}$ approaches the singleton $\{B\}$ as $a \to \infty$, so (p_n^a) is *approximately* Benford for large a, a fact repeatedly recorded in the literature; see, e.g., [4] and [26, Cor.4].

In conclusion, and to put the results of this article in perspective, the reader may find the following two comments helpful: On the one hand, a similar analysis can be carried out for $(x_n) = (n^a)$, with $a \in \mathbb{R}^+$. It turns out that again Q_N approaches $N^a * B_{1/a}$ as $N \to \infty$, just like in Theorem 3.4, but does so at a much faster rate since

$$\lim_{N\to\infty} \frac{N}{\sqrt{\ln N}}\, d_{\mathbb{S}}(Q_N[n^a], N^a * B_{1/a}) = \sqrt{\frac{a \log e}{6\pi}}\, \sigma(a) \qquad \forall a \in \mathbb{R}^+\,, \tag{4.1}$$

where the function $\sigma : \mathbb{R}^+ \to \mathbb{R}^+$ is somewhat delicate and discontinuous, but $1 \le \sigma(a) \le \sqrt{3}$ for all $a \in \mathbb{R}^+$, and $\sigma(a) = 1$ for all but countably many a; see [1] for details. For example, $\sigma(1) = \frac{1}{3}\sqrt{11}$, and hence (4.1) for $a = 1$ reads

$$\lim_{N\to\infty} \frac{N}{\sqrt{\ln N}}\, d_{\mathbb{S}}(Q_N[n], N * U) = \frac{1}{3}\sqrt{\frac{11 \log e}{6\pi}} = 0.1678\,.$$

In addition to this, and in contrast to (3.4),

$$\lim_{N\to\infty} \frac{N}{\sqrt{\ln N}} \int_{\mathbb{S}} \ell \,\mathrm{d}(Q_N[n] - N * U) = 0 \qquad \forall \ell \in \mathrm{Lip}_1(\mathbb{S})\,,$$

so $Q_N[n]$ does not approach $N * U$ in any one "direction".

On the other hand, it is illuminating to once again contrast (p_n) with (F_n). Although the latter sequence is Benford, in $\lim_{N\to\infty} Q_N[F_n] = B$ neither the precise rate of convergence nor its "direction" (if any) appear to be known. As a theoretical

bound on the rate of convergence for *every* sequence (x_n) in $\mathbb{R}^+$, it follows from [18] that

$$\limsup_{N\to\infty} \frac{N}{\sqrt{\ln N}}\, d_{\mathbb{S}}(Q_N[x_n], B) > 0\,.$$

In addition, if the transcendental number $\log\phi$, where $\phi = \frac{1}{2}(1+\sqrt{5})$, had bounded partial quotients, then [21, Thm. 2.3.4] would imply that

$$\limsup_{N\to\infty} \frac{N}{\ln N}\, d_{\mathbb{S}}(Q_N[F_n], B) < \infty\,. \tag{4.2}$$

As far as the authors have been able to ascertain, it is unknown whether or not $\log\phi$ has bounded partial quotients. Experimental evidence seems to suggest that (4.2) indeed is valid, even with $\sqrt{\ln N}$ instead of $\ln N$; see Fig. 7. However, due to the presence of a factor $\sqrt{\ln N}$ which may be more or less undetectable experimentally, any such evidence should be viewed with a healthy dose of scepticism.

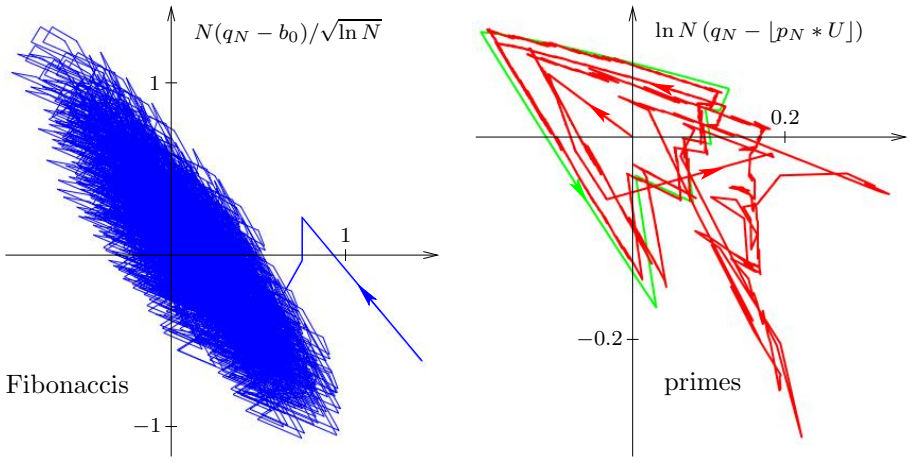


Fig. 7. For the Fibonaccis, the first digit distributions (q_N) converge to b_0, and the sequence $\big(N(q_N - b_0)/\sqrt{\ln N}\big)$ appears to be bounded, yet rather complicated. By contrast, for the primes (q_N) diverges, with $\Omega[q_N] = \lfloor \mathcal{U} \rfloor$, and $\ln N\,(q_N - \lfloor p_N * U \rfloor)$ approaches the green closed curve $\big\{\lfloor s * U * U \rfloor - \lfloor s * U \rfloor : s \in \mathbb{S}\big\}$ as $N \to \infty$; note that this curve is not a polygon. For graphical depiction, the same projection $\mathbb{R}^9 \to \mathbb{R}^2$ as in Fig. 2, 4, and 6 has been used, and successive points are joined by straight line segments, similarly to Fig. 2.

Acknowledgments

The first author was partially supported by an NSERC Discovery Grant. Both authors are grateful to an anonymous referee whose constructive criticism greatly helped improve the presentation.

Appendix A. Proof of a technical lemma

The purpose of this appendix is to provide a proof of Lemma 3.3. As presented here, the argument involves a modicum of straightforward but lengthy calculations, not unlike the comparable analyses in [6,25,26]. To capture the essence of the argument, it is helpful to first analyze a slightly more general situation: Let $f : \mathbb{R}^+ \to \mathbb{R}$ be any smooth function given by

$$f(t) = a_1(\ln t + \ln\ln t) + a_2 + a_3\frac{\ln\ln t}{\ln t} + \frac{a_4 + \widetilde{f}(t)}{\ln t} \qquad \forall t \geq e\,, \tag{A.1}$$

where $a_1 \in \mathbb{R}^+$, $a_2, a_3, a_4 \in \mathbb{R}$, and $\lim_{t\to\infty} \widetilde{f}(t) = 0$. As detailled below, the essence of Lemma 3.3 is contained in the following auxiliary result.

Lemma A. 1. *Assume the sequence (x_n) in $\mathbb{R}^+$ satsfies $\log x_n = f(n)$ for every $n \in \mathbb{N}$, with an increasing smooth function $f : \mathbb{R}^+ \to \mathbb{R}$ as in* (A.1)*. Then* (3.1) *holds with $a = a_1/\log e$.*

Proof. Denote by g the inverse of f, and deduce from (A.1) that, with some $t_0 \in \mathbb{R}^+$,

$$g(t) = \frac{a_1 e^{(t-a_2)/a_1}}{t-a_2}\left(1 + (a_1-a_3)\frac{\ln(t-a_2)}{t-a_2} - \frac{a_4 + (a_1-a_3)\ln a_1 + \widetilde{g}(t)}{t-a_2}\right) \tag{A.2}$$

for all $t \geq t_0$, where $\lim_{n\to\infty} \widetilde{g}(t) = 0$. Notice that for all $n \in \mathbb{N}$ and $s \in \mathbb{S}$,

$$\begin{aligned} NF_{x_N^{-1}*Q_N}(s) &= \#\{1 \leq n \leq N : S(x_N^{-1}) * S(x_n) \in [1,s]\} \\ &= \#\{1 \leq n \leq N : f(n) - f(N) \in [0, \log s] + \mathbb{Z}\}\,. \end{aligned}$$

For convenience, let $\ell_n = \ln n$ for every $n \in \mathbb{N}$, and $r = \log s$ for every $s \in \mathbb{S}$; also write $F_{x_N^{-1}*Q_N}(10^r)$ simply as $G_N(r)$. Note that $\lim_{N\to\infty} f(N)/\ell_N = a_1$. Since $f(n) - f(N) \in [0,r] - j$ with $j \in \mathbb{N}$ if and only if

$$n \in [g(f(N)-j), g(f(N)-j+r)] \cap \{1, 2, \ldots, N\}\,,$$

provided that $f(N) - j \geq t_0$, there exist $N_0 \in \mathbb{N}$ and $c_1 \in \mathbb{R}^+$ so that

$$\left| NG_N(r) - \sum\nolimits_{j=1}^{\lfloor f(N)-t_0\rfloor} \big(g(f(N)+r-j) - g(f(N)-j)\big)\right| \leq c_1\ell_n \tag{A.3}$$

for every $N \geq N_0$ and $r \in [0,1]$, and also $f(N) - t_0 \geq (\ln \ell_N)^2$. From (A.1) and (A.2), it is readily deduced that, given $\varepsilon > 0$, there exists $N_\varepsilon \geq N_0$ so that

$$\left|\frac{g(f(N)-t)}{N} - e^{-t/a_1}\left(1 + \frac{t}{a_1\ell_N}\right)\right| < \frac{\varepsilon e^{-t/a_1}}{\ell_N} \qquad \forall N \geq N_\varepsilon, t \in [0, (\ln\ell_N)^2]\,,$$

but also

$$\left|\frac{g(f(N)-t)}{N}\right| \leq \ell_N^2 e^{-t/a_1} \qquad \forall N \geq N_\varepsilon, t \in [(\ln\ell_N)^2, f(N)-t_0]\,.$$

Combining these estimates with (A.3) yields, for every $N \geq N_\varepsilon$ and uniformly in $r \in [0,1[$,

$$\begin{aligned}
&\left| G_N(r) - \sum\nolimits_{j=1}^{\infty} \left(e^{(r-j)/a_1} \left(1 + \frac{j-r}{a_1 \ell_N} \right) \right) - e^{-j/a_1} \left(1 + \frac{j}{a_1 \ell_N} \right) \right| \\
&\qquad \leq \frac{c_1 \ell_N}{N} + \frac{2\varepsilon}{\ell_N} \sum\nolimits_{j=1}^{\lfloor (\ln \ell_N)^2 \rfloor} e^{(1-j)/a_1} + 2\ell_N^2 \sum\nolimits_{j=\lfloor (\ln \ell_N)^2 \rfloor + 1}^{\lfloor f(N) - t_0 \rfloor} e^{(1-j)/a_1} + \\
&\qquad\qquad + 2 \sum\nolimits_{j=\lfloor (f(N)-t_0 \rfloor + 1}^{\infty} e^{(1-j)/a_1} \left(1 + \frac{j}{a_1 \ell_N} \right) \\
&\qquad \leq \frac{c_1 \ell_N}{N} + \frac{2\varepsilon}{\ell_N (e^{1/a_1} - 1)} + \frac{2e^{2/a_1}(\ell_N^2 + 1)}{e^{1/a_1} - 1} e^{-(\ln \ell_N)^2/a_1} + \\
&\qquad\qquad + \frac{2e^{1/a_1}}{a_1 \ell_N} \sum\nolimits_{j=\lfloor (\ln \ell_N)^2 \rfloor + 1}^{\infty} e^{-j/a_1} j \\
&\qquad < \frac{2\varepsilon}{\ell_N (e^{1/a_1} - 1)} + \frac{c_2}{\ell_N^2} ,
\end{aligned}$$

where $c_2 \in \mathbb{R}^+$ depends on a_1, but is independent of N, r, and ε. Finally, observe that for every r and $N \geq 2$, and with $\tau = \log e / a_1$ for convenience,

$$\begin{aligned}
&\sum\nolimits_{j=1}^{\infty} \left(e^{(r-j)/a_1} \left(1 + \frac{j-r}{a_1 \ell_N} \right) - e^{-j/a_1} \left(1 + \frac{j}{a_1 \ell_N} \right) \right) \\
&\qquad\qquad = \frac{e^{r/a_1} - 1}{e^{1/a_1} - 1} + \frac{1}{a_1 \ell_N} \left(\frac{e^{1/a_1}(e^{r/a_1} - 1)}{(e^{1/a_1} - 1)^2} - \frac{r e^{r/a_1}}{e^{1/a_1} - 1} \right) \\
&\qquad\qquad = F_{B_\tau}(10^r) + \frac{1}{\ell_N} \left(F_{B_\tau * B_\tau}(10^r) - F_{B_\tau}(10^r) \right) .
\end{aligned}$$

Thus, for every $N \geq N_\varepsilon$ and $r \in [0,1[$,

$$\left| \ell_N \big(G_N(r) - F_{B_\tau}(10^r) \big) - \big(F_{B_\tau * B_\tau}(10^r) - F_{B_\tau}(10^r) \big) \right| < \frac{2\varepsilon}{e^{1/a_1} - 1} + \frac{c_2}{\ell_N}$$

and hence, uniformly in $r \in [0,1[$,

$$\lim\nolimits_{N \to \infty} \ell_N \big(F_{x_N^{-1} * Q_N}(10^r) - F_{B_\tau}(10^r) \big) = F_{B_\tau * B_\tau}(10^r) - F_{B_\tau}(10^r) .$$

With $a = a_1 / \log e$, recalling that $\tau = 1/a$ completes the proof. □

In order to utilize Lemma A.1 for a proof of Lemma 3.3, denote by π the prime counting function, i.e., $\pi(t) = \#\{n \in \mathbb{N} : p_n \leq t\}$ for all $t \in \mathbb{R}^+$, and notice that, as a quantitative refinement of the prime number theorem established in [15],

$$\limsup\nolimits_{t \to \infty} (\ln t)^2 \left| \frac{\pi(t) \ln t}{t} - \left(1 + \frac{1}{\ln t} \right) \right| < \infty . \tag{A.4}$$

Proof of Lemma 3.3. Deduce from (A.4) that

$$\frac{p_n}{n} = \ln n + \ln \ln n - 1 + \mathcal{O}\left(\frac{\ln \ln n}{\ln n} \right) \quad \text{as } n \to \infty , \tag{A.5}$$

and consequently, as $n \to \infty$,

$$\log x_n = a \log p_n = a \log e \left(\ln n + \ln\ln n + \frac{\ln\ln n}{\ln n} - \frac{1}{\ln n} + \mathcal{O}\left(\frac{(\ln\ln n)^2}{(\ln n)^2} \right) \right) .$$

Thus $\log x_n = f(n)$ for every $n \in \mathbb{N}$, with f as in (A.1), where $a_1 = a_3 = -a_4 = a \log e$, $a_2 = 0$, and $\widetilde{f}$ is appropriately chosen. An application of Lemma A.1 therefore completes the proof. □

Remark A. 2. The attentive reader will have noticed that none of the parameters a_2, a_3, a_4, and $\widetilde{f}$ in (A.1) appears explicitly in the conclusion of Lemma A.1. This suggests that Lemma 3.3, and hence its many consequences presented in Section 3 as well, remain valid for sequences (x_n) more general than (p_n^a); cf. [21,25,26,32].

References

[1] A. Berger, A note on the distribution of $(\log n)$ mod 1, *Unif. Distrib. Theory* **17**(2022), 77–100.
[2] A. Berger, L.A. Bunimovitch, and T.P. Hill, One-dimensional dynamical systems and Benford's law, *Trans. Amer. Math. Soc.* **357**(2005), 197–219.
[3] A. Berger and T.P. Hill, Benford's law strikes back: no simple explanation in sight for mathematical gem, *Math. Intelligencer* **33**(2011), 85–91.
[4] A. Berger and T.P. Hill, *An Introduction to Benford's Law*, Princeton University Press (2015).
[5] A. Berger, T.P. Hill, and E. Rogers, Benford Online Bibliography, `www.benfordonline.net`, 2009. (Last accessed 25 October 2024.)
[6] A. Berger and A. Rahmatidehkordi, Circling the uniform distribution, *J. Math. Anal. Appl.* **527**(2023), to appear.
[7] S. Bobkov and M. Ledoux, One-dimensional empirical measures, order statistics, and Kantorovich transport distances, *Mem. Amer. Math. Soc.* **261**(2019), 1259.
[8] L. Brown and S. Steinerberger, On the Wasserstein distance between classical sequences and the Lebesgue measure, *Trans. Amer. Math. Soc.* **373**(2020), 8943–8962.
[9] C.A. Cabrelli and U.M. Molter, The Kantorovich metric for probability measures on the circle, *J. Comput. Appl. Math.* **57**(1995), 345–361.
[10] Z. Cai, M. Faust, A.J. Hildebrand, J. Li, and Y. Zhang, The Surprising Accuracy of Benford's Law in Mathematics, *Amer. Math. Monthly* **127**(2020), 217-237.
[11] J. Cigler, Asymptotische Verteilung reeller Zahlen mod 1, *Monatsh. Math.* **64**(1960), 201–225.
[12] A. Deitmar and S. Echterhoff, *Principles of Harmonic Analysis*, Springer, 2009.
[13] R.M. Dudley, *Real analysis and probability*, Cambridge Studies in Advanced Mathematics **74**, Cambridge University Press, 2002.
[14] R.L. Duncan, An application of uniform distributions to the Fibonacci numbers, *Fibonacci Quart.* **5**(1967), 137–140.
[15] P. Dusart, The kth prime is greater than $k(\ln k + \ln\ln k - 1)$ for $k \geq 2$, *Math. Computation* **68**(1999), 411–415.
[16] B.J. Flehinger, On the probability that a random integer has initial digit A, *Amer. Math. Monthly* **73**(1966), 1056–1061.
[17] A.L. Gibbs and F.E. Su, On choosing and bounding probability metrics, *International Statistical Review* **70**(2002), 419–435.

[18] C. Graham, Irregularity of distribution in Wasserstein distance, *J. Fourier Anal. Appl.* **26**(2020), 21pp.
[19] T.P. Hill, Base-invariance implies Benford's law, *Proc. Amer. Math. Soc.* **123**(1995), 887–895.
[20] J.H.B. Kemperman, Distributions modulo 1 of slowly changing sequences, *Nieuw Arch. Wisk.* **21**(1973), 138–163.
[21] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience, 1974.
[22] B. Luque and L. Lacasa, The first-digit frequencies of prime numbers and Riemann zeta zeros, *Proc. R. Soc. A* **465**(2009), 2197–2216.
[23] B. Massé and D. Schneider, A survey on weighted densities and their connection with the first digit phenomenon, *Rocky Mountain J. Math.* **41**(2011), 1395–1415.
[24] S.J. Miller, *Benford's Law: Theory and Applications*, Princeton University Press, 2015.
[25] Y. Ohkubo and O. Strauch Distribution of leading digits of numbers, *Unif. Distrib. Theory* **11**(2016), 23–45.
[26] Y. Ohkubo and O. Strauch, Distribution of leading digits of numbers II, *Unif. Distrib. Theory* **14**(2019), 19–42.
[27] L. Pietronero, E. Tosatti, V. Tosatti, and A. Vespigniani, Explaining the uneven distribution of numbers in nature: the laws of Benford and Zipf, *Physica A* **293**(2001), 297–304.
[28] R.A. Raimi, The first digit problem, *Amer. Math. Monthly* **83**(1976), 521–538.
[29] P. Schatte, On the uniform distribution of certain sequences and Benford's law, *Math. Nachr.* **136**(1988), 271–273.
[30] S. Steinerberger, Wasserstein distance, Fourier series and applications, *Monatsh. Math.* **194**(2021), 305–338.
[31] O. Strauch, *Distribution of sequences: a theory*, Veda a Academia Nakladatelství, 2019.
[32] O. Strauch and Š. Porubský, *Distribution of sequences: a sampler*, Schriftenreihe der Slowakischen Akademie der Wissenschaften **1**, Peter Lang, 2005.
[33] L.C. Washington, Benford's law for Fibonacci and Lucas numbers, *Fibonacci Quart.* **19**(1981), 175–177.
[34] A. Wintner, On the cyclical distribution of the logarithms of the prime numbers, *Quart. J. Math. Oxford Ser.* **6**(1935), 65–68.
[35] R. Wong, Asymptotics of linear recurrences, *Anal. Appl.* **12**(2014), 463–484.