Pure and Applied Functional Analysis
Volume 11, Number 1, 2026, 1–23

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL
© Copyright 2026

A BRIEF SURVEY OF BENFORD'S LAW IN DYNAMICAL SYSTEMS

ARNO BERGER AND THEODORE P. HILL

Abstract. This article provides a brief overview on a range of basic dynamical systems that conform to the logarithmic distribution of significant digits known as Benford's law. As presented here, most theorems are special cases of known, more general results about dynamical systems whose orbits or trajectories follow this logarithmic law, in one way or another. These results span a wide variety of systems: autonomous and non-autonomous; discrete- and continuous-time; one- and multi-dimensional; deterministic and stochastic. Illustrative examples include familiar systems such as the tent map, Newton's root-finding algorithm, and geometric Brownian motion. The treatise is informal, with the goal of showcasing to the specialists the generality and universal appeal of Benford's law throughout the mathematical field of dynamical systems. References to complete proofs are provided for each known result, while one new theorem is presented in some detail.

1. Introduction

The purpose of this article, written in honor of the authors' friend and colleague Professor Leonid Bunimovich on the occasion of his 75th birthday, is to present an informal overview of the remarkable ubiquity of the statistical phenomenon called **Benford's law** (henceforth abbreviated as **BL**) in the mathematical field of dynamical systems. This topic is especially fitting to honor Leonid, since it was in joint work with him [2] that the authors first explored some of its many intriguing aspects, in response to the question "Do dynamical systems follow Benford's law?" raised by Tolle, Budzien, and LaViolette [19] after they had found experimental evidence of BL in several familiar physical dynamical systems; see also [11, 16, 18].

The tone of this survey is intentionally informal with the goal of introducing researchers in dynamical systems to a variety of mathematical examples of BL. As such, most of the theorems presented here are not new, nor are they stated in the most general form possible. One new theorem (Theorem 6.1 below) is included, however, and references to the proofs of all other results are given. As the reader is going to see, the results primarily take the form of sufficient conditions for a variety of dynamical systems to conform to BL, in one way or another.

The article is organized as follows: Section 2 provides a brief introduction to BL. Sections 3 to 5 present a selection of typical results regarding BL for various basic

2020 *Mathematics Subject Classification.* Primary 37-02, 37A50; Secondary 37H05, 60-02.
Key words and phrases. Dynamical system, Benford's law, differential equation, iterated maps, significant digits.

classes of dynamical systems. Examples are given to illustrate each theorem, though the presentation of these examples is brief, with most of the (usually elementary) details being left to the interested reader. In keeping with the informal tone, the selection of material in this survey is somewhat narrow; for a more comprehensive overview the reader may want to turn, e.g., to [7]. Finally, Section 6 establishes conformance to BL for the solutions of a class of highly non-linear two-step recursions, utilizing dynamical systems ideas and techniques.

Throughout, standard notation and terminology is used. Specifically, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}_+$, and $\mathbb{R}$ denote the sets of all positive integers, integers, positive reals, and reals, respectively, equipped with their usual arithmetic, order, and topology. Half-open intervals in $\mathbb{R}$ are written as $[a, b[= \{x \in \mathbb{R} : a \leq x < b\}$, and similarly for open or closed intervals. Also, $\lfloor x \rfloor = \max\{k \in \mathbb{Z} : k \leq x\}$ for every $x \in \mathbb{R}$. The cardinality of any finite set A is $\#A$. As usual, a function $T : \mathbb{R} \to \mathbb{R}$ being C^m, with $m \in \mathbb{N}$, means that T is m times continuously differentiable.

2. Basic definitions

For the purpose of this article, by a *dynamical system* is simply meant any mathematical model of how a particular process unfolds over time. Given is a fixed set $\mathbb{X}$ ("phase space"), together with a rule T that specifies how the process unfolds in $\mathbb{X}$. The phase space may consist of real numbers, vectors, or random variables, etc., and the changes may occur in discrete or continuous time. Given $x_0 \in \mathbb{X}$ and a rule T, the orbit $O_T(x_0)$ is the trajectory of the process as it moves through the state space $\mathbb{X}$, starting at x_0 and proceeding according to the rule specified by T.

In this survey, the emphasis is on *decimal* representations of numbers, the classical setting of BL, so here and throughout $\log t$ denotes the base-10 logarithm of $t \in \mathbb{R}_+$, and all digits are *decimal* digits. For other bases such as binary or hexadecimal, analogous results hold with very little change, simply by replacing log with $\log_b$ for the appropriate base $b > 1$; the interested reader is referred to [7, p. 9] for details.

The basic notion underlying BL concerns the *first* (or *leading*) *significant digit* and, more generally, the *significand* of a real number. Formally, for every x in $\mathbb{R}_+$, the (**decimal**) **significand** of x, denoted $S(x)$, is given by $S(x) = t$, where t is the unique number in $[1, 10[$ with $x = 10^k t$ for some (necessarily unique) $k \in \mathbb{Z}$. For negative x, let $S(x) = S(-x)$, and also $S(0) = 0$ for convenience. The integer $D_1(x) = \lfloor S(x) \rfloor$ is the **first significant** (**decimal**) **digit** of x, so $D_1(x) \in \{1, 2, \ldots, 9\}$ for every x in $\mathbb{R} \setminus \{0\}$. For example, $S(2025) = 2.025 = S(0.02025) = S(-20.25)$ and $D_1(2025) = 2 = D_1(0.02025) = D_1(-20.25)$.

Informally, a real-valued orbit, trajectory, or dataset is *Benford* if its significands are logarithmically distributed, that is, if for every $t \in [1, 10[$ the (limiting) proportion of significands with value $\leq t$ is exactly $\log t$. For instance, if a sequence of numbers is Benford (precise definition below), then exactly $\log 2 \cong 30.10\%$ of these numbers have first significant digit $D_1 = 1$, exactly $\log \frac{3}{2} \cong 17.60\%$ have first significant digit $D_1 = 2$, etc. Figure A shows the statistics of D_1 for three familiar integer sequences. As suggested by this figure, it is straightforward to perform a preliminary check on whether or not a dataset is Benford, simply by looking at the frequencies of the first significant digits. For this reason, BL or, more precisely, the

first-digit law derived from it, is widely used to check for indications of anomalies in data; see, e.g., Chapter 10 in [7], or search [9] for "fraud".

D_1	1	2	3	4	5	6	7	8	9
(2^n)	30.10	17.61	12.49	9.70	7.91	6.70	5.79	5.12	4.58
(F_n)	30.11	17.62	12.50	9.68	7.92	6.68	5.80	5.13	4.56
$(n!)$	29.56	17.89	12.76	9.63	7.94	7.15	5.71	5.10	4.26
exact BL	30.10	17.60	12.49	9.69	7.91	6.69	5.79	5.11	4.57

FIGURE A. Proportions (in percent) of first significant digits $D_1 = 1, 2, \ldots, 9$ for the first $N = 10^4$ entries of (x_n), for the *first-order autonomous* system $x_n = 2x_{n-1}$ and $x_0 = 1$ (powers of 2); the *second-order linear system* $x_n = x_{n-1} + x_{n-2}$ and $x_1 = x_2 = 1$ (Fibonacci numbers); and the *non-autonomous* system $x_n = nx_{n-1}$ and $x_0 = 1$ (factorials). Each proportion matches quite closely the corresponding precise proportion in BL (bottom row), in the case of (2^n) and (F_n) quite strikingly so; see also Example 5.4 below. This match is indicative of the fact that all three sequences indeed are Benford, according to Definition 2.1.

The following definition formalizes conformance to BL for various mathematical objects that may arise from dynamical systems.

Definition 2.1. (i) A (real) *sequence* $(x_n) = (x_1, x_2, x_3, \ldots)$ is **Benford** if

$$\lim_{N\to\infty} \frac{\#\{1 \le n \le N : S(x_n) \le t\}}{N} = \log t \qquad \forall t \in [1, 10[\,;$$

(ii) a (Borel) *function* $f : [0, \infty[\to \mathbb{R}$ is **Benford** if, with λ denoting Lebesgue measure,

$$\lim_{N\to\infty} \frac{\lambda\big(\{x \in [0, N[\,: S\big(f(x)\big) \le t\}\big)}{N} = \log t \qquad \forall t \in [1, 10[\,;$$

(iii) a (real-valued) *random variable* X is **Benford** if

$$P\{S(X) \le t\} = \log t \qquad \forall t \in [1, 10[\,;$$

(iv) a *sequence of random variables* $(X_n) = (X_1, X_2, X_3, \ldots)$ **converges in distribution to BL** if

$$\lim_{n\to\infty} P\{S(X_n) \le t\} = \log t \qquad \forall t \in [1, 10[\,,$$

and is **Benford with probability one** if

$$P\{(X_1, X_2, X_3, \ldots) \text{ is Benford}\} = 1;$$

(v) a *stochastic process* $(X_t)_{t\geq 0}$ **converges weakly to BL** if (X_{t_n}) converges in distribution to BL for every increasing sequence (t_n) with $t_n \to \infty$, and is **Benford with probability one** if the function (or *path*) $t \mapsto X_t$ is Benford with probability one.

As an immediate consequence of Definition 2.1, if a sequence (x_n) is Benford then

$$\lim_{N\to\infty} \frac{\#\{1 \leq n \leq N : D_1(x_n) = j\}}{N} = \log \frac{j+1}{j} \qquad \forall j \in \{1, 2, \ldots, 9\}\,, \tag{2.1}$$

and similarly for a Benford function, random variable etc. Though BL is most often encountered in this first-digit form, notice that (2.1) is weaker than the Benford property; for instance, the sequence $(x_n) = \big(D_1(2^n)\big)$ satisfies (2.1) but clearly is not Benford.

Many intriguing and not immediately obvious characterizations for the Benford properties of Definition 2.1 can be found in the literature; see, e.g., Chapters 4 and 5 of [7]. For later use, only one such characterization pertaining to property (i) is stated here; throughout, let $\log 0 = 0$ for convenience, and abbreviate *uniformly distributed modulo one* as *u.d. mod* 1.

Theorem 2.2. *For every sequence (x_n) in $\mathbb{R}$ the following are equivalent:*

(i) *(x_n) is Benford;*
(ii) *(ax_n) is Benford for every $a \in \mathbb{R} \setminus \{0\}$;*
(iii) *$(\log |x_n|)$ is u.d. mod* 1.

Proof. See [7, Thms. 4.2 and 4.4]. □

3. One-dimensional discrete-time systems

Throughout, given any set $\mathbb{X}$, any map $T : \mathbb{X} \to \mathbb{X}$ or sequence of maps $T_n : \mathbb{X} \to \mathbb{X}$, and any point $x_0 \in \mathbb{X}$, denote by $O_T(x_0)$ the unique sequence in $\mathbb{X}$ with $x_n = T(x_{n-1})$ or $x_n = T_n(x_{n-1})$ for all $n \in \mathbb{N}$. Refer to $O_T(x_0)$ as the **orbit** of x_0 under T or (T_n). A key object of interest throughout this article is what may be referred to as the *Benford set* $\mathbb{B}$: If $\mathbb{X} \subset \mathbb{R}$ then simply

$$\mathbb{B} = \big\{x_0 \in \mathbb{X} : O_T(x_0) \text{ is Benford}\big\}\,.$$

If $\mathbb{B}$ constitutes a (measure-theoretically or topologically) large part, or even all of $\mathbb{X}$ then BL can rightly be viewed as a prominent feature of the dynamics of T or (T_n).

3.1. Autonomous systems.

Theorem 3.1. *Let $T : \mathbb{R}_+ \to \mathbb{R}_+$ satisfy $T(x) = ax + g(x)$ for some $a > 1$ and $g = g(x) \geq 0$. If $\log a$ is irrational and $g(x) = o(x)$ as $x \to \infty$ then $\mathbb{B} = \mathbb{R}_+$; if $\log a$ is rational and $g(x) = o(x/\log x)$ as $x \to \infty$ then $\mathbb{B} = \varnothing$.*

Proof. See [7, Thm. 6.13]. □

Example 3.2. The orbit of x_0 under $T(x) = ax$ is $(ax_0, a^2x_0, a^3x_0, \ldots)$. By Theorem 3.1, the sequences (2^n) and (5^n) are Benford; similarly $(5 \cdot 2^n)$ and $(2 \cdot 5^n)$ are Benford. Clearly (10^n) is not Benford.

Example 3.3. By Theorem 3.1, the orbit of x_0 under $T(x) = 2x + e^{-x}$ is Benford for every $x_0 \in \mathbb{R}_+$, and indeed for every $x_0 \in \mathbb{R}$ — notwithstanding the fact that no simple explicit formula for T^n is known to the authors.

Theorem 3.4. *Let $T : \mathbb{R} \to \mathbb{R}$ be C^2 with $T(0) = 0$ and $0 < |T'(0)| < 1$, and $c \in \mathbb{R}_+$ sufficiently small. If $\log |T'(0)|$ is irrational then $\mathbb{B} \supset [-c, c] \setminus \{0\}$; if $\log |T'(0)|$ is rational then $\mathbb{B} \cap [-c, c] = \varnothing$.*

Proof. See [7, Cor. 6.18]. □

Example 3.5. Given any $a \in \mathbb{R}$, the map $T(x) = x + ae^{-x} - a$ is smooth with $T(0) = 0$ and $T'(0) = 1 - a$. By Theorem 3.4, the orbit $O_T(x_0)$ is Benford for all x_0 sufficiently close to 0 if $a = 0.1$, but for no such x_0 if $a = 0.9$. In fact, $O_T(x_0)$ is Benford unless $T(x_0) = 0$ if $a = 0.1$, but is not Benford for any $x_0 \in \mathbb{R}$ if $a = 0.9$.

For the statement of the next theorem, as well as several results in a similar spirit later on, say that $\mathbb{B}$ has **full measure** in a set $\mathbb{Y} \subset \mathbb{X}$ if $\mathbb{Y} \setminus \mathbb{B}$ is a nullset. Thus, if $\mathbb{B}$ has full measure in $\mathbb{Y}$ then (Lebesgue) almost every element of $\mathbb{Y}$ also belongs to $\mathbb{B}$. Say that $\mathbb{B}$ is **meagre** (or **of first category**) in $\mathbb{Y}$ if $\mathbb{B} \cap \mathbb{Y}$ is a union of countably many nowhere dense sets in $\mathbb{Y}$. Note that if $\mathbb{B}$ is meagre in $\mathbb{Y}$, and if $\mathbb{Y}$ is a Baire space (with the topology inherited from $\mathbb{X}$) with no isolated points, e.g., if $\mathbb{Y}$ is any non-degenerate interval, then $\mathbb{Y} \setminus \mathbb{B}$ is dense in $\mathbb{Y}$ and uncountable; in particular, $\mathbb{Y} \setminus \mathbb{B} \neq \varnothing$ in this case. As is well known, it is not uncommon in dynamics for sets of interest (or their complements) to have full measure and yet be meagre. The following simple result indicates how such a situation may occur naturally in any study of BL based on Theorem 2.2; see, e.g., [1, Lem. 3.2].

Proposition 3.6. *Given any number $b > 1$ and sequence (c_n) in $\mathbb{R}$, the set*

$$\big\{y \in \mathbb{R} : (b^n y + c_n) \text{ is u.d. mod } 1\big\}$$

has full measure but is meagre in $\mathbb{R}$.

Theorem 3.7. *Let $T : \mathbb{R}_+ \to \mathbb{R}_+$ be C^1 with $T(x) = ax^b + g(x)$ for some $a \in \mathbb{R}_+, b > 1$. If $g'(x) = o(x^{b-1}/\log x)$ as $x \to \infty$ then there exists $c \in \mathbb{R}_+$ so that $\mathbb{B}$ has full measure but is meagre in $[c, \infty[$.*

Proof. See [7, Thm. 6.23]. □

In the context of Theorem 3.7, note that while the set $\mathbb{B} \cap [c, \infty[$ is large measure-theoretically, the set $[c, \infty[\setminus \mathbb{B}$ is large topologically as it is *residual*, i.e., it contains an intersection of countably many open dense sets, and hence is dense in $[c, \infty[$ and uncountable. Thus the Benford set and its complement are intertwined in a very intricate way indeed. Such a state of affairs is going to be observed repeatedly throughout the remainder of this article.

Example 3.8. The orbit of x_0 under $T(x) = x^2 + 1$ is Benford for almost all $x_0 \in \mathbb{R}$, but the set $\{x_0 \in \mathbb{R} : O_T(x_0) \text{ is not Benford}\}$ is residual in $\mathbb{R}$, hence in particular dense and uncountable. It is unknown whether or not the integer sequence $O_T(1)$, or in fact $O_T(k)$ for any $k \in \mathbb{Z}$, is Benford; see [7, Exp. 6.25].

Theorem 3.9. *Let $T : \mathbb{R} \to \mathbb{R}$ be real-analytic with $T(0) = T'(0) = 0$. If T is not constant then there exists $c \in \mathbb{R}_+$ so that $\mathbb{B}$ has full measure but is meagre in $[-c, c]$.*

Proof. See [7, Cor. 6.28]. □

Example 3.10. The orbit of x_0 under $T(x) = x - 1 + e^{-x}$ is Benford for almost all, but not all, $x_0 \in \mathbb{R}$.

As the reader will have noticed, the previous theorems address Benford behavior of orbits generated by maps that grow (at most) polynomially as $x \to \infty$ or $x \to 0$. The next result presents a simple condition guaranteeing Benford behavior in orbits under exponential functions.

Theorem 3.11. *Let $T : \mathbb{R}_+ \to \mathbb{R}_+$ be such that $x \mapsto \log T(10^x)$ is convex. If $T(10^x) > 10^x T(1)$ for all $x \in \mathbb{R}_+$ then $\mathbb{B}$ has full measure but is meagre in $\mathbb{R}_+$.*

Proof. See [7, Prop. 6.31]. □

Example 3.12. The orbit $O_T(x_0)$ of x_0 under $T(x) = e^x$ is Benford for almost all, but not all, $x_0 \in \mathbb{R}$.

While the results in this section so far pertain to systems that are rather unexciting dynamically, the following is about a classical "chaotic" system.

Theorem 3.13. *Let $T : \mathbb{R} \to \mathbb{R}$ be the tent map $T(x) = 1 - |2x - 1|$. Then $\mathbb{B} \supset \mathbb{R} \setminus [0, 1]$, and $\mathbb{B} \cap [0, 1]$ is a dense, uncountable nullset in $[0, 1]$.*

Proof. See [7, Exp. 6.52]. □

One of the most widely used numerical root-finding algorithms is *Newton's method*, and as the next theorem shows, the differences between successive Newton approximations conform to BL, and so do the differences between these approximations and the (usually) unknown root.

Theorem 3.14. *Let $\mathbb{I} \subset \mathbb{R}$ be a non-empty open interval, and assume that $f : \mathbb{I} \to \mathbb{R}$ is real-analytic but not linear, with $f(x^*) = 0$ for some $x^* \in \mathbb{I}$. Let $T : \mathbb{I} \to \mathbb{R}$ be the Newton root-finding map associated with f, that is, $T(x) = x - f(x)/f'(x)$ if $f'(x) \neq 0$ and $T(x) = x$ if $f'(x) = 0$.*

(i) *If x^* is a simple root of f then the sequences $\big(T^{n+1}(x_0) - T^n(x_0)\big)$ and $(T^n(x_0) - x^*)$ both are Benford for almost all, but not all, x_0 in a neighbourhood of x^*.*

(ii) *If x^* is a root of f with multiplicity at least 2 then the sequences $\big(T^{n+1}(x_0) - T^n(x_0)\big)$ and $(T^n(x_0) - x^*)$ both are Benford for all $x_0 \neq x^*$ in a neighbourhood of x^*.*

Proof. See [7, Thm. 6.35]. □

Example 3.15. The Newton approximations associated with $f(x) = e^x - 2$ and $g(x) = (e^x - 2)^3$ both are Benford in the sense of Theorem 3.14.

3.2. **Non-autonomous systems.**

Theorem 3.16. *For every $n \in \mathbb{N}$, let $T_n : \mathbb{R}_+ \to \mathbb{R}_+$ be given by $T_n(x) = a_n x$ for some $a_n \in \mathbb{R}_+$. If $O_T(1) = (a_1, a_1a_2, a_1a_2a_3, \ldots)$ is Benford then $\mathbb{B} = \mathbb{R}_+$; otherwise $\mathbb{B} = \varnothing$.*

Proof. See [7, Thm. 6.40]. □

Example 3.17. Let $T_n(x) = (2 + 1/n)x$ for each $n \in \mathbb{N}$. By Theorem 3.16, $O_T(x_0)$ is Benford for all $x_0 \in \mathbb{R}_+$ since, as can be shown, the sequence $O_T(1) = (3, 3 \cdot 5/2!, 3 \cdot 5 \cdot 7/3!, \ldots)$ is Benford.

Theorem 3.18. *For every $n \in \mathbb{N}$, let $T_n : \mathbb{R}_+ \to \mathbb{R}_+$ be given by $T_n(x) = a_n x^{b_n}$ for some $a_n \in \mathbb{R}_+, b_n \in \mathbb{R} \setminus \{0\}$. If $\liminf_{n\to\infty} |b_n| > 1$ then $\mathbb{B}$ has full measure but is meagre in $\mathbb{R}_+$.*

Proof. See [7, Thm. 6.46]. □

Example 3.19. Let $T_n(x) = 2^n x^2$ for every $n \in \mathbb{N}$. By Theorem 3.18, $O_T(x_0)$ is Benford for almost all, but not all, $x_0 \in \mathbb{R}$.

Theorem 3.20. *For every $n \in \mathbb{N}$, let $T_n : \mathbb{R}_+ \to \mathbb{R}_+$ be such that $x \mapsto \log T_n(10^x)$ is convex on $\mathbb{R}_+$, and $T_n(10^x) \geq 10^{b_n x} T_n(1)$ for some $b_n \in \mathbb{R}_+$ and all $n \in \mathbb{N}$, $x \in \mathbb{R}_+$. If $\liminf_{n\to\infty} b_n > 1$ then $\mathbb{B}$ has full measure but is meagre in $[c, \infty[$ for some $c \in \mathbb{R}_+$.*

Proof. See [1, Thm. 3.3] and [7, Thm. 6.49]. □

Example 3.21. Let $T_n(x) = x^{2n} + 1$ for every $n \in \mathbb{N}$. By Theorem 3.20, $O_T(x_0)$ is Benford for almost all, but not all, $x_0 \in \mathbb{R}$.

4. One-dimensional continuous-time dynamical systems

This short section complements the results of the previous section by considering two continuous-time scenarios. Given any continuously differentiable function $T : \mathbb{R} \to \mathbb{R}$ and any point $x_0 \in \mathbb{R}$, assume that the initial value problem

$$\dot{x} := \frac{\mathrm{d}x}{\mathrm{d}t} = T(x), \quad x(0) = x_0,$$

has a unique solution $x = x(t)$ for all $t \geq 0$; to emphasize the analogy to the previous section, denote this solution by $O_T(x_0)$.

Theorem 4.1. *Let $T : \mathbb{R} \to \mathbb{R}$ be C^1 with $T(x) > 0$ for all $x \in \mathbb{R}_+$. If $T(x)/x$ either converges to a (finite) positive limit as $x \to \infty$ or is non-decreasing on $]c, \infty[$ for some $c \geq 0$, then $\mathbb{B} \supset \mathbb{R}_+$.*

Proof. See [7, Thm. 6.58]. □

Example 4.2. Clearly the solution $x(t) = x_0 e^t$ of the initial value problem $\dot{x} = x$, $x(0) = x_0$ is Benford for all $x_0 \neq 0$. Similarly, by Theorem 4.1, the solution of

$$\dot{x} = \sqrt{x^2 + 1}, \quad x(0) = x_0$$

is Benford for every $x_0 \in \mathbb{R}$.

Theorem 4.3. *Let $T : \mathbb{R} \to \mathbb{R}$ be C^1 with $T(0) = 0$. If $T'(0) < 0$ then $\mathbb{B} \supset [-c, c] \setminus \{0\}$ for some $c \in \mathbb{R}_+$.*

Proof. See [7, Thm. 6.60]. □

Example 4.4. For every $x_0 \neq 0$, the solution of

$$\dot{x} = -\frac{x}{x^2 + 1}, \quad x(0) = x_0$$

eventually gets close to 0, and hence is Benford by Theorem 4.3.

5. Multi-dimensional and stochastic systems

This section presents three scenarios where the phase space $\mathbb{X}$ is not one-dimensional, and may in fact be infinite-dimensional. Though they are stated here in their most elementary and concrete forms, the results in all three scenarios can also be understood more abstractly as pertaining to the dynamics of $T : \mathbb{X} \to \mathbb{X}$ given by $T(x) = x \bullet x_0$ for all $x \in \mathbb{X}$, where $x_0 \in \mathbb{X}$ is fixed and $\bullet$ denotes an appropriate binary operation on $\mathbb{X}$, such as, e.g., the multiplication of square matrices or the convolution of distribution functions. As the reader will no doubt appreciate, these results typically require for their proof more sophisticated tools than the results in previous sections; in keeping with the informal nature of this survey, however, proofs are again provided solely via precise references.

For a first scenario, fix $d \in \mathbb{N}$ and let $\mathbb{X} = \mathbb{R}^{d\times d}$, the set (in fact, $\mathbb{R}$-algebra) of all real $d \times d$-matrices. Given $A \in \mathbb{X}$, denote by $\rho(A)$ and $|A|$ the spectral radius and the (spectral or Euclidean) norm of A, respectively, and by $[A]_{k\ell}$ the entry of A in the kth row and ℓth column, with $k, \ell \in \{1, 2, \dots, d\}$; as ususal, $A > 0$ means that $[A]_{k\ell} > 0$ for all k, ℓ.

Theorem 5.1. *Let $A \in \mathbb{X}$, and assume that $A^m > 0$ for some $m \in \mathbb{N}$. Then the following are equivalent:*

(i) $\log \rho(A)$ *is irrational;*
(ii) *the sequence $(|A^n|)$ is Benford;*
(iii) *for every $k, \ell \in \{1, 2, \dots, d\}$ the sequence $([A^n]_{k\ell})$ is Benford.*

Proof. See [4, Thm. 3.2] and [7, Thm. 7.11]. □

Example 5.2. Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. Since $A^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} > 0$ and $\log \rho(A) = \log \frac{1}{2}(1 + \sqrt{5})$ is irrational, the sequence $(|A^n|)$ is Benford, and so is $([A^n]_{k\ell})$ for every $k, \ell \in \{1, 2\}$.

A simple corollary of Theorem 5.1 pertains to linear recursions.

Theorem 5.3. *Let $d \in \mathbb{N}$, and $a_1, a_2, \dots, a_d, x_1, x_2, \dots, x_d \in \mathbb{R}_+$. Assume that (x_n) is the (unique) solution of*

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \ldots + a_d x_{n-d} \qquad \forall n \geq d + 1.$$

Then (x_n) is Benford if and only if $\log \zeta$ is irrational, where $z = \zeta$ is the (unique) root of $z^d = a_1 z^{d-1} + a_2 z^{d-2} + \ldots + a_d$ with the largest real part.

Proof. See [7, Thm. 7.39]. □

Example 5.4. For the Fibonacci recursion $x_n = x_{n-1} + x_{n-2}$, $n \geq 3$, the root of $z^2 = z + 1$ with the largest real part is $\zeta = \frac{1}{2}(1 + \sqrt{5})$. As in Example 5.2, therefore, (x_n) is Benford unless $x_1 = x_2 = 0$. In particular, the sequence (F_n) of Fibonacci numbers is Benford, and so is the sequence of Lucas numbers. As has been reported repeatedly, the distribution of D_1 for $F_1, F_2, \dots, F_N$ matches the right-hand side of (2.1) remarkably well even for not-so-large N; see, e.g., [10, 12] and Figure B.

A prominent class of matrices amenable to Theorem 5.1 are irreducible and aperiodic stochastic matrices, often encountered in homogeneous finite-state Markov

D_1	1	2	3	4	5	6	7	8	9
$N = 10^2$	30 *30*	18 *18*	13 *12*	9 *10*	8 *8*	6 *7*	5 *6*	7 *5*	4 *4*
$N = 10^3$	301 *301*	177 *176*	125 *125*	96 *97*	80 *79*	67 *67*	56 *58*	53 *51*	45 *46*
$N = 10^4$	3011 *3010*	1762 *1761*	1250 *1249*	968 *969*	792 *792*	668 *669*	580 *580*	513 *512*	456 *458*

FIGURE B. Occurrences of D_1 for the Fibonacci numbers $F_1, F_2, \ldots, F_N$, for three different N. For comparison, the corresponding best approximation of the "Benford vector" $N\big(\log 2, \log \frac{3}{2}, \ldots, \log \frac{10}{9}\big) \in \mathbb{R}^9$ by $(N_1, N_2, \ldots, N_9) \in \mathbb{N}^9$ with $\sum_{j=1}^{9} N_j = N$ is shown in *italics*.

chains. As the following result shows, conformance to BL is prevalent in this context.

Theorem 5.5. *Let $d \in \mathbb{N}\setminus\{1\}$, and assume that the random stochastic $d\times d$-matrix P has an absolutely continuous distribution (w.r.t. Lebesgue measure). Then, with probability one, P is irreducible and aperiodic. Moreover, with $P^* = \lim_{n\to\infty} P^n$ and probability one, the sequences $([P^{n+1}-P^n]_{k\ell})$ and $([P^n-P^*]_{k\ell})$ both are Benford or eventually zero for every $k, \ell \in \{1, 2, \ldots, d\}$.*

Proof. See [8, Thm. 12]. □

Example 5.6. Let X, Y be independent $U(0,1)$-random variables. By Theorem 5.5, for the two-state Markov chain with one-step transition matrix

$$P = \begin{bmatrix} 1-X & X \\ Y & 1-Y \end{bmatrix} \in \mathbb{R}^{2\times 2},$$

with probability one, P is irreducible and aperiodic,

$$P^* = \lim_{n\to\infty} P^n = \frac{1}{X+Y}\begin{bmatrix} Y & X \\ Y & X \end{bmatrix},$$

and the sequences $([P^{n+1} - P^n]_{k\ell})$ and $([P^n - P^*]_{k\ell})$ both are Benford for every $k, \ell \in \{1, 2\}$.

Without the assumption, in Theorem 5.1, that $A \in \mathbb{X}$ satisfy $A^m > 0$ for some $m \in \mathbb{N}$, the Benford properties of the sequences $(|A^n|)$ and $([A^n]_{k\ell})$ are considerably more delicate. For a comprehensive analysis of the latter, as well as for studies on BL in more general multi-dimensional dynamical systems, the reader is referred, e.g., to [4, 5].

For a second scenario, let $\mathbb{X}$ be the set of all (real-valued) random variables (on a fixed probability space). Clearly, if $X, Y \in \mathbb{X}$ then also $XY \in \mathbb{X}$, and hence in particular $X^n \in \mathbb{X}$ for every $n \in \mathbb{N}$.

Theorem 5.7. *Let X be a (real-valued) random variable. Then:*

(i) *(X^n) converges in distribution to BL if and only if $\mathbb{E}e^{2\pi i n \log|X|} \to 0$ as $n \to \infty$;*

(ii) *(X^n) is Benford with probability one if and only if $\log|X|$ is irrational with probability one.*

Proof. See [7, Thm. 8.5 and 8.6]. □

Corollary 5.8. *If the random variable X has a density (w.r.t. Lebesgue measure) then (X^n) converges in distribution to BL and is Benford with probability one.*

Example 5.9. If the random variable X is uniformly distributed (on some non-degenerate interval $\mathbb{I} \subset \mathbb{R}$), exponentially distributed, or normally distributed (with variance $\sigma^2 > 0$) then the sequence $(X, X^2, X^3, \ldots)$ converges in distribution to BL and is Benford with probability one. By contrast, if $X = 10^Y$ and Y is uniformly distributed on the classical Cantor middle thirds set, then (X^n) does not converge in distribution to BL, even though X is continuous, i.e., the event $\{X = x\}$ has probability zero for every $x \in \mathbb{R}$.

For a third and final scenario, let X again be a (real-valued) random variable, and consider the iid sequence (X_n) with $X_1 = X$. Clearly, if X, Y are independent then $(X_n Y)$ is an iid sequence as well. Say that X is **non-discrete** if the event $\{X \in C\}$ has probability < 1 for every countable set $C \subset \mathbb{R}$.

Theorem 5.10. *Let $X_1, X_2, X_3, \ldots$ be iid random variables. If X_1 is non-discrete then the sequence $(X_1, X_1X_2, X_1X_2X_3, \ldots)$ converges in distribution to BL and is Benford with probability one.*

Proof. See [7, Cor. 8.19]; to appreciate how the conclusion ultimately hinges on an ergodic theorem, see also [6] and [7, Thm. 4.24]. □

For a continuous-time variant of Theorem 5.10, recall that a *Brownian motion* (or *Wiener process*) is a continuous stochastic process on $\mathbb{R}$ with independent and stationary increments that are normally distributed. Widely used as a model, e.g., for stock prices, a *geometric Brownian motion* (or *Black–Scholes process*) is a stochastic dynamical system on $\mathbb{R}_+$ whose logarithm is a Brownian motion [13].

Theorem 5.11. *Let $(X_t)_{t\geq 0}$ be a geometric Brownian motion. Then $(X_t)_{t\geq 0}$ converges weakly to BL and is Benford with probability one.*

Proof. See [6, Ex. 3.3], [17, Thms. 2.4 and 3.11] as well as [7, Cor. 8.19]. □

6. A new theorem

This final section presents one new result that may help to further illustrate the ubiquity of BL in dynamical systems. As in all other sections of this survey, the result is presented in a simple concrete form rather than in the greatest possible generality. Unlike in other sections, however, a detailed proof is given here. This

is not only because no proof is available yet in the literature but also, and perhaps more importantly, because from the proof the reader will get a clear idea of some of the tools naturally employed for the study of BL in dynamical systems.

To motivate the result, recall from Section 5 that one of the simplest, most prominent examples of a Benford sequence is any non-trivial solution of the linear two-step recursion

$$x_n = x_{n-1} + x_{n-2} \qquad \forall n \geq 3\,. \tag{6.1}$$

With a view towards applications, e.g., in economics [3], it is natural to consider non-linear variants of (6.1) such as, for instance,

$$x_n = a_1 x_{n-1}^{b_1} + a_2 x_{n-2}^{b_2} \qquad \forall n \geq 3\,, \tag{6.2}$$

with parameters $a_1, a_2 \in \mathbb{R}_+$ and $b_1, b_2 > 1$. The dynamics on $\mathbb{X} = \mathbb{R}_+^2 := \mathbb{R}_+ \times \mathbb{R}_+$ of recurrences like (6.2) have been studied systematically in [3], and will be recalled below to the extent necessary for the present article. To concisely state the main result, and in analogy to previous sections, let

$$\mathbb{B} = \big\{(x_1, x_2) \in \mathbb{X} : (x_n) \text{ is Benford}\big\}\,,$$

where (x_n) denotes the unique solution of (6.2), given (x_1, x_2). As it turns out, this solution is Benford for almost all, but not all, choices of (x_1, x_2).

Theorem 6.1. *Let $a_1, a_2 \in \mathbb{R}_+$ and $b_1, b_2 > 1$. Then $\mathbb{B}$ has full measure but is meagre in $\mathbb{X}$.*

Though straightforward to state, Theorem 6.1 requires for its proof a few preparations, aimed primarily at connecting it to dynamical systems in a meaningful way. To this end, consider the map $T : \mathbb{X} \to \mathbb{X}$ given by

$$T(u, v) = (v, a_2 u^{b_2} + a_1 v^{b_1}) \qquad \forall (u, v) \in \mathbb{X}\,.$$

In fact, T is a diffeomorphism from $\mathbb{X}$ onto $T(\mathbb{X}) \subset \mathbb{X}$, and clearly

$$(x_{n-1}, x_n) = T(x_{n-2}, x_{n-1}) = T^{n-2}(x_1, x_2) \qquad \forall n \geq 3\,.$$

While it is common practice to thus associate a two-dimensional map with a two-step recursion like (6.2), the reader may want to appreciate that such an approach does not per se guarantee a satisfactory understanding of the latter. This merely reflects the well-documented fact that non-linear two- or multi-step recursions may exhibit an immense variety of dynamical behaviors for which no comprehensive theory exists at present; see, e.g., [3] and the extensive literature cited there. What makes an approach to Theorem 6.1 utilizing T viable, then, is that the dynamics of (6.2) actually is quite simple; see [3, Thm. 2.2 and Exp. 4.2].

Proposition 6.2. *Let $a_1, a_2 \in \mathbb{R}_+$ and $b_1, b_2 > 1$. Then there exists a unique non-empty, open and bounded set $\mathbb{A}_0 \subset \mathbb{X}$ with the following properties:*

- (i) *$\mathbb{A}_0$ is convex, and $\mathbb{A}_0 \supset \{x \in \mathbb{X} : x_1^2 + x_2^2 < r^2\}$ for some $r > 0$;*
- (ii) *the boundary $\partial\mathbb{A}_0$ of $\mathbb{A}_0$ is smooth, and if $(x_1, x_2) \in \partial\mathbb{A}_0$ then (x_n) is asymptotically 2-periodic, i.e., (x_{2n-1}) and (x_{2n}) both converge;*
- (iii) *if $(x_1, x_2) \in \mathbb{A}_0$ then $\lim_{n\to\infty} x_n = 0$;*
- (iv) *if $(x_1, x_2) \in \mathbb{X} \setminus (\mathbb{A}_0 \cup \partial\mathbb{A}_0)$ then $\lim_{n\to\infty} x_n = \infty$.*

Informally put, Proposition 6.2 asserts that the dynamics of (6.2) neatly partitions the quadrant $\mathbb{X}$ into precisely three parts: the domains of attraction of 0 and ∞, given by the non-empty open sets $\mathbb{A}_0$ and $\mathbb{A}_\infty := \mathbb{X} \setminus (\mathbb{A}_0 \cup \partial\mathbb{A}_0)$, respectively, as well as their common boundary $\partial\mathbb{A}_0 = \partial\mathbb{A}_\infty$ which is a smooth curve. Regarding assertion (ii), note that if $(x_1, x_2) \in \partial\mathbb{A}_0$ and $(x_{2n-1}, x_{2n}) \to (p, q) \in \partial\mathbb{A}_0$ then $T(p, q) = (q, p)$ and $T^2(p, q) = (p, q)$, with $p = q$ and $p \neq q$ both being possibilities, depending on the values of the parameters a_1, a_2, b_1, b_2; see also Figure C and Example 6.5 below.

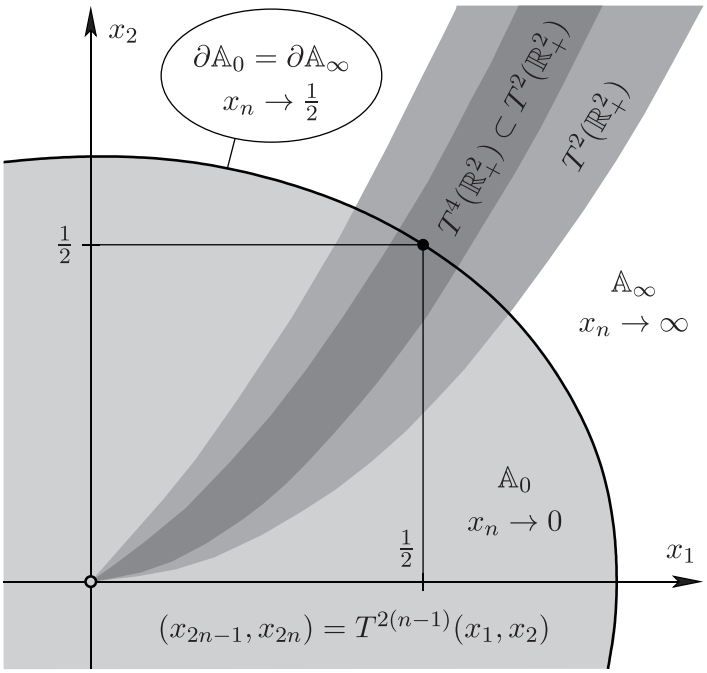


FIGURE C. For the "quadratic Fibonacci" recursion of Example 6.5, that is, for $x_n = x_{n-1}^2 + x_{n-2}^2$, $n \geq 3$, the set $\mathbb{A}_0 = \{(x_1, x_2) \in \mathbb{R}^2 : \lim_{n\to\infty} x_n = 0\}$ is non-empty, open, bounded and convex. Its boundary $\partial\mathbb{A}_0$ is a smooth oval containing the (unique) fixed point $(\frac{1}{2}, \frac{1}{2})$ of T. If $(x_1, x_2) \in \partial\mathbb{A}_0$ then $\lim_{n\to\infty} x_n = \frac{1}{2}$, and clearly (x_n) is *not* Benford in this case. By contrast, Corollary 6.4 shows that (x_n) is Benford for almost all, but not all, $(x_1, x_2) \in \mathbb{A}_0 \cup \mathbb{A}_\infty$. The figure only displays slightly more than one quarter of $\mathbb{A}_0$ (light grey) and $\mathbb{A}_\infty$ (white), in light of their obvious symmetries; note, however, that neither set is symmetric w.r.t. the line $x_1 = x_2$.

In addition to Proposition 6.2, the proof of Theorem 6.1 makes use of the following elementary calculus fact which may also be viewed as a special case of a *shadowing lemma*; see, e.g., [7, Lem. 6.24].

Proposition 6.3. *Let (y_n) be a sequence in $\mathbb{R}$, and $b > 1$. Assume that $\sup_{n\geq 2} |y_n - by_{n-1}| < \infty$. Then $\sup_{n\in\mathbb{N}} |y_n - b^n y| < \infty$ with $y \in \mathbb{R}$ if and only if*

$$y = \widehat{y} := \frac{y_1}{b} + \sum\nolimits_{k=2}^{\infty} \frac{y_k - by_{k-1}}{b^k}\,.$$

Moreover, if $\lim_{n\to\infty}(y_n - by_{n-1}) = c$ then $\lim_{n\to\infty}(y_n - b^n\widehat{y}) = -c/(b-1)$.

Proof of Theorem 6.1. By Proposition 6.2, except for the nowhere dense nullset $\partial\mathbb{A}_0$, the two non-empty open sets $\mathbb{A}_0$ and $\mathbb{A}_\infty$ make up all of $\mathbb{X}$. Thus it suffices to show that $\mathbb{B}$ has full measure but is meagre in each of these two sets. A detailed argument is presented here for $(x_1, x_2) \in \mathbb{A}_\infty$, that is, for the case of $x_n \to \infty$. The argument for $(x_1, x_2) \in \mathbb{A}_0$ is completely analogous and will be commented on only briefly towards the very end of the proof, with all details being left to the interested reader.

Assume from now on that $(x_1, x_2) \in \mathbb{A}_\infty$. Recall from Theorem 2.2 that replacing (x_n) by (ax_n) with any $a \in \mathbb{R}_+$ does not affect the Benford property. When rescaled in this way, (6.2) reads

$$x_n = a_1 a^{b_1-1} x_{n-1}^{b_1} + a_2 a^{b_2-1} x_{n-2}^{b_2} \qquad \forall n \geq 3\,.$$

Thus, taking $a = a_1^{-1/(b_1-1)}$ or $a = a_2^{-1/(b_2-1)}$, it can be assumed that $a_1 = 1$ or $a_2 = 1$, respectively. In either case, clearly $(x_{n-1}, x_n) \in]1,\infty[^2$ for all sufficiently large $n \in \mathbb{N}$, and hence $\mathbb{A}_\infty = \bigcup_{n\in\mathbb{N}} T^{-n}(]1,\infty[^2)$. Since T is a diffeomorphism, it suffices to show that $\mathbb{B}$ has full measure but is meagre in $]1,\infty[^2$. This will now be done via a careful analysis of the sequence (y_n), where $y_n = \log x_n$ for every $n \in \mathbb{N}$. By Theorem 2.2, (x_n) is Benford if and only if (y_n) is u.d. mod 1. Note that $y_n \to \infty$ as well, and

$$\mathbb{U} := \big\{(y_1, y_2) \in \mathbb{R}_+^2 : (y_n) \text{ is u.d. mod } 1\big\} \subset \mathbb{R}_+^2$$

is a Borel set. So, what really will be shown is that $\mathbb{U}$ has full measure but is meagre in $\mathbb{R}_+^2$. For the convenience of the reader, three disjoint cases regarding the possible values of the parameters $b_1, b_2 > 1$ are going to be considered separately.

Case I: $b_1^2 > b_2$. As will become clear shortly, in this case the dynamics of (6.2) is dominated by the term $a_1 x_{n-1}^{b_1}$. Assume w.l.o.g. that $a_1 = 1$, i.e., consider

$$x_n = x_{n-1}^{b_1} + a_2 x_{n-2}^{b_2} \qquad \forall n \geq 3\,. \tag{6.3}$$

Note that $x_n > 1$ for every n, provided that $x_1, x_2 > 1$, and then $x_n \to \infty$. In this situation, deduce from $x_n > x_{n-1}^{b_1}$ and $b_1^2 > b_2$ that

$$\frac{x_n^{b_1}}{x_{n-1}^{b_2}} > x_{n-1}^{b_1^2-b_2} \to \infty \quad \text{as } n \to \infty\,.$$

With the abbreviation

$$\delta_n = \delta_n(y_1, y_2) := b_2 y_{n-1} - b_1 y_n \qquad \forall n \geq 2\,,$$

therefore, $\delta_n \leq -(b_1^2 - b_2)y_{n-1} < 0$ for $n \geq 3$, and $\delta_n \to -\infty$ whenever $y_1, y_2 > 0$. Note that (y_n) is given recursively as

$$y_n = b_1 y_{n-1} + \log\big(1 + a_2 10^{\delta_{n-1}}\big) \qquad \forall n \geq 3\,.$$

By Proposition 6.3, $\lim_{n\to\infty}\big(y_n - b_1^{n-2}h(y)\big) = 0$, where for every $y \in \mathbb{R}^2_+$,

$$h(y) = h(y_1, y_2) = y_2 + \sum\nolimits_{k=1}^{\infty} b_1^{-k} \log\big(1 + a_2 10^{\delta_{k+1}}\big)\,. \tag{6.4}$$

The function $h : \mathbb{R}^2_+ \to \mathbb{R}$ thus defined is continuous and positive. Moreover, (y_n) is u.d. mod 1 if and only if $\big(b_1^{n-2}h(y)\big)$ is. As will now be shown, the latter is the case precisely for all $y \in \mathbb{R}^2_+$ belonging to a meagre set of full measure. Since that set equals $\mathbb{U}$ by definition, this will establish the desired properties of $\mathbb{U}$.

To this end, note first that $\lim_{n\to\infty} \delta_n b_1^{-n} = -(b_1^2 - b_2)b_1^{-3}h(y) < 0$, so $\delta_n \to -\infty$ exponentially fast as $n \to \infty$. Also,

$$\delta_n = \delta_n(y) = b_2 y_{n-1} - b_1 y_n = b_2 f^{n-2}(y)_1 - b_1 f^{n-2}(y)_2 \qquad \forall n \geq 2\,,$$

with the smooth map $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\begin{aligned} f(u,v) &= \big(\log T(10^u, 10^v)_1, \log T(10^u, 10^v)_2\big) \\ &= \big(v, b_1 v + \log(1 + a_2 10^{b_2 u - b_1 v})\big)\,. \end{aligned}$$

Termwise partial differentiation of (6.4) yields, at least formally,

$$\frac{\partial h}{\partial y_2} = \frac{1}{1 + a_2 10^{\delta_2}} + \sum\nolimits_{k=1}^{\infty} b_1^{-(k+1)} \frac{a_2 10^{\delta_{k+2}}}{1 + a_2 10^{\delta_{k+2}}} \cdot \frac{\partial \delta_{k+2}}{\partial y_2}\,, \tag{6.5}$$

where, for every $k \geq 1$,

$$\begin{aligned} \frac{\partial \delta_{k+2}}{\partial y_2} &= b_2 \frac{\partial}{\partial y_2} f^k(y)_1 - b_1 \frac{\partial}{\partial y_2} f^k(y)_2 \\ &= b_2 \left(\prod_{\ell=0}^{k-1} Df \circ f^\ell(y) \right)_{21} - b_1 \left(\prod_{\ell=0}^{k-1} Df \circ f^\ell(y) \right)_{22}\,. \end{aligned}$$

Together with

$$Df \circ f^\ell(y) = \begin{bmatrix} 0 & \dfrac{a_2 b_2 10^{\delta_{\ell+2}}}{1 + a_2 10^{\delta_{\ell+2}}} \\ 1 & \dfrac{b_1}{1 + a_2 10^{\delta_{\ell+2}}} \end{bmatrix} \qquad \forall \ell \geq 0\,,$$

and consequently $|Df \circ f^\ell| \leq \sqrt{1 + b_1^2 + b_2^2}$, this shows that the series on the right in (6.5) does converge, in fact absolutely as well as uniformly in $y \in \mathbb{R}^2_+$. Hence $\partial h/\partial y_2$ is continuous on $\mathbb{R}^2_+$, and so is $\partial h/\partial y_1$, by a virtually identical argument. Thus h is C^1 on $\mathbb{R}^2_+$. Moreover, deduce from (6.5) that by taking $c \in \mathbb{R}_+$ sufficiently large, one can guarantee that $\partial h/\partial y_2(y) > \frac{1}{2}$ whenever $\delta_2 < -c$ or, equivalently, $b_1 y_2 - b_2 y_1 > c$. (Note that c may depend on a_2, b_1, b_2, but is independent of y_1, y_2; also, the value $\frac{1}{2}$ for the lower bound on $\partial h/\partial y_2$ is arbitrary and could be replaced by any positive number < 1.) For convenience, let

$$\mathbb{V} = \big\{y \in \mathbb{R}^2_+ : b_1 y_2 - b_2 y_1 > c\big\} = \big\{y \in \mathbb{R}^2_+ : \delta_2 < -c\big\}\,.$$

Now, recall that $\lim_{n\to\infty}\big(y_n - b_1^{n-2}h(y)\big) = 0$, and deduce from Proposition 3.6 that $(b_1^{n-2}z)$ is u.d. mod 1 for almost all $z \in \mathbb{R}$, but the set $\big\{z \in \mathbb{R} : (b_1^{n-2}z) \text{ is u.d. mod } 1\big\}$

is meagre (in $\mathbb{R}$). Since $\partial h/\partial y_2 > \frac{1}{2}$ on $\mathbb{V}$, this implies that for every $y_1 \in \mathbb{R}_+$ the y_1-sections of $\mathbb{V} \setminus \mathbb{U}$ and $\mathbb{V} \cap \mathbb{U}$, i.e., the sets

$$(\mathbb{V} \setminus \mathbb{U})_{y_1} = \big\{y_2 \in \mathbb{R}_+ : (y_1, y_2) \in \mathbb{V},\, (y_n) \text{ is } \textit{not} \text{ u.d. mod } 1\big\}\,,$$
$$(\mathbb{V} \cap \mathbb{U})_{y_1} = \big\{y_2 \in \mathbb{R}_+ : (y_1, y_2) \in \mathbb{V},\, (y_n) \text{ is u.d. mod } 1\big\}\,,$$

are a nullset and meagre in $](c+b_2y_1)/b_1, \infty[$, respectively. This is because $h(y_1, \cdot)$ is a local diffeomorphism on that open interval, and hence maps nullsets and meagre sets to nullsets and meagre sets, respectively. By Fubini's theorem, $\mathbb{V} \setminus \mathbb{U}$ is a nullset as well, and by the Kuratowski–Ulam theorem [15, Thm. 15.4], $\mathbb{V} \cap \mathbb{U}$ is meagre in $\mathbb{V}$. Recall that $\delta_n \to -\infty$ as $n \to \infty$ for every $y \in \mathbb{R}^2_+$, so $f^n(y) \in \mathbb{V}$ for all sufficiently large n. In other words, $\mathbb{R}^2_+ \subset \bigcup_{n\in\mathbb{N}} f^{-n}(\mathbb{V})$. On the one hand, observe that $f(\mathbb{R}^2_+ \setminus \mathbb{U}) \subset \mathbb{R}^2_+ \setminus \mathbb{U}$, and hence

$$\mathbb{R}^2_+ \setminus \mathbb{U} = \bigcup_{n\in\mathbb{N}} \big((\mathbb{R}^2_+ \setminus \mathbb{U}) \cap f^{-n}(\mathbb{V})\big) \subset \bigcup_{n\in\mathbb{N}} f^{-n}\big((\mathbb{R}^2_+ \setminus \mathbb{U}) \cap \mathbb{V}\big) = \bigcup_{n\in\mathbb{N}} f^{-n}(\mathbb{V} \setminus \mathbb{U})\,.$$

On the other hand, letting $\mathbb{V} \cap \mathbb{U} \subset \bigcup_{m\in\mathbb{N}} V_m$, where $V_m \subset \mathbb{V}$ is nowhere dense for each $m \in \mathbb{N}$, observe that $f(\mathbb{U}) \subset \mathbb{U}$, and hence

$$\mathbb{U} = \bigcup_{n\in\mathbb{N}} \big(\mathbb{U} \cap f^{-n}(\mathbb{V})\big) \subset \bigcup_{n\in\mathbb{N}} f^{-n}(\mathbb{V} \cap \mathbb{U}) \subset \bigcup_{m,n\in\mathbb{N}} f^{-n}(V_m)\,.$$

Since f is a diffeomorphism of $\mathbb{R}^2_+$ onto $f(\mathbb{R}^2_+)$, it follows that the set $\mathbb{U}$ has full measure but is meagre in $\mathbb{R}^2_+$. Consequently, $\mathbb{B} = \big\{(10^{y_1}, 10^{y_2}) : y \in \mathbb{U}\big\}$ has full measure but is meagre in $]1, \infty[^2$ as well. As explained earlier, this proves the assertion of the theorem in case $(x_1, x_2) \in \mathbb{A}_\infty$ and $b_1^2 > b_2$.

Case II: $b_1^2 = b_2$. In this case, both terms on the right in (6.2) contribute equally to the dynamics. As in Case I, assume w.l.o.g. that $a_1 = 1$. Rewriting (6.3) in terms of the positive quantity $r_n := x_n/x_{n-1}^{b_1}$, $n \geq 2$, yields

$$r_n = \frac{x_n}{x_{n-1}^{b_1}} = 1 + a_2 \left(\frac{x_{n-2}^{b_1}}{x_{n-1}}\right)^{b_1} = 1 + \frac{a_2}{r_{n-1}^{b_1}} \qquad \forall n \geq 3\,.$$

Thus $r_n = R(r_{n-1}) = R^{n-2}(r_2)$, with the smooth map $R : \mathbb{R}_+ \to \mathbb{R}_+$ given by $R(r) = 1 + a_2 r^{-b_1}$. Let $\Phi : \mathbb{X} \to \mathbb{X}$ be the diffeomorphism with

$$\Phi(u, v) = \big((v/u)^{1/b_1}, v\big) \qquad \forall x \in \mathbb{X}\,.$$

By Theorem 3.18, with $a_n = R^{n-2}(r)$ and $b_n \equiv b_1$, for every $r \in \mathbb{R}_+$ the r-section of $\Phi^{-1}(\mathbb{X} \setminus \mathbb{B})$, i.e., the set

$$\begin{aligned}\big(\Phi^{-1}(\mathbb{X} \setminus \mathbb{B})\big)_r &= \big\{x_2 \in \mathbb{R}_+ : \Phi(r, x_2) \in \mathbb{X} \setminus \mathbb{B}\big\} \\ &= \big\{x_2 \in \mathbb{R}_+ : x_1 = (x_2/r)^{1/b_1},\, (x_n) \text{ is } \textit{not} \text{ Benford}\big\}\,,\end{aligned}$$

is a nullset, whereas

$$\Phi^{-1}(\mathbb{B})_r = \big\{x_2 \in \mathbb{R}_+ : x_1 = (x_2/r)^{1/b_1},\, (x_n) \text{ is Benford}\big\}$$

is meagre (in $\mathbb{R}_+$). Thus $\Phi^{-1}(\mathbb{X} \setminus \mathbb{B})$ is a nullset, whereas $\Phi^{-1}(\mathbb{B})$ is meagre. Since Φ is a diffeomorphism, $\mathbb{B}$ has full measure but is meagre in $\mathbb{X}$.

Note that unlike in Case I above and Case III below, this argument does not require separate analyses for $\mathbb{A}_\infty$ and $\mathbb{A}_0$, nor does it involve taking logarithms. There is a dynamical reason for this relative simplicity: Since

$$\Phi^{-1} \circ T \circ \Phi(u,v) = (R(u), R(u)v^{b_1}) \qquad \forall (u,v) \in \mathbb{X}\,,$$

the two-dimensional map T is, up to the diffeomorphism Φ, a *skew-product* over the (decreasing) one-dimensional map R.

Case III: $b_1^2 < b_2$. In this case, the dynamics of (6.2) is dominated by the term $a_2 x_{n-2}^{b_2}$. Assume w.l.o.g. that $a_2 = 1$, i.e., consider

$$x_n = a_1 x_{n-1}^{b_1} + x_{n-2}^{b_2} \qquad \forall n \geq 3\,. \tag{6.6}$$

Using the same abbreviation $\delta_n = \delta_n(y_1, y_2) = b_2 y_{n-1} - b_1 y_n$ for $n \geq 2$ as in Case I, note that

$$y_n = b_2 y_{n-2} + \log\bigl(1 + a_1 10^{-\delta_{n-1}}\bigr) \qquad \forall n \geq 3\,. \tag{6.7}$$

With the goal of applying Proposition 6.3 in a similar way as in Case I, it will first be shown that, given any $y \in \mathbb{R}_+^2$,

$$(\delta_{2n-1}),\ (\delta_{2n}) \text{ both converge in } \mathbb{R} \cup \{\infty\}\,. \tag{6.8}$$

As a first step towards establishing (6.8), deduce from (6.7) that

$$\begin{aligned} \delta_{2n+1} = {} & b_2 \log\bigl(a_1 + 10^{\delta_{2n-1}}\bigr) - \\ & - b_1 \log\bigl(1 + a_1 10^{-(b_2 - b_1^2) y_{2n-1}} (a_1 + 10^{\delta_{2n-1}})^{b_1}\bigr) \qquad \forall n \geq 2\,. \end{aligned} \tag{6.9}$$

Motivated by (6.9), for every $\varepsilon \geq 0$ consider the smooth map $R_\varepsilon : \mathbb{R} \to \mathbb{R}$ given by

$$R_\varepsilon(r) = b_2 \log(a_1 + 10^r) - b_1 \log(1 + \varepsilon(a_1 + 10^r)^{b_1})\,.$$

It is readily checked that $0 < R_\varepsilon'(r) < b_2$ for all $r \in \mathbb{R}$, and

$$R_\varepsilon(-\infty) = b_2 \log a_1 - b_1 \log(1 + \varepsilon a_1^{b_1})\,, \quad R_\varepsilon'(-\infty) = 0\,.$$

Since R_ε is strictly increasing, the sequence $\bigl(R_\varepsilon^n(r)\bigr)$ is monotone for every $r \in \mathbb{R}$, and hence either converges to a fixed point of R_ε, or else $\lim_{n\to\infty} R_\varepsilon^n(r) = \infty$. (Note that $\lim_{n\to\infty} R_\varepsilon^n(r) = -\infty$ is impossible because $R_\varepsilon(-\infty) > -\infty$.) Moreover, the map R_0 is strictly convex, with $R_0'(\infty) = b_2 > 1$, and so has at most two fixed points. Given any $c \in \mathbb{R}_+$, note that $R_\varepsilon \leq R_0$ and $R_\varepsilon \uparrow R_0$ uniformly on $]-\infty, c]$ as $\varepsilon \downarrow 0$. Thus, if R_0 does have fixed points and if c is larger than each of these fixed points, then R_ε has fixed points in $]-\infty, c]$ also, and with $r_\varepsilon^-, r_\varepsilon^+$ denoting the smallest and (not necessarily different) largest fixed point of R_ε in $]-\infty, c]$, respectively, the limits $\lim_{\varepsilon\downarrow 0} r_\varepsilon^-$, $\lim_{\varepsilon\downarrow 0} r_\varepsilon^+$ both exist and are fixed points of R_0. Now, recall that $b_1^2 < b_2$ and $y_n \to \infty$. Given $\varepsilon > 0$, therefore, there exists $N_\varepsilon \geq 2$ so that $0 < a_1 10^{-(b_2 - b_1^2) y_{2n-3}} < \varepsilon$ for all $n \geq N_\varepsilon$, and (6.9) yields

$$R_\varepsilon(\delta_{2n-1}) < \delta_{2n+1} < R_0(\delta_{2n-1}) \qquad \forall n \geq N_\varepsilon\,.$$

It follows that

$$R_\varepsilon^n(\delta_{2N_\varepsilon - 1}) < \delta_{2(n+N_\varepsilon)-1} < R_0^n(\delta_{2N_\varepsilon - 1}) \qquad \forall n \in \mathbb{N}\,. \tag{6.10}$$

At this point, it is convenient to consider the three possibilities regarding the fixed points of R_0 individually.

First, if R_0 has *no fixed point* then, given any $c \in \mathbb{R}_+$, take $\varepsilon > 0$ so small that R_ε has no fixed point in $]-\infty, c]$ either. In this case, letting $n \to \infty$ in (6.10) yields

$$c \leq \lim_{n\to\infty} R_\varepsilon^n(\delta_{2N_\varepsilon-1}) \leq \liminf_{n\to\infty} \delta_{2n-1}\,.$$

Since $c \in \mathbb{R}_+$ has been arbitrary, $\delta_{2n-1} \to \infty$.

Second, if R_0 has *a single fixed point* r_0 (which necessarily is attracting at the left and repelling at the right) then, given any $c \in \mathbb{R}_+$ with $c > r_0$, take $\varepsilon > 0$ so small that $r_0 < r_\varepsilon^+ < c$. From (6.10), it is clear that

$$r_\varepsilon^- \leq \lim_{n\to\infty} R_\varepsilon^n(\delta_{2N_\varepsilon-1}) \leq \liminf_{n\to\infty} \delta_{2n-1}\,.$$

If $\limsup_{n\to\infty} \delta_{2n-1} > r_\varepsilon^+$ then there exists $N \geq N_\varepsilon$ so that $\delta_{2N-1} > r_\varepsilon^+$, and hence $\delta_{2n-1} > R_\varepsilon^{n-N}(\delta_{2N-1})$ for all $n \geq N$, which in turn yields $\liminf_{n\to\infty} \delta_{2n-1} \geq c$, and since $c > r_0$ has been arbitrary, $\delta_{2n-1} \to \infty$. By contrast, if $\limsup_{n\to\infty} \delta_{2n-1} \leq r_\varepsilon^+$ then clearly

$$r_\varepsilon^- \leq \liminf_{n\to\infty} \delta_{2n-1} \leq \limsup_{n\to\infty} \delta_{2n-1} \leq r_\varepsilon^+\,,$$

and letting $\varepsilon \downarrow 0$ yields $\lim_{n\to\infty} \delta_{2n-1} = r_0$.

Third, if R_0 has *two fixed points* $r_0^- < r_0^+$ (where r_0^- and r_0^+ necessarily are attracting and repelling, respectively) then pick any $c > r_0^+$ and take $\varepsilon > 0$ so small that $r_0^+ < r_\varepsilon^+ < c$. On the one hand, if $\liminf_{n\to\infty} \delta_{2n-1} < r_0^+$ then there exists $N \geq N_\varepsilon$ so that $\delta_{2N-1} < r_0^+$, and hence $\delta_{2n-1} < R_0^{n-N}(\delta_{2N-1})$ for all $n \geq N$, so $\limsup_{n\to\infty} \delta_{2n-1} \leq \lim_{n\to\infty} R_0^{n-N}(\delta_{2N-1}) = r_0^-$. Since clearly $\liminf_{n\to\infty} \delta_{2n-1} \geq r_\varepsilon^-$, letting $\varepsilon \downarrow 0$ yields $\lim_{n\to\infty} \delta_{2n-1} = r_0^-$. On the other hand, if $\limsup_{n\to\infty} \delta_{2n-1} > r_\varepsilon^+$ then $\delta_{2n-1} \to \infty$, just as in the case of a single fixed point. The only remaining possibility is that

$$r_0^+ \leq \liminf_{n\to\infty} \delta_{2n-1} \leq \limsup_{n\to\infty} \delta_{2n-1} \leq r_\varepsilon^+\,,$$

where letting $\varepsilon \downarrow 0$ yields $\lim_{n\to\infty} \delta_{2n-1} = r_0^+$.

In summary, it has thus been shown that (δ_{2n-1}) converges in $\mathbb{R} \cup \{\infty\}$ for every choice of $y \in \mathbb{R}_+^2$. In fact, every *finite* limit of (δ_{2n-1}) has been identified as a fixed point of R_0. Moreover, it is clear that the set

$$\mathbb{V}_o := \big\{y \in \mathbb{R}_+^2 : \lim_{n\to\infty} \delta_{2n-1} = \infty\big\}$$

is non-empty and open. If R_0 has no fixed point then clearly $\mathbb{V}_o = \mathbb{R}_+^2$, but otherwise $\mathbb{V}_o$ may be smaller.

To establish (6.8), it remains to consider the sequence (δ_{2n}). For this, simply notice that (6.9) remains valid with n replaced by $n + \frac{1}{2}$. Thus, the same argument as above shows that (δ_{2n}) either converges to a finite limit (which again must be a fixed point of R_0) or else $\delta_{2n} \to \infty$. This fully proves (6.8). Again, the set

$$\mathbb{V}_e := \big\{y \in \mathbb{R}_+^2 : \lim_{n\to\infty} \delta_{2n} = \infty\big\}$$

is non-empty and open. In addition, the identity

$$b_2\delta_{2n-1} + b_1\delta_{2n} = b_2(b_2 - b_1^2)y_{2n-2} - b_1^2 \log(1 + a_1 10^{-\delta_{2n-1}}) \qquad \forall n \geq 2\,,$$

shows that $\lim_{n\to\infty} \delta_{2n-1}$ and $\lim_{n\to\infty} \delta_{2n}$ cannot both be finite. In other words, $\mathbb{V}_o \cup \mathbb{V}_e = \mathbb{R}_+^2$.

The scene is now set for a ("shadowing") argument utilizing Proposition 6.3 which is similar to the one in Case I, yet slightly more involved. Thus, assume for

instance that $y \in \mathbb{V}_o$, i.e., $\lim_{n\to\infty} \delta_{2n-1} = \infty$. As in Case I, deduce from (6.7) that $\lim_{n\to\infty}\big(y_{2n} - b_2^{n-1}h(y)\big) = 0$, where for every $y \in \mathbb{R}^2_+$,

$$h(y) = y_2 + \sum\nolimits_{k=1}^{\infty} b_2^{-k} \log\big(1 + a_1 10^{-\delta_{2k+1}}\big)\,, \tag{6.11}$$

and hence (y_{2n}) is u.d. mod 1 if and only if $\big(b_2^{n-1}h(y)\big)$ is. In analogy to Case I, it is readily seen that $\lim_{n\to\infty} \delta_{2n-1} b_2^{-n}$ is finite and positive, so $\delta_{2n-1} \to \infty$ exponentially fast as $n \to \infty$. Letting

$$\mathbb{E}_o = \left\{ y \in \mathbb{V}_o : (y_{2n}) \text{ or } \left(\frac{b_1}{b_2}\, y_{2n}\right) \text{ is } \textit{not} \text{ u.d. mod } 1 \right\},$$

it will now be shown that

$$\mathbb{E}_o \subset \mathbb{V}_o \text{ is a nullset}\,, \tag{6.12}$$

whereas later it will also be shown that

$$\mathbb{V}_o \cap \mathbb{U} \text{ is meagre (in } \mathbb{V}_o)\,. \tag{6.13}$$

As a first step towards (6.12), and analogously to Case I, observe that

$$\delta_{2n+1} = b_2 y_{2n} - b_1 y_{2n+1} = b_2 g^{2n-1}(y)_1 - b_1 g^{2n-1}(y)_2 \qquad \forall n \geq 1\,,$$

with the smooth map $g : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$g(u,v) = \big(v, b_2 u + \log(1 + a_1 10^{b_1 v - b_2 u})\big)\,.$$

Note that $h : \mathbb{R}^2_+ \to \mathbb{R}$ is continuous, and $h > 0$ on $\mathbb{V}_o$. (Formal) termwise partial differentiation of (6.11) on $\mathbb{V}_o$ yields

$$\frac{\partial h}{\partial y_2} = 1 - \sum\nolimits_{k=1}^{\infty} b_2^{-k} \frac{a_1 10^{-\delta_{2k+1}}}{1 + a_1 10^{-\delta_{2k+1}}} \cdot \frac{\partial \delta_{2k+1}}{\partial y_2}\,, \tag{6.14}$$

where, for every $k \geq 1$,

$$\begin{aligned} \frac{\partial \delta_{2k+1}}{\partial y_2} &= b_2 \frac{\partial}{\partial y_2} g^{2k-1}(y)_1 - b_1 \frac{\partial}{\partial y_2} g^{2k-1}(y)_2 \\ &= b_2 \left(\prod_{\ell=0}^{2k-1} Dg \circ g^{\ell}(y) \right)_{21} - b_1 \left(\prod_{\ell=0}^{2k-1} Dg \circ g^{\ell}(y) \right)_{22}\,. \end{aligned}$$

Observe now that

$$Dg \circ g^{\ell}(y) = \begin{bmatrix} 0 & \dfrac{b_2}{1 + a_1 10^{-\delta_{\ell+2}}} \\[2ex] 1 & \dfrac{a_1 b_1 10^{-\delta_{\ell+2}}}{1 + a_1 10^{-\delta_{\ell+2}}} \end{bmatrix} \qquad \forall \ell \geq 0\,.$$

Recall that $\delta_{2n-1} \to \infty$, while the sequence (δ_{2n}) either converges to a finite limit, or else $\delta_{2n} \to \infty$ as well. In either case, and similarly to Case I, $\sup_{\ell\geq 0} |Dg \circ g^{\ell}| < \infty$, and since $\delta_{2n-1} \to \infty$ exponentially fast, (6.14) shows that $\partial h/\partial y_2$ is continuous on $\mathbb{V}_o$, as is $\partial h/\partial y_1$. Again, pick $c \in \mathbb{R}_+$, possibly depending on a_1, b_1, b_2 but independent of y_1, y_2, so large that $\partial h/\partial y_2 > \frac{1}{2}$ whenever $\delta_3 > c$. With this, let

$\mathbb{W}_o = \{y \in \mathbb{V}_o : \partial h/\partial y_2 > \frac{1}{2}\}$. Since $h(y_1, \cdot)$ is a local diffeomorphism on every non-empty y_1-section of $\mathbb{W}_0$, the y_1-section

$$(\mathbb{E}_o \cap \mathbb{W}_o)_{y_1} = \left\{ y_2 \in \mathbb{R}_+ : y \in \mathbb{W}_o,\, (y_{2n}) \text{ or } \left(\frac{b_1}{b_2}\, y_{2n}\right) \text{ is } \textit{not} \text{ u.d. mod } 1 \right\}$$

is a nullset. Thus, $\mathbb{E}_o \cap \mathbb{W}_o$ is a nullset as well. Since $\delta_{2n-1} \to \infty$ as $n \to \infty$, $g^{2n}(y) \in \mathbb{W}_o$ for all sufficiently large n. Thus $\mathbb{V}_o \subset \bigcup_{n\in\mathbb{N}} g^{-2n}(\mathbb{W}_o)$, and since clearly $g^2(\mathbb{E}_o) \subset \mathbb{E}_o$,

$$\mathbb{E}_o = \mathbb{E}_o \cap \bigcup_{n\in\mathbb{N}} g^{-2n}(\mathbb{W}_o) = \bigcup_{n\in\mathbb{N}} \left(\mathbb{E}_o \cap g^{-2n}(\mathbb{W}_o)\right) \subset \bigcup_{n\in\mathbb{N}} g^{-2n}(\mathbb{E}_o \cap \mathbb{W}_o)\,.$$

Since g is a diffeomorphism of $\mathbb{R}_2^+$ onto $g(\mathbb{R}_+^2)$, this proves (6.12).

Now, to establish (6.13), observe first that, as in the proof of [7, Thm. 6.46], given any $y \in \mathbb{W}_o$, $\varepsilon > 0$, and $0 < t < \frac{1}{4}$ so small that

$$10t\left(\frac{\log 2 - \log t}{\log b_2} + 1\right) < 1\,,$$

there exists $z \in \mathbb{W}_o$ with $|z - y| < \varepsilon$ and

$$\liminf_{N\to\infty} \frac{\#\{1 \le n \le N : z_{2n} \in [\frac{1}{2} - t, \frac{1}{2} + t] + \mathbb{Z}\}}{N} > 10t\,,$$

where $z_n = \log x_n$ for all n, and (x_n) is the solution of (6.6) with $(x_1, x_2) = (10^{z_1}, 10^{z_2})$. It follows that

$$\liminf_{N\to\infty} \frac{\#\{1 \le n \le N : z_n \in [\frac{1}{2} - t, \frac{1}{2} + t] + \mathbb{Z}\}}{N} > 5t\,. \tag{6.15}$$

Consider the continuous, 1-periodic function $\psi : \mathbb{R} \to [0, 1]$ with

$$\psi(s) = \begin{cases} 0 & \text{if } \operatorname{dist}(s, \frac{1}{2} + \mathbb{Z}) \ge 2t\,, \\ 2 - \operatorname{dist}(s, \frac{1}{2} + \mathbb{Z})/t & \text{if } t \le \operatorname{dist}(s, \frac{1}{2} + \mathbb{Z}) < 2t\,, \\ 1 & \text{if } \operatorname{dist}(s, \frac{1}{2} + \mathbb{Z}) < t\,, \end{cases}$$

and note that $\int_0^1 \psi(s)\,\mathrm{d}s = 3t$. By (6.15),

$$\liminf_{N\to\infty} \frac{1}{N} \sum\nolimits_{n=1}^{N} \psi(z_n) > 5t\,,$$

and hence

$$\liminf_{N\to\infty} \left| \frac{1}{N} \sum\nolimits_{n=1}^{N} \psi(z_n) - \int_0^1 \psi(s)\,\mathrm{d}s \right| > 2t\,. \tag{6.16}$$

Now, for every $m \in \mathbb{N}$ define $V_m \subset \mathbb{W}_o$ as

$$V_m = \left\{ y \in \mathbb{W}_o : \left| \frac{1}{N} \sum\nolimits_{n=1}^{N} \psi(y_n) - \int_0^1 \psi(s)\,\mathrm{d}s \right| \le t\ \forall N \ge m \right\}.$$

Note that the set V_m is closed (in $\mathbb{W}_o$) by continuity, but has empty interior,

by (6.16). Also, $V_1 \subset V_2 \subset \cdots$, and $\mathbb{W}_o \cap \mathbb{U} \subset \bigcup_{m\in\mathbb{N}} V_m$ by Weyl's criterion [14, Cor. I.1.2]. Recall from above that $\mathbb{V}_o \subset \bigcup_{n\in\mathbb{N}} g^{-2n}(\mathbb{W}_o)$, and clearly $g^2(\mathbb{U}) \subset \mathbb{U}$. Thus

$$\mathbb{V}_o \cap \mathbb{U} \subset \bigcup_{n\in\mathbb{N}} (g^{-2n}(\mathbb{W}_o) \cap \mathbb{U}) \subset \bigcup_{n\in\mathbb{N}} g^{-2n}(\mathbb{W}_o \cap \mathbb{U}) \subset \bigcup_{m,n\in\mathbb{N}} g^{-2n}(V_m)\,,$$

which shows that $\mathbb{V}_o \cap \mathbb{U}$ is meagre, and hence proves (6.13).

An entirely analogous argument utilizing the set

$$\mathbb{E}_e := \left\{ y \in \mathbb{V}_e : (y_{2n-1}) \text{ or } \left(\frac{b_1}{b_2}\, y_{2n-1}\right) \text{ is } \textit{not} \text{ u.d. mod } 1 \right\},$$

shows that, similarly to (6.12) and (6.13),

$$\mathbb{E}_e \subset \mathbb{V}_e \text{ is a nullset, and } \mathbb{V}_e \cap \mathbb{U} \text{ is meagre}\,.$$

Clearly, therefore, $\mathbb{U} = (\mathbb{V}_o \cap \mathbb{U}) \cup (\mathbb{V}_e \cap \mathbb{U})$ is meagre as well, and it only remains to be shown that $\mathbb{U}$ has full measure in $\mathbb{R}^2_+$. To this end, note that $\mathbb{E}_o \cup \mathbb{E}_e$ is a nullset, and pick any $y \in \mathbb{R}^2_+ \setminus (\mathbb{E}_o \cup \mathbb{E}_e)$. Assume, for instance, that $y \in \mathbb{V}_o$. Then (y_{2n}) and $(b_1 b_2^{-1} y_{2n})$ both are u.d. mod 1. On the one hand, if $y \in \mathbb{V}_e$ then (y_{2n-1}) is u.d. mod 1 as well. On the other hand, if $y \notin \mathbb{V}_e$ then (δ_{2n}) converges to a finite limit. In this case, recall that

$$y_{2n-1} = \frac{b_1}{b_2} y_{2n} + \frac{1}{b_2}\delta_{2n} \qquad \forall n \geq 1\,,$$

which shows that, again, (y_{2n-1}) is u.d. mod 1. In either case, therefore, (y_{2n}) and (y_{2n-1}) both are u.d. mod 1, and hence (y_n) is as well. By completely analogous reasoning, the same conclusion holds whenever $y \in \mathbb{V}_e$. Thus (y_n) is u.d. mod 1 whenever $y \in \mathbb{R}^2_+ \setminus (\mathbb{E}_o \cup \mathbb{E}_e)$. In other words, $(\mathbb{R}^2_+ \setminus \mathbb{U}) \subset \mathbb{E}_o \cup \mathbb{E}_e$, and so $\mathbb{R}^2_+ \setminus \mathbb{U}$ is a nullset. Thus, $\mathbb{U}$ has full measure but, as seen earlier, is meagre in $\mathbb{R}^2_+$. This proves the assertion of the theorem in case $(x_1, x_2) \in \mathbb{A}_\infty$ and $b_1^2 < b_2$.

To conclude the overal proof, the case of $(x_1, x_2) \in \mathbb{A}_0$ still needs to be considered. In essence, the analysis in this case proceeds exactly as in the case of $(x_1, x_2) \in \mathbb{A}_\infty$, but with Cases I and III interchanging their roles. To see this, fix $0 < \xi < 1$ so that $T(]0, \xi[^2) \subset]0, \xi[^2$. Thus, if $0 < x_1, x_2 < \xi$ then $0 < x_n < \xi$ for every n, and $\lim_{n\to\infty} x_n = 0$. In analogy to before, it suffices to show that $\mathbb{B}$ has full measure but is meagre in $]0, \xi[^2$. Assume for instance that $b_1^2 < b_2$, and consider (6.2) in its rescaled form (6.3). With δ_n defined as before,

$$10^{\delta_n} = \frac{x_{n-1}^{b_2}}{x_n^{b_1}} < x_{n-1}^{b_2 - b_1^2} \to 0 \quad \text{as } n \to \infty\,.$$

Thus, $\delta_n \to -\infty$, and the same technique as in Case I applies. By contrast, if $b_1^2 > b_2$ then consider (6.6), with (6.9) remaining valid as before, and again $(b_2 - b_1^2) y_{2n-1} \to \infty$ as $n \to \infty$. Consequently, the same technique as in Case III applies. Finally, as observed earlier, the argument in Case II applies equally if $(x_1, x_2) \in \mathbb{A}_0$. □

Corollary 6.4. *Let $a_1, a_2 \in \mathbb{R}_+$. Then the solution (x_n) of the recursion $x_n = a_1 x_{n-1}^2 + a_2 x_{n-2}^2$, $n \geq 3$, is Benford for almost all $(x_1, x_2) \in \mathbb{R}^2$, but the set $\{(x_1, x_2) \in \mathbb{R}^2 : (x_n) \text{ is not Benford}\}$ is residual in $\mathbb{R}^2$.*

Proof. This follows immediately from Theorem 6.1, since $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(u, v) = (v, a_2u^2 + a_1v^2)$ is a diffeomorphism when restricted to any one of the four open quadrants, and $T^2(\mathbb{R}^2 \setminus \{(0, 0)\}) \subset \mathbb{R}^2_+$. Notice that Case I in the above proof of Theorem 6.1 applies always, since $b_1^2 = 4 > 2 = b_2$. □

Example 6.5. Consider the "quadratic Fibonacci" recursion

$$x_n = x_{n-1}^2 + x_{n-2}^2 \qquad \forall n \geq 3\,. \tag{6.17}$$

By Corollary 6.4, (x_n) is Benford for almost all, but not all $(x_1, x_2) \in \mathbb{R}^2$. The boundary of $\mathbb{A}_0$, the domain of attraction of 0, is a smooth oval containing $(\frac{1}{2}, \frac{1}{2})$. In fact, $\lim_{n\to\infty} x_n = \frac{1}{2}$ whenever $(x_1, x_2) \in \partial\mathbb{A}_0$; see also Figure C. Note that $x_n \to \infty$ for every $(x_1, x_2) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$, but the authors do not know whether or not any of the resulting (doubly-exponentially growing) integer sequences are Benford.

As a slightly modified version of (6.17), consider the recursion

$$x_n = x_{n-1}^2 + 4x_{n-2}^2 \qquad \forall n \geq 3\,.$$

Again, $\partial\mathbb{A}_0$ is a smooth oval, and $(\frac{1}{5}, \frac{1}{5}) \in \partial\mathbb{A}_0$. Unlike for (6.17), however, picking $(x_1, x_2) \in \partial\mathbb{A}_0$ now yields a divergent sequence (x_n) unless $(|x_1|, |x_2|) = (\frac{1}{5}, \frac{1}{5})$ since, for instance, $\lim_{n\to\infty} x_{2n-1} = \frac{1}{30}(5 + \sqrt{5})$ and $\lim_{n\to\infty} x_{2n} = \frac{1}{30}(5 - \sqrt{5})$ whenever $(x_1, x_2) \in \partial\mathbb{A}_0$ and $|x_1| > \frac{1}{5}$.

Example 6.6. The conclusion of Theorem 6.1 may remain valid, at least partially, even if one of the parameters b_1, b_2 is ≤ 1. For example, it is readily seen that for the recursion

$$x_n = x_{n-1}^2 + \sqrt{x_{n-2}} \qquad \forall n \geq 3\,,$$

$x_n \to \infty$ for every $(x_1, x_2) \in \mathbb{X} = \mathbb{R}^2_+$, with $\mathbb{B}$ having full measure but being meagre in $\mathbb{X}$.

By contrast, for the slightly modified recursion

$$x_n = x_{n-1}^2 + \frac{1}{4}\sqrt{x_{n-2}} \qquad \forall n \geq 3\,,$$

it can be shown that $\lim_{n\to\infty} x_n = \zeta_-^2 \approx 0.07268$ whenever $0 < x_1, x_2 < \zeta_+^2 \approx 0.7015$, where ζ_- and ζ_+ is the smaller and larger positive root of $4z^3 - 4z + 1 = 0$, respectively; in particular, (x_n) is *not* Benford for any such (x_1, x_2). Even in this example, however, $\mathbb{A}_\infty \supset]\zeta_+^2, \infty[^2$, and $\mathbb{B}$ has full measure but is meagre in $\mathbb{A}_\infty$.

In conclusion, the reader may observe that even though Theorem 6.1 does not apply to them, the recursions of Example 6.6 do generate an abundance of Benford sequences. So, indeed, do many other non-linear recursions and more general dynamical systems. The study of BL in dynamical systems will undoubtedly remain an intriguing and multi-faceted area of research for many years to come.

Acknowledgements

The first author was partially supported by an NSERC Discovery Grant. Both authors are grateful to the editors for the invitation to contribute to this special issue honoring Professor Bunimovich, and to an anonymous referee for several helpful comments and corrections.

References

[1] A. Berger, *Some dynamical properties of Benford sequences*, J. Difference Equ. Appl. **17** (2011), 137–159.
[2] A. Berger, L. Bunimovich and T. P. Hill, *One-dimensional dynamical systems and Benford's law*, Trans. Amer. Math. Soc. **357** (2005), 197–219.
[3] A. Berger and A. Duh, *Global saddle-type dynamics for convex second-order difference equations*, J. Difference Equ. Appl. **23** (2017), 1807–1823.
[4] A. Berger and G. Eshun, *Benford solutions of linear difference equations*, in: Theory and Applications of Difference Equations and Discrete Dynamical Systems, Z. AlSharawi et al. (eds.), vol. 102, Springer Proceedings in Mathematics & Statistics, 2014, pp. 23–60.
[5] A. Berger and G. Eshun, *A characterization of Benford's law in discrete-time linear systems*, J. Dynam. Differential Equations **28** (2016), 431–469.
[6] A. Berger and S. N. Evans, *A limit theorem for occupation measures of Lévy processes in compact groups*, Stoch. Dyn. **13** (2013): 1250008.
[7] A. Berger and T. P. Hill, *An Introduction to Benford's Law*. Princeton University Press, Princeton, NJ, 2015.
[8] A. Berger, T. P. Hill, B. Kaynar and A. Ridder, *Finite-state Markov chains obey Benford's law*, SIAM J. Matrix Analysis **32** (2011), 665–684.
[9] A. Berger, T. P. Hill and E. Rogers, Benford Online Bibliography, `http://www.benfordonline.net`. (Last accessed January 22, 2025.)
[10] A. Berger and A. Rahmatidehkordi, *The primes perform a Benford dance*, To appear in *Anal. Appl.*, 2025.
[11] D. Biau, *The first-digit frequencies in data of turbulent flows* Physica A: Statistical Mechanics and its Applications **440** (2015), 147–154.
[12] 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.
[13] J. Eells and K. D. Elworthy, *Stochastic dynamical systems*, in: International Seminar Course on Control Theory and Topics in Functional Analysis, International Atomic Energy Agency (IAEA), 1976, pp. 179–185.
[14] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974
[15] J. C. Oxtoby, *Measure and Category*, Graduate Texts in Mathematics 2 (2nd ed.), Springer-Verlag, New York-Heidelberg-Berlin, 1980.
[16] M. Sambridge, H. Tkalčič and A. Jackson, *Benford's law in the natural sciences*, Geophysical Research Letters **37** (2010): L2230.
[17] K. Schürger, *Extensions of Black-Scholes processes and Benford's law*, Stochastic Processes and their Applications **118** (2008), 1219–1243.
[18] M. A. Snyder, J. H. Curry and A. M. Dougherty, *Stochastic aspects of one-dimensional discrete dynamical systems: Benford's law*, Physical Review E **64** (2001): 026222.
[19] C. R. Tolle, J. L. Budzien and R. A. LaViolette, *Do dynamical systems follow Benford's law?* Chaos **10** (2000), 331–336.

Manuscript received September 24 2024
revised January 23 2025

A. Berger
Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Canada
E-mail address: berger@ualberta.ca

T. P. Hill
School of Mathematics, Georgia Institute of Technology, Atlanta, USA
E-mail address: theodore.hill@math.gatech.edu